Iranian Reasearch Organization
for Science and Technology
(IROST)

Advances in Environmental Technology 4 (2021) 263-273

Advances in Environmental Technology

Journal home page: https://aet.irost.ir/

Modelling turbidity removal by poly-aluminium chloride coagulant using gene expression

Ruba D. Alsaeed

Faculty of Civil Engineering, Al-Wataniya Private University, Hama, Syria

ARTICLE INFO

Article history:
Received 12 November 2021
Received in revised form
19 January 2022
Accepted 19 January 2022

Keywords:
Coagulant
Gene expression
Modelling
Poly-aluminium chloride
Turbidity

ABSTRACT

Coagulants are used in drinking water treatment plants to increase the size of particles and to help make them bigger and more able to settle at the later stages of the process. Poly-aluminium Chloride (PACL) was used in this study to evaluate its coagulation effectivity in different conditions. Three sets of experiments were done to determine the relationship between some raw water characteristics, including raw turbidity level, pH, and the temperature with optimum doses of PACL, in order to form a mathematical equation that could predict the removal effectivity. The experiments were performed under different seasonal circumstances. Four levels of turbidity were studied, 10, 50, 100, 150 NTU, with six different PACL doses from 5 to 35 mg/L. The results were used to build up a gene expression model (GEP). The GEP model gave very good results with a correlation coefficient equals to (0.91), and a root mean square error of 0.046.

1. Introduction

The term surface water refers to bodies of water, including estuaries, streams, rivers, and lakes. Rivers and lakes are the most common source of drinking water. Due to the rapid economic development as a result of population growth, the scarcity of water resources has been a serious issue for several decades; it has become an urgent issue in the formulation of sustainable development policies [1]. Water is usually unsafe to be used without treatment. The conventional treatment processes typically include the coagulation-flocculation step, followed by sedimentation and gravity filtration. These complex processes are constantly monitored and modified with the aim of providing consumers with high-quality drinking water at the lowest cost [2]. A variety of impurities are found in natural water, and the majority are colloids. Most of these collides have a negative charge, and their colloidal dispersions are stabilized due to electrostatic repulsion, which prevents particle aggregation and overcomes van der Waals forces. Such colloids cannot be removed by normal precipitation processes. The presence of colloids and natural organic matter (NOM) in raw water can cause many problems for the plants and consumers, including increasing the difficulty and cost of processing as well as many health problems [3]. In order to overcome the diffusion forces, accelerate agglomeration, and increase the effectiveness of sedimentation and filtration,

---

*Corresponding author. Tel: +963992151741
E-mail: rureta89@yahoo.com
DOI: 10.22104/AET.2021.5303.1433

substances called coagulants are used to coagulate water impurities. The coagulation ends with the formation of flocs separated from the medium aqueous, and the cohesion unit of the impurities is destroyed when electrolytes are added. Coagulant agents, such as alum, are usually added during the initial stage of the coagulation process to enhance the removal of suspended solids, including colloidal particles and natural organic matter (NOM) [4]. Recently, polymerized forms of coagulants such as poly-aluminum chloride (PACL) have been used increasingly in Europe, Japan, and North America for drinking water treatment due to their availability, thus reducing cost. Such products claim to be more advantageous over conventional coagulants because of their higher removal of particulate and organic matters as well as natural advantages of lower alkalinity consumption and lesser production of sludge [5]. PACL is produced by adding a certain amount of sodium hydroxide to aluminum solutions in the presence of chloride or nitrate ions. Usually, an OH/Al ratio between 2.4 - 2.7 is used. The predominant species in the commercially produced PACL are primarily $Al_{13}\ O_4\ (OH)_{24}$ followed by $Al(H_2\ 0)_6^3$ and colloidal $Al(OH)_3$. Very few results have been presented in the literature studies regarding the use of this coagulant [6]. Coagulation with different coagulants has been studied in the literature, but the coagulation behavior of poly aluminum chloride (PACL), especially the coagulation process at various coagulant dosages and pH values and turbidity of raw water, has not been investigated and is not well understood [7]. In a study done by Wei et al. (2015), coagulation mechanisms of (PACL) were studied. The optimal final pH and dosages for PACL were obtained between 7 to 8. Both charge neutralization and sweep coagulation could achieve high efficiency under the alkaline condition ranging from the final pH 7.0 to 10.0. The study found that both the charge neutralization and sweep coagulation zones of PACL were broader in the ranges of coagulant dosage and pH than those of alum [8]. In another study, the use of PACL was tested for initial turbidity 10 to 30 NTU; the optimal coagulant dose is 8 milligrams per liter. The alkalinity of raw water 100 mg/L of calcium carbonate. The turbidity removal efficiency was 83.02%, and the cost was 0.096 baht per cubic meter. The initial turbidity of the raw water was 30-50 NTU. A dose of 6mg/L was needed, alkaline 100mg/L of calcium carbonate. The efficiency of turbidity removal was 92.16%, and the cost was 0.137 baht per cubic meter [9]. A recent study in 2019 investigated turbidity removal by PACL; different initial turbidities between 20 to 300 NTU were studied, and the optimal dose was 5 ppm [10]. In recent years, artificial intelligence techniques have been used to model drinking water treatment processes, and research on controlling coagulant dosage has been implemented for several decades [11]. One of the most promising tools is gene expression programming, which is still young in the drinking water field. It could provide a nonlinear equation that is accurate and easy to use for different types of customers [12]. This study investigated coagulating raw water turbidity with PACL as coagulant at various pH and temperature values. The relationship between the optimal PACL dosages and final pH values was explored; the coagulation of PACL was also developed and compared with those of alum at different coagulant dosages. At the end of the study, a model using Gene expression was obtained that could be used for determining the percentage of turbidity removal when PACL is used in a Qusayr plant or any other plant that uses the same methods. The resulted model could be used at any other conventional plant that uses PACL as a coagulant. The model is presented as an easy-to-use mathematical equation, in contrast with other AI methods that need to get the code or the application made by the researcher.

2. Materials and methods

2.1. Data used

The research took place in a water purification plant in Qusayr, Homs, Syria. The raw water from the Orontes River enters the plant. The plant is a conventional plant, with four circular sedimentation tanks and 20 sand filters. The turbidity levels of the raw water defer from day to day and from month to month during the year. The raw water turbidity varied between 6 and 60 during the study and is described in Figure 1.

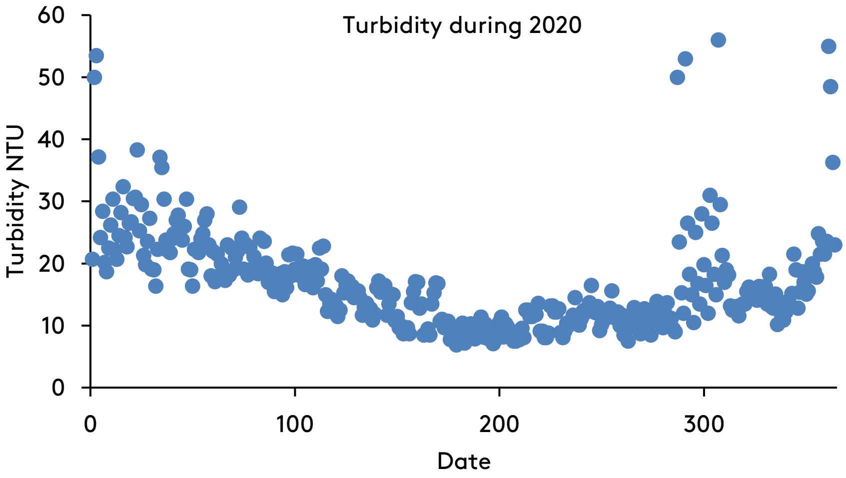

Fig.1. Raw turbidity variation in the plant during the study.

2.2. Jar Test Experiments

A series of chemical experiments were conducted using jar test experiments. The study focused on four aspects: the difference in the raw water turbidity, the effect of pH, and the temperature on the suitable PACL doses. The following steps describe how each experiment was conducted:

1- First, a Jar Test was performed for each initial turbidity level (10, 50,100,150 NTU ± 3). The PACL dose was changed between 5 and 35 mg/L. When needed, synthetic samples were made using kaolinite clay $AL_2Si_2\,O_5(OH)_4$. At this part of the experiments, the pH value was about 7 ± 0.3 and T = 20 ±2℃ .

2- Rapid mixing was carried out at 150 rpm for two minutes. The slow mixing was done at a speed of 40 rpm for 10 minutes, then the sedimentation stage for 60 minutes.

3-The second set of experiments evaluated the effect of the pH differences on each dose. The turbidity values were (10, 50, 100, and 150 ± 3), and the dose varied from 5 to 35. The pH values studied were 6, 8, and 9, adding the values from stage 1 with pH 7. Then, step 2 was repeated.

4-The temperature in the study area was deferred from one month to another during the year. Therefore, a set of experiments were studied for determining the optimal dose in each specific temperature.

This collection of experiments was done at different times of the year. The temperature of the raw water was (14, 25±1), the initial turbidity were (10, 50,100,150 NTU ± 3), and the dose changed between 5 to 35 mg/L with a step of 5 mg/L. Then, step 2 was repeated. Table 1 summarizes the steps of the experiments.

Table 1. The details of the three sets of experiments.

	Turbidity	**PACL dose**	**pH**	**T**
Set 1	10 , 50 , 100, 150	0, 5, 10, 15, 20, 25, 30, 35	7 ± 0.4	20 ±2
Set 2			6,7,8,9	20 ±2
Set 3			7.5 ± 0.4	14,20,25 ± 1
	Rapid Mix	Slow Mix	Sedimentation time	
	150	30	60	

2.3. Gene Expression

Gene expression is one of the artificial intelligence models. GEP is a type of genetic algorithm proposed in 2001 by Ferreira et al. Gene expression's basic search technique is a genetic algorithm. The GEP can solve problems in different fields with high performance. It presents a solution in the form of a tree structure. The first step of the GEP operation is the fitness function Determination, which can be determined as Equation 1:

$$F_i = \sum_{i=1}^{C_i} \left(S - \left|C_{i,j} - T_j\right|\right) \qquad (1) \qquad (1)$$

Hence, S is the range of selection, $C_{i,j}$ is the value returned by the individual chromosome i for fitness case j (out of Ct fitness cases), and T_j is the target value for fitness case j. The gene expression programming method is presented in the flowchart shown in Figure 2. First, the initial population is created. Then, the chromosomes are expressed, then excluded; after that, the fitness is evaluated. The individual is then selected according to their fitness. This process goes in a repetition loop several times until a good solution is found. The datasets were divided into training and validation sets; they were developed via the GeneXpro software to generate the models. Several runs and iterations were done to achieve the best fitness.

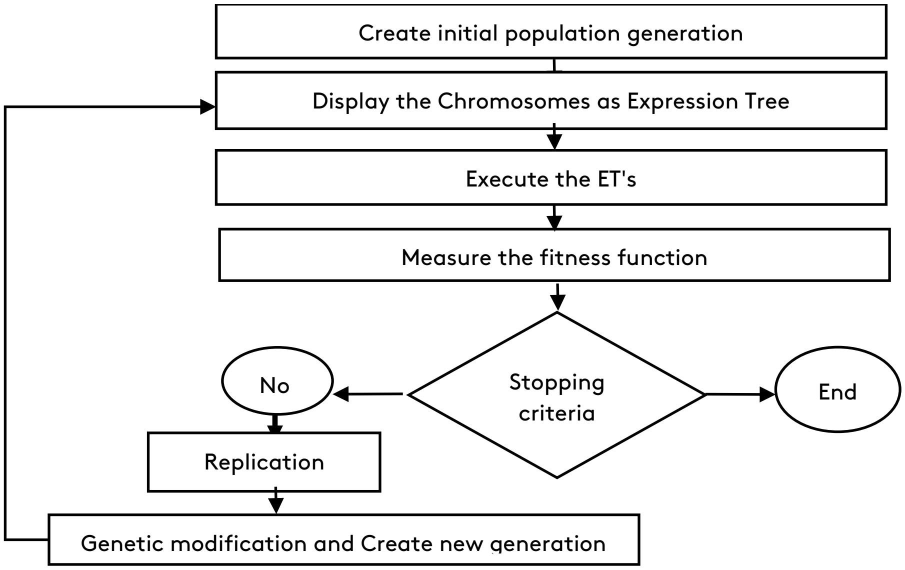

Fig. 2. Flowchart of Gene Expression Programming.

3. Results and discussion

*3.1. Experimental works*1-The first set of experiments include turbidity levels at 10, 50,100, and 150 NTU ± 3, PACL doses between 5 and 35, pH set constant at 7 ∓ 0.2 , and temperature at 20 $\mp$ 02. The same doses of alum were used for the comparting. The results are presented in Figure 3. The coagulant dosage is one of the most important factors in evaluating the performance of the coagulation process. With the best coagulation performance, the coagulation at the optimal coagulant dosage reduces the amount of coagulant used in water treatment. The optimization of the coagulant dosage is particularly important for the charge neutralization coagulation process because the dosage range in this process is narrow. From Figure 4, it is obvious that PACL has more affectivity than alum in regard to turbidity removal at the four turbidity levels. At the raw turbidity 10 NTU, 5 mg/L of PACL was enough to achieve 60% removal with 4 NTU turbidity residual, when a dose of 10 mg/L of alum was needed to give the same results. For the turbidity 50 NTU, 10 mg/L of PACL was needed to achieve a 93.8% removal efficiency, with turbidity residual of about 5 NTU. The optimal dose for the initial turbidity 100,150 NTU were 15, 20 mg/L, respectively. Figure 4 shows a contouring map for the turbidity removal percentage at different PACL doses and turbidity levels.

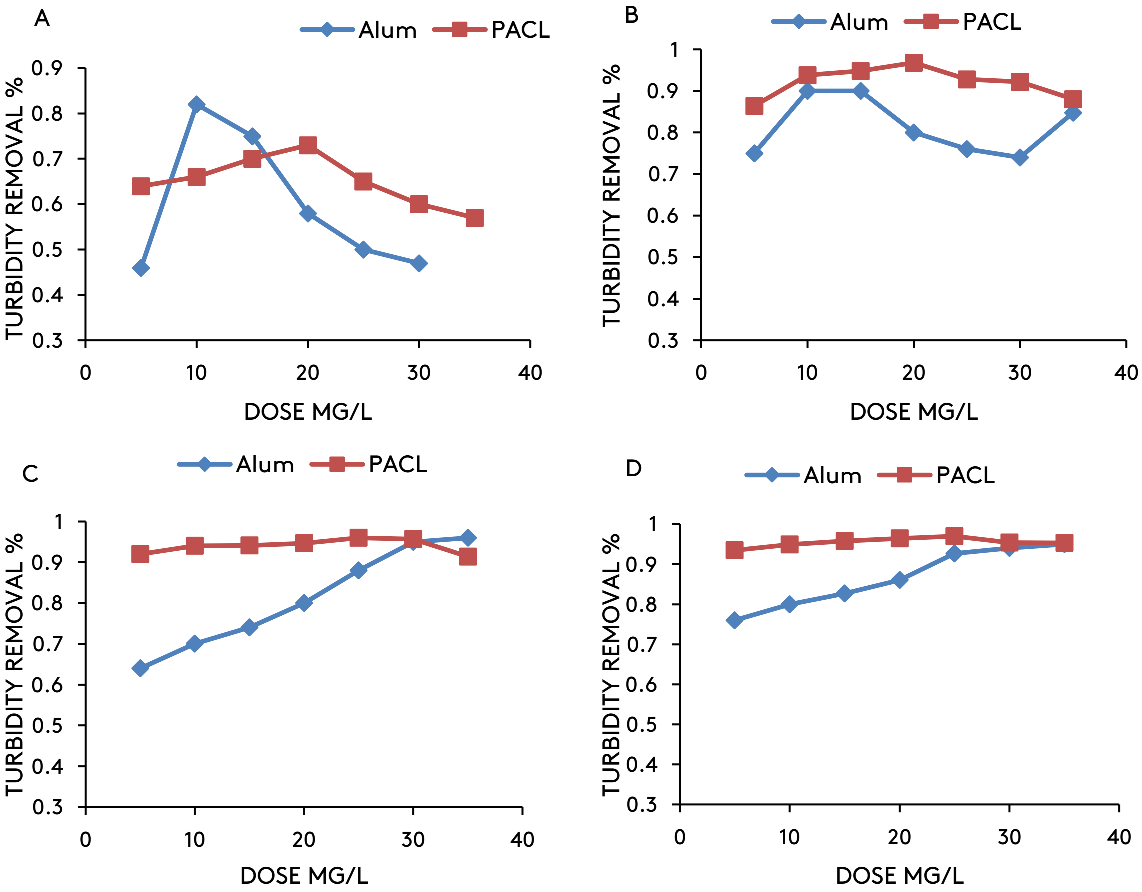

Fig. 3. Turbidity removal percentage at different alum, PACL doses for (a) Initial turbidity 10 NTU, (b) Initial turbidity 50 NTU, (c) Initial turbidity 100 NTU, and (d) Initial turbidity 150 NTU.

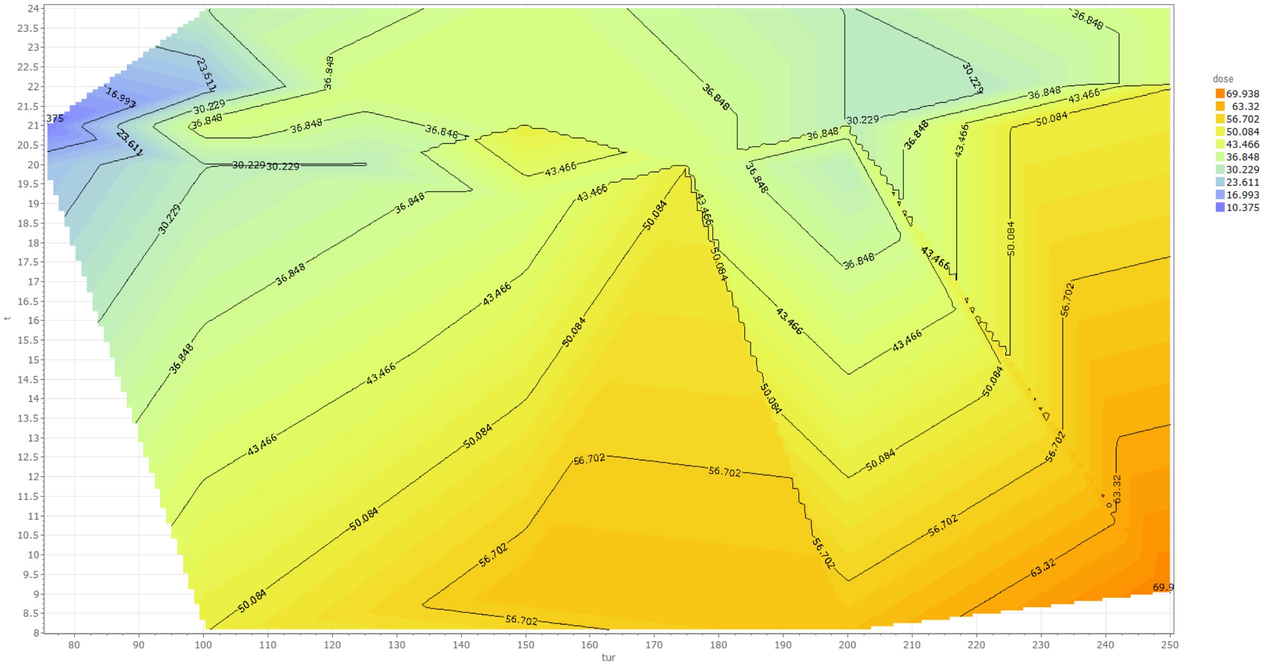

Fig. 4. Contouring map for Turbidity removal percentage at different PACL doses and turbidity levels.

2- In the second set of experiments, turbidity were set at (١0, ٥0, 100, and 150) ∓ 2 NTU, pH was between 6-8.5 for six different doses of PACL (5, 10, 15, 20, 25, 30, and 35) mg/L, and temperature at 20∓ 2. The results are shown in Figure 5. When the pH was raised to eight, the removal effectivity increased from 86 % to 90% for the initial turbidity 10NTU and from 64% to 68 % for the initial turbidity 50 NTU. When the pH value of the treated water was reduced, the coagulation efficiency became less, and the residual turbidity was higher. The effectivity of the PACL kept improving until the pH value of 8, after this value, the effectivity decreased for all of the doses at the four turbidity levels of the raw water. The turbidity removal percentage dropped from 93 % to 89 %, when the pH changed from 8 to 9 for the dose 5 mg/L at the raw water turbidity 100 NTU. And from 95% to 93%, 96 % to 93 %, 96 % to 93%, 97% to 95%, 97% to 94%, and 92% to 90% for the doses 10, 15, 20, 25, 30, and 35 mg/L, respectively.

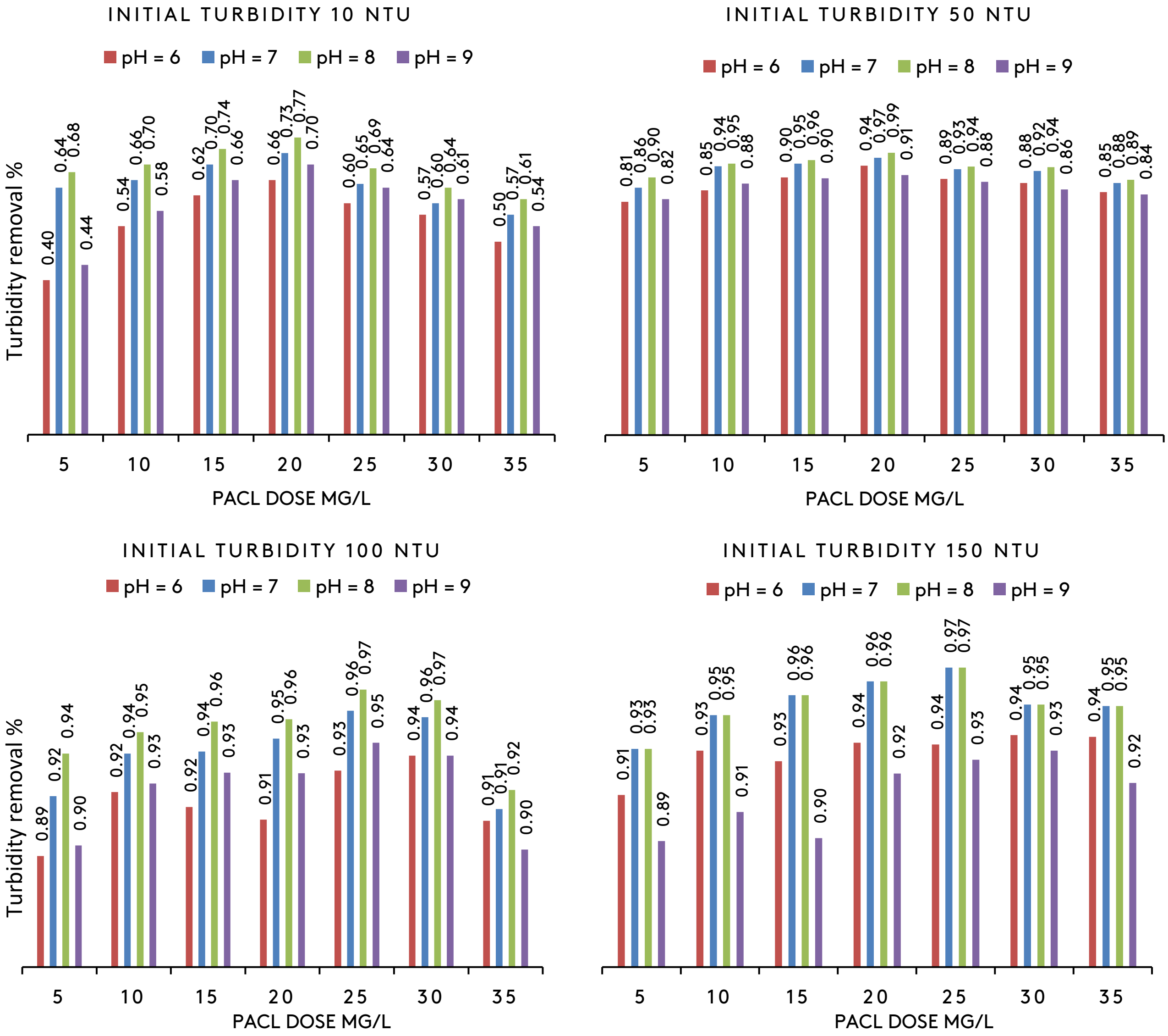

Fig. 5. Residual turbidity for different pH levels at different initial turbidity.

3- The third set of experiments studied the effect of temperature. In this set of experiments, the pH was as constant as possible at 7 ± 0.3. The experiments were done at different times of the year, and the temperatures were 14, 20, and 25 ±1, with study doses of (5, 10, 15, 20, 25, 30, 35) mg/L. The results are described in Table 2. The PACL was not affected in a significant way by the temperature differences; the removal percentages of turbidity for initial turbidity 50 and dose 15 mg/L were 91, 95, and 98 % when the temperature changed as the following 14, 20, 24 ℃ respectively, while when raw water turbidity were 150 NTU and for the same PACL dose the removal changed as 93, 95, 98 % respectively for the same temperatures. Higher turbidity levels were less affected by the temperature differences than the low levels.

Table 2. Turbidity removal at different temperatures.

Initial Turbidity	Dose mg/L	Temperature		
		14	**20**	**25**
10	5	4.6	3.6	1.6
	10	4.2	3.4	1.4
	15	3.6	3	1
	20	3.5	2.7	0.7
	25	4.2	3.5	1.5
	30	5.6	4	2
	35	6	4.3	2.3
50	5	9.3	6.8	3.3
	10	5.6	3.1	0.6
	15	4.6	2.6	1.1
	20	5.1	1.6	0.1
	25	8.1	3.6	2.1
	30	9.4	3.9	2.4
	35	7.5	6	4.5
100	5	12.3	8	5
	10	11.33	6	3
	15	9.5	5.9	2.5
	20	9.1	5.3	2.3
	25	7	4	2
	30	9.5	4.3	3.3
	35	11.6	8.6	5.6
150	5	16.1	10.1	7.1
	10	14.3	8	3.5
	15	10.5	7	3.5
	20	12	6.1	4.6
	25	9	4.5	0.2
	30	9.4	6.9	2.9
	35	12.5	7	2

3.2. Gene Expression model

In order to make a model that could predict the residual turbidity after treating the water using PACL, the results of the previous experiments were used as a dataset; the Genexpro program was used to build a model, where 70% of the data were used to train the model and 30% for validation. Fitness function used to evaluate the models was the root mean squared error (RMSE). There are a variety of parameters related to the Gene expression models; the most important ones are the number of chromosomes, mutation, and functions. This study used two different numbers of chromosomes and mutation rates. The values of the final GEP parameters are shown in Table 3. The performance of various models was evaluated using the following statistical indices:

- Root Mean Squared Error (RMSE).

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(T_i - O_i)^2}{n}} \quad (2)$$

- Correlation Coefficient: R.

$$R=\frac{\sum_{i=1}^{n}(P_{obs}-\overline{P_{obs}})(P_{pre}-\overline{P_{pre}})}{\sqrt{\sum_{i=1}^{n}(P_{obs}-\overline{P_{obs}})^2\times\sum_{i=1}^{n}(P_{pre}-\overline{P_{pre}})^2}} \qquad (3)$$

- Mean Absolute Error

$$MAE=\frac{1}{n}\sum_{i=1}^{n}|x_i-x| \qquad (4)$$

A different number of chromosomes was studied. The results are listed in Figure 6 (A and B). From these two figures, it can be seen that the minimum RMSE and the maximum R were obtained when the number of chromosomes was 600. Therefore, 600 chromosomes were selected as the optimal number of chromosomes for the problem under investigation. The GEP model that gave the best results is presented in Figure 7, and the results for the best model are shown in Table 3. The GEP model that gave the best results is presented in Figure 6. This figure shows the relationships between the input and output data by using the GEP algorithm as a mathematical formula. This feature makes the results more applicable in comparison to other AI methods.

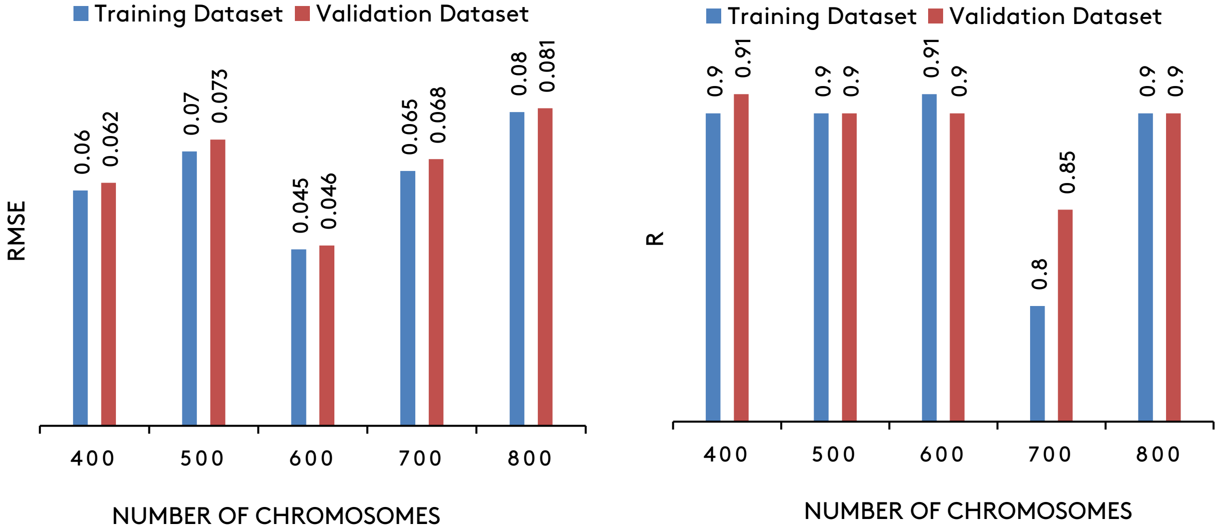

Fig. 6. Comparison between (A- RMSE) and (B- R) of training and validation for a different numbers of chromosomes.

Table 3. Values of GEP Control Parameters.

General		Genetic operators	
Function set	+, -, ×, /, √, exp, ln, log, 10^		
Number of Chromosomes	600	Mutation	0.00138
		Inverse Rate	0.00546
Function Fitness	RMSE	IS transposition rate	
Linking Function	+	RIS	0.00546
Mutation	0.00138		
Head size	8	One point recombination rate	0.00277
Tail size	3		
Numerical constants		two point recombination rate	0.00277
Type of Data	Floating Type	Gene recombination rate	0.00277
Maximum Complexity	10		
Constants per gene	10	Uniform recombination	0.00755

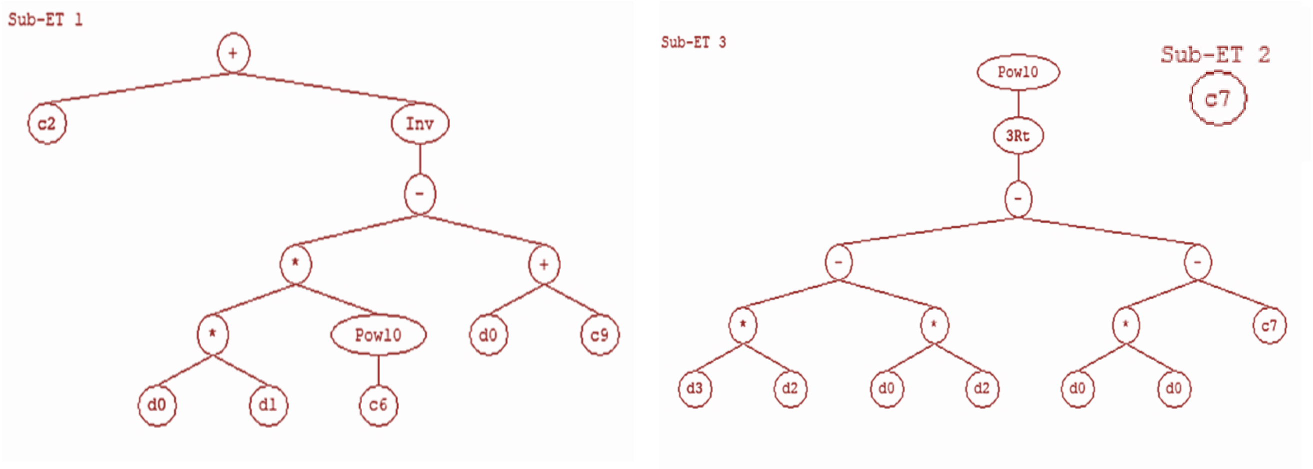

Fig. 7. Sub gene-expression tree for the achieved model.

The constants and the parameter used in the equation are listed in Table 4.

Table 4. Values of parameters used in the equations.

d_0	d_1	d_2	d_3
Turbidity	PACL DOSE	pH	T
G1C1			-4.77167546617023
G1C9			-5.58711172826319
G2C6			-2.27798908352916
G2C7			5.70761781604358

Table 5. The accuracy table of selected gene-expression model.

The number of the Model		RMSE	R	MAE
1	Train	0.045	0.91	0.038

Figure 8 represents a comparing chart of the real data with the results gained from the GEP model. As it is shown, the GEP can be used as a trusted modeling method in this case.

The variable importance of the inputs is done, and the results are presented in Figure 9. As shown, the most affecting parameters were PACL dose, initial turbidity, and temperature. And this is because the turbidity affects the needed dose, and they are much related parameters. The third most affected parameter was the temperature, as it affected the rate of reactions and the resulting species during coagulation.

4. Conclusions

Many articles have studied different types of coagulants like alum, ferric chloride, etc., but there are no enough articles that studied PACL to determine its optimal dose and removing effectivity. Many aspects of the treatment process were studied: the effect of different doses, pH differences, and the effect of the temperature on the turbidity removal. Four levels of raw turbidity were studied, 10, 50, 100, and150 NTU, and the optimal dose for each level was determined. Then, the effect of both pH and temperature was determined. The results of the experiments were used to build a GEP model. GEP was a good and reliable method to determine the effectivity of PACL for removing turbidity under different circumstances with an accuracy of R =0.91. The most affecting parameters on the GEP model were the PACL dose and initial turbidity, with the importance of 34.02 % and 33.87%.

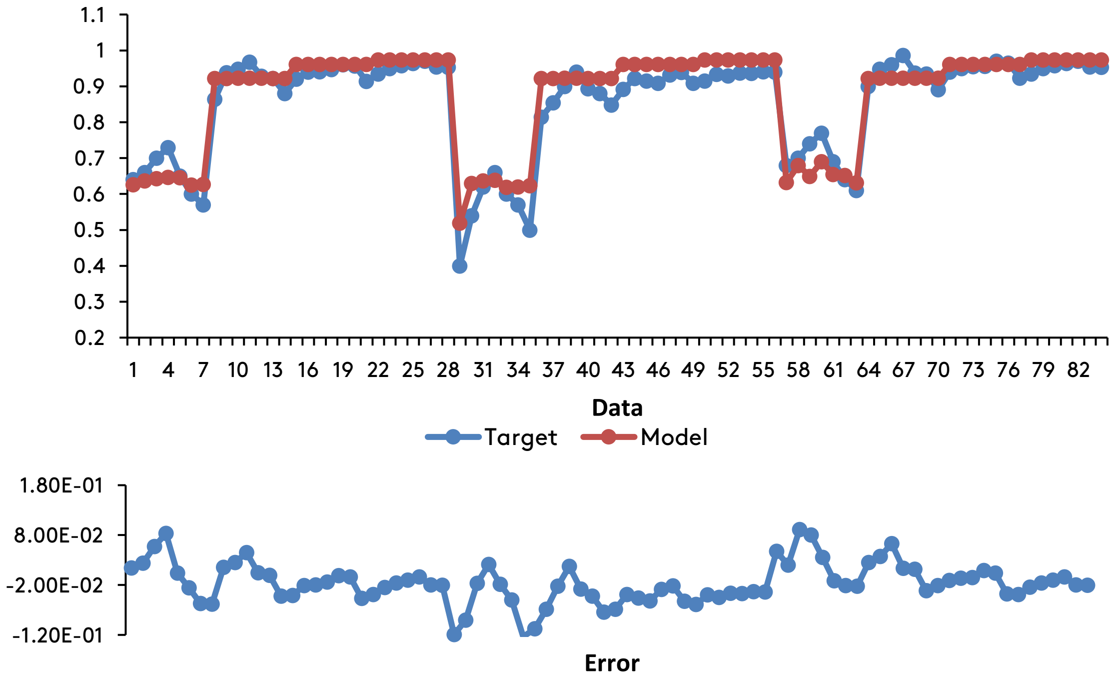

Fig. 8. Comparing the observed data with the results GEP models.

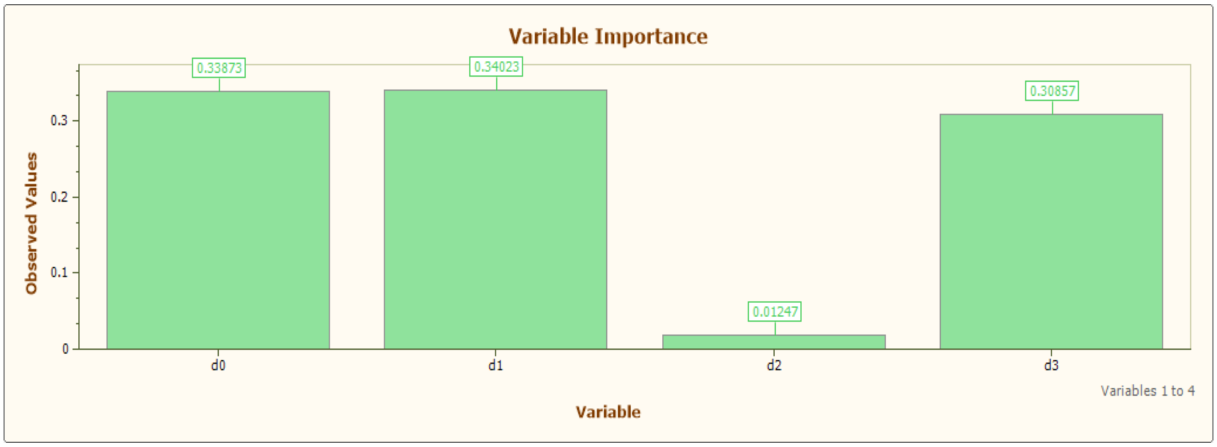

Fig. 9. Variable importance of the inputs.

References

[1] Haghiri, S., Daghighi, A., Moharramzadeh, S.(2018). "Optimum coagulant forecasting by modeling jar test experiments using ANNs". *Drinking water engineering and science, 11*(1), 1-8.

[2] Onyelowe, K. C., Jalal, F. E., Onyia, M. E., Onuoha, I. C., Alaneme, G. U. (2021). Application of gene expression programming to evaluate strength characteristics of hydrated-lime-activated rice husk ash-treated expansive soil. *Applied computational intelligence and soft computing, 1-17*.

[3] ABIDEEN, M. B. Z. (2016). "Optimization of coagulation process in water treatment plant using statistical approach". Ph.D. dissertation, Universiti Teknologi Malaysia.

[4] Zemmouri, H., Drouiche, M., Sayeh, A., Lounici, H., Mameri, N. (2012). Coagulation flocculation test of Keddara's water dam using chitosan and sulfate aluminum. *Procedia engineering*, *33*, 254-260.

[5] Alshikh, O. (2007). Parameters affecting coagulation/flocculation of drinking water under cold temperatures. University of Windsor (thesis), Canada.

[6] Zouboulis, A., Traskas, G., Samaras, P. (2008). Comparison of efficiency between poly-aluminum chloride and aluminum sulphate coagulants during full-scale experiments in a drinking water treatment plant. *Separation science and technology*, *43*(6), 1507-1519.

[7] Liu, W. (2016). Enhancement of coagulant dosing control in water and wastewater treatment processes. Ph.D. dissertation, Norwegian University of Life Sciences.

[8] Wei, N., Zhang, Z., Liu, D., Wu, Y., Wang, J., Wang, Q. (2015). “Coagulation behavior of polyaluminum chloride: Effects of pH and coagulant dosage”. *Chinese journal of chemical Engineering, 23(6),* 1041-1046.

[9] Tantipalakul, Y., Palawatwichai, K., Detchakan, T., & Khaisan, J. (2018). “The study of optimal coagulants for water treatment process of Metropolitan Waterworks Authority”. *Burapha science journal*, *23(1),* 207-220.

[10] Almatin, E., Gholipour, A. (2019). Estimating of optimal dose of PACL for turbidity removing from water. *arXiv e-print, arXiv:1904.06421*.

[11] Al-Baidhani, J. H., Alameedee, M. A. (2017). Optimal alum dosage prediction required to treat effluent water turbidity using artificial neural network. *International journal of current engineering and technology*, *7*(4), 1552-1558.

[12] Alsaeed, R., Alaji, B., Ebrahim, M. (2021). Predicting turbidity and Aluminum in drinking water treatment plants using Hybrid Network (GA-ANN) and GEP. *Drinking water engineering and science discussions*, 1-17.