Hindawi
Journal of Advanced Transportation
Volume 2019, Article ID 3879385, 13 pages
https://doi.org/10.1155/2019/3879385

Research Article

Exploring the Influence of Truck Proportion on Freeway Traffic Safety Using Adaptive Network-Based Fuzzy Inference System

Shiwen Zhang,[1] **Yingying Xing**,[1] **Jian Lu**,[1] **and H. Michael Zhang**[1,2]

[1]*The Key Laboratory of Road and Traffic Engineering, Ministry of Education College of Transportation Engineering, Tongji University, No. 4800, Cao'an Highway, Shanghai 201804, China*
[2]*Civil and Environmental Engineering, University of California, Davis 3145 Engr III, Davis, CA 95616, USA*

Correspondence should be addressed to Yingying Xing; yingying199004@tongji.edu.cn

Received 21 February 2019; Revised 12 June 2019; Accepted 17 August 2019; Published 29 December 2019

Academic Editor: Filomena Mauriello

The truck operation of freeway has an impact on traffic safety. In particular, the gradually increasing in truck proportion will inevitably affect the freeway traffic operation of different traffic volume. In this paper, VISSIM simulation is used to supply the field data and orthogonal experimental is designed for calibrate the simulation data. Then, SSAM modeling is combined to analyze the impact of truck proportion on traffic flow parameters and traffic conflicts. The serious and general conflict prediction model based on the Adaptive Network-based Fuzzy Inference System (ANFIS) is proposed to determine the impact of the truck proportion on freeway traffic safety. The results show that when the truck proportion is around 0.4 under 3200 veh/h and 0.6 under 2600 veh/h, there are more traffic conflicts and the number of serious conflicts is more than the number of general conflicts, which also reflect the relationship between truck proportion and traffic safety. Under 3000 veh/h, travel time and average delay increasing while mean speed and mean speed of small car decreases with truck proportion increases. The mean time headway rises largely with the truck proportion increasing above 3000 veh/h. The speed standard deviation increases initially and then fall with truck proportion increasing. The lane-changing decreases while truck proportion increasing. In addition, ANFIS can accurately determine the impact of truck proportion on traffic conflicts under different traffic volume, and also validate the learning ability of ANFIS.

1. Introduction and Literature Review

With the prosperity of logistics industry in China, the volume of trucks on freeway has been increasing and the freeway traffic safety is more and more obvious [11, 30]. The factors impacting traffic safety on freeway are numerous, including drivers, vehicles, environment, and road alignment. According to statistics from the Traffic Management Bureau of the Public Security Ministry [25], a total of 50,400 road accidents involving trucks were reported, causing 25,000 deaths and 46,800 injuries, which accounted for 30.5%, 48.23%, and 27.81% of the total number of auto liability accidents, respectively. It is much higher than the truck proportion to total number of vehicles. And a total of 5088 traffic crashes occurred in Shanxi province in China, causing 2131 deaths and 5278 injuries. More than 50% of these crashes were involving trucks. The high truck proportion on freeway is an important reason for this phenomenon. Therefore, how to alleviate and eliminate excessive influence of trucks and truck proportion on freeway operational safety is an urgent problem needed to be solved.

A number of studies had examined the impact of trucks on traffic safety based on history crashes data [26] by using various statistical models such as univariate Poisson-lognormal (UVPLN), multivariate Poisson (MVP), and multivariate Poisson-lognormal (MVPLN) regression [5]. It suggested that car crashes involvement rate decreased while the frequencies of truck involved crashes and car–truck involved crashes increased as the truck percentage increasing. Two random parameters ordered probit models [33] were used to explore the influencing factors of single-vehicle and multi-vehicle crashes separately. The multinomial logit (MNL) and negative binomial (NB) models [7] were used to analyze the influences of risk factors on frequency and severity of large truck–involved crashes. A mixed (random parameters) logit model [16] was used to study the injury-severity distributions of accidents on highway segments combined with traffic flow characteristics, which all showed that the presence of higher truck percentages and the sheer number of trucks might have a slowing effect on travel speeds which would tend to decrease the injury-severity of accidents. The literatures mentioned the impact of traffic safety combined with truck proportion via traffic crashes data a little rather than center on the safety via truck proportion.

Then, some literatures analyzed the impact of trucks on traffic flow parameters in order to explore the relationship between truck proportion and traffic safety on freeway. Islam and El-Basyouny [12] quantified the effect of posted speed limit (PSL) reductions in an urban then revealed that there was a high proportion of vans, buses, and trucks on nighttime and weekends under PSL. Li et al. [15] concluded that as the heavy truck proportion grows, the flow and velocity would decrease. The empirical Bayes observational before-and-after study also carried out that the decrease in the standard deviation of speed was 26% while the proportion of light and heavy vehicles exceeding the speed limits more than 20 km/h was reduced respectively by 84% and 77% [18]. The percentage of trucks in the traffic contributed to the average speed.

Apart from that, traffic conflict technic is another useful method for analyzing traffic safety on freeway when traffic crashes data are limited [29, 32]. It showed that the conflict risk degree can reveal the actual safety level of different types of conflicts more comprehensive to some extent. Moreover, the Surrogate Safety Assessment Model (SSAM) [20] also could be used to analyzed the safety performance of conflicts. That the truck volume percentage and the traffic distribution have the highest impact on the conflict frequency is given. Taking the measures of lane-changing, merging and rear-end conflicts as evaluation indicators, the restricted truck lanes and dedicated truck lanes via VISSIM were studied [8]. The proportion of conflicts involving trucks increases as the truck percentage increases and truck lane strategies are most effective when truck percentage exceeded 15%. Time-to-collision (TTC) is also a surrogate safety measure of the safety risk [19] which can predict the probability of safety. Bachmann et al. [3] redefined the algorithm previously implemented for conducting conflict analysis in freeway studies. And, trucks and different vehicles were involved to verify the modified results. It showed that while providing a separate highway for trucks does reduce truck-related conflicts but car lane change conflicts increase. Different literatures tell that the traffic conflict could be surrogate safety measure of traffic safety, so most of conflict techniques could be used for indicating traffic safety under lacking of historical crashes data.

Above all, some literatures analyzed the influential factors such as the weight of truck based on traffic accident data, otherwise, statistic and simulation methods were used to analyze the impact of traffic parameters such as truck performance or conflicts on traffic safety. However, there are a few literatures that discuss the impact of truck proportion on freeway traffic safety and traffic crashes data are used for explore the relationship between truck proportion and traffic crashes mostly. And truck proportion is also the one of most important factors of traffic safety. As a consequence, the purpose of this paper is to explore the impact of truck proportion on traffic flow parameters and traffic conflicts and lay the foundation of studying the correlation between safety and truck proportion. First, the field data extracting from traffic video monitoring system of Shanxi freeway will be used to build simulation model to obtain the simulation data. And then, the traffic conflicts data will be got from SSAM to analyze the relationship between traffic conflicts and truck proportion and the Adaptive Network-based Fuzzy Inference System (ANFIS) will be used for establishing the model of traffic conflicts related to truck proportion in different traffic volume. Therefore, the impact of truck proportion on freeway traffic safety can be reflected.

2. Data Description

First, according to previous literature, traffic flow parameters and traffic conflict parameters were chosen to explore the impact of truck proportion on traffic flow. Then, the traffic parameters such as the mean speed, the average speed difference of truck and small car and the speed standard deviation were selected to explore the impact of truck proportion on traffic conflict and freeway traffic safety.

2.1. Traffic Flow Parameters. Traffic data were collected from surveillance video of three freeway in Shanxi province. There are 7 selected sections and each section is approximately 100 m, linear and straight. Also, the sunny weather was chosen as study condition. The type of vehicles was classified into two categories, small car (include pickup truck) and truck whose body length is more than 6 meters (include passenger car for about 40 passengers) due to specificity of high proportion of trucks on Shanxi freeway. All data that we collected and calculated include truck proportion, traffic volume, average speed, standard deviation, coefficient of variation, number of lane-changing, time headway, and time to collision. And the basic information is as follows (Table 1).

A few common traffic parameters are listed in (Table 1), which are related to truck proportion more or less. The different truck proportion was collected per hour, which is the value that the number of trucks divided by the total number of vehicles. The range of truck proportion covers from 0.176 (17.6%) to 0.905 (90.5%) per hour in different volume on different section, which are mostly in low traffic volume and most focus on the range from 0.2 to 0.4. The total sample collected from the surveillance video is 81 hours. Besides, the standard deviation of the traffic volume per hour is relatively large, which means a greater degree of dispersion between individuals within the group. According to the basic theory, when analyzing the correlation among these parameters, traffic volume was divided into five kinds: low, slightly low, medium, slightly high, high.

Standard deviation was defined as the arithmetic square root of variance, reflecting the degree of dispersion among individuals in the group. Truck proportion would affect the traffic volume, the mean speed, and the speed standard deviation even further, which represents the influence of truck proportion on the stability of speed standard deviation under a certain flow rate. In most cases, the speed standard deviation was estimated by taking a random sample from overall samples and calculated by the sample. It lays the foundation for the further study of the relationship between traffic safety and the truck proportion.

Table 1: The base information of collected data.

Statistical data	Range	Mean value	Standard deviation
Truck proportion (/h)	17.6–90.5%	36.4%	0.198
The volume of truck (h)	82–515	180.78	61.047
The volume of vehicle (veh/h/ln)	240–816	553.69	145.512
The mean speed (km/h)	47.136–118.722	88.231	12.803
The average speed difference of truck and small car (km/h)	7.21–39.519	23.749	7.664
Speed standard deviation (km/h)	10.129–26.065	18.609	3.534
Coefficient of variation	0.12–0.34	0.215	0.0506

2.2. Traffic Conflict Parameters. Under observable conditions, traffic conflicts are the approach of each other spatially that two or more road users at the same time. If one of them takes abnormal traffic behavior, such as changing lane, changing speed, stopping suddenly and so on, a collision will occur unless the other takes the corresponding safety behaviors. This phenomenon is traffic conflicts. Because the object of this paper is the truck on freeway, so there mainly are two types of conflicts on freeway. These are collisions with rear-end on the same lane and lane changing conflicts between vehicles.

The sizes and operating performance of truck have a significant impact on road traffic safety. For larger trucks, the permissible speed is lower than those smaller cars of smaller mass. Larger trucks will affect the visibility of the rear vehicles as well as the lateral distance of the vehicles on adjacent lane, which makes the driver feel oppressed. Most of parameters involved in conflicts are the time to collision (TTC), post encroachment time (PET), deceleration to avoid crash (DRAC), and so on [31]. The truck proportion also has a great impact on traffic conflicts. Jeong and Oh [14] used the TTC, the warning index (WI) thresholds and market penetration rate (MPR) to explore active vehicle safety in a certain truck proportion. TTC is the most frequently used surrogate measure and usually utilized as a benchmark on SSAM imported into [9]. TTC will be utilized in this paper to explore the relationship between truck proportion and traffic conflict parameters. Because the angle of vehicle is not convenient to observe and measure, so the TTC is calculated by velocity and vehicle length (25 frames represent for 1 second). So, the Equation (1) is as follows [28]:

$$TTC_i = \frac{x_{i-1}(t) - x_i(t) - l_{i-1}}{V_i(t) - V_{i-1}(t)}, \ \forall V_i(t) > V_{i-1}(t). \qquad (1)$$

TTC_i—The collision time (s) of vehicle i collide to the vehicle $(i-1)$; $x_{i-1}(t)$—the position of vehicle $i-1$ on the road (m) at time t; $x_i(t)$—the position of vehicle i on the road (m) at time t; $V_i(t)$—the spot speed of vehicle i at time t; $V_{i-1}(t)$—the spot speed of vehicle $i-1$ at time t; l_{i-1}—the length of front vehicle $i-1$.

And, $x_{i-1}(t) - x_i(t) = V_{i-1}(h_i - h_{i-1})$, h_i, h_{i-1} is the headway distance of vehicle i and vehicle $i-1$ respectively. So, the Equation (2) is as follows.

$$TTC_i = \frac{V_{i-1}(h_i - h_{i-1}) - l_{i-1}}{V_i(t) - V_{i-1}(t)}, \ \forall V_i(t) > V_{i-1}(t). \qquad (2)$$

The indicator can be used as both the calculation of the rear-end collision (collision angle $\theta \in [0, 30]$) and the lane-changing collision (collision angle $\theta \in [30, 85]$).

3. Method

Due to the actual conditions, the traffic data obtained through field surveys can hardly cover the traffic flow status under different traffic conditions of the freeway. The main purpose of this paper is to study the influence of truck proportion on freeway traffic safety, so the more and more traffic parameters under different traffic conditions are needed to be obtained. Therefore, modeling of data validation and supplement via VISSIM simulation and SSAM are necessary [1, 21].

3.1. Simulation Modeling. The study segment was simulated using VISSIM, a stochastic, behavior based microscopic simulation platform. To capture the effect of truck's proportion on traffic flow parameters, a basic scenario was constructed according to the characteristics of freeway in Shanxi. In the basic scenario, the divided freeway has two 3.5 m lanes in each direction and the length of the road section is 1000 m. To capture the traffic characteristics, data collection points were set up every 100 m in each direction. As vehicle in China is a right-hand drive, so the driver could only overtake from the left side of the vehicle in the model. Besides, the vehicle parameters derived from the video were set as (Table 2).

Ten different traffic flow rates were considered, i.e., 2200 veh/h, 2400 veh/h, 2600 veh/h, 2800 veh/h, 3000 veh/h, 3200 veh/h, 3400 veh/h, 3600 veh/h, 3800 veh/h, 4000 veh/h. The truck percentage was divided by 10% into 9 grades, 10–90% respectively. Different traffic flow states were combined with different traffic flow and different mixing rate of vehicle in field survey. Vehicles were generated randomly according to Poisson distribution in VISSIM. The traffic flow would have little difference in simulation model and field survey, since the relative error is within 0.1%, the simulation result will not be affected. Vehicle speed design principles are based on the actual situation of the normal distribution, and in accordance with the distribution of the interval via 3σ principle. Wiedemann 99 model was selected as following car behavior parameters which is suitable for intercity road (freeway) traffic.

Many parameters were used in car-following model and lane-changing model of VISSIM. In order to simulate the traffic behavior accurately, it needs to adjust the parameters for the study section [17]. Besides, these parameters have a direct effect on the interactions of vehicles and result in significantly different simulation results even with small adjustments to the parameter values, including the parameters of car-following

Table 2: Parameter setting of vehicles.

Type of vehicle	Length of vehicle (m)	Acceleration (m/s^2)		Deceleration (m/s^2)		Power (kW)	Mass (t)	Desired speed (km/h)		
		Max.	Expected value	Max.	Expected value			15%	50%	85%
Small car	4.76	2	1.8	3	2.8	55–118	1.5–2	84	100	115
Truck	10.215	1	0.8	1.5	1.3	100–177	10–40	67	78	93

Table 3: Factor level table of the parameters to be determined.

Level	CC1 headway time (s)	CC2 "following" variation (m)	CC3 threshold for entering "following"	Min. headway (front/rear) (m)
1	0.5	3	−9	0.3
2	0.9	4	−8	0.5
3	1.3	5	−7	0.7
4	1.7	6	−6	0.9

and lane-changing behavior [2, 23]. Therefore, the average speed of vehicles on freeway was selected as the index to evaluate whether the simulation model meets the field traffic conditions. The calibrated parameters are shown in Table 3.

Four parameters, each of which took 4 different levels of value, were calibrated, and 256 experiments were required if each level of each factor was tested on a case-by-case basis. In order to reduce the number of tests effectively and obtain the required results at the same time, the parameters of the model were calibrated by orthogonal experiment in this simulation. In general, orthogonal table recording experimental programs and results should meet two basic requirements. First, different numbers in each column appear the same number of times. This characteristic indicates that each level of each factor is exactly the same as the probability of participating in the test at each level of the other factors, thereby it ensures that interference from other levels is maximally excluded at each level and it can effectively compare the test results and find the optimal test conditions. Second, in any two pairs of their horizontal composition of pairs, each number of pairs appear equal. This feature ensures that the test points are evenly dispersed in the complete combination of factors and levels, and therefore they have a strong representation. In order to reduce the number of experiments, smaller orthogonal tables are generally chosen. If the experiment requires a higher accuracy, a larger orthogonal table could be chosen according to the experimental conditions. Therefore, in this paper, the design of experimental program was designed as follows.

From Table 4, the influence degree of each factor is ranked as $A > C > B > D$. However, the experimental scheme with the smallest error value may not be the optimal scheme. The optimal scheme should be a combination of the optimal value of each factor. Experimental indicators determine the optimal level of each factor. If the purpose of the experiment is to obtain the maximum experimental index value, the corresponding level will be the largest value among the factors, and vice versa. The model being consistent with the actual traffic conditions is the purpose of paper, that is, to make the absolute difference between the experimental results and the actual data as small as possible. Therefore, the appropriate level for each factor is $A_4B_3C_1D_3$, the simulation error of which is less.

The Surrogate Safety Assessment Model (SSAM) is a traffic safety evaluation software developed by the Federal Highway Administration (FHWA) based on surrogate model for traffic safety evaluation [6, 20, 32]. It can be used to analyze traffic conflicts in vehicle output track files from microscopic traffic simulations. In the paper, the number of TTC was used for exploring the effect of truck proportion on traffic safety. Analyzing the cumulative frequency distribution curve of the TTC sample data, the 85% quantile is taken as the traffic conflict threshold, which means the general conflict (TTC = 6 s) and through literature review [4, 14], TTC less than 2 s were taken as serious conflict, and TTC greater than 2 s and less than 6 s were regarded as general conflict. Apart from that, the conflict angle of rear end and lane change are defined as [0, 30] and [30, 85], respectively. Then, the track files from VISSIM simulations will be imported to SSAM to determine the severity of traffic conflict and the relationship between number of TTC and truck proportion. Thus, the number and the value of TTC will be calculated for analyzing.

In order to ensure the consistency between the simulation data and field data, the root mean square error (*RMSE*) was used to verify the error. Taking the volume of 2200 veh/h as an example, the number of lane-changing was selected as the evaluation parameters to determine the root mean square error (Equation (3)). The observed data, simulation results and root mean square error are shown in Table 5.

$$RMSE = \sqrt{\frac{\sum_{i-1}^{n} \left(X_{obs,i} - X_{model,i}\right)^2}{n}}, \tag{3}$$

where, $X_{obs,i}$ is the field data and $X_{model,i}$ is the simulation result. The deviation between the observed value and the simulated value is evaluated by *RMSE*. The *RMSE* can reflect the well precision of the measurement. The smaller the *RMSE* is, the smaller the difference between the observed value and the simulated value is, and the closer the difference between observed data and the field value is.

Table 4: Orthogonal experimental design.

Test number	Headway time A	"Following" Variation B	Threshold for entering "following" C	Min. headway (front/rear) D	Test scheme	Speed difference between small cars and trucks
1	1(0.5)	1(3)	1(−9)	1(0.3)	$A_1B_1C_1D_1$	1.853
2	3(1.3)	3(5)	1	3(0.7)	$A_3B_3C_1D_3$	1.459
3	4(1.7)	4(6)	1	4(0.9)	$A_4B_4C_1D_4$	1.06
4	2(0.9)	2(4)	1	2(0.5)	$A_2B_2C_1D_2$	1.594
5	2	4	3(−7)	1	$A_2B_4C_3D_1$	1.631
6	4	3	2(−8)	1	$A_4B_3C_2D_1$	1.059
7	3	2	4(−6)	1	$A_3B_2C_4D_1$	1.474
8	1	4	4	3	$A_1B_4C_4D_3$	1.859
9	4	1	4	2	$A_4B_1C_4D_2$	1.248
10	1	3	3	2	$A_1B_3C_3D_2$	1.854
11	2	3	4	4	$A_2B_3C_4D_4$	1.631
12	2	1	2	3	$A_2B_1C_2D_3$	1.614
13	3	1	3	4	$A_3B_1C_3D_4$	1.471
14	3	4	2	2	$A_3B_4C_2D_2$	1.458
15	4	2	3	3	$A_4B_2C_3D_3$	1.084
16	1	2	2	4	$A_1B_2C_2D_4$	1.856
$\overline{K}_1$	1.8555	1.5465	1.4915	1.50425		
$\overline{K}_2$	1.6175	1.502	1.49675	1.5385		
$\overline{K}_3$	1.4655	1.50075	1.51	1.504	$\sum = 24.205$	
$\overline{K}_4$	1.11275	1.502	1.553	1.5045		
Excellent level	A_4	B_3	C_1	D_3		
R_j	0.74275	0.04575	0.0615	0.0345		
Order			A C B D			

Table 5: The estimation of *RMSE*.

The proportion of truck (2200 veh/h)	The number of lane-changing (simulation)	The number of lane-changing (field)	$\lvert(X_{\text{obs}} - X_{\text{model}})\rvert/(X_{\text{obs}} * 100\%)$	*RMSE*
8.86%	188	167	12.57%	
17.71%	183	174	5.17%	
29.43%	204	200	2.00%	
38.57%	207	192	7.81%	
48.29%	166	175	5.14%	10.1882
58.86%	146	143	2.10%	
71.05%	125	119	5.04%	
79.77%	116	109	6.42%	
91.45%	111	101	9.90%	

Therefore, according to Equation (3), the *RMSE* is 10.1882 under the volume of 2200 veh/h, 5.2536 under the volume of 2400 veh/h, 8.4794 under the volume of 2600 veh/h, 15.6525 under the volume of 2800 veh/h and the average error between the observed value and the simulated value is under 20%, which is acceptable. The observed data is incomplete under the high volume, but the error between the measured number of lane-changing and the simulated simulation value is still within 20%, and the error is in the acceptable range. Therefore, the VISSIM simulation model data and observed data are consistency.

3.2. Adaptive Network-Based Fuzzy Inference System. Because there is no direct relationship between the truck proportion and traffic conflicts, the application of traditional mathematics methods is time-consuming and laborious to establish the relationship between the two, and it is not well adapted to modeling under various assumptions. The ANFIS (Adaptive Network-based Fuzzy Inference System) proposed by Jang [13] is a new type of fuzzy inference system that combines fuzzy logic and neuron networks. ANFIS has the advantages of expression of fuzzy logic and self-learning ability of neural network easily, which has gradually become an important research direction of computational intelligence in recent years. And it has the characteristics of simple calculation and mathematical analysis, which provides an effective tool for the modeling and control of complex systems. This also

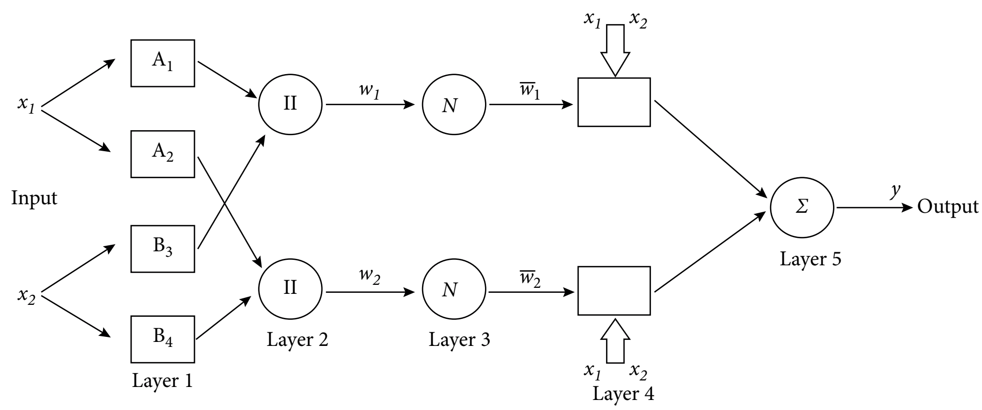

FIGURE 1: The diagram of adaptive network-based fuzzy inference system.

provides a good reference for the use of transportation field. The back-propagation algorithm and the least-squares method are used to adjust the precondition parameters and conclusion parameters, and the If-Then rules can be automatically generated.

ANFIS uses fuzzy neural network to realize the three basic processes of fuzzy control, fuzzy inference and anti-fuzzification. It uses neural network learning mechanism to automatically extract rules from the input and output sample data to form an adaptive neuro-fuzzy controller. Through offline training and online learning algorithms, the fuzzy inference control rules are self-adjusted, and the system itself develops in the direction of self-adaptation, self-organization, and self-learning. Rahimi [22] studied the quantitative assessment of the effects of intelligent transportation systems and technologies on road fatalities. Modification of user behavior, exposure, modal choice, and accident consequences were used for explore the effect of ITS technologies on road safety. Hosseinpour et al. [10] presented the application of ANFIS technique to estimate road accident frequencies as a function of road geometric and environmental characteristics. Compared with other regression models, it verified that the proposed model had higher prediction performance than the other traditional models. Besides, ANFIS combined with Global Satellite Navigation System also can be used for real-time car-following status identification [24] and lane changing maneuver prediction [27] to make sure the traffic safety. Therefore, ANFIS will be used to model the relationship between truck proportion and traffic conflict. The structure diagram is shown in Figure 1. Each node in the same layer has a similar function ($O_{1,i}$ denotes the output of the i-th node in the first layer).

Layer 1. Fuzzing the input variables, then outputting the degree of membership of the corresponding fuzzy set (the fuzzification layer). The input variables x_1 and x_2 are the input of node i, represent for traffic volume and truck proportion, respectively. A_i or B_{i-2} is the linguistic variable associated with the node function value, That is, $O_{1,i}$ is the membership function of fuzzy set A and B. The membership function is determined by the actual input variables (Equation (3)). There, truck proportion and traffic volume are set as x_1, x_2. A_i or B_{i-2} denotes that fuzzy classification of different variables.

$$\begin{aligned} O_{1,i} &= \mu_{Ai}(x_1), \quad i = 1,2 \\ O_{1,i} &= \mu_{B(i-2)}(x_2), \quad i = 3,4. \end{aligned} \tag{4}$$

Layer 2. The node is represented by II. Each one is the fixed node. The input signal is multiplied and the output of the node represents a rule's excitation intensity (Equation (4)), which denotes that each input component belongs to a membership function of a fuzzy set of linguistic variable values.

$$O_{2,i} = w_i = \mu_{Ai}(x_1)\mu_{Bi}(x_2), \quad i = 1,2. \tag{5}$$

Layer 3. The node is denoted by N, which is also a fixed node. The incentive intensity of each rule is normalized. The ratio that the sum of the w_i of the i-th rule and the sum of all the rules w is calculated at the node I (Equation (5)).

$$O_{3,i} = \overline{w}_i = \frac{w_i}{w_1 + w_2}, \quad i = 1,2. \tag{6}$$

Layer 4. Each node is an adaptive node with a node function that calculates the output of each rule (Equation (6)), which denotes that fuzzy rule of these two variables is normalized.

$$O_{4,i} = \overline{w}_i f_i = \overline{w}_i(p_i x_1 + q_i x_2 + r_i), \quad i = 1,2. \tag{7}$$

Layer 5. A single node is a fixed node. The total output of all input signals is as follows. And the total output is traffic conflict in this paper (Equation (7)).

$$O_{5,i} = \sum_i \overline{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i}. \tag{8}$$

Through actual investigation and experimental simulation results, it is concluded that the factors affecting traffic collisions are mainly traffic volume and vehicle type. Therefore, different traffic volume and different truck proportion are used as input variables, and the number of traffic conflicts is for output variables. The fuzzy inference system considered in this paper has two inputs the traffic volume and truck proportion and a single output traffic conflict.

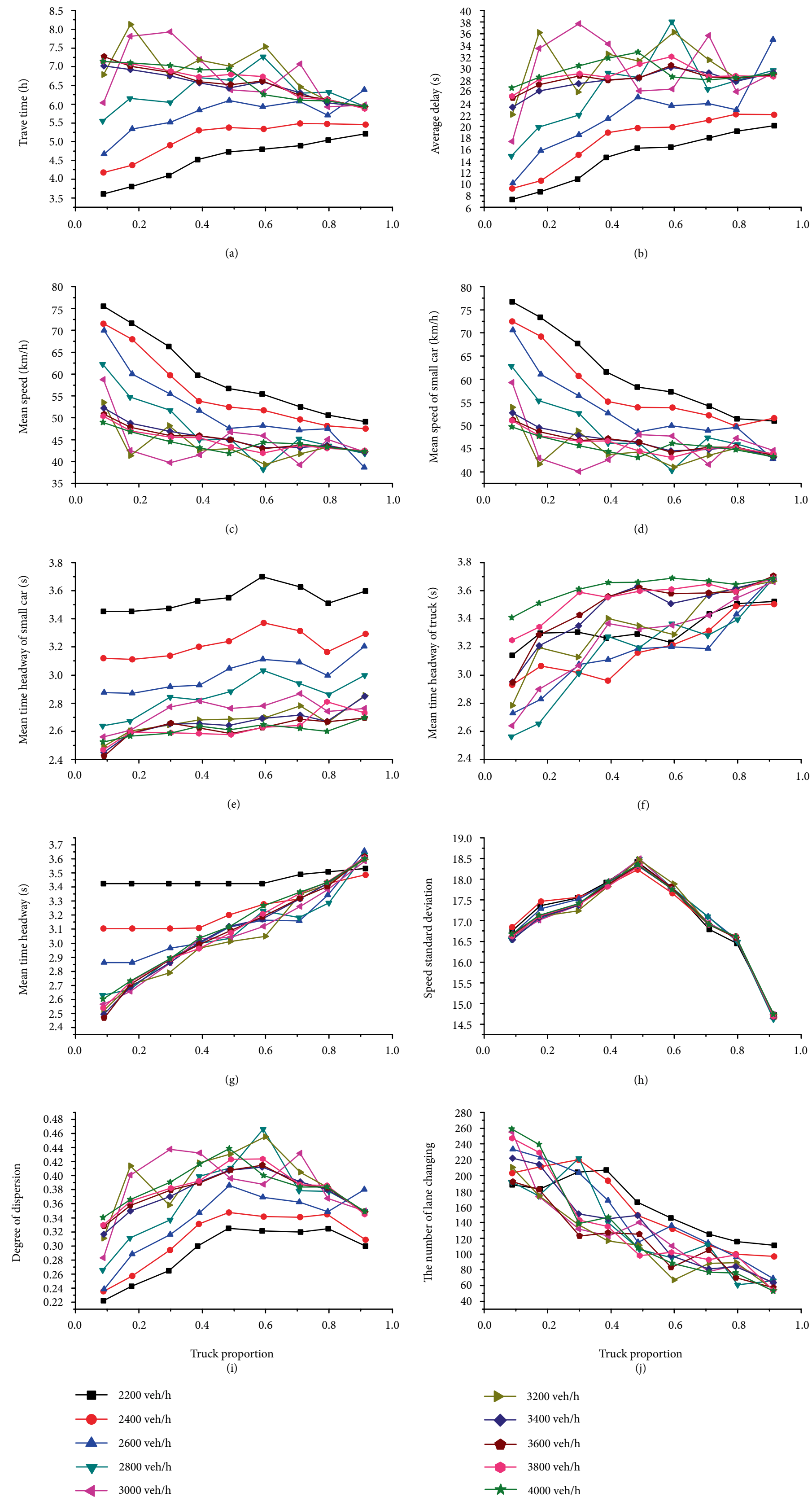

Figure 2: The relationship among truck proportion with other traffic parameters.

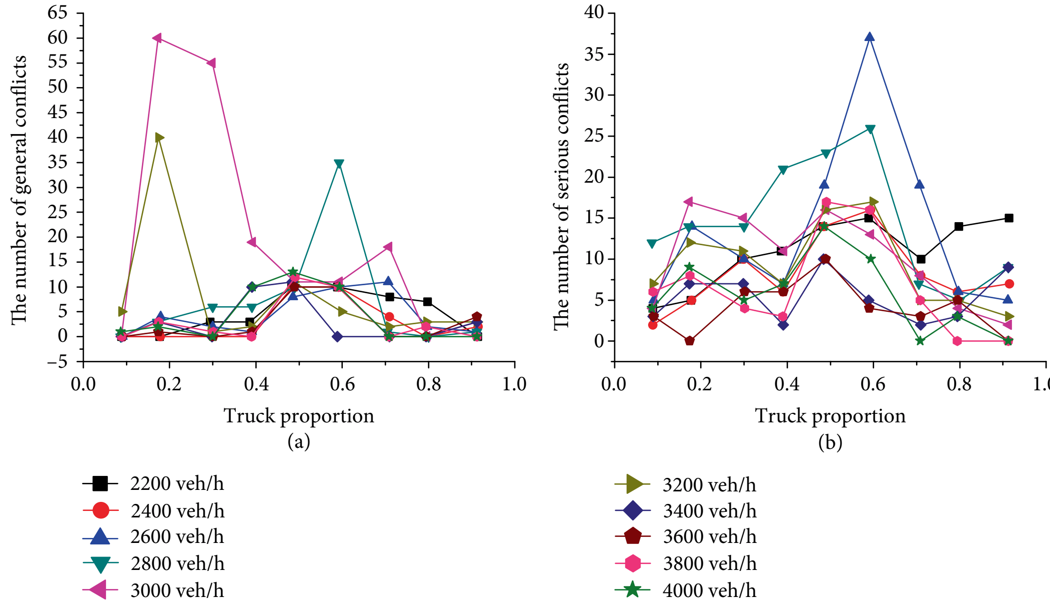

Figure 3: The relationship between truck proportion and TTC.

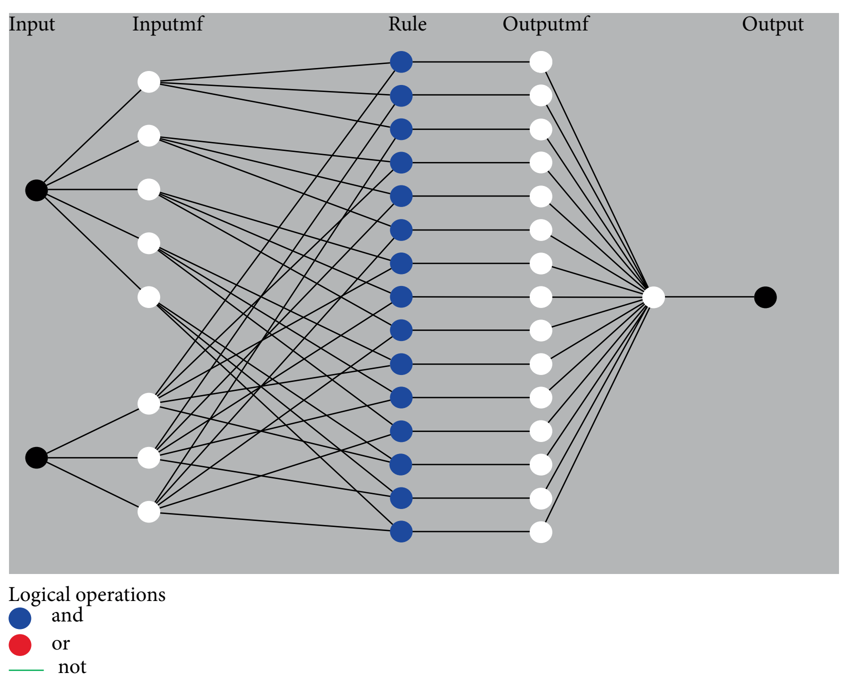

Figure 4: ANFIS model structure.

4. Result and Discussion

4.1. Simulation Results. In the simulation experiment, traffic flow is divided into several kinds of flow, which are the volume of 2200 veh/h, 2400 veh/h, 2600 veh/h, 2800 veh/h, 3000 veh/h, 3200 veh/h, 3400 veh/h, 3600 veh/h, 3800 veh/h, 4000 veh/h, and the relationship of truck proportion with other parameters are shown in Figure 2.

As it can be seen, with the increasing of truck proportion in different traffic volume, different parameters show different trends. In Figure 2(a), travel time rises below the volume of 2800 veh/h due to the truck proportion increasing and slightly decreases above 3400 veh/h while average delay first increases and then stabilize (Figure 2(b)), because with the increase of the traffic volume and truck proportion, the driver's judgment on the road safety depends on driving speed and the road alignment. Based on the safety operation, the driver will choose the relatively safe speed. Therefore, when the truck proportion reaches around 0.6, the delay time tends to be stable. Undoubtedly, the mean speed decreases with the increase in the traffic volume and the truck proportion in whole (Figure 2(c)). By contrast, the mean speed of small car slightly reduces while the truck proportion is less than 0.5 and the volume above 3400 veh/h. And the truck proportion is greater than

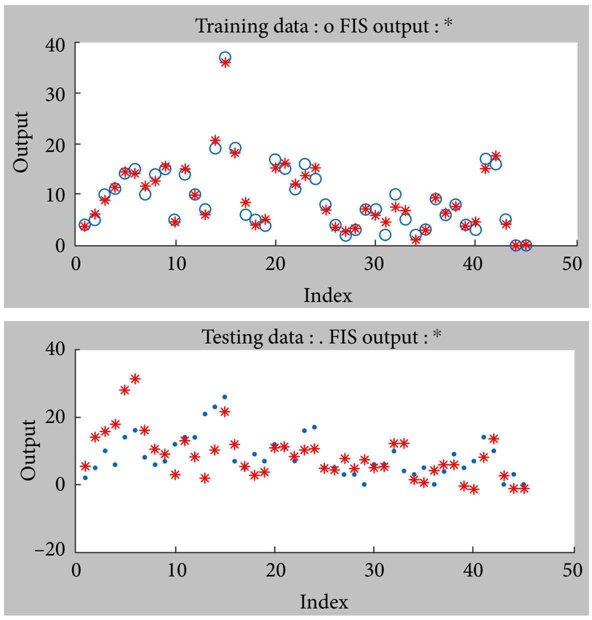

Figure 5: Training data and testing data of ANFIS model.

0.5, the mean speed of small car is significantly descending (Figure 2(d)). The Figure 2(e) shows that the mean time headway of small car is relatively stable and then fluctuates around 0.6. Unlike the mean time headway of small cars, the trend of trucks is exactly the opposite for trucks (Figure 2(f)). Although the numerical value is fluctuated, the trends increasing overall. The reason why it is opposite is that the vehicle driving is relatively free and driver behavior is not constrained due to truck proportion is small, so they maintain normal driving behavior that the trend stays constant. But when the volume of trucks goes up and affects the driving operation of small car drivers, the mean time headway of small car begins to fluctuate. And in the high volume, the fluctuation is relatively steady. The principle of the mean time headway of trucks is the same. When the volume of small cars and the proportion of trucks are small, the interaction and interference between small cars and trucks is much less, so the mean speed of small cars is not marked reduced. When the volume increases, the interaction between the small cars and the trucks rises due to the difference in the vehicle performance. To ensure safety driving, the speed of small cars gradually approaches the speed of trucks to reach equilibrium in the end. When the state is reached, the drivers of small cars will increase the speed overtaking the truck by the judgment on safety. Therefore, when the proportion of trucks is the largest, there will be a slight increase.

The mean time headway increases following the increase of truck proportion from truck proportion at 0.5 (Figure 2(g)) in general. Below the volume of 2600 veh/h and less than truck proportion about 0.6, the mean time headway keeps steady. Furthermore, the speed standard deviation (Figure 2(h)) and the degree of dispersion (Figure 2(i)) are gradually increasing with the increase in the truck proportion in different traffic volume and then falling due to the difference in vehicle performance between small cars and trucks. This is consistent with the actual situation. The number of lane-changing (Figure 2(j)) is declining with the increase of truck proportion. Above all, at the volume of 3000 veh/h and 3200 veh/h, the all parameters except mean time headway of small cars, mean time headway and speed standard deviation are fluctuated. Besides, the parameters are slightly changed when the volume above 3400 veh/h while changed a lot below 2800 veh/h. The reason why it is fluctuated is that the volume at 3000 veh/h is the stable flow at the designed speed 80 km/h on the freeway.

The results of traffic conflict are shown in Figure 3. Generally, serious conflicts (Figure 3(b)) are more than general conflicts (Figure 3(a)), which indicates that the truck proportion has an impact on traffic safety to some degree. When TTC is less than 6 s but more than 2 s, the number of TTC increases with the truck proportion at 0.4 under different traffic flow while the number of serious conflicts is large at 0.5 of truck proportion. Similarly, with the increasing truck proportion, the number of general conflict and serious conflict are rising initially and then fall around 0.6 of the truck proportion in general. As the traffic volume and truck proportion increase, the interaction between different vehicles increases, resulting in limited freedom of travel. Moreover, when the traffic volume and truck proportion continue to increase, the driving speeds of the vehicles tend to be the same, which makes the speed dispersion small, and the vehicles are in a car-following state, leading to the reduction of conflict. Due to the traffic bottleneck, the volume around 3000 veh/h is tending to the stabilizing flow state and the parameters of traffic flow and other parameters are little fluctuated in the period.

Above all, truck proportion has a great impact on traffic flow parameters and traffic conflict. With increasing truck proportion in different traffic volume, travel time and average delay are increasing; the mean speed and time headway are generally decreasing; speed standard deviation, degree of desperation, the number of lane-changing and the number of serious and general conflicts are increasing initially and then decreasing. These all do a great impact on traffic safety.

4.2. ANFIS Modeling Results. The first step in ANFIS modeling is fuzzification. The purpose is to establish an initial fuzzy inference system, determine the number of membership functions and types of each input variable and select the structural rules of the ANFIS model. Because the data distribution is much obvious, the grid partition method is used to divide the loaded data. Besides, the fuzzy system is established according to the set parameters according to the fuzzy C-means clustering method. The purpose of clustering is to determine the minimum number of fuzzy rules required for FIS construction and their associated member parameters. The dataset will be input into multiple cluster groups by cluster partitioning so that the similarity among members of the group is higher and the similarity of members within the group is lower. The input membership function is a Gaussian function, and the number of fuzzy subsets is 3 for truck proportion and 5 for traffic volume to cover the input variables, and the membership function of the output variables is selected as a linear function, so the initial FIS is generated. In the learning process, both training and inspection datasets are used to avoid overfitting. The margin of error is set closely to 0 and the number of iterations is 50. The selection of the best

Table 6: The estimated input membership parameters and the estimated consequent parameters of sugeno linear function.

Type of conflict	Input	Type of membership function	Membership	Estimated parameters [σ c]
General conflict	Traffic volume (x)	Guassmf	Low	(169.9 2200)
			Slightly low	(169.9 2600)
			Medium	(169.9 3000)
			Slightly high	(169.9 3400)
			high	(169.9 3800)
	Truck proportion (y)		low	(0.2141 0.07092)
			middle	(0.1942 0.5095)
			high	(0.2349 0.903)
Serious conflict	Traffic volume (x)		Low	(169.9 2200)
			Slightly low	(169.9 2600)
			Medium	(169.9 3000)
			Slightly high	(169.9 3400)
			high	(169.9 3800)
	Truck proportion (y)		low	(0.14 0.1887)
			middle	(0.1669 0.3561)
			high	(0.1379 0.7708)

$u = a * \text{Traffic volume } (x) + b * \text{Truck proportion } (y) + c$

Type of conflict	u (Output)	a	b	c (Constant)
General conflict	u1	0.0003174	−11.85	1.385e-07
	u2	0.0009358	12.29	5.308e-07
	u3	0.01711	−43.66	7.958e-06
	u4	−0.0007657	87.62	−1.918e-7
	u5	−0.03563	215	−1.309e-05
	u6	−0.0527	131.6	−1.96e-05
	u7	−0.03184	1604	−1.06e-05
	u8	−0.1856	1076	−6.207e-05
	u9	−0.1834	553.2	−6.134e-05
	u10	0.004113	−237.3	1.132e-06
	u11	0.03614	−201.1	1.024e-05
	u12	0.0132	−33.25	3.551e-06
	u13	−0.0004658	107.8	−1.101e-07
	u14	−0.02468	244.8	−6.408e-06
	u15	−0.05842	226.9	−1.532e-05
Serious conflict	u1	0.001342	17.4	4.692e-07
	u2	−0.001864	33.97	−1.111e-07
	u3	−0.005815	31.57	−2.654e-06
	u4	0.03195	576	1.226e-05
	u5	−0.1677	846	−6.429e-05
	u6	0.0007217	−0.7736	2.536e-07
	u7	0.007582	332	2.606e-06
	u8	−0.0504	300.4	−1.724e-05
	u9	−0.001501	5.846	−4.463e-07
	u10	0.007593	173.4	2.226e-06
	u11	−0.03741	254.7	−1.1e-05
	u12	−0.01781	76.92	−5.225e-06
	u13	0.01283	173.4	3.387e-06
	u14	−0.04989	397	−1.32e-05
	u15	−0.005832	23.13	−1.556e-06

ANFIS model is based on achieving a minimum *RMSE* for both the *RMSE* controlled and nonincremented training and testing datasets. The number of membership functions in this paper are truck proportion and traffic volume while the truck proportion type is divided into three sets and traffic volume is divided into five sets. For the rules Rj: if x_1 (traffic volume) is Ai $(i = 1, 2, 3, 4, 5)$ and x_2 (truck proportion) is Bi $(i = 1, 2, 3)$, then $y_j = p_{j0} + p_{j1}x_1 + p_{j2}x_2$. The ANFIS network structure of

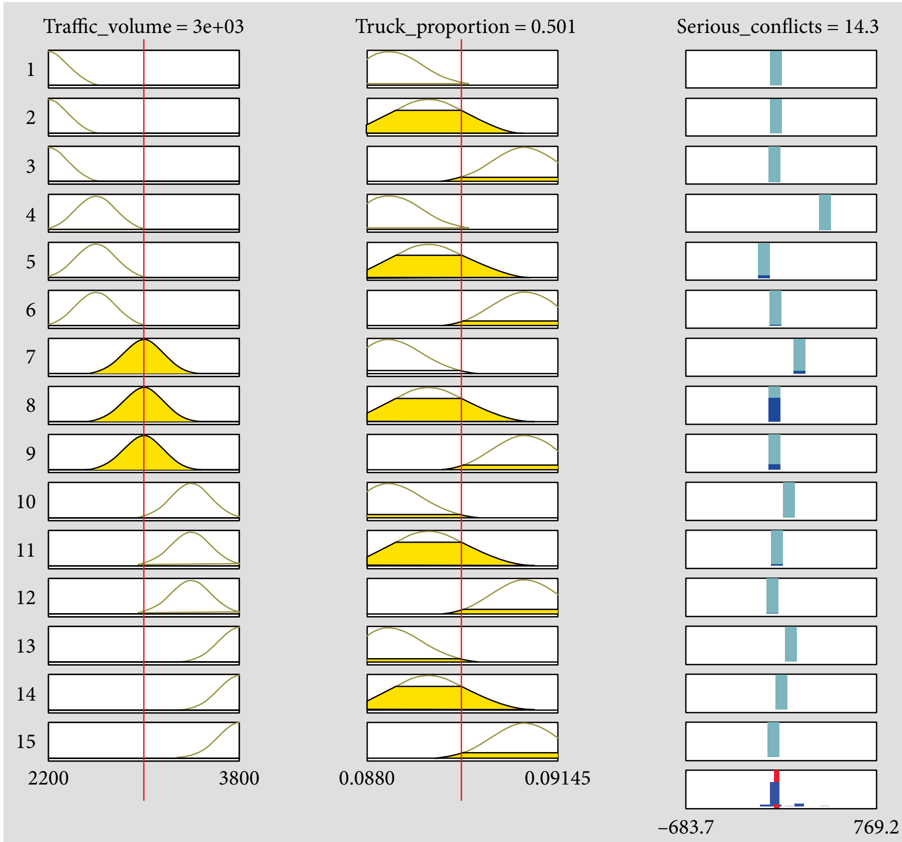

Figure 6: The schematic diagram of fuzzy rules on general conflict.

fuzzy inference system is shown in Figure 4. In Figure 4, the input variables are the traffic volume and truck proportion.

The fuzzy membership function subset with sharper shape function curve has higher resolution and higher control sensitivity. On the contrary, the membership function curve has a flatter shape, the control characteristics are more gradual, and the stability performance is better. Therefore, when selecting the membership function of the fuzzy variable, the low-resolution fuzzy set is used in the region with large error, and the high-resolution fuzzy set is selected in the region with small error. When the error is close to zero, the higher control selection is used. The blurring of the resolution makes it possible to achieve a control effect with high control precision and good stability. So the Guass membership function is used for fuzzing these two input variables. Traffic volume and truck proportion are fuzzed as five subsets and three subsets, respectively. Then 15 rules have been constructed to put out 15 membership functions and the output is the number of general conflicts. Thus, the *RMSE* for the training datasets converged to a minimum of 1.2457 ultimately, which led to the completion of the learning process. Then the testing data is loaded, which means the remaining data in the dataset is used as testing data. The *RMSE* for the testing data converged to a minimum of 6.6159 ultimately. The training data, testing data are as following Figure 5. Therefore, according to the parameters of each function, the output function of the system can be summarized.

The relationship among traffic volume, truck proportion and serious conflict is established as the same method as above. And the Gaussian membership function expression covering the fuzzy subset $traffic\ volume \in [2200, 4000]$ can be written and the estimated parameters is shown in Table 6 according to the Gaussian membership function formula $F(x, \sigma, c) = exp\left(-(x-c)^2/2\sigma^2\right)$. Similarly, the Gaussian membership function covering the fuzzy subset $truck\ proportion \in [0.1379, 0.903]$ can be written as well as traffic volume.

Taking serious conflict as an example, it can be seen intuitively that the number of serious conflicts occurring under different traffic volume and different truck proportion in Figure 6. The output is the last line shown in Figure 6. Besides, the output of training data is approximately close to the output of simulation data.

In general, when the traffic volume is around 3000 veh/h, the number of general conflicts is large in general. When the truck proportion is below 0.5, the number of general conflicts is initially increasing then falling. While the truck proportion above 0.5, the number of serious conflicts is fluctuated but overall it is lower when the traffic volume is high. Meanwhile, when the traffic volume is around 2600 veh/h, the number of serious conflicts is large in general. When the truck proportion is below 0.5, the number of general conflicts is initially increasing then falling. While the truck proportion above 0.5, the number is lower than the truck proportion is above 0.5. When the truck proportion is constant, the number of serious conflicts decreases as the traffic volume increasing in general. When the traffic volume is low, the interaction between the vehicles is small because of the free driving, and the impact of the truck proportion, whether large or small, on traffic conflicts is not obvious. However, when the traffic volume is high or stable, the interaction between vehicles increases, especially the interaction between trucks and small cars. Trucks could narrow down the vision of other vehicles, especially small ones. The low speed of trucks and the decrease of sight distance lead to the overtaking behavior of small cars. This may have an impact on traffic safety and traffic conflicts, leading to an increase in serious conflicts. And when the truck proportion is more than 50%, the traffic flow is gradually stable, and the mean time headway has also reached a stable state.

5. Conclusion

This paper mainly studies the impact of truck proportion on traffic safety by analyzing its effect on traffic flow parameters and traffic conflicts. The traffic data of the straight section of freeway in Shanxi province was collected, and simulation data was corrected by orthogonal experiment and supplemented by VISSIM to obtain the trajectory of the vehicle. Then, the number of traffic conflicts was calculated through SSAM and the impact of truck proportion on traffic conflicts was established through ANFIS. Finally, the impact of the truck proportion on traffic flow parameters and traffic conflicts under different traffic flows was analyzed. The results mainly concluded that with the increase of traffic volume and truck proportion, travel time and average delay are increasing whilst mean speed and mean time headway are decreasing. What's more, with the increase of truck proportion, speed standard deviation, degree of dispersion, the number of lane changing and the number of traffic conflict are initially increasing then falling down and the higher the traffic volume is, the larger these parameters are. These all explain the impact of the truck proportion on freeway traffic safety. The traffic state will be in danger when truck proportion is around 0.6 at traffic volume 2600 veh/h while the traffic state will be in danger when truck proportion is around 0.4 at traffic volume 3200 veh/h, where the number of traffic conflicts is large.

Actually, little works have been done in traffic data analysis using ANFIS. But in this paper, ANFIS establishes the relationship among parameters quite well, and also indicates the influence of truck proportion on traffic safety by the severity of surrogate indicators indirectly. ANFIS is the control system that combines the functional advantages of fuzzy system and neural network, and overcomes the traditional fuzzy logic reasoning process, which relies on operator thinking and human intervention to design the membership function graph, so as to reduce the weakness of system efficiency, and uses the system's own training and learning process to solve complex adaptive problems. The algorithm is high efficiency, fast convergence and high model accuracy. It has the advantages of expression of fuzzy logic and self-learning ability of neural network easily, which has gradually become an important research direction of computational intelligence in recent years. And it has the characteristics of simple calculation and mathematical analysis, which provides an effective tool for the modeling and control of complex systems. This also provides a good reference for the use of transportation field. The back-propagation algorithm and the least-squares method are used to adjust the precondition parameters and conclusion parameters, and the If-Then rules can be automatically generated. From the typical representative input and output data, the mapping rules of input and output are directly concluded, that is, given some sample points of input and output, the model's input and output characteristics are obtained by learning. This method greatly simplifies the modeling problem, avoids the modification of the original model in theory, and directly uses the trained ANFIS model to describe the law of input and output, which is its greatest advantage. What's more, ANFIS is not based on experience or knowledge, but obtains the membership function and fuzzy rules of the system through learning a large number of sample data, and adjusts and optimizes the parameters of the front and back parts according to the self-learning characteristics of the neural network, so as to improve the performance of the fuzzy system. This laid the foundation for the analysis of traffic safety data.

Data Availability

The data (excel) used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research is supported by the Department of Shanxi Communications Science and Technology (Project No.: 16002380370): Research in the Impact of Truck Proportion on Highway Traffic Safety, the People's Republic of China. This project is also supported by the China Scholarship Council.

References

[1] J. Archer and W. Young, "Signal treatments to reduce heavy vehicle crash-risk at metropolitan highway intersections," *Accident Analysis & Prevention*, vol. 41, no. 3, pp. 404–411, 2009.

[2] A. Agarwal, M. Zilske, K. R. Rao, and N. Kai, "An elegant and computationally efficient approach for heterogeneous traffic modelling using agent based simulation," *Procedia Computer Science*, vol. 52, pp. 962–967, 2015.

[3] C. Bachmann, M. J. Roorda, and B. Abdulhai, "Simulating traffic conflicts on truck-only infrastructure using an improved time-to-collision definition," *Transportation Research Board Meeting*, Transportation Research Board, Washington, DC, 2010.

[4] C. Chai and Y. D. Wong, "Comparison of two simulation approaches to safety assessment: cellular automata and SSAM," *Journal of Transportation Engineering*, vol. 141, no. 6, 05015002 pages, 2015.

[5] C. Dong, D. B. Clarke, S. H. Richards, and B. Huang, "Differences in passenger car and large truck involved crash frequencies at urban signalized intersections: an exploratory analysis," *Accident Analysis & Prevention*, vol. 62, pp. 87–94, 2014.

[6] U. Durrani, C. Lee, and H. Maoh, "Calibrating the wiedemann's vehicle-following model using mixed vehicle-pair interactions," *Transportation Research Part C: Emerging Technologies*, vol. 67, pp. 227–242, 2016.

[7] C. Dong, Q. Dong, B. Huang, W. Hu, and S. S. Nambisan, "Estimating factors contributing to frequency and severity of large truck–involved crashes," *Journal of Transportation Engineering, Part A: Systems*, vol. 143, no. 8, 04017032 pages, 2017.

[8] S. El-Tantawy, S. Djavadian, M. Roorda, and B. Abdulhai, "Safety evaluation of truck lane restriction strategies using microsimulation modeling," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2099, pp. 123–131, 2009.

[9] D. Gettman, L. Pu, T. Sayed, S. Shelby, and I. T. S. Siemens, *Surrogate safety assessment model and validation (No. FHWA-HRT-08-051)*, Office of Safety Research and Development, Federal Highway Administration US, 2008.

[10] M. Hosseinpour, A. S. Yahaya, S. M. Ghadiri, and J. Prasetijo, "Application of adaptive neuro-fuzzy inference system for road accident prediction," *KSCE Journal of Civil Engineering*, vol. 17, no. 7, pp. 1761–1772, 2013.

[11] Q. Hou, A. P. Tarko, and X. Meng, "Analyzing crash frequency in freeway tunnels: a correlated random parameters approach," *Accident Analysis & Prevention*, vol. 111, pp. 94–100, 2018.

[12] M. T. Islam and K. El-Basyouny, "Full Bayesian mixed-effect intervention model for before–after speed data analysis," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2513, pp. 11–20, 2015.

[13] J. S. Jang, "ANFIS: adaptive-network-based fuzzy inference system," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 23, no. 3, pp. 665–685, 1993.

[14] E. Jeong and C. Oh, "Evaluating the effectiveness of active vehicle safety systems," *Accident Analysis & Prevention*, vol. 100, pp. 85–96, 2017.

[15] X. Li, X. Li, Y. Xiao, and B. Jia, "Modeling mechanical restriction differences between car and heavy truck in two-lane cellular automata traffic flow model," *Physica A: Statistical Mechanics and its Applications*, vol. 451, pp. 49–62, 2016.

[16] J. C. Milton, V. N. Shankar, and F. L. Mannering, "Highway accident severities and the mixed logit model: an exploratory empirical analysis," *Accident Analysis & Prevention*, vol. 40, no. 1, pp. 260–266, 2008.

[17] A. Mehar, S. Chandra, and S. Velmurugan, "Highway capacity through VISSIM calibrated for mixed traffic conditions," *KSCE Journal of Civil Engineering*, vol. 18, no. 2, pp. 639–645, 2014.

[18] A. Montella, L. L. Imbriani, V. Marzano, and F. Mauriello, "Effects on speed and safety of point-to-point speed enforcement systems: evaluation on the urban motorway A56 tangenziale di Napoli," *Accident Analysis & Prevention*, vol. 75, pp. 164–178, 2015.

[19] G. Mwesige, H. Farah, and H. N. Koutsopoulos, "Risk appraisal of passing zones on two-lane rural highways and policy applications," *Accident Analysis & Prevention*, vol. 90, pp. 1–12, 2016.

[20] A. M. Molan and J. E. Hummer, "Safety analysis of the new synchronized and milwaukee B interchanges in comparison to existing designs," *Accident Analysis & Prevention*, vol. 109, pp. 29–35, 2017.

[21] B. Németh, Z. Bede, and P. Gáspár, "Modelling and analysis of mixed traffic flow with look-ahead controlled vehicles," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 15639–15644, 2017.

[22] A. M. Rahimi, "Neuro-fuzzy system modelling for the effects of intelligent transportation on road accident fatalities," *Technical Gazette*, vol. 24, no. 4, pp. 1165–1171, 2017.

[23] S. P. Siddharth and G. Ramadurai, "Calibration of VISSIM for Indian heterogeneous traffic conditions," *Procedia-Social and Behavioral Sciences*, vol. 104, pp. 380–389, 2013.

[24] R. Sun, D. Xue, and Y. Zhang, "Rear-end collision detection based on GNSS/compass fusion and adaptive neuro fuzzy inference system," in *In Cooperative Positioning and Service (CPGP)*, IEEE, pp. 273–277, Harbin, China, 2017.

[25] Traffic Management Bureau of the Public Security Ministry, "Traffic Accident Statistics in China," https://www.sohu.com/a/223922497_116921, [EB/OL], 2017.

[26] A. Theofilatos, G. Yannis, P. Kopelias, and F. Papadimitriou, "Predicting road accidents: a rare-events modeling approach," *Transportation Research Procedia*, vol. 14, pp. 3399–3405, 2016.

[27] J. Tang, F. Liu, W. Zhang, R. Ke, and Y. Zou, "Lane-changes prediction based on adaptive fuzzy neural network," *Expert Systems with Applications*, vol. 91, pp. 452–463, 2018.

[28] K. Vogel, "A comparison of headway and time to collision as safety indicators," *Accident Analysis & Prevention*, vol. 35, no. 3, pp. 427–433, 2003.

[29] X. Wang, D. Zhao, H. Peng, and D. J. LeBlanc, "Analysis of unprotected intersection left-turn conflicts based on naturalistic driving data," in *Intelligent Vehicles Symposium (IV)*, IEEE, pp. 218–223, CA, USA, 2017.

[30] K. Xie, K. Ozbay, H. Yang, J. Holguín-Veras, and E. F. Morgul, "Modeling safety impacts of off-hour delivery programs in urban areas," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2478, no. 1, pp. 19–27, 2015.

[31] W. Young, A. Sobhani, M. G. Lenné, and M. Sarvi, "Simulation of safety: a review of the state of the art in road safety simulation modelling," *Accident Analysis & Prevention*, vol. 66, pp. 89–103, 2014.

[32] W. Yin, Q. Zou, C. Lv, Z. Wu, and G. Tan, "A quantitative model of conflict risk degree at non-signalized intersections," *Procedia Engineering*, vol. 137, pp. 171–179, 2016.

[33] W. Zou, X. Wang, and D. Zhang, "Truck crash severity in New York city: an investigation of the spatial and the time of day effects," *Accident Analysis & Prevention*, vol. 99, pp. 249–261, 2017.