

Adaptive neuro-fuzzy approach for modeling equilibrium speed–density relationship

Md. Hadiuzzaman, Mohammad Rayeedul Kalam Siam, Nazmul Haque, Tahmida Hossain Shimu & Fahmida Rahman

To cite this article: Md. Hadiuzzaman, Mohammad Rayeedul Kalam Siam, Nazmul Haque, Tahmida Hossain Shimu & Fahmida Rahman (2018): Adaptive neuro-fuzzy approach for modeling equilibrium speed–density relationship, Transportmetrica A: Transport Science, DOI: [10.1080/23249935.2018.1435589](https://doi.org/10.1080/23249935.2018.1435589)

To link to this article: <https://doi.org/10.1080/23249935.2018.1435589>



Accepted author version posted online: 31 Jan 2018.
Published online: 15 Feb 2018.



Submit your article to this journal [↗](#)



Article views: 7





View related articles [↗](#)



View Crossmark data [↗](#)



Adaptive neuro-fuzzy approach for modeling equilibrium speed–density relationship

Md. Hadiuzzaman ^a, Mohammad Rayeedul Kalam Siam^a, Nazmul Haque^a, Tahmida Hossain Shimu ^b and Fahmida Rahman^a

^aDepartment of Civil Engineering, Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh; ^bDepartment of Civil Engineering, Military Institute of Science and Technology (MIST), Dhaka, Bangladesh

ABSTRACT

This paper endeavors to model the equilibrium speed–density relationship using a new fuzzy logic-based approach. To capture the randomness in traffic flow dynamics, it develops a single-input Adaptive Neuro-Fuzzy Inference System (ANFIS) trained with a hybrid algorithm. Furthermore, it proposes a new Premise–Consequent Conjugate Effect (PCCE) relationship to estimate fundamental diagram (FD) parameters from the ANFIS model. Several options (i.e. optimal split of data into training and testing data, selection of suitable membership function) offered by ANFIS are then illustrated. An experiment is performed to determine the maximum achievable goodness of fit without the occurrence of overfitting phenomenon by changing the number of clusters. The calibrated ANFIS model is compared against traditional and advanced speed–density models for five different freeway locations. Results show that the proposed model outperforms other preceding models by attaining goodness-of-fit values within a range of 0.82–0.92. Finally, the proposed PCCE relationship shows the ANFIS model's robustness in accurate estimation of FD parameters for all freeway locations.

ARTICLE HISTORY



Received 19 May 2017
Accepted 29 January 2018

KEYWORDS

Adaptive Neuro-Fuzzy Inference System; hybrid algorithm; overfitting; fundamental diagram; equilibrium flow

1. Introduction

The fundamental diagram (FD), that is bivariate equilibrium relationships of traffic flow (q), speed (v), and density (ρ), is of great importance in analysis of traffic flow and modeling of traffic behavior. For example, the concept of level of service for a highway relies on input from the speed–flow relationship (TRB 2010). The flow–density relationship is central in many investigations associated with traffic flow dynamics (Daganzo 1995). In high-order traffic flow models (Messmer and Papageorgiou 1990; Kotsialos et al. 2002), the speed–density relationship plays an important role. However, among the three 'pair-wise' relationships (e.g. speed–density, flow–density, and speed–flow), the speed–density relationship appears to be challenging as it encapsulates a direct connection to everyday driving behavior. Traffic flow in the real world is characterized by heterogeneous driving

CONTACT Md. Hadiuzzaman  mhadiuzzaman@ce.buet.ac.bd  Department of Civil Engineering, Bangladesh University of Engineering and Technology (BUET), Dhaka 1000, Bangladesh

experience, i.e. perturbation in speed choice influenced by the surrounding vehicles. Therefore, a speed–density model that better captures these effects is demanded for traffic flow analysis and efficient traffic control.

Modeling of the speed–density relationship began with the Greenshields' linear model in the seminal paper: *A Study in Highway Capacity* (Greenshields 1935). This seminal work has inspired a slew of new models (Greenberg 1959; Underwood 1961; Pipes 1967; Drew 1968) which propose a variety of mathematical models containing several parameters. The values of these parameters are estimated by fitting these models to empirical traffic data. A comprehensive explanation of the parameters in terms of traffic flow dynamics is then developed to endow the mathematical model with a phenomenological interpretation. However, these models are limited in the sense that they try to fit the traffic dynamics to a definite shape (i.e. logarithmic, exponential, and exponential to the quadratic and various forms of polynomials). Nevertheless, a live transportation system is dictated by traffic streams with dynamical randomness effects, which result from heterogeneity in both vehicle attributes and driver preferences. For example, in a multimodal system, passenger cars, trucks, and public transportation travel in the same right-of-way. The difference in vehicle dynamics necessarily endows them with heterogeneous gap maintaining behavior, lateral movement and longitudinal motion, and lanes to travel on. As another example, drivers demonstrate heterogeneous driving behavior, e.g. minimum headway distances to keep when following leading vehicles, frequency in applying brakes, fluctuation in arrival pattern, etc. These lead to a complex traffic dynamics, which cannot be replicated by conventional speed–density models. To address this phenomenon, the proposed model must have sufficient shape flexibility to be fitted accurately to the measured field data (MacNicholas 2008).

A number of studies suggest that changes in behavior of the leading vehicle cause fluctuations in the reaction of drivers; this phenomenon causes traffic flow instability (Chandler, Herman, and Montroll 1958; Kometani and Sasaki 1961). At low densities, vehicle interactions are occasional and the drivers travel at their most comfortable speed. This leads to a stable traffic flow and, therefore, there are no substantial speed drops. This type of traffic flow pattern corresponds to a density range from 0 to 32 vehicle/mile/lane (del Castillo 2001). Near critical density, the interaction between drivers becomes pivotal to maintain stability of traffic flow. In particular, it has been observed that drivers vary their choice in speed and spacing selection somewhat randomly between two branches in the flow–density FD. It is assumed that drivers continuously vary their spacing and seek for lane-changing opportunities (Kerner 2004). Thus the propagation of a speed perturbation is not fully predictable because of the random nature of driving. Empirical observation from I-80 detectors in Berkeley, CA (Figure 1(a)), shows that traditional speed–density models perform poorly to capture random driver behavior influenced by the shockwave in the transition zone. The transition zone is also associated with curvature change in the speed–density curve. In addition, a number of researchers (Newman 1963; Mita, Kreer, and Yuan 1969) demonstrate that traffic flow dynamics is influenced by different geometric patterns of the road (i.e. on-ramp/off-ramp on a freeway). Figure 1(b) represents the aforementioned phenomena showing that mainline traffic maintains its stability for a larger density value.

However, for other freeway locations, there has been a variation on the origin of the traffic flow instability due to the change in geometrical pattern of the road and traffic

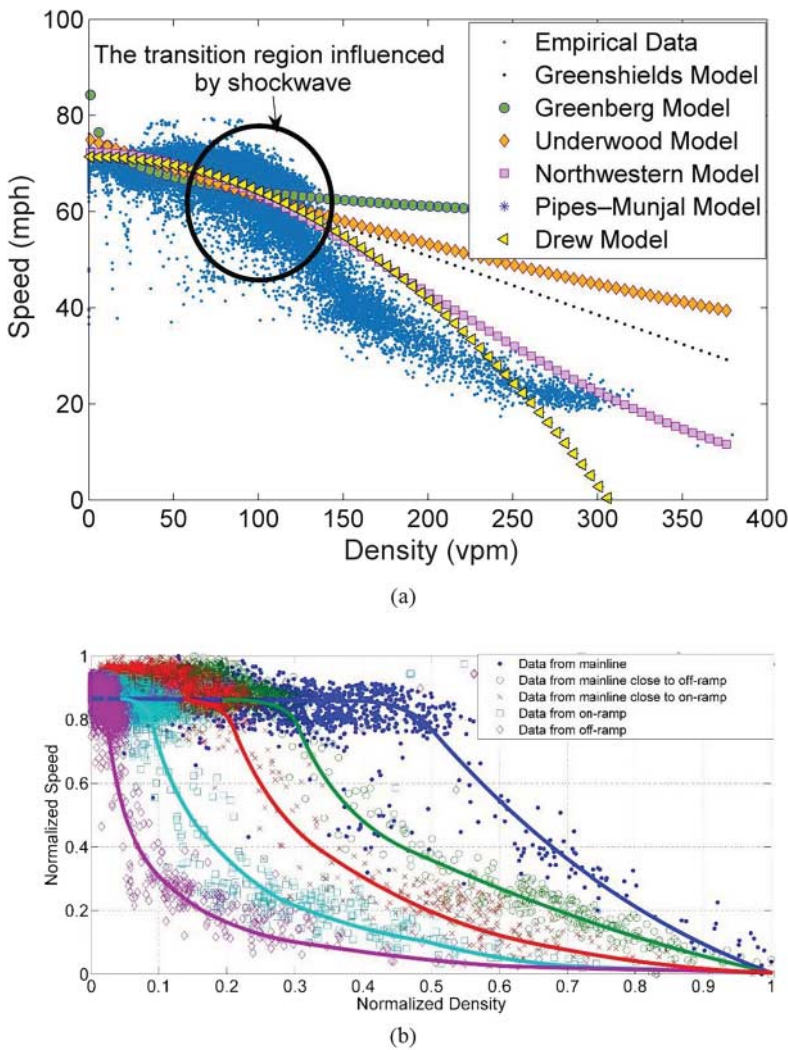


Figure 1. (a) Traditional speed–density models perform poorly to capture random driver behavior influenced by the shockwave in the transition zone and (b) variation in the origin of traffic instability for different locations.

operation. A developed speed–density relationship must acknowledge the effects of: (a) random driver behavior in the transition zone, and (b) different geometrical and traffic flow patterns of basic and nonbasic roadway segments on the shape of FD.

Traffic flow dynamics is characterized by uncertainty, subjectivity, fuzziness, and ambiguity. Human operators, dispatchers, drivers, and passengers take decisions based on subjective knowledge or linguistic information (Teodorovic and Vukadinovic 2012). Human decisions about lane change, gap acceptance, and acceleration and deceleration process affect the equilibrium speed–density relationship. The equilibrium traffic state assumes that for a given traffic density, the mean speed converges (with a lag time) to the FD. Thus, on the FD, all the drivers have the same desired speed, and the corresponding point

will not move unless the density on that roadway changes. However, humans make decisions based on complex reasoning and it becomes challenging to formulate an appropriate model for it, making the application of fuzzy logic systems justified in developing the speed–density relationship. The application of neural network based learning for various transportation problems is not fresh at all. Teodorovic and Vukadinovic (2012) described a number of potential applications of fuzzy logic and neural networks in transportation problems. Andrade, Uchida, and Kagaya (2006) developed an adaptive network based fuzzy logic transportation choice model. Pahlavani and Delavar (2014) described the application of a locally linear neuro-fuzzy system to perform multicriteria route planning based on drivers' preferences in multicriteria route selection.

In this paper, we have demonstrated that traditional speed–density models perform poorly in capturing the randomness in driver behavior as well as the effects of different freeway locations (on-ramp, off-ramp) on the shape of FD. To address the problem, we have developed a new calibration approach, integrating the abilities of ANNs self-learning with the linguistic expression function of fuzzy inference, that can replicate speed–density relation over all possible traffic states, i.e. light-traffic/free-flow conditions, transition region as well as congested/jam conditions. We have also developed a new premise-consequent relationship to extract FD parameters. A multitude of options (i.e. optimal split of data into a training sample and a testing sample, selection of suitable membership function (MF) offered by ANFIS) are then illustrated using I-80 dataset near Berkeley. The calibrated ANFIS model is compared against traditional as well as advanced speed–density models for five different highway locations.

The remaining sections of this paper are outlined as follows. A literature survey on speed–density models is provided in Section 2. This is followed by the proposed ANFIS-based speed–density model and its parameters and properties in Section 3. Calibration and validation of the proposed model is conducted using field observations from I-80 in California in Section 4. Finally, Section 5 concludes with some remarks and future directions.

2. Literature review

Considering the focus of this study, the extent of the literature review has been confined within different traditional and advanced speed–density models proposed by various researchers. Particularly, a brief discussion on these models including their characteristics and mathematical formulation has been presented in this section.

2.1. Traditional speed–density models

It has been almost 82 years since Greenshields first rigorously documented a linear relationship between speed and density (Greenshields 1935). Consequently, corresponding flow–density relationship considers a parabolic transition around capacity condition with overestimated optimal density and underestimated optimal speed. Nevertheless, this seminal work led to a new era of transportation science and engineering. There has been a number of follow-up studies devoted to reviewing or improving such a simplified relationship. These studies include single regime models: Greenberg's Model (Greenberg 1959), the Underwood Model (Underwood 1961), Northwestern (Drake, Schofer, and May 1967; Drew 1968), and the Pipes–Munjal Generalized Model (Pipes 1967). There are also multi-regime

Table 1. Single regime speed–density models.

Single regime models	Functions	Parameters
Greenshields (1935)	$v = v_f \left(1 - \frac{\rho}{\rho_j}\right)$	v_f, ρ_j
Greenberg (1959)	$v = v_m \log \frac{\rho_j}{\rho}$	v_m, ρ_j
Underwood (1961)	$v = v_f \exp \left(-\left(\frac{\rho}{\rho_o}\right)\right)$	v_f, ρ_o
Northwestern (Drake, Schofer, and May 1967)	$v = v_f \exp \left(-\frac{1}{2} \left(\frac{\rho}{\rho_o}\right)^2\right)$	v_f, ρ_o
Drew (1968)	$v = v_f \left[1 - \left(\frac{\rho}{\rho_j}\right)^{n+\frac{1}{2}}\right]$	v_f, ρ_j
Drake, Schofer, and May (1967)	$v = v_f \exp \left[\frac{1}{2} \left(\frac{\rho}{\rho_j}\right)^2\right]$	v_f, ρ_j
Pipes (1967)	$v = v_f \left(1 - \left(\frac{\rho}{\rho_j}\right)^n\right)$	v_f, ρ_j

Table 2. Multi-regime speed–density models.

Multi-regime models	Free-flow regime	Transitional-flow regime	Congested-flow regime
Eddie (1961)	$v = 54.9 \exp \left(-\frac{\rho}{163.9}\right) (\rho \leq 50)$	–	$v = 26.8 \ln \left(\frac{162.5}{\rho}\right) (\rho \geq 50)$
Two-regime model (May 1990)	$v = 60.9 - 0.51\rho (\rho \leq 65)$	–	$v = 40 - 0.265\rho (\rho \geq 65)$
Modified Greenberg (May 1990)	$v = 48 (\rho \leq 35)$	–	$v = 32 \ln \left(\frac{145.5}{\rho}\right) (\rho \geq 35)$
Three-regime linear model (May 1990)	$v = 50 - .098\rho (\rho \leq 40)$	$v = 81.4 - 0.913\rho (40 \leq \rho \leq 65)$	$v = 40.0 - 0.26\rho (\rho \geq 65)$

models which include two-regime models such as Eddie Model (Eddie 1961), multi-regime model by cluster analysis (Sun and Zhou 2005), two-regime linear model (May 1990), modified Greenberg, and three-regime models (Drake, Schofer, and May 1967; May 1990). Table 1 and Table 2 list most of the well-known speed–density models.

The fact that traditional single- and multi-regime speed–density models match empirical observations poorly has been well documented in a number of studies (May 1990; Hall, Hurdle, and Banks 1992; Qu, Wang, and Zhang 2015). For example, Greenshields (1935) collected only seven data points from one lane in a two-way rural road in which six of the data points were under 60 mile/hour and the seventh data point was taken from a different road (Hall, Hurdle, and Banks 1992). Seven data points are unlikely to generate a whole picture of a speed–density relationship. Greenberg (1959) bridged the gap between macroscopic models and the third General Motor car-following model (Gazis, Herman, and Potts 1959; Gazis, Herman, and Rothery 1961). However, this model fails to predict speed at lower densities, because as density approaches zero, speed tends to increase to infinity. In addition, optimum speed and jam density are difficult to observe from empirical data (May 1990). Later, Underwood (1961) proposed an exponential model that tried to overcome the drawback of the Greenberg model. The main limitation of the Underwood

model is that speed becomes zero only when density reaches infinity and so this model cannot be used for predicting speeds at high densities (May 1990). This deficiency also appears in the Northwestern model (May 1990) coupled with its assumption of a concave flow–density relationship for the congested regime that is not attractive to dynamic traffic analysis. The Drew model (Drew 1968) and Pipe–Munjal model (Pipes 1967) are the varied forms of Greenshields’ model and so they inherited the shortcomings of the origin model, both severely predicting underestimated optimal speed.

The inability of single regime models to fit the empirical data consistently well both in free-flow regime and congested regime paved the way for the development of multi-regime models in spite of the fact that this approach lacks mathematical elegance. Multi-regime models usually include two or three regimes, for example two-regime models such as the Edie Model (Edie 1961), two-regime model (May 1990), multi-regime model by cluster analysis (Sun and Zhou 2005), modified Greenberg, and three-regime models (Drake, Schofer, and May 1967; May 1990). The two-regime model is developed on the notion that two different curves are used to model the free-flow regime and the congested-flow regime separately. Edie (1961) first used this phenomenon and therefore proposed a discontinuous exponential form using the Underwood model for the free-flow regime and the Greenberg model for the congested-flow regime. In addition, there is a three-regime model that applied three linear curves to model the free-flow, transitional-flow and congested-flow regimes; each of the regimes is characterized by a Greenshields model (May 1990). Table 2 lists most of the well-known multi-regime speed–density models (Drake, Schofer, and May 1967; May 1990). The main drawback of the above multi-regime models is that they cannot determine breakpoints in a precise way. To address this problem, Barua et al. (2017) adopted the likelihood maximization approach for identifying multi-regime models’ breakpoint to calibrate various FDs under different road geometric and traffic operational conditions. However, multi-regime models still lack mathematical elegance.

2.2. Advanced speed–density models

Mahnke and Kaupužs (1999) demonstrated that a stochastic relationship could better characterize this fundamental relationship in consideration of the stochastic nature of scattered samples. In view of this phenomenon, a few pioneering attempts have been made to capture this stochastic property. Muralidharan, Dervisoglu, and Horowitz (2011) proposed a probabilistic graphical approach to account for the probabilistic nature of FD parameters based on the triangular deterministic model. Fan and Seibold (2013) introduced a new varying parameter, the empty-road velocity, to capture the randomness in the Aw–Rasclé–Zhang (ARZ) model. Jabari, Zheng, and Liu (2014) explored the derivation of probabilistic stationary speed–density relation based on Newell’s simplified microscopic car-following model. Muralidharan, Dervisoglu, and Horowitz (2011), Fan and Seibold (2013), and Jabari, Zheng, and Liu (2014) are among the few researchers that have developed models to generate the distributions of speed or flow as a function of density and percentile based fundamental (flow/speed–density) relations are attainable (Qu, Zhang, and Wang 2017). However, the fact that these models are developed based on a specific traffic flow model limits their applicability to deal with the stochasticity of traffic flow using other traffic flow models.

Wang et al. (2009) have attempted to overcome the well-known shortcomings of traditional models by developing a stochastic speed–density relationship model. It considers second-order statistics (i.e. mean and variance) to establish a stochastic speed–density relationship. Moreover, an exponential correlation between different densities within the empirical observations has also been considered to incorporate the dependency of driver collective behavior. Later, Wang et al. (2011) proposed a five-parameter logistic speed–density model (5PL) in consideration of the fact that the empirical speed–density observations exhibit a reversed ‘S’ shape. The proposed logistic model did match the empirical speed–density observations well and also included a parameter k_t which aids in capturing the inflection point, the point where there is no curvature or where the curve shifts from being concave to convex. These attractive features make it possible for the 5PL model to overcome almost all the drawbacks traditional speed–density models suffer from. However, the model assumes that even if the density approaches infinity, there exists a non-zero speed ($\rho \rightarrow \infty, V \rightarrow v_b$); this makes it difficult to obtain one of the important FD parameters, jam density. Later, Qu, Wang, and Zhang (2015) demonstrated that the inaccuracy of the well-known three-parameter logistic (3PL) model (Wang et al. 2011) can be contributed not only to their functional forms but also by the biasness in sample selection. Afterwards, a weighted least square method (WLSM) was applied that addressed this biasness. Besides, Ni et al. (2015) proposed the longitudinal control model (LCM) considering the perspective of physics and human factors. Particularly, it is a single regime model with a large set of parameters that is applied to the entire range of speed–density. Ultimately, the research concluded that the model flexibility and goodness-of-fit increases with an increase in the number of model parameters.

3. Fuzzy logic-based equilibrium speed–density relationship

The Adaptive Neuro-Fuzzy Inference System (ANFIS) was first originally developed by Jang in 1993 (Jang 1993). ANFIS is a functional data processing technique that applies fuzzy logic to map the input variables to output dependents (Jang 1993; Jang and Gulley 1994; Jang and Sun 1995). As a neuro-fuzzy system is based on linguistic rules, it offers the advantage of integrating prior knowledge into the system and this can substantially shorten the learning process. Fuzzy logic provides a simple way to arrive at a definite conclusion even if the input information is vague, ambiguous, imprecise, noisy, or missing (Kaehler 1998). Typically, fuzzy systems rely on a set of IF–THEN rules paired with MFs modified and adjusted to emulate the trend of the dataset through connected nodes that have different parametric functions. Afterwards, the learning algorithms are implemented to bring in the adjustment of the MF parameters to capture the attributes of the input data and map output accordingly. These IF–THEN rules cannot share the same output MF. The number of MFs must be equal to the number of rules. To present the single-input ANFIS architecture, fuzzy IF–THEN rules based on a first-order Sugeno model are considered:

Rule₍₁₎: **IF** ρ is A_1 , **THEN**

$$f_1 = p_1 \rho + r_1$$

Rule₍₂₎: **IF** ρ is A_2 , **THEN**

$$f_2 = p_2 \rho + r_2$$

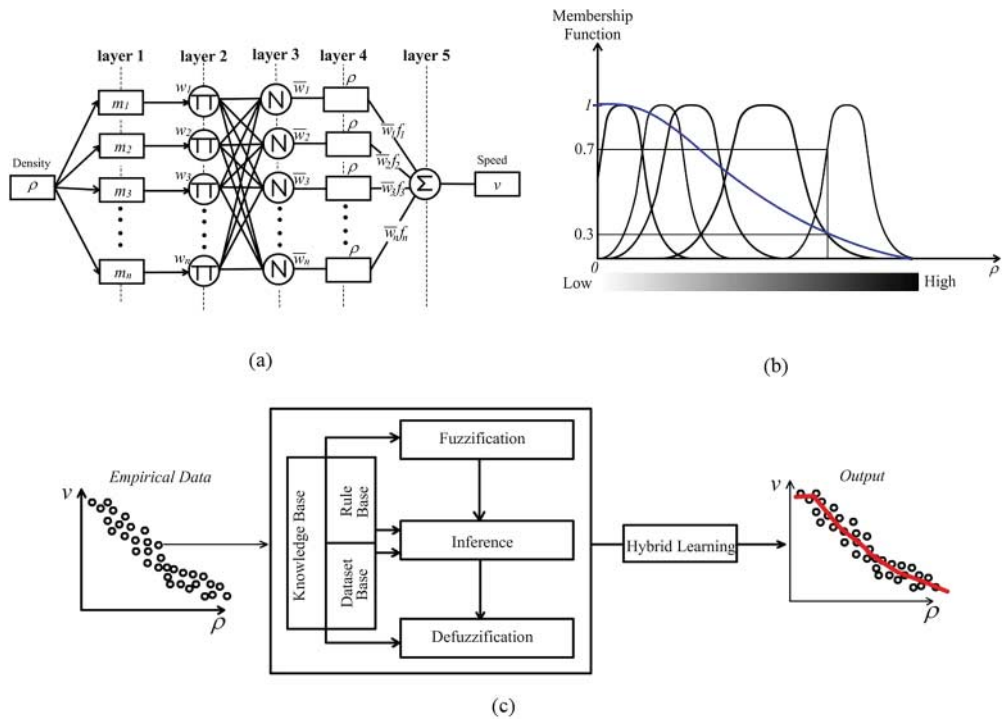


Figure 2. ANFIS structure: (a) layer representation; (b) linguistic representation; (c) block diagram.

....

Rule_(n): **IF** ρ is A_n , **THEN**

$$f_n = p_n \rho + r_n$$

where ρ is the density, A_i is the fuzzy set, f_i are the outputs within the fuzzy region specified by the fuzzy rule, and p_i and r_i are the design parameters that are determined during the training process.

Figure 2(a) illustrates the reasoning mechanism for this Sugeno model, which is the basis of the ANFIS model. The ANFIS architecture implements the rules within this mechanism. In this figure, a circle indicates a fixed node, whereas a rectangle indicates adaptive node. Moreover, ANFIS has a five-layer architecture. Each layer is explained in detail below.

Layer 1 is made up of input parametric MFs, and it provides the input values to the following layer. All nodes here are considered as adaptive nodes. This layer takes density as input and the outputs of Layer 1 are the fuzzy membership grade of the inputs, which are given in Equation (1),

$$O_{1,i} = m_{A_i}(\rho), \quad i = 1, 2, \dots, n \tag{1}$$

where ρ is the density input to node i and A_i is the linguistic labels (high, low, etc.) associated with this node function that is shown in Figure 2(b). m_{A_i} can adopt any MF. If A bell-shaped

MF is employed, m_{A_i} is given in Equation (2),

$$m_{A_i} = \frac{1}{1 + \left[\left(\frac{\rho - c_i}{a_i} \right)^2 \right] b_i}, \quad i = 1, 2, \dots, n \quad (2)$$

or the Gaussian MF by Equation (3),

$$m_{A_i}(\rho) = \exp \left[- \left(\frac{\rho - c_i}{a_i} \right)^2 \right], \quad i = 1, 2, \dots, n \quad (3)$$

where a_i, b_i, c_i are the parameters of the MFs. The parameters in this layer are usually referred to as premise parameters.

The membership layer is the second layer and here nodes are fixed nodes. This layer involves fuzzy operators; it uses the **AND** operator to fuzzify the inputs. They are labeled as π , indicating that they perform as a simple multiplier. However, the speed–density relation has only one input. Thus the output of these layers can be represented by Equation (4). This layer exerts the firing strength of the rules. The individual rule possesses individual firing strength.

$$O_{2,i} = w_i = m_{A_i}, \quad i = 1, 2, \dots, n \quad (4)$$

In Layer 3, the nodes are also fixed nodes labeled by N to indicate that they play a normalization role to the firing strengths from the previous layer. The output of this layer can be represented by Equation (5). The output of this layer is called normalized firing strengths.

$$O_{3,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2 + \dots + w_n}, \quad i = 1, 2, \dots, n \quad (5)$$

In Layer 4, the nodes are adjustable. The output of each node in this layer is simply the product of the normalized firing strength and a first-order polynomial (for a first-order Sugeno model). The output of this layer is given in Equation (6),

$$O_{4,i} = \bar{w}_i f_i = \bar{w}_i (p_i \rho + r_i), \quad i = 1, 2, \dots, n \quad (6)$$

where \bar{w}_i is the output of Layer 3 and p_i, r_i are the consequent parameters.

In the Layer 5, there is only single fixed node labeled with \sum . This node performs the summation of all incoming signals. The overall output of the model is given in Equation (7).

$$O_{5,i} = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (7)$$

The learning algorithm for ANFIS is a hybrid algorithm that is a combination of gradient descent and least squares methods. In the forward pass of the hybrid learning algorithm, node outputs go forward until Layer 4 and the consequent parameters are identified by the least squares estimate. In the backward pass, the error rates propagate backward and the premise parameters are updated by the gradient descent algorithm. The hybrid learning approach converges much faster by reducing search space dimensions of the original

back propagation method (Jang 1991). The overall output can be expressed by Equation (8), where p_1, p_2, \dots, p_n and r_1, r_2, \dots, r_n are the linear consequent parameters.

$$\left. \begin{aligned} f &= \frac{w_1}{w_1+w_2+\dots+w_n}f_1 + \frac{w_2}{w_1+w_2+\dots+w_n}f_2 + \dots + \frac{w_n}{w_1+w_2+\dots+w_n}f_n, \\ f &= \bar{w}_1(p_1\rho + r_1) + \bar{w}_2(p_2\rho + r_2) + \dots + \bar{w}_n(p_n\rho + r_n), \\ f &= (\bar{w}_1\rho)p_1 + (\bar{w}_1)r_1 + (\bar{w}_2\rho)p_2 + (\bar{w}_2)r_2 + \dots + (\bar{w}_n\rho)p_n + (\bar{w}_n)r_n, \end{aligned} \right\} \quad (8)$$

3.1. The motivation for using ANFIS

The functional form of ANFIS is the Sugeno model (see Equation (6)), which is trained over empirical data using hybrid algorithm mentioned in the previous section. Therefore, the Sugeno model and its parameters (premise parameters and consequent parameters) are directly related with the FD parameters. Equation (8) can be rearranged as

$$v = \rho \sum_i^n \bar{w}_i p_i + \sum_i^n \bar{w}_i r_i \quad (9)$$

where \bar{w}_i is the firing strength, which depends upon the premise parameters (a_i and c_i) and p_i, r_i are the consequent parameters. However, Equation (9) can also be rewritten as

$$\begin{aligned} v &= \mathbf{A}\rho + \mathbf{B} \\ &= \mathbf{B} \left(1 + \frac{\mathbf{A}}{\mathbf{B}}\rho \right) \end{aligned} \quad (10)$$

where $\mathbf{A} = \sum_i^n \bar{w}_i p_i$ and $\mathbf{B} = \sum_i^n \bar{w}_i r_i$. Equation (10) and the functional form proposed by Greenshields (1935) for speed–density are equivalent. However, unlike the Greenshields model, the parameters in Equation (10) are not constant as follows: $\mathbf{A} = f(\rho, a_i, c_i, p_i)$ and $\mathbf{B} = f(\rho, a_i, c_i, r_i)$.

ANFIS provides shape flexibility incorporating three factors: (a) number of MF; (b) premise parameters; and (c) consequent parameters. Among these, premise parameters provide the required curvature to the speed–density FD, whereas the number of MF controls the extent of the curvature. Individual MF covers a distinct region of the speed–density curve. It opens a scope for agglomeration of the empirical data according to their characteristics into a particular MF. Moreover, the overlapping among the MFs causes superimposition of the curves, which ultimately adds further flexibility to the ANFIS model. Finally, the consequent parameters control direction of baseline of the curvature. Combination of these three functionalities enables ANFIS model to capture the actual trend of empirical speed–density relationship.

3.2. FD parameters estimation from ANFIS model

In general, free flow speed (v_f) is the speed at $\rho = 0$ and the jam density (ρ_j) can be termed as the density at $v = 0$. Using these conditions, v_f and ρ_j can be estimated from Equation (10). Consequently, putting $(\rho, v) \in \{(0, v_f), (\rho_j, 0)\}$, the following relations can be

obtained as

$$v_f = \mathbf{B}$$

$$\rho_j = -\frac{\mathbf{B}}{\mathbf{A}}$$

For homogenous MFs and the Sugeno model, the free flow speed (v_f) and jam density (ρ_j) can be extracted from the following equations:

$$v_f = \sum_{i=1}^n \overline{w}_i r_i \quad (11)$$

$$\rho_j = \frac{\sum_{i=1}^n \overline{w}_i r_i}{\left| \sum_{i=1}^n \overline{w}_i p_i \right|} \quad (12)$$

where \overline{w}_i is the firing strength of MF i and n is the total number of MFs.

Equations (11) and (12) entail that both v_f and ρ_j are related to premise (a_i and c_i) and consequent parameters (p_i and r_i). v_f includes only the constant term of the consequent parameters. On the other hand, ρ_j considers both gradient (p_i) and constant term (r_i) while estimation. It illustrates that the slope of the speed–density relationship also affects ρ_j . This Premise-Consequent Conjugate Effect (PCCE) aids in providing the desired calibrated FD. Premise parameters control the shape of the FD. On the other hand, consequent parameters are liable for providing the final output value.

The gradient p_i can be described as the representative of the gradient of the speed–density relationship $\left(\frac{dv}{d\rho}\right)$, where $\frac{dv}{d\rho} = \frac{1}{\rho} \frac{dq}{d\rho} - \frac{v}{\rho} = f\left(\frac{dq}{d\rho}, v, \rho\right)$, then $\exists ! i | \rho_j \rightarrow p_i$, which is similar to ‘backward wave speed’ phenomenon (Mehran, Kuwahara, and Naznin 2012). It also indicates that there exists only one MF that contains the consequent parameters for estimating the jam density. However, practical application is not straight forward as the consequent parameters in Sugeno models, premise parameters in MFs are not homogenous. This heterogeneity can be observed from the dissimilarity in geometrical shape among the MFs and Sugeno models.

In practice, different MFs cover different portions of the data (as shown in Figure 3(b)). It makes the parameter values different for different MFs which introduce heterogeneity. Thus complexity arises when formulations based on homogeneity, such as Equations (11) and (12) are used, as the value of the FD parameters become dependent on the density. Thus, to solve this, conditions need to be adopted which will also ensure the accuracy in estimation of FD parameters. In this regard, two basic properties of the FD: (i) v_f can be obtained at $\rho = 0$ and (ii) ρ_j can be obtained at $v = 0$ are considered. These entail that v_f must comprise of the premise and consequent parameters of the first MF and ρ_j is included in the last MF; the first and last MFs cover the data associated with v_f and ρ_j respectively (see Figure 3(b)). Moreover, the premise and consequent parameters needed for estimating q_{\max} reside in one of the MFs apart from the first and the last one.

To accommodate all these above-mentioned considerations, modified formulations are necessary to obtain FD parameters. These considerations reduce the complexity and

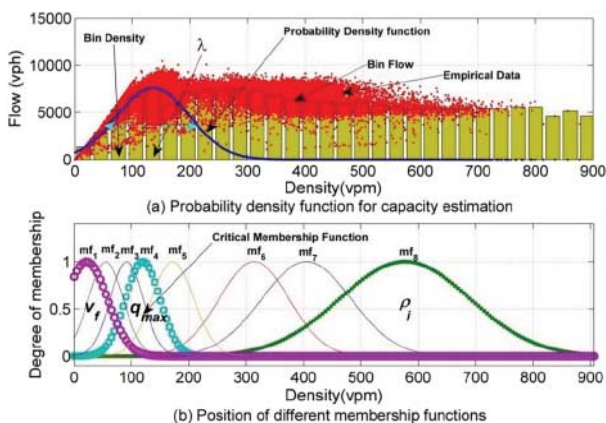


Figure 3. (a) Probability density function for capacity estimation and (b) position of different MFs.

dependency of the formulations on density and ensure an accurate result in FD parameter estimation. v_f can be estimated from the following equation:

$$v_f = \sum_{i=1}^n \bar{w}_i(c_1) r_i \quad (13)$$

where $\bar{w}_i(c_1) = w_i(c_1) / \sum_i w_i(c_1)$ and $w_i(c_1) = \exp[-((c_1 - c_i)/a_i)^2]$.

c_1 is the premise parameter of the first MF. Inclusion of c_1 eliminates the effect of the other MFs by reducing firing strength for them. The reduced firing strength reduces the effect of other MFs on v_f estimation, whereas infiltration of the unwanted parameters significantly affects the estimation of parameter v_f . Similarly, the estimation of ρ_j can be done using the second condition which uses the final MF. Thus ρ_j can be estimated from

$$\rho_j = \frac{\sum_{i=1}^n \bar{w}_i(c_n) r_i}{\left| \sum_{i=1}^n \bar{w}_i(c_n) p_i \right|} \quad (14)$$

where $\bar{w}_i(c_n) = w_i(c_n) / \sum_i w_i(c_n)$ and $w_i(c_n) = \exp[-((c_n - c_i)/a_i)^2]$.

However, the determination of capacity (q_{\max}) is not straightforward. The required MF for estimation of q_{\max} is in which the max flow resides. This particular MF determined from the probability distribution function is as follows:

$$\Pr(\rho) = e^{-\frac{(\rho-\lambda)^2}{2\sigma^2}} \quad (15)$$

In Equation (15), λ is determined using the empirical data after the data is aggregated into bins. It is noted that the flow data is estimated by multiplying the estimated speed data with the empirical density data. Afterwards, the bin density (λ_j) and bin flow (f_j) are determined from the following formula:

$$\text{Bin density, } \lambda_j = \frac{\rho_{e_k} + \rho_{s_k}}{2}$$

Bin flow, $f_j = \text{median}\{f_{s_k(j)}, f_{s_k+1(j)}, f_{s_k+1(j)}, \dots, f_{e_k(j)}\}$, $j = 1, 2, \dots, B$ where s_k and e_k denote the start and end density of a particular bin, $M = e_k - s_k + 1$ is the number of data in the bin, and B is the total number of bins. The bin flow over the different bin density is shown in Figure 3(a).

From the aggregation, the parameter λ is obtained considering $\lambda = \lambda_j, \exists! J | f_j = \max\left(\bigcup_{j=1}^B f_j\right), J \in j$. On the other hand, the value of σ is determined from the following equation:

$$\sigma = \left(\frac{\lambda}{\rho_j}\right) (Q_3 - Q_1) \left(1 - \frac{\sqrt{\frac{1}{B} \sum_{i=1}^B (f_j^{\max} - f_j)^2}}{f_j^{\max}}\right), \quad \exists! f_j^{\max} | f_j^{\max} \in \max(f_i)$$

where ρ_j is the jam density obtained from Equation (14). Q_3 and Q_1 are the 75th and 25th percentile of the λ_j . The $\left(\frac{\lambda}{\rho_j}\right) (Q_3 - Q_1)$ term dominates the width of the distribution curve. However, it is reduced by the term between $\left(1 - \frac{\sqrt{\frac{1}{B} \sum_{i=1}^B (f_j^{\max} - f_j)^2}}{f_j^{\max}}\right)$ which stands for the deviation in the flow from f_j^{\max} . Its value ranges from 0 to 1. Values towards 1 indicate skewed parabola-shaped flow–density diagram and values towards 0 indicate inverse lambda shape FD. The probability distribution generated by Equation (15) is shown in Figure 3(a). Afterwards the critical density for each MF is determined from the following equation (16) and the probability of each of the critical densities to become the actual critical density is determined using Equation (15). For which, the probability is maximized, is called the critical MF. Using the firing strength and the value of the consequent parameters of that MF, the capacity can be found from Equation (17),

$$\rho_{c_i} = \frac{r_i}{|2p_i|} \quad (16)$$

$$q_{\max} = \left(\frac{r_c}{2p_c}\right)^2 \sum_{i=1}^n w_i p_i - \left(\frac{r_c}{2p_c}\right) \sum_{i=1}^n w_i r_i \quad (17)$$

where r_c and p_c are consequent parameters of critical MF.

4. Model calibration and validation

There are a variety of options and settings to explore while proposing an ANFIS model. Authors provide an illustration of the available options and their effects on a number of modeling aspects. For example, an experiment was performed to identify the optimal split of data into a training and a testing sample. In addition, three other aspects are the type of MFs, number of epochs and number of clusters. These aspects are also illustrated here. The developed ANFIS model is further validated using a different dataset and against other conventional models and the results are accumulated in later part of this paper.

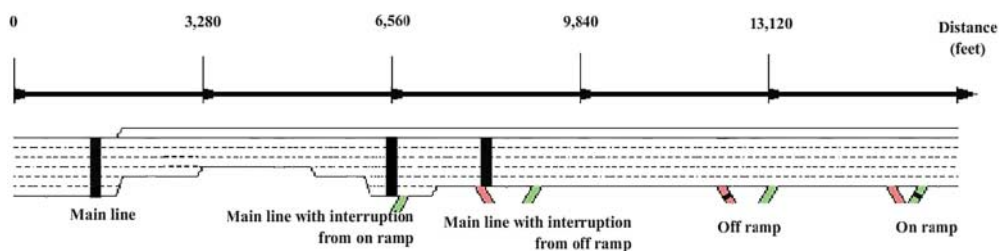


Figure 4. Illustration of layout of study site I-80, Berkeley (not to scale).

4.1. Data sources

Data collection process for model calibration, validation and comparison is based on five different locations. These locations represent different driving behavior and consist of mainline, mainline with interruption from off ramp, mainline with interruption from on ramp, off-ramp and on-ramp. Four datasets were collected for each location from different locations. One dataset is used for calibration and the other three are used for validation. These datasets were collected from PeMS (<http://pems.dot.ca.gov/>) which provides flow, speed and density data across different vehicle detector stations in the form of time series data over days of operation. The data used in this research were aggregated over an interval of 5 minutes. The data was collected from I-80 near Berkeley which is a freeway with on-ramps and off-ramps. At each location, 1 year continuous observations were collected. This time interval is long enough for calibration and validation of FD. A general setting of I-80 detectors from the study site in Berkeley, CA, is shown in Figure 4.

4.2. Calibration results

To implement and test the proposed single-input ANFIS architecture, MATLAB Fuzzy Logic Toolbox from MathWorks was selected as the development tool. This tool offers an environment to build and evaluate fuzzy systems using a graphical user interface (GUI). It consists of a Fuzzy inference system (FIS) editor, the rule editor, an MF editor, the fuzzy inference viewer, and the output surface viewer.

4.2.1. Effect of size of training and testing datasets

In neural networks, a subset of the sample (training dataset) is used to train the neural network. Later, to test the ability of the model to reproduce other realities, the remaining portion of the sample (testing dataset) that was not used for model calibration is used as a testing device. In practice, the optimal composition of the original sample (separation of the data into training and testing sets) needs to be known. The effects of different sizes of the training and the testing datasets on goodness-of-fit value are summarized in Table 3.

For this study, five different compositions of datasets for each location were considered to assess the effect. With an increase in training dataset, the fitness value (R^2) obtained from Equation (18) improves as the generated FIS structure is fed with larger number of training dataset to optimize its parameter values. On the other hand, decreasing testing dataset causes the 'overfitting' phenomenon to occur. Overfitting is a common problem in ANFIS model building, which occurs when the data is over trained by ANFIS. Every dataset that

Table 3. Effect of different composition of training dataset on fitness of data (R^2 values) for different locations.

Highway locations	Training dataset and testing dataset composition				
	50–50	60–40	70–30	80–20	90–10
Mainline	0.86825	0.85964	0.8663	0.87682	0.87399
Mainline close to off-ramp	0.88327	0.88350	0.88411	0.88459	0.86604
Mainline close to on-ramp	0.91362	0.91608	0.91814	0.92058	0.91554
Off-ramp	0.90985	0.91711	0.91360	0.92654	0.91132
On-ramp	0.94493	0.94596	0.94713	0.95118	0.94361

Table 4. Effect of type of MF on fitness of data (R^2 values) for five different highway locations.

Highway locations	Type of input MFs					Clustered based gaussian
	Triangular	Symmetric gaussian	Gaussian combination	Trapezoidal	Generalized bell shaped	
Mainline	0.86772	0.87208	0.86729	0.8670	0.86781	0.87463
Mainline close to off-ramp	0.88127	0.88413	0.88465	0.88405	0.88473	0.88488
Mainline close to on-ramp	0.91977	0.91773	0.91434	0.91307	0.92041	0.92063
Off-ramp	0.85731	0.88766	0.88696	0.89194	0.87546	0.90988
On-ramp	0.94236	0.94319	0.94317	0.94273	0.94343	0.94509

is trained using ANFIS has its maximum number of epochs before overfitting occurs; this causes the predicted output to be over its accuracy. From Table 3, it can be observed 80-20 composition demonstrated the best fit over others. As, 80-20 split of original dataset for training and checking produced best results for all locations, this size is adopted as optimum composition for ANFIS modeling in this paper.

$$R^2 = 1 - \frac{SS_E}{SS_T} \quad (18)$$

where $SS_E = \sum_i^N (v_{Actual,i} - v_{ANFIS,i})^2$, $SS_T = \sum_i^N (v_{Actual,i} - \bar{v}_{Actual})^2$, $v_{Actual,i}$ = Actual speed for i th observation, and $v_{ANFIS,i}$ = ANFIS estimated speed for i th observation.

4.2.2. Effect of MF type

MFs determine the shape of the projection of input–output mapping and the parameters of these MFs are adjusted through the fuzzy system to capture the trend of the empirical data. In this study, six types of MFs were considered: (1) triangular (trimf); (2) symmetric gaussian (gaussmf); (3) gaussian combination (gauss2mf); (4) trapezoidal (trapmf); (5) generalized bell shaped (gbellmf); and (6) clustered based gaussian.

The results for different MFs are summarized in Table 4 for different locations. In all cases, clustered based gaussian achieved the best fitness value (R^2) as determined from Equation (18). As clustering of a given dataset serves the purpose of discerning its natural classification and generate an incisive representation of the data, a clustered based approach reproduces better results in all locations. In addition, from the visual evaluations

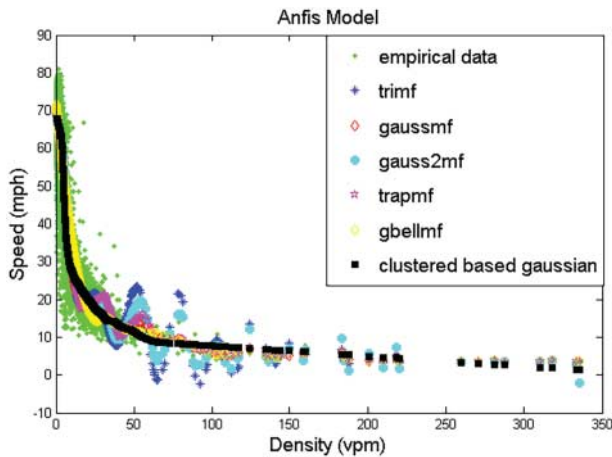


Figure 5. FD of speed–density model upon empirical scattered plot for on-ramp location using different type of MFs: (i) Triangular (trimf), (ii) Symmetric gaussian (gaussmf), (iii) Gaussian combination (gauss2mf), (iv) Trapezoidal (trapmf), (v) Generalized bell shaped (gbellmf), and (vi) clustered based gaussian.

(see Figure 5), it is noticed that FIS from clustered based gaussian MFs best project the trend of the speed–density relationship.

4.2.3. Effect of number of epoch

Epoch number controls the number of times the combined execution of back propagation and least square estimation (hybrid learning) occurs for training dataset. The training process terminates once the designated epoch number or the training goal is reached. However, the number of epoch should be chosen such that overfitting does not occur. In this research, experiments with ANFIS training on six different types of MFs were conducted with varying epoch numbers (i.e. 3, 5, 10, 50, and 100). Analysis shows that no notable variation in results was found (R^2 value varying less than 0.002% on average). Only a minor change in computation time was noticed if all other calibration parameters were kept unchanged. For example, an average runtime for a 1-year dataset with epoch number 3 was found to be 32.04 seconds, with epoch number 5, runtime increased to be 36.87 seconds and with epoch number 100, the runtime was 72.12 seconds. Later, basing on recommendation from some of the literatures (Petković et al. 2012), a fixed epoch number of 100 is chosen for analysis all through this research.

4.2.4. Effect of number of clusters

In the subclustered ANFIS model, each input variable is clustered into several class values to build up fuzzy rules and each fuzzy rule would be constructed through several parameters of MF. The number of MFs and rules are mutually related; more MFs imply more rules and determine the level of detail of the model. Fuzzy C-means (FCM) clustering is used in this study to identify the antecedent MFs and a set of rules. The results of the effects of number of clusters on the mainline dataset are summarized in Table 5. From the analysis, it is observed that with increasing number of clusters, R^2 value improves and the absolute error ($|e|$) of the fitted models (see Equation 19) decreases but these improvements in results are

Table 5. Improvement in fitness of data (R^2 values) and absolute error of fitted output with increasing no. of clusters.

Number of clusters	R^2 value	Absolute error
3	0.86669	0.025507
5	0.86688	0.025458
7	0.86753	0.025357
9	0.86761	0.025341
12	0.86776	0.025316
15	0.86776	0.025312

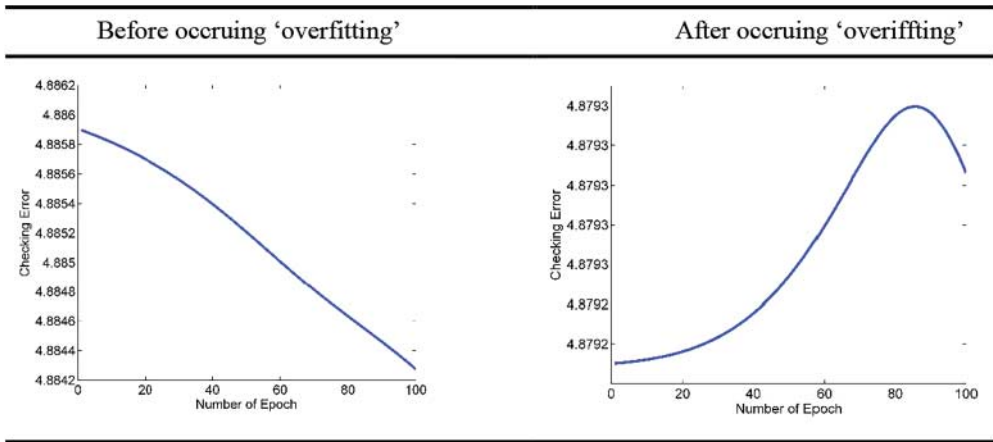


Figure 6. Overfitting phenomenon on root mean square error versus epochs curve (mainline with interruption from off-ramp location is illustrated here).

imperceptible.

$$|e| = \frac{1}{N} \sum_i^N \frac{|v_{\text{Actual},i} - v_{\text{ANFIS},i}|}{v_{\text{Actual},i}} \quad (19)$$

where N = Total number of observations.

Although it is better to take a large number of clusters as it indicates better classification of varying data points, it may cause ‘overfitting’ (see Figure 6). Overfitting is analyzed by setting the number of clusters to different values for 100 epochs. Considering these two phenomena, an optimum number of clusters has to be determined for which maximum R^2 value will be obtained without the occurrence of overfitting. In this regard, the following objective function (Equation 20) has been used to determine the optimum number of clusters. The objective function gives the optimum number of clusters considering two aspects: whether (1) inflection point in error–epoch curve exists, (2) the change in R^2 is less than a particular significance threshold τ or not.

Using Equation 20, optimal number of clusters are obtained for different freeway locations. From analysis, it was obtained that the proposed ANFIS model works best without causing the overfitting phenomenon for all the study locations, if the number of MFs is in between 4 and 7. Particularly, the optimum number of MFs are: 8 for freeway mainline, 7 for mainline close to off-ramp, 6 for mainline close to on-ramp, 4 for off-ramp, and 5 for on-ramp. The analysis also shows that for different highway locations, overfitting occurs at

different number of clusters. This change in optimum number of clusters can be contributed to the traffic stability phenomenon. As mentioned in Section 1, geometrical and environmental changes have an influence on the variation of the origin of the traffic instability. At different locations, variance in the reaction of drivers to changes in behavior of the leading vehicle causes instability. This variation leads to a different number of cluster points.

$$\left. \begin{aligned} N_o &= \text{Min}(C_O, \text{Max}(C_O, C_{R^2})) \\ C_O &= \{n|n \in Z, E_{n,i+1} \geq E_{n,i}\} \\ C_{R^2} &= \{n|n \in Z, R^2_n - R^2_{n-1} \geq \tau, R^2_n > R^2_{n-1}\} \end{aligned} \right\} \quad (20)$$

where E = relative absolute error, C_O = maximum number of clusters without causing overfitting, C_{R^2} = minimum number of clusters attaining saturated R^2 , and τ = significance threshold. This equation gives the optimum number of clusters for any location considering that overfitting must not occur before attaining a marginal value of goodness-of-fit. The iteration proceeds for different number of clusters n until it finds either overfitting phenomena or marginal goodness-of-fit value. The overfitting phenomena is defined as a change in gradient in error–epoch curve. Equation 20 represents the discretized form of finding this gradient change. When error $E_{n,i+1}$ for a certain n number of clusters becomes greater than the error $E_{n,i}$ generated in the previous epoch i , the iteration stops and gives the corresponding n number of clusters.

4.3. Model validation result

Validation is indicative of how the proposed model replicates system behavior with enough reliability to serve analysis objectives. The ANFIS-based speed–density model is fitted from an I-80 dataset. In order to test whether the model also fits well with empirical data from other highway locations, in total 15 independent datasets (3 datasets for each five locations) are fitted using the ANFIS model as shown in Figure 7. Five locations considered for validation are: (a) Mainline (LC-1); (b) Mainline close to off-ramp (LC-2); (c) Mainline close to on-ramp (LC-3); (d) Off-ramp (LC-4); and (e) On-ramp (LC-5). LC-1 contains uninterrupted traffic flow having very light amount of data below the speed 20 mph in comparison to other locations. LC-2 and LC-3 represent traffic flow influenced by either off-ramp or on-ramp, respectively. Furthermore, LC-4 and LC-5 are isolated flows of off-ramp and on-ramp having different characteristics than mainline.

4.3.1. Comparison with single regime models

A graphical representation of the numerical comparison of different model performance in terms of goodness-of-fit (R^2) at various locations using the testing data is plotted in Figure 7. The models considered are Greenshields model, Greenberg model, Underwood model, Northwestern model, Pipes Munjal model, Drew model, and the proposed ANFIS model. All the aforementioned models are included in a single plot to see the effectiveness in capturing the traffic state along with the empirical trend for different locations. The results compared to empirical observations show that:

1. There exists a curvature change in the speed–density relationship (see Figure 7). However, the relative positions of the point of curvature change (inflection point) with respect to the $(v_f, 0)$ and $(0, \rho_j)$ are not same for all locations, where v_f and ρ_j are free

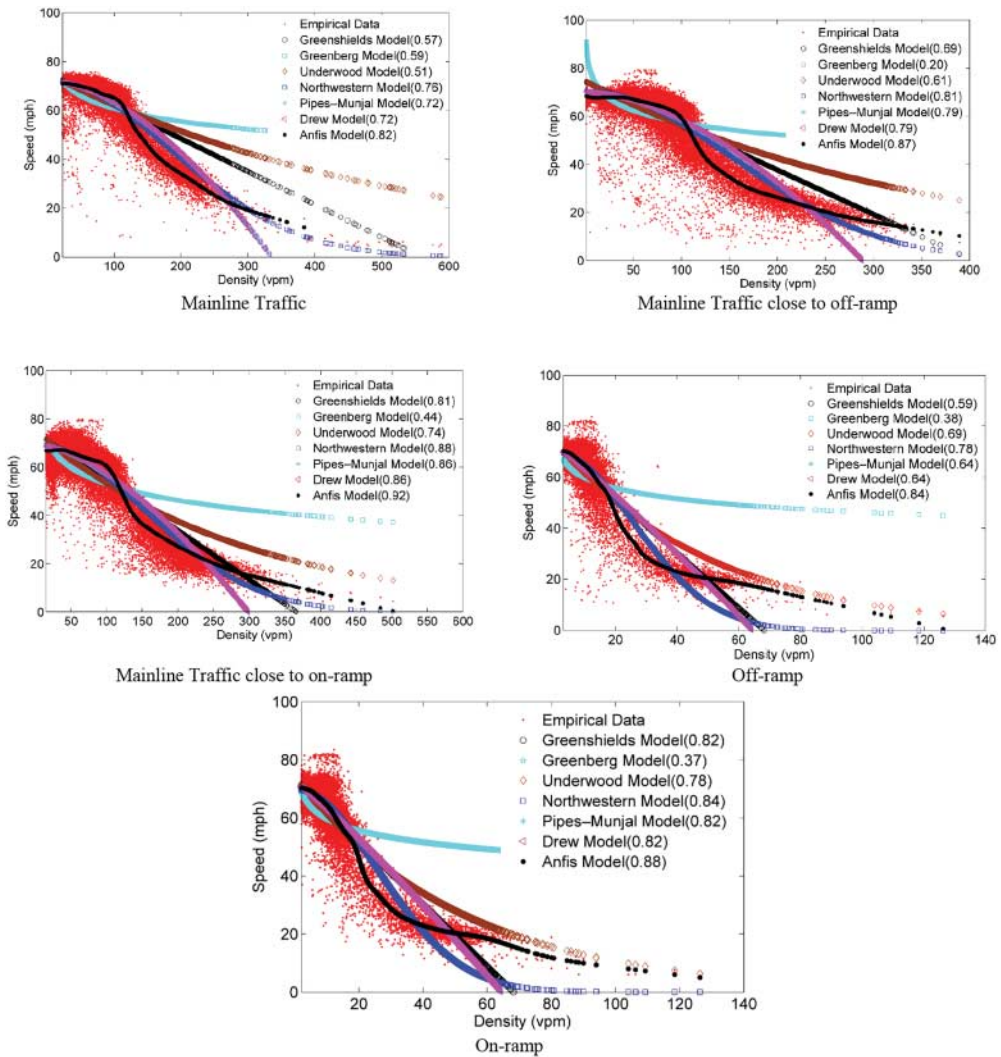


Figure 7. Graphical representation of different model performances at I-80.

flow speed and jam density, respectively. It can be seen that the inflection point shifts to the left from LC-1 to LC-5. It is due to the origin of instability near the inflection point. This instability occurs due to sudden speed variation caused by heterogeneous driving behavior. It can be seen from Figure 7 that only the ANFIS model captures the exact inflection point in all locations. Interestingly, the left part of the inflection point is concave and the other is convex. The left part is partially skewed downward and the right part becomes asymptotic near the zero speed. ANFIS model can also sense the shift in the inflection point and therefore show different inflection points for different locations. Although, Northwestern model shows the inflection point, the point is not in exactly same position as the empirical one. Other speed–density models cannot capture this kind of trend in data.

Table 6. Performance comparison in estimation of FD parameter.

Highway locations	v_f (mile/hour)			q_{max} (veh/hour)			ρ_j (veh/mile)		R^2 value	
	ANFIS	5PL	%Error	ANFIS	5PL	%Error	ANFIS	5PL	ANFIS	SPL
Mainline	68.84	69.86	1.46	7218.58	7158.75	0.84	299.98	–	0.82	0.80
Mainline close to on-ramp	67.39	67.44	0.08	3263.77	3259.42	0.13	423.40	–	0.92	0.91
Mainline close to off-ramp	67.94	67.86	0.12	3338.05	3058.01	9.16	756.50	–	0.87	0.86
On-ramp	65.14	68.38	4.74	3886.30	3969.12	2.08	454.56	–	0.88	0.87
Off-ramp	70.48	68.97	2.20	4030.74	4062.10	0.77	358.70	–	0.84	0.84

2. The proposed ANFIS model achieved the best goodness-of-fit (R^2) in all the locations (see Figure 7). Although Northwestern, Drew and Pipes-Munjal models achieve good R^2 value, Greenberg, Greenshields and Underwood perform poorly to capture the trend of empirical data. Greenshields, being a linear model, fails to produce good R^2 as none of locations has data with linear trend. In addition, both Greenberg and Underwood are convex in nature and so these models fail to capture the concave portion of the data. However, the performance of the underwood model increases from LC-1 to LC-5. It is due to decrease in the length of concave part of the curve.
3. Another observation from Figure 7 is that ANFIS generated curve does not have any definite shape, rather its shape changes with the change in the trend of the dataset. In other words, the gradient of the equilibrium curve generated by ANFIS model is data driven as it does not have any deterministic gradient. However, other models have a deterministic gradient in nature meaning that they try to fit the traffic dynamics to a definite shape forcing the empirical data into the model without capturing the actual trend.
4. Change in shape, orientation and location of the transition region in speed–density is observed from LC-1 to LC-5. Interestingly, the proposed ANFIS model has captured all the changes perfectly, whereas the other models have failed to capture the changes. The superimposed effect of premise parameters from different MFs enables the ANFIS model to capture the curvature of the transition region. On the other hand, consequent parameters determine the target gradient of the superimposed curvature.

In order to demonstrate the proposed ANFIS model's efficacy in estimation of FD parameters, we consider the 5-parameter logistic speed–density model (5PL) of Wang et al. (2011). The choice of the 5PL model is supported by its discerning similarities with the ANFIS model in capturing inflection point and also achieving good R^2 value in comparison to the traditional models. Table 6 lists the comparison of performance of the ANFIS model in estimation of the FD parameters with the 5PL model. The results reveal that both models produce similar result in parameter estimation as well as achieve good R^2 value for all five highway locations. In addition, ANFIS model provides estimation of jam density which 5PL model cannot determine.

4.3.2. Comparison with multi-regime models

An illustrative numerical performance comparison with Multi-Regime models at LC-1 is plotted in Figure 8. Additionally, the goodness-of-fit values for different location are provided in Table 7. The Multi-Regime models considered in this research work are the Edie model and Two-Regime Linear model. Quandt's likelihood estimation technique (Quandt

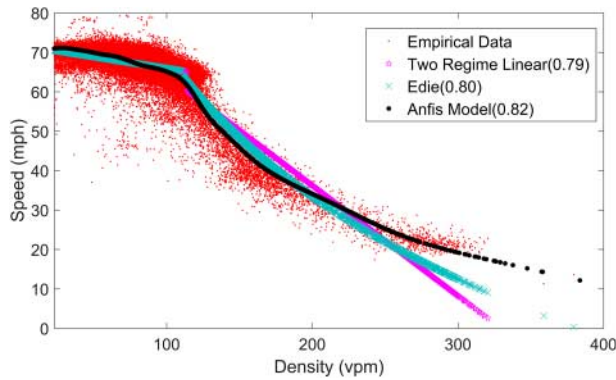


Figure 8. Comparison between ANFIS and multi-regime models (mainline location is illustrated here).

Table 7. Performance comparison with multi-regime models.

Highway locations	Eddy	Two regime linear	ANFIS
Mainline	0.80	0.79	0.82
Mainline close to on-ramp	0.85	0.83	0.92
Mainline close to off-ramp	0.79	0.75	0.87
On-ramp	0.81	0.79	0.88
Off-ramp	0.78	0.72	0.84

1958) is adopted in these two models to identify breakpoints between the regimes for freeway traffic (May 1990). Analysis shows that Edie and Two-Regime Linear model have performed well and achieved goodness-of-fit value 0.80 and 0.79, respectively, for mainline location. However, the ANFIS model has achieved the maximum goodness-of-fit value (0.82) and has outperformed them. The existence of estimable breakpoint allows the multi-regime models to reduce the residual while performing the regression. However, it forcefully fits the data into the predefined shapes identical to the single regime models. Thus it also inherits the limitations of the single regime models such as inability to capture the transition point, the inflection point, and the gradient of the empirical data. In contrast, the data-driven ANFIS model is able to address these limitations; therefore, the ANFIS model performed best.

5. Conclusion and future directions

FDs have been the subject of comprehensive study due to its importance in both the design of traffic facilities and for the control of traffic operations. Traditional speed–density models try to fit the traffic dynamics to a definite shape (i.e. logarithmic, exponential, and exponential to the quadratic and various forms of polynomials) rather than capturing the variations in drivers' behavior caused by the shockwave in the transition zone. In addition, they fail to acknowledge the effect of different geometric patterns of the roadway (i.e. on-ramp/off-ramp on a freeway) on the shape of the speed–density FD.

We have applied the ANFIS in modeling equilibrium speed–density relationship because of its efficacy in being shape-flexible. ANFIS provides shape flexibility incorporating three factors: (a) premise parameters; (b) number of MF; and (c) consequent parameters. Among

these, premise parameters provide the required curvature to the speed–density FD, whereas the number of MF controls the extent of the curvature. Individual MF covers a particular region of the speed–density curve. It opens a scope for agglomeration of the empirical data according to their characteristics into a particular MF. The overlapping among the MFs causes superimposition of the curves, which ultimately adds further flexibility to the ANFIS model. In contrast, the consequent parameters control the direction of the baseline of this curvature. Combination of these three factors enables ANFIS model to capture the actual trend in empirical speed–density data. Moreover, in this research, the Sugeno model has been correlated with the equivalent Greenshields model, which proves the eligibility of ANFIS model to represent speed–density relation. It helped to reveal the contribution of premise and consequent in the speed–density relation. Incorporating this knowledge, the PCCE relationship with the speed–density equilibrium is established. Accordingly, the MFs for estimating free flow speed, jam density and capacity have been revealed.

A number of options (i.e. optimal split of data into an estimation sample and a validation sample, selection of suitable MF) offered by ANFIS are illustrated using data from I-80 near Berkeley. Hybrid learning algorithm has been used to optimize the parameters within the ANFIS structure. The study results show that with an increase in training dataset, the fitness value improves as the generated FIS structure is fed with larger number of input–output data pairs to optimize its parameter values. It is observed that 80-20 split of the original dataset for training and checking produced best results for all locations. For MFs, clustered based gaussian MFs best project the trend of the speed–density relationship in all locations as clustering of a given dataset serves the purpose of discerning its natural classification and generate an incisive representation of the data. An epoch number of 100 is chosen so that ‘overfitting’ phenomenon does not occur. Epoch number controls the number of times the combined execution of back propagation and least square estimation (Hybrid learning) occurs for training dataset. Model results show that with increasing number of clusters, fitness of data (R^2 value) improves and the absolute error of the fitted models decreases but these improvements in results are imperceptible. Although it is better to take a large number of clusters as it indicates better classification of varying data points, it may cause ‘overfitting’. To incorporate this contrasting behavior, we propose an objective function to obtain the optimal number of clusters so that ‘overfitting’ phenomenon does not occur.

Calibrated ANFIS models are validated using 15 different datasets and against other conventional speed–density models. Results show proposed ANFIS-based speed–density model achieves the best R^2 value compared to the other five models that indicate the model’s efficacy in capturing the traffic state along with the empirical trend for all five locations. Empirical observation shows that the relative positions of the point of curvature change (inflection point) with respect to free flow speed and jam density are not same for all locations. This variation is due to the origin of instability near the inflection point and occurs due to sudden speed variation caused by drivers’ random decision about lane changing. ANFIS model captures this variation most effectively. We then compare the performance of the ANFIS model in matching the empirical speed–density observations and also in estimation of the FD parameters with the SPL proposed by Wang et al. (2011). The results reveal that the ANFIS model achieves R^2 value in the range of 0.82–0.92 for

all five highway locations and produces similar result to the 5PL model in FD parameter estimation.

In high-order traffic models, the speed–density or speed–spacing FD also plays a critical role. However, the speed–density relation used on these models are as similar to conventional models. ANFIS will be more suitable as it is data driven and capture real-time flow dynamics more efficiently. In this regard, authors have a plan to investigate the performance of ANFIS-based higher order traffic state estimation model.

Acknowledgements

The authors would like to express thanks to the Committee for Advanced Studies and Research (CASR) of Bangladesh University of Engineering and Technology (BUET) for the financial support.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Committee for Advanced Studies and Research (CASR), Bangladesh University of Engineering and Technology (BUET) [Grant Number 69].

ORCID

Md. Hadiuzzaman  <http://orcid.org/0000-0002-5690-3351>

Tahmida Hossain Shimu  <http://orcid.org/0000-0002-2468-0085>

References

- Andrade, K., K. Uchida, and S. Kagaya. 2006. "Development of Transport Mode Choice Model by Using Adaptive Neuro-Fuzzy Inference System." *Transportation Research Record: Journal of the Transportation Research Board* 1977: 8–16.
- Barua, S., N. Haque, A. Das, M. Hadiuzzaman, and S. Hossain. 2017. "Automatic Multi-Regime Fundamental Diagram Calibration Using Likelihood Estimation." *International Journal for Traffic and Transportation Engineering* 7 (1): 79–92.
- Chandler, R. E., R. Herman, and E. W. Montroll. 1958. "Traffic Dynamics: Studies in car Following." *Operations Research* 6: 165–184.
- Daganzo, C. F. 1995. "A Finite Difference Approximation of the Kinematic Wave Model of Traffic Flow." *Transportation Research Part B: Methodological* 29 (4): 261–276.
- del Castillo, J. M. 2001. "Propagation of Perturbations in Dense Traffic Flow: a Model and its Implications." *Transportation Research Part B: Methodological* 35 (4): 367–389.
- Drake, J. S., J. L. Schofer, and A. D. May. 1967. "A Statistical Analysis of Speed–Density Hypotheses." *Highway Research Record* 156: 53–87.
- Drew, D. R. 1968. *Traffic Flow Theory and Control*. New York: McGraw Hill.
- Edie, L. C. 1961. "Car-following and Steady-State Theory for Noncongested Traffic." *Operations Research* 9 (1): 66–76.
- Fan, S., and B. Seibold. 2013. "Data-Fitted First-Order Traffic Models and Their Second-Order Generalizations." *Transportation Research Record: Journal of the Transportation Research Board* 2391: 32–43.
- Gazis, D. C., R. Herman, and R. B. Potts. 1959. "Car-following Theory of Steady-State Traffic Flow." *Operations Research* 7 (4): 499–505.
- Gazis, D. C., R. Herman, and R. W. Rothery. 1961. "Nonlinear Follow-the-Leader Models of Traffic Flow." *Operations Research* 9 (4): 545–567.

- Greenberg, H. 1959. "An Analysis of Traffic Flow." *Operations Research* 7 (1): 79–85.
- Greenshields, B. D. 1935. "A Study in Highway Capacity." *Highway Research Board Proceedings* 14: 448–477.
- Hall, F. L., V. F. Hurdle, and J. H. Banks. 1992. "A Synthesis on Recent Work on the Nature of Speed Flow and Flow-Occupancy (or Density) Relationships on Freeways." *Transportation Research Record: Journal of the Transportation Research Board* 1365: 12–18.
- Jabari, S. E., J. Zheng, and H. X. Liu. 2014. "A Probabilistic Stationary Speed–Density Relation Based on Newell's Simplified car-Following Model." *Transportation Research Part B: Methodological* 68: 205–223.
- Jang, J.-S. R. 1991. "Fuzzy Modeling Using Generalized Neural Networks and Kalman Filter Algorithm." In Proceedings of the ninth national conference on artificial intelligence (AAAI-gl), 762–767.
- Jang, J.-S. R. 1993. "ANFIS: Adaptive-Network-Based Fuzzy Inference System." *IEEE Transactions on Systems, Man, and Cybernetics* 23 (3): 665–685.
- Jang, J.-S. R., and N. Gulley. 1994. "Gain Scheduling Based Fuzzy Controller Design." Fuzzy information processing society biannual conference, 101–105.
- Jang, J.-S. R., and C.-T. Sun. 1995. "Neuro-fuzzy Modeling and Control." *Proceedings of the IEEE* 83 (3): 378–406.
- Kaehler, S. D. 1998. *Fuzzy logic tutoria – An introduction*. Encoder: The Newsletter of the Seattle Robotics Society. Accessed April 2, 2017. <http://www.seattlerobotics.org/encoder/Mar98/fuz/flindex.html>.
- Kerner, B. S. 2004. *The Physics of Traffic*. Berlin, NY: Springer.
- Kometani, E., and T. Sasaki. 1961. *Dynamic Behavior of Traffic with a Nonlinear Spacing-Speed Relationship*. Amsterdam: Elsevier.
- Kotsialos, A., M. Papageorgiou, C. Diakaki, Y. Pavlis, and F. Middelham. 2002. "Traffic Flow Modeling of Large-Scale Motorway Networks Using the Macroscopic Modeling Tool METANET." *IEEE Transactions on Intelligent Transportation Systems* 3 (4): 282–292.
- MacNicholas, M. J. 2008. "A Simple and Pragmatic Representation of Traffic Flow." Symposium on The Fundamental Diagram: 75 years, Transportation Research Board, Woods Hole, MA.
- Mahnke, R., and J. Kaupužs. 1999. "Stochastic Theory of Freeway Traffic." *Physical Review E* 59 (1): 117–125.
- May, A. D. 1990. *Traffic Flow Fundamentals*. Prentice Hall.
- Mehran, B., M. Kuwahara, and F. Naznin. 2012. "Implementing Kinematic Wave Theory to Reconstruct Vehicle Trajectories From Fixed and Probe Sensor Data." *Transportation Research Part C: Emerging Technologies* 20 (1): 144–163.
- Messmer, A., and M. Papageorgiou. 1990. "METANET: A Macroscopic Simulation Program for Motorway Networks." *Traffic Engineering & Control* 31 (9): 466–470.
- Mita, H. S., J. B. Kreer, and L. S. Yuan. 1969. "Dual Mode Behavior of Freeway Traffic." *Highway Research Record* 279: 1–12.
- Muralidharan, A., G. Dervisoglu, and R. Horowitz. 2011. "Probabilistic Graphical Models of Fundamental Diagram Parameters for Simulations of Freeway Traffic." *Transportation Research Record: Journal of the Transportation Research Board* 2249: 78–85.
- Newman, L. 1963. "Traffic Operation at two Interchanges in California." *Highway Research Record* 27: 14–43.
- Ni, D., J. D. Leonard, C. Jia, and J. Wang. 2015. "Vehicle Longitudinal Control and Traffic Stream Modeling." *Transportation Science* 50 (3): 1016–1031.
- Pahlavani, P., and M. R. Delavar. 2014. "Multi-Criteria Route Planning Based on a Driver's Preferences in Multi-Criteria Route Selection." *Transportation Research Part C: Emerging Technologies* 40: 14–35.
- Petković, D., M. Issa, N. D. Pavlovic, N. T. Pavlovic, and L. Zentner. 2012. "Adaptive Neuro-Fuzzy Estimation of Conductive Silicone Rubber Mechanical Properties." *Expert Systems with Applications* 39 (10): 9477–9482.
- Pipes, L. A. 1967. "Car Following Models and the Fundamental Diagram of Road Traffic." *Transportation Research* 1: 21–29.

- Qu, X., S. Wang, and J. Zhang. 2015. "On the Fundamental Diagram for Freeway Traffic: a Novel Calibration Approach for Single-Regime Models." *Transportation Research Part B: Methodological* 73: 91–102.
- Qu, X., J. Zhang, and S. Wang. 2017. "On the Stochastic Fundamental Diagram for Freeway Traffic: Model Development, Analytical Properties, Validation, and Extensive Applications." *Transportation Research Part B: Methodological* 104: 256–271.
- Quandt, R. E. 1958. "The Estimation of the Parameters of a Linear Regression System Obeying two Separate Regimes." *Journal of the American Statistical Association* 53 (284): 873–880.
- Sun, L., and J. Zhou. 2005. "Development of Multiregime Speed–Density Relationship by Cluster Analysis." *Transportation Research Record: Journal of the Transportation Research Board* 1934: 64–71.
- Teodorovic, D., and K. Vukadinovic. 2012. *Traffic Control and Transport Planning: a Fuzzy Sets and Neural Networks Approach*. Boston, NY: Springer Science & Business Media.
- Transportation Research Board. 2010. *Highway Capacity Manual*. Washington, DC: Transportation Research Board.
- Underwood, R. T. 1961. *Speed, Volume, and Density Relationship: Quality and Theory of Traffic Flow*. New Haven, CT: Yale Bureau of Highway Traffic.
- Wang, H., J. Li, Q. Y. Chen, and D. Ni. 2009. "Speed–density Relationship: From Deterministic to Stochastic." In *Transportation Research Board 88th Annual Meeting*, 10.
- Wang, H., J. Li, Q. Y. Chen, and D. Ni. 2011. "Logistic Modeling of the Equilibrium Speed–Density Relationship." *Transportation Research Part A: Policy and Practice* 45 (6): 554–566.