IAENG International Journal of Computer Science

Method for Identifying and Analyzing Factors Influencing Traffic Efficiency at Lane Level

Huasheng Liu, Kui Dong, Kai Ding*, Haoran Deng, Zhongpeng Wang, Jin Li, and Sha Yang

Abstract—**This study proposes a method for analyzing the factors influencing traffic efficiency at both section and lane levels. A traffic state differentiation index quantifies the variations in traffic states at the lane level. The findings indicate that higher traffic volumes result in greater disparities in traffic conditions across lanes in our study scenario. The relationship between traffic efficiency and its influencing factors is examined at the lane level, with a sensitivity analysis based on the elasticity coefficient to identify sensitive ranges. Multiple linear regression quantifies the effects of these factors, while a genetic algorithm-optimized backpropagation neural network and a gated recurrent neural network predict traffic indicators. The best prediction results are achieved with the traffic flow dataset, followed by parking vehicles and stopping time datasets, with the speed-limits dataset yielding the least accurate results. Prediction errors are more pronounced within the calculated sensitive intervals, warranting closer attention. The proposed method improves prediction accuracies for queue length, travel time, lane-level flow, and lane-level occupancy by 11.36%, 12.03%, 3.80%, and 6.71%, respectively. This approach effectively identifies key factors and sensitivity ranges impacting traffic status at both section and lane levels, which enhances the interpretability of machine-learning methods in traffic prediction.**

Index Terms—**urban traffic, data analysis, traffic simulation, neural network**

I. Introduction

URBAN traffic is a complex network system where local bottlenecks form due to congestion. If not quickly resolved, these bottlenecks can further spread along lines and surfaces, affecting the entire network [1]–[3]. Traffic flow characteristics refer to the patterns and relationships within a traffic system, describing the laws of change in traffic flow under different conditions through both quantitative and qualitative descriptions. Hence, the study of traffic flow characteristics is the basis for an in-depth analysis of how traffic problems emerge on road sections, which can guide the proposal of improvement strategies and the optimization of traffic design. Several countries (regions) have conducted research on traffic flow characteristics and prepared corresponding road access manuals for road traffic conditions. The most widely used manual is the US Highway Capacity Manual, and its latest results are published in the 2016 edition [4]. Scholars have continuously improved this manual, mainly from two perspectives: traffic data collection technology and calculation methods.

Urban road traffic flow characteristics must be studied based on traffic flow data from actual scenarios. Thus, the efficient and realistic collection of traffic data and the construction of a simulation environment for the interweaving zone based on the collected data are particularly important. Sherief *et al.* [5] used video-based measurement techniques to collect and analyze traffic data such as flow, speed, and density. By collecting experimental data, Yuan *et al.* [6] analyzed forced lane-changing behaviour in the intertwining zones of expressway entrances and exits. Anas *et al.* [7] studied the relationship between the number of lanes, heavy vehicles, and saturation flow using actual data. However, the high cost, difficulty, and low precision of actual data collection caused significant difficulties in studying traffic characteristics. With continuous advancements in traffic simulation technology, the accuracy and diversity of computer simulation results have gradually improved. These simulations can be tailored to different road factors according to the researcher's requirement to model the traffic conditions. Scholars have largely employed simulation technology to study and analyze various transportation-related issues. Bharadwaj *et al.* [8] used VISSIM to simulate traffic flow operations on a multilane highway, generating speed flow curves from the simulation data, which were then used to calibrate the capacity model parameters. Marczak *et al.* [9] used FOSIM software to simulate the traffic state of urban roads. Nagel and Schreckenberg [10] proposed a classical one-dimensional meta-cellular model, the Na–Sch model, which is a simple model that can simulate actual road traffic phenomena. VISSIM exhibited better accuracy and reliability. It can obtain accurate operational efficiency data of sections and lanes and simulate complex traffic flow scenarios. Therefore, this software is widely used by researchers.

Manuscript received July 5, 2024; revised May 6, 2025.

This work was supported by the Scientific Research Project of the Jilin Education Department in 2023 (JJKH20231189KJ); Jilin Province Transportation Innovation Development Support (Technology) Project (2024-1-6); Graduate Innovation Fund of Jilin University (2023CX197); National Natural Science Foundation of China (71871103).

Huasheng Liu is an associate professor at the College of Transportation, Jilin University, Changchun 130022, China (corresponding author to provide email: liuhuasheng521@163.com).

Kui Dong is a postgraduate student at the College of Transportation, Jilin University, Changchun 130022, China (email: dongkui23@mails.jlu.edu.cn).

Kai Ding is an engineer at Jilin Road and Bridge Construction (Group) Co., Ltd., Changchun 130013, China (corresponding author to provide email: 13904315920@139.com).

Haoran Deng is a postgraduate student at the College of Transportation, Jilin University, Changchun 130022, China (email: deng19963009040@163.com).

Zhongpeng Wang is an engineer at Jilin Provincial transportation administration., Changchun 130021, China (email: 280009869@qq.com).

Jin Li is an associate professor at the College of Transportation, Jilin University, Changchun 130022, China (email: li_jin@jlu.edu.cn).

Sha Yang is a postgraduate student at the College of Transportation, Jilin University, Changchun 130022, China (email: ys18886335213@163.com).

Calculation methods for traffic flow characteristics primarily include statistical analysis and traffic forecasting. Statistical analysis is mainly applied to quantitative or qualitative analyses of traffic status. Adams [11] first proposed a traffic flow theory based on mathematical and statistical theories. Subsequently, scholars simultaneously focused on the traffic flow characteristics of urban roads and the interplay between different traffic elements and their use in related traffic data analysis. Traffic forecasting is primarily used to assess and identify future traffic states. Fedotkin *et al.* [12] proposed a numerical imitational model that not only allowed the user to watch the traffic at a crossing in video mode but also computed the basic characteristics of the system using the minimum average weighted waiting time in stationary mode.

Qi *et al.* [13] used a speed-based grading method to identify bottlenecks, which had better bottleneck identification performance. To reveal the interrelationships between traffic flow parameters and their spatiotemporal patterns, they integrated difference and sensitivity analyses with statistics to quantitatively analyze the traffic characteristics of urban road sections and lanes. Variability analyses have been widely applied in traffic efficiency research. Liu *et al.* [14] studied the regional variability of road traffic efficiency in space. Karlaftis *et al.* [15] collected panel data over 11 years from 15 European cities and investigated the variability of different methods used to evaluate urban road traffic efficiency. Barnum *et al.* [16] analyzed the variability of the operational efficiency of public transport using the data envelopment analysis method. To identify the key factors and sensitive nodes influencing the efficiency of traffic operations on road sections, sensitivity analysis was used to calculate their impact on the evaluation of indicators. The magnitude of this impact is referred to as the sensitivity factor. Sensitivity analyses are widely employed in engineering applications. Tobin *et al.* [17] conducted a sensitivity analysis with variational inequalities on urban traffic-balancing network flows. Kitamura *et al.* [18] identified the main factors affecting patronage through a sensitivity analysis.

Both theory and practice demonstrate that traffic management and control can effectively reduce traffic congestion and improve the operational efficiency of transportation systems. Jelena *et al.* [19] demonstrated that on-road parking for guidance improved resource utilization. Traffic management and control are generally classified into traffic demand and system management. The road traffic state supports the input parameters of the traffic control management system and is the basis for optimized decision-making. It has several applications in ramp control, travel time estimation, accident detection, transportation planning, and transport infrastructure evaluation [20], [21]. Research on traffic status focuses on the macroscopic grasp parameters (flow, speed, and density) of the traffic status of road sections. A few scholars have studied lane traffic status, which is a fine-grained traffic-status estimation problem that can meet the needs of fine-grained management decision-making. Lane-level traffic status has several applications in traffic control management, such as freeway ramp control, lane change recommendations, and lane-splitting speed limits [22].

According to this research perspective, the traffic status can be categorized into two levels: section and lane. This distinction has seldom been made in previous studies. However, with advances in traffic control and management, road-section level traffic state estimates no longer meet the requirements for tasks such as lane-splitting speed limit and flow allocation. Section-level approaches have dominated traffic-status-related studies. Anand *et al.* [23] used a Kalman filter algorithm to fuse video-obtained traffic with travel-time data from the GPS to estimate traffic density at the road-section level and validated the model using both field and simulation data. Deng *et al.* [24] proposed a method for estimating traffic density at the road-section level using a variety of data in a least-squares framework. Duret *et al.* [25] proposed a framework for assessing the traffic status at the road-section level that fits both Eulerian and Lagrangian data. Only a few traffic-status estimation methods have been explored at the lane level. Wright *et al.* [26] explored multiple traffic status estimation methods for single and multiple lanes and compared their effectiveness. Bekiaris-Liberis *et al.* [27] proposed a lane-level traffic-state estimation method that combined networked vehicles with fixed-coil data.

Student commuting is an important part of urban traffic and is characterized by high complexity, regularity, and short-term aggregation [28]. These factors also contribute to the vulnerability of school roads to traffic accidents [29]. Therefore, its impact on the general operational efficiency of urban traffic cannot be ignored. In cities of developing countries, increasing car ownership is making private cars the preferred mode of travel because of their convenience, speed, and other advantages. This shift has led to a profound change in how students commute to school. Vehicle-stopping behaviour is commonly observed on school roads. The frequent movement of vehicles into and out of parking spaces results in large variations in traffic conditions across lanes on these roads. Mouronte-López *et al.* [30] studied the factors influencing the choice of transportation mode for student commuting. Hu *et al.* [31] investigated changes in traffic flow status caused by temporary parking behaviour on roadways. Cao *et al.* [32] showed that on-street parking reduced the lateral residual width of lanes, which affected the road capacity, mainly in terms of the impact on parking lanes and their adjacent lanes. Song *et al.* [33] proposed a cyclic reservation and allocation model for improving the utilization efficiency of parking spaces. Their main research objective was to construct a simulation scenario based on the actual traffic flow data of school road sections. A combination of macro- and micro-approaches was used to investigate the traffic status of a school scene at the road section and lane levels, exploring the mechanism of traffic problems and proposing improvement strategies for optimizing traffic operation organizations.

Intelligent transportation systems (ITS) are essential technologies for advancing urban traffic management. Consequently, traffic state prediction has emerged as a key development area in transportation [34]. Urban road traffic conditions are highly susceptible to external factors. Moreover, the numerous traffic parameters combined with complex mapping relationships, especially in congestion-prone areas, where the number of lanes and lane

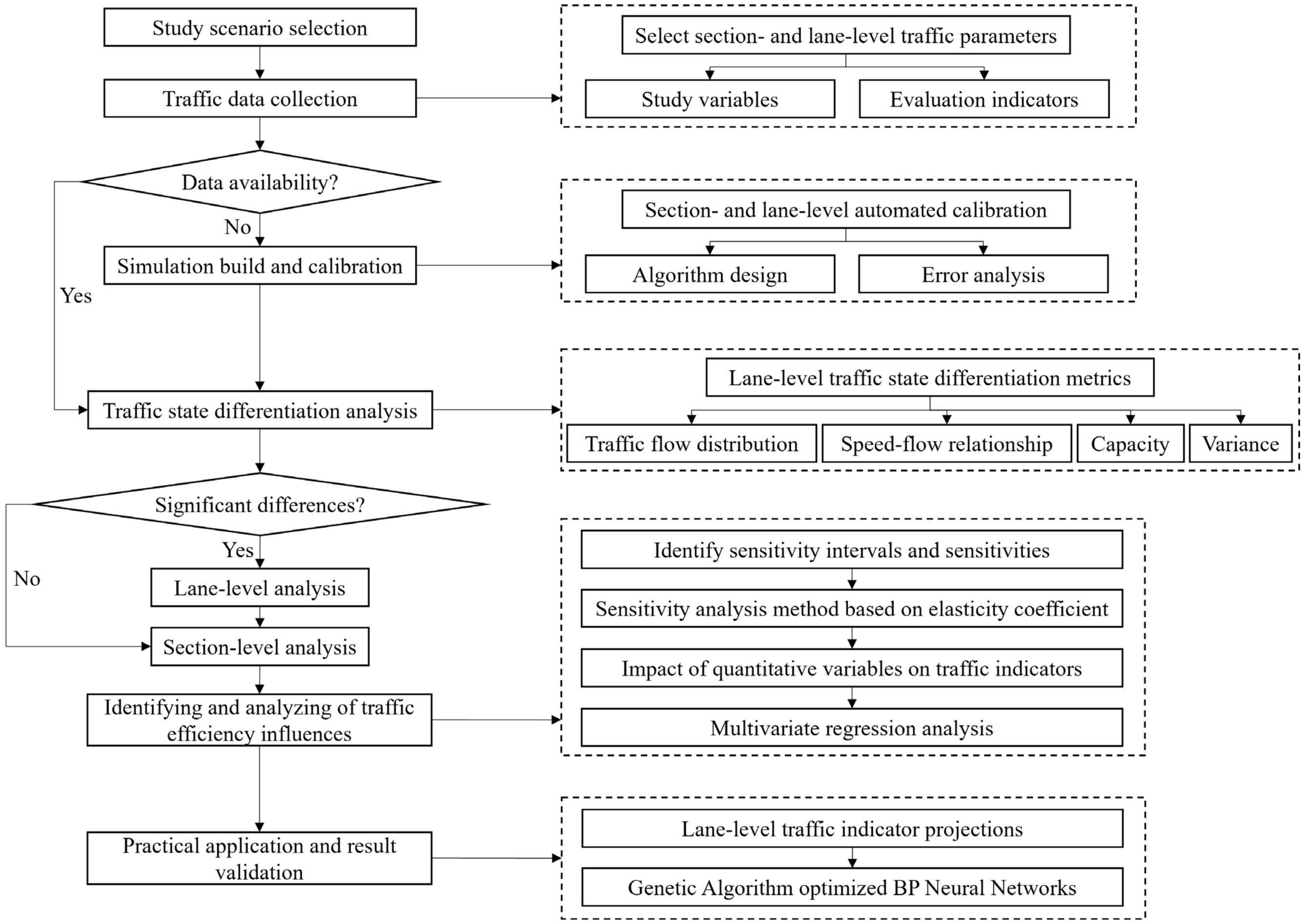

Fig. 1. Study framework for identification and analysis of traffic factors.

types can increase drastically, significantly increasing the dimensionality and complexity of traffic predictions [35]. Therefore, identifying traffic parameters that strongly correlate with predictive indices remains a pressing challenge.

Bai *et al.* [36] introduced a novel two-branch, multitemporal-resolution convolutional network with an adaptive graph NN to predict traffic states in urban road systems. This spatiotemporal resolution convolutional network utilized optimized Tujin frequent differential equations and adaptive correlation adjacency graphs to ensure accurate feature propagation throughout the network. Osorio-Arteaga *et al.* [37] proposed a robust identification and control method based on an NN using a recursive adaptive training algorithm. This approach optimizes a NN by quantifying the impacts of multiple variables on a system. Du *et al.* [38] focused on extracting and learning complex dynamic spatiotemporal features from raw traffic flow data by applying deep learning models to improve predictive accuracy. Papasani *et al.* [39] demonstrated that effective feature extraction can significantly enhance the operational efficiency and estimation accuracy of machine learning algorithms. Zhang *et al.* [40] demonstrated that optimizing combinations of system-related parameters could improve the robustness and accuracy of NNs. They began with traffic prediction to verify the feasibility of the proposed method for identifying factors affecting traffic efficiency at the lane level.

In this study, starting with the traffic congestion problem of mixed and intertwined roads in urban areas, a comprehensive simulation model was established using representative school road traffic flows during school hours. The section-level and lane-level traffic factors affecting road efficiency were identified and analyzed. The proposed framework is shown in Fig. 1.

The structure of the rest of the paper is as follows. In Section II, the school road simulation model is constructed based on the composition of traffic flow and the behavioural characteristics of traffic participants in the actual road section. Driving-behavior parameters were calibrated using a genetic algorithm (GA) considering section- and lane-level indicators. Section III presents the lane-level traffic state differentiation, describing the traffic flow distribution, speed-flow relationship, and capacity of each lane. Section IV proposes a sensitivity-analysis method for exploring the mechanisms of school road traffic problems. A multiple linear regression model is used to quantify the extent to which different factors influence traffic states. In Section V, GA-optimized BPNNs are applied to predict section- and lane-level traffic indicators and compare the results with those analyzed in this study.

II. Simulation Modeling and Parameter Calibration

Traffic microsimulation has significant advantages in analyzing the mechanism of the supply-demand balance of complex traffic networks and deducing the spatiotemporal operation state of traffic. Numerous parameters and models are embedded and fused to describe better traffic system operations, traffic flow characteristics, and driver behaviour.

PTV VISSIM is a widely used microsimulation program for modelling multimodal traffic processes using the Vision Traffic Suite software package. It enables detailed and highly flexible modelling of road design, vehicle performance, and driver behaviour. This study, PTV VISSIM 2021 was used to construct the simulation model, and MATLAB R2016a was used to calibrate the simulation parameters to ensure its accuracy and reliability. The computer hardware was equipped with an 11th Gen Intel (R) Core (TM) i7-11800H CPU, with 16 GB of RAM, 8 cores, and 16 logical processors.

A. Simulation Modeling

Construction of a school road-section simulation model requires the collection of actual geometric features, vehicle operating statuses, vehicle types, and traffic evaluation indicators of the road section. Geometric data for the simulation modelling were measured in situ using rangefinders and other devices. The operating status of vehicles, vehicle composition, and traffic evaluation indicators were collected using videos captured by unmanned aerial vehicles. This method has the advantages of rich data extraction and affordability [41], [42]. Because the data-collection site is in Northeast China, a region with a cold climate, non-motorized travel options were minimal during the survey period. Consequently, this study mainly considered the impacts of motorized vehicles. Basic data, such as speed, acceleration, and traffic flow, were further processed using TRACKER software and manual statistics. The basic data types and names are listed in Table I.

TABLE I
TYPE OF DATA COLLECTED

Geometric data	Traffic Statue	Evaluation Indicator
Lane number	Velocity	Queue length
Road width	Acceleration	Travel time
Lane width	Car number	Occupancy
Nonmotorized lane width	Number of stops	Cross-section flow
Number of parking spaces	Stopping time	-

VISSIM provides two psychophysical healing models: Wiedemann (1974) and Wiedemann (1999) for longitudinal vehicle motion and a rule-based algorithm for lateral motion [43]. Wiedemann (1974) was used in this study because it represents urban scenarios. The modelling steps were as follows: First, a road was built; based on an actual scenario, a unidirectional three-lane roadway was built, and a pedestrian crosswalk was set up. Second, the traffic-operation organization of the road was restored. Traffic facilities, such as parking spaces, school buses, pedestrian entrances and exits, signal lights, and other traffic management measures, such as signs and markings, speed reduction zones, priority rules, and vehicle paths, were installed. Finally, traffic flow data were input. Additionally, the vehicle composition, velocity, acceleration, deceleration, headway time, distance, and other relevant parameters were set.

In this study, we compared and analyzed traffic status differences across two dimensions: road cross-sections and lanes. Nine sections were selected and labelled, as shown in Fig. 2.

B. Parameter Calibration

The simulation model parameter calibration involves various parameters and complex combinations and requires an optimization algorithm to obtain the closest combination of driving behaviour parameters. The calibration problem in this study did not have a strictly mathematical analytical solution, and the simulation model must be called frequently during the calibration process. A genetic algorithm was used to reduce the calibration time and prevent local optimization.

To verify the reliability of the simulation model and the accuracy of the calibration parameters, evaluation indicators that can reflect the operation of vehicles in both actual and simulated traffic environments must be selected. Owing to the uneven distribution of traffic in school section lanes during school days, both section-level and lane-level indicators were selected for this study. Travel time and queue length can be obtained from real-world scenarios, and driving behaviour parameters with which evaluation indicators are correlated exist; therefore, they were chosen as section-level indicators [44]. The main indicators of lane traffic distribution are lane saturation flow rate and utilization [45]. In VISSIM, lane-level volumes and occupancy can reflect these indicators and can be obtained statistically; therefore, they were selected as lane-level indicators.

Parameter calibration aims to minimize the error between the output value of the simulation and the actual measured value. Based on the research requirements of this study, the root mean square relative error was used to construct an objective function [46].

$$MinF(X)=\alpha\sqrt{(Q_r-Q_s)^2}+\beta\sqrt{(T_r-T_s)^2}+\sum_i\sum_j\eta_{ij}\sqrt{(O_{ij}^r-O_{ij}^s)^2}+\sum_i\sum_j\varepsilon_{ij}\sqrt{(V_{ij}^r-V_{ij}^s)^2} \quad (1)$$

where α, β, η_{ij}, and ε_{ij} are the coefficients of queue length, travel time, occupancy rate, and cross-sectional flow; Q_r and Q_s are the observed and simulated values of queue length; T_r and T_s are the observed and simulated values of travel time; O_{ij}^r and O_{ij}^s are the observed and simulated values of lane i cross-section j occupancy rate; and V_{ij}^r V_{ij}^s are the observed and simulated values of lane i cross-section j flow, respectively.

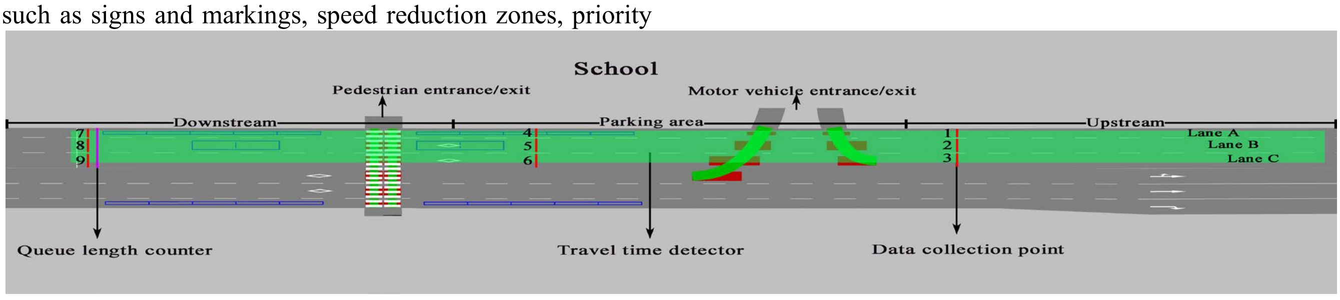

Fig. 2. Simulation model of a roadway by a school

III. Analysis of Lane-level Traffic State Differentiation

To reflect the difference in traffic status of each lane on the school road, we analyzed the distribution of traffic flow in the lanes, speed–flow relationship, and capacity of the lanes. First, a set of simulation scenarios was created. The school-road cross-section was divided into upstream, parking area, and downstream, with lanes divided into a, b, and c. For example, the parking area reflects the process of pick-up and drop-off vehicles pulling in and out of lane A, which provides parking spaces. Second, the lane traffic status characteristics were analyzed. The differences between lanes were discussed in terms of trends, intervals, and key nodes. Finally, a one-way analysis of variance (ANOVA) was applied to test lane differentiation.

To represent the traffic flow distribution and speed–flow relationship of the school section lanes, the section traffic flow was set as a variable with a range of [100,2000], step size of 100 pcu/h, speed limit of 70 km/h, percentage of parked cars of 10%, and a stopping time of 30 s. Each set of simulations was repeated 10 times (10 simulation seeds). It had a runtime of 3600 s and a warm-up time of 600 s. When the traffic flow on the road reaches 1400, 1500, and 1600 pcu/h, the upstream, parking area, and downstream lanes are saturated.

Fig. 3 shows the traffic flow distributions at various locations. Differences in traffic flow distribution are observed between lanes of the same cross section, with lanes A and B having similar and significantly smaller flows than lane C, owing to the impact of the stops. In addition, differences in traffic flow are present at different cross sections of the same lane, with the downstream section having the highest flow and the upstream section having the lowest. Differences in traffic flow between the cross sections of the same lane are significantly smaller than those between lanes of the same cross section.

Fig. 4 shows the relationship between lane-level flow and speed. Variation in vehicle speed was observed between lanes of the same cross-section, with lane A having the highest rate of change and lane C having the least. Variations were also observed between cross-sections of the same lanes. The upstream area showed the highest variation, with an interval of [70,10]. The variation interval in the parking section is [70,18]. The downstream section had the least variability and faster vehicle speeds, with an approximate interval of [63,23]. The variation in vehicle speed at different cross sections of the same lane was less than that between lanes of the same cross section.

Fig. 5 presents the capacities of different locations on the school road. The capacity significantly varies between lanes of the same cross-section. Lane C has the most significant variation, whereas Lane A has the smallest. The variation of traffic flow in the cross-sections for the same lanes is slight. The upstream and parking sections have similar and lower capacities than the downstream sections. The variation of traffic flow between cross-sections of the same lane is significantly smaller than the variation in capacity between lanes of the same cross section.

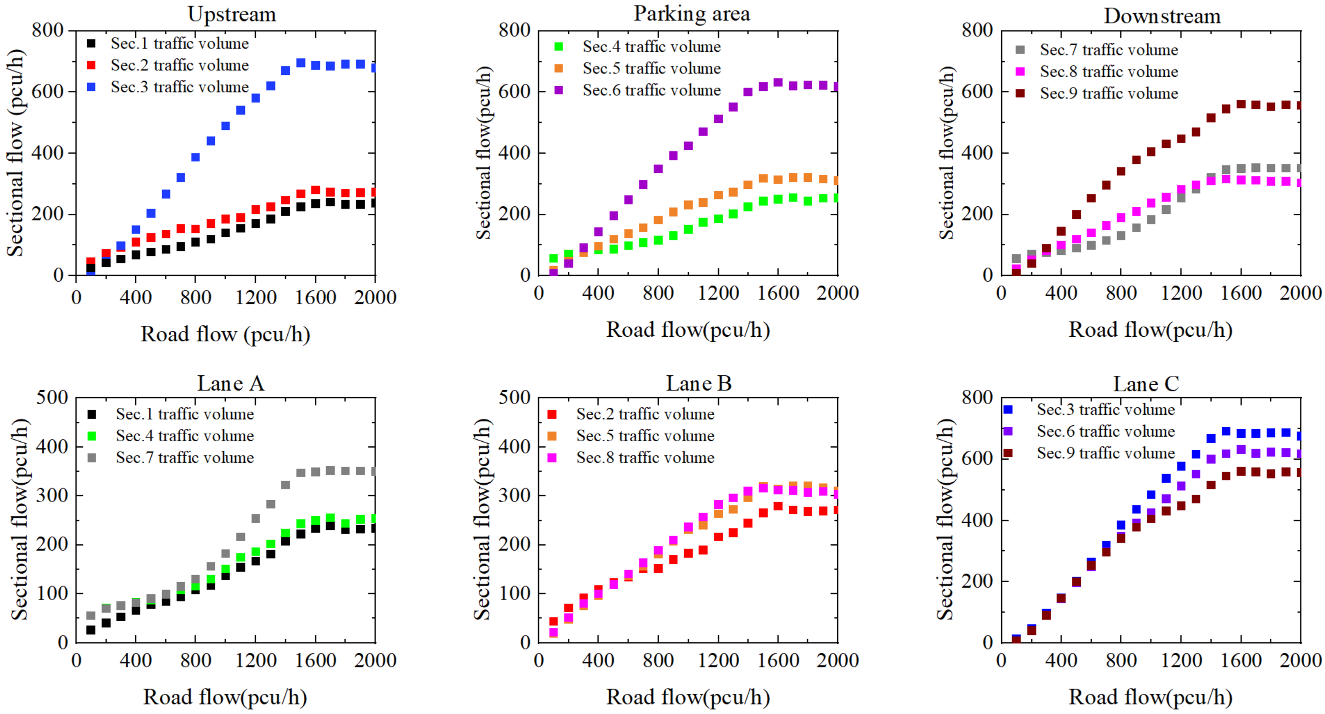

Fig. 3. Lane-level traffic flow distribution

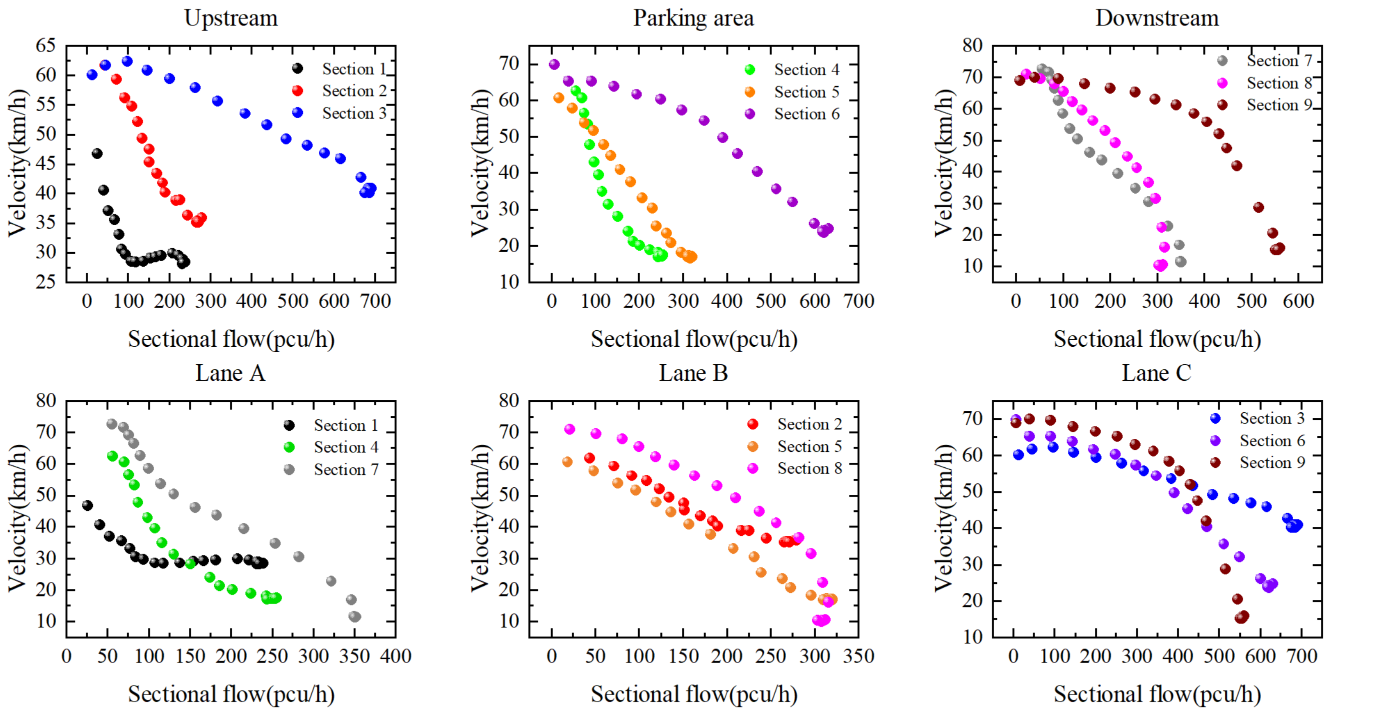

Fig. 4. Lane-level flow–velocity relationship

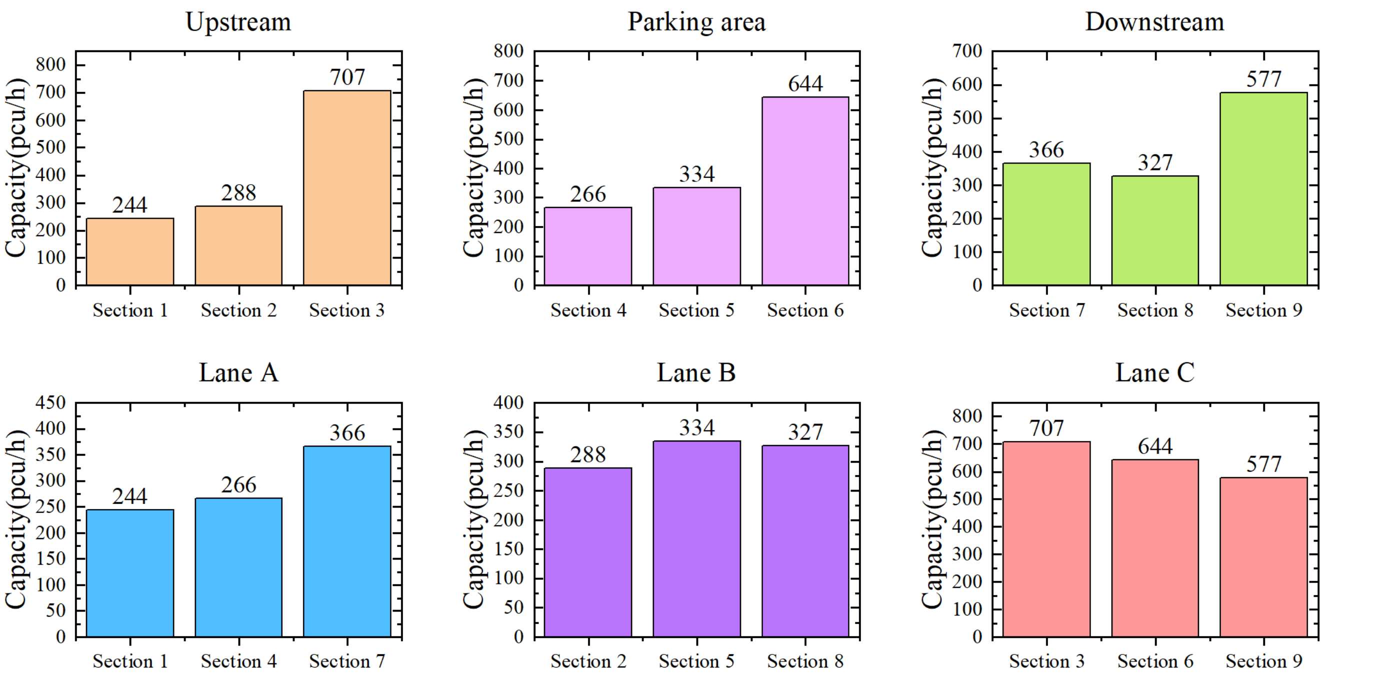

Fig. 5. Lane-level traffic capacity

Variance at various locations was selected as an indicator of differentiation. Fig. 6 shows the relationship between capacity variance and traffic flow. As the traffic flow increased, the differentiation between lanes increased, and they exhibited a positive correlation. Traffic lane choice should be guided to rationalize right-of-way assignments when vehicles gather rapidly on school roads during school hours.

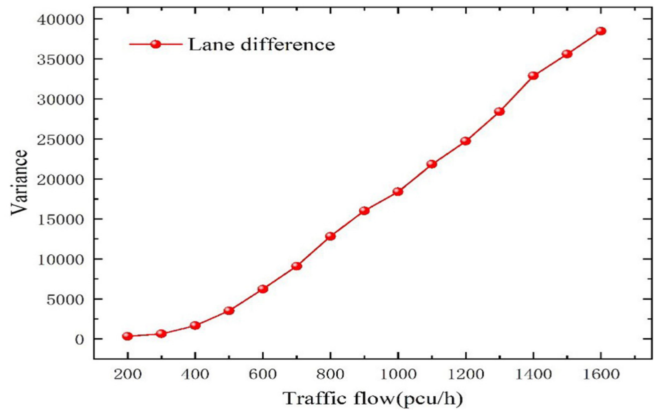

Fig. 6. Lane-level traffic capacity

The results of the one-way ANOVA test showed that the significances between the lane cross sections were 0.00, except for the lane B parking area and upstream, where the significance was >0.05 (0.054 for traffic distribution, 0.93 for vehicle speed, and 0.108 for capacity).

IV. Analysis of Factors Affecting Traffic Status

On-road parking is a major contributor to the variance of the traffic status of school-section lanes. Based on the results of the lane differentiation analysis and related literature [31], the speed limit, traffic flow, number of parked vehicles, and stopping time were selected as the key variables affecting the turnover of parking spaces and the operational efficiency of the road. The steps to quantitatively analyze the relationship between the traffic status of different road locations and the variables were as follows: First, a simulation experiment was designed. Second, the sensitivity and relations between the variables and evaluation indicators were analyzed. Finally, the key factors and sensitive nodes influencing the operational efficiency of school sections were identified.

A. *Experiment Design*

Four sets of experiments were designed to analyze the relationship between the variables and evaluation indicators. Based on the previous discussion, road queue length, travel time, lane-level flow, and occupancy were selected as the evaluation indicators.

Experiment 1: Effects of road speed limits on school sections and lanes.

Experiment 2: Effects of traffic flow on school sections and lanes.

Experiment 3: Effect of the number of parked vehicles on school sections and lanes.

Experiment 4: Impact of stopping time on school sections and lanes.

Based on the results of the lane differentiation analysis, the road capacity was probably 1500 pcu/h. As the traffic flow changed, the maximum value of the vehicle speed was distributed above and below 70 km/h, and the minimum value was distributed above and below 15 km/h. At a stopping time of 30 s, the number of parked vehicles accommodated by the road section was approximately 200. Table II lists the specific setups of the four sets of experiments used to analyze their impact on the evaluation indicators.

TABLE II
EXPERIMENTAL DESIGN

Expt.	Speed limits	Traffic flow	Parking vehicles	Stopping time	Step
1	[10,70]	1500	150	30	5
2	30	[200,1600]	150	30	100
3	30	1500	[100,250]	30	15
4	30	1500	150	[10,120]	10

B. *Results Analysis*

In a school road-transport system, the road operational efficiency is mainly influenced by the speed limit, traffic flow, number of parked vehicles, and stopping time. We varied them within their respective ranges and analyzed the trends and extent of changes in the evaluation indicators. Sensitivity and multiple linear regression analyses were also performed. Calculations were performed on MATLAB 2016a.

In this study, based on a single-factor sensitivity analysis, we proposed a sensitivity function and sensitivity factor in dimensionless forms, which makes the multifactor sensitivity analysis comparable. First, the functional relationship between the evaluation indicators and variables was established. Modelling the actual scenario as closely as possible is crucial to an effective parameter sensitivity analysis. Because cubic spline interpolation is not only characterized by excellent stability, guaranteed convergence, and smooth curves but also requires only function value information, it can avoid the Longe phenomenon produced by higher interpolation. In this study, based on the simulation data, a cubic spline interpolation function was established, which is represented by solid lines in Fig. 7–15.

The above analysis provides only an idea of the sensitivities of the evaluation indicators to a single factor. In real-world scenarios, traffic status variables are generally physical quantities with different units. The above analysis did not allow a comparison of sensitivity between the factors. Therefore, rendering the sensitivity function and sensitivity factor dimensionless is necessary. Second, a sensitivity function was established. By setting the cubic spline interpolation function established in the previous section as p and the variable as a_k, we defined the sensitivity function $S_k(a_k)$ as the rate of change in the evaluation indicators with respect to changes in the variable.

$$S_k(a_k) = (\frac{|\Delta p|}{p}) / (\frac{|\Delta a_k|}{a_k}) = \left|\frac{\Delta p}{\Delta a_k}\right| \frac{a_k}{p} \tag{2}$$

If $|\Delta a_k| / a_k$ is small, $S_k(a_k)$ can be approximated as:

$$S_k(a_k) = \left|\frac{d\varphi_k(a_k)}{da_k}\right| \frac{a_k}{p} \tag{3}$$

Finally, the sensitivity factor is calculated. From (2), the sensitivity function curve of a_k can be plotted, represented by the dotted line in Fig. 10–18. We set $a_k = a_k^*$ when the maximum value S_k^* is achieved. The sensitivity factor can be obtained from this value. The sensitivity of each variable is analyzed by comparing their S_k^*.

$$S_k^* = S_k(a_k^*) = \left|(\frac{d\varphi_k(a_k)}{da_k})_{a_k = a_k^*}\right| \frac{a_k^*}{p^*} \tag{4}$$

Fig. 7 shows the variables as functions of the section-level evaluation indicators and sensitivity functions. In Experiment 1, this indicator decreased as the speed limit increased. The sensitive intervals for the queue lengths were [20,30] and [50,70]. The travel time interval was [10,40]. In Experiment 2, the queue length increased with traffic flow on the road and had a sensitivity interval of [650,1600]. The travel time increased as the traffic flow increased at [800,1600]. [1200,1600] is its sensitive interval. In Experiments 3 and 4, the indicators increased with the number of parked vehicles and the parking time. The queue length increased more rapidly. The sensitive intervals for the section-level indicators were [100,220] and [10,40].

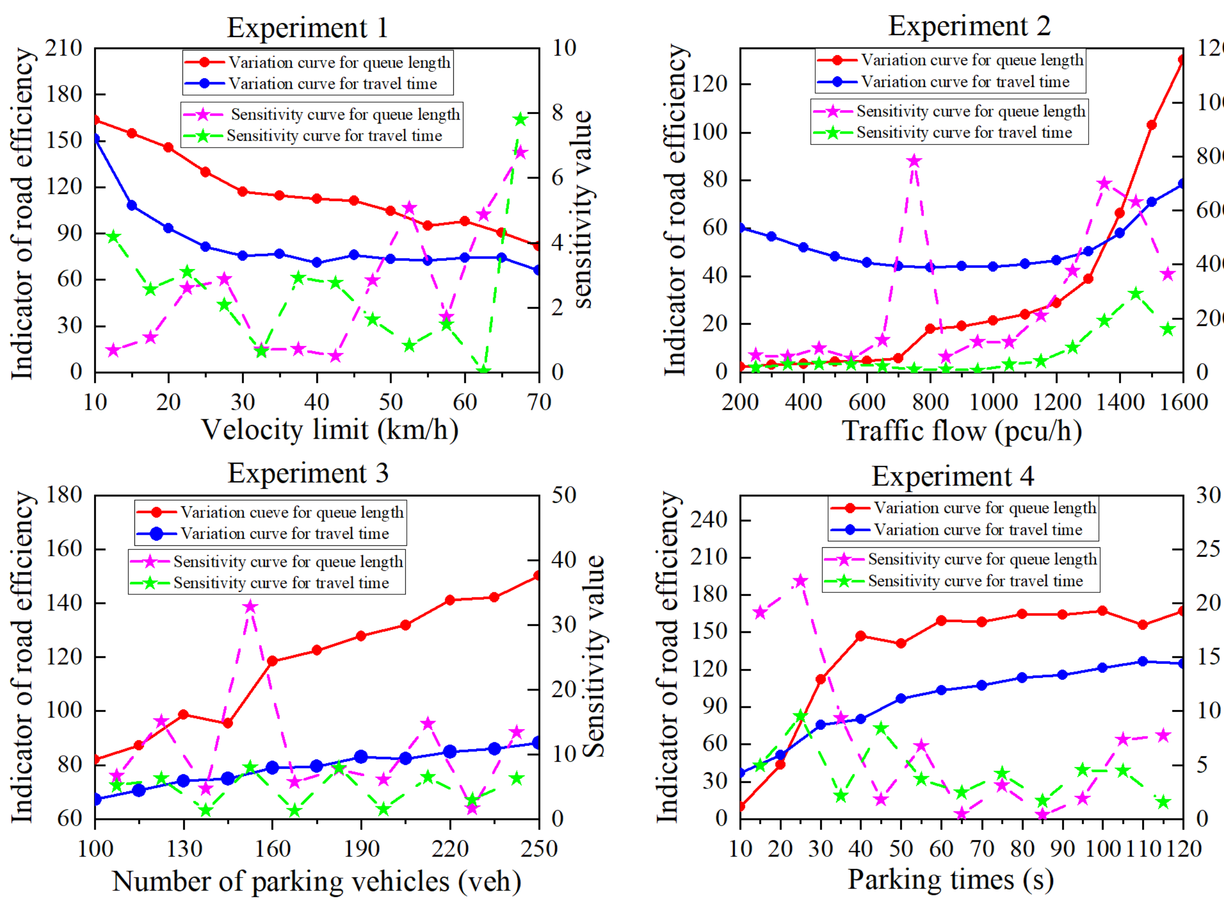

Fig. 7. Section-level evaluation indicator variations and sensitivity curves

Table III lists the section-level sensitivity factors for the different experiments. According to the size of the sensitivity coefficients, the factors are traffic flow, number of parked vehicles, stopping time, and speed limit.

TABLE III
SENSITIVITY FACTORS OF SECTION-LEVEL EVALUATION INDICATORS

Indicator	Expt. 1	Expt. 2	Expt. 3	Expt. 4
Queue length	6.78	781.56	32.76	22.04
Travel time	7.80	291.21	7.94	9.53

Fig. 8 shows the speed limit as a function of the lane-level volumes and sensitivity curves. Table IV lists the characteristics of the variations of these curves. The traffic flow intervals in the downstream section and lane C were more affected by the speed limit. In addition, they had relatively high concentrations in the sensitive intervals. Table V lists the flow sensitivity factors for different road locations. The downstream section and lane C were the most sensitive to road speed-limit values, whereas lanes A and B were less sensitive. According to the purpose of vehicle travel, lanes A and B were primarily for vehicles that dropped off and picked up students, whereas lane C was primarily for fast traffic on the road. Calculations showed that the speed limit had a weak effect on pick-up and drop-off traffic but a larger effect on passing traffic. Hence, when setting speed limits on school roads, emphasis should be placed on the traffic volume and speed.

Fig. 9 illustrates the speed limit as a function of occupancy and sensitivity curves. Table VI lists the characteristics of the variations in the turns. The occupancy intervals of the downstream section and lane B were strongly influenced by this phenomenon. The downstream section and lane C exhibited relatively high concentrations at the sensitive intervals. Table VII presents the occupancy sensitivity factors for different locations of the road. The downstream section and lane C were more sensitive to this. These results suggest that the lane choices for both types of traffic should be guided in the upstream section. A varying speed limit affects the efficiency of traffic dissipation and should be divided by lanes and promptly lifted. Specifically, setting lower speed limits on downstream sections and higher speed limits on upstream sections can improve the dissipation efficiency of traffic flow.

TABLE IV
CHARACTERIZATION OF LANE-LEVEL FLOW VARIATIONS IN EXPERIMENT 1

Characteristic	Cross Section			Lane		
	D	P	U	A	B	C
Tendency	It increases first, then stabilizes.					
Variation	252–413	253–409	259–419	210–285	213–301	341–654
Sensitive	10–30	10–30	10–20	10–30	10–20	10–30

Note: D = Downstream; P = Parking area; U = Upstream.

TABLE V
SENSITIVITY FACTORS FOR LANE-LEVEL FLOW IN EXPERIMENT 1

Cross-section	Lane			
	A	B	C	Average
Downstream	3.62	3.47	6.5	4.53
Parking area	2.39	3.00	5.05	3.48
Upstream	2.89	4.47	3.94	3.77
Average	2.97	3.65	5.16	3.93

TABLE VI
CHARACTERIZATION OF LANE-LEVEL FLOW VARIATIONS IN EXPERIMENT 1

Characteristic	Cross-section			Lane		
	D	P	U	A	B	C
Tendency	It increases first, then stabilizes.					
Variation	0.05–0.15	0.25–0.35	0.26–0.3	0.2–0.27	0.16–0.26	0.2–0.29
Sensitive	10–30	15–45	15–40	15–40	10–30	15–30

Note: D, Downstream; P, parking area; U, upstream.

TABLE VII
SENSITIVITY FACTORS FOR LANE-LEVEL OCCUPANCY IN EXPERIMENT 1

Cross-section	Lane			
	A	B	C	Average
Downstream	7.12	7.38	6.79	7.10
Parking area	4.81	4.36	3.77	4.31
Upstream	5.49	6.69	3.92	5.37
Average	5.81	6.14	4.83	5.59

Fig. 10 shows the traffic flow as a function of the lane-level volume and sensitivity curves. Table VIII lists the characteristics of the curve variations. The traffic volume intervals in the upstream section and lane C were more affected. The parking section and lane B exhibited relatively high concentrations at the sensitive intervals. Table IX provides the different locations of the road volume sensitivity factors for the road. The parking section and Lane C were the most sensitive. The results indicated that the parking section had the lowest capacity. Additional signals should be rationalized to control the number of vehicles entering a section. The downstream and lane C exhibited the highest capacity. The traffic flow should be rationally directed through signage and markings to accelerate vehicle dissipation.

TABLE VIII
CHARACTERIZATION OF LANE-LEVEL FLOW VARIATIONS IN EXPERIMENT 2

Characteristic	Cross-section			Lane		
	D	P	U	A	B	C
Tendency	It increases rapidly, then slows down.					
Variation	53–410	54–410	54–423	72–298	61–307	27–637
Sensitive	200–1500	200–1400	200–1600	200–1600	200–1300	200–1500

Note: D, Downstream; P, parking area; U, upstream.

TABLE IX
CHARACTERIZATION OF LANE-LEVEL FLOW VARIATIONS IN EXPT. 2

Cross-section	Lane			
	A	B	C	Average
Downstream	124.06	136.92	130.40	158.27
Parking area	130.40	123.32	244.15	165.96
Upstream	135.19	108.00	200.68	147.96
Average	129.88	122.75	219.56	157.40

Fig. 11 shows traffic flow as a function of occupancy and sensitivity curves. Table X lists the characteristics of the curve variations. The occupancy intervals of the upstream section and Lane A are strongly influenced by this phenomenon. Downstream and lane A had relatively high concentrations at the sensitive intervals. Table XI lists the occupancy sensitivity factors for different road locations. The upstream and lane A were more sensitive to this. The results indicate that the signal timing scheme should be further optimized to reduce the traffic pressure of upstream and through traffic in lane A.

TABLE X
CHARACTERIZATION OF LANE-LEVEL OCCUPANCY VARIATIONS IN EXPERIMENT 2

Characteristic	Cross-section			Lane		
	D	P	U	A	B	C
Tendency	It increases rapidly, then slows down.					
Variation	0.01 – 0.08	0.02–0.28	0.09–0.31	0.02–0.31	0.01–0.19	0.01–0.23
Sensitive	200–1500	200–1600	200–1600	200–1500	200–1600	200–1600

Note: D = Downstream; P = Parking area; U = Upstream.

TABLE XI
THE SENSITIVITY FACTORS FOR LANE-LEVEL OCCUPANCY IN EXPERIMENT 2

Cross-section	Lane			
	A	B	C	Average
Downstream	309.92	354.93	231.00	298.62
Parking area	179.80	271.13	395.18	282.04
Upstream	647.00	501.46	200.68	449.71
Average	378.91	375.84	275.62	343.46

Fig. 12 shows the number of parked vehicles as a function of the lane-level volume and sensitivity curves. Table XII lists the characteristics of the curve variations. The traffic volume intervals in the parking section and lane A are affected. The parking section and lane B exhibited relatively high concentrations at the sensitive intervals. Table XIII lists the volume sensitivity factors for the different road locations. The parking section and Lane B were the most sensitive. The results indicate that parking should be rationalized in terms of time and space.

TABLE XII
CHARACTERIZATION OF LANE-LEVEL FLOW VARIATIONS IN EXPERIMENT 3

Characteristic	Cross-section			Lane		
	D	P	U	A	B	C
Tendency	S	D	–	D	D	I
Variation	383–400	375–413	379–425	255–293	274–306	608–640
Sensitive	145–235	155–235	130–175	115–235	145–235	115–235

Note: D = Downstream; P = Parking area; U = Upstream; S = Stabilize; D = Decrease; I = Increase.

TABLE XIII
SENSITIVITY FACTORS FOR LANE-LEVEL FLOW IN EXPERIMENT 3

Cross-section	Lane			
	A	B	C	Average
Downstream	5.22	8.48	1.79	5.16
Parking area	10.13	7.78	5.32	7.74
Upstream	4.23	6.23	5.15	5.20
Average	6.53	7.50	4.09	6.03

Fig. 13 shows the number of parked vehicles as a function of occupancy and sensitivity curves. Table XIV lists the characteristics of variations in the curves. The occupancy intervals of the upstream area and lane C were strongly influenced by the number of parked vehicles. The parking section and lane B exhibited relatively high concentrations at the sensitive intervals. Table XV lists the occupancy sensitivity factors for different road locations. The parking section and lane B were the most sensitive. The results show that frequent vehicle movements into and out of parking spaces can severely affect the traffic operating conditions in lane B. Vehicle priorities should be rationalized to achieve orderly operation. For example, school buses and private cars jointly complete students' commuting tasks, which can effectively reduce this phenomenon.

TABLE XIV
CHARACTERIZATION OF LANE-LEVEL OCCUPANCY VARIATIONS IN EXPERIMENT. 3

Characteristic	Cross-section			Lane		
	D	P	U	A	B	C
Tendency	S	I	I	I	S	I
Variation	0.07-0.08	0.26-0.3	0.25-0.32	0.21-0.26	0.17-0.19	0.2-0.24
Sensitive	175-235	130-235	115-250	115-235	160-220	115-195

Note: D, Downstream; P, parking area; U, Upstream; S, stabilized; D, decreased; I, increased.

TABLE XV
SENSITIVITY FACTORS FOR LANE-LEVEL OCCUPANCY IN EXPERIMENT 3

Cross-section	Lane			
	A	B	C	Average
Downstream	50.42	51.64	14.49	38.85
Parking area	18.63	19.23	11.01	16.29
Upstream	14.44	10.45	11.72	12.20
Average	27.83	27.11	12.41	22.45

Fig. 14 shows the stopping time as a function of the lane-level flow and sensitivity curves. Table XVI lists the characteristics of variations of the curves. The traffic volume intervals in the parking section and Lane A were strongly affected. Lane C exhibited a relatively high concentration during the sensitive interval. Table XVII lists the volume sensitivity factors for the different road locations. The parking section and Lane A were the most sensitive. The results show that the stopping time can significantly impact both the aggregation and dissipation of pick-up and drop-off traffic.

TABLE XVI
CHARACTERIZATION OF LANE-LEVEL FLOW VARIATIONS IN EXPT. 4

Characteristic	Cross-section			Lane		
	D	P	U	A	B	C
Tendency	-	-	-	D	D	I
Variation	289-433	245-461	269-453	122-340	169-364	511-643
Sensitive	50-110	50-110	50-110	50-120	50-120	10-30

Note: D, Downstream; P, parking area; U, Upstream; S, stabilized; D, decreased; I, increased.

TABLE XVII
SENSITIVITY FACTORS FOR LANE-LEVEL FLOW IN EXPT. 4

Cross-section	Lane			
	A	B	C	Average
Downstream	9.16	11.01	3.09	7.75
Parking area	26.45	10.54	4.07	13.69
Upstream	8.50	9.05	2.47	6.67
Average	14.70	10.20	3.21	9.37

Fig. 15 shows the stopping time as a function of the occupancy and sensitivity curves. Table XVIII lists the characteristics of variations in the curves. The parked sections and lane-A intervals were significantly affected. The parking section and lane B exhibited relatively high concentrations at the sensitive intervals. Table XIX lists the occupancy sensitivity factors for the different road locations. The parking section and lane B were the most sensitive. The results show that the longer the stopping time, the fewer parking spaces lane A can provide, and the higher the parking section occupancy. Hence, restrictions should be imposed on the duration of vehicle parking based on the supply and demand of parking spaces to improve the turnover efficiency of parking spaces.

TABLE XVIII
CHARACTERIZATION OF LANE-LEVEL OCCUPANCY VARIATIONS IN EXPERIMENT 4

Characteristic	Cross Section			Lane		
	D	P	U	A	B	C
Tendency	D	I	I	I	I	I
Variation	0.05-0.08	0.16-0.36	0.1-0.3	0.11-0.3	0.1-0.21	0.1-0.26
Sensitive	10-110	10-40	10-40	10-40	10-30	0-30

Note: D = Downstream; P = Parking area; U = Upstream; S = Stabilize; D = Decrease; I = Increase.

TABLE XIX
SENSITIVITY FACTORS FOR LANE-LEVEL OCCUPANCY IN EXPERIMENT 4

Cross Section	Lane			
	A	B	C	Average
Downstream	8.87	10.21	3.51	7.53
Parking area	10.46	21.06	5.33	12.28
Upstream	12.76	9.98	13.72	12.15
Average	10.70	13.75	7.52	10.65

Multiple linear regression analysis is a mathematical method based on the correlation between the independent and dependent variables. In this study, the research variables were used as independent variables, and the evaluation indicators were used as dependent variables. The independent variables are not of the same order of magnitude and can significantly impact the dependent variable. Hence, the data for the independent variables are dimensionless, according to (5).

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{5}$$

where x' is the normalized data, x is the original data, and $x_{\min}$ and $x_{\max}$ are the minimum and maximum values of the original data, respectively.

If the dependent variable y correlates with the m variables, the generalized form of the multiple regression model is shown in Equation (6).

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \ldots + \beta_i x_m + \varepsilon \tag{6}$$

where $x_1, x_2 \ldots x_m$ is an independent variable, $\beta_0, \beta_1, \beta_2 \ldots \beta_m$ is a regression coefficient, and ε is a random error that follows the normal distribution $N\left(0, \sigma^2\right)$.

According to the sensitivity analysis, the variables and evaluation indicators were approximately linear within the sensitive intervals. Therefore, these data were used for analysis. In this study, 35 instances were designed as the dataset, and each instance contained four attributes: speed limits (x_1), traffic flow (x_2), number of parked vehicles (x_3), and stopping time (x_4). The multiple linear regression equations can be judged as significant by testing the regression coefficients and Equation (7).

$$F = \frac{SSR / p}{SSE / \left(n - p - 1\right)} \sim F\left(p, n - p - 1\right) \tag{7}$$

Where $SSR = \sum_{i=1}^{n} \left(\hat{y}_i - \bar{y} \right)^2$ and $SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$ are the regression and residual sums of squares, respectively.

To verify how well the multiple linear regression equation reflected the original data. The values in Tables XX and XXI are greater than 0.6, indicating a favourable level of response. The significance F values in Tables XX and XXII are less than 0.05, and the parameter constants are not zero. This shows that the regression equation is significant, and that it is credible to analyze the research variables and evaluation indicators using linear regression.

TABLE XX
ADJUSTED R SQUARE AND SIGNIFICANCE F AT THE SECTION LEVEL

Indicator	Adjusted R Square	Significance F
Queue length	0.874	0.000
Travel time	0.804	0.000

TABLE XXI
ADJUSTED R SQUARE AND SIGNIFICANCE F OF THE SQUARE OF LANE-LEVEL FLOW

Position	Adjusted R Square	Significance F
Cross-section	0.931	0.000
Lane A	0.737	0.000
Lane B	0.792	0.000
Lane C	0.721	0.000

TABLE XXII
ADJUSTED R SQUARE AND SIGNIFICANCE F FOR LANE-LEVEL OCCUPANCY

Position	Adjusted R Square	Significance F
Downstream	0.641	0.002
Parking area	0.725	0.000
Upstream	0.830	0.000
Lane A	0.806	0.000
Lane B	0.602	0.048
Lane C	0.810	0.000

Table XXIII shows the parameters of the multiple linear regression model for the section-level evaluation indicators. Queue length, except for x_1 the other variables, are significantly and linearly correlated. x_2 and x_4 have a large effect on it. For travel time, the variables are all significantly and linearly related to it. Parking on-road is one of the significant factors contributing to vehicle delays on this road. The ranking based on the impact on the section-level evaluation indicators are x_2 , x_3 , x_4 , and x_1 . This is consistent with the results of the sensitivity analysis. It illustrates the plausibility of the multiple linear regression equation. Controlling traffic volumes and optimizing the traffic organization of parking spaces can effectively enhance the efficiency of traffic operation on this road section.

TABLE XXIII
MULTIPLE LINEAR REGRESSION MODEL PARAMETERS AT THE SECTION LEVEL

Indicator	Variable	Regression coefficient	Standard error	t Stat	P-value
Queue length	x_1	–24.184	12.509	–1.93	0.077
	x_2	68.960	13.985	4.93	0.000
	x_3	27.149	12.381	2.19	0.05
	x_4	26.573	12.509	2.12	0.050
	C	–20.410	22.175	0.92	0.375
Travel time	x_1	–15.261	5.806	–2.68	0.022
	x_2	32.939	8.335	3.95	0.002
	x_3	28.432	5.747	4.98	0.000
	x_4	23.447	5.806	4.04	0.001
	C	37.827	10.293	3.68	0.003

Note: P-value is used to determine the level of significance; C = constant.

Table XXIV lists the parameters of the multiple linear regression model for lane-level flow. For the road cross section, except for x_3, the other variables were significantly and linearly related to lane volumes. For lanes A and B, except for x_1, the remaining variables were significantly and linearly related to the lane flow. x_2 and x_4 have the greatest impacts. This indicates that vehicle parking significantly affects traffic efficiency. For Lane C, except for x_3, all the other variables were significantly and linearly correlated. x_3 and x_4 were positively correlated with volume. This indicates that the traffic efficiency is less affected by on-road parking. Increasing the speed limit can improve the traffic efficiency. Lane C has a higher capacity and traffic efficiency. Lanes A and B were affected more by vehicle parking. Traffic guidance and lane-splitting speed limits are effective means of improving traffic efficiency.

TABLE XXIV
MULTIPLE LINEAR REGRESSION MODEL PARAMETERS FOR LANE-LEVEL FLOW

Position	Variable	Regression coefficient	Standard error	t Stat	P-value
Cross-section	x_1	7.615	3.229	2.36	0.036
	x_2	64.811	4.636	13.98	0.000
	x_3	–5.806	3.229	–1.80	0.097
	x_4	–7.326	3.196	–2.29	0.041
	C	349.782	5.725	61.10	0.000
Lane A	x_1	2.362	9.588	0.25	0.810
	x_2	41.656	13.764	3.03	0.011
	x_3	–37.020	9.588	–3.86	0.009
	x_4	–47.243	9.489	–4.98	0.000
	C	296.461	16.996	17.44	0.000
Lane B	x_1	1.858	8.671	0.21	0.834
	x_2	57.724	8.582	6.73	0.000
	x_3	–20.231	8.671	–2.33	0.036
	x_4	–26.834	12.448	–2.96	0.034
	C	333.103	15.371	21.67	0.000
Lane C	x_1	27.816	18.338	1.52	0.155
	x_2	128.755	26.326	4.89	0.000
	x_3	29.957	18.338	1.63	0.097
	x_4	82.647	18.150	4.55	0.001
	C	429.199	32.508	13.20	0.000

Note: *P*-value is used to determine the level of significance; C = constant.

Table XXV lists the parameters of the multiple linear regression model for lane-level occupancy. Downstream, x_1 and x_2 are significantly and linearly related to lane occupancy. This indicates that on-road parking has less impact. Increasing the speed limit can accelerate the dissipation rate of the traffic flow. For the parking area and upstream, except for x_1, the variables were significantly and linearly correlated. x_3 has the greatest impact. The spatial and temporal rationalization of the number of parked vehicles can alleviate congestion. For lanes A and B, except x_1 , the remaining variables were significantly and linearly related to occupancy. Although vehicle parking had a significant impact on occupancy, the impact on Lane A was greater. For Lane C, the variables were significantly and linearly correlated with occupancy. On-road parking has a significant impact on occupancy; however, increasing the speed limit can reduce congestion. In summary, different key factors affect the congestion levels at different locations. Vehicle parking has a large impact on lanes A and B but has minimal impacts downstream. Optimizing the traffic organization of parking spaces and timely release of speed limits can effectively reduce traffic congestion.

TABLE XXV
MULTIPLE LINEAR REGRESSION MODEL PARAMETERS FOR LANE-LEVEL OCCUPANCY

Position	Variable	Regression coefficient	S-error	t Stat	P-value
Downstream	x_1	–0.024	0.005	–5.10	0.000
	x_2	0.030	0.007	4.45	0.000
	x_3	0.003	0.005	0.68	0.508
	x_4	0.006	0.005	1.29	0.221
	C	0.093	0.008	11.00	0.000
Parking area	x_1	–0.029	0.017	–1.72	0.110
	x_2	0.071	0.024	2.98	0.012
	x_3	0.090	0.017	5.44	0.000
	x_4	0.080	0.016	4.85	0.000
	C	0.172	0.029	5.87	0.000
Upstream	x_1	–0.045	0.025	–1.79	0.098
	x_2	0.216	0.036	6.00	0.000
	x_3	0.149	0.025	5.92	0.000
	x_4	0.124	0.025	4.98	0.000
	C	0.050	0.045	1.13	0.282
Lane A	x_1	–0.018	0.015	–1.22	0.248
	x_2	0.100	0.021	4.81	0.000
	x_3	0.052	0.015	3.59	0.001
	x_4	0.081	0.014	5.64	0.000
	C	0.099	0.026	3.85	0.002
Lane B	x_1	–0.002	0.015	–0.15	0.885
	x_2	0.075	0.021	3.52	0.004
	x_3	0.040	0.015	2.70	0.021
	x_4	0.039	0.015	2.65	0.021
	C	0.082	0.026	3.10	0.009
Lane C	x_1	–0.045	0.016	–2.80	0.016
	x_2	0.118	0.023	5.11	0.000
	x_3	0.072	0.016	4.51	0.001
	x_4	0.087	0.016	5.42	0.000
	C	0.105	0.028	3.70	0.003

Note: *P*-value is used to determine the level of significance; C = constant; S-error = Standard error.

V. TRAFFIC INDICATOR PROJECTIONS

To verify the reliability of the analysis of factors influencing traffic status, we applied BP and gated recurrent NNs to predict school road section-level and lane-level traffic indicators. We analyzed the results by applying different variables for the prediction based on the calculations above.

A. Genetic Algorithm (GA)-Optimized Backpropagation (BP) NN

Typically, the weights and thresholds of the BPNN are initialized with random values, as the initialization parameters significantly influence network training. However, this can restrict the consistency and generalization of the training results. The GA, a global search-optimization algorithm based on the population, uses individual fitness as a criterion and includes a variation factor. This gives it strong global search capabilities while also enhancing local search. Hence, the GA is used to optimize the initial weights and thresholds of the BPNN to improve the network performance.

The GA–BP algorithm has three main components: (1) BPNN structure determination. The structure of the network includes nodes of the input and output layers, as well as a hidden layer and its nodes. The number of input and output nodes are assigned according to the specific problem. The number of hidden layer nodes is determined by (8); (2) The GA optimizes the weights and thresholds. The selection operator uses a roulette-wheel selection strategy. The crossover and variation operators are determined adaptively, i.e., the crossover and variation rates change automatically with the fitness of the population; (3) BPNN training and prediction. Network training is a process of constantly correcting the weights and thresholds to reduce output errors. Fig. 16 illustrates the flow of the GA-BP algorithm.

$$k = \sqrt{q+p} + a \tag{8}$$

where k is the number of hidden layer nodes, and q and p are the numbers of input and output nodes, respectively; and a usually ranges from 1 to 10.

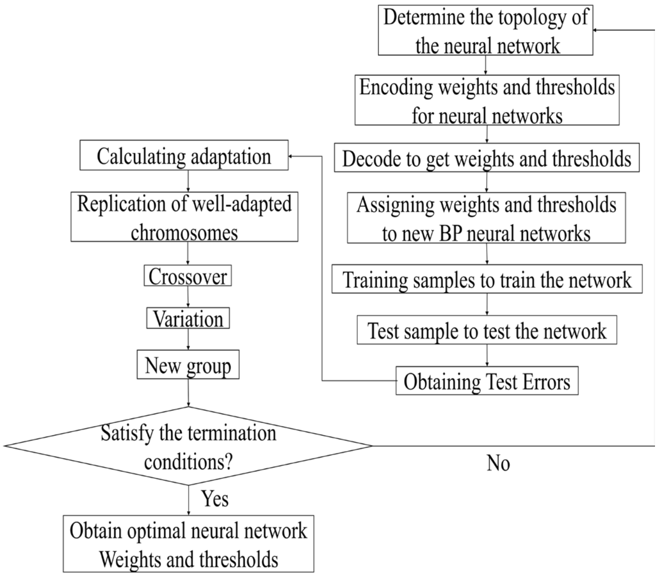

Fig. 16. GA–BP flowchart

B. Gated Recurrent Unit (GRU) NN

Recurrent neural networks (RNNs) are also suitable for traffic flow [47]. However, RNNs suffer from gradient vanishing or explosion, making their application in practical scenarios more challenging. The long short-term memory (LSTM) network addresses the gradient-explosion problem of the original RNN. However, compared to the LSTM, the gated recurrent unit (GRU) simplifies the model by merging the internal self-loop cell and hidden layer, reorganizing the input and forget gates into a single update gate, and introducing a new reset gate, t_r. This modification allows for a more straightforward computation of the hidden state, thereby effectively reducing the model's prediction time. Therefore, GRU was adopted in this study.

The update and reset gates in the GRU combine the inputs of the current time step with the hidden state h_{t-1} from the previous time step. The output values are computed using a fully connected layer with sigmoid activation functions. Here, W_z and W_r represent the weights of the update and reset gates, respectively. The update process is as follows:

Step 1: Update the gate z_t: Determine the influence of the previous state on the current state.

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] \tag{9}$$

Step 2: Reset gate: r_t determines the extent to which the previous state is ignored.

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] \tag{10}$$

Step 3: Update hidden state: $\bar{h}_t$ represents the candidate hidden state, whereas h_t is the updated hidden state.

$$\bar{h}_t = \tanh(W_{\bar{h}_t} \cdot [r_t * h_{t-1}, x_t]) \quad (11)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \bar{h}_t \quad (12)$$

When the value of the parameter in the reset gate approaches zero, the hidden state from the previous time step should be discarded, whereas a value close to 1 indicates that it should be retained. The candidate hidden state formula shows that the reset gate uses the hidden state of the previous time step to update the candidate hidden state of the current time step, thereby effectively discarding irrelevant historical information. Meanwhile, the update gate controls the update pattern of the hidden state, allowing the capture of dependencies within the time series.

C. Traffic Indicator Projections

In this study, we established a simulation environment to gather data for predicting section-level and lane-level traffic indicators. The simulations were iterated 10 times (10 simulation seeds) to ensure an adequate number of training samples. The average of these ten runs was predicted, and the discrepancy between the expected and predicted values served as the predictive effectiveness measure (error). We used BPNNs and a GRU to compare and analyze the predictive impact of various variables, juxtaposing them with the regression analysis results to validate the viability of the proposed method.

Table XXVI lists the average errors in predicting queue length for each influencing factor. The results from both methods consistently demonstrated that the smallest prediction error occurred when traffic flow was used as the input parameter, whereas the largest error arose when the road-section speed limit was used. This aligns with the findings that the traffic state influences the factor identification method proposed in this study. Consequently, incorporating traffic factors that significantly affect the queue length can enhance the prediction accuracy.

Fig. 17 and 18 illustrate the queue length-prediction results obtained using the two methods. For the GA–BPNN, the prediction errors using the traffic flow, parked vehicles, stopping time, and speed limit datasets were 6.21%, 15.99%, 16.75%, and 17.57%, respectively. These values reflect reductions of 9.78%, 10.54%, and 11.36%, respectively, compared with the other methods. For the GRU, the prediction errors using the same datasets were 10.37%, 11.58%, 13.36%, and 16.41%, reflecting reductions of 1.20 %, 2.98 %, and 6.03 %, respectively. These findings further indicate that larger prediction errors occur when data fluctuations are significant, particularly within the sensitivity intervals of each influencing factor. Therefore, minimizing errors within these sensitive intervals should be a key focus for improving prediction accuracy.

TABLE XXVI
AVERAGE ERRORS OF QUEUE LENGTH PREDICTION FOR DIFFERENT TRAINING SETS

Method	Training set			
	Traffic flow	Parking vehicles	Stopping time	Speed limits
GA–BP	1.96	5.05	5.29	5.55
GRU	12.24	13.66	15.76	19.36

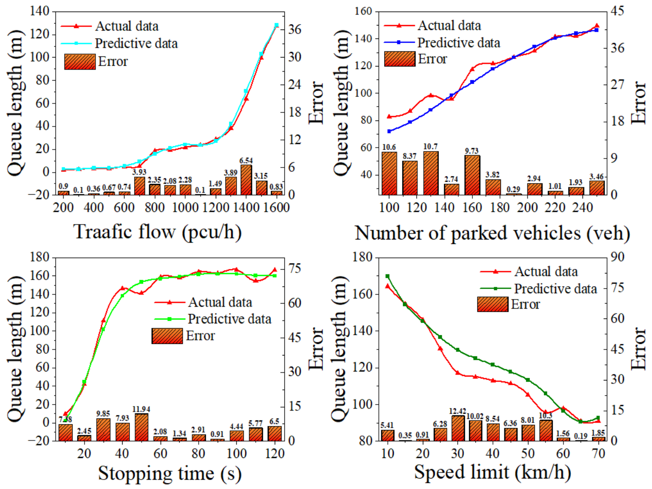

Fig. 17. Queue-length prediction errors when using the BPNN

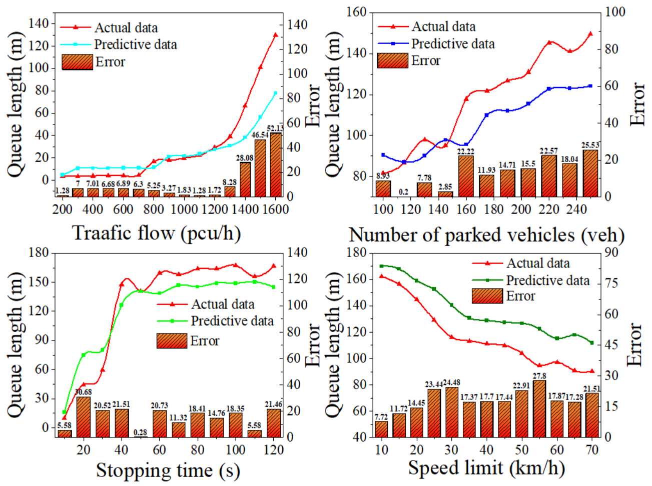

Fig. 18. Queue-length prediction error results when using the GRU

TABLE XXVII
AVERAGE ERRORS IN TRAVEL-TIME PREDICTION FOR DIFFERENT TRAINING SETS

Method	Training set			
	Traffic flow	Parking vehicles	Stopping time	Speed limits
GA–BP	0.84	1.04	1.83	2.00
GRU	4.01	6.57	10.33	13.73

Table XXVII presents the average error in predicting travel time for each influencing factor. The effectiveness of the predictions was ranked as follows: traffic flow, parked vehicles, stopping time, and speed limits. This ranking aligns with the findings of the traffic state-influencing factor identification method proposed in this study. Therefore, utilizing the traffic factors that most significantly affect travel time can improve prediction accuracy.

Fig. 19 and 20 illustrate the travel-time prediction results of the two methods. For the GA–BP NN, the prediction errors using traffic flow, parked vehicles, stopping time, and speed limit datasets were 1.60%, 1.98%, 3.48%, and 3.81%, respectively, reflecting reductions of 0.38%, 1.88%, and 2.21%, respectively. For the GRU, the prediction errors using the same datasets were 5.08%, 8.32%, 13.08%, and 17.38%, with corresponding reductions of 3.24%, 8.00%, and 12.03%, respectively. The results also indicate that larger prediction errors occurred when there were significant data fluctuations, particularly within the sensitivity intervals of each influencing factor. Therefore, reducing errors within these sensitive intervals can significantly enhance the accuracy of travel-time predictions.

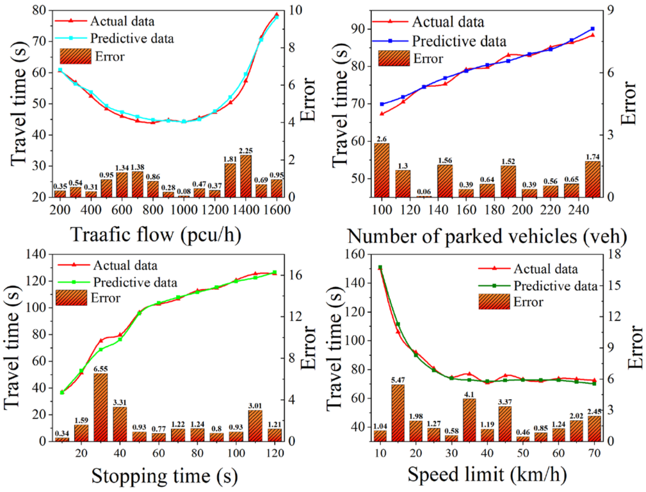

Fig. 19. Travel time prediction error results when using the BPNN

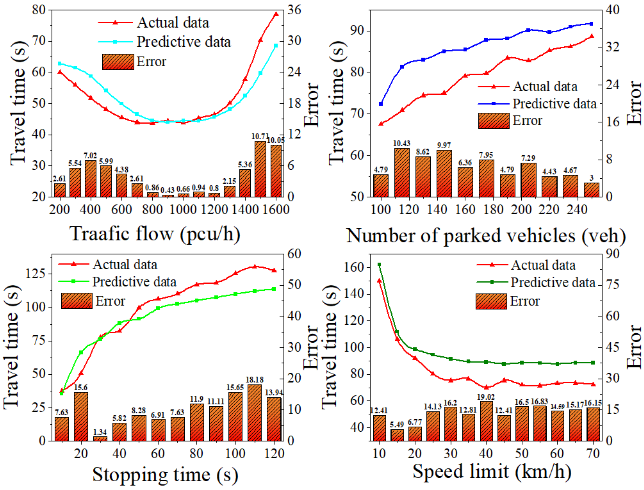

Fig. 20. Travel-time prediction error results when using the GRU

TABLE XXVIII
AVERAGE ERRORS IN LANE-LEVEL FLOW PREDICTION FOR DIFFERENT TRAINING SETS

Method	Training set			
	Traffic flow	Parking vehicles	Stopping time	Speed limits
GA-BP	2.73	2.78	4.76	8.01
GRU	16.29	17.99	21.60	22.67

Table XXVIII presents the average error in predicting lane-level flow for each influencing factor. The smallest prediction error was observed for traffic flow, whereas the largest error was associated with speed limits. This finding is consistent with the results of the traffic state influencing the factor identification method proposed in this study. Therefore, incorporating traffic factors that significantly impact lane-level flow can improve prediction accuracy.

Fig. 21 and 22 show the lane-level flow prediction results obtained using the two methods. For the GA-BP NN, the prediction errors with traffic flow, parking vehicles, stopping time, and speed limit datasets were 1.29%, 1.32%, 2.26%, and 3.80%, respectively, representing reductions of 0.02%, 0.96%, and 2.50%, respectively. For the GRU, the prediction errors with the same datasets were 5.42%, 5.99%, 7.19%, and 7.55%, with corresponding reductions of 0.57 %, 1.77 %, and 2.12 %, respectively. The results indicated that larger errors occurred when data fluctuations were significant, particularly within the sensitivity intervals of each influencing factor. Therefore, minimizing errors within these sensitive intervals is crucial for improving prediction accuracy.

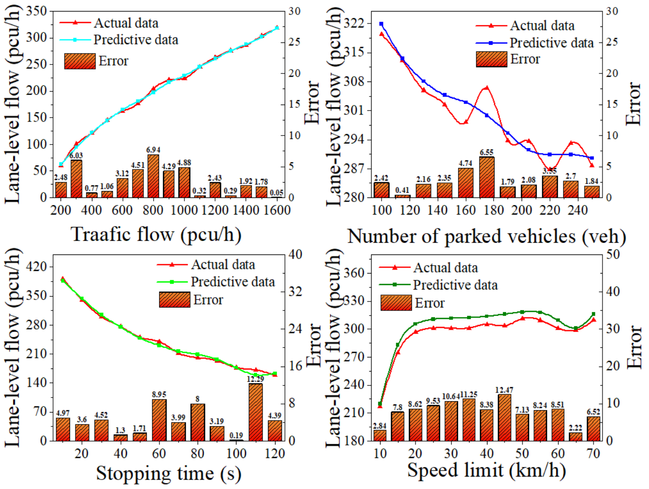

Fig. 21. Lane-level flow prediction error results when using the BPNN

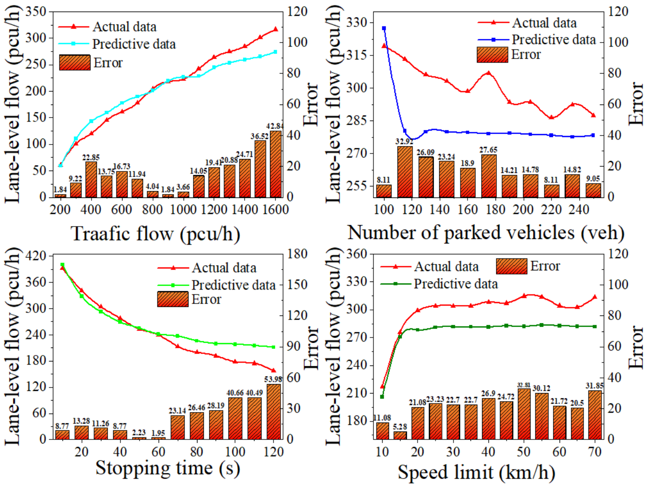

Fig. 22. Lane-level flow prediction error results when using the GRU

Table XXIX presents the mean error in predicting the lane-level occupancy for each influencing factor. The effectiveness of the predictions was ranked as follows: Traffic flow > Parking vehicles > Stopping time > Speed limits. This ranking aligns with the findings of the traffic state-influencing factor identification method proposed in this study. Therefore, incorporating the most significant traffic factors influencing lane-level occupancy can improve the prediction accuracy.

Fig. 23 and 24 display the lane-level occupancy prediction results obtained using the two methods. For the GA–BPNN, the prediction errors for the traffic flow, parking vehicles, stopping time, and speed limit datasets were 19.66%, 50.44%, 52.38%, and 62.88%, with reductions of 30.79%, 32.72%, and 43.22%, respectively. For the GRU, the prediction errors using the same datasets were 15.95%, 17.01%, 20.86%, and 22.66%, with reductions of 1.06%, 4.91%, and 6.71%, respectively. The results indicated that greater data fluctuations led to larger prediction errors, particularly within the sensitivity intervals of each influencing factor. Thus, minimizing the errors within these sensitive intervals is crucial for improving the accuracy of lane-level occupancy predictions.

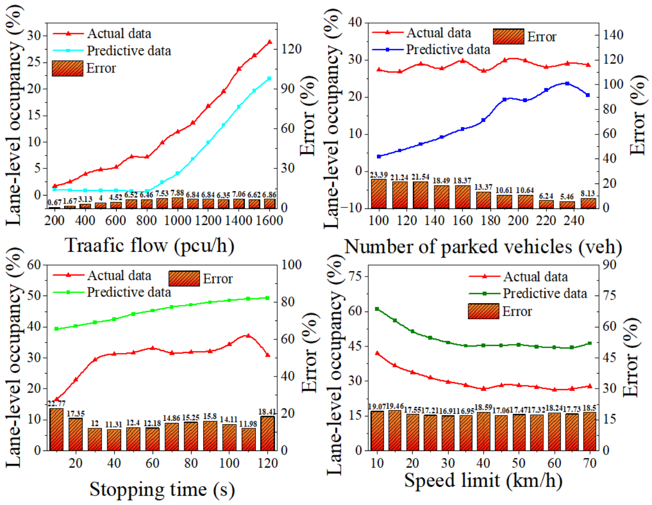

Fig. 23. Lane-level occupancy prediction error results when using the BPNN

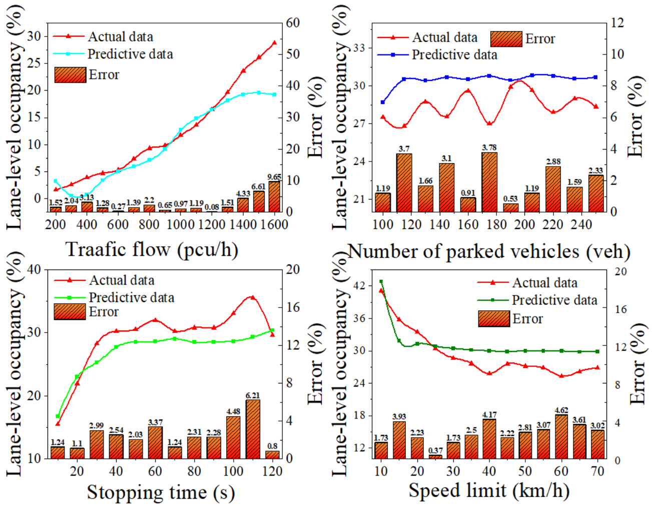

Fig. 24. Lane-level occupancy prediction error results when using the GRU

TABLE XXIX
AVERAGE ERRORS IN LANE-LEVEL OCCUPANCY PREDICTION FOR DIFFERENT TRAINING SETS

Method	Training set			
	Traffic flow	Parking vehicles	Stopping time	Speed limits
GA–BP	5.58%	14.32%	14.87%	17.85%
GRU	1.95%	2.08%	2.55%	2.77%

VI. CONCLUSIONS

This study proposed an analytical methodology that can identify the key factors and sensitive ranges affecting traffic status at the section and lane levels. Consequently, optimization strategies for improving traffic efficiency were designed. We applied these findings to traffic forecasting. The predicted effects were generally consistent with the analysis results. This enhances the precision of the improvement strategy and improves interpretability in machine learning, such as traffic prediction. The main findings of the study are as follows:

Differentiation analysis compared the traffic conditions across the lanes in the model. Lane differentiation was demonstrated in terms of traffic distribution, speed–flow relationship, and capacity. The analysis results show that, except for the upstream area and parking area of lane B, they all show significant differences. In particular, the differences between lane C and downstream and other locations are more evident. At the same time, we also found that the higher the volume of traffic, the greater the variation among the lanes, and the less efficiently they operate on the road. In our study scenario, on-road parking is the main reason for the differentiation of traffic conditions among different lanes, which can be further interpreted as the impact of infrastructure on the surrounding traffic flow.

The sensitivity analysis method was used to obtain the overall trend between the sensitive factors affecting the traffic efficiency at the section level and lane level and the evaluation index. This study proposed a dimensionless sensitivity function and a sensitivity factor to enable a comparable multifactor sensitivity analysis. Secondly, we also use a multiple linear regression model to quantify the relationship between factors and evaluation indicators. The results show that the sensitive interval of the same factor is different in different road locations. In the sensitive interval, the relationship between factors and indexes is approximately linear. The key factors affecting the traffic conditions are different in each location. For example, the key factors affecting the flow of lane A and lane B are parking time and traffic flow, respectively. These results provide optimization directions for us to improve the traffic efficiency of the road section around the school. In response to this, we have also put forward some traffic control measures for different sections. For example, you can consider guiding the separation of commuter traffic from pick-up and drop-off traffic in the upstream section.

The GA–BPNN and GRU were employed to forecast the section-level and lane-level traffic indicators. Various variables were used to predict these indicators and were ranked based on their predictive effectiveness. Traffic flow was the most accurately predicted factor, followed by parking vehicles and stopping time, whereas speed limits showed the least accuracy, exhibiting a large prediction error within the calculated sensitivity interval of the traffic-influencing factors. Using the proposed methodology, the prediction accuracies for the queue length, travel time, lane-level flow, and lane-level occupancy improved by 11.36%, 12.03%, 3.80%, and 6.71%, respectively. These results are consistent with those of the proposed method, enhancing the interpretability of machine learning in transportation. This approach is beneficial not only for forecasting but also for establishing a scientific foundation for identifying traffic issues and developing optimization strategies.

The research method we proposed in this paper is not only applicable to the sections where infrastructure facilities (such as schools, hospitals, shopping malls, etc.) have a large impact on the surrounding traffic flow but also to the sections where there are obvious differences in lane-level traffic status due to road construction, traffic accidents, landslides and other factors. This method has certain portability. In addition, the influencing factors and evaluation indicators selected in this paper are suitable for the urban road scene around the school. In the future, other factors such as construction period and traffic accident grade can be added to other scenes for research. Therefore, the research method in this paper has a certain scalability.

This can be further examined through the following avenues: (1) Examining more intricate heterogeneous traffic patterns, such as the proportion of nonmotorized vehicles and pedestrians and analyzing the impact of variables on traffic efficiency; (2) Verifying the relevance of research findings in areas such as traffic safety within the proposed research framework. (3) Employing various deep learning algorithms to scrutinize the prediction errors based on the study results and exploring additional machine learning prediction methods for increased applicability.

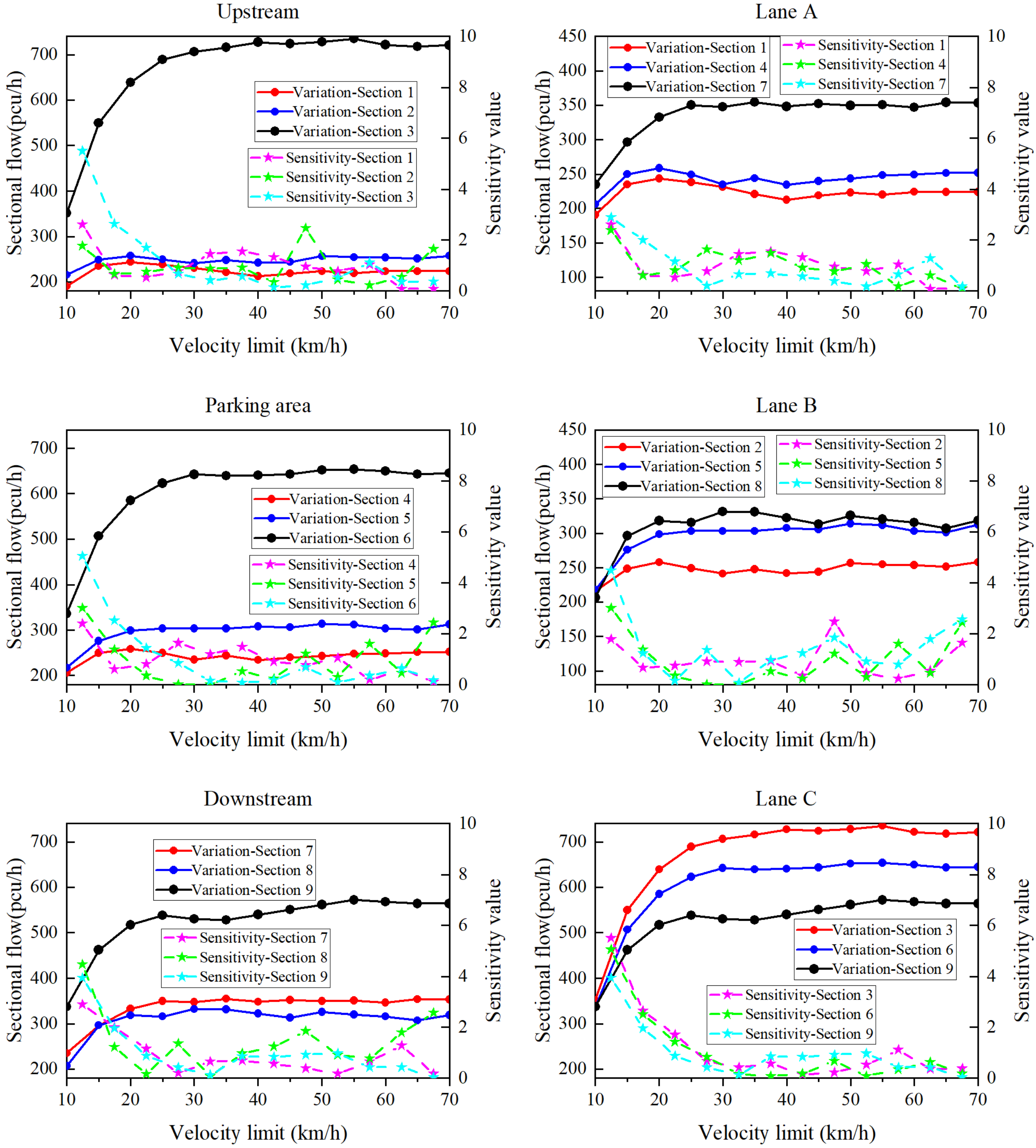

Fig. 8. Lane-level flow variations and sensitivity curves for Expt. 1

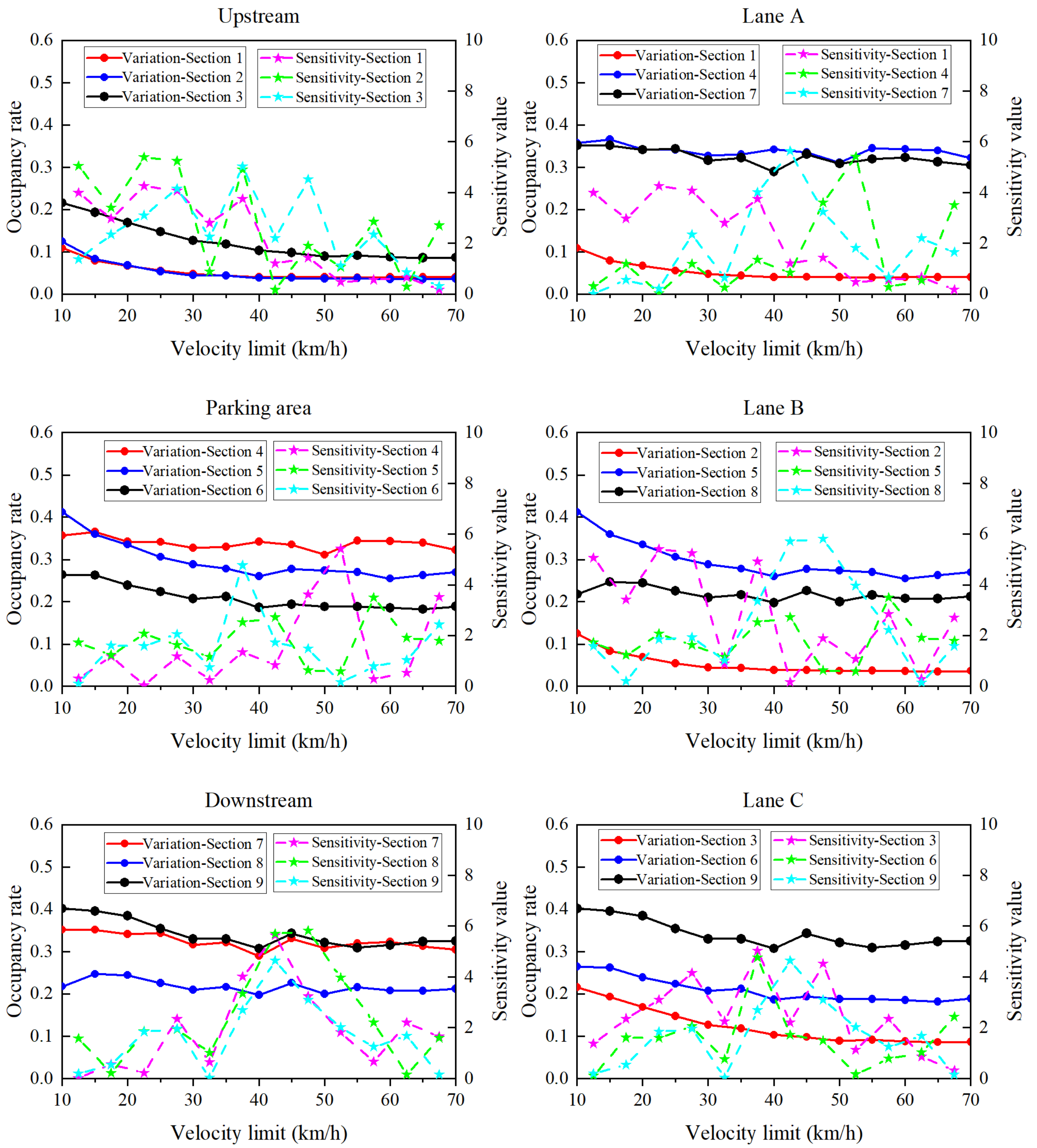

Fig. 9. Lane-level occupancy variations and sensitivity curves for Expt. 1

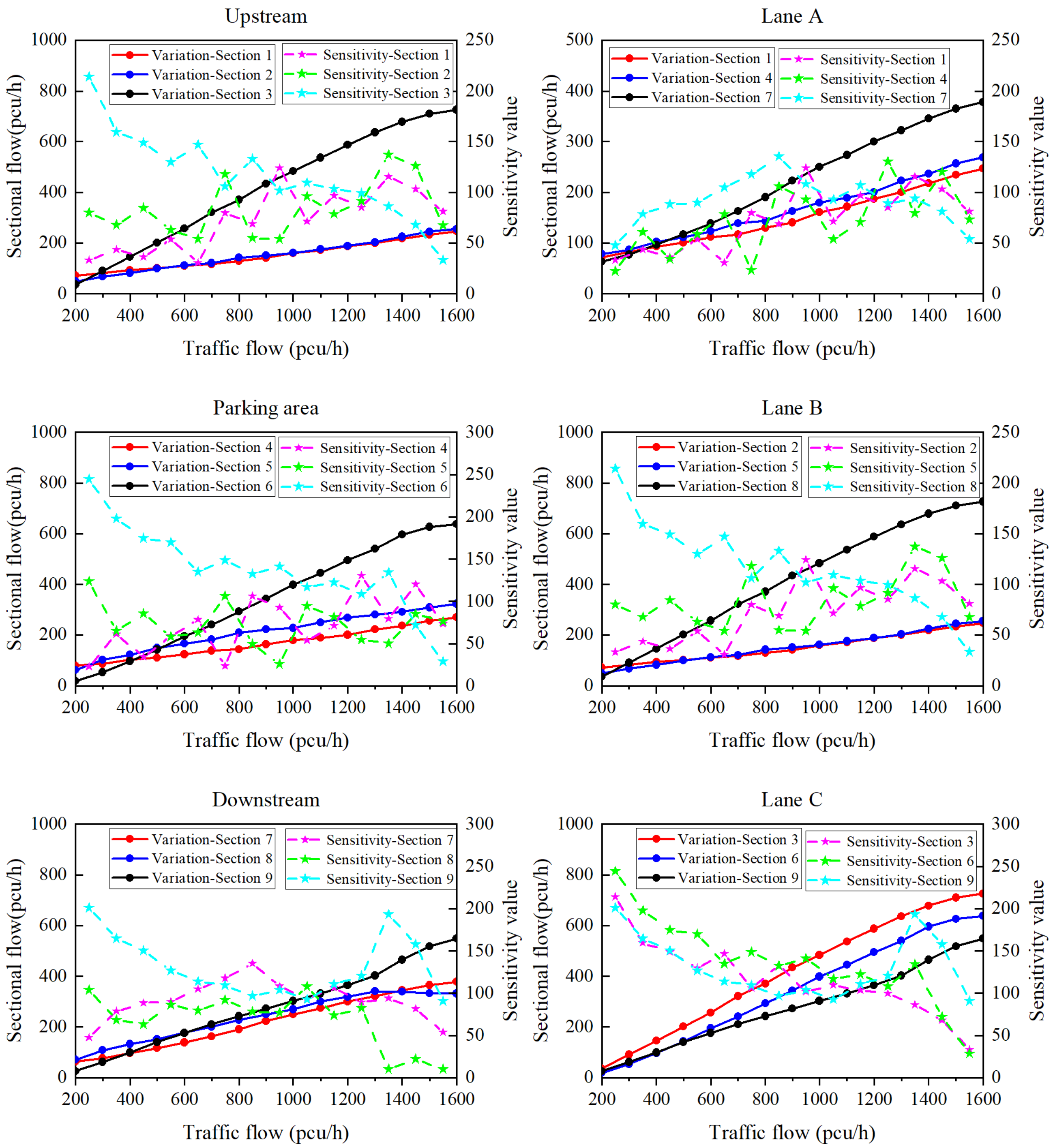

Fig. 10. Lane-level flow variations and sensitivity curves for Expt. 2

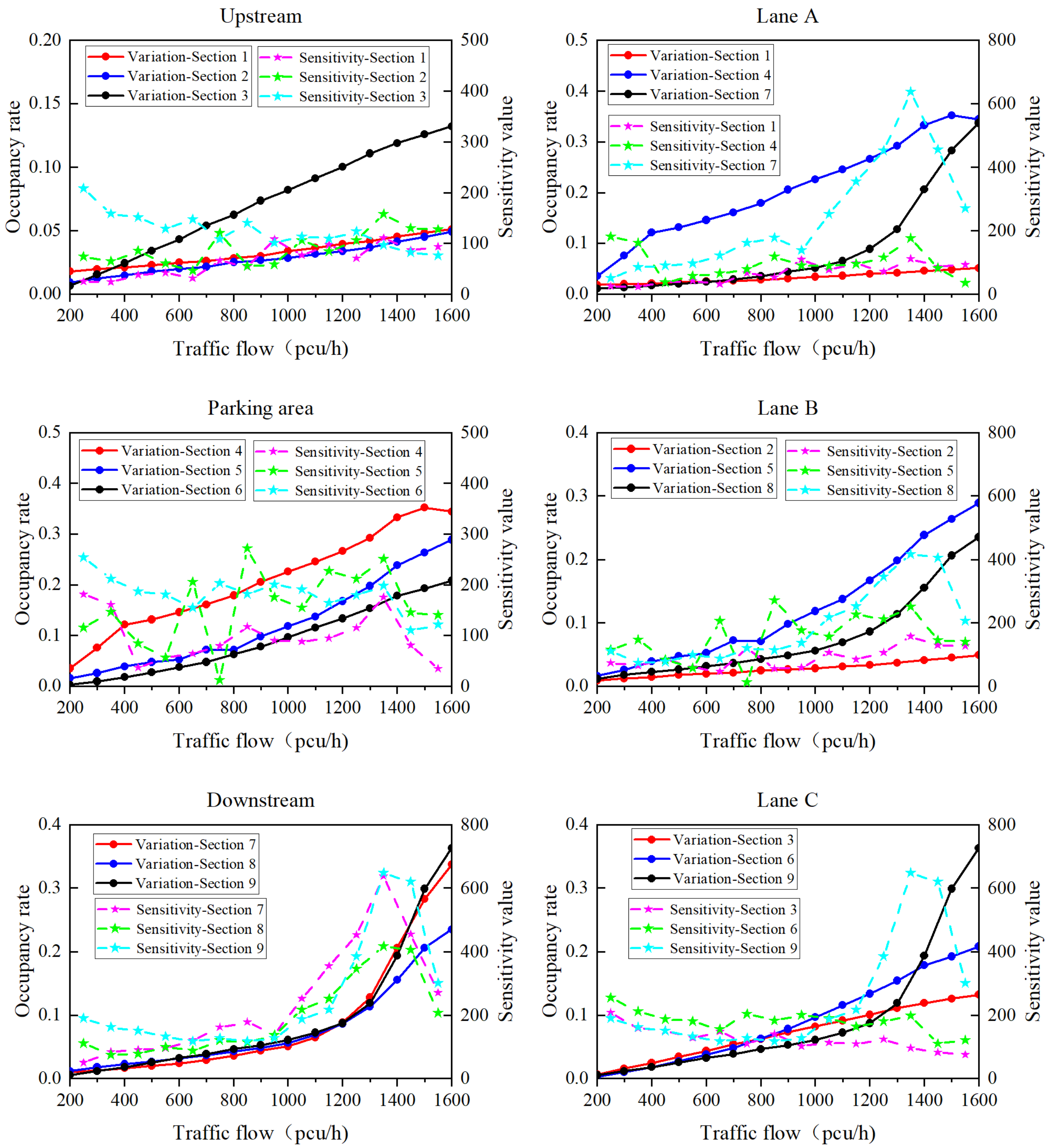

Fig. 11. Lane-level occupancy variations and sensitivity curves for Exp. 2

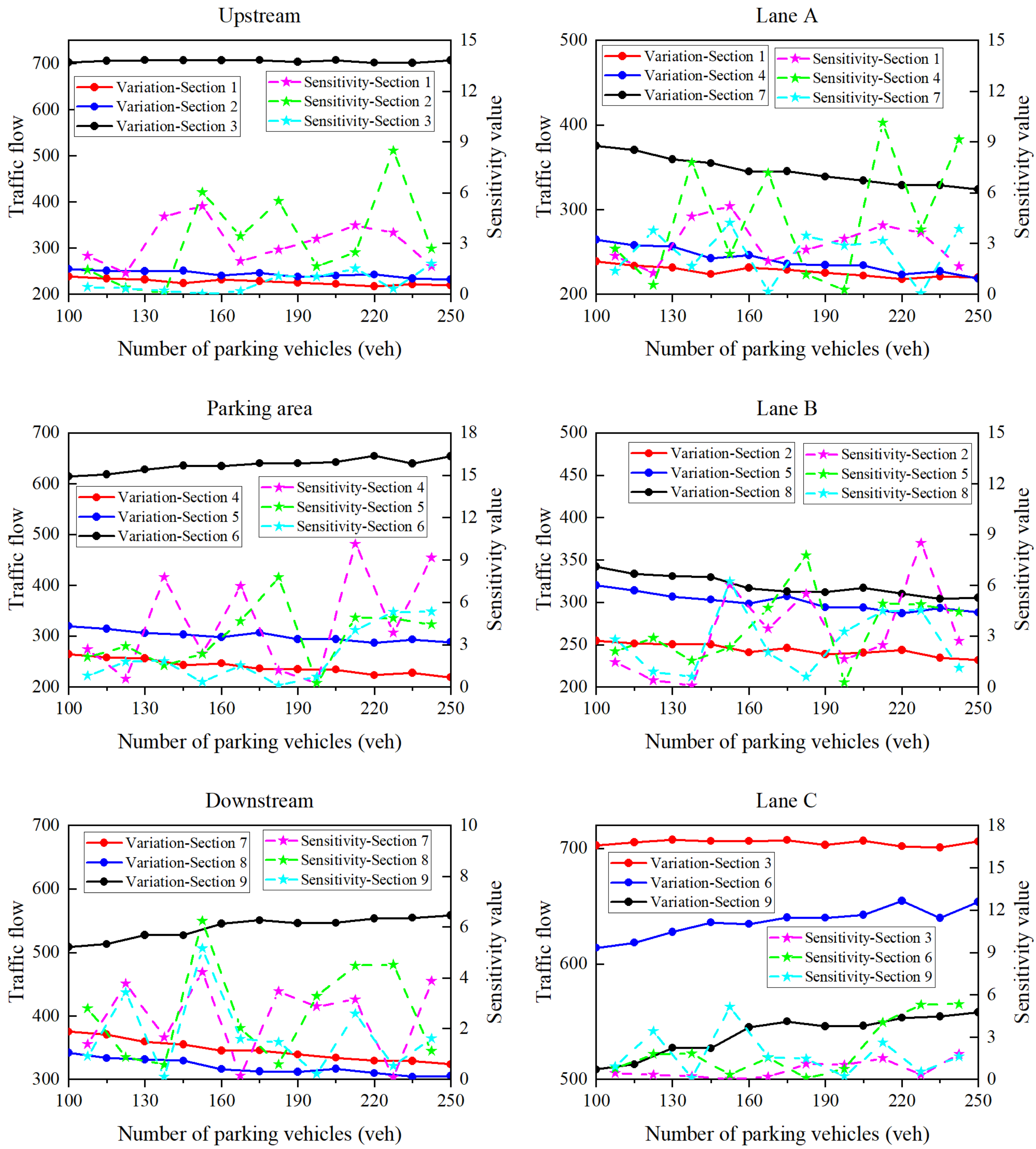

Fig. 12. Lane-level flow variations and sensitivity curves for Expt. 3

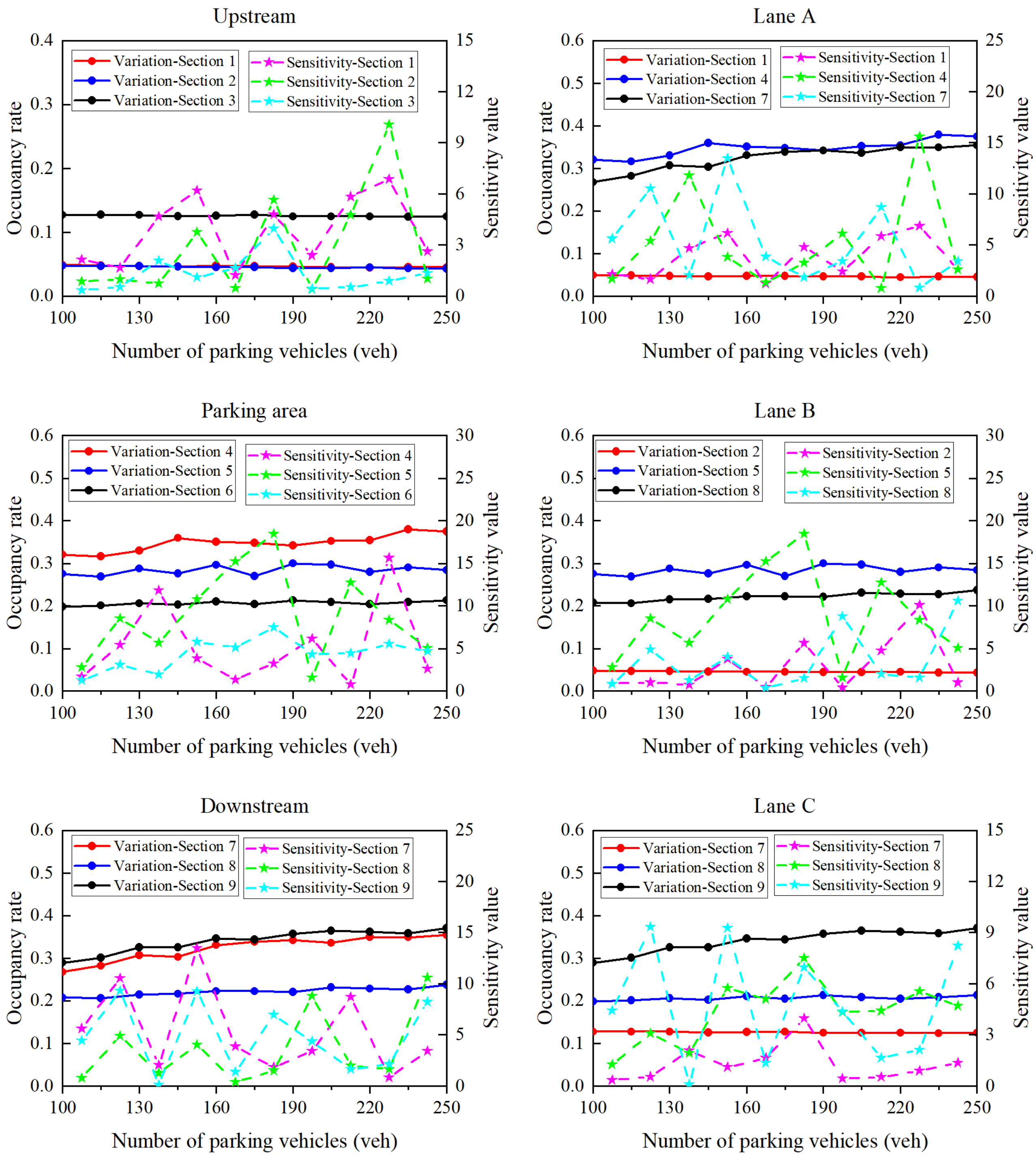

Fig. 13. Lane-level occupancy variations and sensitivity curves for Expt. 3

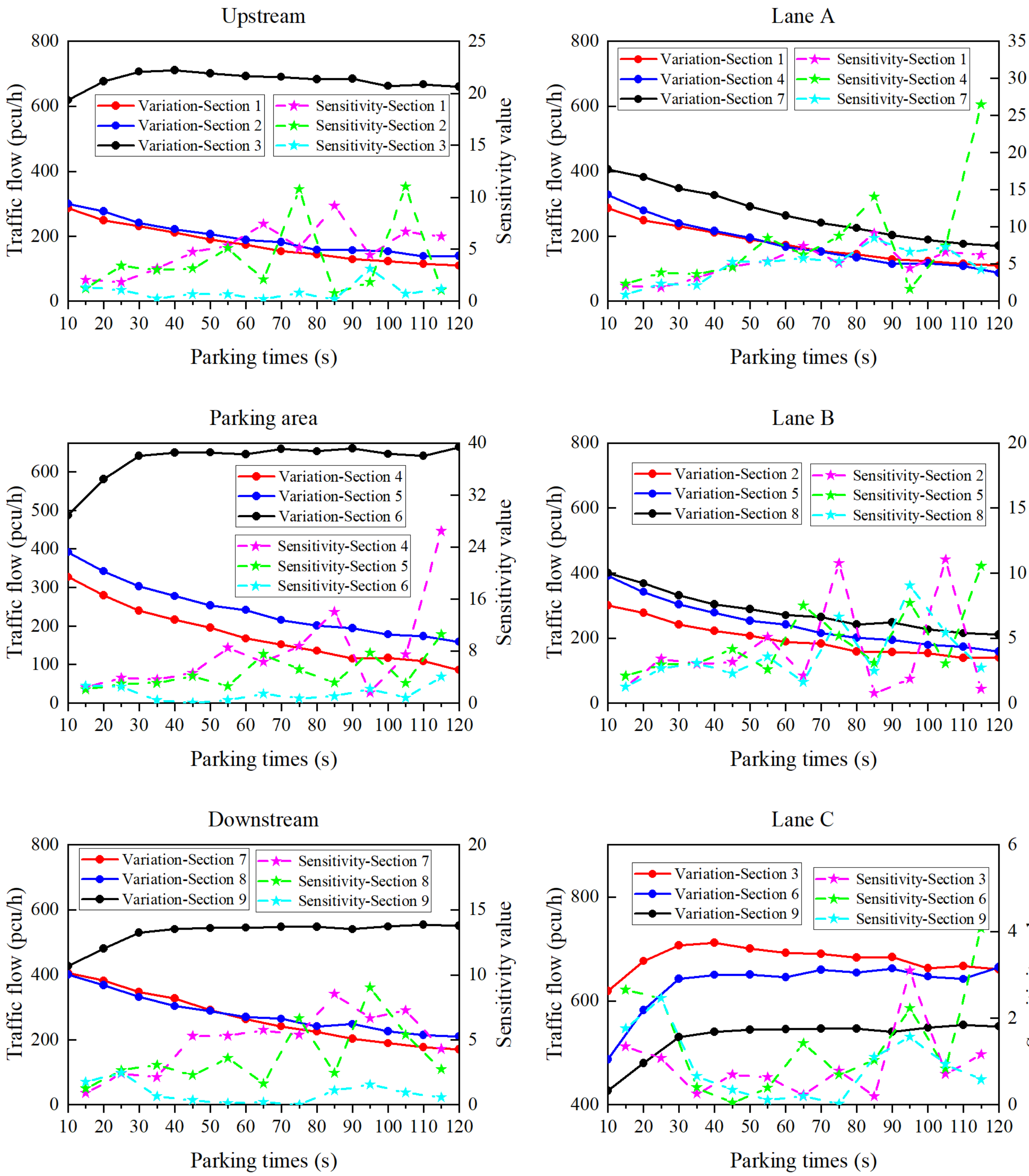

Fig. 14. Lane-level flow variations and sensitivity curves for Exp. 4

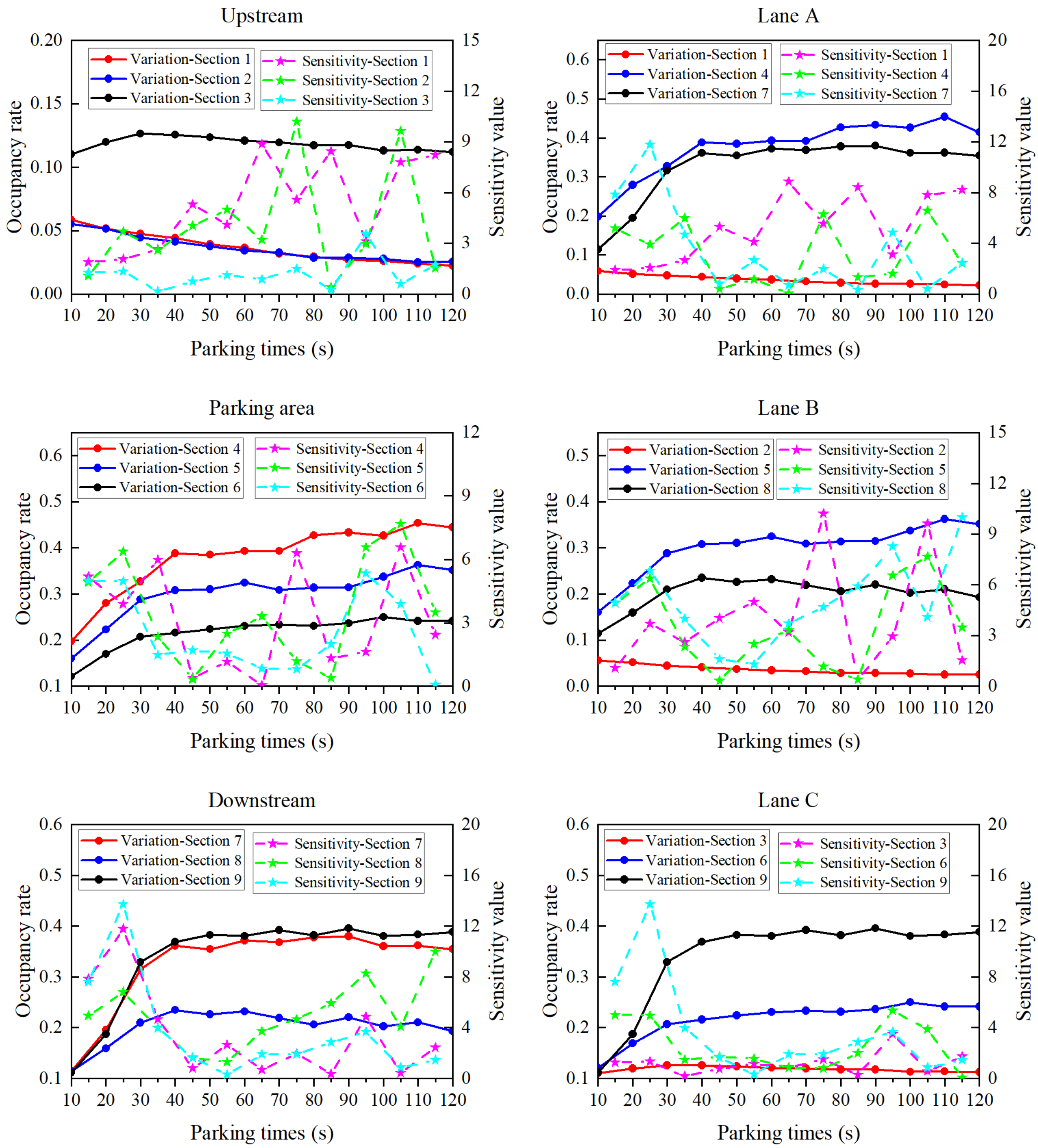

Fig. 15. Lane-level occupancy variations and sensitivity curves for Exp. 4

REFERENCES

[1] M. Fred, "An empirical analysis of driver perceptions of the relationship between speed limits and safety," *Transportation Research Part F*, vol.12, no. 2, pp. 99–106, 2008.

[2] G. Nikolas and C. F. Daganzo, "Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings," *Transp. Res. B Methodol*, vol. 42, no. 9, pp. 759–770, 2008.

[3] A. Soyoung, V. Sravani, and L. Jorge, "A method to account for non-steady state conditions in measuring traffic hysteresis," *Transp. Res. C*, vol. 34, pp. 138–147, 2011.

[4] Lily A, "The Highway Capacity Manual," *The National Academies Press*, vol. 6, p. 41, 2016.

[5] M. M. E. Sherief, I. M. I. Ramadan, and A. M. Ibrahim, "Development of traffic stream characteristics models for intercity roads in Egypt," *Alex. Eng. J.*, vol. 55, no. 3, pp. 2765–2770, 2016.

[6] J. Yuan, M. Abdel-Aty, Q. Cai, J. Lee, "Investigating drivers' mandatory lane change behavior on the weaving section of free way with managed lanes: A driving simulator study," *Transp. Res. Part F Traffic Psychol. Behav.*, vol. 62, pp. 11–32, 2019.

[7] A. M. Anas, M. A. N. Hazem, K. A. Wael, D. Charitha, and M. Babak, "Lane-based analysis of the saturation flow rate considering traffic composition," *Transp. Plann. Technol.*, vol. 46, no. 5, pp. 653–671, 2023.

[8] N. Bharadwaj, P. Kumar, S. S. Arkatkar, and G. Joshi, "Deriving capacity and level-of service thresholds for intercity expressways in India," *Transp. Lett.*, vol. 12, pp. 1–15, 2019.

[9] F. Marczak, L. Leclercq, and C. Buisson, "A macroscopic model for freeway weaving sections," *Comput. Aided Civ. Infrastruct. Eng.*, vol. 30, pp. 464–477, 2015.

[10] K. Nagel and M. A. Schreckenberg, "Cellular automaton model for freeway traffic," *J. Phys (France) I*, vol. 2, pp. 2221–2229, 1992.

[11] W. F. Adams, "Road traffic considered as a random series," *Oper. Res. Q.*, vol. 1, no. 1, pp.121-130,1936.

[12] M. A. Fedotkin and A. M. Fedotkin, "Analysis and optimization of output processes of conflicting Gnedenko-Kovalenko traffic streams under cyclic control," *Autom. Remote Control.*, vol. 70, no. 12, pp. 2024–2038, 2010.

[13] H. Qi, M. Chen, and D. Wang, "Recurrent and non-recurrent bottleneck analysis based on traffic state rank distribution," *Transportmetrica B*, vol. 7, pp. 1–20, 2017.

[14] H. Liu, R. Yang, Y. Wang, and Q. Zhu, "Measuring performance of road transportation industry in China in terms of integrated environmental efficiency in view of Streaming Data," *Sci. Total Environ.*, vol. 727, pp.1-13, 2020.

[15] M. G. Karlaftis and D. Tsamboulas, "Efficiency measurement in public transport: Are findings specification sensitive?," *Transp. A*, vol. 46, pp. 392–402, 2011.

[16] D. T. Bamum, M. G. Karlafl, and S. Tandon, "Improving the efficiency of metropolitan area transit by joint analysis of its multiple providers," *Transp. Res. E*, vol. 47, no. 6, pp. 60–1176, 2011.

[17] R. L. Tobin and T. L. Friesz, "Sensitivity ananlysis for equilibrium network flow," *Transp. Sci.*, vol. 22, no. 4, pp. 242–249, 1988.

[18] R. Kitamura, S. Fujii, and E. I. Pas, "Time-use data, analysis and modeling: Toward the next generation of transportation planning methodologies," *Transport Policy*, vol. 4, no. 4, pp. 225–235, 1997.

[19] S. Jelena and M. Nada, "The impact of on-street and off-street parking regulations on parking type choice," *Transp. Plann. Technol.*, vol. 46, no. 7, pp. 912–928, 2023.

[20] L. J. Van, "Online learning solutions for freeway travel time prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 9, no. 1, pp. 38–47, 2007.

[21] N. E. E. Faouzi, H. Leung, and A. Kurian, "Data fusion in intelligent transportation systems: Progress and challenges — A survey," *Inf. Fusion*, vol. 12, no. 1, pp. 4–10, 2010.

[22] W. J. Schakel and B. V. Arem, "Improving Traffic Flow Efficiency by In-Car Advice on Lane, Speed, and Headway," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 4, pp. 1597–1606, 2014.

[23] A. Anand, G. Ramadurai, and L. Vanajakshi, "Data fusion-based traffic density estimation and prediction," *J. Intell. Transp. Syst.*, vol. 18, no. 4, pp. 367–378, 2013.

[24] W. Deng, H. Lei, and X. Zhou, "Traffic state estimation and uncertainty quantification based on heterogeneous data sources: A three detector approach," *Transp. Res. B Methodol.*, vol. 57, pp. 132–157, 2013.

[25] A. Duret, and Y. Yuan, "Traffic state estimation based on Eulerian and Lagrangian observations in a mesoscopic modeling framework," *Transp. Res. B Methodol.*, vol. 101, pp. 51–71, 2017.

[26] M. Wright, and R. Horowitz, "Fusing loop and GPS probe measurements to estimate freeway density," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 12, pp. 3577–3590, 2016.

[27] N. Bekiaris-Liberis, C. Roncoli, and M. Papageorgiou, "Highway traffic state estimation per lane in the presence of connected vehicles," *Transp. Res. B Methodol.*, vol. 106, pp. 1–28, 2017.

[28] H. Liu, H. Deng, Y. Li, and Y. Zhao, "School surrounding region traffic commuting analysis based on simulation," *Int. J. Environ. Res. Public Health*, vol. 19, pp. 6566, 2022.

[29] K. Mohammadreza, G. Mahyar, E. O. Eren, and E. Alireza, "A conceptualization of the spatial relationship associated with school-related crashes: A case study in Northwest Florida," *Transp. Plann. Technol.*, vol. 46, no. 6, pp. 795–817, 2023.

[30] M. L. Mouronte-López and A. López, "Commuting to college: An analysis of a suburban campus on the outskirts of Madrid," *J. Adv. Transp.*, vol. 2023, no. 1, pp. 1868826, 2023.

[31] X. Hu, X. Hao, H. Wang, Z. Su, and F. Zhang, "Research on on–street temporary parking effects based on cellular automaton model under the framework of Kerner's three-phase traffic theory," *Physica A*, vol. 545, pp. 123725, 2019.

[32] Y. Cao, Z. Z. Yang, and Z. Y. Zuo, "The effect of curb parking on road capacity and traffic safety," *Eur. Transp. Res. Rev.*, vol. 9, no. 1, 2016.

[33] X. Song, T. Zhan, H. Li, B. Liu, Y. Zhang, and X. Liu, "Optimal parking slots reservation and allocation problem for periodic parking platforms with preference constraints," *J. Adv. Transp.*, vol. 2023, no. 1, pp. 881506, 2023.

[34] H. Yuan and G. Li, "A survey of traffic prediction: from spatio–temporal data to intelligent transportation," *J. Adv. Transp.*, vol. 6, no. 1, pp. 63–85, 2021.

[35] L. Wang, D. Guo, H. Wu, K. Li, and W. Yu, "TC–GCN: Triple cross–attention and graph convolutional network for traffic forecasting," *Inf. Fusion*, vol. 105, pp. 102229, 2024.

[36] L. Bai, Z. Huang, S. Wang, and H. Dai, "Adaptive correlation graph neural ordinary differential equation for traffic flow forecasting," *Eng. Lett.,* vol. 32, no. 9, pp. 1770–1782, 2024.

[37] F. Osorio-Arteaga and E. Giraldo, "Adaptive neural network identification for robust multivariable systems," *IAENG International Journal of Applied Mathematics*, vol. 54, no. 1, pp. 68–76, 2024.

[38] A. Papasani, R. Durgam, and N. Devarakonda, "Adaptive neighborhood adjustment strategy based on MOHHO and NSGA-III algorithms for feature selection," *IAENG International Journal of Applied Mathematics*, vol. 54, no. 5, pp. 917–935, 2024.

[39] S. Du, T. Yang, F. Teng, J. Zhang, T. Li, and Y. Zheng, "Multi-scale feature enhanced spatio-temporal learning for traffic flow forecasting," *Knowl.-Based Syst.,* vol. 294, pp. 111787, 2024.

[40] H. Zhang, J. Gan, J. Zhou, and W. Gao, "Knee point identification based decision making for parameter selection in network designing," *IAENG International Journal of Applied Mathematics*, vol. 54, no. 4, pp. 760–783, 2024.

[41] L. Meng, W. X. Han, and S. Ke, "Traffic conflict identification technology of vehicle intersection based on vehicle video trajectory extraction," *Procedia Comput. Sci.*, vol. 109, pp. 963–968, 2017.

[42] D. T. Ho, E. I. Grøtli, P. B. Sujit, T.A. Johansen, and J. B. Sousa, "Optimization of wireless sensor network and UAV data acquisition," *J. Intell. Robot. Syst.*, vol. 78, pp. 159–179, 2015.

[43] F. Huang, P. Liu, H. Yu, and W. Wang, "Identifying if VISSIM simulation model and SSAM provide reasonable estimates for field measured traffic conflicts at signalized intersections," *Accid. Anal. Prev.*, vol. 50, pp. 1014–1024, 2013.

[44] J. Olstam and A. Tapani, "A review of guidelines for applying traffic simulation to level-of-service analysis," *Procedia Soc. Behav. Sci.*, vol. 16, pp. 771–780, 2011.

[45] T. Sando and R. N. Mussa, "Site characteristics affecting operation of triple left-turn lanes," *Transp. Res. Rec.*, vol. 1852, pp. 55–62, 2003.

[46] H. Liu, H. D, J. Li, S. Yang, K. Dong, and Y. Z, "Calibration method for microscopic traffic simulation considering lane difference," *Simulation*, vol. 10, pp. 1–20, 2024.

[47] G. Li and X. Zhang, "Parking demand forecasting based on improved complete ensemble empirical mode decomposition and GRU model," *Eng. Appl. Artif. Intell.*, vol. 119, pp.105717, 2023.