

## Article

# A Novel Approach to Calculate the Spatial–Temporal Correlation for Traffic Flow Based on the Structure of Urban Road Networks and Traffic Dynamic Theory

Mao Du, Lin Yang \* and Jiayu Tu

School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China; dumaosjtu@sjtu.edu.cn (M.D.); tujiayu@sjtu.edu.cn (J.T.)

\* Correspondence: yanglin@sjtu.edu.cn

**Abstract:** Determining the spatial–temporal correlation (STC) between roads can help clarify the operation characteristics of road traffic. Moreover, this correlation affects the utilization quality of traffic data in related research fields. Therefore, it is of significance to provide more reasonable correlation information for other research, such as in traffic speed prediction. Most of the traditional correlation calculation methods for traffic are based on only statistical theory. These methods are simple, but their ability to explain the actual phenomenon is limited due to the lack of consideration of the actual traffic operation characteristics. Therefore, to provide more reasonable correlation information between roads, this paper analysed the influence mechanism of urban traffic based on the traffic dynamic model, and two parameters, traffic complete influence time and traffic correlation strength, were proposed to bring physical meaning to the calculation of STC. Then, an improved calculation model of the STC between different roads considering the adjacency between roads was proposed in this paper. Finally, this paper verified this method against two common traditional methods through different experiments. The verification results show that the calculation method proposed in this paper has better interpretability for the STC between different roads and can better reveal the internal traffic operation characteristics of the road network.

**Keywords:** dynamic programming; intelligent transportation systems; operation characteristics; spatial–temporal correlation; urban road network



**Citation:** Du, M.; Yang, L.; Tu, J. A Novel Approach to Calculate the Spatial–Temporal Correlation for Traffic Flow Based on the Structure of Urban Road Networks and Traffic Dynamic Theory. *Sensors* **2021**, *21*, 4725. <https://doi.org/10.3390/s21144725>

Academic Editor: Giovanni Pau

Received: 3 June 2021

Accepted: 6 July 2021

Published: 10 July 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

With the continuous development of technology, intelligent transportation systems (ITSs) have emerged as an effective means to alleviate congestion [1]. The use of ITSs can help enhance the operational efficiency of traffic and energy efficiencies of vehicles [2]. The correlation of road traffic mainly reflects the degree of interaction between different roads in the past period, which is of great significance for the future development of traffic state and traffic control research [3–5]. Owing to the time-varying characteristics of traffic flow and the spatial distribution characteristics of the road structure, urban traffic data exhibit a notable spatial–temporal characteristic.

The establishment of a scientific and reasonable calculation method for the spatial–temporal correlation (STC) is critical to realize urban traffic research, for example, for the identification of key roads, analysis of the road network topology and development of short-term traffic prediction algorithms [6–8].

However, after the investigation, it can be found that the current research did not provide in-depth research on the correlation between traffic states. Most studies in the field of transportation use statistical principles to measure the correlation between roads.

In fact, in the field of transportation, most of the calculation methods only calculate the similarity of two traffic series data but cannot truly judge whether there is traffic correlation between them. For example, a certain road is congested at 8:00 a.m., while its adjacent

road is congested at 8:00 p.m. According to the principle of statistics, the two traffic states are very similar, but there is obviously no correlation between them. Alternatively, two roads with a distance of 20 km are congested in the same period of time, but in fact, the traffic states between them do not affect each other, so there is also no correlation between them. Thus, studies based on only the principle of statistics cannot truly reveal the internal correlation characteristics of traffic states, and the correlation results may be consistent with the actual situation.

The traffic network of human society has very complex characteristics, and it is impossible to model it completely. However, with the continuous development of science and theory, various traffic flow theories and modelling methods have been proposed [9]. These models and methods bring more physical meaning to the related research in the field of transportation, and they are helpful to explore the origin and evolution of urban traffic flow.

The major shortcomings of the current research on the correlation of traffic flow are the lack of analysis on the interaction between traffic statuses and the lack of physical support and interpretation of the calculation results. Therefore, in view of the prominent role of traffic models in revealing the dynamic changes in traffic flow, this paper considers the introduction of a traffic dynamic model to improve the calculation method of traffic flow correlation.

In view of the different levels of research and application, these methods can be divided into three categories according to the space metric of the road networks: microscopic, mesoscopic and macroscopic traffic flow analysis. Microscopic traffic flow analysis focuses on the traffic flow characteristics on a point or an end face, including the car-following model [10], lane changing model [11], cellular automaton model [12], etc. Mesoscopic traffic flow analysis studies the traffic flow characteristics of road sections. For example, headway distribution model [13], queue model [14], and so on.

Different from microscopic and mesoscopic traffic flow analysis, macroscopic traffic flow analysis (MTFA) pays attention to the traffic flow characteristics of areas of road networks, such as the average speed and traffic flow over a period of time [15]. Compared with the former two methods, MTFA has the advantages of fewer parameters and simple calculation. It can effectively describe the operation characteristics of traffic flow [16] and has the ability to explore the overall regional traffic operation law [17]. This is of great significance to future traffic state prediction and traffic flow control of urban road networks [18,19]. Among them, the LWR model introduces the theoretical knowledge of fluid mechanics into the study of traffic flow, which makes the study of macrotraffic more physically meaningful, so it has been widely used in the field of traffic research [20]. AliA et al. used macroscopic fundamental diagrams (MFDs) to investigate perimeter traffic flow control for traffic congestion control in heterogeneously congested cities. A robust control of the two-region MFD system is presented to achieve good performance in perimeter traffic flow control [21].

In addition, the operation of traffic flow on urban road networks is closely related to the structural characteristics of road networks. It is undeniable that the urban road network has the characteristics of multiple branches and lanes. The traffic state of the same road will be affected by the traffic state of multiple roads at the same time. It is worth noting that there is no direct impact between the non-directly adjacent roads, and the traffic transfer between them is mainly realized through the transfer path between the roads. Therefore, the correlation between indirect adjacent roads is bound to be affected by the correlation between two adjacent roads on the whole transmission path.

Therefore, considering that the macro traffic model has the advantages of simple calculation and can reveal the dynamic changes of traffic flow, combined with the structural characteristics of urban road networks, this paper proposes an STC calculation method based on a traffic dynamic model for the first time. Experiments are conducted to verify the proposed method systematically.

Three primary contributions of this paper are as follows:

First, based on the traffic dynamic model (LWR model), this paper explores the influence characteristics of urban traffic flow. The parameter of traffic complete influence time is proposed to measure the influence time between different roads for the first time.

Second, considering the multibranch characteristics of the urban road network structure, combined with the separation and confluence mechanism of intersections, the traffic correlation strength is proposed to effectively characterize the influence strength of different roads on the same road.

Third, based on these two parameters, in view of the different influence mechanisms of adjacent roads and non-adjacent roads, this paper creatively put forward the idea of establishing the corresponding STC calculation method according to whether the roads are directly adjacent. To integrate the structural characteristics of urban roads into the calculation process of STC, the calculation results are more explanatory.

The proposed method was verified through different experiments. Compared with the calculation method based on statistical principles, the calculation method proposed in this paper has better interpretability for the STC between different roads and can better reflect the internal traffic operation characteristics of the road network. In addition, the results calculated by the proposed method in this paper show that the traffic correlation decreases with increasing time delay between traffic states and increases with increasing road importance and traffic circulation frequency. Therefore, it indirectly verifies the rationality of this method.

The remainder of this paper is organized as follows. In Section 2, the current research progress of STC is discussed. In Section 3, combined with the characteristics of traffic flow and urban road network structure, this paper proposes two physical parameters to improve the measurement of STC. In Section 4, the corresponding STC calculation method is constructed according to whether the roads are directly adjacent. Section 5 compares the proposed method in this paper with two statistical-based methods. Finally, Section 6 concludes the paper.

## 2. Related Work

The time-varying characteristics of traffic flow and the spatial distribution characteristics of roads cause these traffic data to have obvious spatial-temporal characteristics. Therefore, in the development of a prediction algorithm that is based on multi-sensors, an increasing number of scholars also consider the spatial-temporal correlation (STC) characteristics of traffic in the prediction algorithm [22]. A scientific and reasonable correlation of urban traffic networks is very important for urban traffic research, identification of key roads, analysis of road network topology and short-term traffic prediction algorithms [23].

Based on complex networks, Tang, J. and Wang, Y. analyses the time characteristics of traffic data generated by urban road networks [24]. Complex networks are constructed by using the correlation coefficient among days to measure the complexity of traffic time series at different temporal scales.

Yang and Gar-on investigated the spatiotemporal dependency of traffic flow using cross-correlation analysis and then discusses its implications in terms of traffic forecast ability and real-time data effectiveness [25]. Duan et al. [26] proposed a unified spatial-temporal model for short-term road traffic prediction. This model incorporated the physical factors that potentially affect the variation in the STC into a series of parameters. Another study [27] investigated the inclusion of the STC and interaction into a multivariate random-parameter Tobit model with different severity outcomes. The estimation results show that spatial and temporal effects and their interactive effects are significant and that the spatial and interactive effects have strong correlations across injury severities. A Bayesian spatiotemporal model was proposed to measure the association between the crash frequency and possible risk factors [28]. The proposed model can also accommodate the unstructured random effect, and spatio-temporal correlation and interactions. Guo et al. [29] utilized the weighted degree and impact distance as two major measures to identify the most influential locations. A road segment with a larger weighted degree or larger impact distance

suggests that its traffic flow can strongly influence neighboring road sections driven by the congestion propagation.

In regard to the STC calculation of the traffic state, the following correlation calculation methods are often used for analysis:

- (1) Time autocorrelation coefficient. This statistical index is used to analyse the correlation degree between the states of the same road at any two instances. However, this index describes only the correlation degree of time and disregards the spatial properties [30].
- (2) Spatial autocorrelation coefficient. This coefficient reflects the correlation degree of the attribute variables with a certain regularity in different spatial positions. The common statistical indexes for the spatial autocorrelation analysis are Moran's I index and Geary's C index [31].
- (3) STC. The change characteristics of the traffic information considering both the time and space dimensions are analysed. To date, many scholars have extended several indexes to measure the STC, such as the simple spatial-temporal autocorrelation index (ST-ACF) and cross-correlation function (CCF) [32–34]. The common calculation methods of STC include:

- (a) ST-ACF

$$\rho_{lk}(s) = \frac{\sum_{i=1}^N \sum_{t=1}^T L^l x_i(t) L^k x_i(t+s)}{\sqrt{\sum_{i=1}^N \sum_{t=1}^T (L^l x_i(t))^2 \sum_{i=1}^N \sum_{t=1}^T (L^k x_i(t+s))^2}} \quad (1)$$

$$L^0 x_i(t) = x_i(t), L^l x_i(t) = \sum_{j=1}^N \omega_{ij}^l x_j(t), \sum_{j=1}^N \omega_{ij}^l = 1$$

where  $\rho_{lk}(s)$  is ST-ACF.  $t$  is the statistical period, and  $s$  is the time delay.  $L^l$  is the  $l$ -order spatial delay operator of the road network.  $L^l x_i(t)$  is the comprehensive observation value of traffic with all  $l$ -order adjacent roads of the TR.  $\omega_{ij}^l$  is the weight of  $l$ -order adjacent roads.

- (b) CCF

$$\rho_{XY}(s) = \frac{E[(x(t) - \mu_x)(y(t+s) - \mu_y)]}{\sqrt{\sigma_X^2 \sigma_Y^2}} \quad (2)$$

where  $\rho_{XY}(s)$  is the correlation value of data series.  $x(t)$  and  $y(t+s)$  are the corresponding values,  $\mu_x$  and  $\mu_y$  are their average values, and  $\sigma_X$  and  $\sigma_Y$  are their standard deviations.

The first two indexes can calculate the correlation considering only one dimension (time or space). The ST-ACF reflects the overall correlation between different orders of the road network. The CCF can calculate the correlation between two single roads. It is worth noting that most of these methods are based on the Pearson coefficient or its variant structures [35]. Since only several statistical indicators, mean and variance, are used to calculate the correlation of different data series, this method has high generalization ability and has been applied to many fields, including traffic state analysis. However, a fact that cannot be ignored is that the Pearson coefficient based on statistical theory can only reveal the correlation between calculated objects at the data level rather than the physical level.

At the same time, based on the advantages of current technology development, many data collection and calculation models have been proposed. The heterogeneous sensor networks and distributed computing system proposed in related research [36–38] could provide new help for the collection and analysis of road network traffic data, and help to deeply mine the road network traffic correlation.

Therefore, based on the traffic dynamic theory and the structural characteristics of urban road networks, this paper innovatively proposes a novel approach calculation model of STC.

### 3. Physical Parameters for STC Calculation

Specifically, STC is the correlation between the traffic statuses of different roads in different time periods. Without loss of generality, a typical urban road network can be illustrated, as shown in the left subgraph in Figure 1. Note that the traffic state must be treated from the perspective of time and space. The traffic state may correspond to a spatial-temporal network composed of different spatial-temporal regions (STRs), as shown in the right subgraph of Figure 1. The ordinate and abscissa represent different roads and different times, respectively. The STR is the traffic data collected in different time periods. The STC is the correlation between two different STRs, as shown in Figure 1. For convenience, we call the road affected by other roads the target road (TR) in this paper.

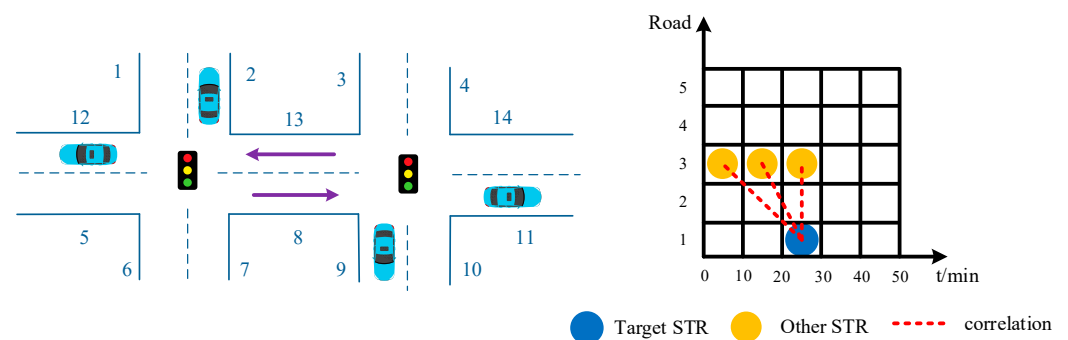


Figure 1. Schematic of a traffic road network.

Considering that the traffic flow has the characteristics of dynamic update, the influence time of the target road affected by the historical traffic state of other roads must be limited. Therefore, in Section 3.1, the parameter of traffic complete influence time and its calculation method is proposed, which is based on the transmission of traffic flow and traffic waves, to measure the influence time of the target road affected by the traffic state of other roads in the historical period.

Moreover, considering the characteristics of urban road networks with multiple branches, this paper proposes the parameter of traffic correlation strength calculated by the flow ratio to measure the correlation intensity of different roads to the same target road based on the diversion combined mechanism of intersections in Section 3.2.

#### 3.1. Traffic Complete Influence Time and Its Calculation Method

With the continuous flow of newly entering traffic, the composition of the traffic flow on the TR changes with time. Assume traffic condition of two roads, road  $i$  and road  $j$ , at different sampling times, let road  $j$  is upstream of road  $i$ . Assume that all the traffic on road  $i$  during a period is generated by road  $j$ . Then, the traffic state of road  $i$  in the sampling period (A) might be very similar to that of road  $j$  in the sampling period (C). However, in the sampling period (B), all the vehicles from road  $j$  in the sampling period (C) drive out of road  $i$ . Therefore, in fact, there is no correlation between these two STRs. Thus, it leads to a very important characteristic, that is, the influence time of traffic state in different spatial-temporal regions (STRs) is limited.

##### 3.1.1. The Traffic Complete Influence Time (TCIT)

The specified correlation refers to the correlation between two traffic data for the same length. Suppose the correlation between the traffic state of road  $j$  in time period  $(\tau_{j,1}, \tau_{j,n})$  and that of road  $i$  in time period  $(\tau_{i,1}, \tau_{i,n})$  need to be calculated. The time delay between them is  $\tau_{i,1} - \tau_{j,1} \geq 0$ .

Theoretically, as long as there is a connection between the two roads, the traffic state of road  $j$  will affect the traffic state of road  $i$  at every moment. However, due to the existence of road length, it takes time for the influence to transfer to or out of road  $i$ . Therefore, the influence of road  $j$  on road  $i$  has time delay characteristics at each time, and the time delay

may be different. That is, the maximum influence time of road  $j$  on road  $i$  will be different when the starting time is different.

Therefore, to determine the maximum influence time of the traffic state of road  $j$  in time period  $(\tau_{j,1}, \tau_{j,n})$  on that of road  $i$  in time period  $(\tau_{i,1}, \tau_{i,n})$ , it is necessary to analyse the influence time of every time in this period.

In period  $(\tau_{j,1}, \tau_{j,n})$ , the maximum influence time of road  $j$  on road  $i$  starting from time  $\tau_{j,s}$  was called as the local influence time (LIT), that is,  $\{t_{j,i}(\tau_{j,s}) | 1 \leq s \leq n\}$ . We call the maximum value of the LITs in time sequence  $[t_{j,i}(\tau_{j,1}), \dots, t_{j,i}(\tau_{j,n})]$  the traffic complete influence time (TCIT) from road  $j$  to road  $i$  in time period  $(\tau_{j,1}, \tau_{j,n})$ , that is,  $\xi_{j,i}(\tau_{j,1})$ . Once the time exceeds the TCIT, with the addition of new information, the influence of road  $j$  on road  $i$  generated in period  $(\tau_{j,1}, \tau_{j,n})$  is expected to gradually change and eventually disappear. Therefore, determining the local influence time is a critical aspect in the analysis.

### 3.1.2. TCIT Calculation Method

To calculate the LIT, it is necessary to explore the causes of the traffic state change. The LWR model is also called the kinematic wave model since it describes the evolution of traffic dynamics as a combination of characteristic, shock, and rarefaction waves. These waves result from the movements of vehicles [39]. As the fundamental diagram, the LWR model describes the average traffic dynamics. This model mainly analyses the changes in traffic flow, speed and density with spatial location and time in the macro state.

According to the theory of the LWR model, there are two main ways to change the traffic state: traffic flow and traffic wave. Only when the traffic flow or traffic wave exists in the target road will it affect its traffic state. Therefore, the time of the traffic flow and traffic wave passing out of the target road was calculated to measure the LIT. Then, the TCIT to the target road in the whole time series is obtained by calculating the maximum value of LITs at different times.

The influence principle of traffic flow and traffic waves on traffic states and the detailed calculation process of TCIT are as follows.

First, according to the traffic dynamic model, the traffic between any two points on the road follows the law of conservation of traffic flow:

$$\frac{\partial q}{\partial x} + \frac{\partial k}{\partial t} = g(x, t) \quad (3)$$

where  $q$  is the traffic flow,  $k$  is the traffic density,  $x$  is the road location variable, and  $t$  is the time variable.  $g(x, t)$  is the vehicle generation (or departure) rate. In the conservation equation, the traffic flow and speed can be regarded as a function of density, and then the dynamic equation of traffic flow can be derived:

$$\frac{dv}{dt} = -k \left( \frac{dv}{dk} \right)^2 \frac{\partial k}{\partial x} \quad (4)$$

where  $v$  is the traffic speed. Formula (4) shows that the average acceleration of vehicles,  $\frac{dv}{dt}$ , depends on the density gradient of traffic flow,  $\frac{\partial k}{\partial x}$ . If  $\frac{\partial k}{\partial x} > 0$ , that is, when the traffic density tends to increase,  $\frac{dv}{dt} < 0$ , the vehicle decelerates; if  $\frac{\partial k}{\partial x} < 0$ , that is, when the traffic density tends to decrease,  $\frac{dv}{dt} > 0$ , the vehicle accelerates.

This dynamic equation proved that the traffic flow has a direct impact on the road traffic state. Therefore, the calculation method for LIT caused by the transmission of traffic flow from road  $j$  starting from time  $\tau_j$  to road  $i$  was:

$$t_{j,i}(\tau_j, 1) = \min\{\tau | L_v(\tau_j, \tau) \geq L_{j,i}\} \quad (5)$$

where  $t_{j,i}(\tau_j, 1)$  represents the end time of the impact of adjacent road  $j$  on road  $i$  from time  $\tau_j$ .  $L_{j,i}$  is the distance between the two roads, and  $L_{j,i} = l_j + l_i$ ,  $l_j$  and  $l_i$  are the lengths of

roads  $j$  and  $i$ , respectively.  $L_v(\tau_j, \tau)$  is the transmission distance of the traffic flow from time  $\tau_j$  to time  $\tau$ , calculated as:

$$L_v(\tau_j, \tau) = \int_{\tau_j}^{\tau_1} v_j(t) dt + \int_{\tau_1}^{\tau} v_i(t) dt \quad (6)$$

where  $\tau_1$  represents the moment at which all the traffic flows out of road  $j$ .  $v_i$  is the average traffic speed of road  $i$ .

Second, according to LWR theory, the problem of solving the traffic state of two roads or two locations satisfies the condition of the Riemann problem [40]. When the traffic density of the road changes, a traffic wave similar to a “shock wave” will be generated and transmitted between the roads. There will be a gradual change of traffic state. Traffic wave theory provides a good explanation for the formation and development of the macrotraffic state of roads. For example, Refs. [41–43] carried out relevant research on road congestion, which estimates the queue length by predicting the propagation speed of traffic waves and then characterizes the traffic state of the road. Inspired by this, this paper proposed a method to calculate the influence time between different roads through the transmission of traffic waves.

Considering the directivity of traffic wave transmission, the positive direction of the traffic wave was specified as the same as the direction of its traffic flow in this work. If the traffic wave disappears or reverses direction in the entire transmission process, the transmission of the current traffic wave is considered to be completed.

Therefore, the calculation method of LIT caused by the transmission of traffic waves from road  $j$  starting from time  $\tau_j$  to road  $i$  could be:

$$\begin{aligned} t_{j,i}(\tau_j, 2) &= \min\{\tau_1, \tau_2\} \\ \tau_1 &= \min\{\tau | L_\omega(\tau_j, \tau) \geq L_{j,i}\} \\ \tau_2 &= \min\{\tau | \omega(\tau)\omega(\tau - 1) \leq 0\} \end{aligned} \quad (7)$$

where  $L_\omega(\tau_j, \tau)$  is the transmission distance of the traffic wave from time  $\tau_j$  to time  $\tau$ , which can be expressed as

$$L_\omega(\tau_j, \tau) = \int_{\tau_j}^{\tau} \omega(t) dt \quad (8)$$

where  $\omega$  is the speed of the traffic wave. According to the classical first-order steady-state LWR model, the traffic wave speed can be calculated as:

$$\omega = \frac{q_j - q_i}{k_j - k_i} \quad (9)$$

where  $q_j$  and  $q_i$  are the traffic flows of roads  $j$  and  $i$ , respectively, and  $k_j$  and  $k_i$  are their traffic densities.

However, when considering the possibility of traffic flow interruption between two roads in the transmission of traffic waves, it is necessary to modify Formula (8).

- (1) If the traffic wave is not transmitted out of road  $j$  when the interruption occurs, the transmission will terminate.  $L_\omega(\tau_j, \tau)$  can be calculated as:

$$L_\omega(\tau_j, \tau) = \int_{\tau_j}^{\tau_{end}} \omega_j(t) dt \quad \tau_{end} = \min(\tau_0, \tau_2) \quad (10)$$

where  $\tau_0$  is the moment when the traffic flow between roads becomes 0.  $\tau_2$  is the moment when the traffic wave passes out of road j.  $\omega_j$  is the wave speed on road j. Due to the interruption of the traffic flow, the LIT is  $t_{j,i}(\tau_j, 2) = \tau_{end}$ .

- (2) If the traffic wave has passed through road j but not through road i when the interruption occurs, the whole process will be divided into two stages. Before the interruption, the traffic wave is transmitted according to speed  $\omega_j$ . After the interruption, road j is equivalent to a road with a traffic flow and density of 0. Therefore,  $L_\omega(\tau_j, \tau)$  at this stage can be calculated as:

$$L_\omega(\tau_j, \tau) = \int_{\tau_j}^{\tau_0} \omega_j(t) dt + \int_{\tau_0}^{\tau} \omega_i(t) dt \quad (11)$$

where  $\tau_0 \geq \tau_2$ . According to the LWR model, the speed is calculated as:

$$\omega_i(x) = \frac{|q_i - 0|}{|k_i - 0|} \quad (12)$$

where  $\omega_i$  is the transmission speed of the traffic wave on road i.

Finally, note that for the upstream road, both the traffic flows and traffic waves may enter the downstream road. Therefore, combining Formulas (5) and (7), the LIT starting at time  $\tau_j$  can be calculated as follows:

- (1) When the TR is the downstream road,

$$t_{j,i}(\tau_j) = \min\{t_{j,i}(\tau_j, 1), t_{j,i}(\tau_j, 2)\} \quad (13)$$

- (2) When the TR is the upstream road,

$$t_{j,i}(\tau_j) = t_{j,i}(\tau_j, 2) \quad (14)$$

Therefore, the TCIT of period  $(\tau_{j,1}, \tau_{j,n})$  can be calculated as follows:

$$\xi_{j,i}(\tau_{j,1}) = \max(t_{j,i}(t)), t \in (\tau_{j,1} \quad \tau_{j,2} \quad \cdots \quad \tau_{j,n}) \quad (15)$$

### 3.2. Traffic Correlation Strength and Its Calculation Method

As shown in Figure 2, the multibranch structure with different guidance directions determines that the intersection has the function of diversion and confluence [44]. The traffic flow of different upstream roads to the downstream will be different. There is no doubt that the most direct influence of this characteristic is that different roads may have different impacts on the same road. Therefore, it is necessary to consider this when calculating the STC of different roads on the target road. To this end, first, the instantaneous correlation strength between two roads is calculated. Then, on this basis, the traffic correlation strength (TCS) between the two STRs in the whole time series is calculated.

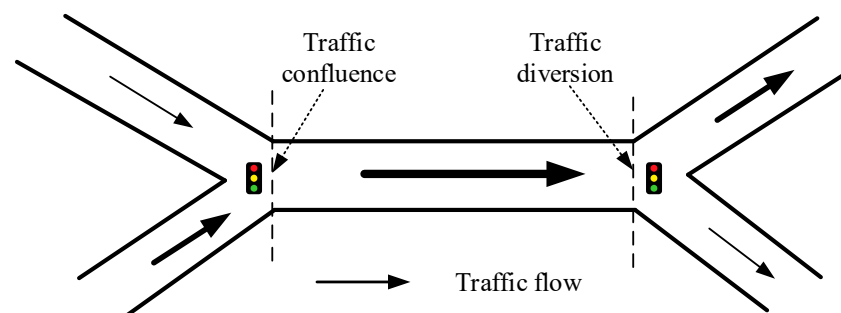


Figure 2. Schematic diagram of the diversion and confluence mechanism of urban road intersections.



### 3.2.1. Instantaneous Correlation Strength

In a certain period, as more traffic flows are transferred between two roads, the relationship between them will be more significant. To measure the influence intensity between two roads, this paper introduces the traffic flow ratio to determine the correlation strength between the roads. The traffic flow ratio is calculated as:

$$\alpha_{j,i}(\tau) = \begin{cases} \frac{q_{j,i}(\tau)}{q_i(\tau)} & j \text{ is the upstream} \\ \frac{q_{i,j}(\tau)}{q_j(\tau)} & i \text{ is the upstream} \end{cases} \quad (16)$$

where  $\alpha_{j,i}(\tau)$  is the traffic flow ratio at time  $\tau$  between road  $i$  and road  $j$ .  $q_{j,i}(\tau)$  is the traffic flow from road  $j$  to road  $i$  at time  $\tau$ , and  $q_i(\tau)$  is the total traffic flow at time  $\tau$  to road  $i$  from its adjacent roads.

Due to the different relative positions of the road, the direction of the influence between two roads will be different. Therefore, this paper evaluates whether road  $j$  influences road  $i$  according to its influence direction. Based on the different ways of influence, the influence direction will be calculated as:

$$\beta_{j,i}(\tau) = \begin{cases} \begin{cases} 1 & \text{if } \alpha_{j,i}(\tau) \neq 0 \\ 0 & \text{if } \alpha_{j,i}(\tau) = 0 \end{cases} & j \text{ is the upstream} \\ \begin{cases} 1 & \text{if } \omega_j(\tau) > 0 \\ 0 & \text{if } \omega_j(\tau) \leq 0 \end{cases} & i \text{ is the upstream} \end{cases} \quad (17)$$

where  $\beta_{j,i}(\tau)$  is used to assess whether road  $j$  has a direct impact on road  $i$  at time  $\tau$ .

The product  $\delta_{j,i}(\tau) = \alpha_{j,i}(\tau)\beta_{j,i}(\tau)$  is considered to measure the instantaneous correlation strength (ICS) between two adjacent roads at time  $\tau$ . Generally, a larger value of  $\delta_{j,i}(\tau)$  corresponds to a stronger correlation.

Note that the correlation between two roads will not change significantly in two adjacent periods because the traffic flow is continuous. Thus,  $\alpha_{j,i}(\tau) = 0$  does not mean that no correlation exists between the roads. For example, even if there is no traffic flow between the congested road and its upstream road, a strong correlation may exist between them. Thus, it is necessary to consider the traffic state of roads when calculating the correlation strength.  $\eta_i(\tau)$  is adopted to express the state of road  $i$  at sampling time  $\tau$ .

$$\eta_i(\tau) = \begin{cases} 1 & \text{if } q_i(\tau) = 0 \\ 2 & \text{if } 0 < q_i(\tau) \& v_i(\tau) \neq 0 \\ 3 & \text{if } v_i(\tau) = 0 \end{cases} \quad (18)$$

Thus, without losing generality, supposing traffic flows from road  $j$  to road  $i$ , the calculation of instantaneous correlation strength could then be modified, as shown in Listing 1.

**Listing 1.** Progress for the calculation of the Instantaneous Correlation Strength.

---

```

if  $q_{j,i}(\tau) = 0$  //if there is no traffic flow between two adjacent roads
  if  $\eta_i(\tau) = 1$  or  $\eta_j(\tau) = 1$  //if at least one of them is empty
     $\delta_{j,i}(\tau) = 0$  //then the ICS between them is 0
  else //if both roads are not empty
    if  $q_i(\tau) = 0$  //if the upstream road i is congested
      if  $q_i(\tau - 1) = 0$  //and it is also in the congestion state at the previous time step
         $\delta_{j,i}(\tau) = \delta_{prev}$  //then the ICS remains the same value during the congestion period
      else //Otherwise, if the upstream road i was not congested at the
        //previous time step
         $\delta_{j,i}(\tau) = \alpha_{j,i}(\tau - 1)\beta_{j,i}(\tau - 1)$  //calculate the ICS according to the data in the
        //previous sampling time
         $\delta_{prev} = \alpha_{j,i}(\tau)\beta_{j,i}(\tau)$  //update the ICS value during this congestion period
        //for later use
      end
    else //if road i is not congested
       $\delta_{j,i}(\tau) = \alpha_{j,i}(\tau)\beta_{j,i}(\tau)$  //calculate the current ICS directly
    end
  end
else //if there exists traffic flow between two roads
   $\delta_{j,i}(\tau) = \alpha_{j,i}(\tau)\beta_{j,i}(\tau)$  //calculate the current ICS directly
end

```

---

### 3.2.2. TCS Calculation Method

The instantaneous correlation strength only reflects the correlation strength of two roads at a certain moment, while the calculation of STC is for the whole time series. In addition, with the continuous updating of time, the traffic state in the specified STR has different effects on that of the target road with different time periods. Therefore, it is necessary to consider the influence of time delay when calculating the TCS.

Assume that the TCIT from the traffic state of road  $j$  with time period  $(\tau_{j,1}, \tau_{j,n})$  to that of road  $i$  in time period  $(\tau_{i,1}, \tau_{i,n})$  is  $\xi_{j,i}(\tau_{j,1})$ . As the time period of road  $i$  slowly exceeds  $\xi_{j,i}(\tau_{j,1})$ , with the influx of the traffic flow from other roads, the TCS owing to road  $j$  with time period  $(\tau_{j,1}, \tau_{j,n})$  will gradually decrease. Therefore, the TCS of road  $j$  on  $i$  in different periods could be determined as:

$$\text{If } \tau_{i,n} \leq \xi_{j,i}(\tau_{j,1}), \quad f_{j,i}(\tau_{j,1}, \tau_{i,1}) = \lambda_j \quad (19)$$

Otherwise,

$$f_{j,i}(\tau_{j,1}, \tau_{i,1}) = \begin{cases} \lambda_j \gamma_1 \gamma_2 & \text{if } \xi_{j,i}(\tau_{j,1}) - \tau_{i,1} > 0 \\ 0 & \text{otherwise} \end{cases} \quad (20)$$

$$\lambda_j = \frac{1}{(\tau_{j,n} - \tau_{j,1})} \sum_{\tau=\tau_{j,1}}^{\tau_{j,n}} \delta_{j,i}(\tau), \quad \gamma_1 = \frac{\xi_{j,i}(\tau_{j,1}) - \tau_{i,1}}{\xi_{j,i}(\tau_{j,1}) - \tau_{j,1}}, \quad \gamma_2 = \frac{\sum_{\tau=\tau_{i,1}}^{\xi_{j,i}(\tau_{j,1})} \delta_{j,i}(\tau)}{\sum_{\tau=\tau_{i,1}}^{\tau_{i,n}} \delta_{j,i}(\tau)} \quad (21)$$

where  $f_{j,i}(\tau_{j,1}, \tau_{i,1})$  is the TCS between the traffic state of road  $j$ , with  $\tau_{j,1}$  as the starting time, and the traffic state of road  $i$ , with  $\tau_{i,1}$  as the starting time.  $\lambda_j$  is the initial influence intensity within the TCIT of road  $j$ .  $\gamma_1$  and  $\gamma_2$  represent the proportion of time and strength for the influence of road  $j$  on road  $i$ . These two parameters reflect the impact of the time delay on the influence intensity. A larger traffic flow ratio corresponds to a greater initial influence intensity. A greater time delay corresponds to a smaller influence intensity.

#### 4. STC Calculation Method for Urban Roads

Compared with adjacent roads with direct impact, there is no direct correlation between the traffic statuses of two non-adjacent roads. In view of this, we creatively propose a method to calculate the correlation between roads according to road adjacency.

First, according to the analysis presented above, the calculation method of the STC between two adjacent roads was established in Section 4.1. Then, based on this, the calculation method between two non-adjacent roads is established in Section 4.2. Finally, we summarize the overall calculation process in Section 4.3.

##### 4.1. Calculation Method for Adjacent Roads

Assume that the traffic data of the TR (road  $i$ ) and its adjacent road  $j$  can be expressed as follows:

$$\begin{aligned} d_i(\tau_{i,1}) &= \begin{pmatrix} x_{\tau_{i,1}} & x_{\tau_{i,2}} & \cdots & x_{\tau_{i,n}} \end{pmatrix} \\ d_j(\tau_{j,1}) &= \begin{pmatrix} y_{\tau_{j,1}} & y_{\tau_{j,2}} & \cdots & y_{\tau_{j,n}} \end{pmatrix} \end{aligned} \quad (22)$$

where  $x_{\tau_{i,1}}$  and  $y_{\tau_{j,1}}$  are the traffic data, for example, the traffic speed, with times of  $\tau_{i,1}$  and  $\tau_{j,1}$ , respectively.  $d_j(\tau_{j,1})$  and  $d_i(\tau_{i,1})$  are the traffic data sequences of roads  $j$  and  $i$ . Here, without loss of generality, assume that  $\tau_{j,1} \leq \tau_{i,1}$ . The STC between the roads can be calculated as:

$$\begin{aligned} r_{j,i}(\tau_{j,1}, \tau_{i,1}) &= \rho_{j,i}(\tau_{j,1}, \tau_{i,1}) f_{j,i}(\tau_{j,1}, \tau_{i,1}) \\ \rho_{j,i}(\tau_{j,1}, \tau_{i,1}) &= \frac{E[(d_j(\tau_{j,1}) - \mu_j)(d_i(\tau_{i,1}) - \mu_i)]}{\sigma_j \sigma_i} \end{aligned} \quad (23)$$

where  $\mu_i$  and  $\sigma_i$  are the average values and standard deviations, respectively, of data  $d_i(\tau_{i,1})$ .  $r_{j,i}(\tau_{j,1}, \tau_{i,1})$  is the correlation between the traffic state of road  $j$  starting from time  $\tau_{j,1}$  and that of road  $i$  starting from time  $\tau_{i,1}$ .

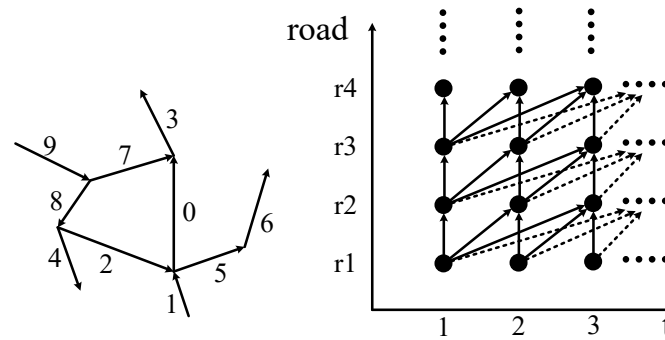
Note that the calculation includes two parts:  $\rho_{j,i}(\tau_{j,1}, \tau_{i,1})$  and  $f_{j,i}(\tau_{j,1}, \tau_{i,1})$ . The former is derived from the cross-correlation function (CCF) of statistical theory, which reveals the similarity between two traffic data and is used to measure the consistency of the change trend of two traffic statuses. The latter reveals the degree of internal correlation between the traffic states.

Therefore, this calculation method considers not only the similarity between the traffic information (as an external performance feature) but also the impact of different traffic flow proportions and time delays (as an internal feature). It introduces physical meaning to the calculation of STC between adjacent roads.

##### 4.2. Calculation Method for Non-Adjacent Roads

Different from adjacent roads, there is no direct influence between non-adjacent roads. However, due to the connectivity of roads, an exchange of traffic information might occur between any two non-adjacent roads [45]. The traffic flow between these roads is realized by the transmission path between the roads [46]. Hence, the correlation between each pair of adjacent roads in the path will affect the overall correlation between the two non-adjacent roads. The traditional calculation method does not consider the physical structure of the road network and ignores the actual transmission characteristics of traffic flow. As a result, regardless of how far away the two roads are, they are directly adjacent by default in the whole calculation process.

For a deeper analysis, a schematic of the road network topology was constructed in Figure 3. The left diagram shows a road network structure, and the right diagram depicts the relationship of the STRs. The abscissa and ordinate represent different starting times and roads, respectively. The arrows represent the connection between two STRs. As shown in the right subgraph of Figure 3, each road will have an impact on the current and future traffic states of its adjacent roads. The specific impact time is calculated according to different traffic conditions.



**Figure 3.** Schematic of the correlation of the road network.

Through analysis, it is worth noting that the influence between traffic states of non-adjacent roads could be realized on the basis of adjacent roads in the path between them. For example, if there is a traffic jam on road 6 in Figure 3, the traffic state of road 5 will change first, and the change of traffic state of road 5 will then affect the traffic state of road 1 or road 2. Therefore, based on the continuity in time and transitivity in space, this paper proposed a calculation method for the correlation between any two non-adjacent roads by considering the transitivity in the path between them. Without loss of generality, we take the correlation calculation of road 9 at time 1 to the TR (road 0) at time 3 as an example to illustrate the process.

Two main influence paths exist, as shown in Figure 4A,B: the left figures show the path extracted from the road network, and the right figures show the transmission diagram of correlation. Each dot represents an STR, and the arrow represents the correlation between the two STRs. Even if the path is determined, many transmission paths exist for the correlation between the STRs due to the existence of time delays. Taking the calculation of  $r_{9,0}(1,3)$  in Figure 4B as an example, there are six influence transmission paths:

$$\begin{aligned}
 & r_{9,7}(1,1) \rightarrow r_{7,3}(1,1) \rightarrow r_{3,0}(1,3) \\
 & r_{9,7}(1,1) \rightarrow r_{7,3}(1,2) \rightarrow r_{3,0}(2,3) \\
 & r_{9,7}(1,1) \rightarrow r_{7,3}(1,3) \rightarrow r_{3,0}(3,3) \\
 & r_{9,7}(1,2) \rightarrow r_{7,3}(2,2) \rightarrow r_{3,0}(2,3) \\
 & r_{9,7}(1,2) \rightarrow r_{7,3}(2,3) \rightarrow r_{3,0}(3,3) \\
 & r_{9,7}(1,3) \rightarrow r_{7,3}(3,3) \rightarrow r_{3,0}(3,3)
 \end{aligned} \tag{24}$$

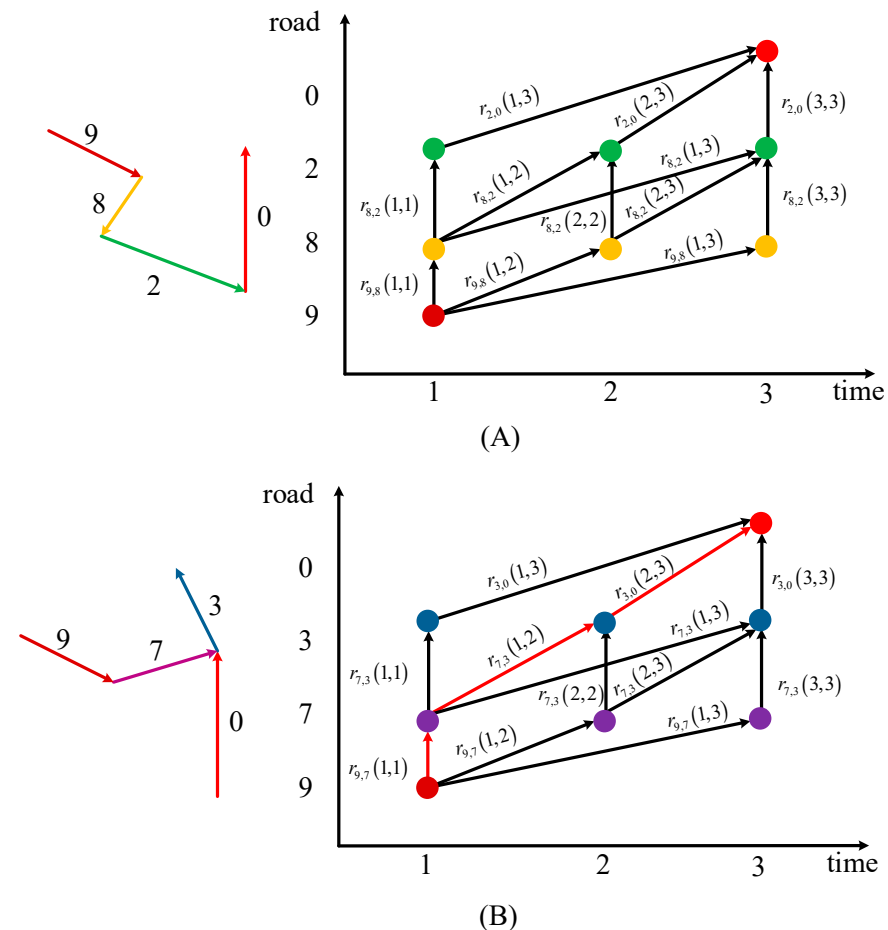
Take the path of the red arrows in Figure 4B as an example, that is,  $r_{9,7}(1,1) \rightarrow r_{7,3}(1,2) \rightarrow r_{3,0}(2,3)$ . Considering the transfer characteristics of traffic, a calculation method for the correlation between road 1 and road 3 was proposed as:

$$r_{9,0}(1,3) = r_{9,7}(1,1)r_{7,3}(1,2)r_{3,0}(2,3) \tag{25}$$

As multiple correlations may be calculated using different transmission paths, to express the correlation more clearly, this paper considers the maximum value as the correlation between the roads. To automatically search the possible path and calculate this maximum value, this paper builds the algorithm based on the idea of the dynamic programming (DP) algorithm. Thus, the correlation calculation between non-adjacent roads can be transformed into an iterative calculation of the STC between adjacent roads. The objective function of this algorithm is:

$$\begin{aligned}
 J(g_f, \tau_f, \tau_i) &= \max \left( \left( J(g, \tau_k, \tau_i) r_{g_f, g}(\tau_f, \tau_k) \right) \right) \\
 \tau_f &\leq \tau_k \leq \tau_i
 \end{aligned} \tag{26}$$

where  $J(g_f, \tau_f, \tau_i)$  is the STC between road  $g_f$  starting at time  $\tau_f$  and the TR starting at time  $\tau_i$ .  $r_{g_f, g}(\tau_f, \tau_k)$  is the STC between road  $g_f$  and its adjacent road  $g$  starting with  $\tau_f$  and  $\tau_k$ , respectively.



**Figure 4.** Two different transmission paths of road 9 to the TR (road 0) extracted from the road network.

This formula incorporates the spatial characteristics to calculate the correlation between the non-adjacent roads. Therefore, compared with those methods that directly calculate the similarity of two traffic state sequences without considering the adjacency of roads, this formula reflects the transmission characteristics of the traffic flow on the road.

#### 4.3. Overall Calculation Progress of STC

Considering the connectivity characteristics of the road network, infinite noneffective paths may be generated between two non-adjacent roads. However, most of the transmitted correlations may be extremely low or even 0. Therefore, to avoid the generation of redundant associated paths, the overall calculation progress of STC with the TR as the starting road and the adjacent relationship as the basis was built.

The pseudocode for the overall calculation of STC for the urban road network is shown in Listing 2. Figure 5 shows the process flow. The blue line represents the TR. The red lines represent the roads that need to be considered in this step. The black lines are the roads that were calculated in the previous step and will not be calculated in subsequent steps.

**Listing 2.** Pseudo code for the overall calculation of STC. Note: where  $g_0$  is the TR, and  $g_j$  is the current road.  $G_{Link1}$  is the set of adjacent roads of road  $g_j$ .  $F$  is the set of historical roads to prevent repeated conflicts in the correlation calculation.  $G$  records the next calculated road, which is updated every iteration.  $\tau_0$  is the starting time of the TR. Based on this analysis, the correlation between different roads and the TR can be obtained.

**Initialisation:**  $G\{1\}\{1\} = g_0, num = 0, flag = 1, F\{1\} = \emptyset J(g_0, \tau_0, \tau_0) = 1$

**Iterative computation**

**while (flag)**

$num = num + 1$

//record the number of iterations

$k = length(G\{num\})$

//calculates the number of the current roads involved in this iteration

$n = 0$

**for**  $j = 1 : k$

$g_j = G\{num\}\{j\}$

// $g_j$  is the current road

$G_{Link1} \cap F\{num\}\{j\} = \emptyset$

//remove the historical roads from its adjacent roads

$G_{Link1} = (g_1, g_2, \dots, g_m)$

$m = length(G_{Link1})$

//calculate the number of the remaining adjacent roads

**for**  $f = 1 : m$

//calculate the correlation between the remaining adjacent roads and the TR

**for**  $\tau_s = 1 : \tau_0$

// $g_f$  is one of the adjacent roads,  $\tau_s$  is the starting time of road  $g_f$

$S(g_f, \tau_s, \tau_0) = \max\{J(g_j, \tau_k, \tau_0)r_{g_f, g_j}(\tau_s, \tau_k) \mid \tau_s \leq \tau_k \leq \tau_0\}$  //calculated the STC between  $g_f$  and

//the TR. See description of Formula (26) for details.

**if**  $S(g_f, \tau_s, \tau_0) > J(g_f, \tau_s, \tau_0)$

$J(g_f, \tau_s, \tau_0) = S(g_f, \tau_s, \tau_0)$

//select the maximum value as the influence from  $g_f$  with

//starting time  $\tau_s$  to the TR with starting time  $\tau_0$

**end**

**end**

$n = n + 1$

$F\{num + 1\}\{n\} = [F\{num\}\{j\}, g]$  //update the collection of historical roads for the adjacent roads

$G\{num + 1\}\{n\} = G_{Link1}(f)$

//update the new collection of current roads

**end**

**end**

**if**  $G\{num + 1\} = \emptyset$

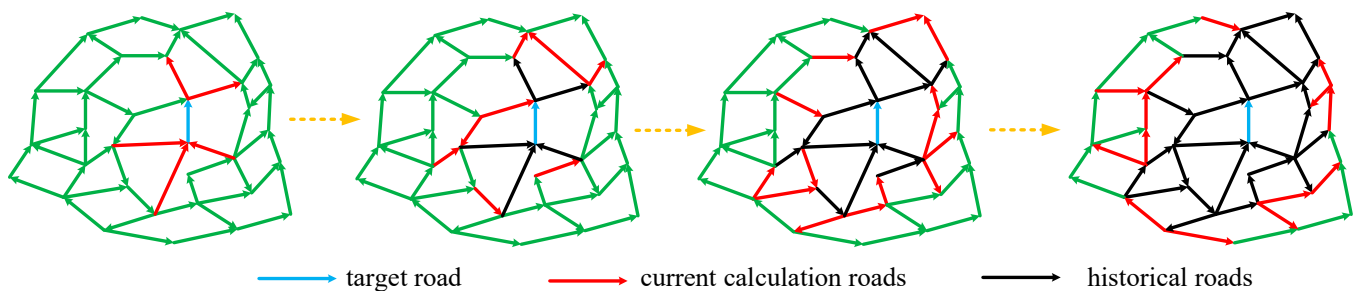
//if there is no new calculation

$flag = 0$

//the iterative calculation process will be terminated

**end**

**end**

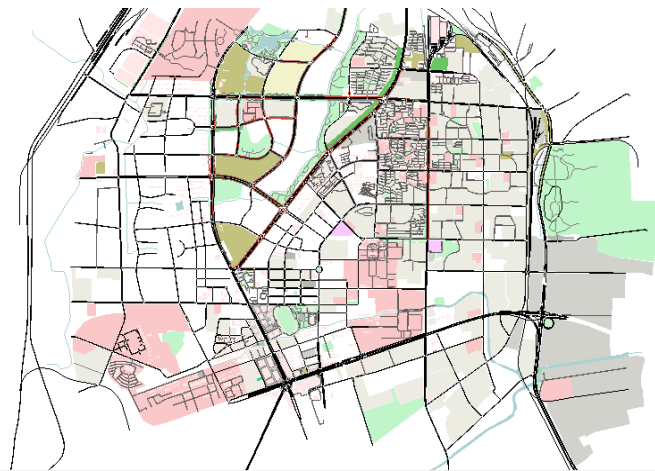


**Figure 5.** Calculation process of the STC between different roads and the TR.

## 5. Results and Discussion

To verify the rationality of the STC calculation method proposed in this paper, a comparative analysis is performed in this section. The verification data are obtained from the operation results of the traffic simulation software Simulation of Urban Mobility (SUMO). The SUMO software, developed by the German Aerospace Centre, is a popular software designed to process large road networks. SUMO can scientifically realize the overall simulation of urban road network traffic with high reliability. The operation map shows an area of Shanghai, China, as shown in Figure 6, and the simulation time is 10,000 s. This map was downloaded from the OpenStreetMap website, including 2167 roads and 5814

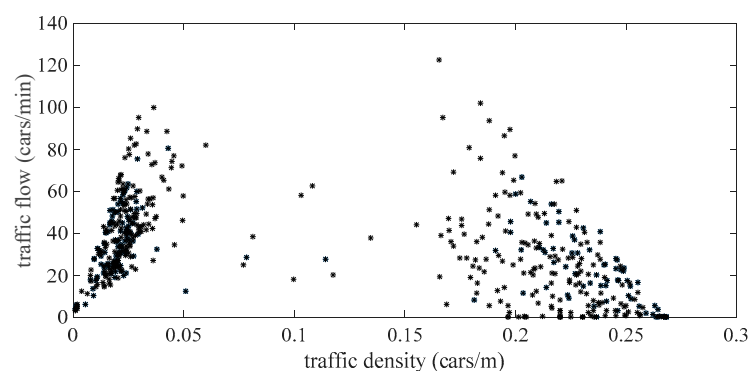
intersections with traffic signals. In fact, many advanced simulation models are integrated into SUMO software. Most of the traffic control rules in this paper adopt the default settings of SUMO. The shortest distance is used in vehicle travel path planning. Car following model is based on the Krauss model proposed by German scholar Krauss. The maximum acceleration and maximum deceleration of the simulation vehicles are respectively set by SUMO, that is,  $2.6 \text{ m/s}^2$  and  $-4.5 \text{ m/s}^2$ . The simulation record parameters include: system time, vehicle speed, location, road network information and traffic light status.



**Figure 6.** Map of the simulation calculation.

The data employed in macro traffic research are generally data with low and medium sampling frequencies. Thus, to be closer to the real data collection conditions, the research organizes the simulation data with an interval time of 30 s for calculation.

The relationship between traffic density and traffic flow of different roads obtained by simulation is shown in Figure 7. The abscissa and ordinate are the traffic density and traffic flow of different roads, respectively. The distribution of the traffic data obtained by simulation is basically consistent with the MFD diagram [21], which conforms to the characteristics of the macro traffic operation law and can provide support for comparative verification.



**Figure 7.** Relationship between traffic density and traffic flow simulated by SUMO in this paper.

The length of the data sequence for the STC calculation has is 10 in this paper. To verify the proposed method, two classical models, the original CCF model and the improved CCF

model, are utilized to calculate the correlation for comparison. The calculation formulas of the improved CCF is as follows:

$$\rho_{XY}(s) = \frac{\sum_{t=1}^T (x_i(t) - \bar{x}_i) (L^k x_i(t+s) - \overline{L^k x_i})}{\sqrt{\sum_{t=1}^T (x_i(t) - \bar{x}_i)^2 \sum_{t=1}^T (L^k x_i(t+s) - \overline{L^k x_i})^2}} \quad (27)$$

$$L^0 x_i(t) = x_i(t), L^l x_i(t) = \sum_{j=1}^N \omega_{ij}^l x_j(t), \sum_{j=1}^N \omega_{ij}^l = 1$$

where  $t$  is the statistical period and  $s$  is the time delay.  $L^k$  is the  $k$ -order spatial delay operator of the road network.  $L^k x_i$  is the comprehensive observation value of traffic with all  $l$ -order adjacent roads of the TR.  $\overline{L^k x_i}$  is the their average value.  $\omega_{ij}^l$  is the weight of  $l$ -order adjacent roads calculated by distance.

For the convenience of expression, these two models are called CCF1 and CCF2, and the model proposed in this paper is the dynamic correlation function (DCF) in the following content.

To provide a more reliable comparison and highlight the calculation characteristics of different models, this paper chooses three aspects to carry out comparative experiments.

First, a local road network to demonstrate the overall calculation results of the correlation between different roads and the TR in Section 5.1 was extracted. This experiment can highlight the calculation characteristics of different models as a whole level.

Second, the results of different models for adjacent and non-adjacent roads is compared in Sections 5.2 and 5.3, respectively, to demonstrate the scientific value and rationality of the proposed model.

Finally, multiple analysis of the results of 180 different road networks was performed. The indexes involved in these comparative tests come from the summary of the time delay (Section 5.4), road network topology (Section 5.5) and operation characteristics of roads (Section 5.6), which can indirectly verify the rationality of different models.

### 5.1. Overall Calculation Results of a Road Network

The extracted road network for this calculation from the map shown in Figure 6, its topological structure is shown in Figure 8, and road 0 is the TR. The average traffic speed data of the TR with a time period of (250–259) was considered as the target STR and calculated the STC with other roads with different time delays. Figures 9 and 10 show the calculation results of CCF1 and CCF2, respectively. Figure 11 shows the calculation results of the proposed DCF method.

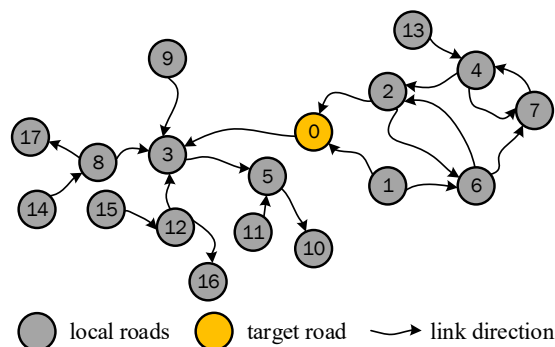


Figure 8. Structure of the road network considered in the test.



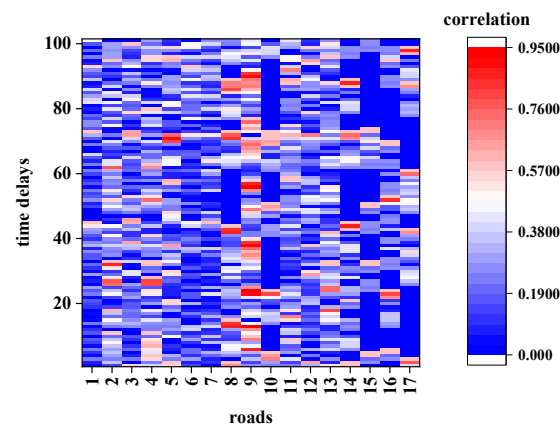


Figure 9. Correlation determined using the original CCF.

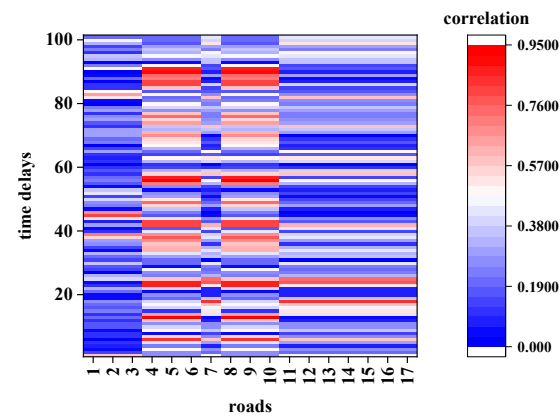


Figure 10. Correlation determined using the improved CCF.

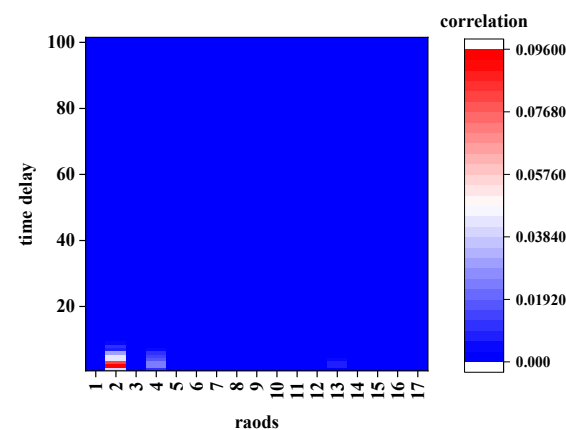


Figure 11. Correlation determined using the DCF.

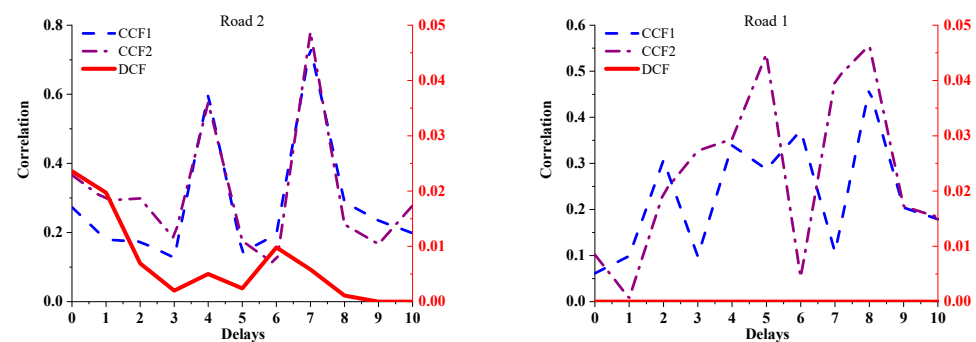
According to Figures 9 and 10, the results calculated using the first two methods exhibit the following characteristics:

- In terms of space, the traffic data in the specified period (250–259) of the TR are related to all the roads, and the correlation of different roads is irregular.
- In terms of time, no matter how much the time delays, the other roads could have a significant impact on the traffic state of road 0 in the time period (250–259). Even when the time delays extend over more than 90 sampling intervals (2700 s or 45 min), a strong correlation could be observed, which is not consistent with the actual situation. However, the calculation result of the DCF shown in Figure 11 shows that:



**Table 2.** Traffic correlation strength between road 2 and TR with different time delays.

Time Delays	$\xi_{2,0}(240)$	$\lambda_2$	$\gamma_1$	$\gamma_2$	$f_{2,0}$
0			1	1	0.1
1			0.9	1	0.09
2			0.8	0.5	0.04
3			0.7	0.5	0.035
4			0.6	0.5	0.03
5	249	0.1	0.5	0.33	0.0165
6			0.4	0.33	0.0132
7			0.3	0.33	0.0099
8			0.2	0.25	0.005
9			0.1	0	0
10			0	0	0

**Figure 13.** Comparison of correlation between the two roads, as determined using the three algorithms.

According to Tables 1 and 2, the following contents can be obtained:

- No vehicles enter the TR during the period (240–249) of road 1. Road 1 does not affect the TR in the period (240–249), as shown in Table 1.
- With the increase in the time delay, the proportion of time and strength that the TR is affected by road 2 in period (240–249) decreases gradually.
- The TCS between TR and road 2 in period (240–249) decreases with an increase in the time delay.

From the calculation results shown in Figure 13, it can be seen that:

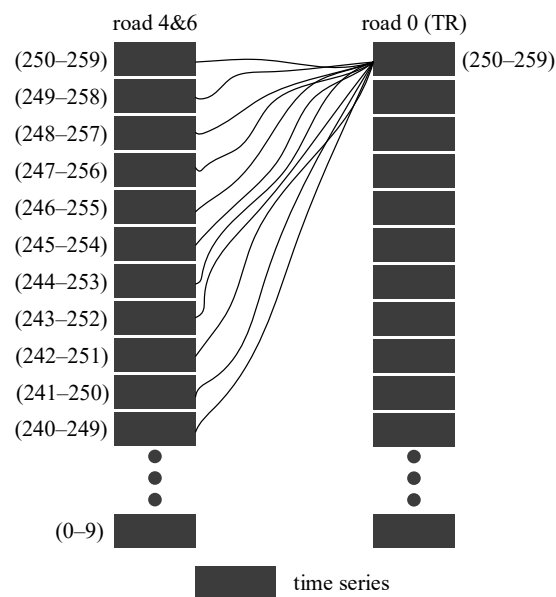
- The results of the proposed DCF method are consistent with the content above.
- However, the calculated results of CCF1 and CCF2 do not show any regularity.

Thus, the correlation between two adjacent roads determined by the proposed improved DCF model is more closely related to the actual situation.

### 5.3. Calculation Results between a Pair of Non-Adjacent Roads

To highlight the differences of the three models in the correlation calculation between non-adjacent roads, 11 segments (240–249), (241–250), . . . , (250–259) were selected for road 4 and road 6 in Figure 8 for the calculation. The time period of the TR is (250–259). The time-series position of roads participating in this calculation is shown in Figure 14.

To make a reliable comparison, a statistical analysis on the traffic flow of roads 4 and 6 with their adjacent roads in the time period (240–259) was performed. The traffic flow ratios are shown in Tables 3 and 4.



**Figure 14.** Schematic diagram of the time series between roads participating in the STC calculation and target roads.

**Table 3.** Flow ratio (%) of roads 4 and 6 to road 2.

Time	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259
road 4	0	0	0	0	0	0	0	0	100	0	0	100	0	100	100	0	100	100	0	0
road 6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

**Table 4.** Flow ratio (%) of roads 1 and 2 to road 6.

Time	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259
road 1	100	0	0	100	0	100	0	0	100	0	0	100	0	0	100	0	0	100	0	0
road 2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

After analysing the contents of Tables 1, 3 and 4, it can be seen that:

- The traffic flow of road 4 in time period (240–259) reaches road 0 through road 2;
- There is no vehicle flow exchange between road 6 and road 2 during this period, but there are some vehicles flowing from road 1 to road 6. However, there is no traffic information exchange between road 1 and road 0 during this period. Therefore, road 6 cannot affect the traffic state of road 0 during this period.

The calculation results of road 4 and road 6 by the three models are shown in Figure 15. From the result of Figure 15, it can be seen that:

- The correlation of road 4 calculated by DCF first increases and later decreases to 0 as the time delay increases, and it is less than the influence of road 2 with the same or fewer delay conditions.
- The correlation of road 6 calculated by DCF remains at 0 within different time delays.
- The results of road 4 and road 6 calculated by CCF1 and CCF2 show irregular changes with time delay. Moreover, the calculation results show that the correlation between these roads and the TR has no relationship with the correlation between road 2 and the TR.

However, the correlation between non-adjacent roads will be affected by the correlation of adjacent roads between them, as analysed. As road 2 is directly adjacent to the target road (road 0), it will affect the target road more easily than road 4 and road 6. Therefore, the calculation results of CCF1 and CCF2 are obviously inconsistent with the actual situation.

In comparison, the results of the proposed DCF model are more reasonable than those of CCF1 and CCF2.

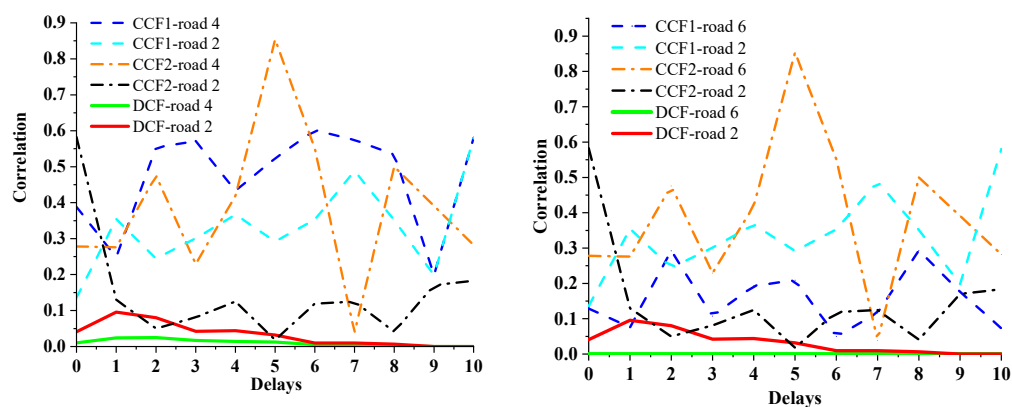


Figure 15. Calculation results for roads 4 and 6.

#### 5.4. Relationship between the Average Correlation and Time Delays

To clarify the relationship between the time delay and correlation, the average correlation with different time delays was calculated as follows:

- (1) Different roads in Figure 6 are selected as the target roads, and then the surrounding road networks centred on these target roads are constructed as the calculation samples.
- (2) The average value of the correlation of all samples under different time delays is calculated as:

$$\mu(k, d) = \frac{1}{n_k} \sum_{j=1}^{n_k} r_{j,k}(d) \quad (28)$$

$$\bar{\mu}(d) = \frac{1}{m} \sum_{i=1}^m \mu(i, d) \quad (29)$$

where  $r_{j,k}(d)$  is the correlation between TR  $k$  and its surrounding road  $j$  with the time delay  $d$ .  $\mu(k, d)$  is the average value of the correlation.  $n_k$  is the number of the surrounding roads.  $\bar{\mu}(d)$  is the average value of  $\mu(k, d)$ , and  $m = 180$  is the total number of target roads. The calculation results are shown in Figure 16.

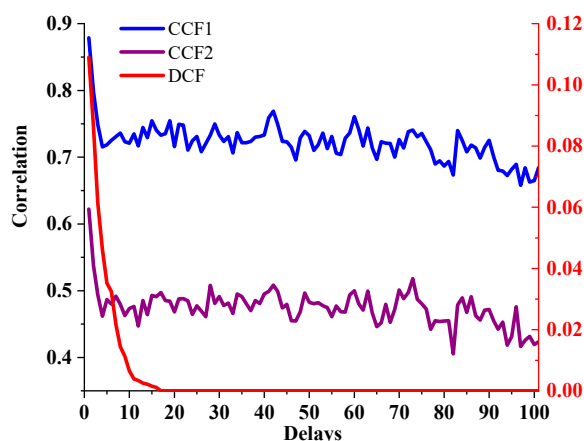


Figure 16. Relationship between the average correlation and time delays.

Figure 16 shows that the results of the CCF-based calculation methods first exhibit a decrease with the increase in the time delay and then still maintain a strong correlation with a considerable time delay from the TR's "specified time period" with a little change.

The proposed DCF model shows that the correlation between the TR and the surrounding roads gradually decreases to 0 with an increase in the time delay.

However, due to the constant update of the traffic flow on the TR, the correlation between the traffic statuses of two different time and space regions (STRs) cannot last forever. Therefore, the result obtained by DCF could be more reasonable than CCF1 and CCF2.

### 5.5. Relationship between the Maximum Correlation and the Importance of Roads

Road network analysis based on complex theory has become an important method in the field of traffic flow research, especially in relation to spatial structure. In complex network theory, the more important the node (road) in the topology structure is, the greater the impact on other nodes (roads) [47]. Therefore, to verify the correlation reliably, the relationship between the correlation calculated by the three models and the importance of roads was analysed. The calculation process is as follows:

- (1) The page rank algorithm (PRA) [48] and signal propagation algorithm (SPA) [49] were used to rank the surrounding roads according to the positions to the TR. The more important the road is, the higher it ranks with a smaller ID.
- (2) Select the maximum correlation of the traffic state between each surrounding road and the TR with a finite number of time delays.

$$\psi_k(j) = \max( r_{j,k}(0) , r_{j,k}(1) , \dots , r_{j,k}(d_{\max}) ) \quad (30)$$

where  $\psi_k(j)$  is the maximum value of the correlation between TR  $k$  and road  $j$  with a time delay from 0 to  $d_{\max}$ ,  $d_{\max} = 30$ .

The relationship between the maximum correlation calculated by the three methods and the ranking of roads are shown in Figure 17.

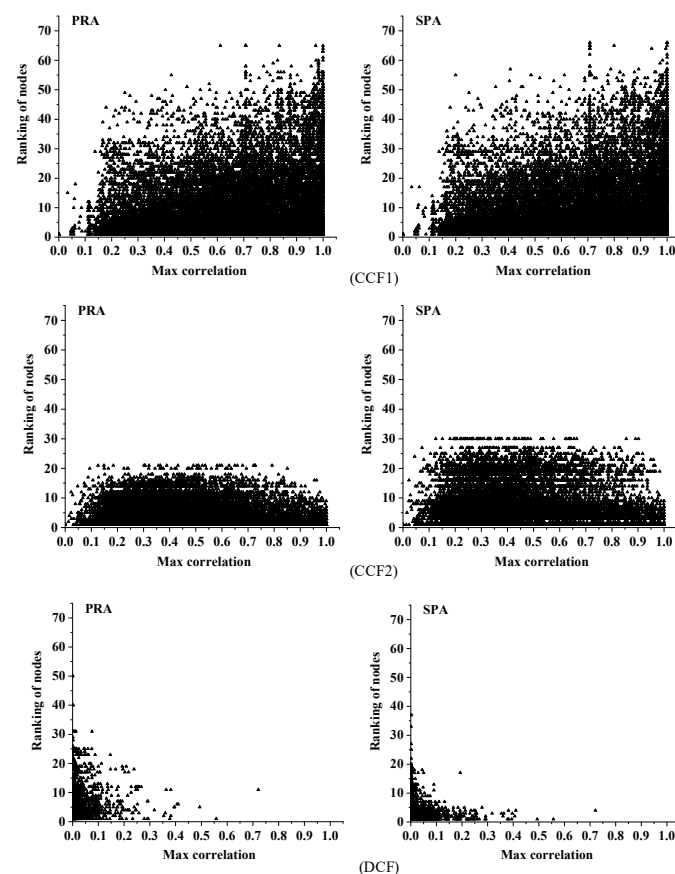


Figure 17. Relationship between the importance of nodes and their maximum relevance to the TR.

In Figure 17, the maximum correlation calculated by the first CCF-based methods has no obvious relationship with the importance of roads. However, the DCF results show that the more important the road is, the more likely the maximum correlation will be.

According to complex network research [48,49], the more important a node is, the greater the impact it will bring when its state changes, and the closer its relationship with the surrounding nodes. Therefore, compared with the former two models, the results of the DCF proposed in this paper are more consistent with the understanding of node importance in complex network theory.

### 5.6. Relationship between the Correlation and Traffic Conditions

Urban traffic networks are complex network systems composed of physical road networks (static characteristics) and traffic demand networks (dynamic characteristics). The existing study has shown that in addition to the static physical structure of the road network, the dynamic operation of traffic flow can also reflect the degree of correlation between different roads [50]. To verify the models from the perspective of dynamic requirements, the relationship between the correlation and operation status of the TR was studied. A parameter named the non-vacant ratio was proposed, which is calculated as:

$$\zeta_i(t) = \frac{1}{n} \sum_{\tau=t}^{t+n-1} \chi_i(\tau), \quad \chi_i(\tau) = \begin{cases} 1 & \text{if } \eta_i(\tau) \neq 1 \\ 0 & \text{if } \eta_i(\tau) = 1 \end{cases} \quad (31)$$

where  $\zeta_i(t)$  is the non-vacant ratio between time  $t$  and  $t + n - 1$  of target road  $i$ .  $\eta_i(\tau)$  is the state of the TR introduced in Section 3. This parameter reflects the degree to which the TR continuously participates in the information exchange of the road network.

The average value of the correlation between the TRs and their surrounding roads under each time delay condition was calculated. Then, the relationship between the average correlation and time delays with different non-vacant ratios of the TRs is shown in Figure 18. The abscissa and ordinate indicate the time delays and average correlation, respectively, and the dotted lines of different colours represent different non-vacant ratios.

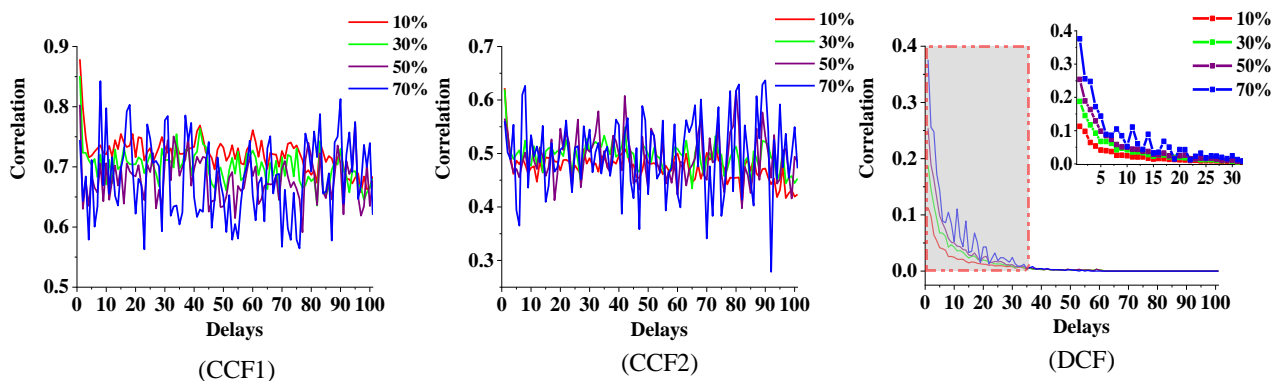


Figure 18. Relationship between the correlation and time delays for different non-vacant ratios.

The calculation results obtained by the two models, CCF1 and CCF2, show that the average correlation between the TR and its surrounding road network is not only irregular with the time delays but also unrelated to the non-vacant ratio. However, the DCF results show that the average correlation decreases with an increase in the time delay and increases with an increase in the non-vacant ratios.

According to the definition and related research, a larger non-vacant ratio in a certain period means that traffic flow exchange between the road and surrounding roads will be more frequent. Therefore, the TR is more likely to be affected by the traffic flow of the surrounding roads, which renders it more relevant to the surrounding road network. Thus, the DCF model proposed in this paper is more reasonable than the former two.

## 6. Conclusions

The calculation of the STC between roads affects not only the selection of the road networks but also the quality of data utilization of each road. Therefore, the provision of more reasonable correlation information for research, such as traffic speed prediction, is of significance. Traditional statistical-based calculation methods do not consider the actual operation regularity of the road network and dynamic characteristics of the traffic data. In addition, they seldom consider the characteristics of urban road network structure. Therefore, these studies have congenital deficiencies when explaining traffic relevance at the physical level.

Therefore, drawing support from the macro traffic model with the advantages of simple calculation and good physical interpretation, combined with the structural characteristics of urban road networks, this paper proposes an STC calculation method based on a traffic dynamic model in this paper for the first time.

Based on the traffic dynamic model (LWR model), the research explores the influence characteristics of urban traffic flow. The parameter of traffic complete influence time (TCIT) is proposed to measure the influence time between different roads for the first time. Another parameter of traffic correlation strength (TCS) was proposed based on the separation and confluence mechanism of intersections to effectively characterize the influence strength of different roads on the same road. Based on these two parameters, this paper creatively put forwards the STC calculation method according to whether the roads are directly adjacent. To integrate the structural characteristics of urban roads into the calculation process of STC, the calculation results are more explanatory.

Finally, the proposed method was verified against two common traditional methods through different experiments. The verification results show that the calculation method proposed in this paper has better interpretability for the STC between different roads and can better reveal the internal traffic operation characteristics of the road network.

This paper mainly focuses on the study of urban traffic correlation at the macro traffic level, but lacks the mechanism research on the correlation between non Lane level traffic within the road. The research of lane-level correlation can provide more detailed information for ITS, which is also the main content of future research for this paper. At the same time, combined with a more reasonable STC, traffic information prediction and path planning can also be further studied.

**Author Contributions:** Conceptualization, L.Y. and M.D.; methodology, M.D.; software, M.D.; validation, M.D. and J.T.; formal analysis, M.D.; investigation, M.D. and L.Y.; resources, M.D. and L.Y.; data curation, M.D.; writing—original draft preparation, M.D.; writing—review and editing, M.D. and L.Y.; visualization, M.D. and J.T.; supervision, L.Y.; project administration, L.Y.; funding acquisition, L.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This project was supported by the National Natural Science Foundation of China (Grant No. 51875339).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Conflicts of Interest:** The author declares no conflict of interest.

## References

1. Xu, Z.; Tang, J.; Yin, C.; Wang, Y.; Xue, G. Experience-Driven Congestion Control: When Multi-Path TCP Meets Deep Reinforcement Learning. *IEEE J. Sel. Areas Commun.* **2019**, *37*, 1325–1336. [[CrossRef](#)]
2. Liu, C.; Wang, J.; Cai, W.; Zhang, Y. An Energy-Efficient Dynamic Route Optimization Algorithm for Connected and Automated Vehicles Using Velocity-Space-Time Networks. *IEEE Access* **2019**, *7*, 108866–108877. [[CrossRef](#)]
3. Ermagun, A.; Levinson, D. Spatiotemporal short-term traffic forecasting using the network weight matrix and systematic detrending. *Transp. Res. Part C Emerg. Technol.* **2019**, *104*, 38–52. [[CrossRef](#)]
4. Yuan, J.; Yang, L.; Chen, Q. Intelligent energy management strategy based on hierarchical approximate global optimization for plug-in fuel cell hybrid electric vehicles. *Int. J. Hydrogen Energy* **2018**, *43*, 8063–8078. [[CrossRef](#)]



5. Zhou, W.; Yang, L.; Ying, T.; Yuan, J.; Yang, Y. Velocity Prediction of Intelligent and Connected Vehicles for a Traffic Light Distance on the Urban Road. *IEEE Trans. Intell. Transp. Syst.* **2018**, *20*, 4119–4133. [[CrossRef](#)]
6. Vlahogianni, E.; Golias, J.C.; Karlaftis, M.G. Short-term traffic forecasting: Overview of objectives and methods. *Transp. Rev.* **2004**, *24*, 533–557. [[CrossRef](#)]
7. Vlahogianni, E.; Karlaftis, M.G.; Golias, J.C. Short-term traffic forecasting: Where we are and where we're going. *Transp. Res. Part C Emerg. Technol.* **2014**, *43*, 3–19. [[CrossRef](#)]
8. Yuan, J.; Yang, L. Predictive energy management strategy for connected 48V hybrid electric vehicles. *Energy* **2019**, *187*, 115952. [[CrossRef](#)]
9. Kontorinaki, M.; Spiliopoulou, A.; Roncoli, C.; Papageorgiou, M. First-order traffic flow models incorporating capacity drop: Overview and real-data validation. *Transp. Res. Part B Methodol.* **2017**, *106*, 52–75. [[CrossRef](#)]
10. Gipps, P. A behavioural car-following model for computer simulation. *Transp. Res. Part B Methodol.* **1981**, *15*, 105–111. [[CrossRef](#)]
11. Hidas, P. Modelling lane changing and merging in microscopic traffic simulation. *Transp. Res. Part C Emerg. Technol.* **2002**, *10*, 351–371. [[CrossRef](#)]
12. Nagel, K.; Schreckenberg, M. A cellular automaton model for freeway traffic. *J. Phys.* **1992**, *2*, 2221–2229. [[CrossRef](#)]
13. Wasielewski, P. Car-Following Headways on Freeways Interpreted by the Semi-Poisson Headway Distribution Model. *Transp. Sci.* **1979**, *13*, 36–55. [[CrossRef](#)]
14. Jain, R.; Smith, J.M. Modeling Vehicular Traffic Flow using M/G/C/C State Dependent Queueing Models. *Transp. Sci.* **1997**, *31*, 324–336. [[CrossRef](#)]
15. May, A.D. *Traffic Flow Fundamentals*; Prentice Hall: Hoboken, NJ, USA, 1990.
16. Wang, X.; Yin, D.Z.; Qiu, T.; Yan, X. Applicability analysis of a macroscopic traffic flow model in traffic state prediction. In Proceedings of the Transportation Research Board 95th Annual Meeting, Washington, DC, USA, 10–14 January 2016.
17. Guyonnet, R.; Monteil, J.; Ghosh, B. Empirical and simulation analysis of the relation between microscopic and macroscopic traffic considering trajectory data sets. In Proceedings of the Transportation Research Board 96th Annual Meeting, Washington, DC, USA, 8–12 January 2017.
18. Tang, J.; Liu, F.; Zhang, W.; Zhang, S.; Wang, Y. Exploring dynamic property of traffic flow time series in multi-states based on complex networks: Phase space reconstruction versus visibility graph. *Phys. A Stat. Mech. Appl.* **2016**, *450*, 635–648. [[CrossRef](#)]
19. Zhang, Z.; Li, M.; Lin, X.; Wang, Y.; He, F. Multistep speed prediction on traffic networks: A deep learning approach considering spatio-temporal dependencies. *Transp. Res. Part C Emerg. Technol.* **2019**, *105*, 297–322. [[CrossRef](#)]
20. Lighthill, M.J.; Whitham, G.B. On kinematic waves II. A theory of traffic flow on long crowded roads. In *Royal Society of London. A. Mathematical and Physical Sciences*; The Royal Society: London, UK, 1955; Volume 229, pp. 317–345.
21. Aalipour, A.; Hamed, K.; Mohsen, R. Robust traffic flow control in networks with nonlinear MFD dynamics. In Proceedings of the 97th Transportation Research Board Annual Meeting, Washington, DC, USA, 7–11 January 2018.
22. Tang, J.; Li, L.; Hu, Z.; Liu, F. Short-Term Traffic Flow Prediction Considering Spatio-Temporal Correlation: A Hybrid Model Combining Type-2 Fuzzy C-Means and Artificial Neural Network. *IEEE Access* **2019**, *7*, 101009–101018. [[CrossRef](#)]
23. Turochy, R.E. Enhancing Short-Term Traffic Forecasting with Traffic Condition Information. *J. Transp. Eng.* **2006**, *132*, 469–474. [[CrossRef](#)]
24. Tang, J.; Wang, Y.; Wang, H.; Zhang, S.; Liu, F. Dynamic analysis of traffic time series at different temporal scales: A complex networks approach. *Phys. A Stat. Mech. Appl.* **2014**, *405*, 303–315. [[CrossRef](#)]
25. Yue, Y.; Yeh, A.G.-O. Spatiotemporal traffic-flow dependency and short-term traffic forecasting. *Environ. Plan. B Plan. Des.* **2008**, *35*, 762–771. [[CrossRef](#)]
26. Duan, P.; Mao, G.; Liang, W.; Zhang, D. A Unified Spatio-Temporal Model for Short-Term Traffic Flow Prediction. *IEEE Trans. Intell. Transp. Syst.* **2019**, *20*, 3212–3223. [[CrossRef](#)]
27. Zeng, Q.; Guo, Q.; Wong, S.C.; Wen, H.; Huang, H.; Pei, X. Jointly modeling area-level crash rates by severity: A Bayesian multivariate random-parameters spatio-temporal Tobit regression. *Transp. A Transp. Sci.* **2019**, *15*, 1867–1884. [[CrossRef](#)]
28. Wen, H.; Zhang, X.; Zeng, Q.; Sze, N.N. Bayesian spatial-temporal model for the main and interaction effects of roadway and weather characteristics on freeway crash incidence. *Accid. Anal. Prev.* **2019**, *132*, 105249. [[CrossRef](#)]
29. Guo, S.; Zhou, D.; Fan, J.; Tong, Q.; Zhu, T.; Lv, W.; Li, D.; Havlin, S. Identifying the most influential roads based on traffic correlation networks. *EPJ Data Sci.* **2019**, *8*, 1–17. [[CrossRef](#)]
30. Anselin, L. Local Indicators of Spatial Association-LISA. *Geogr. Anal.* **1995**, *27*, 93–115. [[CrossRef](#)]
31. Getis, A. *Spatial Interaction and Spatial Autocorrelation: A cross-Product Approach in Perspectives on Spatial Data Analysis*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 23–33.
32. Cliff, A.; Ord, J. Spatial processes: Models & applications. *Q. Rev. Biol.* **1982**, *57*, 236.
33. López-Hernández, F.; Chasco, C. Time-trend in spatial dependence: Specification strategy in the first-order spatial autoregressive model. *Estud. Econ. Apl.* **2007**, *25*, 1–22.
34. Su, F.; Dong, H.; Jia, L.; Tian, Z.; Sun, X. Space-time correlation analysis of traffic flow on road network. *Int. J. Mod. Phys. B* **2017**, *31*, 1750027. [[CrossRef](#)]
35. Yang, Y.; Qin, Y.; Li, X.; Tian, Y.; Jia, L. Correlation patterns of highway segment travel times. In Proceedings of the Transportation Research Board 94th Annual Meeting, Washington, DC, USA, 11–15 January 2015; pp. 1–16.

36. De Freitas, E.P.; Heimfarth, T.; Wagner, F.R.; Pereira, C.E.; Larsson, T. Exploring geographic context awareness for data dissemination on mobile ad hoc networks. *Ad Hoc Netw.* **2013**, *11*, 1746–1764. [[CrossRef](#)]
37. De Freitas, E.P.; Heimfarth, T.; Pereira, C.E.; Ferreira, A.M.; Wagner, F.R.; Larsson, T. Evaluation of coordination strategies for heterogeneous sensor networks aiming at surveillance applications. In Proceedings of the SENSORS, 2009 IEEE, Christchurch, New Zealand, 25–28 October 2009; pp. 591–596.
38. De Freitas, E.P.; Wehrmeister, M.A.; Silva, E.T.; Carvalho, F.C.; Pereira, C.E.; Wagner, F. DERAFA: A High-Level Aspects Framework for Distributed Embedded Real-Time Systems Design. In Proceedings of the 10th International Workshop on Early Aspects—Current Challenges and Future Directions, Vancouver, BC, Canada, 13 March 2007.
39. Richards, P.I. Shock Waves on the Highway. *Oper. Res.* **1956**, *4*, 42–51. [[CrossRef](#)]
40. Andreianov, B.; Donadello, C.; Rosini, M.D.; Andreianov, B.; Donadello, C.; Rosini, M.D. Entropy solutions for a two-phase transition model for vehicular traffic with metastable phase and time depending point constraint on the density flow. *Nonlinear Differ. Equ. Appl.* **2021**, *28*, 1–37. [[CrossRef](#)]
41. Muck, J. Using detectors near the stop-line to estimate traffic flows. *Traffic Eng. Control* **2002**, *43*, 429–434.
42. Skabardonis, A.; Geroliminis, N. Real-Time Monitoring and Control on Signalized Arterials. *J. Intell. Transp. Syst.* **2008**, *12*, 64–74. [[CrossRef](#)]
43. Liu, H.X.; Wu, X.; Ma, W.; Hu, H. Real-time queue length estimation for congested signalized intersections. *Transp. Res. Part C Emerg. Technol.* **2009**, *17*, 412–427. [[CrossRef](#)]
44. Flötteröd, G.; Rohde, J. Operational macroscopic modeling of complex urban road intersections. *Transp. Res. Part B Methodol.* **2011**, *45*, 903–922. [[CrossRef](#)]
45. Zhu, C.; Wang, X. Research on evaluation algorithm of key nodes in urban road traffic network based on complex network. *J. Phys. Conf. Ser.* **2020**, *1629*, 012021. [[CrossRef](#)]
46. Liang, C.; Wang, J.; Zhou, Z.; Li, Q.; Chen, M. Evaluations of node importance of urban road network based on transmission contribution matrix. *Sci. Technol. Rev.* **2018**, *36*, 105–111.
47. Wang, J.-F.; He, X.; Si, S.-Z.; Zhao, H.; Zheng, C.; Yu, H.; Wang, J.-F.; He, X.; Si, S.-Z.; Zhao, H.; et al. Using complex network theory for temporal locality in network traffic flows. *Phys. A Stat. Mech. Appl.* **2019**, *524*, 722–736. [[CrossRef](#)]
48. Langville, A.N.; Meyer, C.D. Google’s PageRank and Beyond: The science of search engine rankings. *Math. Intell.* **2008**, *30*, 68.
49. Hu, Y.; Li, M.; Zhang, P.; Fan, Y.; Di, Z. Community detection by signaling on complex networks. *Phys. Rev. E* **2008**, *78*, 016115. [[CrossRef](#)] [[PubMed](#)]
50. Feng, H.; Bai, F.; Xu, Y. Identification of critical roads in urban transportation network based on GPS trajectory data. *Phys. A Stat. Mech. Appl.* **2019**, *535*, 122337. [[CrossRef](#)]