OPTIMISATION AND MODELING

THE IMPACT OF ROAD CHARACTERISTICS ASSESSMENT ON THE TRAFFIC ACCIDENTS OCCURRENCE

Marjana Čubranić-
Dobrodolac[1, *],

Tatiana Molková[2],

Libor Švadlenka[2]

[1]Faculty of Transport and Traffic Engineering,
University of Belgrade,
Belgrade, Serbia
[2]Faculty of Transport Engineering,
University of Pardubice,
Pardubice, Czech Republic

Abstract:

This paper proposes a model for assessment of driver's propensity toward traffic accidents based on the implementation of fuzzy logic. The proposed Fuzzy inference system is based on three input variables and one output. Input variables are: The assessment of dangerous places on the observed road section (nine dangerous spots are selected; however, the scores from these nine spots are sublimated in one variable), Assessment of road characteristics (seven road characteristics are chosen; however, the scores from these seven spots are sublimated in one variable) and Frequency (Input variable Frequency is based on the criterion how many times a week or a month the examinee drives on the observed road section). Output variable is The number of accidents that a driver had experienced. The model is tested on the sample of 305 drivers and most of them are professional drivers. The results are perceived through the cumulative error that fuzzy inference system makes in description of empirical data.

Keywords:

traffic accidents, road safety, dangerous spots assessment, road characteristics assessment, fuzzy systems.

1. INTRODUCTION

The vulnerability of people in traffic accidents is one of the biggest threat of nowadays. According to the World Health Organization [1], around 3700 people die on the roads every day. In addition to human losses, traffic accidents also cause significant economic costs. Based on various reports, these costs are in range from 1 to 2% of some country's Gross Domestic Product – GDP.

Some general causes of traffic accidents relate to the following: driver, i.e. human factor, vehicle and road [2,3]. According to the police reports in the Republic of Serbia, the most common circumstances that lead to the occurrence of accidents are: unadjusted speed, wrong assessment of the traffic situation and the performance of improper actions in traffic, as well as the psycho-physical condition of the driver, for example, driving under the influence of alcohol [4].

Correspondence:

Marjana Čubranić-Dobrodolac

e-mail:

marjana@sf.bg.ac.rs

The previous stated indicate the importance of studying the nature of risk behavior of drivers in order to take appropriate measures to increase traffic safety. However, it is also important to analyze the factor of road, and particularly the interaction of driver-road factor. This is the reason why the subject of interest in this paper is the question how the road characteristics assessed by a driver impact the propensity toward traffic accidents. For the purpose of this phenomenon quantification, a model proposed in this paper is based on the implementation of fuzzy logic [5]. The model is tested on the sample of 305 drivers and most of them are professional drivers.

2. MODEL DEVELOPMENT – USED VARIABLES

The model proposed in this paper is based on the implementation of fuzzy logic. Fuzzy inference system is based on three input variables and one output. Input variables are: The assessment of dangerous places on the observed road section, Assessment of road characteristics and Frequency of driving on the road section. Output variable is The number of accidents that a driver had experienced.

The first variable x_1 named *Dangerous spots* is formed based on 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.

The chosen road section is in the Republic of Serbia, road category IB, number 22, which is colloquially known as „Ibarska magistrala". Each of 305 examinees from the sample gave scores about the following nine dangerous spots: Žarkovo, Orlovača, Ripanj, Šiljakovac, Stepojevac, Šopić, Lazarevac roundabout and Ćelije. The final value of the first input variable is a mean of all nine scores. The scores are an assessment given by a driver about to what extent are the considered dangerous spots really dangerous.

Input variable *Dangerous spots* refers 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 the 5 fuzzy sets and their membership functions (Fig. 1): VNPR – very low assessment of risk, NPR – low assessment of risk, SPR – medium assessment of risk, VPR – high assessment of risk, VVPR – very high assessment of risk.

As can be seen from Fig. 1, 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. 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 VNPR has the largest interval compared to the remaining 4 fuzzy sets.
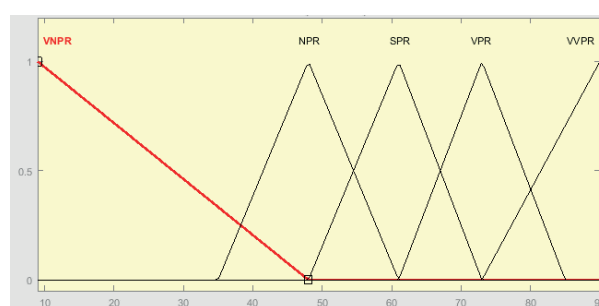

Figure 1. Input variable Dangerous spots

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 the value of 61 was taken as part of fuzzy set SPR with the value of 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. 2: VNPK - very low assessment of road characteristics, NPK - low assessment of road characteristics, SPK - medium assessment of road characteristics, VPK - high assessment of road characteristics, VVPK - very high assessment of road characteristics.
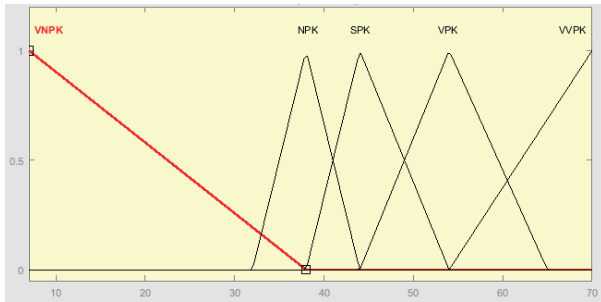
27

Figure 2. 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; on the basis of which the value of 44 was taken as the value of the fuzzy set SPK with the highest degree equal to 1.

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 desribed by the seven fuzzy sets as follows: VVU - very high frequency, VU - high frequency, SVU - medium high frequency, SU - medium frequency, SNU - medium low frequency, NU - low frequency and VNU - very low frequency.
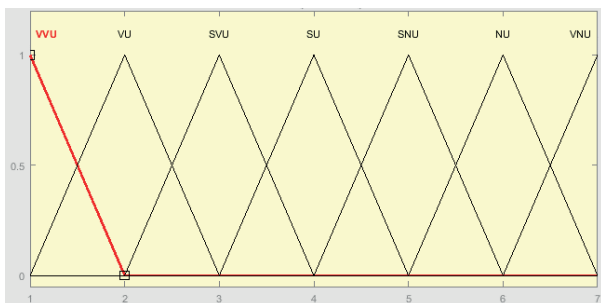

Figure 3. Input variable Frequency of driving

The output variable *y* named *Accidents* relates to the number of traffic accidents that respondents had experienced. In the sample, the examinees reported the number of accidents 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. 4. The following fuzzy sets are introduced: VMBN – very small number of accidents, MBN – small number of accidents, SMBN – moderately small number of accidents, SBN – medium number of accidents, SVBN – moderately high number of accidents, VBN – high number of accidents, VVBN – very high number of accidents.
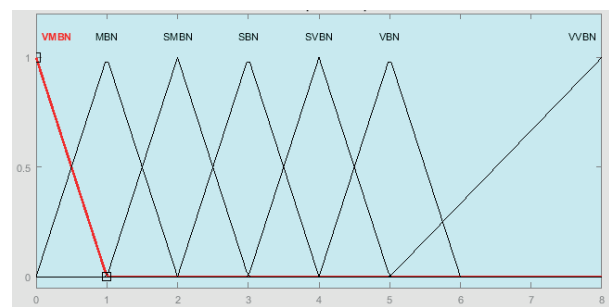

Figure 4. Output variable Accidents

3. MODEL DEVELOPMENT – FORMING FUZZY RULES BASED ON EMPIRICAL DATA

The fuzzy rule base is essential for the performance of fuzzy inference system. In this paper we use well-known approach for defining fuzzy rules proposed by Wang and Mendel [6].

The Wang-Mendel method consists of five steps. Step 1 divides the input and output spaces of the given numerical data into fuzzy regions. In this paper, the implementation of Step 1 is illustrated in the section I. 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 Fig. 1 to 4.

Step 2 generates fuzzy rules from the collected data. The data set is structured as shown in Table I, where letter *i* represents one of 305 respondents from the sample.

Table 1. Data set of input and output values

Examinee	$x_1^{(i)}$	$x_2^{(i)}$	$x_3^{(i)}$	$y^{(i)}$
1	82	39	1	8
2	55	44	1	0
3	52	46	1	0
4	59	47	1	0
305	64	49	3	3

First, one data pair is used for construction of one fuzzy rule. For example, the degrees of a given pair ($x_1^{(i)}$, $x_2^{(i)}$, $x_3^{(i)}$; $y^{(i)}$) should be determined in different regions. Then, this data pair should be assigned to the regions with maximum degree. Finally, one fuzzy rule from one pair of desired input-output data is obtained. IF part is composed of the names of regions with maximum degree for input variables and THEN part from the name of region with maximum degree for output variables. By this, we obtained 305 fuzzy rules. However, there are some rules that are the same. After excluding the same fuzzy rules, there are 131 of them left.

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".

$$D(Rule) = \mu_A(x_1) * \mu_B(x_2) * \mu_A(y) \qquad (1)$$

where $D(Rule)$ is a degree of a rule, $\mu_A(x_1)$ is a value of membership function of the region A when input value is x_1, $\mu_B(x_2)$ is a value of membership function of the region B when 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 maximum degree should be accepted. By implementing the proposed procedure, in our case, from 131 different rules, there are 64 non-conflict fuzzy rules obtained from the collected data.

Step 4 makes a combined fuzzy rule base which consists of rules obtained from empirical data and linguistic rules of human expert. In this case, the complete fuzzy rules base consists of 175 fuzzy rules. Finally, Step 5 determines a mapping from input to output space using a defuzzification procedure.

4. RESULTS AND DISCUSSION

Finally, the defined fuzzy inference system should be tested. This is performed based on Eq. 2. Cumulative deviation (*CD*), as a measure that describes how good the fuzzy inference system describes empirical data, is calculated as an absolute value of difference between the number of accidents which drivers from the sample experienced and corresponding result of fuzzy inference system based on the same input values for the particular driver. The result of fuzzy inference system for an examinee number *i* in the Eq. (2) is marked as *Propensity(i)*.

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

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 fuzzy inference system.

After all the calculations, the final result is CD=844.945. To get a conclusion about the quality of this result, it is necessary to perform further research about causes of traffic accidents and to test new fuzzy inference systems with new variables.

Table 2. Testing different shapes of membership functions

Fuzzy inference system (input membership function – output function)	CD
I (*trimf; trimf; trimf – trimf*)	844.945
II (*trapmf; trapmf; trapmf – trimf*)	853.037
III (*trapmf; trapmf; trapmf – trapmf*)	856.085
IV (*gaussmf; gaussmf; gaussmf – trimf*)	835.211
V (*gaussmf; gaussmf; gaussmf – gaussmf*)	833.732
VI (*gauss2mf ; gauss2mf; gauss2mf – trimf*)	854.587
VII (*gauss2mf; gauss2mf; gauss2mf – gauss2mf*)	856.714
VIII (*gbellmf; gbellmf; gbellmf – trimf*)	848.843
IX (*gbellmf; gbellmf; gbellmf – gbellmf*)	851.726
X ((*zmf, dsigmf, sigmf*); (*zmf, dsigmf, sigmf*); (*zmf, dsigmf, sigmf*) – *trimf*)	829.578
XI ((*zmf, dsigmf, sigmf*); (*zmf, dsigmf, sigmf*); (*zmf, dsigmf, sigmf*) – (*zmf, dsigmf, sigmf*))	828.415

XII ((*zmf, psigmf, sigmf*); (*zmf, psigmf, sigmf*); (*zmf, psigmf, sigmf*) – *trimf*)	829.293
XIII ((*zmf, psigmf, sigmf*); (*zmf, psigmf, sigmf*); (*zmf, psigmf, sigmf*) – (*zmf, psigmf, sigmf*))	828.163
XIV ((*zmf, pimf, smf*); (*zmf, pimf, smf*); (*zmf, pimf, smf*) – *trimf*)	865.620
XV ((*zmf, pimf, smf*); (*zmf, pimf, smf*); (*zmf, pimf, smf*) – (*zmf, pimf, smf*))	873.090
XVI (*trimf*; (*zmf, psigmf, sigmf*); (*zmf, psigmf, sigmf*) – *trimf*)	827.565
XVII (*trimf*; (*zmf, psigmf, sigmf*); (*zmf, psigmf, sigmf*) – (*zmf, psigmf, sigmf*))	826.807

Besides, it is possible to optimize the results of the current fuzzy inference system. This should be examined by changing fuzzy rules, intervals and shapes of membership functions, modifying the method of defuzzification. In this paper, we test the impact of changing the shape of membership functions and other testing may be seen as a direction for further research.

The various membership functions tested in this research are shown in Table II. The first is triangular membership function, and it has the function name *trimf*. The trapezoidal membership function is named *trapmf*.

Two membership functions are built on the Gaussian distribution curve: a simple Gaussian curve and a two-sided composite of two different Gaussian curves. Their names are *gaussmf* and *gauss2mf*.

The generalized bell membership function is specified by three parameters and has the function name *gbellmf*. The next is sigmoidal membership function, which is either open left or right. Asymmetric and closed membership functions can be synthesized using two sigmoidal functions, so in addition to the basic *sigmf*, there is a difference between two sigmoidal functions, *dsigmf*, and the product of two sigmoidal functions *psigmf*.

Polynomial based curves that make three related membership functions are the Z, S, and Pi curves, all named because of their shape. The function *zmf* is the asymmetrical polynomial curve open to the left, *smf* is the mirror-image function that opens to the right, and *pimf* is zero on both extremes with a rise in the middle.

The results of testing various shapes of membership functions are shown in Fig. 5. As it can be noticed, the best result is achieved by the fuzzy inference system No.

XVII. The results of this system are compared with the empirical data which is illustrated in Figure 6. By analyzing the obtained results, an assumption is that there is a space for improvement of the proposed model by introducing new variables that impact the occurrence of traffic accidents.
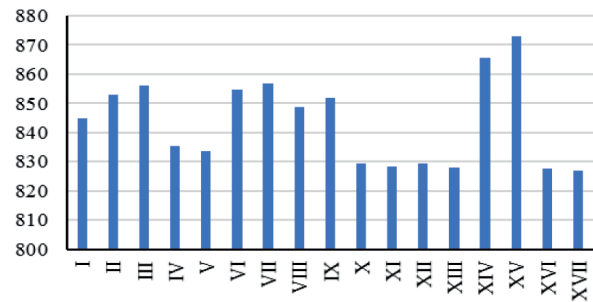

Figure 5. Comparison of various fuzzy inference systems

A possible explanation about the fact that the fuzzy inference system No. XVII does not fit "so well" to the empirical data may be found in the fact that the causes of traffic accidents are very heterogeneous. For example, the personality traits may have a significant role [7]. Therefore, to explain the considered phenomenon more precise, further research is needed.
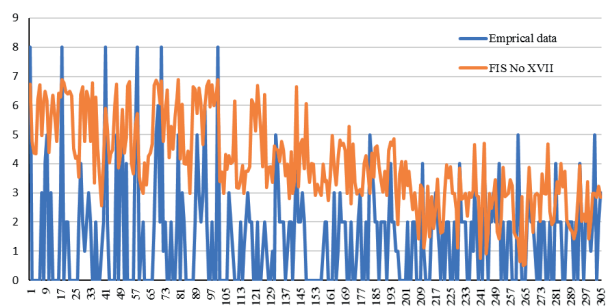

Figure 6. Comparison of empirical data and results of FIS No. XVII

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/ Accessed 12 March 2019.

[2] L. Evans, Traffic safety. Michigan: Science Serving Society, 2004.

[3] W. Vanlaar, G. Yannis, "Perception of road accident causes", Accident Analysis & Prevention, vol. 38, issue 1, pp. 155-161, 2006. https://doi.org/10.1016/j.aap.2005.08.007

[4] M. Cabarkapa, M. Cubranic-Dobrodolac, S. Cicevic, B. Antic, "The influence of aggressive driving behavior and impulsiveness on traffic accidents", vol. 8, issue 3, pp. 372 - 389. 2018. https://doi.org/10.7708/ijtte.2018.8(3).09

[5] L.A. Zadeh, "Fuzzy sets", Information and Control, vol. 8, issue 3, pp. 338-353, 1965. https://doi.org/10.1016/S0019-9958(65)90241-X

[6] L. Wang, and J. Mendel, "Generating fuzzy rules by learning from examples", IEEE Transactions on Systems, Man and Cybernetics, vol. 22, isuue 6, pp. 1414–1427. 1992. https://doi.org/10.1109/21.199466

[7] M. Cubranic-Dobrodolac, K. Lipovac, S. Cicevic, and B. Antic, "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