ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20230713121334

Original scientific paper
Received: 2023-07-13, Accepted: 2025-09-08

Neuro Fuzzy – Regression Based Concept for Dosing Determination

Mirche Stojkov*, Zoran Pandilov

Abstract: This research aims to optimize dosing through the application of intelligent tools-methods. Here, the case study focuses on the flocculation phase of drinking water processing in a real system. Adaptive Neuro Fuzzy Inference System (ANFIS), regression, and "know-how" of the operators represented through linguistic expression as experiential knowledge are integrated to create a dosing calculation program. The goal is to establish a functional relationship between the process variables.

Keywords: Adaptive Neuro Fuzzy Inference System (ANFIS); dosing; experiential knowledge; linguistic expression; regression

1 INTRODUCTION

When treating drinking water, it is important to meet two key requirements: to ensure that the water is safe for consumption and to minimize the costs of the treatment process. Overdose is increasing costs, while insufficient dosing disrupts water treatment process. By using the correct amount of coagulant-flocculent, two important goals are achieved. The first objective is to improve the efficiency of flocculation, which is an important step for water treatment. The second objective is to reduce the costs associated with the process. Typically, operators rely on their experience to determine the appropriate amount of dosing in processing facilities.

The JAR test it simulates the coagulation-flocculation and sedimentation process in a water treatment plant and helps operators determine if they are using the right amount of treatment chemicals. However, in reality, the determination is delayed because it relies on the already changed properties of the raw water. Operators use their experience to verbally group certain values of the variables noted in the worksheets. In past in most cases, intelligent tools-methods were used for the purpose of optimization and modelling in water processing systems. This research integrates the use of these methods. In addition, these methods have application in other various fields. In researches [1-9] are shown the use of methods: ANN, Fuzzy Logic and Regression like valuable tools for optimization, prediction and computing. In a similar studies like our research, input data from surface water and operational data from dosage, are used to create models for the dosing prediction and control also using: ANN, Fuzzy Logic and Regression Analysis [10-16]. The primary goal of this study is to create a model (program) that could serve as a control unit. In this study, flow and turbidity of raw water and dosage of flocculent solution are used as input-output variables. To establish the model, 240 pairs or 960 data were collected from the "Ilovica" treatment plant in Bosilovo, Republic of Macedonia, over of four years. This was done to provide a wide range of values for the variables and to cover different processing periods.

2 METHODOLOGY

2.1 Water Treatment Process

The water treatment plant "Ilovica" has the ability to treat 3840 cubic meters of water per day or a maximum of 44 litters per second. The treatment process is fully automated and controlled by a Programmable Logic Controller, the stages (phases) are presented in Fig. 1.

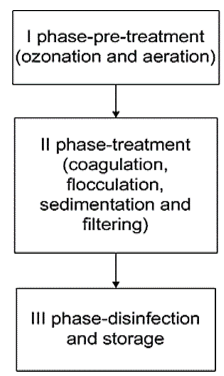

Figure 1 Different phasess of water treatment

In order to effectively control the process, it is important to measure and record the variables being monitored. The way these variables are measured is determined by the existing monitoring system. Some variables are measured in real time using sensors embedded in the system, while others are measured manually using portable measuring devices.

a) Flow-meter b) Turbidity-meter c) Dosing pump

Figure 2 Built-in sensors

2.2 Knowledge by Operators

The goal of using experiential knowledge to find a solution is to relate the values of the input variables to the values of the output variables [17]. The hidden knowledge, represented by historical data can be expressed as a vector (x_i, y_i), i = 1, 2, ..., n (n - number of pairs). Experienced operators can use their knowledge to determine processing rules and define the values of variables in groups, through linguistic expression. The number of selected if-then rules depends on the number of selected groups (subsets). This is shown in Fig. 3. Four distinct groups or subsets represent the qualitative definition of the range of values for the two input variables: small (C_1), medium-small (C_2), medium (C_3), big (C_4) for the flow of the raw water (x_1) and low (D_1), medium-low (D_2), medium (D_3), high (D_4) for the turbidity of the raw water (x_2). The optimal connection of input-output variables (flow and turbidity of raw water like input variables and amount of dosage of flocculent solution like output variable) in our case is set with "if-then" rules (Fig. 3). The calculations for each rule are done using Eq. (1) separately.

If x_1 is in A_i and x_2 is in B_j and i = 1, 2, ..., n and j = 1, 2, ..., n (n = 4) then,

$$y = x_1 B_{1(i,j)} + x_2 B_{2(i,j)} + B_{0(i,j)} \tag{1}$$

Where: A_i - groups (flow of raw water); B_j - groups (turbidity of raw water); x_1, x_2 - input variables (flow and turbidity of raw water); y - output variable (amount of dosage of flocculent solution); B_0, B_1, B_2 – coefficients; n - number of groups (subsets).

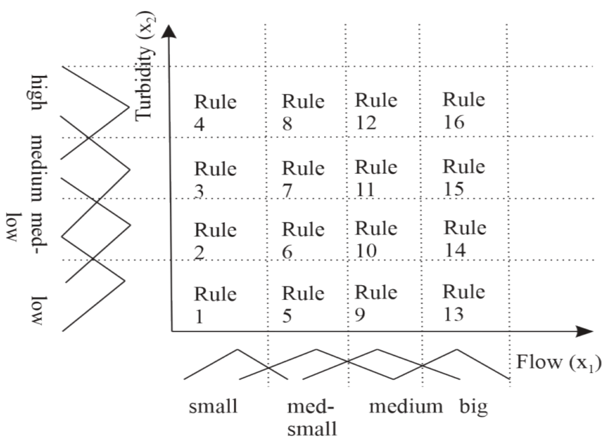

Figure 3 Selected "if-then" rules

Before creating a non-uniform network partition, the number of variables and groups (subsets) to be used are determined. Then, learning methods are used to refine and optimize the previous parameters [18]. This "fine-tuning" capability is used here, in order to determine the boundaries of each of the groups (subsets), as an intersection between two membership functions (Fig. 3). The boundaries of each group (subset) are used in dosing calculation program and are determined by arithmetic calculation using the values of the membership function parameters. Obtaining the optimized parameters was done by using an intelligent ANFIS (Adaptive Neuro Fuzzy Inference System) tool.

2.3 Data Preparation

If any of the variables listed in Tab. 1 change, it implies a change in the water treatment process in real time. In our case, water treatment is done by rapid treatment and pressure filtration, and as emphasized earlier, any change in the input variables directly affects the water treatment. The relationship between the input and output variables is shown by a functional dependency:

$$y = B_0 + x_1 B_1 + x_2 B_2 \tag{2}$$

Where: y - output variable; x_1, x_2 - input variables; B_0, B_1, B_2 - coefficients.

Table 1 Observed values of variables

	Flow	Turbidity	Dosing of flocculent
	m^3/h	NTU	l/h
Maximum	161	99	130
Average	83	53	51
Minimum	29	6	16

The collected input-output pairs for a given criterion (a turbidity value of less than 0.100 NTU) it depends on the processing regime and the total volume in the treatment stage:

$$NS = \frac{VT}{RP} \tag{3}$$

Where: NS - number of samples in one observation; VT - total volume of phase treatment; RP - regime of processing.

Tab. 2 shows the water treatment stage, which consist of coagulation and flocculation, sedimentation and filtration volumes.

Table 2 Total volume of phase treatment

Steps of phase treatment	Volume (m^3)
Coagulation and flocculation	130.4
Sedimentation	120.4
Filtering	26.7
Total:	277.5

2.4 ANFIS (Adaptive Neuro Fuzzy Inference System)

ANFIS (Adaptive Neuro Fuzzy Inference System), introduced by Jang, is a rule-based approximation tool [18]. A specific Sugeno type fuzzy system is presented, which is implemented in a five-layer feed-forward network structure [19]. ANFIS uses "if-then" rules to determine output based on input values [20]. ANFIS as an intelligent tool uses observed and historical data from operational lists to train. Such a tool can also be used in water treatment systems [21]. The ANFIS architecture we are using is a zero-order Sugeno fuzzy model. It includes two inputs, x_1 and x_2, and one output, y. In Fig. 4 is shown the general structure of ANFIS.

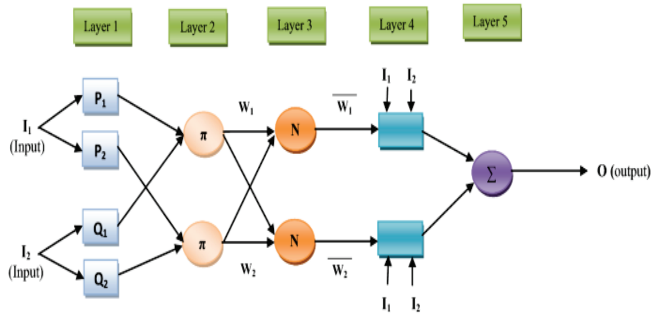

Figure 4 Structure of ANFIS (Adaptive Neuro Fuzzy Inference System)

The hybrid Neuro Fuzzy tool ANFIS optimizes the parameters. These optimized parameters are referred to as "previous" parameters, while the constant values for the output are known as "consequential" parameters. ANFIS learning involves a combination of back propagation and Least Squares Estimation (LSE) [20]. Using the Fuzzy Logic Toolbox, the optimized parameters were obtained. To create a fuzzy approximation system, a network partition is imitated. The values of "previous" parameters that define the triangular fuzzy membership functions, obtained after optimization and which are used in the model shown in Fig. 5, are shown in Tab. 3.

Table 3 Triangular fuzzy membership functions

Linguistic expression	Values of parameters
Small	[3.6; 4.78; 5.94]
Medium-small	[4.76; 5.94; 7.17]
Medium	[5.95; 7.09; 8.30]
Big	[7.13; 8.27; 9.45]
Low	[0.04; 4.92; 9.88]
Medium-low	[4.92; 9.89; 4.84]
Medium	[9.87; 14.84; 19.8]
High	[14.84; 19.8; 24.76]

2.5 Linear Regression

Linear regression is a method-tool used to identify the relationship between variables and the strength of that relationship. Dependence can be functional and statistical.

The relation represents functional dependence:

$$u = f(v) \tag{4}$$

The relationship is shown by the equation:

$$u = f(v) + e \tag{5}$$

Where: $f(v)$ is a function component, and e - error component.

Here, we will use a multiple linear regression model. This model is represented by a general formula:

$$u = q_0 + q_1 p_1 + q_2 p_2 + \ldots + q_i p_i + e \tag{6}$$

The variable u is dependent in the given relationship, p_i are independent variables, q_0, q_1, q_2, …, q_i are coefficients and e is stochastic variable (error).

The functional dependence between the variable u and variables p_i is represented as:

$$u = q_0 + q_1 p_1 + q_2 p_2 + \ldots + q_i p_i \tag{7}$$

A linear regression model shows how the variable u is related to the variables p_i, determined using samples (pairs). This model can be expressed as a system of n-relations.

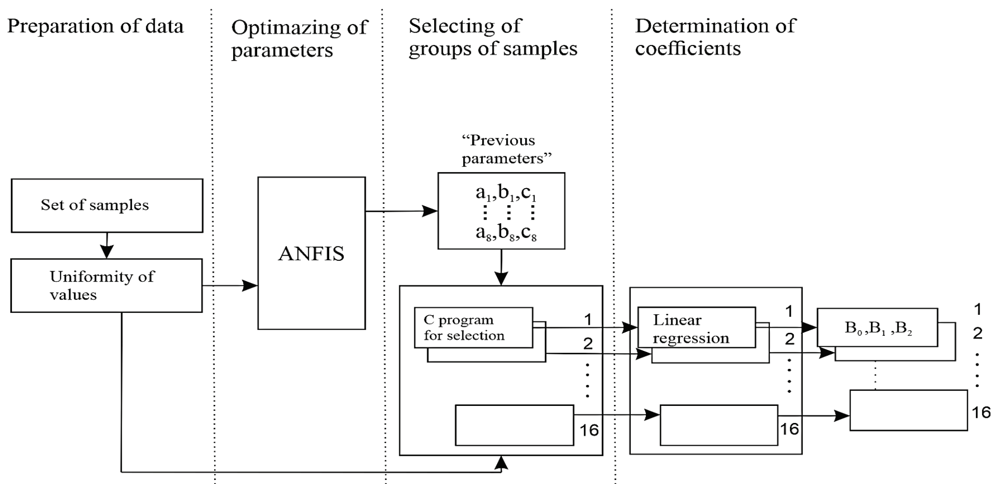

Figure 5 Path-procedure for obtaining the coefficients (ANFIS–Regression model)

In this research is used Regression in Excel to find the coefficients B_0, B_1 and B_2 used in Eq. (1). To determine the number of pairs for each rule, an algorithm-program was developed in the C programming language. Obtaining the coefficients applied in the dosing calculation program (Fig. 13) is show in the Fig. 5.

3 RESULTS

The performance and indication of accuracy of the model is expressed through Mean Absolute Percentage Error (MAPE) Eq. (8) and with coefficient of determination R^2 Eq. (9), using 240 samples (pairs). The results are shown in Tab. 4.

Table 4 Model performance

Model	*MAPE*	R^2
ANFIS-Regression	6.3	0.980

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|o_i - d_i|}{o_i}\times 100 \tag{8}$$

Where: $MAPE$ - Mean Absolute Percentage Error; o - observed values; d - determined (predicted) values; n - number of samples.

$$R^2 = 1 - \left(\frac{\sum_{i=1}^{n}(o_i - d_i)^2}{\sum_{i=1}^{n} o_i^2}\right),\ i = 1, 2, ..., n \tag{9}$$

Where: R^2 - coefficient of determination; o - observed values; d - determined (predicted) values; n - number of samples.

The results of observed values of input and output variables shown in Figs. 6, 7, 8, and the determined (calculated) values of dosing of flocculent (0.15% solution of anionic flocculent) shown in Fig. 9. The calculation of dosing values of flocculent with Rule Viewer from MATLAB Fuzzy Logic Toolbox is shown in Fig. 10**.**

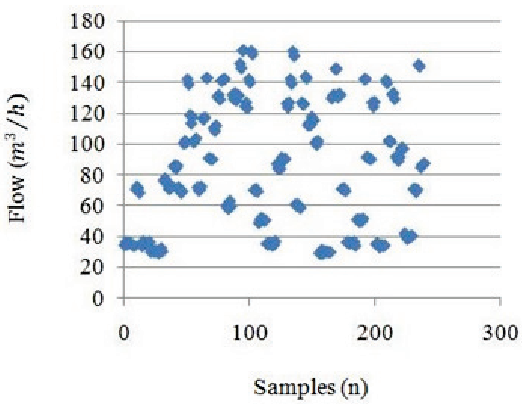

Figure 6 Observed values of flow of row water

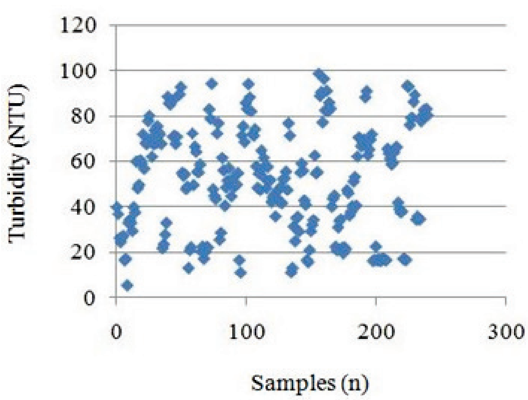

Figure 7 Observed values of turbidity of row water

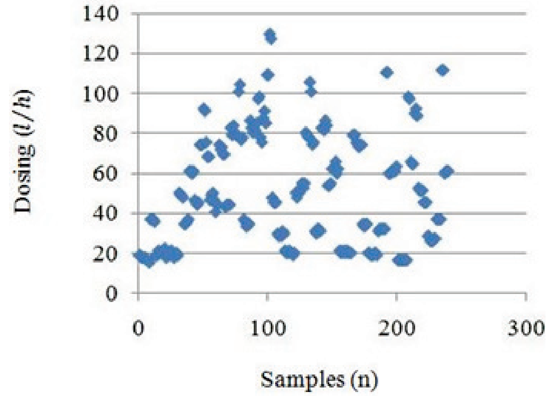

Figure 8 Observed dosing values of flocculent

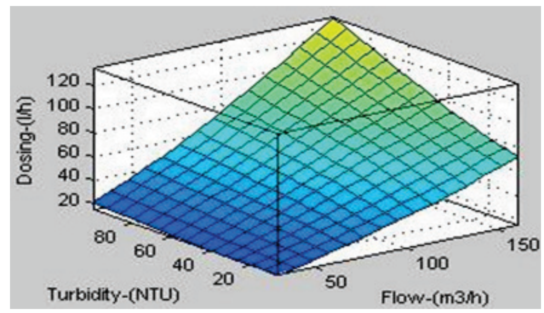

Figure 9 ANFIS surface (determined flocculent)

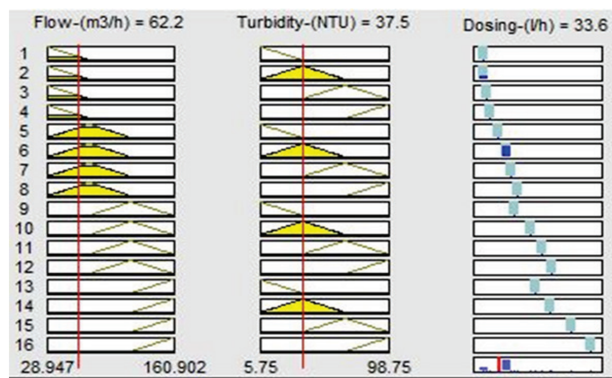

Figure 10 Rule Viewer (determination of flocculent)

4 CONCEPT AND GRAPHIC PRESENTATION

Fig. 11 shows the concept of incorporating a dose calculation program into the dosing system as a control unit. Here the concept for dose determination, which graphically is presented trough diagram in Fig. 12.

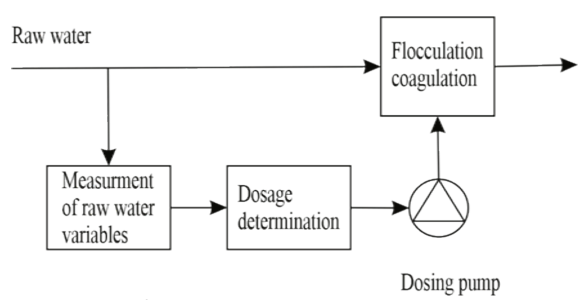

Figure 11 Controlling the dosing based on the input variables

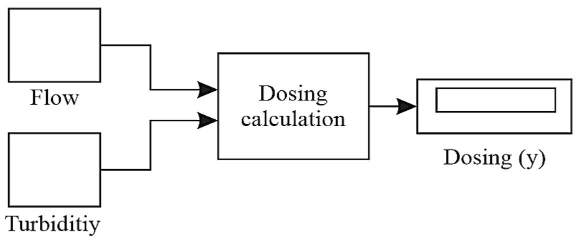

Figure 12 Concept of dosing system

```
IF x1 <= 5.35 AND x2 <= 7.40 THEN
   y := (x2 * 0.066) - (x1 * 0.009) + 5.301;
ELSIF x1 <= 5.35 AND x2 > 7.40 AND x2 <= 12.35 THEN
   y := (x1 * 0.649) + (x2 * 0.064) + 2.076;
 ELSIF x1 <= 5.35 AND x2 > 12.35 AND x2 <= 17.32 THEN
   y := (x1 * 2.455) + (x2 * 0.080) - 7.074;
ELSIF x1 <= 5.35 AND x2 > 17.32 THEN
   y := (x1 * 2.591) + (x2 * 0.068) - 7.582;
ELSIF x1 > 5.35 AND x1 <= 6.53 AND x2 <= 7.40 THEN
   y := (x1 * 1.988) + (x2 * 0.280) - 6.300;
 ELSIF x1 > 5.35 AND x1 <= 6.53 AND x2 > 7.40 AND x2 <= 12.35 THEN
   y := (x1 * 2.335) + (x2 * 0.172) - 7.556;
ELSIF x1 > 5.35 AND x1 <= 6.53 AND x2 > 12.35 AND x2 <= 17.32 THEN
   y := (x1 * 2.706) + (x2 * 0.142) - 9.428;
 ELSIF x1 > 5.35 AND x1 <= 6.53 AND x2 > 17.32 THEN
   y := (x1 * 1.911) + (x2 * 0.111) - 3.832;
ELSIF x1 > 6.53 AND x1 <= 7.70 AND x2 <= 7.40 THEN
   y := (x1 * 2.085) + (x2 * 0.153) - 6.157
ELSIF x1 > 6.53 AND x1 <= 7.70 AND x2 > 7.40 AND x2 <= 12.35 THEN
   y := (x1 * 2.667) + (x2 * 0.324) - 11.548;
ELSIF x1 > 6.53 AND x1 <= 7.70 AND x2 > 12.35 AND x2 <= 17.32 THEN
   y := (x1 * 2.932) + (x2 * 0.297) - 13.172;
ELSIF x1 > 6.53 AND x1 <= 7.70 AND x2 > 17.32 THEN
   y := (x1 * 3.260) + (x2 * 0.004) - 10.366;
 ELSIF x1 > 7.70 AND x2 <= 7.40 THEN
   y := (x1 * 2.251) + (x2 * 0.356) - 9.079;
ELSIF x1 > 7.70 AND x2 > 7.40 AND x2 <= 12.35 THEN
   y := (x1 * 2.147) + (x2 * 0.435) - 8.750;
ELSIF x1 > 7.70 AND x2 > 12.35 AND x2 <= 17.32 THEN
   y := (x1 * 2.217) + (x2 * 0.356) - 8.410
 ELSE
y := (x1 * 3.971) + (x2 * 0.066) - 16.981;

END_IF;
```

Figure 13 Dosing calculation program "if-then" form

5 CONCLUSION

The possibility of "flexible" control and application of the built-in equipment in a modern processing system has a crucial role in the fulfilment of the final task. This research aims to apply the world's scientific tendencies in the use of intelligent tools (methods) in the direction of advancement and improvement of the operation as a whole through optimization and modelling by applying the numerous data in the form of input-output variables. In this paper is presents a concept for determining the dose of flocculent in a real water treatment system. Since there is no generally accepted mathematical correlation, i.e. the characteristics of coagulation-flocculation in water treatment systems are different; an important aspect in this research was the establishment of functional dependence between some of the "input" parameters (variables) with the chemical-physical process of coagulation-flocculation. In this paper, a dosing calculation program ("ELSIF"-statement), is developed in PLC platform "Codesys" in Structured Text (ST), with the possibility of applicability in a logic controller (Fig. 13). The model was obtained by integrating hybrid Neuro Fuzzy tool ANFIS and Regression. The statistical indication, Mean Absolute Error (MAPE) and coefficient of determination R^2 was used in order to determine the performance of the applied model.

The disadvantage of this model is the exponential increase in processing rules with the addition of variables.

Advantages of this research, as well as future recommendations are:

A simplified "if-then" dosing program for use in logic controllers.

Possibility of implementation of other input-output parameters (variables) for water treatment such as conductivity, acidity (pH), temperature, color, UV absorption, i.e. doses of solutions for coagulation-flocculation and disinfection.

Measurements performed in real time by all embedded processes meters-transmitters, and from which operational data is collected (recorded), can be applied in order to obtain conceptual models such as dosing systems.

The implementation of other process measurements in real time, such as conductivity measurement in order to manage coagulant dosing, as well as research on the replacement of coagulants using "electrocoagulation" which is a coagulation process without the use of chemicals for coagulation.

6 REFERENCES

[1] Pingale, R., & Shinde, S. N. (2021). Fuzzy logic approach for routing in Internet of Things network. *Tehnički glasnik*, *15*(1), 18–24. https://doi.org/10.31803/tg-20210204155601

[2] Svalina, I., Galzina, V., Lujic, R., & Simunovic, G. (2013). An adaptive network-based inference system (ANFIS) for the forecasting: The case of close price indices. *Expert Systems with Applications*, *40*(15), 6055–6063. https://doi.org/10.1016/j.eswa.2013.05.029

[3] Valcic, M., Antonic, R., & Tomas, V. (2011). ANFIS based model for ship speed prediction. *Brodogradnja*, *62*(4), 373–382. https://doi.org/10.21278/brod

[4] Dimanoski, K., Vrtanoski, G., & Stojic, G. (2017). Fuzzy logic model for dimensioning of transport capacities in railway freight transport. *Mechanical Engineering Scientific Journal*, *35*(2), 105–114.

[5] Yetkin, M. E., Simsir, F., & Ozfirat, M. K. (2016). A fuzzy approach to selecting roof supports in longwall mining. *South African Journal of Industrial Engineering*, *27*(1), 162–177. https://doi.org/10.7166/27-1-1366

[6] Kahraman, C., Cebeci, U., & Ulukan, Z. (2003). Multi-criteria supplier selection using fuzzy AHP. *Logistics Information Management*, *16*(6), 382–394. https://doi.org/10.1108/09576050310503367

[7] Kahraman, C., Cebeci, U., & Ruan, D. (2004). Multi-attribute comparison of catering service companies using fuzzy AHP: The case of Turkey. *International Journal of Production Economics*, *87*(2), 171–184. https://doi.org/10.1016/S0925-5273(03)00099-9

[8] Kahraman, C., Cevik, S., & Oztaysi, B. (2015). Fuzzy multicriteria decision-making: A literature review. *International Journal of Computational Intelligence Systems*, *8*(4), 637–666. https://doi.org/10.1080/18756891.2015.1046325

[9] Ozifrat, P. M. (2020). A fuzzy event tree methodology modified to select and evaluate suppliers. *South African Journal of Industrial Engineering*, *31*(1). https://doi.org/10.7166/31-1-2027

[10] Wan, J.-Q., Huang, M.-Z., Ma, Y.-W., Guo, W. J., Wang, Y., & Zhang, H.-P. (2010). Control of coagulation process in a

paper–mill wastewater treatment process using a fuzzy neural network. *Chemical and Biochemical Engineering Quarterly*, *24*, 425–435. https://hrcak.srce.hr/62712

[11] Heddam, S., Bermad, A., & Dechemi, N. (2012). ANFIS-based modelling for coagulant dosage in drinking water treatment plant: A case study. *Environmental Monitoring and Assessment*, *184*, 1953–1971. https://doi.org/10.1007/s10661-011-2091-x

[12] Trinh, T. K., & Skang, L. (2010). Response surface methodological approach to optimize the coagulation-flocculation process in drinking water treatment. *Chemical Engineering Research and Design*, *89*, 1126–1135. https://doi.org/10.1016/j.cherd.2010.12.004

[13] Juntunen, P., Liukkonen, M., Lehtola, M., & Hiltunen, Y. (2012). Dynamics of sensors for detecting factors affecting turbidity in drinking water. *Journal of Hydroinformatics*, *15*(2), 416–426. https://doi.org/10.2166/hydro.2012.052

[14] Boruszenko, P., Haley, J., & Thompson, K. C. (2015). On-line analysis: A water company perspective. https://www.rsc.org/images/JohnHaleyYWS5thNovCambridge_tcm18-143639.pdf

[15] Naidoo, P., & Van der Walt, J. J. (2013). Artificial neural networks as chemical dosing budgeting tool. https://hdl.handle.net/10520/EJC143029

[16] Cooperative Research Centre (CRC) for Water Quality and Treatment. (2015). *Natural organic matter – understanding and controlling the impact on water quality and water-treatment processes*. http://www.waterra.com.au/publications/document-search/

[17] Kasabov, N. (1996). *Foundation of neural networks, fuzzy systems and knowledge engineering*. The MIT Press. https://doi.org/10.7551/mitpress/3071.001.0001

[18] Yen, J., & Langari, R. (2013). *Fuzzy logic intelligence, control and information*. Pearson.

[19] Jang, J.-S. R. (1993). ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*, *23*(3), 665–685. https://doi.org/10.1109/21.256541

[20] Gorzalczany, M. B. (2002). *Computational intelligent systems and applications: Neuro fuzzy and fuzzy neural synergisms*. Physica-Verlag.

[21] Wu, G. D., & Lo, S. L. (2008). Predicting real-time coagulant dosage in water treatment by artificial neural networks and adaptive network based fuzzy inference system. *Engineering Applications of Artificial Intelligence*, *21*, 1189–1195. https://doi.org/10.1016/j.engappai.2008.03.015

Authors' contacts:

Mirche Stojkov, MSc
(Corresponding author)
Public Utility Service Company "Ograzhden",
s. Bosilovo No. 257, 2431 Bosilovo,
Municipality of Bosilovo, Republic of North Macedonia
mircestojkov@hotmail.com

Zoran Pandilov, PhD, Full Professor
Institute of Production Engineering and Management,
Faculty of Mechanical Engineering,
Ss. Cyril and Methodius University in Skopje,
P. O. Box 464, 1001 Skopje, Republic of North Macedonia
zoran.pandilov@mf.edu.mk