American Academic Scientific Research Journal for Engineering, Technology, and Sciences
ISSN (Print) 2313-4410, ISSN (Online) 2313-4402
http://asrjetsjournal.org/

Fuzzy Rule-Based System as a Method of Modeling for Estimation Quality of the Vardar River

Vesna Antoska Knights[a*], Zoran Gacovski[b]

[a]*University "St. Kliment Ohridski", Faculty of Technology and Technical Science, 1 Maj bb, Bitola, Republic of North Macedonia*

[b]*Mother Tereza University, Faculty of Technical Science, Mirche Acev No.4, Skopje, Republic of North Macedonia*

[a]*Email: vesna.knights@uklo.edu.mk*

[b]*Email: zoran.gacovski@unt.edu.mk*

Abstract

The paper proposes the use of a fuzzy rule-based system to evaluate the quality of the Vardar river for irrigation purposes, based on the presence of two types of pollutants (bacteriological and chemical) and the contaminated area in the Skopje-Polog region. The input variables of the fuzzy inference system are the quantity of the pollutants detected and the percentage of the contaminated area, while the output variable is the estimated quality of the river Vardar for irrigation. The results of the system are expressed as a fuzzy set, allowing for uncertainty and vagueness in the assessment of the river water quality. The proposed method seeks to provide a more comprehensive and accurate evaluation of the Vardar river water quality for irrigation. Used data obtained from the Institute of Public Health of the Republic of North Macedonia for 2022. The paper presents a fuzzy rule-based system to estimate the levels of chemicals and bacteria in the Vardar River, which is still used for irrigation. The method employs fuzzy inference and defuzzification to generate the output. The estimated values of pollutants in the river have not been measured by either the Ministry of the Environment or any non-governmental organizations. The proposed model provides a valuable basis for simulating future concentrations of these substances in the water for the years to come, making a significant contribution to the understanding of the Vardar River's water quality for irrigation purposes.

Keywords: Fuzzy logic; Fuzzy simulations; Chemicals and bacteriological pollutants.

1. Introduction

One of the main problems of life today is environmental pollution. Despite the appeals of many non-governmental organizations and people concerned about the future of their environment and the planet, pollution is still a major problem. We live in modern times when industry and technology are advancing daily, but this is not always in balance with environmental protection, especially in developing countries.

----

* Corresponding author.

The Vardar River runs along the Republic of North Macedonia from valley to valley. In the 19th century, around it were the most fertile fields from which people were fed. In the last 20-25 years, chemicals from industrial facilities have been pouring into this river, and farmers still use it to irrigate their fields. It is knowns that chemical occurrence in the surface soil and underground waters is due to several interlinked processes like deposition, re-suspension, adsorption, chemical reaction, biological uptake, decay, and percolation. Several attempts for quantitative analysis of the chemical impact on surface waters have been made. Reference [1] have estimated the annual pollution transport as a percent of use (load as a percent of use - LAPU). Reference [3] have developed a regression model for pollution estimation. Binoy and Mujumdar [4] have developed fuzzy model for estimating the pollution in the White River (Indiana, USA).

2. Materials and Methods

This paper aims to estimate the effect of two prevalent pollutants - chemicals and bacteria - on the water quality in the Vardar River in the region of Skopje and Polog. The data used for the study was obtained from the Institute for Public Health for the year 2022. A fuzzy rule-based system is established, where the input variables are the presence of chemicals per unit of water and the percentage of bacteria found. The output of the system is an estimate of the pollutants in 1 liter of river water at the lower stream, which is the site used for irrigation. The use of the fuzzy rule-based system allows for a more precise assessment of the water quality in the Vardar River for irrigation purposes.

The study's objective is to use a fuzzy inference method to model the hazy transport of pollution from water to soil. To calculate the average monthly pollution levels (chemical concentrations) in the Vardar River at the outflows of the Skopje and Polog regions, a fuzzy rule-based model is developed. It is possible to depict the uncertainty and fluctuation in the river's pollution levels more accurately by using fuzzy relations. This model is a useful tool for evaluating the potential effects of pollution on the soil and water quality of the Vardar River.

2.1. Methodology of a fuzzy inference system (FIS)

Instead of using Boolean logic to make decisions about data, a FIS uses a collection of fuzzy membership functions and rules. The rules in a fuzzy inference system often take the following form:

“If the concentration of pollutant X is high and the percentage of contaminated area y is significant then the water quality is poor"

In fuzzy inference systems, the inputs are transformed into fuzzy sets and the rules are used to manipulate these fuzzy sets to produce a final output, which can also be represented as a fuzzy set. The output is then defuzzified to obtain a crisp value that represents the final result of the inference process. The use of fuzzy sets and rules allows for the representation of uncertainty, vagueness, and subjectivity in the data, making it well-suited for complex reasoning and decision-making tasks.

The rule basis or knowledge base refers to the collection of rules in a fuzzy inference system. There are three (or four) steps in the universal inference process.

The first step is *fuzzification*, which maps the crisp input values to degrees of membership in fuzzy subsets.

The second step is *inference*, which uses the rules in the rule base to determine the degree to which each rule applies, based on the fuzzy input.

The third step is *defuzzification*, which aggregates the results of the inference process to produce a crisp output.

The fourth step (optional) is the *evaluation* of the output, where the results are analyzed and compared to the desired results. The FIS structure is usually composed of a fuzzifier, a rule base, an inference engine, a defuzzifier, and an evaluation module.

Assume that the variables *x, y,* and *z (water quality)* all take on values in the interval *[0,10]*, and that the following membership functions and rules are defined:

$low(t) = 1 - (t / 10); \; high(t) = t /10$

Rule 1: If the *concentration of pollutant* x is low and the *percentage of contaminated area y* is low, then the *water quality is* high.

Rule 2: If the *concentration of pollutant* x is low and the *percentage of contaminated area y* is high, then the *water quality is* low.

Rule 3: If the value of *concentration of pollutant* x is high and *percentage of contaminated area y* is low, then *water quality is* low.

Rule 4: If the *concentration of pollutant* x is high and the *percentage of contaminated area y* is high, then the *water quality is* high.

Fuzzy logic is a rule-based system that uses fuzzy logic to reason about knowledge, making it possible to represent imprecise information. It works by assigning each variable a subset of the possible values on a given assignment space. This allows for more flexibility and accuracy when making decisions that involve situations with unclear or uncertain data. For example value to the output variable *water quality*, each rule assigns an entire fuzzy subset as a low or high.

1. In this example, $low(t)+high(t)=1.0$ for all *t*. This is not required, but it is fairly common.

2. The same membership functions are used for all variables. This isn't required and is also 'not' common.

In the fuzzification stage, the actual values of the input variables are transformed into fuzzy membership values by applying the membership functions defined for those variables. These membership values determine the degree to which the rule premise is true, also known as the rule's alpha value. If the alpha value is non-zero, meaning the rule premise is partially or fully true, then the rule is considered to "fire."

The inference sub-process combines the fuzzy sets assigned to each output variable by each rule. This is done using a combination method, such as the "Maximum" method, which outputs the fuzzy set with the highest degree of truth, or the "Weighted Average" method, which outputs a fuzzy set that is a weighted average of the fuzzy sets assigned to each output variable.

MIN and PROD. are two *inference methods* or *inference rules*. In MIN inference, the output membership function is clipped off at a height corresponding to the rule premise's computed degree of truth. PROD. inference, the output membership function is scaled by the rule premise's computed degree of truth. This corresponds to the traditional interpretation of the fuzzy logic AND operation, where the output membership function is "magnified" by the degree to which the premise is true. The choice of which inference method to use will depend on the specific problem and design requirements of the fuzzy inference system.

MAX composition and SUM composition are two *composition rules*. In MAX composition, the combined output fuzzy subset is constructed by taking the point-wise maximum over all of the fuzzy subsets assigned to the output variable by the inference rule.

In SUM composition, the combined output fuzzy subset is constructed by taking the point-wise sum over all of the fuzzy subsets assigned to the output variable by the inference rule. Note that this can result in truth values greater than one! For this reason, SUM composition is only used when it will be followed by a defuzzification method, such as the *centroid* method, that doesn't have a problem with this odd case. Otherwise, SUM composition can be combined with normalization and is therefore a general-purpose method again [6].

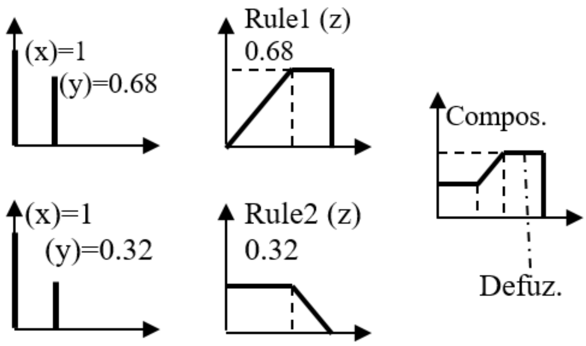

Figure 1: FIS structure: Min-inference, Max-composition [6].

2.2. Description of the model

The model used in this paper is based on information on the presence of contaminants (chemicals and bacteria) in river water. The most significant causes of river pollution are chemicals and bacteria [1]. The Institute of Public Health - Skopje provided the quantities of the current chemicals and microorganisms for the period of January to December 2022. They are used by the fuzzy inference method to forecast the irrigation water quality for the forthcoming season in 2023.

The initial input to our fuzzy model is the chemical concentration rate [%]. The second input for the model is the percentage of bacteria present. Table 1 provides the condensed data for the Skopje region for the months of January to December 2022.

The evaluation of the water quality (of the Vardar River) throughout the year is crucial for the completion of our model. The measurements were taken monthly in 2022 by the Institute for Public Health - Skopje. We used their data to assess the river quality for irrigation in its lower stream and to define the output variable in our model.

We take the average chemical concentration in water [%] as an output variable from our model. At first, we used three-valued triangle membership functions (low, medium, high). Later, to improve accuracy, the number of values was raised and the breadth of the membership functions was altered.

Additionally, we applied other membership functions (including trapezoidal and Gaussian), and they all produced identical results for the output variable (herbicide concentration in the river). The model is trained using the data that is currently available from January through December 2022.

Table 1: Average monthly chemicals and bacteria present in Skopje-Polog region.

Month	Chemicals (%)	Bacteriological (%)
February	23.61	26.57
April	19.37	21.05
June	21.25	23.84
August	16,96	15.44
October	23.77	27.53
December	16.37	19,06

3. Results and Discussion

The results of the study for the first input (chemical concentration) and those for the second are extremely comparable, according to simulations that have been done.

The linguistic values for the input variable Chemicals are "low, med, hi, vh," while the crisp values fall within the range [0, 30] % and membership functions for the first input variable for Chemicals (figure 2).

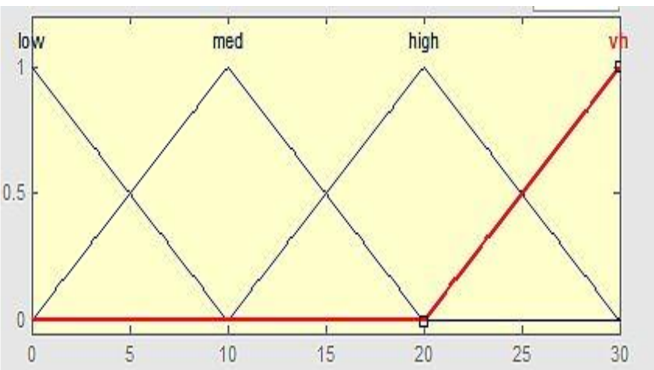

Figure 2: Membership functions for the first input variable (chemicals concentration).

The linguistic values of the input variable Bacteria are "low, med, hi, vh," while the crisp values are within the range [0,30%]. The linguistic values for the output variable Quality are "low, med, hi, vh," and the crisp values fall within the range [0, 0.5]. The membership functions for the output variable (water quality), (figure 3).

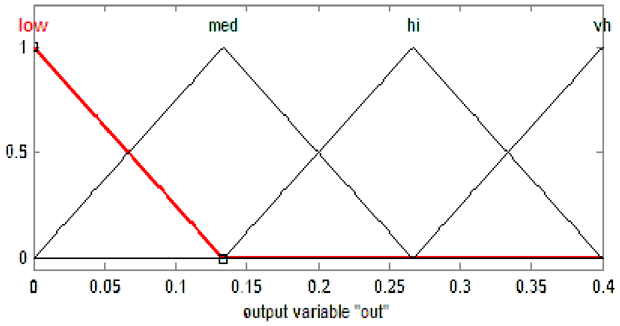

Figure 3: Membership functions for the output variable (water quality for irrigation).

The fuzzy rule base is given with following four rules:

If *chem=med* and *bact=hi* then *qual=low*

If *chem=hi* and *bact=vh* then *qual=low*

If *chem=low* and *bact=med* then *qual=med*

If *chem=low* and *bact=hi* then *qual=low*

Presents the findings for the river's water quality in 2023 for estimated values (figure 4).

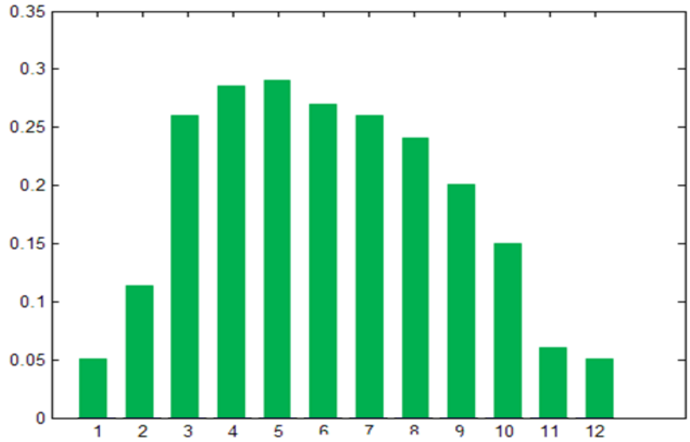

Figure 4: Estimation of the river quality for the year 2023.

Presents the typical river quality results for the upcoming year i.e. 2023 (figure 5).

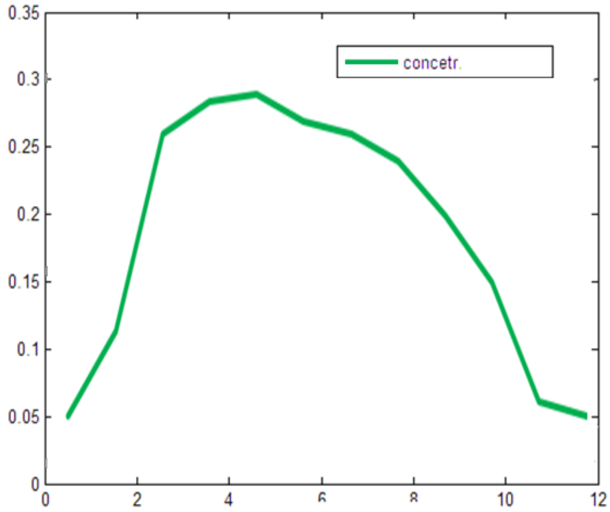

Figure 5: Estimated water quality of the river for the year 2023.

It is obvious that the water quality will never be higher than 0.3 and that it must be above 0.7 in order to be used for irrigation (figure 5).

Thus, our analysis supports the fact that the Vardar River's quality is unfit for irrigation (harmful). Similar methods can be used to get the findings for the bacteria concentration.

4. Conclusion

In this study, is proposed a method to evaluate the pollution level in the Vardar river caused by industrial waste and household sewage in the Skopje-Polog region.

Our approach is based on a fuzzy rule-based system, where the input variables are the percentage of chemicals and bacteria present in the water.

The output variable of the system is the estimated water quality for irrigation in 2023. The study focuses on two common pollutants, chemicals and bacteria, and employs data obtained from the Institute of Public health for the year 2022.

The methodology involves using a fuzzy rule base and defuzzifying the output variable to arrive at the final estimation.

The main contribution of this study is the calculation of pollution levels in the Vardar River, which have not been previously reported by either the Ministry of the Environment or any non-government organizations. Our results indicate that the water quality is low and not suitable for irrigation in 2023.

The model we have developed allows us to predict the future water quality in the river, offering valuable insights for the years to come.

References

[1]. Capel P. and Larson S. (2001), “Effect of scale on the behavior of atrazine in surface waters”, *Environ. Science Tech.*, Vol. 35, No. 4, pp. 648-657.

[2]. Crawfard C. (2001), “Factors affecting pesticide occurrence and transport in a large mid-western river basin”, *Journal of American Water res.*, Vol. 37, No. 5, pp. 1-15.

[3]. Larson S. and Gilliom R. (2001), “Regression models for estimating herbicides concentrations in U.S. streams from watershed characteristics”, *Journal of American Water res.*, Vol. 37, No. 5, pp. 1349-1367.

[4]. Binoy M. and Mujumdar P. (2003), “Fuzzy rule-based model for estimating agricultural diffuse pollution”, In. *Proc. of Diffuse Pollution Conference*, pp. 72-76, Dublin.

[5]. Gacovski Z. and Cvetanovski I. (2005), “Selection of distribution center location via fuzzy decision-making”, In *Proc. of KOREMA Conference*, pp. 87-91, Zagreb, Croatia.

[6]. Gacovski Z. Ilievski I, Gelev S, Deskovski S. (2007), “Fuzzy interference system applied for calculation of Environmental pollution”, ETAI National *Conference*, Ohrid, Macedonia.