Hindawi
Advances in Civil Engineering
Volume 2018, Article ID 8702854, 17 pages
https://doi.org/10.1155/2018/8702854

Research Article

Proposing New Methods to Estimate the Safety Level in Different Parts of Freeway Interchanges

Hamid Behbahani, Sayyed Mohsen Hosseini, Alireza Taherkhani, Hemin Asadi, and Seyed Alireza Samerei

Department of Civil Engineering, Iran University of Science and Technology, P.O. Box 13114-16846, Tehran, Iran

Correspondence should be addressed to Hamid Behbahani; behbahani@iust.ac.ir

Received 18 August 2017; Revised 7 December 2017; Accepted 27 December 2017; Published 29 March 2018

Academic Editor: Alessandro Palmeri

Since attention to the safety of traffic facilities including freeway interchanges has been increased during recent years, accident prediction models are being developed. Simulation-based surrogate safety measures (SSMs) have been used in the absence of real collision data. But, obtaining different outputs from different SSMs as safety indicators had led to a complexity of using them as the collision avoidance system basis. Additionally, applying SSM requires trajectory data which can be hardly obtained from video processing or calibrated microsimulations. Estimating safety level in different parts of freeway interchanges through a new proposed method was considered in this paper. Fuzzy logic was applied to combine the outputs of different SSMs, and an index called no-collision potential index (NCPI) was defined. 13608 calibrated simulations were conducted on different ramps, weaving, merge, and diverge areas with different geometrical and traffic characteristics, and NCPI was determined for every case. The geometrical and traffic characteristics formed input data of two safety estimator models developed by Artificial Neural Network and Particle Swarm Optimization. Ten freeway interchanges were investigated to calibrate the simulations and to ensure the validity of the fuzzy method and accuracy of the models. Results showed an appropriate and accurate development of the models.

1. Introduction

Providing an acceptable safety level of traffic facilities is logically vital due to its consequent effects on prevention of fatality and property damage. On the other hand, the low level of safety shall lead to the low performance of these facilities, as well. Since the freeways have always played a significant role in road transportation, determination of the safety level in different segments of the freeways has been one of the main concerns of researchers. Previous investigations are divided into two categories. These two categories include the real-accident data analyzing and the simulation-based safety study. In the real-accident studies, the effect of some factors (such as mainline speed at the beginning of the weaving segments, the speed difference between the beginning and the end of the weaving segment, the logarithm of the volume, the maximum length of the weaving area [1], the heavy vehicle rate, the hourly traffic volume, the speed differential between cars and heavy vehicles, the number of accesses [2], the number of vehicles that enter and exit the freeway at a specific segment, the length of the speed change lane, the speed of off-ramps [3], parallel-type or taper-type exit ramps [4], left-side or right-side merging and diverging areas [5, 6], the number of lanes on freeways, the number of lanes on ramps, and the speeding-related crashes [7]) on the number and/or severity of accidents was separately investigated and some models were developed. When there are no registered data about the accidents in a specific facility *or* when someone is trying to design that facility *or* when the facility has not yet been built, simulation-based safety studies/estimations such as conflict analysis by microscopic simulation and surrogate safety measures (SSMs) including

time-to-collision (TTC), postencroachment time (PET), proportion of stopping distance (PSD), crash potential index (CPI), unsafe density (UD), max speed (Max S), relative speed (ΔV), kinetic energy (KE), and deceleration rate to avoid collision (DRAC) are often used to estimate the danger or risk of possible collisions. Integrating the rate of TTC variation and the level of hazard associated with TTC as an approximate safety indicator [8] and calculating the risk of sideswipe collisions [9] are some kind of using these measures. An FHWA-sponsored research project has also studied the potential to derive SSMs from existing traffic simulation models. In that research, TTC, PET, and DRAC were used to measure the severity of the conflict, and by use of additional information about the mass of the vehicles, Max S and ΔV were used to measure the severity of the potential collision [10]. These are some examples of using simulation-based SSMs. These measures are so useful because they occur more frequently than accidents and therefore need shorter periods of the study compared with real-accident investigations [11]. It is important to note that the results of using SSMs mostly showed a good relationship between the proposed SSM and actual accident data [12]. So, using the SSM will help to stay far from long-term real-accident data analysis. But, there are some shortcomings in this way. Different surrogate measures for safety result in different outcomes, and unfortunately, an acceptable method to choose the best outcome especially to use as the basis of collision avoidance systems has not yet been presented. On the other hand, using these measures requires some prerequisites such as providing microsimulation software, simulation experts, and trajectory data. In addition, when it is intended to investigate the safety aspects of interchanges, there is not enough literature which focused well on these traffic facilities to review. Here, the questions are that by focusing on freeway interchanges:

(1) Is it possible to combine the outcomes of different SSMs and have an exclusive outcome which could be a safety indicator of the interchange or not?

(2) Could the level of safety be indicated just by having traffic and geometrical characteristics of the interchange instead of simulation attempts, trajectory data, or accident data?

The first purpose of the paper is to present a method to show that it is possible to give a positive answer to the first question. To do this, the fuzzy inference system (FIS) was used to combine the outcomes of the SSM, and fuzzy logic was applied to define an index called no-collision potential index (NCPI) as a safety indicator of the interchange (i.e., the exclusive outcome). Fuzzy logic could be a suitable technique where qualitative definitions cannot be directly quantified. The next purpose is to develop a model to estimate the safety level in the ramps, weaving, merge, and diverge areas of interchanges based on their traffic and geometrical properties. One of the approaches in developing the model is using a powerful tool dealing with the prediction and classification problems, and that is Artificial Neural Network (ANN) which has been a consistent alternative method to analyze the freeway accident frequency and does not require any predefined underlying relationship between dependent and independent variables [13]. Another approach is using an evolutionary computation algorithm motivated by the behavior of organisms and gathered social psychology principles in sociocognition human agents and evolutionary computations. The algorithm is particle swarm optimization (PSO) which has a simple concept and efficiency in computations and is implemented easily [14–17].

2. Methods

2.1. Description of the Proposed Fuzzy Method

2.1.1. NCPI Definition and Determination. In this paper, a fuzzy-based method was proposed to estimate the safety level in different parts of freeway interchanges. In this method, different outputs of SSMs were combined, and the outcome was NCPI which is defined as a safety indicator. The value of NCPI falls within the range of zero to 100. The higher the value of NCPI, the higher the safety level. Since it was intended to indicate the level of safety despite many other types of research that usually determine the level of risk, NCPI was defined to show the likelihood of no collision in the interchange instead of indicating the probability of collisions to occur. To determine the NCPI, the SSM were categorized into two groups:

Number estimation measures: the measures consist of TTC, PET, PSD, CPI, UD, and DRAC which are usually used to measure the severity of the conflicts. These measures can be used to estimate the number of near-crash events.

Severity estimation measures: the measures include Max S, ΔV, and KE which are usually used to measure the severity of the potential collisions.

Due to the complexity and time-consuming of using all measures, in this research study, the measures of TTC and DRAC were selected to estimate the number of possible collisions, and the severity of these potential collisions was estimated by the measures KE and ΔV.

2.1.2. Definition of the Used SSM. Here, the definitions of the measures TTC, DRAC, KE, and ΔV are briefly described.

TTC: this was first defined by Hayward as the remaining collision occurrence time between two vehicles if the collision course and speed difference were maintained constant [18]. When TTC is low, there is an imminent danger of collision [19]. TTC for rear-end conflicts can be calculated by [20]

$$\mathrm{TTC}_{\mathrm{F}}(t) = \frac{X_{\mathrm{L}}(t) - X_{\mathrm{F}}(t) - l_{\mathrm{L}}}{V_{\mathrm{F}}(t) - V_{\mathrm{L}}(t)} \quad \forall V_{\mathrm{F}}(t) > V_{\mathrm{L}}(t), \tag{1}$$

where TTC is the time-to-collision, X is the vehicle position (L: leading and F: following), V is the vehicle speed (L: leading and F: following), and l is the vehicle length.

DRAC: deceleration rate is a good measure to detect dangerous maneuvers. DRAC is the rate at which a vehicle

must decelerate to avoid a probable collision. For vehicles traveling in the same path, DRAC is defined as [11]

$$\mathrm{DRAC}_t = \frac{\left(V_{\mathrm{F}(t)} - V_{\mathrm{L}(t)}\right)^2}{2\left[\left(X_{\mathrm{L}(t)} - X_{\mathrm{F}(t)}\right) - l_{\mathrm{veh\ L}}\right]}, \quad (2)$$

where $l_{\mathrm{veh\ L}}$ is the length of the leading vehicle and other parameters were described previously. For angled conflicts, the equation changes as

$$\mathrm{DRAC}_t = \frac{\Delta V_{ij(t)}^{\ 2}}{2D_{i(t)}}, \quad (3)$$

where $\Delta V_{ij(t)}$ is the relative speed of two vehicles engaged in the conflict and $D_{i(t)}$ is the distance between the current position of the vehicle i and point of the intersection ahead of two vehicles.

KE: from Newtonian physics, we know that a moving vehicle has a kinetic energy as [21]

$$K = \frac{1}{2}mv^2, \quad (4)$$

where K is the kinetic energy, m is the mass, and v is the speed of the vehicle. The kinetic energy transferred to the target vehicle can be calculated by

$$\mathrm{KE}_s = \frac{1}{2}m_s\Delta v_s^{\ 2}, \quad (5)$$

where KE_s is the kinetic energy transferred to the target vehicle, m_s is the mass of the target vehicle, and Δv_s is the change of the target vehicle speed before and after the collision [21].

Δv: Δv is the relative speed of vehicles involved in the conflict as a collision severity reflector [22].

2.1.3. Trajectory Data Analysis. Finding the variables' values of (1–5) requires analyzing the data of trajectory. It was assumed that the vehicles had a linear movement with a constant acceleration or deceleration rate between every two consecutive time intervals in the analysis. So, the coordinates of the intersection ahead of two vehicles i and j could be computed. The acceleration or deceleration rate and speed of vehicles could also be determined using these assumptions. It should be noted that any collision in the study areas could either be a rear-end collision or occurs at an angle of β. Thus, the analysis should be done with respect to angled collisions. In a special case in which the angle of collision is zero, it will be a rear-end collision. The collision of two vehicles at an angle of β is depicted in Figure 1.

On the other hand, it is necessary to check whether the coordinates are within the limits of the study area or not. By assuming no changes in the conditions, it can be concluded that two vehicles i and j will never collide with each other if the coordinates of the intersection ahead of the two vehicles are outside of the limits of the study area. The distance between the position of each vehicle to the intersection ahead and the time required to reach this point can be obtained by simple calculations. Assume that

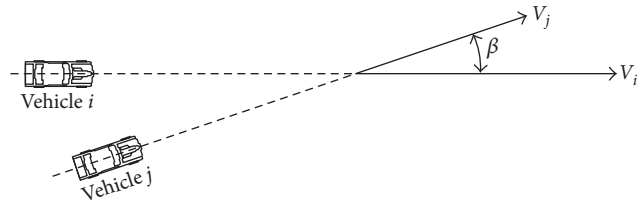

Figure 1: Collision of two vehicles at an angle of β.

T_i and T_j are the time required to reach the point of intersection of vehicles i and j, respectively. While the absolute value of the difference between T_i and T_j is less than the TTC threshold, there will be a near-crash event. The value of the TTC threshold varies in several studies. The proposed values of thresholds are 1.5 seconds for urban areas [23], 4 seconds to distinguish between safe and unsafe situations [12], and 1.0 second for critical situations [24]. But, according to the surrogate safety assessment model, which considers a conflict as an event with TTC less than 1.5 seconds [25], a threshold of 1.5 seconds was used in this paper. Among different pairs of vehicles i and j, the number of those that encountered a near-crash event was counted by

$$C_{\mathrm{TTC}} = \mathrm{Count}_i^j\left[\left|T_i - T_j\right| < \mathrm{TTC}_{\mathrm{threshold}}\right], \quad (6)$$

where C_{TTC} is the number of pairs of vehicles that encountered a near-crash event. Regarding the angled collision of two vehicles i and j, the relative speed of these two vehicles at the time of the collision could be determined by (7) and (8) in each direction of vehicles i and j.

$$\Delta V_{ij} = V_i - V_j\cos\beta, \quad (7)$$

$$\Delta V_{ji} = V_j - V_i\cos\beta, \quad (8)$$

where β is the angle of collision. The minimum required DRAC could be also found by

$$\mathrm{DRAC}_{ij} = 0.5\Delta V_{ij}^{\ 2}\,L_{ij}^{-1}, \quad (9)$$

$$\mathrm{DRAC}_{ji} = 0.5\Delta V_{ji}^{\ 2}\,L_{ji}^{-1}. \quad (10)$$

Maximum deceleration rate for different vehicles was proposed by Maurya and Bokare [26]. If the value of each of the DRACs is more than the maximum deceleration rate, there will be a near-crash event. The DRAC of a couple of vehicles, i and j, is the maximum DRAC of them as written in (11). The number of cases in which a near-crash event takes place will be counted by

$$\mathrm{DRAC} = \max\left\{\mathrm{DRAC}_{ij}\cdot\mathrm{DRAC}_{ji}\right\}, \quad (11)$$

$$C_{\mathrm{DRAC}} = \mathrm{Count}_i^j\left[\mathrm{DRAC} > \text{maximum deceleration rate}\right], \quad (12)$$

where C_{DRAC} is the number of cases in which a near-crash event occurs. The speed of the vehicle i at the time of collision is calculated by

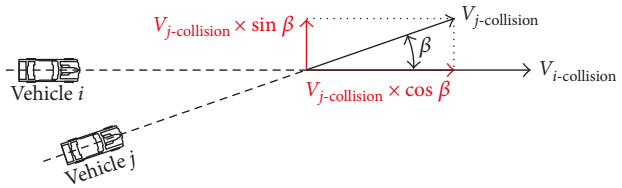

Figure 2: Decomposition of the speed vectors of vehicle j parallel with and perpendicular to the movement direction of vehicle i.

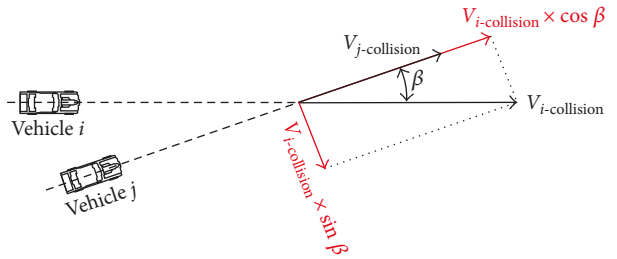

Figure 3: Decomposition of the speed vectors of vehicle i parallel with and perpendicular to the movement direction of vehicle j.

$$V_{i\text{-collision}} = a_i T_i + V_i, \tag{13}$$

where a_i is the acceleration/deceleration rate of the vehicle i and V_i is the speed of the vehicle i at the time of analysis. Decomposition of the speed vectors at the moment of collision is shown in Figures 2 and 3.

So, the relative speed of two vehicles i and j at the moment of the collision could be determined by

$$\overrightarrow{\Delta V} = \overrightarrow{\Delta V_{i\text{-para.}}} + \overrightarrow{\Delta V_{i\text{-perp.}}} = \overrightarrow{\Delta V_{j\text{-para.}}} + \overrightarrow{\Delta V_{j\text{-perp.}}}, \tag{14}$$

$$\left|\overrightarrow{\Delta V_{i\text{-para.}}}\right| = \left|\overrightarrow{V_{i\text{-collision}}}\right| - \left|\overrightarrow{V_{j\text{-collision}}}\right|\cos\beta, \tag{15}$$

$$\left|\overrightarrow{\Delta V_{i\text{-perp.}}}\right| = \left|\overrightarrow{V_{j\text{-collision}}}\right|\sin\beta, \tag{16}$$

$$\left|\overrightarrow{\Delta V_{j\text{-para.}}}\right| = \left|\overrightarrow{V_{j\text{-collision}}}\right| - \left|\overrightarrow{V_{i\text{-collision}}}\right|\cos\beta, \tag{17}$$

$$\left|\overrightarrow{\Delta V_{j\text{-perp.}}}\right| = \left|\overrightarrow{V_{i\text{-collision}}}\right|\sin\beta, \tag{18}$$

$$\begin{aligned} \Delta V &= \left|\overrightarrow{\Delta V}\right| = \left[\left|\overrightarrow{\Delta V_{i\text{-para.}}}\right|^2 + \left|\overrightarrow{\Delta V_{i\text{-perp.}}}\right|^2\right]^{0.5} \\ &= \left[\left|\overrightarrow{\Delta V_{j\text{-para.}}}\right|^2 + \left|\overrightarrow{\Delta V_{j\text{-perp.}}}\right|^2\right]^{0.5}. \end{aligned} \tag{19}$$

If the difference between the speed of vehicle j before and after the collision is equal to the ΔV obtained above, the amount of kinetic energy transferred in the collision will be defined as

$$\mathrm{KE}_{ij} = 0.5\, m_j \Delta V^2, \tag{20}$$

where m_j is the mass of vehicle j.

2.1.4. Probabilistic Framework. Four surrogate measures were used to estimate the number and the severity of possible collisions. But the probability that a near-crash event becomes a real collision should be determined and be considered in the model development. There are two following probabilities:

(1) The probability that after detection of an event as a near-crash event, the event becomes a real collision (it is possible that both or one of the drivers prevents the collision to take place by performing an evasive act).

(2) The probability that the severity of collision does not change (it is possible that both or one of the drivers engaged in a collision reduces the speed to avoid the collision. Even if the collision occurs, the collision severity will become lower).

When both or one of the drivers has a time to react, both mentioned probabilities will appear. In other words, the more the TTC, the less the probability of the collision taking place with a certain severity. A probability density function (PDF) was defined which was sensitive to the perception-reaction time of drivers. For this purpose, an exponential PDF was selected among the most appropriate distributions which satisfy the requirements of the problem and was written as

$$\mathrm{Pr}_i = \lambda e^{-\lambda\left(0.5\,\mathrm{TTC}_i^2\, t_{\text{reaction}}^{-2}\right)}, \tag{21}$$

where Pr_i is the probability that a near-crash event with a certain severity becomes a real collision with the same severity, t_{reaction} is the perception-reaction time of drivers, and λ is a constant value which should be defined based on the problem. Since the probability must be closed to one when TTC approaches zero second, the value of λ should be equal to 1.0.

The American Association of State Highway and Transportation Officials (AASHTO) mandates using a perception-reaction time of 2.7 seconds in most calculations [27]. But, in certain more complex situations, drivers may require more time to react than 2.7 seconds, for example, situations where drivers should detect and react to unexpected events. According to the AASHTO, where a collision avoidance maneuver is needed, the perception-reaction time of 3.0 seconds should be used for rural roads and 9.1 seconds for urban roads. In situations that avoiding a collision needs alterations in the speed, path, and/or direction, the AASHTO recommends a range of perception-reaction time between 10.2 and 11.2 seconds for rural roads, 12.1 and 12.9 seconds for suburban roads, and 14.0 and 14.5 seconds for urban roads [28]. Due to the high speed of vehicles, disorderly movements of vehicles, and a high number of lane changes in a freeway interchange, avoiding a collision certainly requires changing the speed, path, and direction. So, the perception-reaction time of 11.2 seconds was used in this paper. Figure 4 presents the relation between the probability and TTC with $\lambda = 1.0$ and $t_{\text{reaction}} = 11.2$.

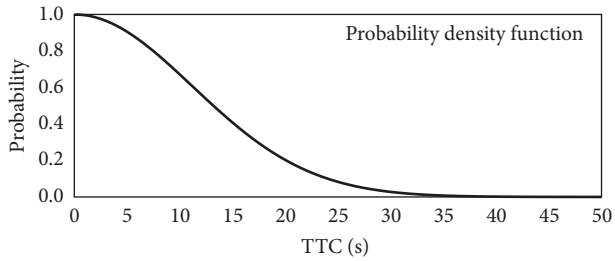

Figure 4: PDF for a near-crash event becomes a real collision.

The number and the severity of collisions in different parts of freeway interchanges could be better estimated by applying the above probability as

$$N_{\mathrm{TTC}} = C_{\mathrm{TTC}}^{-1} \times \sum_{i=1}^{C_{\mathrm{TTC}}} \mathrm{Pr}_i C_{\mathrm{TTC}\,i}, \tag{22}$$

$$N_{\mathrm{DRAC}} = C_{\mathrm{DRAC}}^{-1} \times \sum_{i=1}^{C_{\mathrm{DRAC}}} \mathrm{Pr}_i C_{\mathrm{DRAC}\,i}, \tag{23}$$

$$S_{\mathrm{KE}} = C_{\mathrm{TTC}}^{-1} \times \sum_{i=1}^{C_{\mathrm{TTC}}} \left(\mathrm{Pr}_i \times \mathrm{KE}_i\right), \tag{24}$$

$$S_{\Delta V} = C_{\mathrm{TTC}}^{-1} \times \sum_{i=1}^{C_{\mathrm{TTC}}} \left(\mathrm{Pr}_i \times \Delta V_i\right), \tag{25}$$

where N_{TTC} is the number of predicted collisions by the TTC, N_{DRAC} is the number of predicted collisions by the DRAC, S_{KE} is the severity of predicted collisions by the KE, and $S_{\Delta V}$ is the severity of predicted collisions by the ΔV.

2.1.5. Fuzzy Logic. The estimated number of collisions using the measures of TTC and DRAC is usually different, and the results of the severity of the predicted collisions using the measure of KE are not the same as that of the measure of ΔV, as well. But, all the outputs from the four measures should be considered all together. So, the NCPI must be a function of the outputs of the four mentioned measures as four variables as shown in

$$\mathrm{NCPI} = F\left(N_{\mathrm{TTC}} \cdot N_{\mathrm{DRAC}} \cdot S_{\mathrm{KE}} \cdot S_{\Delta V}\right). \tag{26}$$

The determining function of the NCPI was defined using a Mamdani-type FIS with four inputs (the values of the four measures) and one output (NCPI). At first, a membership function with three qualitative classes was defined for each measure. Then, a *fuzzification* procedure was implemented for the measures. It means that the quantitative value of each measure was automatically placed in one of the three qualitative classes of "low," "medium," and "high" using the membership functions. Between NCPI and its variables, 81 rules were set. These rules determine the qualitative value of NCPI by taking the qualitative values of the four measures into account simultaneously. Table 1 presents the 81 applied

Table 1: The rules of the FIS.

Rule number	Variables				Function
	$S_{\Delta V}^{-1}$	S_{KE}^{-1}	N_{TTC}^{-1}	N_{DRAC}^{-1}	NCPI
1	Low	Low	Low	Low	*Very low*
2	Low	Low	Low	Medium	*Very low*
3	Low	Low	Low	High	*Low*
4	Low	Low	Medium	Low	*Very low*
5	Low	Low	Medium	Medium	*Low*
6	Low	Low	Medium	High	*Medium*
7	Low	Low	High	Low	*Low*
8	Low	Low	High	Medium	*Medium*
9	Low	Low	High	High	*High*
10	Low	Medium	Low	Low	*Very low*
11	Low	Medium	Low	Medium	*Low*
12	Low	Medium	Low	High	*Medium*
13	Low	Medium	Medium	Low	*Low*
14	Low	Medium	Medium	Medium	*Medium*
15	Low	Medium	Medium	High	*Medium*
16	Low	Medium	High	Low	*Medium*
17	Low	Medium	High	Medium	*Medium*
18	Low	Medium	High	High	*High*
19	Low	High	Low	Low	*Low*
20	Low	High	Low	Medium	*Medium*
21	Low	High	Low	High	*Medium*
22	Low	High	Medium	Low	*Medium*
23	Low	High	Medium	Medium	*Medium*
24	Low	High	Medium	High	*High*
25	Low	High	High	Low	*Medium*
26	Low	High	High	Medium	*High*
27	Low	High	High	High	*Very high*
28	Medium	Low	Low	Low	*Very low*
29	Medium	Low	Low	Medium	*Low*
30	Medium	Low	Low	High	*Low*
31	Medium	Low	Medium	Low	*Low*
32	Medium	Low	Medium	Medium	*Low*
33	Medium	Low	Medium	High	*Medium*
34	Medium	Low	High	Low	*Low*
35	Medium	Low	High	Medium	*Medium*
36	Medium	Low	High	High	*High*
37	Medium	Medium	Low	Low	*Low*
38	Medium	Medium	Low	Medium	*Low*
39	Medium	Medium	Low	High	*Medium*
40	Medium	Medium	Medium	Low	*Low*
41	Medium	Medium	Medium	Medium	*Medium*
42	Medium	Medium	Medium	High	*High*
43	Medium	Medium	High	Low	*Medium*
44	Medium	Medium	High	Medium	*High*
45	Medium	Medium	High	High	*Very high*
46	Medium	High	Low	Low	*Low*
47	Medium	High	Low	Medium	*Medium*
48	Medium	High	Low	High	*High*

Table 1: Continued.

Rule number	Variables				Function
	$S_{\Delta V}^{-1}$	S_{KE}^{-1}	N_{TTC}^{-1}	N_{DRAC}^{-1}	NCPI
49	Medium	High	Medium	Low	*Medium*
50	Medium	High	Medium	Medium	*High*
51	Medium	High	Medium	High	*High*
52	Medium	High	High	Low	*High*
53	Medium	High	High	Medium	*High*
54	Medium	High	High	High	*Very high*
55	High	Low	Low	Low	*Very low*
56	High	Low	Low	Medium	*Low*
57	High	Low	Low	High	*Medium*
58	High	Low	Medium	Low	*Low*
59	High	Low	Medium	Medium	*Medium*
60	High	Low	Medium	High	*High*
61	High	Low	High	Low	*Medium*
62	High	Low	High	Medium	*High*
63	High	Low	High	High	*High*
64	High	Medium	Low	Low	*Low*
65	High	Medium	Low	Medium	*Medium*
66	High	Medium	Low	High	*Medium*
67	High	Medium	Medium	Low	*Medium*
68	High	Medium	Medium	Medium	*Medium*
69	High	Medium	Medium	High	*High*
70	High	Medium	High	Low	*Medium*
71	High	Medium	High	Medium	*High*
72	High	Medium	High	High	*Very high*
73	High	High	Low	Low	*Medium*
74	High	High	Low	Medium	*Medium*
75	High	High	Low	High	*High*
76	High	High	Medium	Low	*Medium*
77	High	High	Medium	Medium	*High*
78	High	High	Medium	High	*Very high*
79	High	High	High	Low	*High*
80	High	High	High	Medium	*Very high*
81	High	High	High	High	*Very high*

rules. A membership function for NCPI was defined with five qualitative classes of "very low," "low," "medium," "high," and "very high." So, a qualitative index was attributed to NCPI based on the rules and the membership function. After all, a defuzzification procedure using the centroid method was accomplished, and the quantitative value of NCPI was obtained by the use of NCPI membership function. The FIS settings for AND/OR methods were set to MIN/MAX and for IMP/AGGR methods were set to MIN/MAX, as well. Figure 5 shows the membership functions of NCPI and N_{TTC}, for example, and Figure 6 depicts two samples of 3D profiles of NCPI and its variables.

2.1.6. Field Study. The field study was carried out to reach the following three aims:

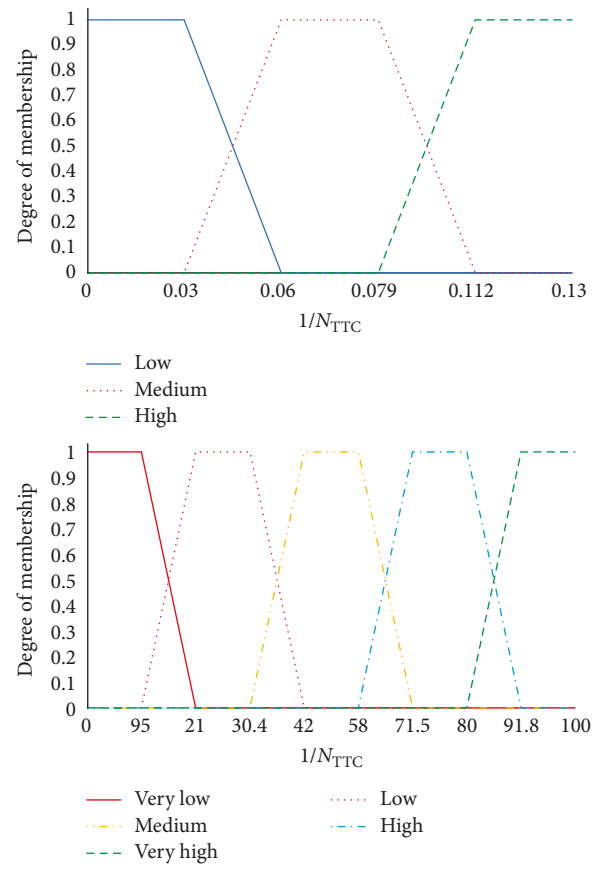

Figure 5: The membership functions for NCPI and N_{TTC}.

(1) To calibrate the simulations

(2) To control the validity of the proposed method in which the safety level is indicated using fuzzy logic and SSM

(3) To check the accuracy of safety estimator models (the safety estimator models will be described in the following sections)

Ten freeway interchanges in Tehran Province in Iran were investigated. All the traffic and geometrical characteristics of the interchanges were gathered and shown in Table 2.

The behavior of the drivers and the trajectory data were extracted by video processing tools. The driving times in different parts of the interchanges, the headways, the real acceleration and deceleration scenarios, the number of lane changes, the route selections of the drivers, the length before the gores which the lane change occurs, and the accepted gap for the lane change were obtained by analyzing the behavior of the drivers. According to the gathered data and analysis on the behavior of the drivers, the simulation was calibrated. The value of the NCPI was also computed for every case by analyzing the real trajectory data and applying the proposed fuzzy method. The interchanges were simulated, and trajectory data were derived. The NCPIs computed using simulation trajectory data were compared with the NCPIs obtained from real trajectory data for controlling the validity

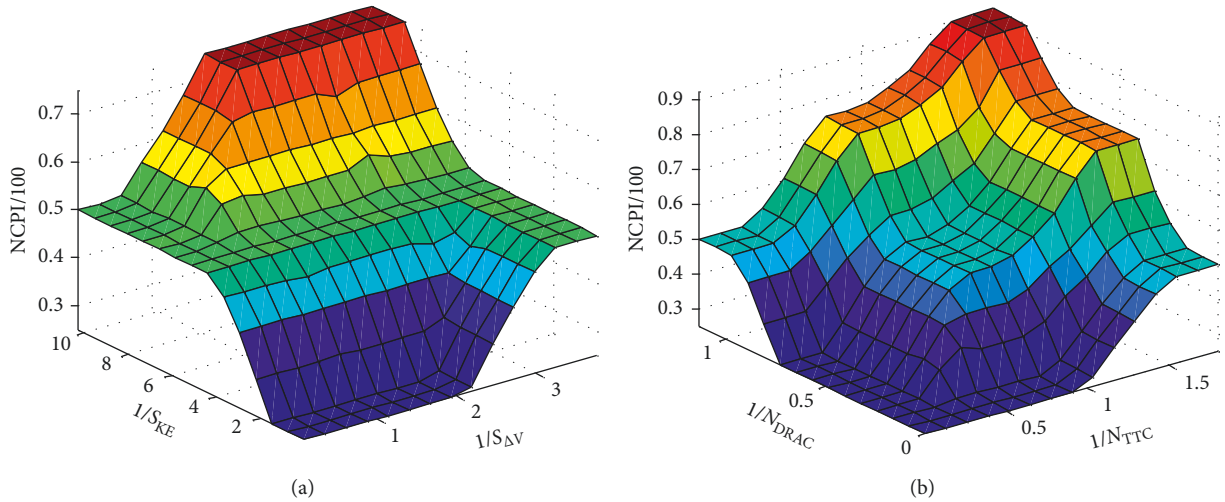

Figure 6: The 3D profiles of NCPI and its variables. (a) Severity estimation variables; (b) number estimation variables.

of the method. Figure 7 presents the aerial photos of the ten studied interchanges. The results were used to check the accuracy of the future developed safety estimator models, too.

2.2. Development of Safety Estimator Models. In the previous section, it was shown that the safety level could be determined using microsimulation, trajectory data analysis and computations, SSM, and the proposed fuzzy method and computing NCPI. But, providing expensive prerequisites like simulation software, simulation experts, trajectory computations, and the time required is difficult and time-consuming. In this paper, it was intended to develop two models to estimate the level of safety based on geometrical and traffic characteristics of freeway interchanges. Developing the models needs generating a database containing geometrical and traffic characteristics of the ramps, weaving, merge, and diverge areas as inputs and computed NCPIs as output. Then, ANN and PSO were used to develop the models.

2.2.1. Generating Database. Every effective geometrical and traffic variable which may affect the safety level of the interchange was recognized, and a list of these variables comprising the length of the weaving area, the number of lanes in the weaving area, the number of lanes in on-ramp, the number of lanes in off-ramp, the free flow speed in freeway, the free flow speed in on-ramp, the free flow speed in off-ramp, the freeway flow rate, the on-ramp flow rate, the off-ramp flow rate, the length of the acceleration lane, the length of the deceleration lane, the length of the interchange ramp, the average slope of the ramp, and the ramp radius was prepared. Table 3 describes the variables and their range used for the simulations. Combining various values of each variable with those of the other variables resulted in the generation of 13608 different ramps, weaving, merge, and diverge areas with different traffic and geometrical characteristics. Different kinds of interchange ramps, weaving, merge, and diverge areas with different geometry and different traffic characteristics were simulated. The trajectory data were derived from every simulation run. By applying the proposed fuzzy-based method, NCPI was computed for every case. Thus, a database could be generated containing 10368, 2160, 720, and 360 rows of information for weaving areas, merge areas, diverge areas, and ramps, respectively.

As described in Table 4, the database comprised geometrical and traffic characteristics of the ramps, weaving, merge, and diverge areas of the interchange as inputs and computed NCPIs as the output. Using this database, the models could be developed.

2.2.2. ANN Approach. ANNs are massive parallel architectures that can determine answers to demanding problems with the participation of simple mutually dependent processing elements (artificial neurons) [29]. An ANN has powerful aspects of learning and data processing, and thus, it is an effective tool for engineering applications. The multilayer backpropagation network, the most popular ANN paradigm [30], was used for efficient generalization competence. ANNs have the ability to carry out with a good amount of generalization from the patterns on which they are trained [31]. Training incorporated processing neural networks with a set of known specific input-output data using the generated database. The backpropagation ANN included layers with neurons: the input layer, the output layer, and the hidden layers. The learning process was continued in the output layer where the error between the model outputs and desired outputs was calculated and then propagated back to the network with updated weights for the direction in which the performance function decreases very rapidly [32]. The entire training process was repeated for a number of epochs until the desired accuracy in the network output was gained. After training and then, validation of the network, the network was tested by using data that have never experienced before.

Table 2: Characteristics of the interchanges investigated in the field studies.

Weaving area										
Location	L_W (m)	N_W (−)	$N_{R\text{-}ON}$ (−)	$N_{R\text{-}OFF}$ (−)	V_{FW} (veh/h)	$V_{R\text{-}ON}$ (veh/h)	S_{FW} (km/h)	$S_{R\text{-}ON}$ (km/h)	$S_{R\text{-}OFF}$ (km/h)	Number of samples
Hemmat W-E: Yadegar ON and Yadegar OFF	304.71	5	2	2	6650	1264	90	40	40	13
Hemmat Freeway W-E: Asharfi ON and Yadegar OFF	705	5	2	2	6850	839	90	40	40	17
Hakim Freeway E-W: Sheikh Bahaee ON and Chamran OFF	296	4	2	2	2990	895	80	40	60	18
Karbala Freeway E-W: Sahne ON and Hamedan OFF	285	4	2	2	1870	185	110	50	60	21
Imam Ali Freeway S-N: N-N U-turn ON and Sabalan OFF	535.04	5	1	2	5860	157	80	40	60	19
Imam Ali Freeway S-N: Khavaran ON and Khavaran OFF	256	4	2	2	878	833	80	30	40	17

Merge area								
Location	L_{ACC} (m)	V_{FW} (veh/h)	N_{FW} (−)	$V_{R\text{-}ON}$ (veh/h)	$N_{R\text{-}ON}$ (−)	S_{FW} (km/h)	$S_{R\text{-}ON}$ (km/h)	Number of samples
Hemmat Freeway W-E: merge Asharfi N	145	5023	4	1253	2	90	50	11
Niayesh Freeway E-W: merge Chamran S	118	3794	3	909	2	80	40	14
Tehran-Qom Freeway N-S: merge Vahnabad E	173	2252	3	169	1	120	60	25
Tehran-Qom Freeway S-N: merge Vahnabad W	154	1266	3	440	1	120	60	19
Tehran-Saveh Freeway W-E: merge Shahriar W	225	2667	3	361	2	120	40	18

Diverge area							
Location	L_{DEC} (m)	N_{FW} (−)	$N_{R\text{-}OFF}$ (−)	V_{FW} (veh/h)	S_{FW} (km/h)	$S_{R\text{-}OFF}$ (km/h)	Number of samples
Hakim Freeway W-E: diverge Sheikh Bahaee S	172	4	1	3188	80	30	17
Hemmat Freeway W-E: diverge Yadegar S	202	4	2	4196	80	60	13
Tehran-Saveh Freeway E-W: diverge Dehshade W	215	3	2	4160	120	60	22
Tehran-Saveh Freeway E-W: diverge Robat Karim E	180	3	2	1895	120	40	14
Yadegar Freeway N-S: diverge Kouhestan E	152	3	1	1930	80	50	9

Interchange ramps						
Location	L_R (m)	N_R (#)	S_{LONG} (%)	V_R (veh/h)	R_R (m)	Number of samples
Chamran S to Hemmat W	438	2	+3.8	1644	60	26
Hemmat W to Chamran S	350	2	+2.3	943	35	21
Imam Ali N to Mahallati W	402	2	+4.2	796	60	16
Mahallati E to Imam Ali N	514	2	+0.5	234	60	11
Niayesh E to Chamran S	470	2	−1.9	681	60	19

Determining the number of hidden neurons is an important task. The number of hidden neurons has a strong effect on the stability of the neural network which is estimated by error (minimal error reflects better stability). The random selection of a number of hidden neurons may cause either overfitting or underfitting of the predicting models [33]. This concern arises when the network corresponds to the data so closely that the ability to generalize over the test

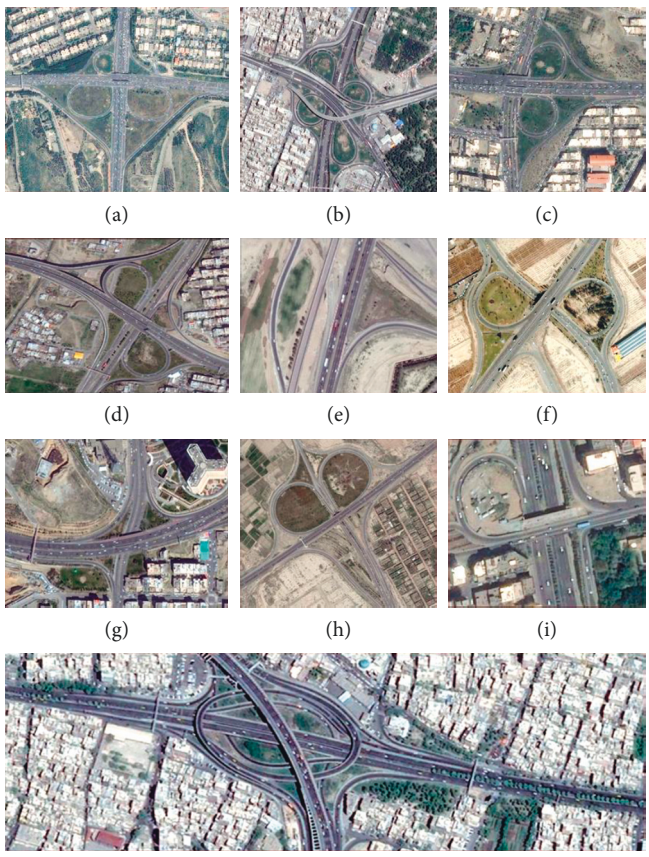

Figure 7: Case studies: (a) Hemmat FW-Yadegar FW; (b) Imam Ali FW-Khavaran FW; (c) Hemmat FW-Ashrafi; (d) Niayesh FW-Chamran FW; (e) Tehran-Qom FW-Vahnabad; (f) Tehran-Saveh FW-Shahryar; (g) Hakim FW-Sheikh Bahaee; (h) Tehran-Saveh FW-Dehshad; (i) Yadegar FW-Kouhestan; (j) Imam Ali FW-Mahallati.

data is lost. An excessive number of hidden neurons will cause overfitting where neural networks overestimate the complexity of the target problem. In this sense, determination of the proper number of hidden neurons to prevent overfitting is critical for problem estimation with the capability for steady generalization with the lowest possible deviation in estimation. Accordingly, the number of hidden neurons must be delimited within a reasonable range. To reach the minimum error, the optimized number of hidden neurons was determined. The ANN model had also a number of neurons in the input layer and one neuron in the output layer. In the trial and error attempts, the number of neurons in the input layer was considered equal to or a multiple of the number of variables in every part of the interchange. Subsequently, the neural network results were evaluated to determine the number of neurons that would provide satisfactory results.

To ensure good generalization under ANN processing, dividing the experimental data into three subdivisions is mandatory: training, validation, and testing. For database partitioning, Looney recommended 25% for testing [34], whereas Swingler proposed 20% [35] and Nelson and

Table 3: Variable description and their range used for simulation.

Variables[a]	Range
Weaving area[a]	
L_W (m)	300 to 900
N_W (#)	4 to 5
$N_{R\text{-}ON}$ (#)	1 to 2
$N_{R\text{-}OFF}$ (#)	1 to 2
V_{FW} (veh/h)	750 to 2970
$V_{R\text{-}ON}$ (veh/h)	600 to 1600
S_{FW} (km/h)	90 to 120
$S_{R\text{-}ON}$ (km/h)	40 to 60
$S_{R\text{-}OFF}$ (km/h)	40 to 60
Merge area[b]	
L_{ACC} (m)	100 to 500
V_{FW} (veh/h)	750 to 2970
N_{FW} (#)	3 to 4
$V_{R\text{-}ON}$ (veh/h)	600 to 1600
$N_{R\text{-}ON}$ (#)	1 to 2
S_{FW} (km/h)	90 to 120
$S_{R\text{-}ON}$ (km/h)	40 to 60
Diverge area[c]	
L_{DEC} (m)	100 to 500
N_{FW} (#)	3 to 4
$N_{R\text{-}OFF}$ (#)	1 to 2
V_{FW} (veh/h)	750 to 2970
S_{FW} (km/h)	90 to 120
$S_{R\text{-}OFF}$ (km/h)	40 to 60
Ramps[d]	
L_R (m)	100 to 500
N_R (#)	1 to 2
S_{LONG} (%)	−3 to 3
V_R (veh/h)	600 to 2200
R_R (m)	60 to 140

[a]L_W is the length of the weaving area, N_W is the number of lanes in the weaving area, $N_{R\text{-}ON}$ is the number of on-ramp lanes, $N_{R\text{-}OFF}$ is the number of off-ramp lanes, V_{FW} is the freeway volume, $V_{R\text{-}ON}$ is the on-ramp volume, S_{FW} is the freeway free-flow speed, $S_{R\text{-}ON}$ is the speed of on-ramp, and $S_{R\text{-}OFF}$ is the speed of off-ramp. [b]L_{ACC} is the length of the acceleration lane, V_{FW} is the freeway volume, N_{FW} is the number of freeway lanes, $V_{R\text{-}ON}$ is the on-ramp volume, $N_{R\text{-}ON}$ is the number of on-ramp lanes, S_{FW} is the freeway free-flow speed, and $S_{R\text{-}ON}$ is the speed of on-ramp. [c]L_{DEC} is the length of the deceleration lane, N_{FW} is the number of freeway lanes, $N_{R\text{-}OFF}$ is the number of off-ramp lanes, V_{FW} is the freeway volume, S_{FW} is the freeway free-flow speed, and $S_{R\text{-}OFF}$ is the speed of off-ramp. [d]L_R is the length of the interchange ramp, N_R is the number of lanes in the interchange ramp, S_{LONG} is the average slope of the ramp, V_R is the ramp flow rate, and R_R is the radius of the ramp.

Table 4: The database description.

Part	Number of rows	Number of variables (inputs)	Function (output)
Weaving areas	10368	9	NCPI
Merge areas	2160	7	NCPI
Diverge areas	720	6	NCPI
Ramps	360	5	NCPI
Total	13608		

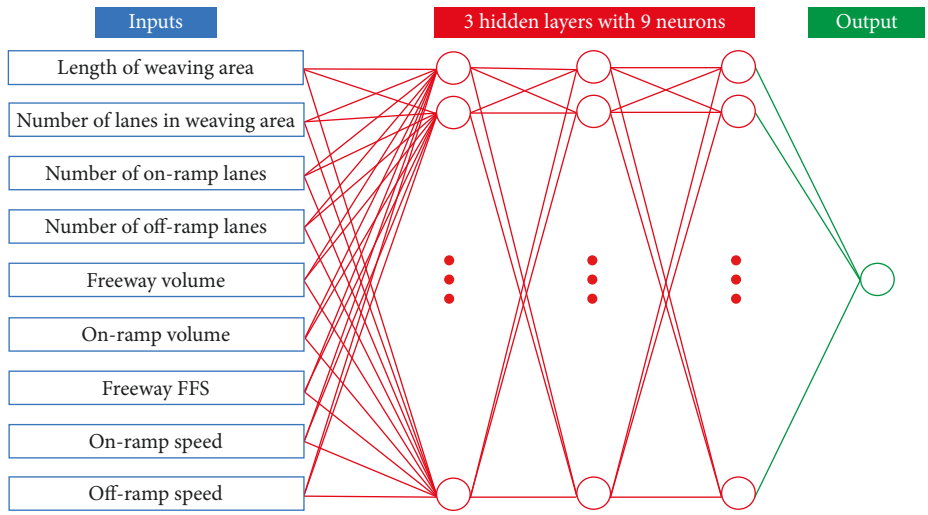

Figure 8: ANN for estimation of the safety level in the weaving area.

Illingworth suggested 20 to 30% [36]. Therefore, in the present work, the training dataset consists of 60% of data entries, and the remaining data entries are divided equally between the validation and testing sets. To test the reliability of the neural network model, 20% of the samples were randomly selected as the validation set and 20% of the samples as the test set. The performance of an ANN-based model primarily depends on the network architecture and parameter settings. In this study, the Matlab ANN toolbox [37] was used for ANN applications. All the networks were trained using the Levenberg–Marquardt algorithm with tansig transfer functions between the input and the hidden layers and the purelin transfer function between the hidden layers and the output layer along with a trainlm training function. Performance evaluation was characterized by plots of root-mean-squared error (RMSE) versus epoch for the training, validation, and test performances of the data. In principle, the error is reduced after more epochs of training but may begin to increase on the validation dataset as the network starts overfitting the training data. To reach the best ANN, its performance was deduced from the epoch with the lowest validation error after which there was no more increase in the RMSE error. Consequently, a reliable ANN-based safety estimator model was developed, and therefore, NCPI could be estimated using adequate information about the values of geometrical and traffic characteristics. The structure of the ANN used for estimation of the safety level in the weaving area is shown in Figure 8, for example.

The properties of the designed ANNs are as follows:

Weaving area:

Number of hidden layers: 3
Number of neurons in each hidden layer: 9

Merge area:

Number of hidden layers: 3
Number of neurons in each hidden layer: 7

Diverge area:

Number of hidden layers: 3
Number of neurons in each hidden layer: 7

Ramps:

Number of hidden layers: 3
Number of neurons in each hidden layer: 5

2.2.3. PSO Approach. Recently, Eberhart et al. proposed a PSO algorithm to be used for global optimization on the basis of random exploration methods and models of simple social systems. They stated that the algorithm is very efficient in solving the nonlinear problems. The PSO algorithm is established based on researches on different communities such as birds' community and can be used to optimize both nonlinear-continuous and discrete problems. In addition, the algorithm does not require plenty of time and memory for calculations because of its simple concept [17, 38, 39].

"Any information can be exchanged among the population" and *"the behavior of each particle is affected by the behavior of other particles in the community"* are the two hypotheses of the PSO algorithm which forms its basic concepts according to the results of a lot of studies on different communities. PSO has been developed in a two-dimensional (x-y) space, and the coordination of each particle is defined in this space. The vectors of the movement of particles are illustrated as V_x and V_y in each direction as velocity vectors. So, the movements of each particle are depicted by its coordination and vectors. Every particle in the community tries to optimize a specific target function. It knows its own best results and current coordination, and the particle is also informed about the best results of the community. Therefore, the movement vector of each particle can be achieved by [40]

$$v_i^{k+1} = w_i \times v_i^k + c_1 \times \text{rand} \times \left(p\text{best} - x_i^k\right) + c_2 \times \text{rand} \times \left(g\text{best} - x_i^k\right), \quad (27)$$

where v_i^k is the movement vector of particle i in iteration k, v_i^{k+1} is the modified movement vector of particle i, rand is a random digit between 0 and 1, x_i^k is the current location of particle i in iteration k, $p$best is the best result of particle i, $g$best is the best result of the community, w_i is the weight coefficient of the velocity vector of particle i, and c_i is the weight coefficient of each component. The location of each particle can be obtained by [40]:

$$x_i^{k+1} = x_i^k + v_i^{k+1}. \tag{28}$$

The concept of modifying the search point is depicted in Figure 9.

To ensure the convergence of the PSO, a contraction factor was applied to (27). Thus, the equation of the movement vector is modified as [40]

$$\begin{aligned} v_{i+1}^{k+1} = K \times \left(w \times v_i^k + c_1 \times \text{rand} \times \left(p\text{best} - x_i^k\right)\right. \\ \left. + c_2 \times \text{rand} \times \left(g\text{best} - x_i^k\right)\right), \end{aligned} \tag{29}$$

where K is the contraction factor and can be computed by [40]

$$K = \frac{2}{\left|2 - \varphi - \sqrt{\varphi^2 - 4\varphi}\right|}, \tag{30}$$

$$\begin{aligned} &\varphi = c_1 + c_2, \\ &\varphi > 4. \end{aligned} \tag{31}$$

All parameters were described previously. Choosing an appropriate w will lead to creating a balance between the general and local search. The coefficient w can be calculated by [40]

$$w = w_{\max} - \frac{w_{\max} - w_{\min}}{\text{iter}_{\max}} \times \text{iter}, \tag{32}$$

where iter is the number of iterations and $\text{iter}_{\max}$ is the maximum number of iterations. Based on the above concepts, the following steps were performed to develop the PSO-based safety estimator model in the four parts of freeway interchanges.

Step 1 (defining the general formulation). Equation (33) depicts the general formulation of the PSO algorithm.

$$\text{NCPI}_{i,\text{PSO}} = \begin{bmatrix} t_1 \\ g_1 \\ \vdots \\ t_n \\ g_n \end{bmatrix} \begin{bmatrix} a_{1,i} \\ . \\ \vdots \\ . \\ a_{n,i} \end{bmatrix}, \tag{33}$$

where $\text{NCPI}_{i,\text{PSO}}$ is the NCPI calculated by the PSO-based model in iteration i, t_1 to t_n are the traffic variables, g_1 to g_n are the geometry variables, and $a_{1,i}$ to $a_{n,i}$ are the constant parameters and coefficients of the PSO basic equation.

Step 2 (defining the input data). The geometry and traffic variables in the database were considered as the input data of the PSO algorithm.

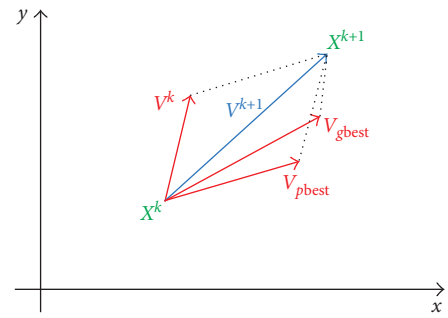

Figure 9: The concept of modifying the search point in PSO.

Step 3 (defining the target output data). The NCPIs (safety level) in the rows of information of the generated database were considered as the target output data.

Step 4 (proposing a basic equation with adequate constant parameters and coefficients). Some equations were proposed, and after a great deal of trial and error attempts, the best equations were chosen. Equations (34–40) are the selected equations. Constant parameters of the basic equations would be defined based on the generated database.

Weaving area:

$$\text{NCPI}_{\text{W}} = b_1 \sin\left(b_2\theta + b_3\right) + b_4 \sin\left(b_5\theta + b_6\right) + b_7, \tag{34}$$

$$\begin{aligned} \theta = 0.111\big[&a_1 L_{\text{W}}^{a_2} + a_3 L_{\text{W}}^{a_4} + a_5 L_{\text{W}}^{a_6} + a_7 N_{\text{W}} + a_8 N_{\text{R-ON}} \\ &+ a_9 N_{\text{R-OFF}} + a_{10} V_{\text{FW}}^{a_{11}} + a_{12} V_{\text{R-ON}}^{a_{13}} \\ &+ a_{14} S_{\text{FW}}^{a_{15}} + a_{16} S_{\text{R-ON}}^{a_{17}} + a_{18} S_{\text{R-OFF}} + a_{19}\big], \end{aligned} \tag{35}$$

where NCPI_{W} is the NCPI in the weaving segment, L_{W} is the length of the weaving area, N_{W} is the number of lanes in the weaving area, $N_{\text{R-ON}}$ is the number of on-ramp lanes, $N_{\text{R-OFF}}$ is the number of off-ramp lanes, V_{FW} is the freeway volume, $V_{\text{R-ON}}$ is the on-ramp volume, S_{FW} is the freeway free-flow speed, $S_{\text{R-ON}}$ is the speed of on-ramp, $S_{\text{R-OFF}}$ is the speed of off-ramp, and a_i and b_i are constant parameters.

Merge area:

$$\begin{aligned} \text{NCPI}_{\text{M}} = b_1 \times 0.143^{b_2} \big|&a_1 e^{a_2 L_{\text{ACC}}} + a_3 e^{a_4 L_{\text{ACC}}} + a_5 V_{\text{FW}}^{a_6} \\ &+ a_7 N_{\text{FW}} + a_8 V_{\text{R-ON}}^{a_9} + a_{10} N_{\text{R-ON}} + a_{11} e^{a_{12} S_{\text{FW}}} \\ &+ a_{13} e^{a_{14} S_{\text{FW}}} + a_{15} S_{\text{R-ON}}^{a_{16}} + a_{17}\big|^{b_2} + b_3, \end{aligned} \tag{36}$$

where NCPI_{M} is the NCPI in the merge area, L_{ACC} is the length of the acceleration lane, N_{FW} is the number of freeway lanes, $N_{\text{R-ON}}$ is the number of on-ramp lanes, V_{FW} is the freeway volume, $V_{\text{R-ON}}$ is the on-ramp volume, S_{FW} is the freeway free-flow speed, $S_{\text{R-ON}}$ is the speed of on-ramp, and a_i and b_i are constant parameters.

Diverge area:

$$\mathrm{NCPI_D} = b_1 \tan\left|b_2\omega + b_3\right| + b_4, \tag{37}$$

$$\omega = 0.167\left(a_1 L_{\mathrm{DEC}}^{a_2} + a_3 N_{\mathrm{FW}} + a_4 N_{\text{R-OFF}} + a_5 V_{\mathrm{FW}}^{a_6} + a_7 e^{a_8 S_{\mathrm{FW}}} + a_9 S_{\text{R-OFF}} + a_{10}\right), \tag{38}$$

where $\mathrm{NCPI_D}$ is the NCPI in the diverge area, L_{DEC} is the length of the deceleration lane, N_{FW} is the number of freeway lanes, $N_{\text{R-OFF}}$ is the number of off-ramp lanes, V_{FW} is the freeway volume, S_{FW} is the freeway free-flow speed, $S_{\text{R-OFF}}$ is the speed of off-ramp, and a_i and b_i are constant parameters.

Ramps:

$$\mathrm{NCPI_R} = a_1 L_{\mathrm{R,mod}}^2 + a_2 L_{\mathrm{R,mod}} + a_3 e^{a_4 N_{\mathrm{R}}} + a_5 V_{\mathrm{R}}^{0.5} + a_6\left(25.4\, R_{\mathrm{R}}\right) + a_7\left(25.4\, R_{\mathrm{R}}\right)^{0.5} + a_8, \tag{39}$$

$$L_{\mathrm{R,mod}} = L_{\mathrm{R}}\left(1 - \frac{S_{\mathrm{LONG}}}{100}\right), \tag{40}$$

where $\mathrm{NCPI_R}$ is the NCPI of the interchange ramp, L_{R} is the length of the ramp, $L_{\mathrm{R,mod}}$ is the modified length of the ramp, N_{R} is the number of lanes in the ramp, S_{LONG} is the average slope of the ramp, V_{R} is the ramp flow rate, R_{R} is the radius of the ramp, and a_i and b_i are constant parameters.

Step 5 (checking the difference between the NCPIs estimated by the PSO-based model and the NCPIs in the database). The results of the NCPI using the basic equations were compared to the NCPI in the database, and the RMSE was calculated by

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{k}\left(\mathrm{NCPI}_{i,\mathrm{PSO}} - \mathrm{NCPI}_{i,\mathrm{DB}}\right)^2}{k}}, \tag{41}$$

where $\mathrm{NCPI}_{i,\mathrm{DB}}$ is the NCPI in the rows of information of the generated database.

Step 6 (obtaining the global best result). After a lot of trial and error attempts and several iterations, the best results of constant parameters (as the global best results) of the basic equations were obtained when the possible minimum RMSE was reached.

2.2.4. Statistical Analysis. The models were evaluated by statistical analysis and field studies. The survey results were compared with the models' outputs. Regarding possible differences between the survey results and the models' outputs, it was necessary to know that these differences were because of either data distribution and their random properties or a significant diversity between the outcomes. Statistical analysis indicated whether there was a significant difference between the surveyed NCPIs and the corresponding NCPIs estimated by the models or not. A pooled t-test was used due to the limited number of samples. Statistical t could be calculated by [41]

$$t = \left(\mu_{\mathrm{m}} - \mu_{\mathrm{s}}\right) {S_p}^{-1}\left({n_{\mathrm{m}}}^{-1} + {n_{\mathrm{s}}}^{-1}\right)^{-0.5}, \tag{42}$$

$$S_p = \left(\left(n_{\mathrm{m}} - 1\right){\sigma_{\mathrm{m}}}^2 + \left(n_{\mathrm{s}} - 1\right){\sigma_{\mathrm{s}}}^2\right)^{0.5}\left(n_{\mathrm{m}} + n_{\mathrm{s}} - 2\right)^{-0.5}, \tag{43}$$

where μ_{m} and μ_{s} are the mean of the models' population and the mean of the survey population, respectively. n_{m} and σ_{m} are the number of samples and standard deviation of the models' results, and n_{s} and σ_{s} are the number of samples and standard deviation of survey results, respectively. Computed t should be compared with the tabulated values of the t-distribution table. The tabulated values of the t-distribution table depend on the degree of freedom, f, which represents the number of independent parts. The degree of freedom is defined by (44) in t-distributions [41]:

$$f = n_{\mathrm{m}} + n_{\mathrm{s}} - 2. \tag{44}$$

Once the statistical t is determined, the tabulated values of the t-distribution table yield the probability of a situation in which the t value is greater than the computed value. In order to limit the probability of a "type I" error to 0.05, the difference in the means will be considered significant only if the probability is less than or equal to 0.05. In other words, if the calculated t value falls within the 5% area of the tail, or in other words, if there is less than a five percent chance that such a difference could be found in the same population, the difference in the means will be considered significant. If the probability is greater than 5% (or the computed t value is less than the tabulated values of the t-distribution table), it could be concluded that such a difference in means could be found in the same population and the difference would not be considered significant [41].

3. Results and Discussion

Estimation of the safety level in the ramps, weaving, merge, and diverge segments of freeway interchanges was the main purpose of this research. A huge number of simulated parts of freeway interchanges and the ten existed freeway interchanges in Tehran Province in Iran were investigated, and by using the proposed fuzzy-based method, the NCPI was determined in every case. The models' required database was generated after the NCPI was calculated in all the cases. The database contained, respectively, nine, seven, six, and five traffic and geometric characteristics for weaving areas, merge areas, diverge areas, and ramps as inputs and the NCPIs as output. Therefore, two ANN-based and PSO-based models were developed utilizing the database.

3.1. ANN Results. The ANN-based development process was accomplished, and data training was stopped when the RMSE in the validation set began to increase, which signifies that the ANN generalization stopped increasing. Further analysis was conducted to test the accuracy of estimation by completely new input data to ensure that the reference model of the neural networks provides reliable results. To verify that the given estimation models are suitable to provide reliable results, a complete set of new data (which

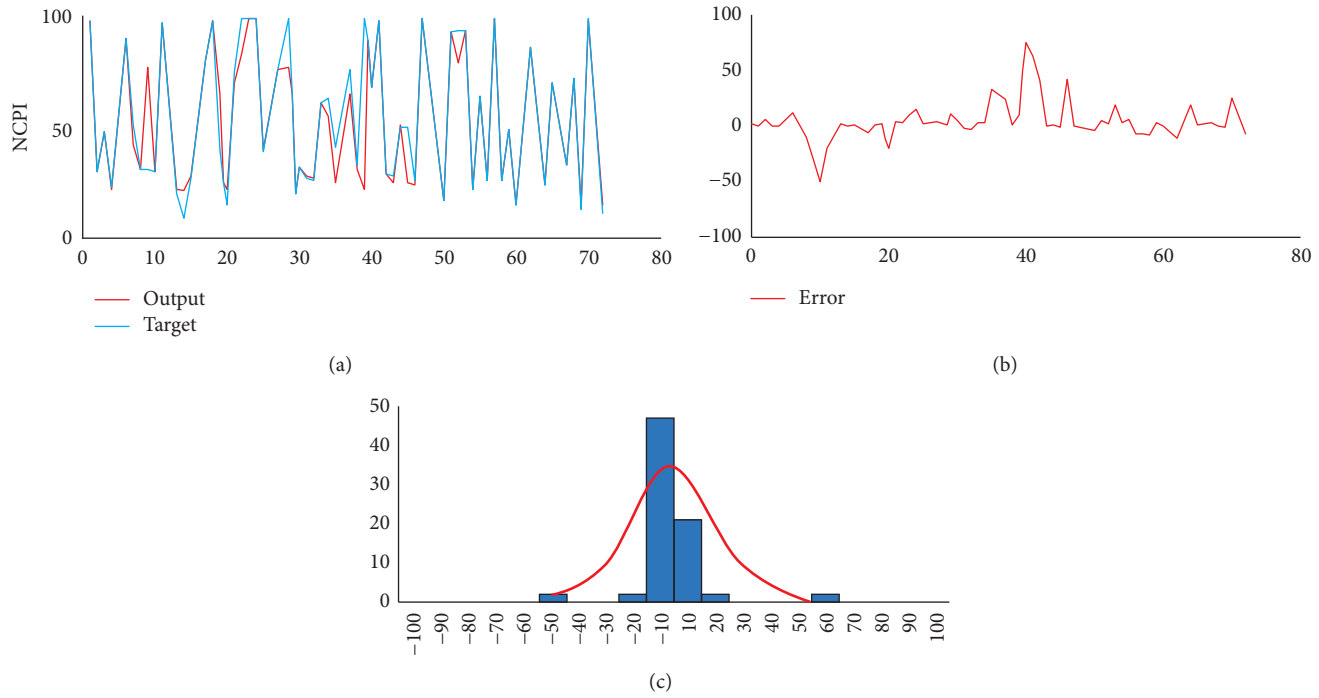

Figure 10: Ramps: (a) test data comparison of the ANN output and the NCPI in rows of information as target, (b) error diagram, and (c) error distribution.

were not used in the evaluation) was used. The results of the ANN-based model were categorized into three collections of all data, validation data, and test data as indicated in Figure 10 for interchange ramps, for instance. For each collection, ANN outputs were compared with the NCPI in the database (as the target), and the results were presented in the first graph. Error diagram and error distribution of the data of each collection were depicted in the second and third graphs, respectively.

The results of using the ANN to estimate the NCPI in the four parts of the interchange are illustrated in Figure 11, and the correlation between the estimated NCPI by the ANN-based model and the NCPI in the generated database in the interchange ramps was represented in Figure 12, for instance. According to the results, considering the fact that all neural network models showed similar small error means and high coefficients of correlation, the difference between the results of training and validation sets was negligible.

As shown in Figures 11 and 12, the analysis indicated a relative agreement between the NCPIs in the database and estimated data, according to all statistical performance measures. The standard deviation, coefficient of correlation, error mean, and RMSE of the three collections demonstrated a proper development of the ANN-based model. The low values of RMSE and error mean and the high value of coefficient of correlation are commonly recognized as a good estimation of the model.

But, comparing the statistical results of the four parts with each other, the highest coefficient of correlation and the minimum RMSE along with the smallest value of standard deviation were found in the diverge area. The lowest coefficient of correlation was detected in the merge area, and the high values of RMSE and standard deviation were obtained in interchange ramps and merge areas. However, the results of conducted analysis indicated that the proposed ANN-based model could generally be used to estimate the safety level of the four parts of the freeway interchange with sufficient accuracy, and the results clearly depicted that, for all the samples, values estimated with the ANN-based model are strongly consistent with the values obtained by analyzing the trajectory data. It is clear that the modeling results are exceptionally correct; therefore, there is no doubt regarding the accuracy of the estimation performance of the ANN-based model.

3.2. PSO Results. After a lot of trial and error attempts, the maximum number of iterations and the best population size (swarm size) were discovered in the PSO-based model development process. When the algorithm reached the minimum MSE as the best cost, the trial and error attempts were stopped. Therefore, the constant parameters of (34–40) were determined based on the results of iterations. The global best results and their position in the basic equations along with MSE and RMSE of the algorithm for the diverge area are presented in Figure 13, for instance, and a brief overview of applying the optimization process on all parts of interchanges is presented in Table 5.

Statistics		Weaving areas	Merge areas	Diverge areas	Ramps
R (Coefficient of correlation)		0.929	0.882	0.999	0.969
All data	RMSE	3.933	7.897	1.824	8.335
	Error mean	−0.049	−0.192	0.148	−0.038
	Standard deviation	3.933	7.897	1.820	8.347
Test data	RMSE	3.616	8.590	0.647	12.714
	Error mean	−0.186	0.436	0.065	1.553
	Standard deviation	3.612	8.588	0.646	12.707
Validation data	RMSE	4.019	7.592	1.215	8.427
	Error mean	−0.090	−0.113	0.149	−0.454
	Standard deviation	4.018	7.600	1.210	8.474

Figure 11: Statistical results of using the ANN-based model.

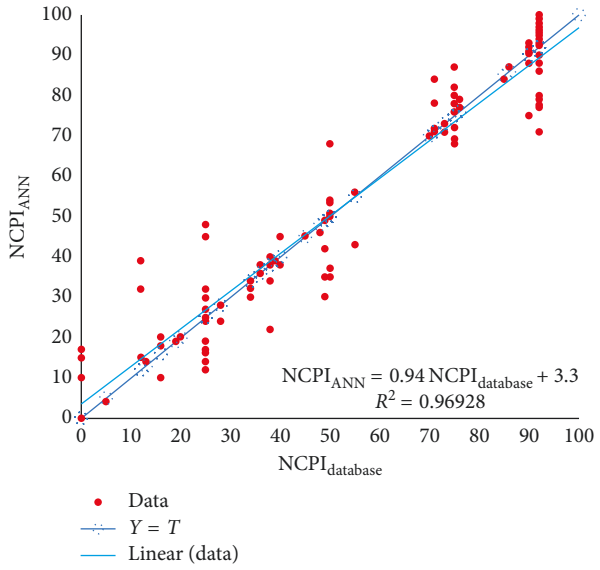

Figure 12: Correlation between the estimated safety level and the safety level from the database in the ramps.

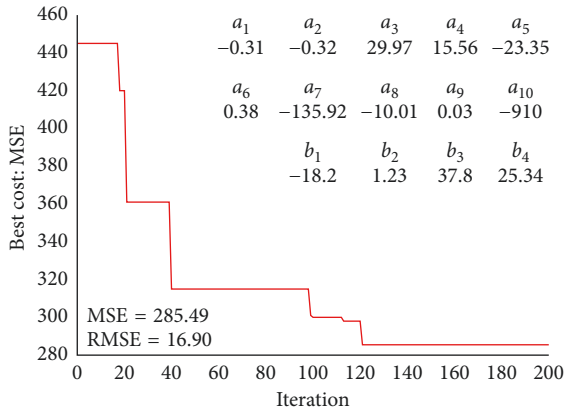

Figure 13: PSO iteration in the diverge area (best cost = MSE).

The values of RMSE indicated a relatively good development of the model. The values are different and vary from 7.66 for the weaving areas to 18.7 for the ramps. Irregular movement of the vehicles in curves may cause a higher value of RMSE in this area.

Although the RMSE values for the PSO-based model are more than those for ANN-based models, it is not enough to decline the application of the PSO-based model in estimating the safety level in freeway interchanges. Constant parameters were determined, and therefore, (34–40) were rewritten as (45–51).

Weaving area:

$$\mathrm{NCPI_W} = 16.48\sin(30\theta - 98.82) + 31.73, \tag{45}$$

$$\begin{aligned}\theta = {} & -0.21L_{\mathrm{W}}^{5.11} - 4.44L_{\mathrm{W}}^{10} + 4.44N_{\mathrm{W}} - 0.11N_{\text{R-ON}} \\ & + 4.08N_{\text{R-OFF}} + 6660000V_{\mathrm{FW}}^{-19.99} + 0.13V_{\text{R-ON}}^{-6.20} \\ & + 6.66S_{\mathrm{FW}}^{1.68} + 557.79S_{\text{R-ON}}^{-4.10} + 18.07.\end{aligned} \tag{46}$$

Merge area:

$$\begin{aligned}\mathrm{NCPI_M} = {} & 0.428 \times 0.143^{0.0816}\left|0.6e^{-5.6L_{\mathrm{ACC}}}\right. \\ & + 7V_{\mathrm{FW}}^{5.204} - 30N_{\mathrm{FW}} - 210V_{\text{R-ON}}^{-5.625} + 46.9N_{\text{R-ON}} \\ & + 355320000000000e^{-1.2S_{\mathrm{FW}}} + 69.564e^{0.302S_{\mathrm{FW}}} \\ & \left. + 1800000S_{\text{R-ON}}^{-14}\right|^{0.0816} + 2.891.\end{aligned} \tag{47}$$

Diverge area:

$$\mathrm{NCPI_D} = -18.2\tan|1.23\omega + 37.80| + 25.34, \tag{48}$$

$$\begin{aligned}\omega = 0.167\Big(& -0.31L_{\mathrm{DEC}}^{-0.32} + 29.97N_{\mathrm{FW}} \\ & + 15.56N_{\text{R-OFF}} - 23.35V_{\mathrm{FW}}^{0.38} \\ & - 135.92e^{-10.01S_{\mathrm{FW}}} + 0.03S_{\text{R-OFF}} - 910\Big).\end{aligned} \tag{49}$$

Table 5: A brief outlook of applying the optimization process on all parts of interchanges.

Part	Effective iterations	MSE	RMSE	Number of constant parameters
Weaving area	62	58.64	7.66	25
Merge area	21	184.245	13.57	20
Diverge area	121	285.49	16.90	14
Ramps	74	349.69	18.70	8

Table 6: Results of statistical analysis between means of the NCPIs estimated by the two safety estimator models versus the NCPIs achieved by field studies.

Part	Location	ANN-based model		PSO-based model		Field study		Statistical pooled t-test				
		μ_m	σ_m	μ_m	σ_m	μ_s	σ_s	$S_{p\text{-ANN}}$	$S_{p\text{-PSO}}$	t_{ANN}	t_{PSO}	t value (t-dist. table)
Waving area	Hemmat W-E: Yadegar ON and Yadegar OFF	43.34	2.23	47.9	5.90	44.88	5.53	4.22	5.72	0.93	1.35	1.711
	Hemmat Freeway W-E: Asharfi ON and Yadegar OFF	28.94	2.89	33.42	6.68	30.54	6.10	4.77	6.39	0.97	1.31	1.695
	Hakim Freeway E-W: Sheikh Bahaee ON and Chamran OFF	29.75	2.58	27.44	7.35	30.43	8.16	6.05	7.77	0.34	1.16	1.694
	Karbala Freeway E-W: Sahne ON and Hamedan OFF	21.87	1.07	19.69	2.76	21.19	2.97	2.23	2.86	1.00	1.69	1.684
	Imam Ali Freeway S-N: N-N U-turn ON and Sabalan OFF	38.02	1.08	37.31	6.64	39.28	6.99	5.00	6.82	0.78	0.89	1.692
	Imam Ali Freeway S-N: Khavaran ON and Khavaran OFF	37.38	4.74	38.8	10.67	38.03	10.46	8.12	10.57	0.23	0.21	1.695
Merge area	Hemmat Freeway W-E: Merge Asharfi N	16.3381	2.18	18.85	4.21	17.20	2.12	2.15	3.33	0.94	1.16	1.725
	Niayesh Freeway E-W: Merge Chamran S	14.7116	3.09	17.06	5.11	15.59	3.12	3.10	4.23	0.75	0.92	1.706
	Tehran-Qom Freeway N-S: Merge Vahnabad E	16.1305	4.48	14.41	5.30	15.73	4.21	4.35	4.79	0.33	0.97	1.687
	Tehran-Qom Freeway S-N: Merge Vahnabad W	15.2322	2.28	13.01	3.12	14.53	2.03	2.16	2.64	1.00	1.78	1.692
	Tehran-Saveh Freeway W-E: Merge Shahriar W	18.1533	3.41	15.17	4.22	17.24	3.07	3.25	3.69	0.85	1.68	1.694
Diverge area	Hakim Freeway W-E: Diverge Sheikh Bahaee S	1.6394	0.18	1.61	0.20	1.65	0.20	0.19	0.20	0.18	0.66	1.694
	Hemmat Freeway W-E: Diverge Yadegar S	32.335	6.01	36.52	7.30	33.83	6.76	6.39	7.03	0.60	0.98	1.711
	Tehran-Saveh Freeway E-W: Diverge Dehshade W	49.555	8.72	57.17	10.86	52.18	9.91	9.34	10.40	0.93	1.59	1.681
	Tehran-Saveh Freeway E-W: Diverge Robat Karim E	5.6069	1.83	4.23	1.44	5.39	1.83	1.83	1.65	0.32	1.86	1.706
	Yadegar Freeway N-S: Diverge Kouhestan E	89.363	14.66	99.00	17.63	92.65	16.50	15.61	17.07	0.45	0.79	1.746
Interchange ramps	Ramp: Chamran S to Hemmat W	30.45	2.62	31.65	3.90	30.88	3.80	3.26	3.85	0.47	0.73	1.676
	Ramp: Hemmat W to Chamran S	57.93	4.98	50.99	10.19	55.71	11.13	8.62	10.67	0.83	1.43	1.684
	Ramp: Imam Ali N to Mahallati W	63.02	5.42	68.99	13.11	65.10	12.37	9.55	12.74	0.62	0.86	1.697
	Ramp: Mahallati E to Imam Ali N	90.06	7.75	97.68	13.67	92.64	12.97	10.68	13.33	0.57	0.89	1.725
	Ramp: Niayesh E to Chamran S	75.77	6.52	69.17	12.32	73.65	13.11	10.35	12.72	0.63	1.09	1.689

Ramps:

$$\mathrm{NCPI_R} = -0.000073L_{\mathrm{R,mod}}^2 + 0.0005L_{\mathrm{R,mod}} + 2.022e^{-29.958N_\mathrm{R}} - 2.856V_\mathrm{R}^{0.5} - 0.838R_\mathrm{R} + 18.632R_\mathrm{R}^{0.5} + 66.16, \quad (50)$$

$$L_{\mathrm{R,mod}} = L_\mathrm{R}\left(1 - \frac{S_{\mathrm{LONG}}}{100}\right). \quad (51)$$

All parameters were described previously.

3.3. Field Studies' Outcomes. As it was mentioned before, field studies were used to evaluate the proposed method and the models. Six weaving areas, five merge areas, five diverge areas, and five ramps were surveyed through the field study on ten freeway interchanges of Tehran Province in Iran. The values of the NCPI were computed using trajectory data analysis and estimated by applying ANN-based and PSO-based models on the interchanges with characteristics brought in Table 2. The results of statistical analysis are illustrated in Table 6. It could be said that no significant differences between the means of the models' population and the real population were found when almost all computed statistical t-values (which are achieved from statistical analysis of the population of the models and studied areas) are less than the tabulated values of the t-distribution table.

4. Conclusion

In this paper, it was intended to propose a new method to have an exclusive safety indicator among different SSMs at the first step and to develop a model to estimate the safety level according to geometrical and traffic characteristics of different parts of freeway interchanges at the second step. Different surrogate measures were combined using fuzzy logic, and an index called NCPI was defined as a safety level indicator. The variables of NCPI including outputs of four surrogate measures of safety were determined by analyzing the trajectory data. The trajectory data could be either achieved from video processing or derived from microsimulation. Then, NCPI was obtained by applying fuzzy rules to the variables. It was done to estimate the level of safety when there are not enough data or information about the number and severity of accidents or the segment is just being designed and has not yet been built. Due to the difficulties of accessing or obtaining the trajectory data, two models were developed by ANN and PSO algorithms to estimate the safety level based on geometrical and traffic characteristics of interchanges. At last, field studies were carried out to calibrate the simulations, controlling the validity of the proposed fuzzy method and checking the accuracy of safety estimator models.

The results indicated an acceptable confidence about the validity of the proposed fuzzy method. The results also showed a good accuracy of the developed models in terms of compliance with the database generated from data analysis and also surveyed data of the field studies. It also became clear that, in most cases, the results of the ANN-based model have more accuracy than the results of the PSO-based model. However, the advantage of using the PSO-based model is that finally, there will be a certain relationship which can be conveniently used, while the ANN-based model has some prerequisites to work, say, database knowledge and MATLAB expertise.

The developed models can also be used for comparing different proposed plans of interchanges from safety aspects and ranking them in the design steps and plan selection process. In these cases, the lower accuracy of some PSO-based models compared with the ANN-based models does not matter.

In general, the proposed models will be valid when the geometric and traffic characteristics of interchange's parts fall within the range of variables' values used for model development. Definitely, the more the difference between the values of these characteristics and the range of mentioned variable values, the more the reduction in the validity of the models. Another conclusion is that the models could be trained to estimate other traffic parameters such as the density, delay, and speed of interchange's parts or other traffic facilities based on their traffic and geometric characteristics.

In this paper, it was proposed to use fuzzy logic and the algorithms of ANN and PSO to estimate the safety level of different parts of freeway interchanges. But, there will be a long way to reach a point that these methods become general and be used in every situation. In this way, we propose that these methods can be applied in a wider range of variables' values and also be used in other segments of the freeway which include traffic conflicts, merging, or diverging.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] L. Wang, M. Abdel-Aty, Q. Shi, and J. Park, "Real-time crash prediction for expressway weaving segments," *Transportation Research Part C: Emerging Technologies*, vol. 61, pp. 1–10, 2015.

[2] J. Zhang, X. Sun, and Y. He, "Safety prediction models of freeway in plain terrain," in *Proceedings of the International Consortium of Critical Theory Programs (ICCTP'10)*, Beijing, China, August 2010.

[3] M. Sarhan, Y. Hassan, and A. O. Abd El Halim, "Safety performance of freeway sections and relation to length of speed-change lanes," *Canadian Journal of Civil Engineering*, vol. 35, no. 5, pp. 531–541, 2008.

[4] J. Xie, Y. Ma, L. Yuan, P. Zhang, and Y. Liu, "Safety and capacity performances of single-lane right exit ramp on freeway: a case study in Jiangsu Province, China," *Procedia Engineering*, vol. 137, pp. 563–570, 2016.

[5] D. Eustace, A. Aylo, and W. Y. Mergia, "Crash frequency analysis of left-side merging and diverging areas on urban freeway segments–a case study of I-75 through downtown Dayton, Ohio," *Transportation Research Part C: Emerging Technologies*, vol. 50, pp. 78–85, 2015.

[6] H. Chen, H. Zhou, J. Zhao, and P. Hsu, "Safety performance evaluation of left-side off-ramps at freeway diverge areas,"

Accident Analysis & Prevention, vol. 43, no. 3, pp. 605–612, 2011.

[7] W. Y. Mergia, D. Eustace, D. Chimba, and M. Qumsiyeh, "Exploring factors contributing to injury severity at freeway merging and diverging locations in Ohio," *Accident Analysis & Prevention*, vol. 55, pp. 202–210, 2013.

[8] H. Behbahani, N. Nadimi, and S. S. Naseralavi, "New time-based surrogate safety measure to assess crash risk in car-following scenarios," *Transportation Letters*, vol. 7, no. 4, pp. 229–238, 2015.

[9] H. Behbahani and N. Nadimi, "A framework for applying surrogate safety measures for sideswipe conflicts," *International Journal for Traffic & Transport Engineering*, vol. 5, no. 4, pp. 371–383, 2015.

[10] D. Gettman and L. Head, "Surrogate safety measures from traffic simulation models," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1840, pp. 104–115, 2003.

[11] J. Archer, *Indicators for Traffic Safety Assessment and Prediction and their Application in Micro-Simulation Modelling: A Study of Urban and Suburban Intersections*, Royal Institute of Technology, Stockholm, Sweden, 2005.

[12] K. Ozbay, H. Yang, B. Bartin, and S. Mudigonda, "Derivation and validation of new simulation-based surrogate safety measure," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2083, pp. 105–113, 2008.

[13] L.-Y. Chang, "Analysis of freeway accident frequencies: negative binomial regression versus artificial neural network," *Safety Science*, vol. 43, no. 8, pp. 541–557, 2005.

[14] P. Angeline, "Evolutionary optimization versus particle swarm optimization: philosophy and performance differences," in *Evolutionary Programming VII*, Springer, Berlin, Germany, 1998.

[15] J. Kennedy, "The particle swarm: social adaptation of knowledge," in *Proceedings of the IEEE International Conference on Evolutionary Computation*, Indianapolis, IN, USA, April 1997.

[16] Y. Shi and R. Eberhart, "A modified particle swarm optimizer," in *Proceedings of the IEEE World Congress on Computational Intelligence Evolutionary Computation Proceedings*, Anchorage, AK, USA, May 1998.

[17] Y. Shi and R. Eberhart, "Parameter selection in particle swarm optimization," in *Evolutionary Programming VII*, Springer, Berlin, Germany, 1998.

[18] J. Hayward, *Near Misses as a Measure of Safety at Urban Intersections*, Department of Civil Engineering, The Pennsylvania State University, University Park, PA, USA, 1971.

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

[20] M. M. Minderhoud and P. H. L. Bovy, "Extended time-to-collision measures for road traffic safety assessment," *Accident Analysis and Prevention*, vol. 33, no. 1, pp. 89–97, 2001.

[21] A. Sobhani, W. Young, D. Logan, and S. Bahrololoom, "A kinetic energy model of two-vehicle crash injury severity," *Accident Analysis and Prevention*, vol. 43, no. 3, pp. 741–754, 2011.

[22] D. Gettman and L. Head, *Surrogate Safety Measures from Traffic Simulation Models*, Federal Highway Administration, Washington, DC, USA, 2003.

[23] C. Hydén, *The Development of a Method for Traffic Safety Evaluation: The Swedish Traffic Conflicts Technique*, Bulletin Lund Institute of Technology, Lund, Sweden, 1987.

[24] J. C. Hayward and Pennsylvania Transportation and Traffic Safety Center, *Near Miss Determination through Use of a Scale of Danger*, The Pennsylvania State University, University Park, PA, USA, 1972.

[25] D. Gettman, L. Pu, T. Sayed, and S. Shelby, *Surrogate Safety Assessment Model and Validation*, Research, Development, and Technology Turner-Fairbank Highway Research Center, McLean, VA, USA, 2008.

[26] A. K. Maurya and P. S. Bokare, "Study of deceleration behaviour of different vehicle types," *International Journal for Traffic and Transport Engineering*, vol. 2, no. 3, pp. 253–270, 2012.

[27] American Association of State Highway and Transportation Officials, *A Policy on Geometric Design of Highways and Streets*, AASHTO, Washington, DC, USA, 2011.

[28] R. P. Roess, E. S. Prassas, and W. R. McShane, *Traffic Engineering*, Pearson Prentice Hall, Upper Saddle River, NJ, USA, 2011.

[29] S. Haykin and R. Lippmann, "Neural networks, a comprehensive foundation," *International Journal of Neural Systems*, vol. 5, no. 4, pp. 363-364, 1994.

[30] R. Lippmann, "An introduction to computing with neural nets," *IEEE Assp Magazine*, vol. 4, no. 2, pp. 4–22, 1987.

[31] M. Bilgehan and P. Turgut, "The use of neural networks in concrete compressive strength estimation," *Computers & concrete*, vol. 7, no. 3, pp. 271–283, 2010.

[32] M. A. Kewalramani and R. Gupta, "Concrete compressive strength prediction using ultrasonic pulse velocity through artificial neural networks," *Automation in Construction*, vol. 15, no. 3, pp. 374–379, 2006.

[33] K. G. Sheela and S. N. Deepa, "Review on methods to fix number of hidden neurons in neural networks," *Mathematical Problems in Engineering*, vol. 2013, pp. 1–11, 2013.

[34] C. G. Looney, "Advances in feedforward neural networks: demystifying knowledge acquiring black boxes," *IEEE Transactions on Knowledge and Data Engineering*, vol. 8, no. 2, pp. 211–226, 1996.

[35] K. Swingler, *Applying Neural Networks: A Practical Guide*, Morgan Kaufmann Publishers, Burlington, MA, USA, 1996.

[36] N. M. McCord and W. Illingworth, *A Practical Guide to Neural Nets*, Addison-Wesley, Boston, MA, USA, 1990.

[37] M. H. Beale, M. T. Hagan, and H. B. Demuth, *Neural Network Toolbox™ User's Guide*, The MathWorks, Inc., Natick, MA, USA, 2012.

[38] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in *Proceedings of the IEEE International Conference on Neural Networks*, Piscataway, NJ, USA, November–December 1995.

[39] J. Kennedy, *Particle Swarm Optimization, in Encyclopedia of Machine Learning*, Springer, Berlin, Germany, 2011.

[40] B. Zhao, C. Guo, and Y. Cao, "A multiagent-based particle swarm optimization approach for optimal reactive power dispatch," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 1070–1078, 2005.

[41] C. Spiegelman, E. S. Park, and L. R. Rilett, *Transportation Statistics and Microsimulation*, Taylor & Francis, Abingdon, UK, 2010.