❐   560

Neural network-based pH and coagulation adjustment system in water treatment

Oscar Iván Vargas Mora[1], Daiam Camilo Parrado Nieto[1], Jairo David Cuero Ortega[1],
Javier Eduardo Martínez Baquero[1], Robinson Jiménez-Moreno[2]

[1]Faculty of Basic Sciences and Engineering, Universidad de los Llanos, Villavicencio, Colombia
[2]Department of Mechatronic Engineering, Universidad Militar Nueva Granada, Bogota, Colombia

Article Info

ABSTRACT

This document presents a machine learning model development as a tool to improve chemical dosing procedure in ariari regional aqueduct (ARA). The supervised learning model has been addressed starting from the knowledge of data color, turbidity and pH at the water inlet to the aqueduct and the dosing results of type A aluminum sulfate and calcium oxide (lime) obtained through jar tests. The construction of the automatic learning model had a comprehensive implementation and improvement field through continuous system training, which allowed an optimal dosage of Aluminum Sulfate and Lime to generate an outlet pH less than 7.5 and outlet turbidity less than 8 nephelometric turbidity unit (NTU). Those outlet water parameters meet the ministry of social protection criteria in Colombia. Also, a virtual jar test was created to reduce the time required to obtain chemical dosing values to less than a minute. In contrast, a laboratory test takes approximately a half-hour to displays results.

Corresponding Author:

Javier Eduardo Martinez Baquero
Faculty of Basic Sciences and Engineering, Engineering School, Universidad de los Llanos
Transversal 25 #13-34, Villavicencio, Colombia
Email: jmartinez@unillanos.edu.co

1. INTRODUCTION

Drinking water treatment systems in aqueducts involve a chemical dosing process to control the key parameters that guarantee the quality of treated water; these processes present a correlation between their different variables, making it incredibly complex to set classical control strategies. Proportional, integral and derivative (PID) controllers and their variants, which are part of classical control, present many advantages at an industrial level, as explained Yu *et al.* [1] Abdullah and Ali [2]. However, they apply well to linear systems, whose models are feasible to establish. For systems that do not meet this condition, there are control strategies based on artificial intelligence [3], [4], called intelligent control techniques [5].

Intelligent control also allows to replicate PID strategies as [6] and [7]; however, it is more complex to implement. Therefore, it is used with non-linear plants in prediction systems presented Mohamed *et al.* [8] or satellite dynamic altitude control systems [9]. Intelligent control applications cover vast and diverse fields from academic settings to industrial processes such as those in [10]–[16]. Neural controllers have even been developed for security management in residential homes [17]. This has allowed systems as complex and delicate as water treatment systems to also choose intelligent control techniques [18]–[20]. In [21], [22], a recent review of state of the art is exposed where they highlight the use of neural networks as an essential technique in water treatment, where one of the critical parameters of the process is the control of pH as stated in [23]–[25].

At the ariari regional aqueduct (ARA) in Meta Colombia, it is desired to adjust the chemical dosing process in the treatment plant. pH is a critical factor in meeting national regulations because the high correlation between the intervening variables affects it. It is observed that the advantages of intelligent control based on neural networks can contribute to automating this process, which due to its complexity, has been carried out manually with the technique known as jar tests [26], which lasts between 20-30 min., making adjustment difficult.

This article is composed of four sections that expose the development of the work done. The first section is the present introduction. In section 2, the methodology where dataset and network architecture are shown. Section 3 presents the analysis of results. Finally, section 4 exposes the conclusions drawn.

2.   METHOD

At the dosage required for drinking water treatment, surface water catchments and raw water physicochemical parameters can change suddenly. Which affects water quality if proper adjustments are not made to the process chemistry. Figure 1 shows the main stages that are an integral part of the development in the intelligent control model, which will be discussed in next sub-sections.
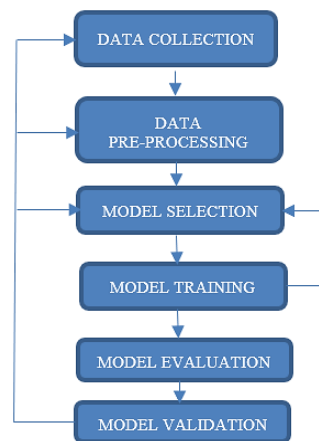

Figure 1. Intelligent control flow chart

2.1. Data collection and pre-processing

In the drinking water treatment process is essential to make adjustments to pH stabilizer and coagulant immediately when variations occur to guarantee the water quality. For data collection, within the analysis of chemical dosing, the following essential input variables were identified: color, turbidity, pH, amount of qualified dosed, and amount of Aluminum sulfate type A dosed, and as output variables, color, turbidity, and pH. Empresa de servicios públicos del meta (EDESA S.A. E.S.P.) in ARA provided a set of 720 jar analysis reports that comprised 488 data vectors.

After the data collection stage, the data processing is started by performing the format assignment in the Excel database. Here, the number format was selected to facilitate loading the data to the modeling process later. The data preprocessing phase was about identification. As a first instance of the amount of lost or missing data, these incomplete or empty data sets are eliminated from the general dataset since these "empty" data affect the behavior of the data machine learning models.

The data collected for the variables of turbidity, color, pH, lime, and Aluminum Sulfate type A, contain values within a wide range and not uniform with each other, which is inherent to the process. Still, the learning model's training process automatically generates inconveniences of data dispersion that increase convergence times of the algorithm and sometimes give poor results. Therefore, scaling converts the data to a uniform range, normalizing original data to a range between 0 and 1; this is performed by maximums and minimums (1) without affecting the dataset and maintaining the proportionality of each data. As an example, Figure 2 shows the graph of the preprocessing for the variable turbidity (input), on the left initial purified data and on the right normalized.

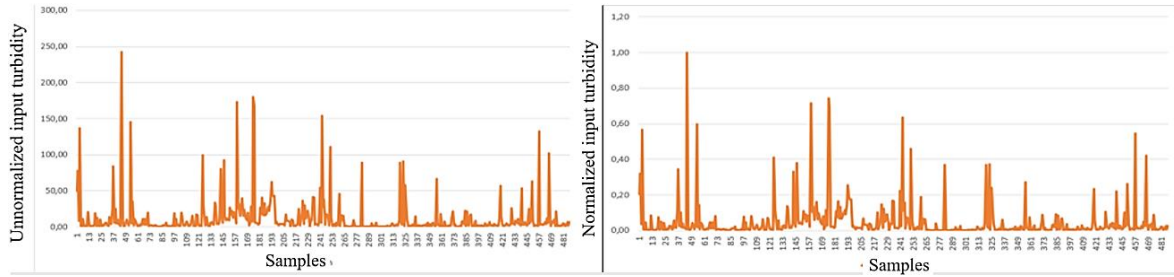$$Xnormalized = \frac{X - Xmin}{Xmax - Xmin} \tag{1}$$

Figure 2. Unnormalized and normalized input turbidity data

For training of neural networks, percentage distribution of the total data sample (488) was defined as in Table 1. The database is distributed in 3 sets for training and validation. Each sample corresponds to the tabulated information of the manual jar testing process.

Table 1. Distribution of samples in the training, testing and validation stages

Stage	Data percentage	Number of samples
Training	70%	342
Testing	15%	73
Validation	15%	73

2.1. Neuronal arquitecture

Chemical dosing is a stage with information on the values of the input variables and output variables of interest, so this problem must be adjusted to a supervised type of machine learning model, such as artificial neural networks. It is essential to clarify that the chemical dosing model is carried out in two stages. First, both pH turbidity and color are measured to adjust and determine lime and aluminum sulfate's initial level. The second stage network takes as input data the turbidity, color, pH, lime, and Aluminum Sulfate. As output variables, it takes the turbidity, color, and pH, so the neural architecture of the model to be implemented will follow the same dosage form. That is, two neural architectures, one model for each stage shown in Figure 3.
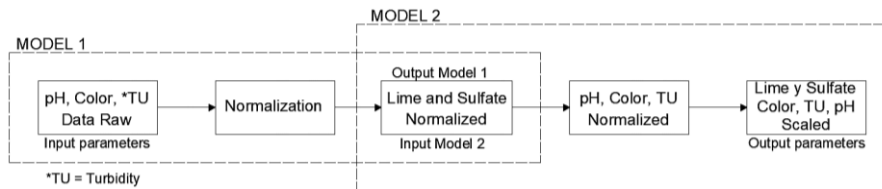

Figure 3. Neural network model

For the training phase of the neural network model, the input and output relationship of each model is straightforward. It allows to directly establish the number of inputs and outputs of the network through an iterative process the number of neurons in the layer is determined hidden. Figure 4 shows a general representation of the proposed models with internal architectures of the neural networks of models 1 (left) and 2 (right), respectively. The architecture of the neural network model 1 had 20 (twenty) neurons in the hidden layer, which is twice that of model 2, due to the amount and complexity of data used for the training process. The output of model 1 becomes inputs for model 2 according to the manual dosing process used.
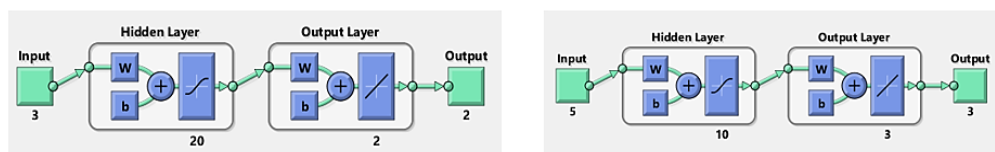

Figure 4. Neural network architecture of model 1 y 2

The number of neurons in the hidden layer of each model was selected by iterating five by five until the best possible approximation was obtained. Although the greater the number of neurons in the hidden layer, the greater the time required for training, validation, and testing of the model, a high computational load was not presented. A fixed number of epochs was used for each training trial, set at 1000 epochs in both neural network models.

Figures 5 and 6 show the training, validation and testing of each neural network model. The value of R indicates the relationship obtained between output data and the target value, a value of R= 1 indicates a perfect fit, for the case under study, in model 1 the value of R was 0.81325 is a higher value. In model 2, the value of R was 0.96158, which also shows an excellent relationship between the values output and objectives.
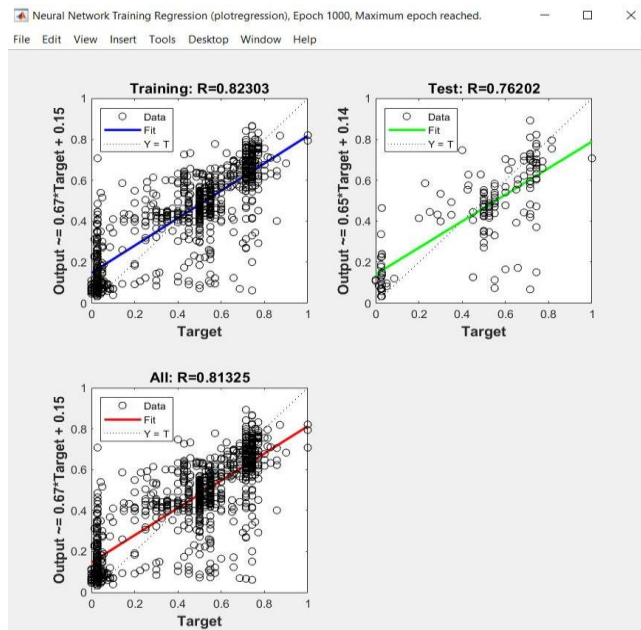

Figure 5. Neural network training for model 1, view of the regression graphs
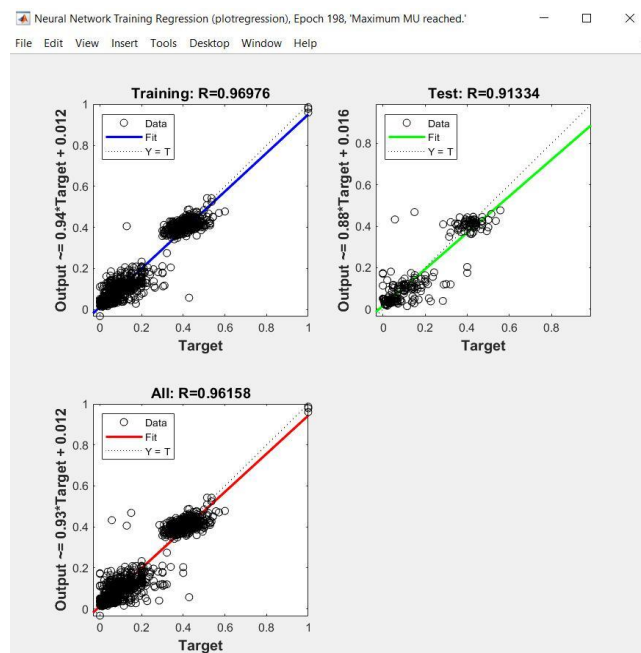

Figure 6. Neural network training for model 1, view of the regression graphs

For the phase of evaluation (model 1 and model 2), a dataset equivalent to 15% of the total available data was reserved. In Figure 7, the error is plotted against each training cycle of model 1. It indicates that the best performance was obtained at time 394 with an error of 0.022209, and the smaller the mean square error, the more approximate they are predicted and observed values.
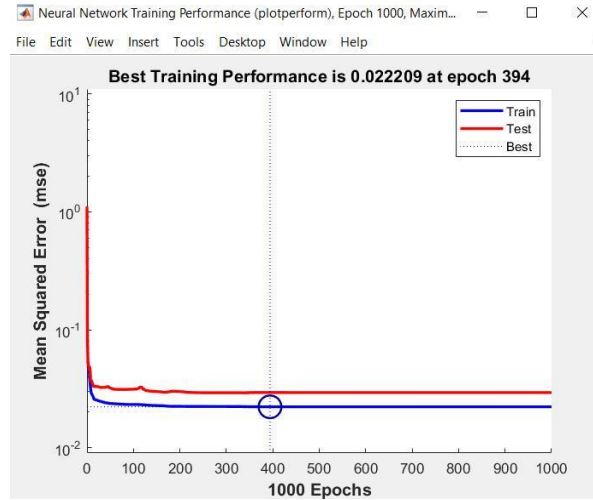

Figure 7. Neural network training performance results for model 1

Figure 8 shows the error against each training cycle of model 2 and indicates that the best performance was obtained in epoch 197 with an error of 0.0017541. The smaller the mean square error (MSE), the more approximate are the predicted and observed values. The low level of error obtained allows to validate this architecture for process automation tests.
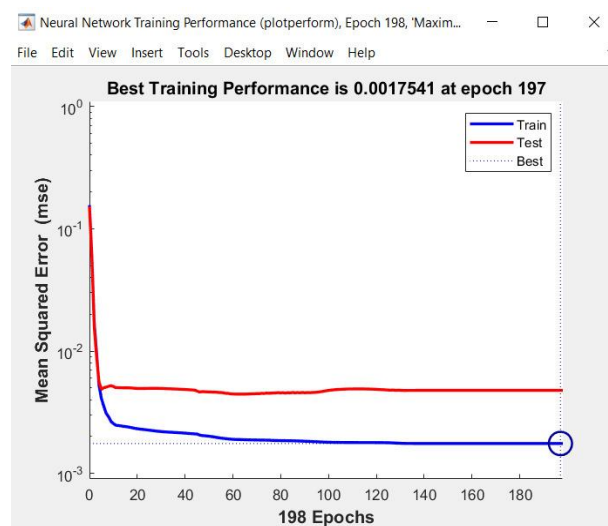

Figure 8. Neural network training performance results for model 2

3. RESULTS AND DISCUSSION

The algorithm delivers the results Figure 9 from the jar test in the machine learning model developed to calculate the dosage of lime and sulfate to treat drinking water at ARA. In other words, it is possible to determine the discharge of lime and sulfate required to adjust the pH level of the treated water. The neural model seeks to replicate the appropriate dosage levels according to the standard, according to the parameterization of the network.

Lime = 19.4984726 ppm
Sulfate = 47.9535313 ppm
Color = 10.8484245 PCU
Turbidity = 1.4653494 NTU
pH = 5.8932976
Flow = 180.00 l/s
Lime discharge = 210.58 g/min
Sulfate discharge = 517.90 g/min

Figure 9. Output of machine learning model

The Data from five of the jar tests performed by the plant operators were used to compare the response of the registered nurse (RN) neural network of the machine learning model with the manual one of the jar tests. The results are shown in Table 2. According to Colombian regulation, the water parameters for human consumption are color 0-15 platinum cobalt units (PCU), turbidity of 0-2 nephelometric turbidity unit (NTU), and pH between 6.5 and 9.0. Analysis of Table 2 shows that the values of the dosage of lime and sulfate obtained through RN are lower than those applied by the operators. Still, these generate a color and turbidity output within the norm and equal or improve the pH value. However, the parameter remains below the norm's values, an aspect that is adjusted by placing an offset to the minimum value obtained in predicting the neural model.

Table 2. Comparison of test results of jugs against neuronal network

	Variable	Test									
		1	2	3	4	5	6	7	8	9	10
Output Data	Color (Jar Test)	10.5	14.9	13.6	6.8	4.7	20.1	9.5	9.8	7	14.1
	Color (Jar Test)	10.83	10.86	10.83	10.72	7.62	17.87	18.35	11.24	8.48	11.87
	Turbidity (Jar Test)	1	1.9	0.6	0.3	0.7	1.4	1.7	0.8	0.9	2.9
	Turbidity (Neuronal Net)	1.67	1.69	1.53	1.49	1.4	2.18	2.3	1.56	1.41	1.92
	pH (Jar Test)	6	6.1	5.3	5	4.8	7.4	5.3	6	5.5	6.2
	pH (Neuronal Net)	5.99	6	5.92	5.89	5.93	5.79	6.15	6	6.03	5.94
	Calcium oxide (lime) (ppm)	18.85	19.43	19.49	19.33	17.64	18.26	14.01	19.96	18.24	19.22
	Aluminum Sulfate (ppm)	46.53	47.32	47.95	47.81	44.56	43.84	32.46	49.23	45.88	44.54
Dosage Reduction (%)	Calcium oxide (lime) (ppm)	5.75%	11.68%	2.55%	7.95%	11.80%	23.92%	29.95%	4.95%	17.09%	3.90%
	Aluminum Sulfate (ppm)	6.94%	9.00%	4.10%	4.38%	10.88%	12.32%	37.58%	5.33%	11.77%	14.35%

4. CONCLUSION

The performance of the algorithms of artificial intelligence in solutions-oriented to the industry was demonstrated with the coagulant and pH control system for water treatment developed through a neural model. The optimal dosage of sulfate and lime obtained by this method generated an output pH lower than 7.5 and output turbidity lower than 8 NTU. The treatment plant's output presents low pH problems, as could be evidenced in the data from the jar tests obtained since they are below the range suitable for human consumption. For this reason, the predictive model created from this data optimizes and standardizes the chemistry of the process. Still, it is necessary to correct pH in the jar tests and thus be able to update the database and retrain the model to correct output error. Then, it is conclusive that to make more efficient models that meet the requirements of drinking water treatment, more effective characterization data must be obtained.

REFERENCES

[1]   X. Yu, X. Yang, C. Yu, J. Zhang, and Y. Tian, "Direct approach to optimize PID controller parameters of hydropower plants," *Renewable Energy*, vol. 173, pp. 342–350, Aug. 2021, doi: 10.1016/j.renene.2021.03.129.

[2]   A. N. Abdullah and M. H. Ali, "Direct torque control of IM using PID controller," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 1, p. 617, Feb. 2020, doi: 10.11591/ijece.v10i1.pp617-625.

[3]   M. Shirzadeh, A. Amirkhani, N. Tork, and H. Taghavifar, "Trajectory tracking of a quadrotor using a robust adaptive type-2 fuzzy neural controller optimized by cuckoo algorithm," *ISA Transactions*, vol. 114, pp. 171–190, Aug. 2021, doi: 10.1016/j.isatra.2020.12.047.

[4]   Y. Zhang, H. Ma, and J. Xu, "Neural network-based fuzzy vibration controller for offshore platform with random time delay," *Ocean Engineering*, vol. 225, p. 108733, Apr. 2021, doi: 10.1016/j.oceaneng.2021.108733.

[5]   H. Sathishkumar and S. S. Parthasarathy, "A novel neural network intelligent controller for vector controlled induction motor drive," *Energy Procedia*, vol. 138, pp. 692–697, Oct. 2017, doi: 10.1016/j.egypro.2017.10.202.

[6]   S. A. Hosseini, A. S. Shirani, M. Lotfi, and M. B. Menhaj, "Design and application of supervisory control based on neural network PID controllers for pressurizer system," *Progress in Nuclear Energy*, vol. 130, p. 103570, Dec. 2020, doi: 10.1016/j.pnucene.2020.103570.

[7]   M. Pirasteh-Moghadam, M. G. Saryazdi, E. Loghman, A. K. E., and F. Bakhtiari-Nejad, "Development of neural fractional order PID controller with emulator," *ISA Transactions*, vol. 106, pp. 293–302, Nov. 2020, doi: 10.1016/j.isatra.2020.06.014.

[8]   A. A. S. Mohamed, H. Metwally, A. El-Sayed, and S. I. Selem, "Predictive neural network based adaptive controller for grid-connected PV systems supplying pulse-load," *Solar Energy*, vol. 193, pp. 139–147, Nov. 2019, doi: 10.1016/j.solener.2019.09.018.

[9]   R. F. da Costa, O. Saotome, E. Rafikova, and R. Machado, "Fast real-time SDRE controllers using neural networks," *ISA Transactions*, vol. 118, pp. 133–143, Dec. 2021, doi: 10.1016/j.isatra.2021.02.019.

[10]   L. Amigues, V. Pommier-Budinger, and J. Bordeneuve, "Hysteresis dispersion compensation with neural network based controller," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 7691–7697, 2020, doi: 10.1016/j.ifacol.2020.12.1513.

[11]   V. de J. da S. Ribeiro, G. F. de M. Oliveira, M. Cristian, A. L. Martins, L. D. Fernandes, and M. P. Vega, "Neural network based controllers for the oil well drilling process," *Journal of Petroleum Science and Engineering*, vol. 176, pp. 573–583, May 2019, doi: 10.1016/j.petrol.2019.01.078.

[12]   C. Lupu, A. Spinu, and A. Udrea, "Intelligent precision control for Air heater systems," *IFAC Proceedings Volumes*, vol. 43, no. 22, pp. 81–85, 2010, doi: 10.3182/20100929-3-RO-4017.00015.

[13]   B. Hasan, S. Alani, and M. A. Saad, "Secured node detection technique based on artificial neural network for wireless sensor network," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 1, p. 536, Feb. 2021, doi: 10.11591/ijece.v11i1.pp536-544.

[14]   M. Hernandez-Gonzalez and E. A. Hernandez-Vargas, "Discrete-time super-twisting controller using neural networks," *Neurocomputing*, vol. 447, pp. 235–243, Aug. 2021, doi: 10.1016/j.neucom.2021.03.060.

[15]   a. G. Abdullah, M. sh. Aziz, and B. A. Hamad, "Comparison between neural network and P&O method in optimizing MPPT control for photovoltaic cell," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 5, p. 5083, Oct. 2020, doi: 10.11591/ijece.v10i5.pp5083-5092.

[16]   A. Zahidi, S. Amrane, N. Azami, and N. Nasser, "Monitoring of solenoid parameters based on neural networks and optical fiber squeezer for solenoid valves diagnosis," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 2, p. 1697, Apr. 2021, doi: 10.11591/ijece.v11i2.pp1697-1708.

[17]   V. Teslyuk, P. Denysyuk, N. Kryvinska, K. Beregovska, and T. Teslyuk, "Neural controller for smart house security subsystem," *Procedia Computer Science*, vol. 160, pp. 394–401, 2019, doi: 10.1016/j.procs.2019.11.075.

[18]   D. Wang and H. Xiang, "Composite control of post-chlorine dosage during drinking water treatment," *IEEE Access*, vol. 7, pp. 27893–27898, 2019, doi: 10.1109/ACCESS.2019.2901059.

[19]   A. Pinto, A. Fernandes, H. Vicente, and J. Neves, "Optimizing water treatment systems using artificial intelligence based tools," Aug. 2009, pp. 185–194, doi: 10.2495/WRM090171.

[20]   S. Al Aani, T. Bonny, S. W. Hasan, and N. Hilal, "Can machine language and artificial intelligence revolutionize process automation for water treatment and desalination?," *Desalination*, vol. 458, pp. 84–96, May 2019, doi: 10.1016/j.desal.2019.02.005.

[21]   L. Li, S. Rong, R. Wang, and S. Yu, "Recent advances in artificial intelligence and machine learning for nonlinear relationship analysis and process control in drinking water treatment: A review," *Chemical Engineering Journal*, vol. 405, p. 126673, Feb. 2021, doi: 10.1016/j.cej.2020.126673.

[22]   H. Maier, "Use of artificial neural networks for predicting optimal alum doses and treated water quality parameters," *Environmental Modelling & Software*, vol. 19, no. 5, pp. 485–494, May 2004, doi: 10.1016/S1364-8152(03)00163-4.

[23]   N. Akshay and D. Subbulekshmi, "Online auto selection of tuning methods and auto tuning PI controller in FOPDT real time process-pH neutralization," *Energy Procedia*, vol. 117, pp. 1109–1116, Jun. 2017, doi: 10.1016/j.egypro.2017.05.235.

[24]   L. Estofanero, R. Edwin, and G. Claudio, "Predictive controller applied to a pH neutralization process," *IFAC-PapersOnLine*, vol. 52, no. 1, pp. 202–206, 2019, doi: 10.1016/j.ifacol.2019.06.062.

[25]   T. P. Rose and G. G. Devadhas, "Detection of pH neutralization technique in multiple tanks using ANFIS controller," *Microprocessors and Microsystems*, vol. 72, p. 102845, Feb. 2020, doi: 10.1016/j.micpro.2019.07.004.

[26]   N. C. Chulluncuy-Camacho, "Tratamiento de agua para consumo humano," *Ingeniería Industrial*, no. 029, p. 153, Mar. 2011, doi: 10.26439/ing.ind2011.n029.232.

BIOGRAPHIES OF AUTHORS

Oscar Iván Vargas Mora 🆔 Ⓖ SC 🅓 is an Electronic Engineer graduated from Universidad de los Llanos in 2008. Specialist in Industrial Instrumentation and Control from Universidad de los Llanos in 2021. His current working as commissioning engineer in Oil & Gas industry, and ten years as Instrument Engineer for O&G and energy industry. He can be contacted at email: oscar.vargas.mora@unillanos.edu.co.

Daiam Camilo Parrado Nieto ⓘ 🗗 SC ⚫ is a Mechatronic Engineer graduated from Universidad San Buenaventura in 2008 and Specialist in Industrial Instrumentation and Control from Universidad de los Llanos in 2021. His current working as a resident engineer in the Automation of the Ariari Regional Aqueduct and eight years as a Mechatronic Engineer for drinking water and water treatment. He can be contacted at email: daiam.parrado@unillanos.edu.co.

Jairo David Cuero Ortega ⓘ 🗗 SC ⚫ is a professor at School of Engineering in Universidad de los Llanos. In 2012 he graduated as an Electronics Engineer and in 2015 as an Industrial Instrumentation and Control Specialist from the same alma mater. He received the Master of Engineering degree in automation from Universidad Nacional de Colombia in 2022. His current research focuses are on Instrumentation, Control, and Robotics. He can be contacted at email: jairo_cuero@unillanos.edu.co.

Javier Eduardo Martinez Baquero ⓘ 🗗 SC ⚫ is an Electronic Engineer graduated from Universidad de los Llanos in 2002. Posgraduated in Electronic Instrumentation from Universidad Santo Tomas in 2004, Posgraduated in Industrial Instrumentation and Control from Universidad de los Llanos in 2020 and M.Sc. in Educative Technology and Innovative Media for Education from Universidad Autonoma de Bucaramanga in 2013. Currently, he works as associated professor in Universidad de los Llanos, his research focuses on Instrumentation, Automation, Control and Renewable Energies. He can be contacted at email: jmartinez@unillanos.edu.co.

Robinson Jiménez-Moreno ⓘ 🗗 SC ⚫ is an Electronic Engineer graduated from Universidad Francisco José de Caldas in 2002. He received a M.Sc. in Engineering from Universidad Nacional de Colombia in 2012 and Ph. D in Engineering at Universidad Francisco José de Caldas in 2018. His current working as assistant professor of Universidad Militar Nueva Granada and research focuses on the use of convolutional neural networks for object recognition and image processing for robotic applications such as human-machine interaction. He can be contacted at email: robinson.jimenez@unimilitar.edu.co.