A vehicular traffic congestion predictor system using Mamdani fuzzy inference

Mehran Amini
Department of Information Technology
Szechenyi Istvan University
Gyor, Hungary
mehran@sze.hu

Miklos F. Hatwagner
Department of Information Technology
Szechenyi Istvan University
Gyor, Hungary
miklos.hatwagner@sze.hu

Gergely Cs. Mikulai
Doctoral School of Regional Sciences and Business Administration
Budapest, Hungary
mikulai.gergely@ga.sze.hu

Laszlo T. Koczy
Department of Information Technology
Szechenyi Istvan University
Gyor, Hungary
Department of Telecommunications and Media Informatics
Budapest, Hungary
koczy@tmit.bme.hu

***Abstract*— The process of traffic control systems significantly relies on the immediate detection of breakdown states. As a result of their crisp (non-fuzzy) based calculation procedures, conventional traffic estimators and predictors cannot effectively model traffic states. In fact, these methods are characterized by exact features, while traffic is defined by uncertain variables with vague properties. Furthermore, typical numerical methodologies have constraints on evaluating the overall system status in heterogeneous and convoluted networks mainly due to the absence of reliable and real-time data. This study develops a fuzzy inference system that uses data from the Hungarian freeway networks for predicting the severity of congestion in this complex network. Congestion severity is considered the output variable, and traffic flow along with the length and the number of lanes of each section are assigned as input variables. Seventy-five fuzzy production rules were generated using accessible datasets, percentile distribution, and experts' consensus. The Matlab Fuzzy Logic Toolbox simulates the designed model and analysis steps. According to available resources, the results demonstrate linkages among input variables. Analyses are also used to construct intelligent traffic modeling systems and further service-related planning.**

Keywords—Fuzzy inference, intelligent transportation, congestion prediction

I. INTRODUCTION

Intelligent systems offer a systematic, structured approach for addressing critical and complex problems and generating consistent and reliable solutions over the practice. From a computational standpoint, intelligent systems have two main characteristics; dealing with complex issues in real-world circumstances and finding the same answers and results for identical categories of problems by a computational analysis model [1], [2]. Traffic is one of these problems in which a combination of involved elements and their interactions turn it into one the most convoluted systems [3] [4]. This complexity might be addressed by improving existing parameters, which lead to expanding the road capacity (e.g., increasing the number of the road lanes) or utilizing intelligent techniques to minimize breakdown and congestion severity more cost-effectively [5]. However, [6] showed that increasing the road capacity will not minimize traffic but cause more congestion. One effective method to control traffic is to establish an efficient and accurate system, which can assist better in allocating transportation resources and dispersing flow before traffic breakdown. The Intelligent Transportation System (ITS) is the most frequently used one among such systems [7], [8] which have been presented in traffic-related problems [9], [10] to deal with them more efficiently by new data inferencing and communication tools [11].

A range of modern technologies, such as transportation communication systems, are integrated with ITS [12]. Furthermore, by leveraging the development of 5G based communication and numerous on-road sensors, ITS may enhance traffic efficiency, reduce congestion severity, boost road capacity, and minimize pollution and traffic accidents [13], [14], [15]. As a critical component of ITS, an effective and precise road traffic modeling is needed to offer continuous and exact road status information based on historical freeway data [16]. Estimating and predicting the level of traffic in a specific section of the freeway is among the common attributes of such models. These characteristics are frequently based on standard mathematical methods, e.g., statistical regression, which is incapable of coping with the complexities of freeway traffic dynamics and connections among its elements. Moreover, engaged variables are primarily connected with uncertain characteristics since they are directly affected by factors relating to human behavior and decisions; as a result, these variables and the level of uncertainty in their quantities must be considered in the computation of the system.

Using human knowledge to tackle problems that require human intelligence is the crucial concept of intelligent systems [17]. The fuzzy inference system is a verified model of human knowledge, with applications in a large variety of domains [18]. The main reason for the diversity of applications is due to the features and functionality of fuzzy rule-based techniques. For example, because of their capabilities to assign connections among input and output parameters in nonlinear and dynamic systems, and more

Cite as: M. Amini, M. F. Hatwagner, G. C. . Mikulai, and L. T. Koczy, “A vehicular traffic congestion predictor system using Mamdani fuzzy inference”, *Syst. Theor. Control Comput. J.,* vol. 1, no. 2, pp. 49–57, Dec. 2021.
DOI: 10.52846/stccj.2021.1.2.27

importantly, to address uncertainty by offering linguistic variables and graded membership [19]. As an extension of the research originally presented at the SACI 2021 conference [20], this study attempts to use fuzzy techniques' characteristics to predict congestion levels in a network of freeways with uncertain observations and causal relationships. Although prior literature has investigated many traffic engineering related problems, this study developed a Mamdani type fuzzy rule-based inference system in preference to employing conventional mathematical techniques in order to model the available variables in predicting the severity of congestion. In the developed model, congestion severity is considered the output variable, and traffic flow values along with the lengths and the number of lanes of each section are selected as input variables. The generated model is simulated and validated by the Matlab Fuzzy Logic Toolbox R2020b.

Accordingly, this study is aimed to demonstrate the efficiency of fuzzy inference in modeling an immensely heterogeneous and convoluted system to predict the level of congestion based on data collected from the Hungarian network of freeways. The following part discusses several traffic-based control methods. The third part explains the procedures for developing fuzzy inference-based systems, focusing on Mamdani-inference. Later, the modeling processes and dataset description are explained before performing the findings and conclusion.

II. Traffic Congestion Modeling

Reducing traffic congestion is considered a vital part of traffic control-related strategies since its role is undeniable in the sustainability and development of transportation and mobility systems [10]. The dynamic nature of traffic behavior in transportation systems and defining involved parameters by precise relations were the main motivations of creating traffic flow models. Mathematical theories of traffic flow began in the early 1950s [11]. Defining connections and relations between human-involved parameters (e.g., drivers, pedestrians) and infrastructures associated parameters (e.g., freeways, number of lanes) is the central part of these theories, which aim to comprehend and evolve transport systems with effectual mobility and avoid traffic breakdown. Traffic breakdown is a crucial concept in the transportation systems analysis because it is regarded as a critical point where increasing traffic volume beyond it can lead from uncongested flow to congested state [21]. Adequate traffic mobility is reachable by considering three factors, continuous traffic flow, traffic monitoring, particularly at known traffic breakdown areas, and identifying and solving accident-related risk factors [22]. For dealing with mentioned factors, three parameters of traffic flow description are highlighted: velocity, density, and flow [5], [23]. These parameters are macroscopic traffic model variables where aggregate traffic parameters or the overall behavior of the traffic stream are modeled [24].

The necessity of using reliable congestion detection and prediction techniques mainly arose from recent advancements in intelligent transportation systems. These techniques are categorized on two primary levels: first, conventional methods formed on statistical approaches (e.g., autoregressive integrated moving average, Kalman filtering) joined with the flow and congestion-related parameters; second, data-driven methods employing machine learning algorithms (e.g., artificial neural network, support vector regression, and fuzzy-based computation). These methods are the most frequently applied techniques in the latest researches [25], [26]. Employing such techniques requires clarifying traffic congestion concepts. Although it has been investigated and developed in various aspects [27], demand–capacity equilibrium is a significant characteristic of congestion that needs to be considered. This category is a relative caliber of traffic flow or a proportion of the best possible condition of the freeway and current condition which any change in equilibrium between traffic flow and approximate capacity of the freeway can affect travel time, economic aspects, and variation of behavior. Approximation plays a significant role in all involved traffic measurements; this means that each involved parameter in the modeling of congestion in respect to the precision of its representation among real-world circumstances needs to be analyzed by a framework that can deal with ambiguity and uncertainty. Therefore, the computation of grading description of congestion levels necessitates being fuzzy (particularly, as the uncertainty involved here is in a large part of non-statistical nature). Fuzzy inference methods not only can be beneficial in determining the degree of congestion but also can homogeneously approximate and model every existing continuous nonlinear system to a subjective degree of exactness [28]. Describing the level of traffic is connected to uncertainty-associated properties. Already in his early work Zadeh illustrated that uncertain rational statements allow the development of methods that employ uncertain information to derive imprecise analysis [29], [30]. Among the first researches, [31] proposed the fuzzy inference-based method to deal with a specific problem of traffic congestion where a fuzzy-based controller was implemented in an intersection to compare the results with the conventional vehicle-actuated controller; consequently, performed analyses indicated that the fuzzy based controller had a better performance.

III. Fuzzy Inference-based Systems

Fuzzy set theory proposed by Zadeh [29] has been employed afterward to deal with numerous scientific and industrial problems in different fields of science and technology. Along with its capability of handling uncertain and imprecise information, since fuzzy theory provides a basis for applying expert supervised customizations in the form of If-Then rules, human knowledge has a central role in engineering and designing procedures [30]. The most significant part of this idea is overcoming and dissolving the crisp set limitations where individuals are dichotomized (divided into two sharply defined classes) as members and non-members. Handling these limitations is achieved by increasing the volume of acceptable and allowable uncertainty through sacrificing some of the accurate information in favor of an ambiguous but more robust represenatation [29]. The membership or non-membership of value x in the binary set A is assigned by function μ_A of A, illustrated by Eq. 1 [32]:

$$\mu_A(x) = \begin{cases} 1, & if\ x \in A \\ 0, & if\ x \notin A \end{cases} \tag{1}$$

As opposed to a crisp set in which a sharp and unambiguous distinction exists between the members and non-members, a fuzzy set allows ambiguity with the aim of reducing complexity by eliminating the sharp boundary separating members of the set from non-members. Therefore, a value (an element within a variable set) can partly be a member of a specific set. These values are computed with linguistic metaphors rather than numerical expressions [31]; an element is assigned to a class with membership function in closed interval 0 and 1; 1 expresses complete membership and 0 means non-membership. Membership function μ_A quantifies the degree of belonging of x to A. In Eq. 2, the fuzzy set A over the universe U is given [33], [34]:

$$A = \{x, \mu_A(x) | x \in U\} \quad (2)$$

These characteristics of fuzzy sets along with their possible representation as linguistic terms offer a computational algorithm for modeling and resolving imprecision and uncertainty associated problems, these latter being inseparable features of nonlinear and complex systems [35]. A fuzzy-rule based system is generally formed by four components containing fuzzification (determination of the degree of matching), knowledge base, fuzzy inference system, and defuzzification (fig. 1).

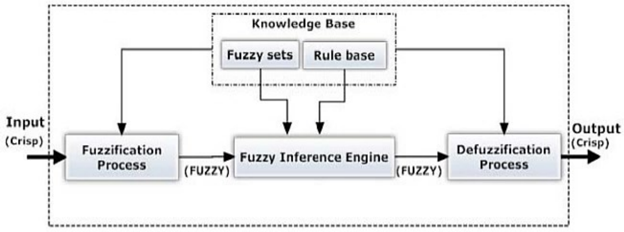


Fig. 1. Fuzzy Inference Architecture [18]

A. *Fuzzification*

The first step in forming a fuzzy inference system is fuzzification. As a mathematical process, it is required for transforming a value in the universe of discourse to a membership function of the fuzzy set [18]. It contains the procedure of converting crisp inputs into a degree of membership through membership functions for producing linguistic metaphors (e.g., high, very high). Membership functions are the pivotal players in the fuzzification step. In the practice, they have various sorts of linear and nonlinear forms (e.g., triangular, trapezoidal, Gaussian) where their types are chosen based on the context and the modeled problem and the experts' perspectives [36]. Triangular and trapezoidal membership functions as the most often employed types in fuzzification are given in Eq. (3) and (4), respectively [37]:

$$f(x; a, b, c) \Rightarrow \mu_A(x) = \max\left[\min\left[\frac{x-a}{b-a}, \frac{c-x}{c-b}\right], 0\right] \quad (3)$$

$$f(x; a, b, c, d) \Rightarrow \mu_A(x) = \max\left[\min\left[\frac{x-a}{b-a}, 1, \frac{d-x}{d-c}\right], 0\right] \quad (4)$$

where μ_A *(x)* is a membership function of a linguistic variable in the fuzzy set; *a, b, c,* and *d* are the parameters of the piecewise linear fuzzy membership functions.

B. *Knowledge base*

This component of the system contains two central units; a rule base unit including the fuzzy if-then rules of the system, and the system database unit that expresses the membership functions which are employed in the fuzzy rules. These rules are characterized either by experts' assessment or are determined by data analysis algorithms (machine learning, i.e., through clustering or evolutionary algorithms) [38]. The connections between inputs and output are assigned by fuzzy conditional functions, called fuzzy if-then (production) rules. Fuzzy conditional rules are formed by a premise or antecedent coupled with a result or consequent part. For instance, if *x* is high, then *y* is low, where 'high' and 'low' are expressed by membership functions [39]. In other words, if the observation matches with some premises, the respective rules are „fired" in the inference system, and they participate to some degree in making the decision [40]. In fuzzy models, every rule is expressed by a relation which is stated as:

$$\mu_{Ri}(x, y) = I(\mu_{Ai}(x), \mu_{Bi}(y)), \quad i = 1, 2, \ldots, n \quad (5)$$

where membership degree of rule *i* in connection to inputs *x* and *y* is shown as $\mu_{Ri}(x, y)$; $\mu_{Ai}(x)$ is the membership degree of input *x* and $\mu_{Bi}(y)$ is the membership degree of input *y*, *I* represents *and* or *or* operators, and *n* is the number of rules [37].

C. *Fuzzy Inference Engine (FIE)*

The procedure of mapping and modeling inputs and outputs derived from the fuzzification step is accomplished in this phase by combining the fuzzy if-then rules with the degrees of matching obtained from the observation [41]. In this procedure, using the reasoning technique or aggregating the defined rules by employing a conjunctive and/or disjunctive method. Various types of fuzzy inference systems (e.g., Mamdani, Sugeno, Tsukamoto-Singleton) are widely applied and accepted in terms of modeling numerous academic and industrial problems [37]. Rule aggregation and defuzzifying algorithms are unique in each fuzzy inference system; for instance, the algorithm in the Mamdani approach is associated with linguistic variables, while in the Takagi-Sugeno method, in the consequent part there is a piecewise linear function connecting the inputs and the output. In this paper, the Mamdani inference system is chosen to be applied due to its simple and advantageous features. The Mamdani fuzzy system has been c applied for dealing with a very wide scope of complex problems, among others, in the field of traffic engineering as well [5], [9], [36]. This model employs fuzzy set representation to convert the utterly unstructured linguistic heuristics into an executable algorithm [37], [38]. The if-then rule process of the Mamdani algorithm (Fig. 2) is:

If x_1 is A_{i1} and x_2 is A_{i2} and ... x_r is A_{ir} then y is B_i (for i = 1, 2, ...k) (6)

where x_i is the input variable and the output variable is *y*, A_{ir} *and* B_i are linguistic terms, and *k* is the number of rules. In determining fuzzy relations in the proposed model, applying proper composition techniques is a crucial step. Among various composition techniques, max-min is the most commonly used [39]. An illustration of a two-rule max-min composition of the Mamdani inference mechanism is shown in Fig. 2. This composition mathematically is formulated as follows:

$$\mu_{C_K}(Z) = \max_{K = 1, 2, \ldots, r}\left[\min\left[\mu_{A_K}(input(x)), \mu_{B_K}(input(y))\right]\right] \quad (7)$$

where the membership functions are μ_{C_K}, μ_{A_K}, and μ_{B_K} of output *z* for rule *k*, input *x*, and *y*, respectively [36], [40].

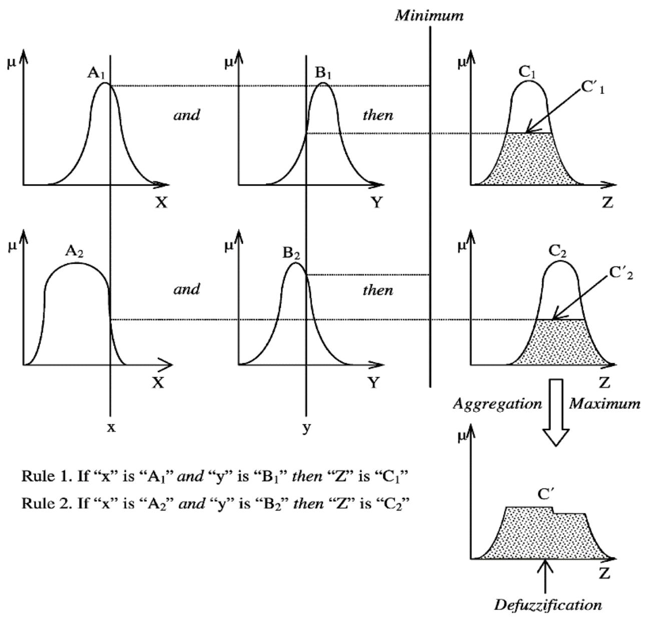


Fig. 2. A two-rule max-min composition [28], [30]

D. *Defuzzification*

In this phase the mathematical procedure of converting fuzzy set results generated by fuzzy inference mechanism into crisp values is implemented. In order to applying this step different translators or defuzzifiers are used (e.g., centroid of area (COA), center of gravity, mean of the maximums, smallest of the maximums) among which the most frequently used technique is the centroid of area operator [43]. This operator is formulated as follows:

$$Z_{COA} = \frac{\int_Z \mu_A(z) z dz}{\int_Z \mu_A(z) dz} \tag{8}$$

Where z is the fuzzy scheme output and aggregated output membership function is assigned as $\mu_A(z)$.

IV. Case Study

In this section, the proposed fuzzy inference-based model is applied to a set of statistics on traffic flow over Hungary's freeways network. The dataset is collected from the online transaction processing server of the Hungarian UD e-toll way system, which is considered an electronic system handled by NÚSZ that allows supporting the confirmation of legally use the network of freeways levying and gathering of the standard road sections tollways [41]. The data are contained seven variables: freeway name, section name, collected sold e-toll over one week in each section of 212 freeways (links) which is considered as the number of vehicles, time (per minute), day, length of the segments, and the number of the lanes in each section. These links include 2446 different segments. Each segment length varies from 100 – 18000 meters, and its number of lanes differs from 2 - 4. The map of the freeways network can be seen in Fig. 3.

Fig. 3. An overview of the chosen network of freeways in the FIS model [41]

A. *Proposed model*

In this study, a fuzzy inference model based on the Mamdani algorithm implemented on the fuzzy logic toolbox R2020b of Matlab is introduced to detect traffic congestion. The proposed model is established to compute and predict the congestion level in a network of freeways. The architecture of the modeled fuzzy inference system with assigned inputs and corresponding output in the Matlab environment is depicted (Fig.4).

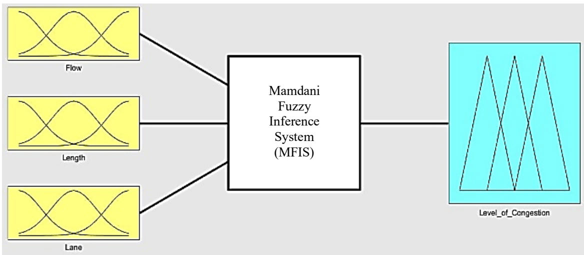


Fig. 4. Schematic illustration of the proposed Fuzzy Inference System (FIS)

The first phase in creating a fuzzy inference model is to specify the model's input and output parameters. Three input parameters (i.e., length, number of lanes, and flow) and one output as the level of congestion (LOC) are selected to be employed in the model. The proposed Mamdani fuzzy inference algorithm has four main design steps:

1) Defining input and output linguistic variables and their corresponding numerical ranges (Table I and II). Input variables are assigned as follows:
   - Flow, which refers to the number of vehicles that passes a specific segment per time unit (the time interval equals 60 minutes),

     $$q = \frac{n}{T} = \frac{n}{\sum_{i=0}^{n} i} \tag{9}$$

     where q is the average number of vehicles *(n)* that pass a section during a unit of time *(T)*.
   - Length of each segment of freeway network per kilometer.
   - Lane, refers the number of lanes in each segment.

TABLE I. INPUT VARIABLES CLUSTERING RANGES

Input variables	Numerical Ranges	Linguistic term	Equivalence (impact on congestion)
Flow	1<F≤300 veh. 200 ≤ F≤600 500 ≤F≤ 1000 800 ≤ F≤ 1400 1200 ≤F< 2000	Very low Low Average High Very high	Very small impact Small impact Steady state Hi. increasing impact Very high increasing impact
Length	0.1 <Le ≤2 1 ≤ Le≤ 6 4≤ Le≤ 12 8 ≤ Le≤ 17 15 ≤ Le < 19	Very short Short Average Long Very long	Very high impact High impact Steady state Reducing impact High reducing impact
Lane	1 < La ≤ 2 2 ≤ La ≤ 3 3 ≤ La < 4	Narrow Average Wide	High increasing impact Reducing impact High reducing impact

TABLE II. OUTPUT VARIABLE CLUSTERING RANGES

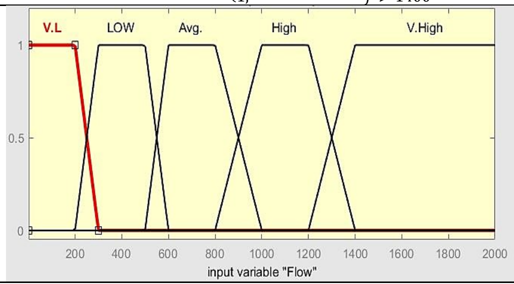
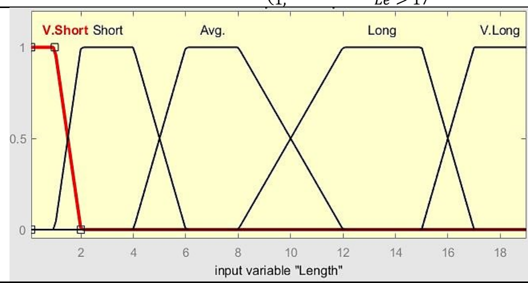
Output variable	Ranges	Linguistic term	Equivalence
LOC	1<LOC ≤ 280 veh.	Completely congestion free	-**IF** Flow is very low/low/avg. **AND** Lane is narrow/avg./wide **AND** length is short/avg./long/very long
	186 ≤ LOC ≤580 veh.	Congestion free	-**IF** Flow is very low/low/avg./high **AND** Lane is narrow/avg./wide **AND** length is avg./long/very long
	470≤ LOC ≤940 veh.	Low	-**IF** Flow is low/avg./high/very high **AND** Lane is narrow /avg./wide **AND** length is avg./long/very long
	750≤ LOC≤1200 veh.	Stable	-**IF** Flow is low/avg./high/very high **AND** Lane is narrow /avg./wide **AND** length is short/avg./long/very long
	1130≤ LOC ≤1500 veh.	Near congestion	-**IF** Flow is avg./high/very high **AND** Lane is narrow/avg./wide **AND** length is length is short/avg.
	1400 ≤ LOC ≤1650 veh.	Congestion	-**IF** Flow is high/very high **AND** Lane is narrow/avg. **AND** length is very short/short/avg.
	1600≤ LOC <2000 veh.	Severe congestion	-**IF** Flow is very high **AND** Lane is narrow/avg. **AND** length is very short/short

2) Input and output parameters are fuzzified (Table III and IV) by triangular and trapezoidal membership functions since they capture and represent the characteristics of the case study's fuzzy set. Preferably, triangular and trapezoidal MFs are shown, respectively in Equations 10 and 11:

$$\mu_A(x) = \begin{cases} 0, & x < \alpha_{min} \\ \dfrac{x-\alpha_{min}}{\beta-\alpha_{min}}, & x \in (\alpha_{min}, \beta) \\ \dfrac{\alpha_{max}-x}{\alpha_{max}-\beta}, & x \in (\beta, \alpha_{max}) \\ 0, & x > \alpha_{max} \end{cases} \quad (10)$$

$$\mu_A(x) = \begin{cases} 0, & x \le \alpha_{min} \\ \dfrac{x-\alpha_{min}}{\beta_1-\alpha_{min}}, & x \in (\alpha_{min}, \beta_1) \\ \dfrac{\alpha_{max}-x}{\alpha_{max}-\beta_2}, & x \in (\beta_2, \alpha_{max}) \\ 0, & x \ge \alpha_{max} \end{cases} \quad (11)$$

TABLE III. MATHEMATICAL AND GRAPHICAL REPRESENTATION OF THE FUZZIFIED INPUT VARIABLES

Input	Membership function
Flow	$\mu_{Flow\_V.low}(f) = \begin{cases} 0, & f > 300 \\ \frac{300-f}{100}, & 200 \le f \le 300 \\ 1, & f < 200 \end{cases}$
	$\mu_{Flow\_Low}(f) = \begin{cases} 0, & (f < 200)_{or}(f > 600) \\ \frac{f-200}{100}, & 200 \le f \le 300 \\ 1, & 300 \le f \le 500 \\ \frac{600-f}{100}, & 500 \le f \le 600 \end{cases}$
	$\mu_{Flow\_Avg.}(f) = \begin{cases} 0, & (f < 500)_{or}(f > 1000) \\ \frac{f-500}{100}, & 500 \le f \le 600 \\ 1, & 600 \le f \le 800 \\ \frac{1000-f}{200}, & 800 \le f \le 1000 \end{cases}$
	$\mu_{Flow\_High}(f) = \begin{cases} 0, & (f < 800)_{or}(f > 1400) \\ \frac{f-800}{200}, & 800 \le f \le 1000 \\ 1, & 1000 \le f \le 1200 \\ \frac{1400-f}{200}, & 1200 \le f \le 1400 \end{cases}$
	$\mu_{Flow\_V.high}(f) = \begin{cases} 0, & f < 1200 \\ \frac{f-1200}{200}, & 1200 \le f \le 1400 \\ 1, & f > 1400 \end{cases}$
Flow MF Plot	
Length	$\mu_{Length\_V.short}(LE) = \begin{cases} 0, & Le > 2 \\ 2-Le, & 1 \le Le \le 2 \\ 1, & Le < 1 \end{cases}$
	$\mu_{Length\_Short}(Le) = \begin{cases} 0, & (Le < 1)_{or}(Le > 6) \\ \mathrm{Le}-1, & 1 \le Le \le 2 \\ 1, & 2 \le L \le 4 \\ \frac{6-Le}{2}, & 4 \le f \le 6 \end{cases}$
	$\mu_{Length\_Avg.}(Le) = \begin{cases} 0, & (Le < 4)_{or}(Le > 12) \\ \frac{Le-4}{2}, & 4 \le Le \le 6 \\ 1, & 6 \le Le \le 8 \\ \frac{12-Le}{4}, & 8 \le f \le 12 \end{cases}$
	$\mu_{Length\_Long}(Le) = \begin{cases} 0, & (Le < 8)_{or}(Le > 17) \\ \frac{Le-8}{4}, & 8 \le Le \le 12 \\ 1, & 12 \le Le \le 15 \\ \frac{17-Le}{2}, & 15 \le f \le 17 \end{cases}$
	$\mu_{Length\_V.long}(Le) = \begin{cases} 0, & Le < 15 \\ \frac{Le-15}{2}, & 15 \le Le \le 17 \\ 1, & Le > 17 \end{cases}$
Length MF Plot	

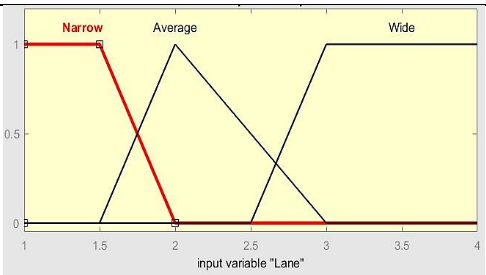
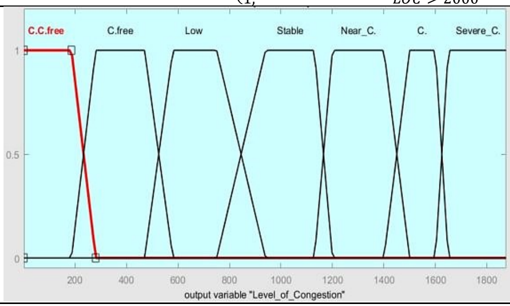
Lane	$\mu_{Lane\_Narrow}(La) = \begin{cases} 0, & La > 2 \\ \frac{2-La}{0.5}, & 1.5 \le La \le 2 \\ 1, & La < 1.5 \end{cases}$ $\mu_{Lane\_Avg.}(La) = \begin{cases} 0, & La \le 1.5 \\ La - 1.5, & 1.5 < La \le 2 \\ 3 - La, & 2 < La < 3 \\ 0, & La \ge 3 \end{cases}$ $\mu_{Lane\_Wide}(La) = \begin{cases} 0, & La < 2.5 \\ \frac{La-2.5}{0.5}, & 2.5 \le La \le 3 \\ 1, & La > 3 \end{cases}$
Lane MF Plot	

TABLE IV. MATHEMATICAL AND GRAPHICAL REPRESENTATION OF THE FUZZIFIED OUTPUT VARIABLE

Output	Membership function
Level of Cong. (LOC)	$\mu_{LOC\_Completely\ Congestion\ free}(LOC) = \begin{cases} 0, & LOC > 280 \\ \frac{280-LOC}{94}, & 186 \le LOC \le 280 \\ 1, & LOC < 186 \end{cases}$
	$\mu_{LOC\_Congestion\ free}(LOC) = \begin{cases} 0, & (LOC < 186)_{or}(LOC > 580) \\ \frac{LOC-186}{94}, & 186 \le LOC \le 280 \\ 1, & 280 \le LOC \le 470 \\ \frac{580-LOC}{110}, & 470 \le LOC \le 580 \end{cases}$
	$\mu_{LOC\_Low}(LOC) = \begin{cases} 0, & (LOC < 470)_{or}(LOC > 940) \\ \frac{LOC-470}{110}, & 470 \le LOC \le 580 \\ 1, & 580 \le LOC \le 750 \\ \frac{940-LOC}{90}, & 750 \le LOC \le 940 \end{cases}$
	$\mu_{LOC\_Stable}(LOC) = \begin{cases} 0, & (LOC < 750)_{or}(LOC > 1200) \\ \frac{LOC-750}{90}, & 750 \le LOC \le 940 \\ 1, & 940 \le LOC \le 1130 \\ \frac{940-LOC}{90}, & 1130 \le LOC \le 1200 \end{cases}$
	$\mu_{LOC\_Near\ congestion}(LOC) = \begin{cases} 0, & (LOC < 1130)_{or}(LOC > 1500) \\ \frac{LOC-1130}{70}, & 1130 \le LOC \le 1200 \\ 1, & 1200 \le LOC \le 1400 \\ \frac{1500-LOC}{100}, & 1400 \le LOC \le 1500 \end{cases}$
	$\mu_{LOC\_Congestion}(LOC) = \begin{cases} 0, & (LOC < 1400)_{or}(LOC > 1650) \\ \frac{LOC-1400}{100}, & 1400 \le LOC \le 1500 \\ 1, & 1500 \le LOC \le 1600 \\ \frac{1650-LOC}{50}, & 1600 \le LOC \le 1650 \end{cases}$
	$\mu_{LOC\_Severe\ congestion}(LOC) = \begin{cases} 0, & LOC < 1600 \\ \frac{LOC-1600}{50}, & 1600 \le LOC \le 1650 \\ 1, & LOC > 2000 \end{cases}$
LOC MF Plot	

3) The input-output relationships are defined by if-then fuzzy rules. A total of 75 rules were assigned based on the available dataset, percentile distribution of the data, and experts' judgment. These rules were implemented in the Matlab Fuzzy Rule Editor to create the inference and nonlinear surface model; 25 rules of the total of 75 rules are illustrated in the Matlab Fuzzy Rule Editor environment in Fig. 5.

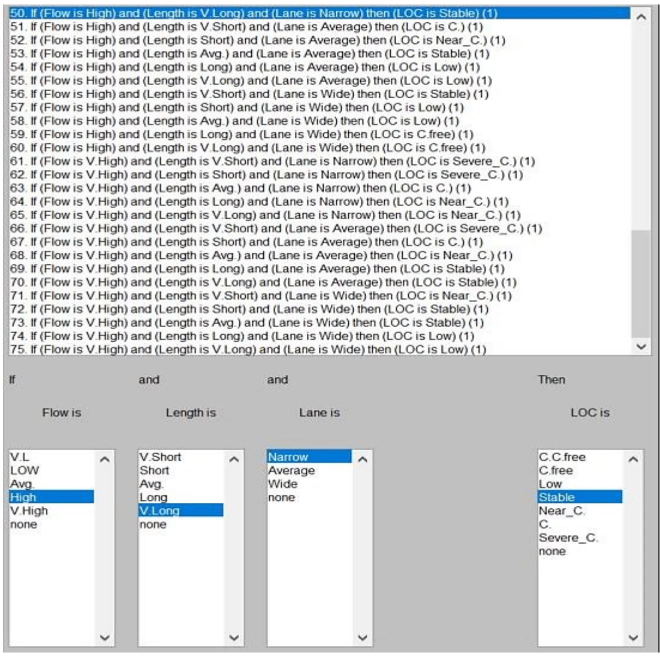


Fig. 5. Input-output rules for determining the level of congestion severity

4) Applying COA as the defuzzification operator to detect the corresponding action (level of congestion) to be executed. This operator is formulated as follows:

$$Z_{COA} = \frac{\int_Z \mu_A(z) z dz}{\int_Z \mu_A(z) dz} \tag{12}$$

where z is the fuzzy system output and aggregated output membership function is assigned as $\mu_A(z)$.

V. RESULTS

As it can be observed from the introduced algorithm steps in the previous section, determining the level of congestion in each segment is derived from three types of available row data, including the number of vehicles in a specific time unit which passed from a certain point along with the number of lanes and length of the given segment. All these input data provided approximate indicators to evaluate the segment's relative capacity in supplying the created demand that can cause alteration in the level of congestion. The obtained results illustrate that the proposed fuzzy inference system is quite efficient in generalizing nonlinear complex relations between congestion levels and the other numerical properties of traffic.

Schematic illustration of fuzzy inferential mechanism can be seen in Fig. 6. The ranges of the input variables are assigned as:

- flow rate from 1 to 2000 vehicles per 60 minutes,
- length of the segments from 0.1 to 19 km,
- number of lanes from 1 to 4;

also, as the output variable:

- level of congestion is considered in the range of 1 to 1875.

The developed Mamdani model in this study can provide an estimation of congestion severity when input data is inserted. As a sample of the proposed model application from Fig.6, it can be observed that if real-time input parameter properties are entered as: flow rate = 253, the segment has two lanes and 5.16 km length, then the level of congestion (LOC) would be predicted as 281, which based on the membership function presented in Table IV, is categorized at the level of congestion-free.

The performed prediction of congestion level was based on historical and real-time observations that play a significant role in various traffic models [42]. As opposed to conventional methods of traffic detection, the proposed mechanism has a sophisticated discipline known as approximate reasoning [19], [43] through which exact traffic connected properties (e.g., geometric features including junctions, bifurcations, off-ramps, and on-ramps) that can be assigned in both microscopic and mesoscopic types of traffic modeling [44], [45] are sacrificed, in order to reach significantly low time and computational efforts. Besides, natural linguistic rules form the executed model, which is aligned with the general concepts of traffic characteristics. Also, the results were obtained from a combined description of the congestion state because of employing multiple and compound rules in the modeled inference system instead of using a single rule.

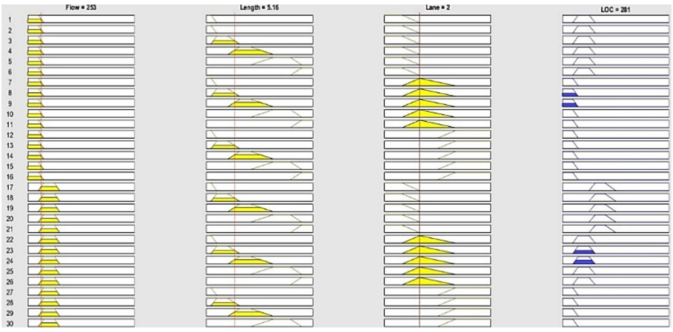
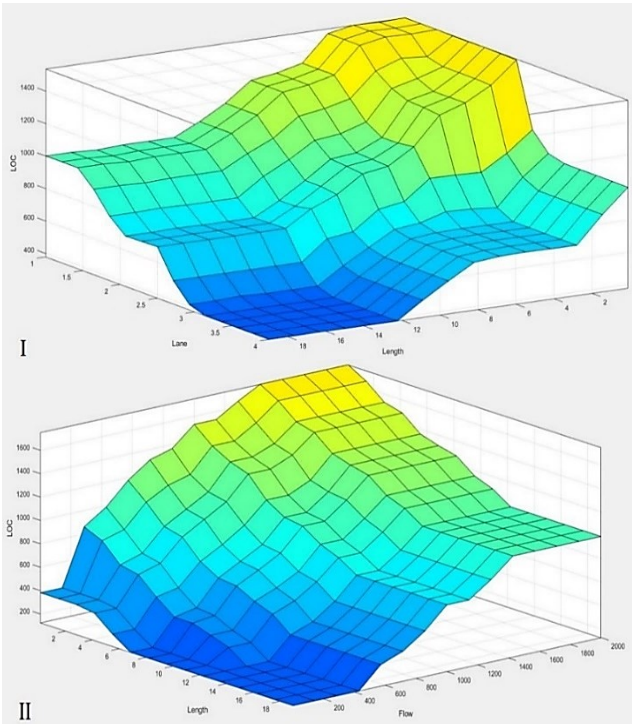


Fig. 6. Lookup diagram of fuzzy rules of the level of congestion

One of the most significant contributions of the simulated results is the fuzzy surface view which can produce practical information extracted from the analyzed system's data, for example, evaluating correlation and strength of the relationship between assigned input and output variables. Although the information provided by the fuzzy surface view mainly focuses on the correlations of the input-output variables, another feature of the provided view is about the system reaction rate to the fluctuations caused by the input variables and the direction of the alteration effects on the output variable. It is a significant advantage since an entirely different effectual view of the analyzed system coupled with having the capability to evaluate a large number of possible scenarios and outcomes at once can be observed by engineers without having to infer the system's mathematical formulations which conventional control models disable to provide. Transparency is one of the main advantages and reasons of wide applicability of fuzzy rule-based models.

The interdependency among input variables and LOC derived from the generated rules in the proposed fuzzy inference system can be demonstrated by employing a fuzzy control surface in a visual perception view (Fig. 7). It shows a correlation between LOC and the input variables. It can be observed that the most severe alteration takes place in the LOC when the length is approximately in the range of 4-6 km and the lane number is 2-3 (Fig.7 part I). Also, if the range is approximately 1-6 km in the length variable, in each segment with an increasing flow rate of more than 200, it can be seen that an intensive reaction (around 50% increase) in the LOC emerges (Fig.7 part II). Altering the number of lanes has the most intensive impact on the LOC. Segments with 3 and 4 lanes will not experience severe congestion, while increasing the flow rate by 200 vehicles in less than two lanes segments can increase LOC by more than 50% (Fig.7 part III).

Besides its tractability in dealing with ambiguity and subjectivity, the above-performed analysis is aligned with traffic modeling purposes, i.e., sustainability development, infrastructure, and spatial planning [10], [46]. Furthermore, various traffic engineering tasks can be supported through the presented correlations and relationships among the involved variables, such as assessing the effect of increased demand or capacity-reducing events (e.g., a lane drop because of maintenance) on the traffic flow behavior. In other words, since the use of the model is mainly real-time, effects of changing in each of the variables in a specific segment on the whole system can be observed, particularly in order to, providing surveillance of the congestion state in complex networks, traffic control, and prevention strategies, further the assessment of the impact of new constructions. Moreover, as the proposed model is designed mainly based on the available research datasets, all applications mentioned above can be significantly improved by involving additional accurate traffic associated parameters.



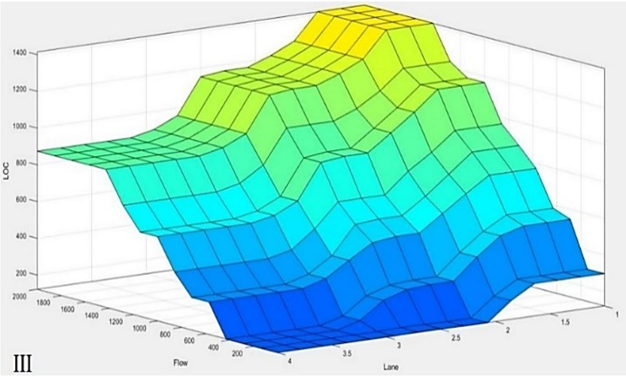


Fig. 7. Rule surface of LOC for length and lane (I), and length and flow (II), and flow and lane (III)

VI. Conclusion

In this study, the fuzzy rule-based inference is proposed for modeling traffic-involved parameters to detect the severity of congestion by available row data. The inferential process time and complexity are within an appropriate time frame based on the computed number of inputs along with the number of produced fuzzy rules. Furthermore, the projected approach not only allows to recognize and analyze traffic congestion with ease, but it also demonstrates effectual performance and reliable traffic congestion control system with exceptionally high noise acceptance. Such a control system can pay the way for creating traffic breakdown-related alerts and intelligent early warning systems with its potential advantages to deal with traffic-connected problems, infrastructure development, and services progression. Besides all of the contributions of the proposed fuzzy-based inference system (e.g., flexibility, computationally efficiency, dealing with imprecision), its characteristic about assuming all of the input variables as equal-weighted values on the level of congestion needs further improvements, particularly concerning associated traffic events which are notoriously heterogeneous. Moreover, the used data only applies to one week. The accuracy of the fuzzy model can be better experienced and estimated if more data is provided to the system.

References

[1] B. Avron and E. A. Feigenbaum, "The handbook of artificial intelligence", volume 2, Butterworth-Heinemann 2014, pp. 3–15, 1982.

[2] Grosan, C., & Abraham, A. Intelligent systems. Springer. 2011, DOI: https://doi.org/10.1007/978-3-642-21004-4

[3] Cascetta, Ennio. Transportation systems engineering: theory and methods. Vol. 49. Springer Science & Business Media, 2013, DOI: https://doi.org/10.1007/978-1-4757-6873-2

[4] Bellouquid, A., De Angelis, E. and Fermo, L., Towards the modeling of vehicular traffic as a complex system: A kinetic theory approach. Mathematical models and methods in applied sciences,22(supp01),2012,DOI:https://doi.org/10.1142/S0218202511400033

[5] M. Kalinic and J. M. Krisp, "Fuzzy inference approach in traffic congestion detection," Ann. GIS, vol. 25, no. 4, pp. 329–336, 2019, DOI: https://doi.org/10.1080/19475683.2019.1675760

[6] Dechenaux, Emmanuel, Shakun D. Mago, and Laura Razzolini. "Traffic congestion: an experimental study of the Downs-Thomson paradox." Experimental economics 17, no. 3 (2014): 461-487, DOI:10.1007/s10683-013-9378-4

[7] Hu W., Feng Z., Chen Z., Harkes J., Pillai P., Satyanarayanan M. Live synthesis of vehicle-sourced data over 4G LTE ACM MSWIM (2017), pp. 161-170, DOI: https://doi.org/10.1145/3127540.3127543

[8] Bao W., Yuan D., Yang Z., Wang S., Zhou B., Adams S., Zomaya A. SFog: Seamless fog computing environment for mobile IoT applications ACM MSWIM (2018), pp. 127-136, DOI: https://doi.org/10.1145/3242102.3242107

[9] K. Iqbal, M. A. Khan, S. Abbas, Z. Hasan, and A. Fatima, "Intelligent transportation system (ITS) for smart-cities using Mamdani Fuzzy Inference System," Int. J. Adv. Comput. Sci. Appl., vol. 9, no. 2, pp. 94–105, 2018.

[10] C. Pasquale, S. Sacone, S. Siri, and A. Ferrara, "Traffic control for freeway networks with sustainability-related objectives: Review and future challenges," Annu. Rev. Control, vol. 48, pp. 312–324, 2019, DOI: https://doi.org/10.1016/j.arcontrol.2019.07.002

[11] J. G. WARDROP, "ROAD PAPER. SOME THEORETICAL ASPECTS OF ROAD TRAFFIC RESEARCH.," Proc. Inst. Civ. Eng., vol. 1, no. 3, pp. 325–362, 1952, DOI: https://doi.org/10.1680/ipeds.1952.11259

[12] Boukerche A, Wang J. Machine Learning-based traffic prediction models for Intelligent Transportation Systems. Computer Networks. 2020 Nov 9;181:107530, DOI: https://doi.org/10.1016/j.comnet.2020.107530

[13] Nya N., Baynat B. Performance model for 4G/5G heterogeneous networks with different classes of users ACM MSWiM (2017), pp. 171-178, DOI: https://doi.org/10.1145/3127540.3127554

[14] Olivieri B., Endler M. DADCA: An efficient distributed algorithm for aerial datacollection from wireless sensors networks by UAVs ACM MSWiM (2017), pp. 129-136, DOI: https://doi.org/10.1145/3127540.3127553

[15] Aljeri N., Boukerche A. Movement prediction models for vehicular networks: An empirical analysis Wirel. Netw., 25 (4) (2019), pp. 1505-1518, DOI: https://doi.org/10.1007/s11276-017-1607-x

[16] Alipour B., Tonetto L., Ketabi R., Yi Ding A., Ott J., Helmy A. Where are you going next? A practical multi-dimensional look at mobility prediction ACM MSWIM (2019), pp. 5-12, DOI: https://doi.org/10.1145/3345768.3355923

[17] F. Hayes-Roth, "Knowledge-based expert systems," Computer (Long. Beach. Calif)., vol. 17, no. 10, pp. 263–273, 1984.

[18] N. Sabri, S. A. Aljunid, M. S. Salim, R. B. Badlishah, R. Kamaruddin, and M. F. Abd Malek, "Fuzzy inference system: Short review and design," Int. Rev. Autom. Control, vol. 6, no. 4, pp. 441–449, 2013.

[19] Mehran Amini, Miklos F. Hatwagner , Gergely Mikulai and Laszlo T. Koczy, "Developing a macroscopic model based on fuzzy cognitive map for road traffic flow simulation", Infocommunications Journal, Vol. XIII, No 3, September 2021, p. 14-23, DOI: 10.36244/ICJ.2021.3.2

[20] M. Amini, M. F. Hatwagner, G. C. Mikulai and L. T. Koczy, "An intelligent traffic congestion detection approach based on fuzzy inference system," 2021 IEEE 15th International Symposium on Applied Computational Intelligence and Informatics (SACI), 2021, pp. 97-104, DOI: 10.1109/SACI51354.2021.9465637.

[21] P. Arnesen and O. A. Hjelkrem, "An estimator for traffic breakdown probability based on classification of transitional breakdown events," Transp. Sci., vol. 52, no. 3, pp. 593–602, 2018, ODI: https://doi.org/10.1287/trsc.2017.0776

[22] B. Krause, C. von Altrock, and M. Pozybill, "Intelligent highway by fuzzy logic: congestion detection and traffic control on multi-lane roads with variable road signs," in Proceedings of IEEE 5th International Fuzzy Systems, 1996, vol. 3, pp. 1832–1837 vol.3, DOI: 10.1109/FUZZY.1996.552649.

[23] S. Logghe and L. H. Immers, "Multi-class kinematic wave theory of traffic flow," Transp. Res. Part B Methodol., vol. 42, no. 6, pp. 523–541, 2008, DOI: https://doi.org/10.1016/j.trb.2007.11.001

[24] A. Messmer and M. Papageorgiou, "METANET: a macroscopic simulation program for motorway networks," Traffic Eng. Control, vol. 31, no. 8–9, pp. 466–470, 1990.

[25] K. A. R. B, R. Moormann, and D. Dierich, "Learning Fuzzy Cognitive Map for Traffic Prediction Using an Evolutionary Algorithm," pp. 170–181, 2015, DOI: https://doi.org/10.1007/978-3-319-26404-2_16

[26] S. Majumdar, M. M. Subhani, B. Roullier, A. Anjum, and R. Zhu, "Congestion prediction for smart sustainable cities using IoT and machine learning approaches," Sustain. Cities Soc., vol. 64, no. March 2020, p. 102500, 2021, DOI: https://doi.org/10.1016/j.scs.2020.102500

[27] M. Aftabuzzaman, "Measuring Traffic Congestion- A Critical Review 30 th Australasian Transport Research Forum 30 th Australasian Transport Research Forum," 30th Australas. Transp. Res. Forum, no. January 2007, pp. 1–16, 2007.

[28] J. M. Mendel, "Fuzzy-Logic Systems for Engineering - a Tutorial (Vol 83, Pg 345, 1995)," Proc. Ieee, vol. 83, no. 9, p. 1293, 1995, DOI: 10.1109/5.364485

[29] L. A. Zadeh, "Fuzzy sets," Inf. Control, vol. 8, no. 3, pp. 338–353, 1965, DOI: https://doi.org/10.1142/9789814261302_0021

[30] L. A. Zadeh, "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes," IEEE Trans. Syst. Man Cybern., vol. SMC-3, no. 1, pp. 28–44, 1973, DOI: 10.1109/TSMC.1973.5408575

[31] 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, DOI: 10.1109/TSMC.1977.4309605

[32] J. Finol, Y. Ke Guo, and X. D. Jing, "A rule based fuzzy model for the prediction of petrophysical rock parameters," J. Pet. Sci. Eng., vol. 29, no. 2, pp. 97–113, 2001, DOI: https://doi.org/10.1016/S0920-4105(00)00096-6

[33] M. Monjezi, M. Rezaei, and A. Yazdian Varjani, "Prediction of rock fragmentation due to blasting in Gol-E-Gohar iron mine using fuzzy logic," Int. J. Rock Mech. Min. Sci., vol. 46, no. 8, pp. 1273–1280, 2009, DOI: https://doi.org/10.1016/j.ijrmms.2009.05.005

[34] A. Yazdani-Chamzini and S. H. Yakhchali, "Tunnel Boring Machine (TBM) selection using fuzzy multicriteria decision making methods," Tunn. Undergr. Sp. Technol., vol. 30, pp. 194–204, 2012, DOI: https://doi.org/10.1016/j.tust.2012.02.021

[35] A. Jamshidi, A. Yazdani-Chamzini, S. H. Yakhchali, and S. Khaleghi, "Developing a new fuzzy inference system for pipeline risk assessment," J. Loss Prev. Process Ind., vol. 26, no. 1, pp. 197–208, 2013, DOI: https://doi.org/10.1016/j.jlp.2012.10.010

[36] E. H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," Int. J. Man. Mach. Stud., vol. 7, no. 1, pp. 1–13, 1975, DOI: https://doi.org/10.1016/S0020-7373(75)80002-2

[37] M. Iphar and R. M. Goktan, "An application of fuzzy sets to the Diggability Index Rating Method for surface mine equipment selection," Int. J. Rock Mech. Min. Sci., vol. 43, no. 2, pp. 253–266, 2006, DOI: https://doi.org/10.1016/j.ijrmms.2005.07.003

[38] Y. Wang and Y. Chen, "A Comparison of Mamdani and Sugeno Fuzzy Inference Systems for Traffic Flow Prediction," J. Comput., vol. 9, no. 1, pp. 12–21, 2014, DOI: :10.4304/jcp.9.1.12-21

[39] T. J. Ross, Fuzzy logic with engineering applications, vol. 2. Wiley Online Library, pp 212-357, 2004.

[40] M. Monjezi and M. Rezaei, "Developing a new fuzzy model to predict burden from rock geomechanical properties," Expert Syst. Appl., vol. 38, no. 8, pp. 9266–9273, 2011, DOI: https://doi.org/10.1016/j.eswa.2011.01.029

[41] GSMPRO Kft., "General Terms and Conditions," https://tracker.gsm-pro.hu/. 2019.

[42] W. Min and L. Wynter, "Real-time road traffic prediction with spatio-temporal correlations," Transp. Res. Part C Emerg. Technol., vol. 19, no. 4, pp. 606–616, 2011, DOI: https://doi.org/10.1016/j.trc.2010.10.002

[43] D. Pradeepkumar and V. Ravi, "Soft computing hybrids for FOREX rate prediction: A comprehensive review," Comput. Oper. Res., vol. 99, pp. 262–284, 2018, DOI: https://doi.org/10.1016/j.cor.2018.05.020

[44] S. K. Zegeye, B. De Schutter, J. Hellendoorn, E. A. Breunesse, and A. Hegyi, "Integrated macroscopic traffic flow, emission, and fuel consumption model for control purposes," Transp. Res. Part C Emerg. Technol., vol. 31, pp. 158–171, 2013, DOI: https://doi.org/10.1016/j.trc.2013.01.002

[45] W. Imran, Z. H. Khan, T. Aaron Gulliver, K. S. Khattak, and H. Nasir, "A macroscopic traffic model for heterogeneous flow," Chinese J. Phys., vol. 63, no. December 2019, pp. 419–435, 2020.

[46] T. Marshall, "Infrastructure planning and spatial planning: Current relationships in the UK," Town Plan. Rev., vol. 91, no. 1, pp. 47–69, 2020.