Desalination and Water Treatment
www.deswater.com

doi: 10.5004/dwt.2023.29497

292 (2023) 60–68
April

Acquisition and analysis of floc images by machine learning technique to improve the turbidity removal process

Fabricio Paredes Larroca[a], Javier Quino-Favero[a], Uwe Rojas Villanueva[b], Erich Saettone Olschewski[a,*]

[a]*Universidad de Lima, Instituto de Investigación Científica, Grupo de Investigación en Soluciones Tecnológicas para el Medio Ambiente, Perú, emails: esaetton@ulima.edu.pe (E.S. Olschewski), fparedes@ulima.edu.pe (F.P. Larroca), jquinof@ulima.edu.pe (J. Quino-Favero)*
[b]*Universidad de Lima, Carrera de Ingeniería Industrial, Perú, email: urojas@ulima.edu.pe*

Received 7 November 2022; Accepted 23 March 2023

A B S T R A C T

This article reports on the implementation and use of a floc image acquisition and analysis system in a pilot water treatment plant to remove kaolin turbidity with a coagulant and flocculant. The system is based on the Hausdorff dimension (d_f) of the images and is used to obtain information about the image texture and to ensure that the flocs could be removed by the filtration system, and to use d_f values for corrections of the dosage of both chemical agents via signals with pulse width modulation that feed and control dosage pumps during treatment, ensuring a continuous adjustment for changing water conditions, which allows for a close on-site process control and a rapid response to changes in the quality of the effluent.

Keywords: Supervised machine learning; Coagulation; Flocculation; Turbidity; Hausdorff dimension

1. Introduction

Urban and industrial development in areas where water is scarce requires the search for sustainable solutions to rationally use water. In this context, one of these solutions that exhibits the greatest potential is the reuse of wastewater [1]. Treatment of wastewater requires a balance between the technological level and the operational complexity required to achieve high-quality effluent while maintaining the type of simplicity that allows infrequent monitoring and maintenance [2]. In addition, approaches that characterise uncertainties in the long term are needed, including micropollutants and pathogenic microorganisms [3]. The coupling of big data and automation, also known as Industry 4.0, has the potential to solve these complex problems by increasing operational efficiency and improving the ability to monitor and control functions at reduced costs [4]. This is achieved by applying interconnection, information transparency, technical assistance, and decentralised decision making [5].

In water applications, artificial intelligence (AI), machine learning (ML), and smart technologies are expected to model and address complex problems, water applications that have seen notable ML utilization include water and wastewater treatment [6], ammonia concentration and total coliform concentration have been predicted using supervised machine learning [7] and to predict water quality using supervised machine learning algorithms [8].

Wastewater treatment plants are subject to extreme dynamics throughout their life expectancy, and since they are designed with fixed performance demands and pollutant loads, new approaches are required to characterize uncertainty through time [3]. In addition, these technologies allow the autonomy of the plant to be operated, reducing energy consumption, and improving competitiveness [9]. In

* Corresponding author.

1944-3994/1944-3986

the literature, there are reports on internet of things (IoT)-based systems-together with mathematics, statistics, neural networks, and sensor integration-for technological interventions to eliminate sudden (negative) controllable changes [10], which indicates this technology is still being developed.

The limitations associated with the use of ML are that these models are highly dependent on the selected data, which sometimes limits the accessibility where there is a lack of data management/storage capabilities necessary for the function of these models [11]. Some ML techniques are subject to poor reproducibility when developed using random weights that work only with characteristics similar to those of the dataset originally trained and tested [12].

Aggregation processes (coagulation/flocculation) are widely used in water treatment. In many cases, different methods may be used to monitor the aggregation process, especially when related to the dosage of additives used in the process. Hydrolyzing metal salts, such as aluminum and ferric salts, are used to promote the aggregate formation by neutralizing the of negatively charged colloids by cationic hydrolysis and the incorporation of pollutants into an amorphous solid precipitate [13]. Additionally, adsorption of ferric ions onto kaolin particles occurs by formation of surface complexes between positive charged ferric ions and hydroxyl groups on the surface of the mineral. As kaolinite turbidity is a problem faced by the clay processing industry, kaolinite has been extensively used to model the presence of turbidity in water [14–16]. Therefore, a kaolin and ferric chloride-based coagulation process is useful to test a process for removal of turbidity. Once aggregates (flocs) are formed, their characteristics such as size, density, and fractal dimension can be measured. Once the flocs have been filtered or sedimented, the efficacy of the process is assessed by measuring the residual turbidity, which is the most widely used method [17].

Significant developments in advanced image analysis have allowed particle/floc analysis in real time and have offered an easy way to measure a variety of properties for these particles [18]. The application of ML technologies in coagulation/flocculation is based on two aspects: predicting effluent turbidity and predicting the amount needed based on influent turbidity. However, its implementation is expensive due to the need for specialized software and equipment. Therefore, it is necessary to take advantage of the many benefits offered by ML technologies to cut costs without sacrificing the expected results.

This article reports on the implementation and use of a floc image acquisition and analysis system in a pilot water treatment plant to remove kaolin turbidity with a coagulant and a flocculant. During a first step, a factorial experiment design is used to test several coagulant and flocculant doses in continuous flow to obtain the Hausdorff dimension (d_f) of the images and to relate them to the turbidity post-treatment, thus identifying the value of the d_f and relating it with the lowest value of turbidity. As a second step, images are then used to create a database with pre-treatment turbidity values (T_{in}) and post-treatment turbidity values (T_{out}), flocculant (S_{fl}) and coagulant (S_{co}) doses. The Hausdorff dimension is then used to obtain information about the image texture to ensure that the flocs could be removed by the filtration system and using d_f values for corrections of the dosage of both chemical agents via signals with pulse width modulation (PWM) that feed and control dosage pumps during treatment, ensuring that d_f values stay within established limits.

2. Materials and methods

2.1. Kaolin suspensions

When injected into running water at a continuous flow, kaolin is used for obtaining a homogeneous mixture with constant turbidity, which is known as pre-treatment water. The kaolin grains (COMACSA, Lima, Peru) were ground for 4 h and then sieved at 25 µm. A sample was characterised by X-ray diffraction (XRD) (Bruker D8 Advance, USA). The suspensions were prepared mixing 2.0 g of kaolin per liter of deionised water at a 6.5 pH level of 6.5 and maintained in constant agitation with a mixer (Lightnin L1U08F, USA).

2.2. Water treatment plant

The treatment plant (Fig. 1) is made up of a speed pump drive (Cole Parmer 75211-15, USA), which provides a 2 L/min tap water flow (FNU < 5). This flow is injected with the kaolin suspension by means of a peristaltic pump (LeadFluid YZ15), obtaining a water mixture with a T_{in} = 60–110 FNU with an average input turbidity of 70 FNU. Treatment to remove turbidity includes the addition of a 24% ferric chloride coagulant followed by a vigorous mixture in a static mixer. Subsequently, 0.2% Magnafloc 394 (BASF) flocculant is added followed by a 12 m-long 3.8 cm diameter PVC tube flocculator, with a Minitwist filtration system with a 25-µm steel filtration element (Pelmar Engineering LTD, ON Canada) at its tip for capturing the flocs that have formed. The residence time in the water treatment plant was approximately 2.5 min. Turbidity is continuously measured using

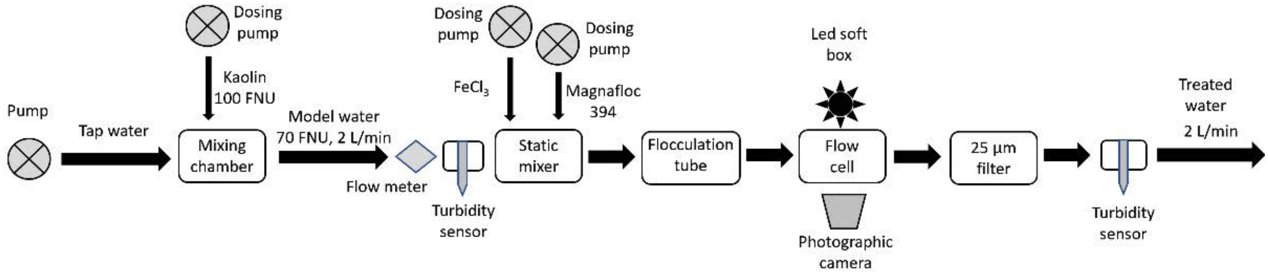


Fig. 1. Diagram for the turbidity removal process and image acquisition system.

two sensors (Turbimax CUS52D, Endress + Hauser) placed after the input of the kaolin suspension and at the 25 μm mesh filter output.

The plant has a SCADA (supervisory control and data acquisition) system [19], but it was only used for data acquisition because the control is carried out by the proposed solution. The plant included two turbidity sensors connected to a transmitter (Endress + Hauser Liquiline CM448), a PLC (S7-1500 Siemens, Germany) that receives information from the transmitter, a Dell 6490 I5 Intel computer, 24 Gb RAM, 64-bit Windows system 10 environment running MATLAB, and the IoT (2040 Siemens) that generates PWM signals to control the two coagulant and flocculant dosing pumps (LeadFluid 50 S and BT101S, respectively) in the plant.

A PWM is a signal consisting of square pulses of two values, zero and a maximum value per cycle. Each cycle has a constant period; however, the duration of zero and maximum can be varied to generate signals with different rms values to control the pump speed, hence the dosing of the chemical agents in the treatment plant. A low dosage corresponds to a PWM signal with more duration of zero value than the maximum one every cycle, and vice versa.

2.3. Image capture and treatment

The flocs formed pass through the inner channel of a flow cell made by two Plexiglas plates lit posteriorly with white, diffuse led light (Philips DL252, The Netherlands). The photographic record of images begins when obtaining constant turbidity values at the treatment plant input (T_{in}). The images were taken every 20 s during 10 min with a Canon T3i camera with a Canon EFS 18–55 mm lens and generated in RAW format. The camera was controlled by the "Image Acquisition Tool" by MATLAB [20]. All photographs were taken 12 cm away from the anterior flow cell side to keep the same scale across images (Fig. 2).

The images registered by the camera were obtained in RAW format and processed to extract the region of interest (ROI), then converted to greyscale, filtered, and improved their contrast. The process has 16 stages, as shown in Fig. A1.

2.4. Experimental design

A preliminary value of the coagulant and flocculant dose was obtained using a jar test (ASTM D2035-08) using a Platypus Jar Tester device (Microfloc Pty Ltd., Australia) device, measuring turbidity before and after the treatment (Lovibond infrared turbidity meter TB210 IR).

The most effective combination of coagulant and flocculant obtained from the jar test was assigned a zero coded value from a completely randomised factorial design 2^3 of two variables and three levels. Noncoded values were obtained by multiplying the concentration of coagulant and flocculant by the pump flow and dividing it according to the flow of treated water flow (Table A1).

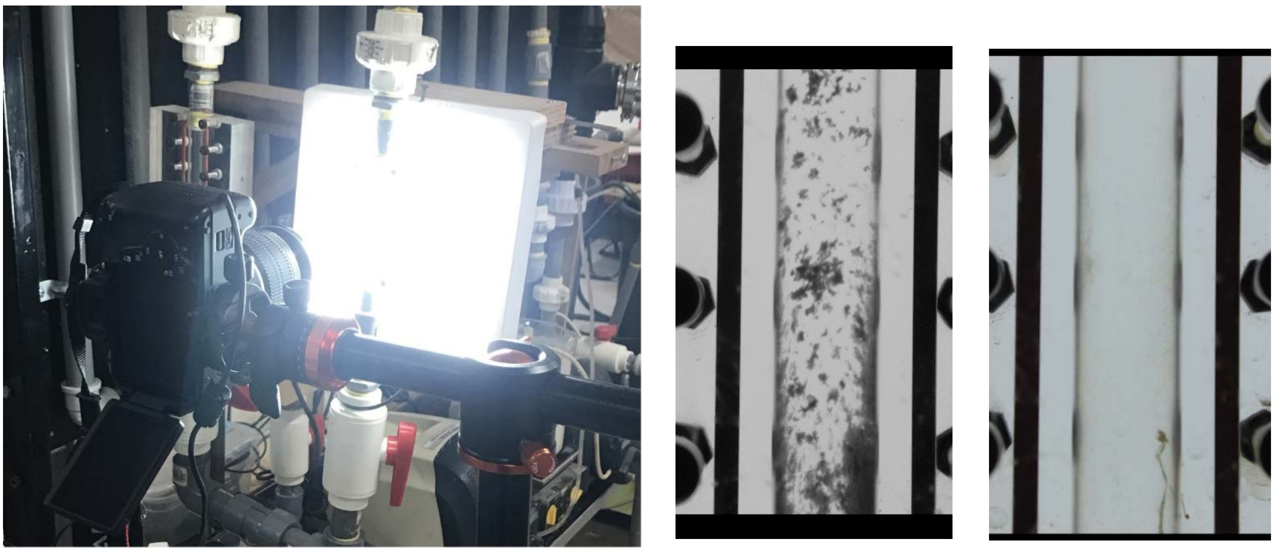
The trials were carried out at the water treatment plant described in Section 2.2 – Water treatment plant. The flow process is shown in Fig. 3.

2.5. Hausdorff dimension (d_f)

The fractal dimension of the flocs allows it to be associated with turbidity removal. This fractal dimension may be defined by different approaches, of which the Hausdorff dimension (d_f) is the most common. In this way, once the image is treated, this dimension is calculated using the box count method [21] through Eq. (1) [22].

$$d_f = \lim_{\varepsilon \to 0} \frac{\log N(\varepsilon)}{\log(1/\varepsilon)}$$

where $N(\varepsilon)$ represents the small boxes in the object with ε dimensions and ε sides [23]. This calculation is performed with the Box-Counting tool [24] by MATLAB [25]. A process diagram for calculating the d_f from the images is shown in Fig. 4.



(a) (b) (c)

Fig. 2. (a) Image acquisition system, (b) flow cell with flocs, and (c) floc-free cell.

2.6. Evaluation of supervised models

Each trial lasts 10 min, in which 30 images are recorded (once every 20 s). Each image has its d_f calculated, which allows for creating a database containing 30 packs of 5 data each: input turbidity (T_{in}), output turbidity (T_{out}), Hausdorff dimension (d_f), flocculant (S_{fl}) and coagulant (S_{co}) dosage values. From this database, five supervised models were assessed: logistic regression (LR), linear discriminant analysis (LDA), *K*-nearest neighbours (KNN), Naive Bayes (NB) gaussian [26], and support vector machine (SVM). These models were chosen because an output (turbidity) is necessary for a supervised model.

These dose values predicted in each model are contrasted with the actual values measured in the trials. Then, the model that produces the best dosage prognosis is selected.

2.7. Dosing adjustment

The distribution of d_f that allows separation in the filter, and which is also associated with the lowest turbidity value, becomes the range of desired values. In this way, a d_f value outside the range modifies the dosage in such a way that the desired d_f range is reached.

In other words, when a desirable turbidity value is established and the data acquisition by the sensors begins, there is a small delay with respect to the image acquisition system used to obtain the Hausdorff dimension. To achieve chemical dosing, a difference is established between the input turbidity and the output turbidity, and the difference is interpreted as an error; since the image acquisition system obtained the image in advance – due to the delay in sensor-mediated acquisition- and has processed the image data using the Naive Bayes algorithm for the required coagulant and flocculant dosage, the system based on the algorithm is ahead. When the sensor finds the new turbidity value, the image acquisition system will already have calculated the new Hausdorff dimension value together with the dosage prediction. Thus, the error continues to be calculated until the system reaches the desired value, at which point it will stop until a new adjustment is required due to the lack of flocculant and coagulant dosage.

Dosing adjustments are achieved through a PWM signal. Set in Firmata mode, the IoT 2040 turns into a peripheral that receives S_{fl} and S_{co} and, via the pyFirmata software, generates the PWM pulses [27] that control dosage pumps.

3. Results

3.1. Kaolin compounds

The XRD spectrum of the kaolin suspension obtained as described in Section 2.1 – Kaolin suspensions is shown in Fig. 5 and shows a typical kaolin spectrum with peaks of kaolinite (K), illite (I) and β-quartz (Q).

3.2. Hausdorff dimension as a proxy for controlling turbidity

Fig. 6 shows the temporal results of two trials with a 2.5 mg/L flocculant dosage, 15 mg/L coagulant dosage (Trial #1, Table A1), 3.75 mg/L flocculant dosage, and 23 mg/L coagulant dosage (Trial #8, Table A1).

As shown, in t = 0 all the input and output values coincide with an initial value of about 5 FNU corresponding to the moment at which running water injection starts.

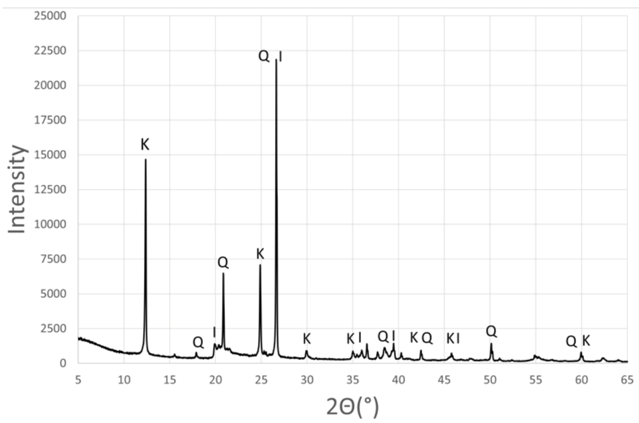


Fig. 5. Kaolin XRD spectrum. K: kaolinite; Q: β-quartz; I: illite.

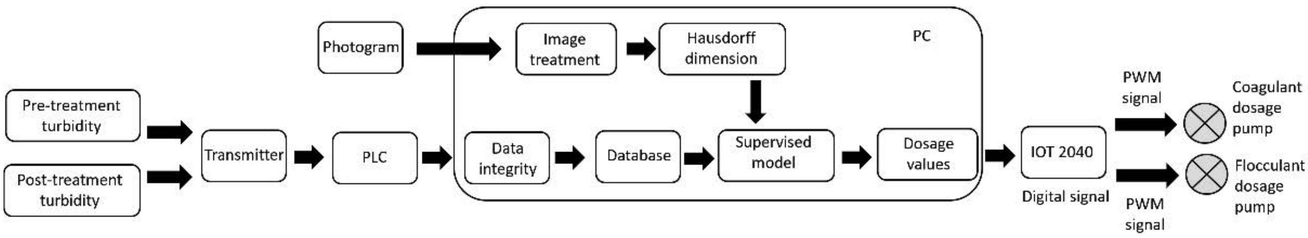


Fig. 3. Flow process of the experimental design.

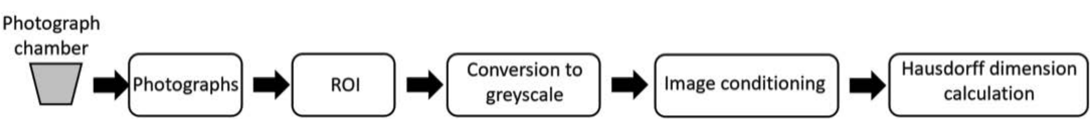


Fig. 4. Diagram of the image treatment process for calculating the Hausdorff dimension.

Immediately after, T_{in} gradually increased due to the injection of a mixture of water and kaolin. T_{out} remains almost unchanged until it abruptly increases as a result of the arrival of water with kaolin at the plant output.

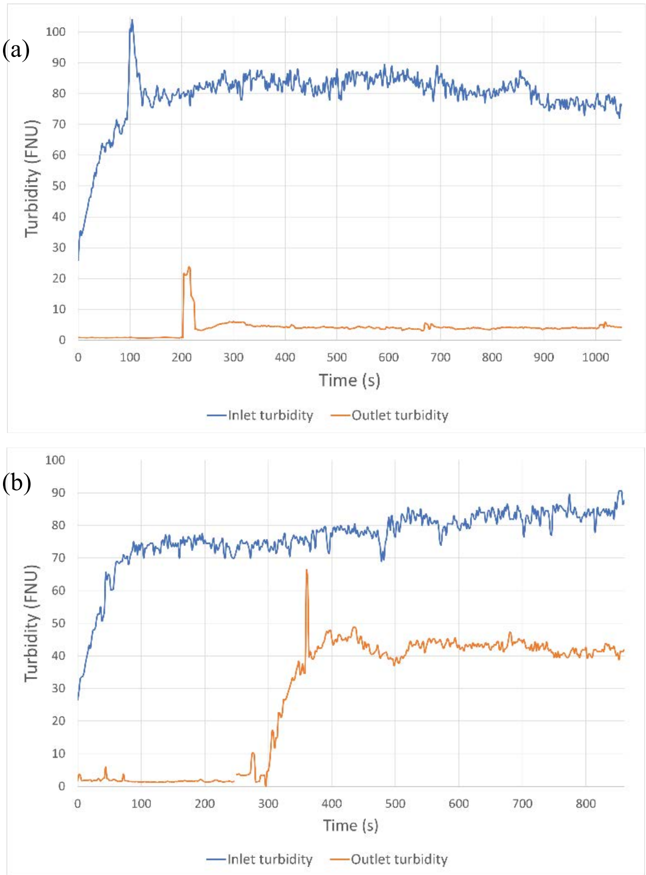


Fig. 6. Values of turbidity of (a) trial 1 and (b) trial 8, as shown in Table 1.

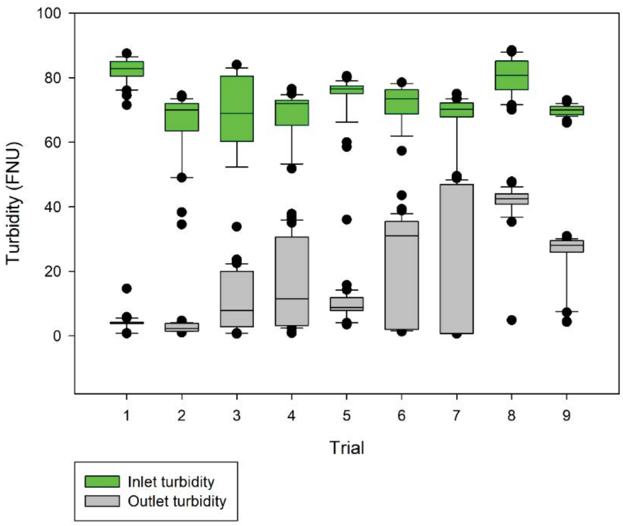


Fig. 7. Box plot for inlet turbidity and outlet turbidity.

Fig. 7 shows the inlet and outlet turbidity measurements for each test. The highest turbidity removals (higher differences between inlet and outlet turbidities) were obtained during trials 1 (2.5 mg/L flocculant and 15 mg/L coagulant dose) and 2 (3.75 mg/L flocculant and 15 mg/L coagulant dosage). Although trials 1 and 2 produced the best removals, the Hausdorff dimension value of trial 2 is more appropriate to act as a control variable because its distribution has less overlapping with the rest of the values of each trial (Fig. 8).

3.3. Evaluation of supervised models

The mean accuracy values obtained with the selected models are 0.0705 for SVM, 0.2002 for LDA, 0.8254 for LR, 0.9297 for KNN, and 0.9916 for Naive Bayes gaussian (NB). In Fig. 10 are shown additional results obtained from the assessment. Results are shown in Fig. 9.

As shown, the algorithm with the best accuracy to predict the dosages of flocculant and coagulant is NB.

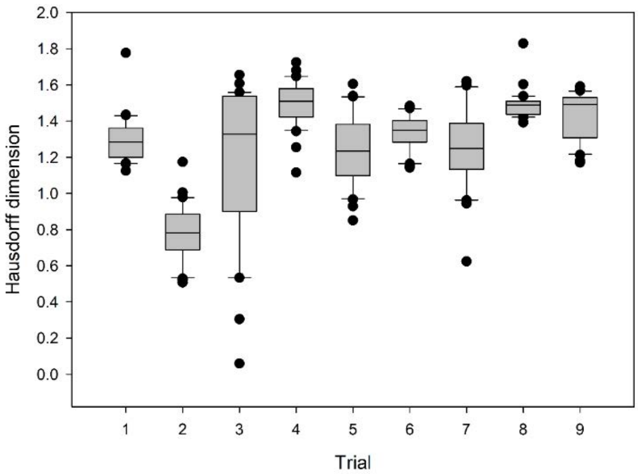


Fig. 8. Box plot of the calculated Hausdorff dimension for each trial.

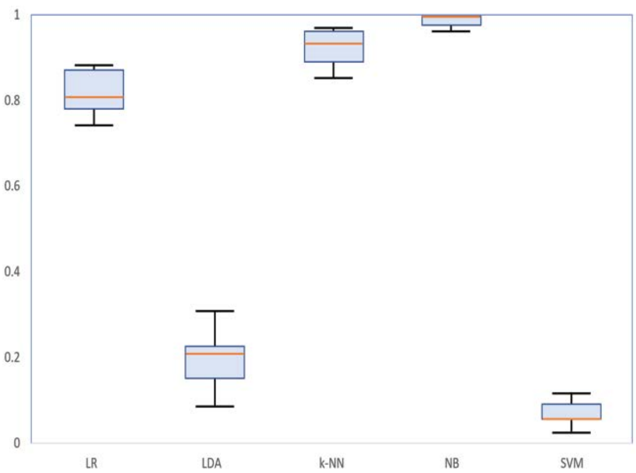


Fig. 9. Box plot of the assessed models.

(a)

(b)

(c)

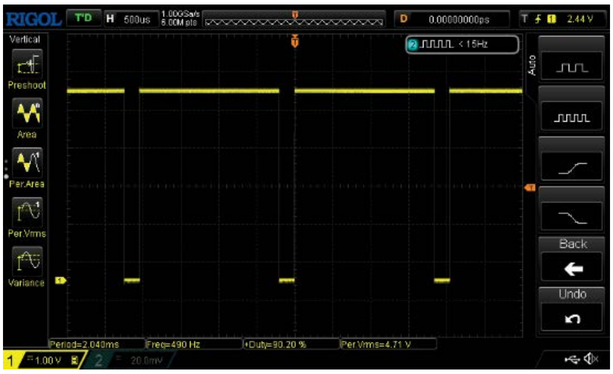


Fig. 10. Some examples of pulse width modulation signals generated to control pump dosage as shown in the oscilloscope: (a) low dosage, (b) medium dosage, and (c) high dosage. pulse width modulation signals are comprised of two pulses, zero voltage and maximum voltage (shown as a discontinuous yellow line at the bottom and top of each graph, respectively); the combination of the two pulses controls the pump speed where a higher duration of zero value means lower dosing, and a higher duration of the maximum value means higher dosing.

Table 1
f-score for prediction using the Naive Bayes model

Variable	Symbol	Value	Percentage (%)
Hausdorff dimension	d_f	5,642	43.1
Turbidity post-treatment	T_{out}	4,520	34.5
Turbidity pre-treatment	T_{in}	2,394	18.4
Flocculant dosage	S_{fl}	303	2.4
Coagulant dosage	S_{co}	227	1.7

3.4. Evaluation of *f*-score for the NB model

An assessment of the predominance between variables (*f*-score) that explains the prediction process for water treatment through the NB model delivered the water results shown in Table 1.

Here, the variable with the highest significance is d_f, with a 43.1% *f*-score, followed by T_{out} with 34.5%. Both contributed to 77.6% of the prediction.

3.5. Dosage pump control

Fig. 10 shows some results for the PWM control pump signal with low, medium, and high dosage [28] (Fig. 10a–c, respectively) as explained in Section 2.2 – Water treatment plant.

4. Discussion

The process used in this study starts with obtaining floc images during water treatment. Images have previously been reported to assess the efficacy of coagulation and flocculation processes [23,29–31]. The most common methodology requires calculating the Hausdorff dimension to establish the characteristics of the obtained [32,33]. The association between d_f and turbidity post-treatment is used as a model to find the dosage of chemical products that are usually found during a jar test. The disadvantages of using the jar test to find better dosages are operator time consumption for the preparation and execution of the test, the use of dosages restricted to certain values, and the fact that the water characteristics may vary a moment after the test was made. Finding the appropriate dosage for coagulant and flocculant establishing a relationship between output turbidity and the Hausdorff dimension for the flocs allows for a closer on-site process control and a continuous adjustment for changing water conditions, which offers very short response times and potential savings in the cost of the chemicals used to remove turbidity.

Of the five supervised algorithms assessed for dosage prediction, NB was clearly superior with an accuracy score of 0.9916 (Fig. 7), with d_f being the best predictor variable (*f*-score 43.1%) followed by turbidity after treatment (*f*-score 34.5%). Both explain 77.6% of the variability observed in the treatment process and can be used to make predictions about the behavior of the turbidity and adjust the dosage of chemical products for treatment.

The d_f associated with the highest turbidity removal was 0.507–1.174, less than 1.225–1.525 previously reported [34], the fractal dimension reflects the degree of compaction degree and the ease of sedimentation ease of flocs [35], where higher values usually describe flocs that are more likely to sediment. Although other studies report higher d_f values (1.659–1.809 [36], 2.45–2.44 [37]), the small values for the fractal dimension obtained by this study did not cause separation issues since we used a filtration system instead of a sedimentation tank.

In this study, the Hausdorff dimension was obtained from the image containing several flocs, such as the one shown in Fig. 2b. This procedure significantly reduces the computing time required to calculate the d_f for each floc in each image. In the latter procedure, turbidity reduction control is likely more efficient, but it will require more computing power, leading to a more expensive control system.

The results show that it is possible to stop using turbidity sensors and replace them with a photographic chamber while working within the described turbidity range (60–110 FNU), which allows reducing the implementation costs for the control system. However, the utility of the system is limited to process with active floc removal, such as filtration. In this work, the turbidity removal process works straightforward because the floc removal is being done by filtration, and the effluent turbidimeter measurements are taken in real time, facilitating a rapid response to changes in the quality of the effluent. When floc removal is carried out by sedimentation, the information based on the effluent turbidimeter is delayed due to the large lag characteristics of the flocculation process, therefore, a prediction model will be more efficient [38].

An automated system for turbidity monitoring requires at least two turbidity probes, a transmitter, and a PLC with proprietary software. The proposed system could run on a small computer like a Jetson nano or a Raspberry Pi IV for image acquisition and processing, a simple HD camera that replaces the turbidity probes, a Siemens 2040 IoT, and open-source software.

5. Conclusions

A turbidity monitoring and control system based on the Hausdorff dimension has been developed using conventional imaging equipment. Chemical dosing that produces floc images with Hausdorff dimension values ranging between 0.507–1.174 was associated with high levels of turbidity removal. The range 0.507–1.174 range is different enough from other ranges to be used to control the water treatment process. The proposed solution could be replicated with simple and cost-effective equipment with an estimated cost around one-magnitude order below the conventional approach.

References

[1] X. Garcia, D. Pargament, Reusing wastewater to cope with water scarcity: economic, social and environmental considerations for decision-making, Resour. Conserv. Recycl., 101 (2015) 154–166.

[2] L.S. Gaulke, On-site wastewater treatment and reuses in Japan, Proc. Inst. Civ. Eng. - Water Manage., 159 (2006) 103–109.

[3] D. Dominguez, W. Gujer, Evolution of a wastewater treatment plant challenges traditional design concepts, Water Res., 40 (2006) 1389–1396.

[4] A.T. Rosário, J.C. Días, How industry 4.0 and sensors can leverage product design: opportunities and challenges, Sensors, 23 (2023) 1165, doi: 10.3390/s23031165.

[5] M. Hermann, T. Pentek, B. Otto, Design Principles for Industries 4.0 Scenarios, 49th Hawaii International Conference on System Sciences (HICSS), 2016, pp. 3928–3937.

[6] M. Lowe, R. Qin, X. Mao, A review on machine learning, artificial intelligence, and smart technology in water treatment and monitoring, Water, 14 (2022) 1384, doi: 10.3390/w14091384.

[7] A. Najah Ahmed, O. Faridah Binti, A. Haitham Abdulmohsin, I. Rusul Khaleel, F. Chow Ming, H. Md Shabbir, M. Ehteram, A. Elshafie, Machine learning methods for better water quality prediction, J. Hydrol., 578 (2019) 124084, doi: 10.1016/j.jhydrol.2019.124084.

[8] N.H.A. Malek, W.F. Wan Yaacob, S.A. Md Nasir, N. Shaadan, Prediction of water quality classification of the Kelantan River Basin, Malaysia, using machine learning techniques, Water, 14 (2022) 1067, doi: 10.3390/w14071067.

[9] A. Korodi, R. Crisan, A. Nicolae, I. Silea, Industrial internet of things and fog computing to reduce energy consumption in drinking water facilities, Processes, 8 (2020) 282, doi: 10.3390/pr8030282.

[10] D.R. Prapti, A.R. Mohamed Shariff, H. Che Man, N.M. Ramli, T. Perumal, M. Shariff, Internet of things (IoT)-based aquaculture: an overview of IoT application on water quality monitoring, Rev. Aquacult., 14 (2022) 979–992.

[11] E. Doherty, G. McNamara, L. Fitzsimons, E. Clifford, Design and implementation of a performance assessment methodology cognisant of data accuracy for Irish Wastewater Treatment Plants, J. Cleaner Prod., 165 (2017) 1529–1541.

[12] R. Lakshmanan, M. Djama, S. Perumal, R. Abdulla, Automated smart hydroponics system using internet of things, Int. J. Electr. Comput. Eng., 10 (2020) 6389–6398.

[13] J. Duan, J. Gregory, Coagulation by hydrolysing metal salts, Adv. Colloid Interface Sci., 100 (2003) 475–502.

[14] R. Divakaran, V.S. Pillai, Flocculation of kaolinite suspensions in water by chitosan, Water Res., 35 (2001) 3904–3908.

[15] M.G. Kılıç, Ç. Hoşten, A comparative study of electrocoagulation and coagulation of aqueous suspensions of kaolinite powders, J. Hazard. Mater., 176 (2010) 735–740.

[16] Z. Ferasat, R. Panahi, B. Mokhtarani, Natural polymer matrix as safe flocculant to remove turbidity from kaolin suspension: performance and governing mechanism, J. Environ. Manage., 255 (2022) 109939, doi: 10.1016/j.jenvman.2019.109939.

[17] J. Gregory, Monitoring particle aggregation processes, Adv. Colloid Interface Sci., 147–148 (2009) 109–123.

[18] R.F. Nielsen, N.A. Kermani, L. la Cour Freiesleben, K.V. Gernaey, S.S. Mansouri, Novel strategies for predictive particle monitoring and control using advanced image analysis, Comput. Aided Chem. Eng., 46 (2019) 1435–1440.

[19] S. Brad, M. Murar, G. Vlad, E. Brad, M. Popanton, Lifecycle design of disruptive SCADA systems for waste-water treatment installations, Sustainability, 13 (2021) 4950, doi: 10.3390/su13094950.

[20] A.V. Tasamá, A.F. Ramírez Sánchez, D.M. Martínez, M.J. Alzate Silva, Aplicación didáctica para el procesamiento de imágenes digitales usando interfaz gráfica de usuario en MATLAB, Instituto Tecnológico Metropolitano de Medellín-Componentes Electrónicas, 2008, pp. 1–5. Available at: https://compelect.com.co/archivos/diamatlab/2008/Aplicacion_Didactica_PRIM.pdf

[21] P. French, Hausdorff Dimension by the Box Counting Method, MATLAB Central File Exchange, 2022. Available at: https://www.mathworks.com/matlabcentral/fileexchange/15918-hausdorff-dimension-by-the-box-counting-method

[22] A. Napolitano, S. Ungania, V. Cannata, Fractal Dimension Estimation Methods for Biomedical Images, V.N. Katsikis, Ed., MATLAB–A Fundamental Tool for Scientific Computing and Engineering Applications, IntechOpen, 2012, pp. 161–178.

[23] H. Zheng, G. Zhu, S. Jiang, T. Tshukudu, X. Xiang, P. Zhang, Q. He, Investigations of coagulation–flocculation process by performance optimization, model prediction and fractal structure of flocs, Desalination, 269 (2011) 148–156.

[24] J. Li, Q. Du, C. Sun, An improved box-counting method for image fractal dimension estimation, Pattern Recognit., 42 (2009) 2460–2469.

[25] F. Moisy, Boxcount, MATLAB Central File Exchange, 2022. Available at: https://www.mathworks.com/matlabcentral/fileexchange/13063-boxcount

[26] J. Ren, S.D. Lee, X. Chen, B. Kao, R. Cheng, D. Cheung, Naive Bayes Classification of Uncertain Data, 9th IEEE International Conference on Data Mining, Miami Beach, FL, USA, 2009, pp. 944–949.

[27] E. Kravaritis, C. Sourkounis, Efficiency Analysis of Pumps Drives for Space Vector PWM and Hysteresis Band PWM With on Operation Transaction of the Control Method, 25th Mediterranean Conference on Control and Automation (MED), Valletta, Malta, 2017, pp. 473–477.

[28] J.S. Kumar, P. Poongodi, P. Balakumaran, Artificial intelligence-based alum dosage control in water treatment plant, Int. J. Eng. Technol., 5 (2013) 3344–3350.
[29] D. Cruz, M. Pimentel, A. Russo, W. Cabral, Charge neutralization mechanism efficiency in water with high color turbidity ratio using aluminium sulfate and flocculation index, Water, 12 (2020) 572, doi: 10.3390/w12020572.
[30] S. Sun, M. Weber-Shirk, L.W. Lion, Characterization of flocs and floc size distributions using image analysis, Environ. Eng. Sci., 33 (2016) 25–34.
[31] M.R. MacIver, M. Pawlik, Analysis of in situ microscopy images of flocculated sediment volumes, Chem. Eng. Technol., 40 (2017) 2305–2313.
[32] A. Vahedi, B. Gorczyca, Application of fractal dimensions to study the structure of flocs formed in lime softening process, Water Res., 45 (2011) 545–556.
[33] J. Yu, D. Wang, X. Ge, M. Yan, M. Yang, Flocculation of kaolin particles by two typical polyelectrolytes: a comparative study on the kinetics and floc structures, Colloids Surf., A, 290 (2006) 288–294.
[34] C. Ying, L. Qian-Jun, Z. Jin-Song, Flocculation control study based on fractal theory, J. Zhejiang Univ. Sci. B, 6 (2005) 1038–1044.
[35] B. Liu, F. Qu, W. Chen, H. Liang, T. Wang, X. Cheng, H. Yu, G. Li, B. Van der Bruggen, Microcystis aeruginosa-laden water treatment using enhanced coagulation by persulfate/Fe(II), ozone and permanganate: comparison of the simultaneous and successive oxidant dosing strategy, Water Res., 125 (2017) 72–80.
[36] F. Niu, H. Zhang, J. Zhang, X. Yu, Temperature variation characteristics and model optimization of flocculation sedimentation of overflow ultra-fine iron tailings, Minerals, 12 (2022) 643, doi: 10.3390/min12050643.
[37] S. Nieto, N. Toro, P. Robles, E. Gálvez, S. Gallegos, R.I. Jeldres, Flocculation of clay-based tailings: differences of kaolin and sodium montmorillonite in salt medium, Materials, 15 (2022) 1156, doi: 10.3390/ma15031156.
[38] W. Dongsheng, C. Chang, M. Kaiwei, L. Zhixuan, D. Deng, Estimating effluent turbidity in the drinking water flocculation process with an improved random forest model, Water Supply, 22 (2022) 1107–1119.

Appendix

Table A1
Dosage values used in trials

Trial number	Dosage			
	Coded values		Non coded values	
	Coagulant	Flocculant	Coagulant (Fe^{+3}, mg/L)	Flocculant (mg/L)
1	–1	–1	15	2.50
2	–1	0	15	3.75
3	–1	+1	15	5.00
4	0	–1	19	2.50
5	0	0	19	3.75
6	0	+1	19	5.00
7	+1	–1	23	2.50
8	+1	0	23	3.75
9	+1	+1	23	5.00

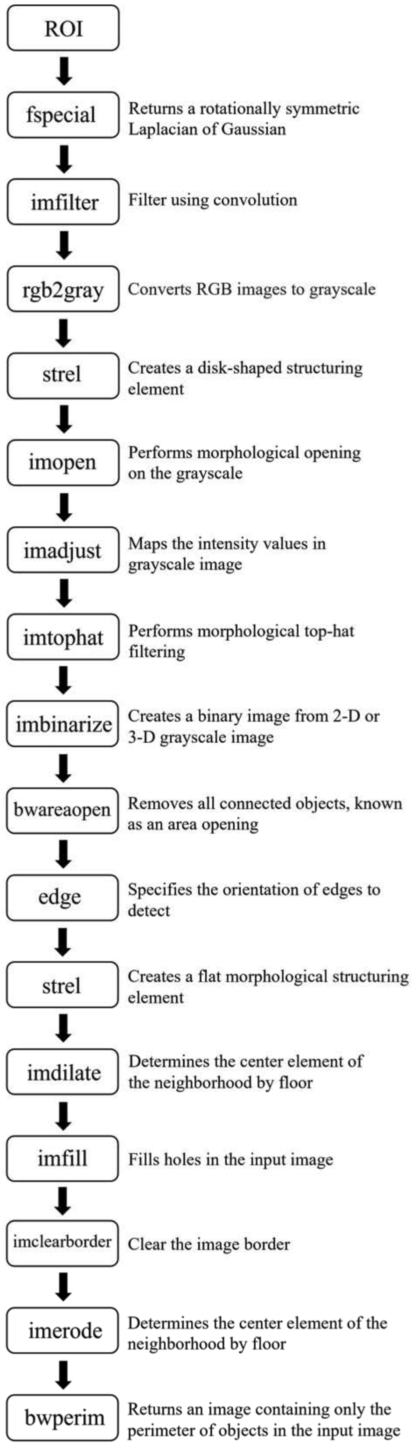


Fig. A1. Image treatment process.