p-ISSN 2083-0157, e-ISSN 2391-6761 IAPGOŚ 2/2026 113

http://doi.org/10.35784/iapgos.8508

received: 03.10.2025 | revised: 11.02.2026 | accepted: 12.02.2026 | available online: 30.06.2026

IRON COAGULATION OPTIMIZATION DURING WATER TREATMENT USING ARTIFICIAL INTELLIGENCE TOOLS

Andriy Safonyk, Ivan Tarhonii, Oleksandr Naumchuk, Vladyslav Danchenkov, Roman Zaichuk
National University of Water and Environmental Engineering, Institute of Energy, Automation and Water Engineering, Rivne, Ukraine

Abstract. *A process of water purification using electrical coagulation and its optimisation using artificial intelligence tools was presented. Experimental data was analysed and correlated. The experimental studies used were developed to optimise iron coagulation during water treatment. A neural network was developed to optimise iron coagulation and machine learning was performed. Neural network tests were conducted and methods for its practical application were proposed.*

Keywords: neural networks, computer vision, water treatment, modelling, process optimization, RGB, HSL

OPTYMALIZACJA KOAGULACJI ŻELAZA PODCZAS UZDATNIANIA WODY Z WYKORZYSTANIEM NARZĘDZI SZTUCZNEJ INTELIGENCJI

Streszczenie. *Przedstawiono proces oczyszczania wody z wykorzystaniem koagulacji elektrycznej oraz jego optymalizację przy użyciu narzędzi sztucznej inteligencji. Przeanalizowano i skorelowano dane eksperymentalne. Przeprowadzono badania eksperymentalne mające na celu optymalizację koagulacji żelaza podczas uzdatniania wody. Opracowano sieć neuronową w celu optymalizacji koagulacji żelaza oraz przeprowadzono proces uczenia maszynowego. Przeprowadzono testy sieci neuronowej i zaproponowano metody jej praktycznego zastosowania.*

Słowa kluczowe: sieci neuronowe, wizja komputerowa, uzdatnianie wody, modelowanie, optymalizacja procesów, RGB, HSL

Introduction

Humanity's technological development tends to add its input to bank of ecological issues. As industrial establishments are constantly growing in power, ecology-treatment levers must synchronously grow in productivity. But highest productivity often causes the process to be complicated and come with a varying stack of problems that require new solutions to be explored. During electrocoagulation, metallic salts are adhered to the substance's environment and, connecting with hydroxyl group coagulant, is being withdrawn from a fluid by surface-directed air currents. As purification level needed is not stationary and varies in real time, as contaminated fluid enters the system. Problem here is that the level of metallic salts must be precisely measured to be added into water, but those properties of system's real state cannot be measured directly. Meanwhile, indirect measuring contributes to final error, that involves too much calculations. The use of artificial intelligence technologies can simplify the process of iron coagulation in water treatment and improve the efficiency and accuracy of adding the appropriate reagents.

In this particular study, artificial intelligence methods were considered to be involved in enhancing iron level estimation algorithm for electrocoagulation water treatment process.

1. Analysing research sources on the subject of the study

Electrocoagulation is a water treatment technology that is highly effective in removing a wide range of contaminants. It does not rely on aggressiveness of the environment (what biological purification does), initial properties (like water hardness, that cannot be removed during boiling), or range of the pollutants (where chlorination cannot be applied), and it is fast enough (that cannot be proposed by ordinary sedimentation) [4, 5].

In an article "Natural organic matter removal by coagulation during drinking water treatment: A review" by Any Matilainen, Mikko Vepsäläinen and Mika Sillanpää, an overview of recent research and findings on natural organic matter abolishment methods was presented [4].

In study "A comprehensive review of electrocoagulation for water treatment: Potentials and challenges" by Dina T. Moussa and Muftah H. El-Nass researchers analysed a topic of stability for colloidal particles, history and application of electrocoagulation for water treatment, development vector of this research current was described [5].

In an article "Coagulation in Drinking Water Treatment: Particles, Organics and Coagulants", author J. K. Edzwald expands a review and adds importance of raw water chemistry in connection with natural organic matter and chemistry of coagulants. Edzwald signs that, in process of water treatment, presence of electrostatic charge, hydrophilic effects and sterilic interactions must be considered and accounted in the system, because they enforce mineral particles bonds with containing fluid [1].

Research titled "The role of coagulation in water treatment" by Jia-Qian Jiang also considers history side of the topic. He described the way approaches to measurements and the treatment process itself varied throughout the history. The author also mentions that coagulant's dose control is crucial for the process [2].

Article: "Coagulation – ultrafiltration system for river water treatment" by Krystyna Konieczny, Dorota Sąkol, Joanna Płonka, Mariola Rajca and Michał Bodzek assumes that significant part of treatment's effectiveness is up to application method. In-line coagulant's application has shown an increase in system's productivity, as it performs significantly better when operating water polluted with natural organic matter. This paper also investigates sorts of reagents for coagulation. In the work, a range of reagents $FeCl_3$, $Fe_2(SO_4)_3$, and $Al_2(SO_4)_3$ have been considered for the best productivity, signing that capillary polyetheruslphone was applied in system during the experiment [3].

2. Results of experimental studies

An experimental setup was developed to explore various applications of electrocoagulation for treating different pollutants in water, enabling thorough research. A schematic of this setup is presented in Fig. 1. It consists of power source, a set of pumps, and a measurement camera with a registration and control device.

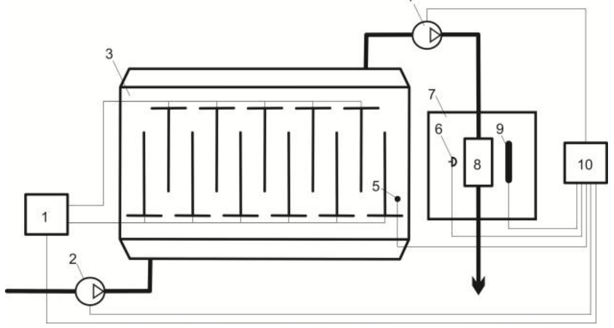

Fig. 1. Experimental setup scheme

Principle of operation and description of the experimental setup. Power supply (1) is providing the whole setup with energy. Pump (2) is running contaminated water into the system. Flow-through-kind electrical coagulator (3) is made of non-conductive material with metal plates inside – no neighbouring metal plate has shared polarity – they represent cathodes and anodes. After passing treatment procedure, water is directed by pump (4) to estimation chamber (7). Temperature control in coagulation reservoir is also being done by respective sensor (5). Estimation chamber (7) includes light source (6), transparent measurement tank (8) and a sensitive element (9). All electrical components of the setup are being controlled by registration and control device (10).

The results of the experimental studies are shown in Table 1. From the studies conducted, a connection between the concentration of components, current strength, and colour of the substance at different times and at different voltages is observed. Analysing these dependencies, it was found that there are stable correlations between these parameters. Thus, it is possible to develop a neural network model to optimize iron coagulation in the water treatment process [6].

Table 1. Experimental data on color change relying on concentration value

No	Time (m)	Concentration (mg/dm^3)	Red	Green	Blue
1	6	0.8	204	207	200
2	12	1.1	214	215	193
3	18	3.7	211	194	75
4	24	4.5	220	173	54
5	30	6.3	210	155	49
6	36	6.9	215	147	41
7	42	9.1	193	116	38
8	48	9.6	189	111	33
9	54	11.8	188	100	27
10	60	12.6	178	74	11

The modelling of the iron coagulation process was performed in MATLAB R2023a. To create a neural network model from a set of data presented in the form of numerical arrays the specialized toolkit of the MATLAB program – Neural Network Toolbox (NNTool) was used. NNTool was also used to gather graphical information on created neural network's R2 score and Mean Square Error values; difference between expected and trained values; damping factor; overall function fit. Data given was expanded from experimental data, using mathematical model generated in Curve Fitter. Computation process was held on a device containing AMD Ryzen 7 5800H 3.2 GHz processor, NVIDIA GeForce RTX 3070 Laptop 8 GB GPU, 16 GB RAM.

3. Features of the research methodology

To identify the mathematical relationships underlying the process, an approximation is applied that enables the representation of data by substituting correlation points with a specific mathematical equation. This kind of substitution is often performed to solve an issue of automatic control. But here, mathematical approximation is being performed to expand dataset. Approximation was involved, as maximum deviation from real value estimated passed requirements set. The results of the approximation of the experimental data are shown in Figs. 2, 3.

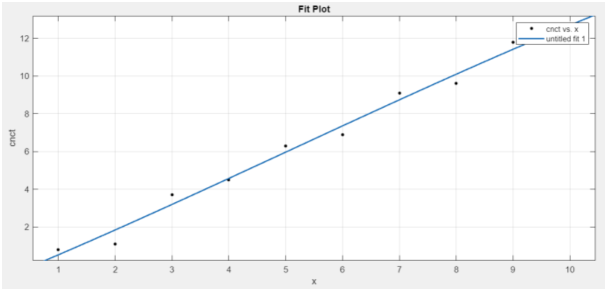

Fig. 2. Concentration data on unit axis

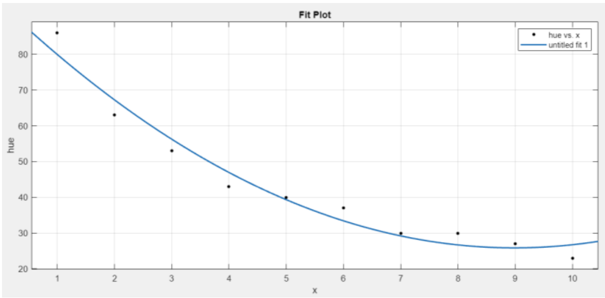

Fig. 3. Hue data on unit axis

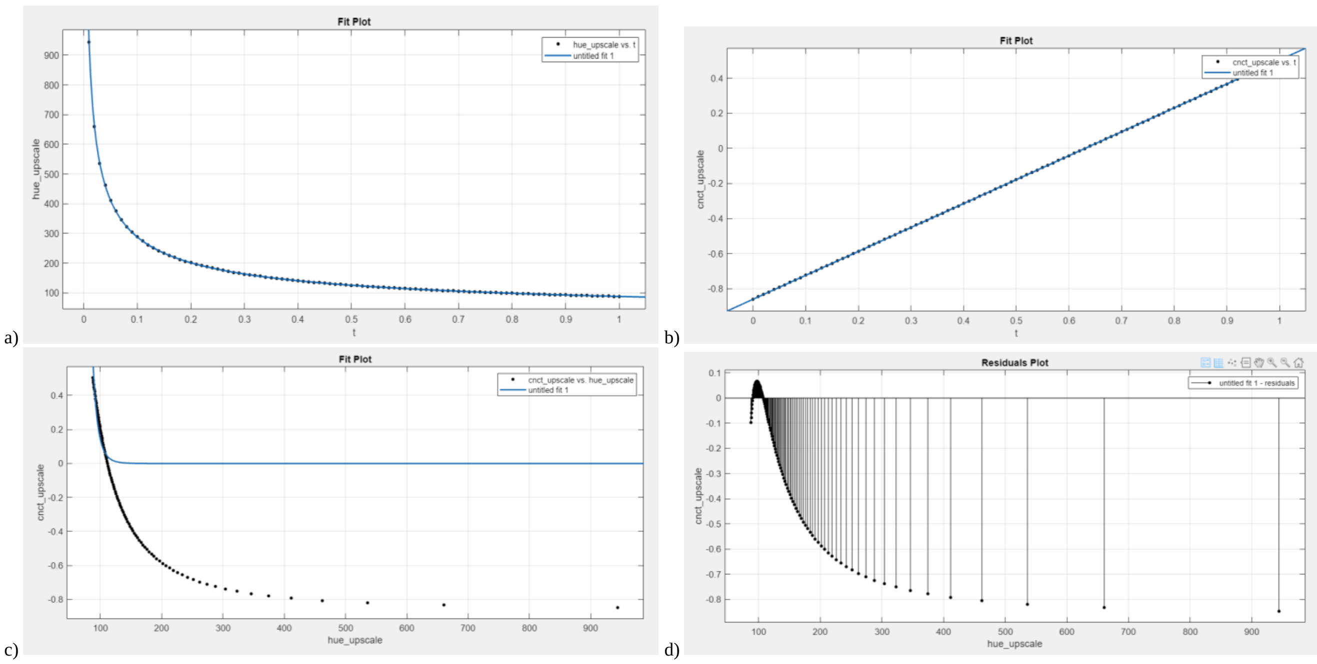

Fig. 4. Hue data on unit axis (a); expanded concentration data on unit axis (b); correlation approximation between hue and concentration data (c, d)

Our experimental results include 10 points. As a result of applying the machine learning algorithms shown in Fig. 4, their high performance was established, which is clearly seen in the graphs presented. It is also noticed that the more input data for modelling, the more effective these algorithms are. Therefore, to expand the dataset during training and testing of the neural network, we approximated the obtained dependencies and obtained expressions that increased the number of points. To do this, the mathematical expressions obtained after the approximation were filled with time values in the ranges from 0 to 3600 seconds in increments of 1 second. At the same time, the resulting approximation curves are fitted to the parameters of our neural network using the MatLab Curve Fitter option, as shown in Fig. 4.

In the course of the experimental studies, 10 samples were taken (Table 1), on the basis of which the content of total and three-valent iron was determined. This made it possible to determine the change in the RGB spectrum in each of the 10 experiments. At the same time, the obtained results of colour consistency and concentration were converted to the HSL spectrum, on the basis of which a neural network was developed to determine the concentration of the nutrient content in water.

The HSL color model was used because, unlike the RGB model, which is based on red, green, and blue color parameters with 256 variations, HSL uses three parameters (hue, saturation, and brightness) to establish a specific output color. Thus, the hue of a particular color is represented not by a numerical value, but by an angular degree on the color palette. The other two parameters, lightness and saturation, are decimal numbers that range from 0 to 1. No additional experiments are required, since the conversion algorithm between the RGB and HSL color models is known.

The neural network structure we used has one hidden layer that contains all the sequences of neurons that are between the input and output of the model, forming a neural network. The number of hidden layers depends on the number of parameter sets that the network takes into account. The connections of the neural network are configured using the neuron weight parameter, which strengthens or weakens the neural connection.

Also, for the effective functioning of a neural network, it is necessary to conduct machine learning, which is a preparatory stage for its operation. This stage is very important for the proper functioning of the neural network, as the neural network is trained for specific purposes and does not have general computing capabilities. At this stage, a data model is formed from the input dataset, which carries information about the properties of the relationships between two not closely related model parameters.

4. Machine learning of a neural network

As a result of model establishment, a set of properties is being mentioned. Epoch is a parameter, that relates to steps in machine learning process. Elapsed time for current dataset volume was less than one second, but validation checks passing as a stopping condition (instead of stopping by the condition of epoch passing) is a sign of successful training. This process is shown in Fig. 5.

Supervised machine learning algorithm, that is used during training, require both input and output values to be matched. Training dataset was represented by two correlated data arrays of the same length, where hue input and concentration output values are connected by pairs (Fig. 6).

To train our neural network the Levenberg-Marquardt algorithm was used. It is a standard algorithm for solving non-linear curve-fitting problems. It takes advantage of gradient-decent and Gauss-Newton optimization methods.

Training results were estimated from that quantity of epochs, in what successful result has been reached. As can be seen from Fig. 7, out of 1000 possible epochs for the model, only 71 of them have passed for training, 15 for validation and 15 for testing, so learning process did not come to an end by final condition. R^2 (coefficient of determination) and mean square error of the model also show signs of a successful training.

Training a neural network is a gradual process, so it is important to determine the effectiveness of training and stop it at the appropriate time. As neural network's gradient represents partial derivatives of its loss function, it can be seen, that loss is slightly falling with every epoch (Fig. 8).

Damping factor here is technically reassuring, that neural network ended training by fulfilling the training condition, so the learning process stopped (Fig. 9).

Training Progress

Unit	Initial Value	Stopped Value	Target Value
Epoch	0	14	1000
Elapsed Time	-	00:00:00	-
Performance	0.22	3.43e-06	0
Gradient	1.61	0.00233	1e-07
Mu	0.001	1e-07	1e+10
Validation Checks	0	6	6

Fig. 5. Training progress details

Data

Predictors: hue_upscale - [101x1 double]

Responses: cnct_upscale - [101x1 double]

hue_upscale: double array of 101 observations with 1 features.

cnct_upscale: double array of 101 observations with 1 features.

Fig. 6. Input data description

Training Results

Training start time: 28-Nov-2023 12:24:34

Layer size: 10

	Observations	MSE	R
Training	71	0.0000	0.9999
Validation	15	0.0033	0.9908
Test	15	0.0003	0.9995

Fig. 7. Training results in form of observation quantity, mean square error, and R^2 score

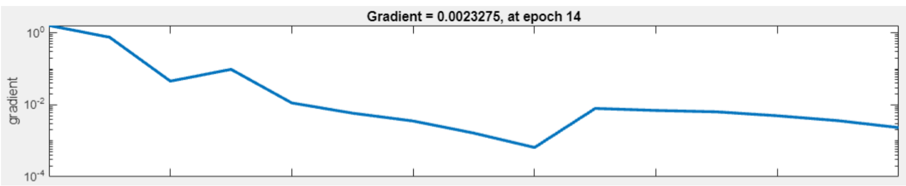

Fig. 8. Neural network's gradient graph

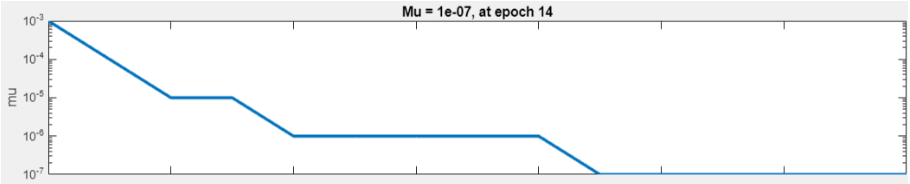

Fig. 9. Neural network's damping factor

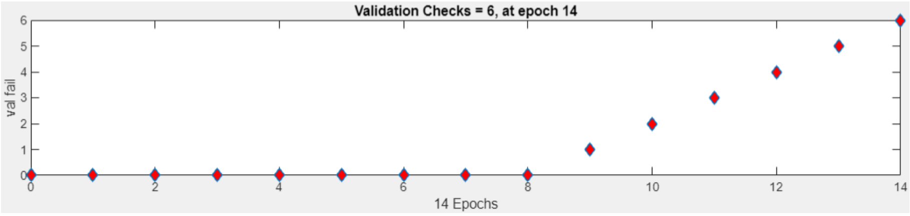

Fig. 10. Validation split by epochs

Validation from eights epoch show linear growth in validation fails, so it was limited by eights epoch as a final (Fig. 10).

As can be seen from Fig. 11, overall scoring by training, validation and testing show, that testing showed its best result on eights epoch out of fourteen.

For a neural network to function efficiently, its errors should be determined. With this in mind, most of the training tests of our neural network were conducted around the value of 0.001809, which is a sign of zero error around the corresponding axis of the value domain. Also, the graphs in Fig. 12 shows how the data alternates between the initial and target values.

Analysing the function fitting data and the error distribution shown in Fig. 13, it can be seen that the largest number of errors is concentrated between the four-hundredth and five-hundredth epochs. It can also be seen, that the most successful epochs from this graph are ending at around three hundred forties, so it was decided to set estimated limit.

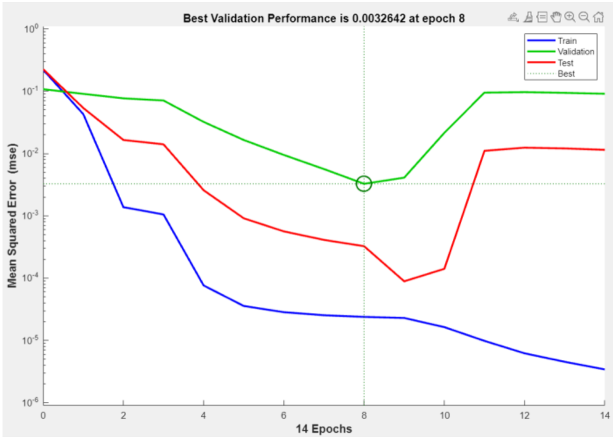

Fig. 11. Validation split by epochs

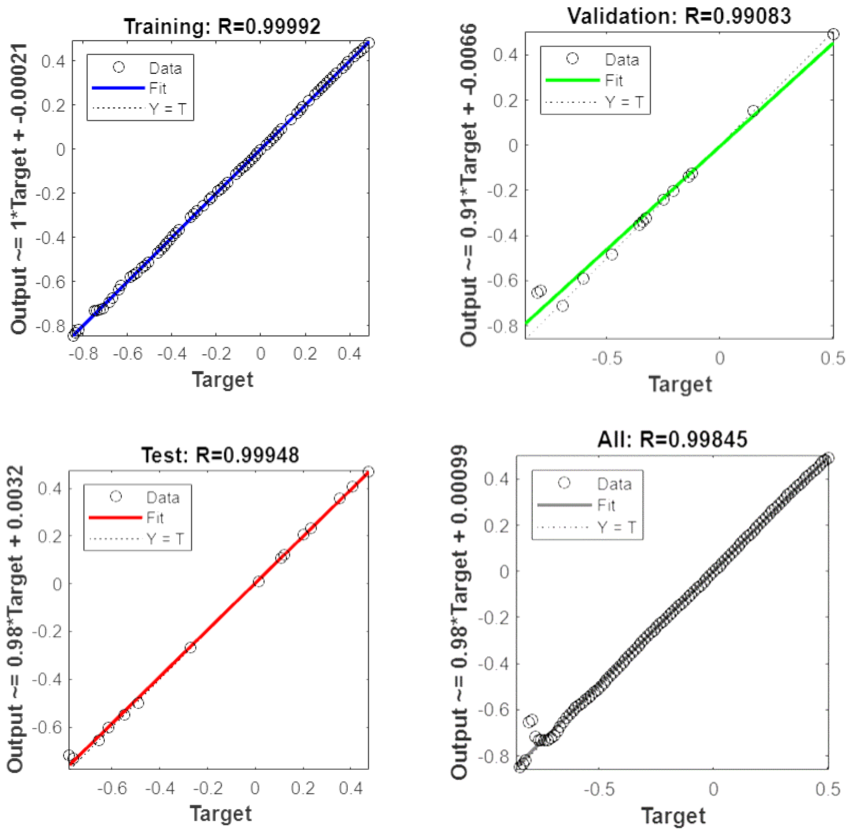

Fig. 12. Training success deviation. Validation success deviation. Testing success deviation. Overall success deviation

An important step in preparing to use a neural network is its testing. In the process of testing the trained neural network, it was found that the best trained value of our network is 0.000379. This is clearly seen in the final testing process, which is shown in Fig. 14.

The fit curve (Fig. 15) during testing indicates that a significant part of the results during testing is concentrated up to 300 epochs, and the largest error is localized at about 660 epochs. At the same time, the error scatter takes into account the error values more accurately, so that the difference can be assessed visually by omitting the training fit scale.

Thus, after our experimental studies and their analysis, the neural network developed by us has been shown to be highly efficient and suitable for further practical use.

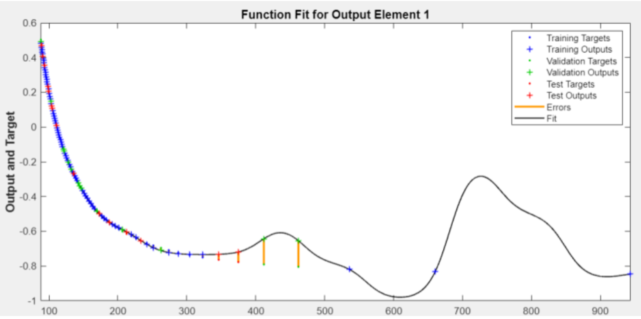

Fig. 13. Key points' spread by epochs

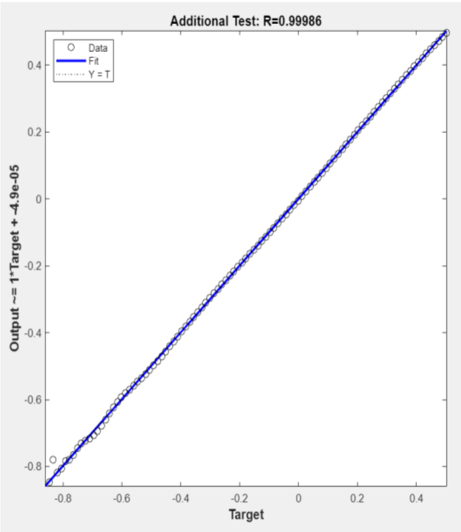

Fig. 14. Additional testing graph

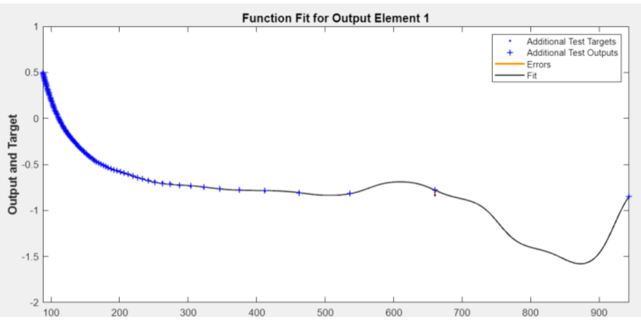

Fig. 15. Target-output deviation on fitting curve

5. Conclusion

In this article an approach to water treatment using electrical coagulation was proposed. Reasons for the research have been given. Also, theoretical basis was described. Neural network was built using MATLAB's nnfit to mimic the way artificial network will behave with experimental data given. During this research, key parameters of the researched model and their correlations were estimated. As a result, optimal parameters for the experimental data were established.

References

[1] Edzwald, J. K. (1993). Coagulation in Drinking Water Treatment: Particles, Organics and Coagulants. *Water Science and Technology*, *27*(11), 21–35. https://doi.org/10.2166/wst.1993.0261

[2] Jiang, J.-Q. (2015). The role of coagulation in water treatment. *Current Opinion in Chemical Engineering*, *8*, 36–44. https://doi.org/10.1016/j.coche.2015.01.008

[3] Konieczny, K., Sąkol, D., Płonka, J., Rajca, M., & Bodzek, M. (2009). Coagulation—Ultrafiltration system for river water treatment. *Desalination*, *240*(1–3), 151–159. https://doi.org/10.1016/j.desal.2007.11.072

[4] Matilainen, A., Vepsäläinen, M., & Sillanpää, M. (2010). Natural organic matter removal by coagulation during drinking water treatment: A review. *Advances in Colloid and Interface Science*, *159*(2), 189–197. https://doi.org/10.1016/j.cis.2010.06.007

[5] Moussa, D. T., El-Naas, M. H., Nasser, M., & Al-Marri, M. J. (2017). A comprehensive review of electrocoagulation for water treatment: Potentials and challenges. *Journal of Environmental Management*, *186*, 24–41. https://doi.org/10.1016/j.jenvman.2016.10.032

[6] Safonyk, A., Hrytsiuk, I., Mishchanchuk, M., & Ilkiv, I. (2021). Information system of electrochemical obtaining of coagulant on the basis of photocolorimetric analysis. *Measuring and Computing Devices in Technological Processes*, (1), 97–104. https://doi.org/10.31891/2219-9365-2021-67-1-14

Prof. Andrii Safonyk
e-mail: a.p.safonyk@nuwm.edu.ua

Director of the Institute of Energy, Automation and Water Engineering, National University of Water and Environmental Engineering, Rivne, Ukraine.
Engaged in scientific mathematical modelling of natural and technological processes, computer techniques and computer technologies, programming.

https://orcid.org/0000-0002-5020-9051

Ph.D. Ivan Tarhonii
e-mail: i.m.tarhonii@nuwm.edu.ua

Associate professor of the Department of Automation, Electrical Engineering and Computer Integrated Technologies of the Institute of Energy, Automation and Water Engineering, National University of Water and Environmental Engineering, Rivne, Ukraine.
Engaged in scientific mathematical modelling of natural and technological processes, computer techniques and computer technologies, programming.

https://orcid.org/0000-0002-0811-952X

Ph.D. Oleksandr Naumchuk
e-mail: o.m.naumchuk@nuwm.edu.ua

Associate professor of the Department of Automation, Electrical Engineering and Computer Integrated Technologies of the Institute of Energy, Automation and Water Engineering, National University of Water and Environmental Engineering, Rivne, Ukraine.
Engaged in scientific research of effects of electromagnetic radiation from the mobile communication equipment on the environment, and automation, mathematical modelling of technological processes.

https://orcid.org/0000-0003-2483-4141

Vladyslav Danchenkov
e-mail: danchenkov_ak20@nuwm.edu.ua

Master's degree student of the Department of Automation, Electrical Engineering and Computer Integrated Technologies of the Institute of Energy, Automation and Water Engineering, National University of Water and Environmental Engineering, Rivne, Ukraine.
Engaged in scientific mathematical modelling, of natural and technological processes, programming.

https://orcid.org/0009-0000-5373-2166

M.Sc. Roman Zaichuk
e-mail: r.m.zaichuk@nuwm.edu.ua

Ph.D. student of the Department of Hydro and Heat Power Engineering and Hydraulic Machines of the Institute of Energy, Automation and Water Engineering, National University of Water and Environmental Engineering, Rivne, Ukraine.
Engaged in scientific mathematical and physical modelling, natural and technological processes.

https://orcid.org/0009-0004-8419-0178