



# USING THE INTERVAL TYPE-2 FUZZY INFERENCE SYSTEMS TO DETERMINE A RELATIONSHIP BETWEEN THE ROAD CHARACTERISTICS ASSESSMENT AND ROAD TRAFFIC ACCIDENTS

Marjana Čubranić-Dobrodolac<sup>1,2\*</sup>,  
Libor Švadlenk<sup>2</sup>,  
Svetlana Čičević<sup>1</sup>,  
Aleksandar Trifunović<sup>1</sup>

<sup>1</sup>University of Belgrade  
Faculty of Transport and Traffic Engineering,  
Belgrade, Serbia

<sup>2</sup>University of Pardubice  
Faculty of Transport Engineering,  
Pardubice, Czech Republic

## Abstract:

There are several Type-2 fuzzy inference systems (T2FISs) tested in the research described in this paper. These models use one, two, or three input variables and all of them are described by one output variable. Input variables relate to the assessment of dangerous places on the observed road section, assessment of road characteristics, and frequency of driving. The assessment of dangerous places is obtained as an average score from assessing nine dangerous spots on the considered road section in Serbia called "Ibarska magistrala". Assessment of road characteristics is based on seven scores, which means that participants assessed seven predefined road characteristics in the same road section. The frequency of driving is an input variable based on the criterion of how many times a week or a month the examinee drives on the observed road section. Output variable is The number of road traffic accidents that a driver had experienced. T2FISs are tested on a sample of 305 drivers and most of them are professional drivers. The results are perceived through the cumulative error that T2FISs make in the description of empirical data.

## Keywords:

Type-2 fuzzy inference systems, road traffic accidents, road safety, dangerous spots assessment, road characteristics assessment.

## INTRODUCTION

Road traffic accidents (RTAs) cause around 3700 deaths in the world every day [1]. Thousands of people are also injured every day. In addition, the RTAs generate significant economic costs. It is evident that there is a need to examine the reasons for the occurrence of RTAs, which would facilitate the design of programs for traffic safety improvements.

By reviewing the literature, it is possible to conclude that there are three general causes of RTAs. They relate to the driver, i.e. human factor, vehicle, and road [2,3]. Sam, Velanganni, and Evangelin [4] report that human errors are recognized as the far most common influential factor causing more than 90 % of RTAs. Some of the most common driver behaviors that lead to the occurrence of RTA are: unadjusted speed, wrong assessment of the traffic situation, and the performance of improper actions in traffic [5].

## Correspondence:

Marjana Čubranić-Dobrodolac

## e-mail:

marjana@sf.bg.ac.rs



It is proven that the drivers who do not respect the traffic rules in one segment, usually do not behave properly also in some other segment. For example, the drivers in Serbia are forbidden to talk on the phone while driving, except when using a hands-free device. A study by Čubranić-Dobrodolac et al. [6] showed that the participants who violate this rule, are prone to drive under the influence alcohol as well, especially the group of drivers who experienced more than three RTAs in their driving experience. This points to the conclusion that the human factor as a cause of RTAs and general driver behavior can be explained to a large extent by the corresponding psychological traits. There are studies that confirm a relationship between certain psychological traits that lead to risky behavior and the propensity for RTAs [2,7]. Besides, there is also evidence that the skills of drivers, such as speed estimation or spatial abilities, affect the likelihood of being involved in an RTA [8].

In this paper, we examine also a relationship between certain skills of drivers and the number of experienced RTAs. Here, the assessment of dangerous places on the road, as well as the road characteristics assessment are considered as skills of drivers. In addition, the frequency of driving is also taken into consideration.

For the purpose of quantification of relationships between the considered variables, both individually and jointly, and the number of experienced RTAs, seven Type-2 fuzzy inference systems (T2FISs) are designed and tested [9,10]. The T2FIS models are tested on a sample of 305 drivers. This paper can be seen as an extension of the research presented in the paper by Čubranić-Dobrodolac et al. [11] where Type-1 fuzzy inference systems were tested; however, with the different aim, to compare the effects of various shapes of membership functions on the performance of the fuzzy system.

## 1. MODEL DEVELOPMENT – USED VARIABLES

The model proposed in this paper is based on the design and testing of the performance of seven T2FISs. These T2FISs use one, two, or three input variables, and all of them have the same output variable. Input variables relate to The assessment of dangerous places on the observed road section, Assessment of road characteristics, and Frequency of driving on the road section. The output variable is The number of accidents that a driver had experienced by his or her fault. An additional explanation of designed T2FISs is offered in Table 1.

TABLE I. TESTED TYPE-2 FUZZY INTERFERENCE SYSTEMS

T2FIS No.	Used variables	Name of used variables
I	$x_1, y$	Dangerous places – Accidents
II	$x_2, y$	Road characteristics – Accidents
III	$x_3, y$	Frequency of driving – Accidents
IV	$x_1, x_2, y$	Dangerous places, Road characteristics – Accidents
V	$x_1, x_3, y$	Dangerous places, Frequency of driving – Accidents
VI	$x_2, x_3, y$	Road characteristics, Frequency of driving – Accidents
VII	$x_1, x_2, x_3, y$	Dangerous places, Road characteristics, Frequency of driving – Accidents

The chosen road section is in the Republic of Serbia, road category IB, number 22, marked as road E-763, which is colloquially known as „Ibarska magistrala“. The first variable  $x_1$  named *Dangerous spots* is formed based on the assessment of nine dangerous places on the considered road section which are previously noticed by the authors. These dangerous places are determined based on the official statistic about the most frequent location of traffic accidents in the past (Fig. 1).



Figure 1. Dangerous places at route E-763 marked with the triangles [12]



Dangerous spots that were assessed by the participants are the following: Žarkovo, Orlovača, Ripanj, Šiljakovac, turn to the Barajevo, Stepojevac, Šopić, Lazarevac roundabout, and Čelije. Each of the 305 examinees from the sample gave scores about the aforementioned nine dangerous places. The participants gave the marks according to the level of risk they perceive considering these places, i.e. to what extent are the considered dangerous spots really dangerous. The final value of the first input variable is a sum of all nine scores.

Input variable *Dangerous spots* refer to the sum of estimates of nine hazardous sites that could be rated by a score of 1 to 10, which means that the domain of this variable is from 9 to 90. It can be displayed using 5 fuzzy sets and their membership functions (Fig. 2): VLAR – very low assessment of risk, LAR – low assessment of risk, MAR – medium assessment of risk, HAR – high assessment of risk, VHAR – very high assessment of risk.

As can be seen from Fig. 2, the fuzzy sets that describe the input variable *Dangerous spots* do not cover the same intervals, as a result of the fact that these fuzzy sets are defined based on the empirical data of 305 respondents (Table 2). Although the minimum value for the assessment of nine points is 9, the observed sample showed that the minimum value was 35. Based on this, it can be concluded that in the average driver population, the lesser values are rarely met and therefore, a relatively large interval for potential values of risk assessment remains uncovered. Accordingly, the fuzzy set VLAR has the largest interval compared to the remaining 4 fuzzy sets.

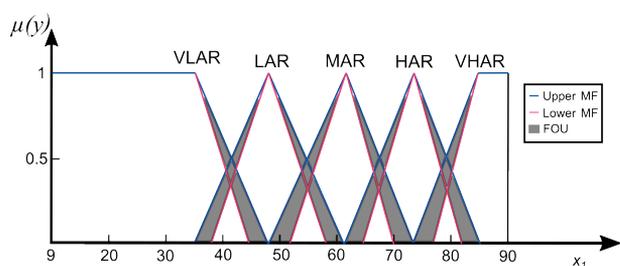


Figure 2. Input variable Dangerous spots

TABLE II. DOMAIN INTERVALS FOR USED VARIABLES AND DESCRIPTIVE STATISTICS OF THE SAMPLE

Variable	Domain	Descriptive statistics of the sample		
		Minimum	Mean	Maximum
$x_1$	[9,90]	35	60.91	85
$x_2$	[7,70]	32	43.84	65
$x_3$	[1,7]	1	2.58	7
$y$	[0,8]	0	1.46	8

On the other hand, the mean value of the assessment of the hazardous locations of all respondents who participated in the study was 60.91; on the basis of which this value was taken as part of fuzzy set MAR with the highest value of upper membership function (equal to 1).

The second input variable  $x_2$  named *Road characteristics* contains an evaluation of seven characteristics of the observed road section. These are: The shape of the route and general perspective, roadway condition, state of horizontal signaling, state of vertical signaling, condition of the protective fence, sideroad part - drainage channels, etc, and traffic safety on connecting roads.

The input variable *Road characteristics* refers to the sum of the estimates of seven road characteristics that could be estimated from 1 to 10, which means that the domain of this variable is from 7 to 70. It can be displayed using the 5 fuzzy sets and their membership functions, as shown in Fig. 3: VLAC – very low assessment of road characteristics, LAC – low assessment of road characteristics, MAC – medium assessment of road characteristics, HAC – high assessment of road characteristics, VHAC – very high assessment of road characteristics.

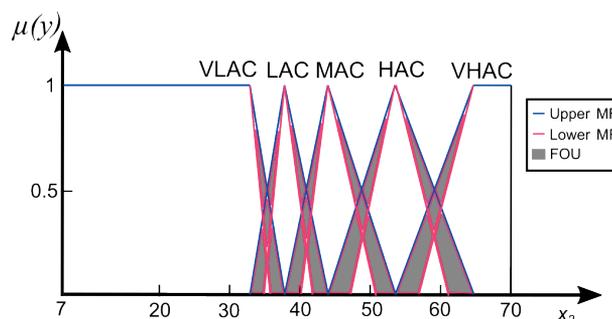


Figure 3. Input variable Road characteristics



It can be noticed that the fuzzy sets that describe the input variable *Road characteristics* do not cover the same intervals. Empirical data showed that the sum of grades ranges from 32 to 65. The mean value of the assessment of all respondents who participated in the study was 43.83; based on which this value was taken as the value of the fuzzy set SPK with the highest membership degree.

The third input variable  $x_3$  named *Frequency of driving* refers to the characteristic of how often the driver drives at the observed road section. The following marks are introduced: 1 – every day, 2 – drives 3-4 times a week, 3 – drives 2 times a week, 4 – drives once a week, 5 – drives 2-3 times a month, 6 – drives once a month, and 7 - drives once in a couple of months. The mentioned values are described by the seven fuzzy sets as follows: VHF – very high frequency, HF – high frequency, MHF – medium-high frequency, MF – medium frequency, MLF – medium-low frequency, LF – low frequency and VLF – very low frequency, as shown in Fig. 4.

The output variable  $y$  named *Accidents* relates to the number of traffic accidents that respondents had experienced. In the sample, the participants reported the number of RTAs from 0 to 8; however, drivers who participated in more than 5 accidents are relatively rare. Therefore, the output variable  $y$  is defined as shown in Fig. 5. The following fuzzy sets are introduced: VSNA – very small number of accidents, SNA – small number of accidents, MSNA – moderately small number of accidents, MNA – medium number of accidents, MHNA – moderately high number of accidents, HNA – high number of accidents, VHNA – very high number of accidents.

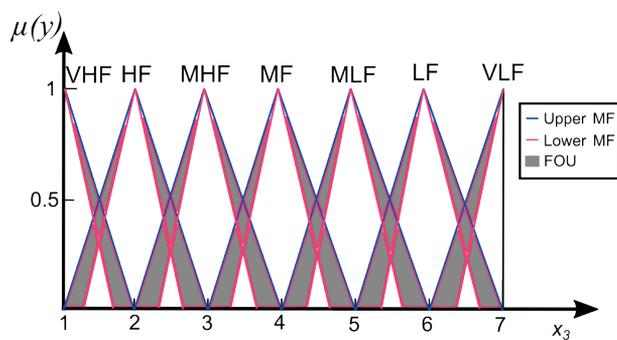


Figure 4. Input variable Frequency of driving

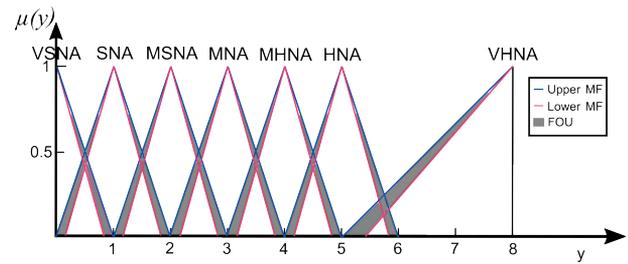


Figure 5. Output variable Accidents

## 2. MODEL DEVELOPMENT – FORMING FUZZY RULES BASED ON EMPIRICAL DATA

To design a T2FIS, it is necessary to set the fuzzy rules base. Here, we use a well-known approach for defining fuzzy rules based on empirical data proposed by Wang and Mendel [13].

The Wang-Mendel method consists of five steps. Step 1 divides the input and output spaces of the given numerical data into fuzzy regions. For each of the used variables, the domain interval is determined, i.e. the interval of the possible values of variables. Each domain interval is divided into  $2N+1$  regions, as shown in figures from 2 to 5.

Step 2 generates fuzzy rules from the collected data. One data pair is used for the construction of one fuzzy rule. Because there 305 respondents in the sample, in this step we generated 305 fuzzy rules. IF part of the fuzzy rule is composed of the names of regions with the maximum membership degree for input variables and THEN part from the name of the region with the maximum degree for the output variable. However, considering 305 designed rules, some of them are the same. The same fuzzy rules should be excluded and remained just one as their representative.

In Step 3, a problem of conflicting rules should be solved. These are the rules that have the same IF part, but a different THEN part. For this purpose, each of the formed rules should be assigned a degree, defined by the Eq. (1) for the case when a rule is defined as following: “IF  $x_1$  is A and  $x_2$  is B, THEN  $y$  is C”.

$$D(Rule) = \mu_A(x_1) * \mu_B(x_2) * \mu_A(y) \tag{1}$$

where  $D(Rule)$  is a degree of a rule,  $\mu_A(x_1)$  is a value of the membership function of the region A when the



input value is  $x_1$ ,  $\mu_B(x_2)$  is a value of the membership function of the region  $B$  when the input value is  $x_2$ , and  $\mu_A(y)$  is a value of membership function of the region  $C$  when output value is  $y$ . In a conflict group, only the rule that has a maximum degree should be accepted.

Step 4 makes a combined fuzzy rule base which consists of rules obtained from empirical data and linguistic rules of a human expert. Finally, Step 5 determines a mapping from input to output space using a defuzzification procedure.

To get fuzzy rules based on the Wang-Mendel approach, we applied the programming code that is prepared based on the code presented in the paper by Čubranić-Dobrodolac et al. [2].

### 3. RESULTS AND DISCUSSION

After the creation of seven T2FISs according to the aforementioned methodology, their performance should be tested. This is done based on Eq. 2 [2,8,14]:

$$CD = \sum_{i=1}^{305} |y^{(i)} - Propensity(i)| \quad (2)$$

where  $CD$  is Cumulative deviation,  $y^{(i)}$  is the number of accidents which drivers from the sample experienced, and  $Propensity(i)$  is the result of T2FIS.

Cumulative deviation ( $CD$ ), as a measure that describes how well the T2FIS describes the empirical data. It is calculated as an absolute value of the difference between the number of RTAs that the drivers from the sample experienced and the corresponding result of T2FIS. The result of T2FIS for a participant number  $i$  in the Eq. (2) is marked as  $Propensity(i)$ .

After the required calculations, the final results are presented in Fig 6. Here, the values of  $CD$  for each considered T2FIS can be noticed.

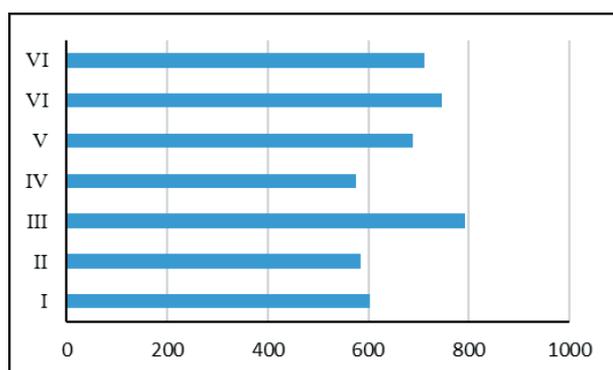


Figure 6. Comparison of various fuzzy inference systems based on the obtained  $CD$  values

It should be kept in mind that smaller values of  $CD$  characterize the T2FIS that gives better results. Therefore, from seven considered T2FIS, the best performance is achieved by T2FIS No. IV ( $CD= 574.61$ ).

This result indicates that the T2FIS with two input variables is the most convenient. These two variables are related to the assessment of dangerous places on the road and the assessment of road characteristics. Therefore, a useful decision-making tool for measuring the propensity for RTAs can be formed by analyzing the driver's characteristics related to these two aforementioned variables.

To discuss the general quality and performance of this T2FIS, it is meaningful to carry out further research about the causes of RTAs, and to test new T2FIS using some other variables.

In this paper, we tested the various combinations for input variables of T2FIS. However, some other methodological approaches related to testing would be welcome as a direction for further research. For example, it is possible to optimize the performance of the current T2FISs. This should be done by using some other method for generating the fuzzy rules, intervals, and shapes of membership functions, and by modifying the method of defuzzification.

Finally, it can be concluded that the proposed research can be very useful in the field of traffic safety. By analyzing the driver's assessment of dangerous places on the road and road characteristics, it is possible to get an impression about his or her propensity towards RTAs. This information would facilitate the design of adequate educational programs tailored to the needs of a particular driver, with the final aim to improve the skills of drivers, and by that to contribute to fewer RTAs on the roads, saving the lives of people and contributing to their wellbeing.

### REFERENCES

- [1] World Health Organization, "The Global status report on road safety 2018." (2018). [https://www.who.int/violence\\_injury\\_prevention/road\\_safety\\_status/2018/en/](https://www.who.int/violence_injury_prevention/road_safety_status/2018/en/) Accessed 18 June 2020.
- [2] M. Čubranić-Dobrodolac, L. Švadlenka, S. Čičević, M. Dobrodolac, "Modelling driver propensity for traffic accidents: a comparison of multiple regression analysis and fuzzy approach", International Journal of Injury Control and Safety Promotion, vol. 27, issue 2, pp. 156-167, 2020. <https://doi.org/10.1080/17457300.2019.1690002>



- [3] M. Karacasu, B. Ergül, A.A. Yavuz, "Estimating the causes of traffic accidents using logistic regression and discriminant analysis", *International Journal of Injury Control and Safety Promotion*, vol. 21, issue 4, pp. 305-313, 2014. <https://doi.org/10.1080/17457300.2013.815632>
- [4] D. Sam, C. Velanganni, T.E. Evangelin, "A vehicle control system using a time synchronized Hybrid VANET to reduce road accidents caused by human error", *Vehicular Communications*, vol. 6, pp. 17-28. 2016. <https://doi.org/10.1016/j.vehcom.2016.11.001>
- [5] M. Čabarkapa, M. Čubranić-Dobrodolac, S. Čičević, B. Antić, "The influence of aggressive driving behavior and impulsiveness on traffic accidents", *International Journal for Traffic and Transport Engineering*, vol. 8, issue 3, pp. 372 - 389. 2018. [https://doi.org/10.7708/ijtte.2018.8\(3\).09](https://doi.org/10.7708/ijtte.2018.8(3).09)
- [6] M. Čubranić-Dobrodolac, S. Čičević, M. Dobrodolac, M. Nešić, "The risks associated with using a mobile phone by young drivers", *Transport*, vol. 26, issue 4, pp. 381-388. 2013. <https://doi.org/10.3846/16484142.2013.865672>
- [7] M. Čubranić-Dobrodolac, K. Lipovac, S. Čičević, and B. Antić, "A model for traffic accidents prediction based on driver personality traits assessment", *Promet - Traffic & Transportation*, vol. 29, issue 6, pp. 631-642. 2017. <https://doi.org/10.7307/ptt.v29i6.2495>
- [8] M. Čubranić-Dobrodolac, L. Švadlenka, S. Čičević, A. Trifunović, M. Dobrodolac, "Using the Interval Type-2 Fuzzy Inference Systems to Compare the Impact of Speed and Space Perception on the Occurrence of Road Traffic Accidents", *Mathematics*, vol. 8, no. 1548. 2020. <https://doi.org/10.3390/math8091548>
- [9] E.H. Mamdani, S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller", *International Journal of Man-Machine Studies*, vol. 7, issue 1, pp. 1-13, 1975. [https://doi.org/10.1016/S0020-7373\(75\)80002-2](https://doi.org/10.1016/S0020-7373(75)80002-2)
- [10] L.A. Zadeh, "The Concept of a Linguistic Variable and Its Application to Approximate Reasoning-1", *Information Sciences*, vol. 8, issue 3, pp. 199-249, 1975. [https://doi.org/10.1016/0020-0255\(75\)90036-5](https://doi.org/10.1016/0020-0255(75)90036-5)
- [11] M. Čubranić-Dobrodolac, T. Molková, L. Švadlenka, "The Impact of Road Characteristics Assessment on The Traffic Accidents Occurrence" *The Proceedings of Sinteza 2019 - International Scientific Conference on Information Technology and Data Related Research*, Belgrade, Singidunum University, Serbia, 2019, pp. 26-31. <https://doi.org/10.15308/Sinteza-2019-26-31>
- [12] JP "Putevi Srbije", Dangerous places at the road sections in the Republic of Serbia, available at: <http://www.putevi-srbije.rs>
- [13] L. Wang, and J. Mendel, "Generating fuzzy rules by learning from examples", *IEEE Transactions on Systems, Man and Cybernetics*, vol. 22, issue 6, pp. 1414-1427. 1992. <https://doi.org/10.1109/21.199466>
- [14] S. Jovčić, P. Průša, M. Dobrodolac, L. Švadlenka, "A proposal for a decision-making tool in third-party logistics (3PL) provider selection based on multi-criteria analysis and the fuzzy approach", *Sustainability*, vol. 11, issue 15, no. 4236, 2019. <https://doi.org/10.3390/su11154236>