2023 8th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS) | 978-1-6654-5530-5/23/$31.00 ©2023 IEEE | DOI: 10.1109/MT-ITS56129.2023.10241428

Application of a Mamdani-Based Fuzzy Traffic State Identifier to a Real Case Study

Gizem Erdinç
Dept. of Civil, Constructional and Environmental Engineering
Sapienza University of Rome
Rome, Italy
gizem.erdinc@uniroma1.it

Chiara Colombaroni
Dept. of Civil, Constructional and Environmental Engineering
Sapienza University of Rome
Rome, Italy
chiara.colombaroni@uniroma1.it

Gaetano Fusco
Dept. of Civil, Constructional and Environmental Engineering
Sapienza University of Rome
Rome, Italy
gaetano.fusco@uniroma1.it

***Abstract*— This paper presents a fuzzy-logic application based on the Mamdani inference method to get the range of road traffic conditions. It was tested with real data extracted from the Padua-Venice motorway in Italy, which contains a dense network of monitoring that provide continuous measurements of flow, occupancy, and speed. The empirical results show that the proposed study functions well in qualitative classification. The experiment can provide another perspective on motorway traffic control.**

Keywords—traffic state identification, fuzzy logic, congestion level, Mamdani inference

I. INTRODUCTION

Estimating traffic congestion has become one of the major problems in transportation engineering. In the related literature, there are many congestion detection techniques that are based on statistical and data-driven approaches. However, describing the level of traffic is complicated because of the inherent random components in the phenomenon and the high instability observed when approaching or exceeding critical conditions. So, it is difficult to find a quantitative deterministic model that satisfactorily fits them. However, it has been argued that a fuzzy qualitative definition may be an appropriate approach to solve traffic and transportation problems with more than forty years of experience [1-4].

Among the recent papers, Hamad and Kikuchi [5] proposed fuzzy inference approaches to evaluate congestion levels by considering speed as the input variable. They used travel speed, free-flow speed, and the proportion of very low speed in the total travel time as input variables to compute the congestion. In [6], Kikuchi and Chakroborty studied a fuzzy approach to handle the uncertainty embedded in the definition of the level of service (LOS). They criticized the current HCM procedure arguing that it does not accurately represent the notion of LOS as a user-perceived measure and questioned whether one measure (e.g., density) can capture all the factors that affect LOS. Thus, they provided a framework that handles uncertainty under the different paradigms: deterministic, probabilistic, or possibilistic. Huang et al. [7] proposed a fuzzy C-Means clustering method to classify traffic data of flow, speed, and occupancy into four traffic states and exemplified this method through a small example on an urban road equipped with two detectors. With the trust of those experiences, a Mamdani-based fuzzy method has been applied to classify traffic state in this study. The idea of detecting traffic congestion with the Mamdani approach is already studied and proven to be effective in previous reference works. Amini et al. [8] developed a fuzzy inference model aimed at predicting the level of congestion in road networks where the paucity of accurate and real-time data can cause problems in using conventional quantitative techniques and took as inputs parameters length, number of lanes, and flow inputs to get the level of congestion output. However, the experiment was based on only a one-week period of data. Kalinic and Krips presented a model containing two inputs (flow and density), as in most classical studies on traffic study [9]. However, the congestion levels provided as outputs depend on a combination of the inputs as independent variables and do not catch the non-linear relationship between the two traffic flow variables. Toan and Wong [10] applied a fuzzy-based methodology for the quantification of congestion level using density and speed and argued that a method based on two variables tends to neutralize in between and scale up in a more stable manner with the levels of service compared to measurements based entirely on density like in the classification of the Highway Capacity Manual.

The present paper is founded on the belief that the fuzzy approach is the most suitable to catch the inherent uncertainty under the drivers' perception of traffic congestion and that the simultaneous observation of the three fundamental quantities of traffic can improve the identification of critical or hypercritical conditions that are characterized by high and rapid variability. Thus, this article presents a Mamdani-based fuzzy model that uses flow, density, and speed values to identify the messages on traffic congestion to provide drivers with information. The proposed model is trained on a large set of traffic data collected from a network of 3 motorways in Northern Italy over more than 8 months.

The present paper is structured as follows. After this Introduction, the Second Section discusses the traffic flow characteristics and defines the traffic state classification. The Third Section introduces the methodology for classification, composed of the variable definition, variable fuzzification, and formation of rules. The Fourth Section describes the case study and discusses the results. The Conclusion resumes the main assumptions of the methodology implemented and the outcomes produced by the method.

II. TRAFFIC STATE CLASSIFICATION

A. Traffic Flow Characteristics

In the traffic theory, three parameters are usually used to describe the traffic flow characteristics; these are the flow (or volume) (q = vehicle/h), speed (v = km/h), and traffic density (k = vehicle/km). The relationship among the three parameters at the stationary state is expressed formally with Equation 1:

$$q = v \cdot k \tag{1}$$

While a traffic state is defined as a given tuple of values of the fundamental variables, the road congestion definition is a vague concept, so it is not easy for a traveler information system to provide clear advice to users. Until now there is still no unified regulation [7]. Nevertheless, a huge mass of traffic observations exists, and the traffic phenomenon is very well known. The following example is presented with the only aim to provide an illustrative explanation of the concepts underlying the fuzzy approach introduced in the paper. As well-known, in the flow-density plane direct traffic measurements illustrate two different trends (Fig.1a): in the range of small values of density, the flow is almost linear with density and presents limited deviations from the average speed; however, the flow exhibits a very noisy and sparse pattern in the range of higher values of density with an average decreasing trend as the density increases. It is well known that the high-density regime is characterized by unstable flow conditions determined by the microscopic mechanism underlying the traffic stream, consisting of even slightly irregular driving maneuvers that lead to a stop-and-go regime.

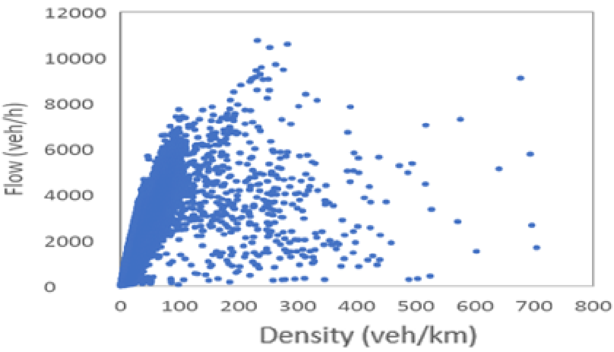

Fig. 1(a). Flow-Density measures

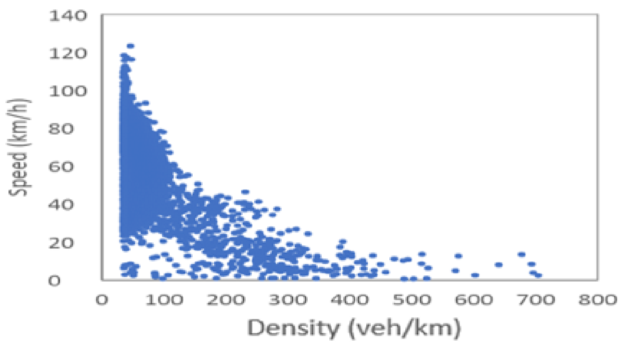

Fig. 1(b). Speed-Density measures

Such an unstable regime produces highly noisy flow-density patterns that change rapidly in a dynamic context. To explain the empirical spatiotemporal features of traffic breakdown and the resulting traffic congestion, multidimensional models were introduced. Kerner and Rehborn [11] proposed a three-phase traffic theory based on the observation of a synchronized traffic phase on multilane motorways. Persaud and Hall proposed a three-variable model based on the catastrophe theory to explain the transitions to and from congested operations upstream of incidents [12].

In a two-dimensional plane, the observation of speed-density measures highlights in a clear way the decreasing trend of speed with density (Fig.1b) and explains the decrease of flow in a high-density regime by virtue of equation (1).

B. *Traffic State Division*

Many studies faced the problem of congestion identification under different approaches and for different purposes. While applications devoted to planning like the US Highway Capacity Manual [13], and management like the EU DATEX [14], consider five traffic states corresponding to many levels of services, incident detection algorithms focus on the simplest distinction between congestion flow and non-congestion. Theoretical studies considered different traffic classes ranging from two [15], three [16], four [17], to five [18].

In our view, as we evaluate the states according to the EU DATEX standard, we divided the range of traffic conditions into five features as given in Table I.

According to the recent observations that showed that three-dimensional models are more appropriate to characterize congested traffic states, a fuzzy model based on speed, flow, and density observations for congestion identification is introduced in the following. Traffic conditions are divided into different states that describe the change from free flow to congested conditions, with particular attention to the critical point where the curve inverts its trend, and the flow reaches its capacity. Table I summarizes qualitative parameter definitions with respect to possible traffic states.

TABLE I. TRAFFIC STATE CLASSIFICATION SUMMARY

State	**q**	**k**	**v**	**Traffic condition index [0-1]**
Smooth	Very low	Very low	Very High	Very low (0.00-0.15)
Intense	Low	Medium	High	Low (0.15-0.35)
Slow	Medium	Medium-High	Low	Medium (0.35-0.65)
Queuing	High	High	Low	High (0.65-0.9)
Stationary	Low	Very high	Very low	Very high (0.9-1.00)

III. METHODOLOGY

A. *Variable Definition*

In this Mamdani method, we define three input variables (flow, density, and speed) and one output parameter (traffic condition) (Fig. 2).

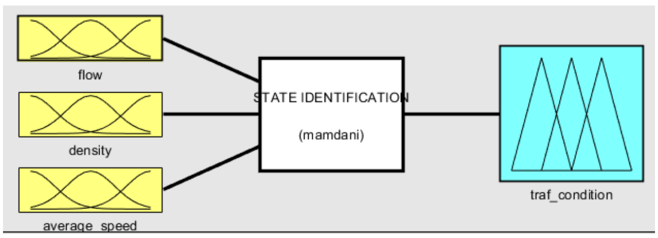

Fig. 2. The fuzzy inference system

The input variable flow is assigned with the following linguistic variables: Free Flow (FF), Reasonably Free Flow (RFF), Average Flow (AF), Congested Flow (CF), and Very Congested Flow (VCF). The input variable density is clustered as: Very Low Density (VLD), Low Density (LD), Medium Density (MD), High Density (HD), Very High Density (VHD). The input variable speed is defined as: Very Slow (VS), Slow (S), Average (A), Fast (F), and Very Fast (VF). The output variable Traffic state (calculated separately for each road section) is also classified with five linguistic variables: Stationary (S), Queuing (Q), Slow (SL), Intense (I), Smooth (SM).

B. *Variable Fuzzification*

After input and output definitions, all variables have been fuzzified by transferring the crisp numerical values of the selected input variables, through membership functions into membership degrees of the fuzzy set. Membership degrees quantify the belongingness of the variable value to the fuzzy set. Even though there are various membership functions commonly used, in this paper triangular membership functions as given in Equation 2 are used since they capture the characteristics of the case study's fuzzy set and it's one of the most used examples.

$$\mu(x) = \begin{cases} 0, x < \alpha min \text{ or } x > \alpha max \\ \frac{x-\alpha min}{\beta-\alpha min}, \ x \in (\alpha min, \beta) \\ \frac{\alpha max - x}{\alpha max - \beta}, \ x \in (\beta, \alpha max) \end{cases} \quad (2)$$

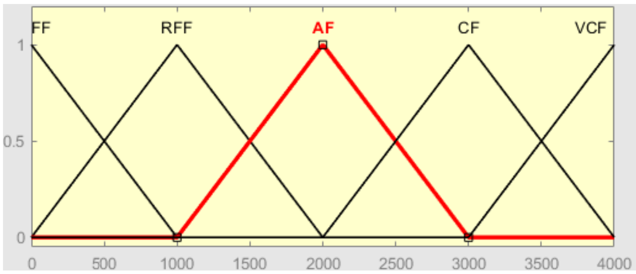

Fig. 3(a). µ of flow in the range between [0-4000]

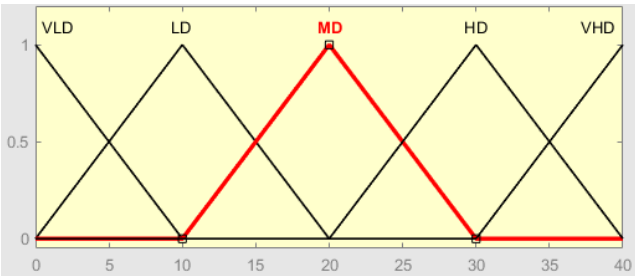

Fig. 3(b). µ of density in the range between [0-40]

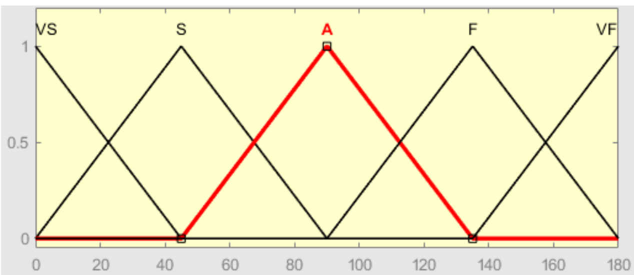

Fig. 3(c). µ of speed in the range between [0-180]

Fig. (3a, 3b, 3c) show the membership function plots of flow, density, and speed input variables respectively. In each plot, the horizontal plane shows the value range of the related variables, while the vertical plane gives the membership function value of it in the range [0-1]. The range intervals of the related variables are set after a close investigation accordingly to the data. It is to note that the terms are imprecisely defined following the fuzzy logic concept, and there are no clear cuts between the fuzzy congestion levels. Membership grade represents the possibility of occurrence that is primarily governed by a subjective degree of belief. So, the overlaps between the fuzzy sets are designed such that the sum of membership grades for adjacent fuzzy sets at any point in the overlapping sections equals one. It follows that the membership grade of a particular fuzzy set approaches zero as that of the adjacent fuzzy set approaches one.

C. *Formation of Rules*

As the core of the method, the input-output relationship should be modeled to build the inference and a nonlinear surface model with specific rules, it indicates how to project input variables onto output space. In this study, all stable and unstable feasible solutions which create traffic states are modeled with different weights using IF-THEN rules (some of them in Table II). The rules are weighted according to the frequency of occurrence; a bigger weight is assigned to rules that occur in dense areas and a smaller weight to the rules that occur infrequently. For this section, seventeen rules have been defined. Usually, the inputs of the fuzzy model are defined with more than one fuzzy set, to combine these membership values and obtain unique results, linguistic information (such as free flow and medium density, and slow speed) relates to the AND operator meaning that a minimum condition must be met for conditional if statement to be fulfilled. The AND operator is one of the most used operators in Fuzzy modeling.

TABLE II. IF-THEN RULES MODELLED

IF	THEN
If Flow is FF	Traffic is Smooth
If Density is VHD	Traffic is Stationary
If Flow is CF and Density is HD	Traffic is Queuing
If Flow is AF and Density is MD	Traffic is Slow
If Flow is CF and Density is VHD and Speed is AS	Traffic is Queuing
If Flow is FF and Density is LD and Speed is FS	Traffic is Intense

The consequent part of the 'IF–THEN' rule is another fuzzy linguistic set defined by the corresponding membership function, the output of each IF–THEN rule is a fuzzy set. To elaborate, aggregation is the process where fuzzy sets that represent the outputs of each rule are combined into a single fuzzy set, this can be done with an operator. The MAX operator is one of the most used operators for this process. After the aggregation process, generated fuzzy sets for each output variable should be de-fuzzy. Among several existing methods centroid method (it finds the centroid of a two-dimensional function) proposed by [9] and [19] is the most applied. Table III gives the average traffic states of each section as the results of the first application in the peak hour in the morning all over the observation period.

TABLE III. IDENTIFIED CONDITIONS

Section number	Road section	Average of the Observed Intervals of			Traffic State
id	id	Flow (pc/lane/hour)	Density (pc/lane/km)	Speed (km/hour)	
41-1	741467	1454	16	97	Slow
42-2	740855	1409	14	107	Intense
42-1	741527	1400	14	108	Intense
43-1	741555	1369	15	93	Slow
44-2	740700	1145	14	106	Intense
44-1	741682	993	12	103	Intense
45-1	741708	1359	14	100	Intense
46-1	741741	930	10	98	Intense

IV. CASE STUDY

In this paper, the proposed model is applied to classify traffic conditions of the motorway between Padua and Venice in Italy using flow, density, and speed fundamental variables on the Mamdani-based fuzzy logic application. In the literature, this approach was applied with only density and flow traffic flow characteristics [8, 9] and the level of congestion was related to the variable density. Differently from those studies, we believe that including speed values can improve the effectiveness of the method because it can better characterize unstable conditions when fundamental variables are affected by rapid changes that violate the relationships that hold in stationary conditions.

The Padua-Venice motorway network comprises 4 branches of 3-lane roads with separate carriageways. It has a total length of about 74 km and includes 16 intersections. Real traffic data collected from the 31st of December 2018 to the ^{3}0th of August 2019 contain the following information: local unit code, code section counting, day type, road (section) id, date, flow, occupancy, and harmonic speed, collected every 1 minute and aggregated every 15 minutes. Six days for each month were selected at random for this analysis. Density values were estimated by applying the state equation (1).

While Fig.4 shows the identification codes of the sections Fig.5 demonstrates monitoring sections of the related segments for the study network. The whole dataset was examined statistically section-by-section; on the other hand, it is important to acknowledge that due to temporary failures of some detectors, some parts of the data are missing. In this study, we studied the congestion for the time interval between 6:30 am and 9 am. The application is run on MATLAB fuzzy logic toolbox R2020b.

The scope of the proposed model is to simulate the general state of congestion of the road and derive a relationship between the three fundamental variables by applying the fuzzy Mamdani inference approach regardless of their values [9]. Since each section could have a different general state, they have been modeled and run separately. The network generally composes two main branches in the North (Highway) and South (Tangential). In this paper, we focus on the only Northern branch. As a given example, Table III presents the average situation of several road segments in the observed 8-month of period from Monday to Saturday between 6:30-9:00 am (in the peak hour in the morning) with the average of the observed data of variables.

The road is in the transition between the *slow-intense* situation in general (Fig. 6) with a smooth shifting into slow traffic on sections 43, and the opposite way at sections 42 and 44.

In order to see the effect of the selected membership function of the proposed fuzzy model a sensitivity analysis has been done in terms of traffic condition range index between [0-1]. For this analysis, five cases with triangular, trapezoidal and Gauss, gbell, and Gauss2 function shapes have been carried out. Results are given in Table IV.

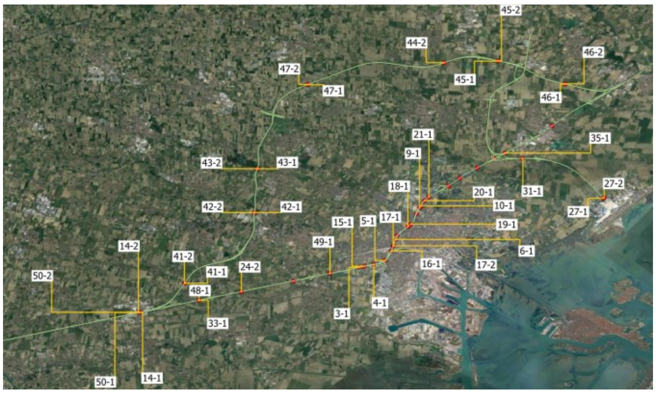

Fig. 4. Identification codes of the sections

Fig. 5. Traffic monitoring sections

Fig 6. Average traffic situation on the highway

TABLE IV. IDENTIFIED CONDITIONS AND TRAFFIC STATES

Section number	**Identified Condition Index [0-1] and Traffic States Accordingly**									
id	**Triangular shape membership function**	**Traffic State**	**Trapezoidal shape membership function**	**Traffic State**	**Gauss shape membership function**	**Traffic State**	**Gbell shape membership function**	**Traffic State**	**Gauss2 shape membership function**	**Traffic State**
41-1	0.3696	Slow	0.3582	Slow	0.3708	Slow	0.3997	Slow	0.3857	Slow
42-2	0.3438	Intense	0.2508	Intense	0.3375	Intense	0.3317	Intense	0.2871	Intense
42-1	0.3442	Intense	0.2508	Intense	0.338	Intense	0.3327	Intense	0.2875	Intense
43-1	0.3501	Slow	0.3104	Intense	0.3445	Intense	0.3435	Intense	0.3034	Intense
44-2	0.3014	Intense	0.2508	Intense	0.3008	Intense	0.2898	Intense	0.2689	Intense
44-1	0.2499	Intense	0.25	Intense	0.2766	Intense	0.2772	Intense	0.2688	Intense
45-1	0.3409	Intense	0.2508	Intense	0.3337	Intense	0.3284	Intense	0.3506	Slow
46-1	0.2489	Intense	0.25	Intense	0.2674	Intense	0.2741	Intense	0.2688	Intense

The proposed model has close condition range indexes with all membership functions. The reason for this is that we kept fixed the numerical ranges of both input and output variables. It is noteworthy that there are no clear cuts between the fuzzy traffic condition ranges. Membership grade represents the possibility of occurrence that is primarily governed by a subjective degree of belief. In this regard, to determine the effect of membership function type we need a benchmark.

For that reason, we compared our results to LOS in HCM [13] and DATEX II standards [14]. In [10, 13], when density is between 7-11 pc/lane/km and speed is lower than 120 km/h, the road has a B level of service. Similarly in our interpretation (Table I), in a condition of *low-density* but *high-speed level*, traffic starts to be getting in an intense state, but drivers can drive freely. According to this, Section 46-1 has been identified as in an *intense* state and all membership functions can get it with close condition indexes for this state. However, when density goes up to the level of 11-16 pc/lane/km, which is the dominant situation for the rest of the network, the level of service gets down at the C level.

The problem with this level is having significant overlaps in the relationship between the condition index and the LOSs [10], and it makes it confusing to classify the situation. In this regard, to be clearer in the separation of intense-slow states, we took the speed variable as a guide as in [14]. According to this, if the average speed is between 25%-75% of its free-flow level (which is 130 km/h in this study), then the traffic moves in a slow state. Under this logic, for Section 45-1, Gauss2 shape function has over calculated the index value and it considered the road as in a *slow* state, while it should be put in the *intense* class as given by all other functions. Last but not least, the states of sections 41-1 and 43-1 have been identified as *slow* (as in Table III) with 97 km/h and 93 km/h speed levels respectively. While all functions could catch it for Section 41-1, only the triangular membership function is good to be able to reflect it for Section 43-1.

V. CONCLUSION

In this paper, we handled the traffic state identification problem and proposed a fuzzy logic-based application. In the application, we aimed to give another implementation into the related literature by using more input variables and a much bigger real dataset. We know that traffic state identification with a limited variables such as only density or the detected speed value may be inadequate. Also, each traffic condition has a level of similarity, which makes traffic state classification fuzzy. That's why we use a qualitative approach -fuzziness-. We believe that the fuzzy methodology seems to be more appropriate to provide qualitative information to drivers than a simple deterministic threshold-based method. In the future works, apart from the analysis of the Tangential part of the network, we will consider other aspects as road quality or weather conditions.

REFERENCES

[1] C. P. Pappis, and E. H. Mamdani, “A Fuzzy Logic Controller for a Traffic Junction,” IEEE Trans. Syst. Man Cybern., vol. 7, no. 10, pp. 707–717, 1977.

[2] M. Sugeno, and M. Nishida, “Fuzzy Control of Model Car. In: Fuzzy Sets and Systems.” North-Holland 16 (2): 103–113, 1985, doi:10.1016/S0165-0114(85)80011-7.

[3] T. Sasaki, and T. Akiyama. “Fuzzy On-ramp Control Model on Urban Expressway and Its Extension.” In Transportation and Traffic Theory, edited by N. Gartner and N. Wilson, 377– 395, New York: Elsevier Science, 1987.

[4] P. Chakroborty, and S. Kikuchi, “Examination of Fuzzy Notion of Levels of Service: Case of Freeway Traffic Flow.” In Proc., Sixth World Conference on Transport Research, Lyon, France, Vol. 2, pp. 1047–1057, 1992.

[5] K. Hamad, and S. Kikuchi, “Developing a measure of traffic congestion - fuzzy inference approach”, Transp. Res. Rec.: J. Transp. Res. Board, 77–85, Paper No. 02-2770, 2002.

[6] S. Kikuchi, and P. Chakroborty, “Frameworks to Represent Uncertainty when Level of Service is Determined”, Transportation Research Record, 1968(1), 53–62, 2006.

[7] Y. Huang, L. Xu, Q. Luo, “The Method of Traffic State Identification and Travel Time Prediction on Urban Expressway”, Research Journal of Applied Sciences, Engineering and Technology 5(4): 1271-1277, 2013.

[8] M. Amini, M. Hatwagner, G. Mikulai, L. Koczy, “An intelligent traffic congestion detection approach based on fuzzy inference system”, SACI 2021, IEEE 15th International Symposium on Applied Computational Intelligence and Informatics, May 19-21, Timişoara, Romania, 2021.

[9] M. Kalinic and J.M. Krisp, “Fuzzy inference approach in traffic congestion detection”, Annals of GIS, 2019. https://doi.org/10.1080/19475683.2019.1675760.

[10] T. D. Toan, and Y.D. Wong, Fuzzy logic-based methodology for quantification of traffic congestion. Physica A—Statistical Mechanics and its Applications, 570, 125784, 2021.

[11] B. S. Kerner, and H. Rehborn, "Experimental properties of phase transitions in traffic flow". Physical Review Letters, 79(20), 4030, 1997.

[12] B. N. Persaud, and F.L. Hall, "Catastrophe theory and patterns in 30-second freeway traffic data—Implications for incident detection". Transportation Research Part A: General, 23(2), 103-113, 1989.

[13] National Academies of Sciences, Engineering, and Medicine, Highway Capacity Manual 7th Edition: A Guide for Multimodal Mobility Analysis. Washington, DC: The National Academies Press, 2022.

[14] DATEX II. Version 3.2. https://docs.datex2.eu/situation/index.html.

[15] G. F. Newell, "A simplified car-following theory: a lower order model" Transportation Research Part B: Methodological, 36(3), 195-205, 2002.

[16] M. Ashifuddin Mondal AND Z. Rehena, "Intelligent traffic congestion classification system using artificial neural network", In Companion Proceedings of The 2019 World Wide Web Conference, pp. 110-116, 2019.

[17] X. Lin, "A road network traffic state identification method based on macroscopic fundamental diagram and spectral clustering and support vector machine", Mathematical Problems in Engineering, 2019.

[18] D. Helbing, M. Treiber, A. Kesting, M. Schönhof, "Theoretical vs. empirical classification and prediction of congested traffic states". The European Physical Journal B, 69(4), 583-598, 2009.

[19] C. C. Lee, "Fuzzy Logic in Control Systems: Fuzzy Logic Controller-Part I and Part H." IEEE Transactions on Systems Man and Cybernetics 20 (2): 404–435, 1990.