Scientia Iranica F (2021) 28(3), 1871–1883

Sharif University of Technology
Scientia Iranica
Transactions F: Nanotechnology
http://scientiairanica.sharif.edu

An integrated model for predicting the size of silver nanoparticles in montmorillonite/chitosan bionanocomposites: A hybrid of data envelopment analysis and genetic programming approach

E. Sarvestani and G.R. Khayati*

Department of Materials Science and Engineering, Shahid Bahonar University of Kerman, Kerman, P.O. Box 76135-133, Iran.

Received 20 December 2018; received in revised form 27 August 2020; accepted 18 October 2020

KEYWORDS
Silver nanoparticle;
Gene expression programming;
Montmorillonite;
Chitosan;
Bionanocomposites;
Sensitivity analysis.

Abstract. Unique chemical and physical properties of silver nanoparticles (AgNPs) contribute to the broader scope of their applications in different fields including medical utilities. Considering the high dependency of AgNPs properties on their size, this study employed Gene Expression Programming (GEP) to develop a quantitative model for estimating the size of AgNPs in montmorillonite/chitosan bionanocomposites prepared by the chemical approach. Generalization capabilities, fault tolerance, noise tolerance, high parallelism, nonlinearity, and significant information processing characteristics are the main advantages of GEP. Accordingly, the practical parameters including reaction temperature, $AgNO_3$ concentration, weight of montmorillonite in aqueous $AgNO_3$/chitosan solution (W_{MMT}), and percentage of chitosan were the input parameters selected through GEP modeling. The accuracy of the proposed models was investigated based on statistical indicators including *Mean Absolute Percentage Error (MAPE)*, *Root Relative Squared Error (RRSE)*, *Root Mean Square Error (RMSE)*, and correlation coefficient (R^2). Finally, the best model was selected by $R^2 = 0.987$, $RMSE = 0.100$, $RRSE = 0.146$, and $MAPE = 0.221$. The sensitivity analysis confirmed that the percentage of chitosan, concentration of $AgNO_3$, W_{MMT}, and reaction temperature were the most effective parameters for the size of AgNPs.

1. Introduction

Nanotechnology has been the main focus of a number of researches in materials science due its administrated effect on chemical and physical properties [1]. In this respect, preparation of noble metal nanoparticles including Ag, Au, Pd, and Pt has been a popular subject of different researches [2]. Generally, there are three approaches to preparing nanoparticles, namely chemical, physical, and biological approaches. The unique potential of chemical methods for preparing finer particles and the narrower distribution size of products distinguish this technique from others. Application of metal nanoparticles plays a key role in enhancing the performance of nanocomposites with extensive utilities in industrials. Silver nanoparticles

*. *Corresponding author.*
E-mail addresses: Elham.sarvestani1993@gmail.com (E. Sarvestani); khayatireza@gmail.com (G.R. Khayati)

doi: 10.24200/sci.2020.52510.2748

(AgNPs) are widely applied in the fields of catalysis, optics, dentistry, photography, mirrors, clothing, food industries, and electronics [3]. Advanced applications of AgNPs contribute to the evolution of different techniques with the preference of the lower size as the main criterion. Accordingly, it would be favorable to prepare AgNPs with a smaller size and lower possibility of aggregation [1].

Another challenge of preparing nanoparticles is their strong tendency for agglomeration [4]. Different approaches have been employed to reduce the agglomeration of nanoparticles such as the applications of layered silicates of montmorillonite (MMT) as the matrix. MMT has a two-to-one layered network including one octahedral aluminum layer between two layers of silicon tetrahedral. The thickness of this layer is about 1 nm and the lateral dimension of this layer ranges between 100–1000 nm [5]. Furthermore, MMT is characterized by the ion exchange properties, swelling, and intercalation. The lamellar structure of MMT raises the possibility of its delamination to the elemental sheets without difficulty and its applications as the substrate for the synthesis of nanoparticles through electroless technique. For instance, the structure of MMT is used as an appropriate substrate for the preparation of biomaterial nanoparticles through the adsorption of cationic ions and anchoring of the transition metal complexes [6].

Chitosan (Cts), a natural polymer, enjoys unique characteristics including non-toxicity, solubility in aqueous medium, biodegradability, excellent biocompatibility, multi-functional groups, artificial skin, and bone substitutes [7]. Cts is added to be intercalated in W_{MMT} through the mechanism of hydrogen bonding processes and cationic exchange, and it produces Bio Nano-Composites (BNCs) with functional and structural properties [8]. BNCs comprise organic/inorganic nano-sized filler and polymeric matrix. To the best of the author's knowledge, metal/clay/polymer compounds, or BNCs, owing to their excellent properties, have drawn considerable academic attention [7,9–11].

Gene Expression Programming (GEP) is used to determine the performance and properties of engineering materials as a mathematical expression. With respect to the classical techniques such as artificially neural network and regression analysis, GEP provides a capable environment. Compared to the classical approach, there is not any predefined function through the GEP modeling [12,13].

In summary, the lower particle size of AgNPs produces synergetic effects on the biological properties of the prepared bio-nanocomposites. However, the agglomeration follows a general trend, while it is in the range of the lower sizes of AgNPs. Accordingly, the main objectives of the present study were: (i) establishing the relationship among the practical variables, including the concentration of $AgNO_3$, chitosan percentage, d-spacing of clay layers, and reaction temperature, as the input values and the size of AgNPs as the output value; (ii) using GEP for modeling; (iii) determining the effect of each variable quantitatively on the size of AgNPs; and (iv) utilizing MMT as the substrate to avoid the agglomeration.

2. Data collection and analysis

2.1. Data collection

Correct selection of practical variables affecting the output is the first step in GEP modeling. Shabanzadeh et al. [14] investigated the size of AgNPs prepared in montmorillonite/chitosan bionanocomposites. They showed that the size of AgNPs was strongly dependent on the concentration of $AgNO_3$, chitosan percentage, d-spacing of clay layers, and reaction temperature. In this study, 30 datasets from the literature were taken into account [14], as shown in Table 1. These datasets were randomly divided into two categories: the training (20 data) and testing (10 data) datasets [15].

2.2. Data analysis

Outlier data complicates the clear understanding of the dependency of AgNPs on the practical parameters. Accordingly, determination and removal of outlier can positively affect the higher accuracy of the GEP modeling [16]. Box plot is a data analysis procedure used for determination of the outlier during the practical dataset. Figure 1 shows the box plot of four practical parameters during the synthesis of AgNPs. As observed, the median of Cts and $MAgNO_3$ data sets are in the box center, indicating the symmetric distribution of these data. However, the box plot of T and W_{MMT} are skewed to the bottom and top, respectively. In other words, the box shifted to the bottom by a whisker and most data are small with some minor exceptionally large ones. The average of T and W_{MMT} is higher than their median and is close to that of T (40) and W_{MMT} (1.99). Of note, there is not any outlier between the selected datasets for all practical parameters. Table 2 summarizes the statistical explanation of dataset in this study.

Another key parameter for the GEP modeling is the independency of practical parameters on each

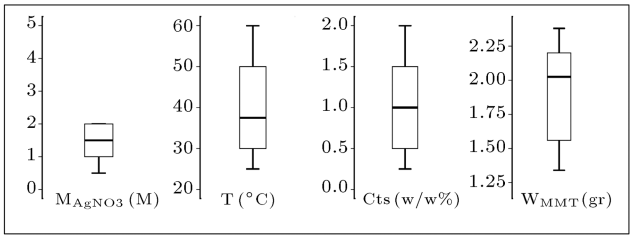


Figure 1. Boxplot for M_{AgNO3}, T, Cts, and W_{MMT} variables during the green synthesis of AgNPs.

Table 1. Illustration of the practical data used for preparing AgNPs through green synthesis.

No.	Factors				Response
	M_{AgNO3} (M)	T (°C)	Cts (w/w%)	W_{MMT} (gr)	AgNPs size (nm)
1	0.5	25	0.25	1.34	3.51
2	0.5	30	0.25	1.35	3.54
3	0.5	40	0.25	1.39	3.58
4	0.5	50	0.25	1.40	3.62
5	0.5	60	0.25	1.41	3.63
6	1.0	25	0.5	1.55	3.85
7	1.0	30	0.5	1.56	3.87
8	1.0	40	0.5	1.60	3.94
9	1.0	60	0.5	1.69	4.08
10	1.5	30	1	1.93	4.14
11	1.5	35	1	1.97	4.15
12	1.5	40	1	2.11	4.18
13	1.5	50	1	2.16	4.21
14	2.0	25	1	2.29	4.25
15	2.0	30	1.5	2.08	4.52
16	2.0	60	1.5	2.33	5.14
17	5.0	30	2	2.17	4.87
18	5.0	40	2	2.25	5.32
19	5.0	50	2	2.35	5.59
20	5.0	60	2	2.38	5.632
21	1.5	60	1	2.26	4.22
22	2	35	1.5	2.12	4.60
23	2	40	1.5	2.18	4.64
24	2	50	1.5	2.29	4.69
25	5	35	2	2.20	4.99
26	0.5	35	0.25	1.37	3.55
27	1	35	0.5	1.58	3.89
28	1	50	0.5	1.64	3.97
29	1.5	25	1	1.88	4.11
30	5	25	2	2.14	4.76

other. The present study employed Bivariate Correlation Analysis (BCA) to find the relationships among the input parameters. To this end, BCA as a comprehensive analysis was first carried out to find the highly correlated pairs. Low negative/positive correlation between the pairs is necessary in GEP modeling. In other words, significant interdependency of the variables may intensify the effect of input data in the modeling, referred to as multi-dependency [17]. Table 3 shows the correlation coefficients of the experimental variables. What is significantly notable in Table 3 is the considerable dependency of M_{AgNO_3} and T, M_{AgNO_3} and Cts, M_{AgNO_3} and W_{MMT}, T and Cts, T and W_{MMT}, and Cts and W_{MMT}.

Principal Component Analysis (PCA) is an appropriate approach to the investigation of the multi-

Table 2. Statistical distribution of M_{AgNO3}, T, Cts, and W_{MMT} used in the development of Gene Expression Programming (GEP) models.

Variable	Min.	Max.	Mean	Median	Std.
M_{AgNO3}	0.50	5.00	2.00	1.50	1.61
T	25.00	60.00	40.00	37.50	12.11
Cts	0.25	2.00	1.04	1.00	0.65
W_{MMT}	1.34	2.38	1.99	21.02	0.36

Table 3. Correlation coefficients of M_{AgNO3}, T, Cts, and W_{MMT} used for preparing AgNPs through green synthesis.

Variable	M_{AgNO3}	T	Cts	W_{MMT}
M_{AgNO3}	1	0	0.921	0.723
T	0	1	0.033	0.193
Cts	0.921	0.033	1	0.891
W_{MMT}	0.723	0.193	0.891	1

dependency of experimental variables. This technique is capable to eliminate the correlation among variables from a multi-dimensional space to a low-dimensional space. The variables in the new space are uncorrelated [18–20]. Since the presence of a higher correlation coefficient among the variables does not necessarily show the multi-dependency of the dataset, first, Kaiser Mayer Olkin (KMO) [21], as a criterion, must be estimated through Eq. (1) to ensure the possibility of PCA:

$$KMO = \frac{\sum\sum r_{ij}^2}{\sum\sum r_{ij}^2 + \sum\sum a_{ij}^2}. \quad (1)$$

In case KMO is lower than 0.7, the dependency is unreal and thus, it should be neglected [16]. Since the KMO factor in this study is equal to 0.516, no PCA is required to eliminate the multi-dependency. In other words, the high dependency of some variables in Table 3 is related to the nature of BCA. Accordingly, the selected practical parameters are independent and, hence, appropriate for further analysis via GEP.

3. Explanation of predictive equation

3.1. GEP

Ferreira (2011) proposed a novel approach, taking into account the population based on an evolutionary strategy to compensate the disadvantageous of Genetic Algorithm (GA) and Genetic Programing (GP) called GEP [22,23]. It provides an advanced strategy for the construction of inferring function among the input parameters with the capability of forecasting the output with acceptable accuracy and minimum error estimation. GEP enjoys several advantages: (a) It raises the possibility of finding the best mathematical equation, and (b) It is able to check the entire search space without trapping in the local minimum. In other words, GEP enjoys the advantages of both GA and GP that make modeling more feasible [24,25]. Like GA, the individuals of GEP have fixed length string characters (chromosomes) that can be connected to each other by Karva language. This language enables the GEP to convert the coded programs to chromosomes. Generally, the coded solution can be expressed as the Expression Tree (ET) structure. GEP comprises a chromosome with different genes consisting of two sections, i.e., tail and head (Figure 2).

While the head in Figure 2 determines the functions and terminals, the tail determines just the terminals. Of note, although the terminals are input variables and constant values, their functions include the basic mathematical element including ($*$, /, $+$, $-$), Boolean operators (and, or, nor, not), and nonlinear

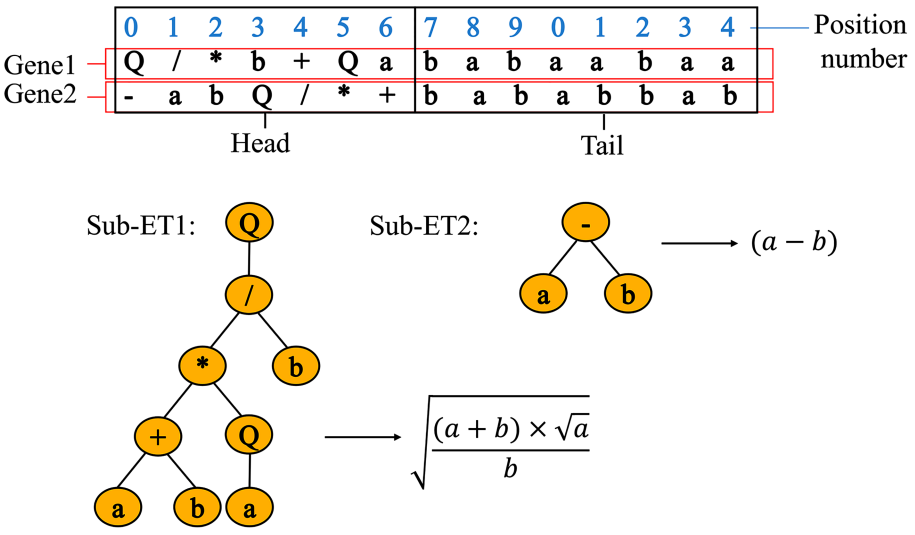


Figure 2. Illustration of Karva language and expression tree for a chromosome with two genes.

functions (*Exp, 3Rt, sqrt, arctan, tan, sin, -cos*). The head length is predefined by the GEP user and the tail length depends on h. Moreover, $n_{\max}$ (the maximum number of argument) is correlated to the function, as shown in Eq. (2):

$$t = h(n_{\max} - 1) + 1. \tag{2}$$

The chromosome of GEP is randomly made, which is illustrated by a subset of sub-ET correlated to the others via a linking function ($*$, /, +, $-$) [26–29]. As shown in Figure 3, GEP, first, generates the chromosomes randomly and then, creates the first population known as Karva language (K-Expression). Then, each depicted chromosome is representative of ET and mathematical equation, respectively. Upon considering the fitness function as the criterion, the cost of individuals can be evaluated. The best chromosomes are selected for the next generation. If $f(S_i(t))$ is defined as the fitness of individual S_i in the population at t generation, the probability of individual S_i would be copied to the next population generation as a result of any one reproduction operation (Eq. (3)) [30] by taking into account the fitness-proportionate selection:

$$\frac{f(S_i(t))}{\sum_{j=1}^{M} f(S_i(t))}. \tag{3}$$

After estimating the probability value of each chromosome, the selection process continues, as shown in Figure 3. In this step, the individuals are selected as the candidates for selection by taking into account the fitness value as an important criterion. In the next phase, genetic operator (transposition, recombination, and mutation inversion) uses chromosomes to generate better individuals for future generations. The same process continues until finding an appropriate solution [31–34]. In the following, a brief summary of the genetic operators is presented:

I. *Mutation.* Mutation is mainly characterized by intrinsic modification power that can be wielded anywhere across the length of the chromosome. While the tail mutation converts a terminal to another through mutation operator, the head mutation provides the possibility of changes in functions or terminals [22,30,35].

II *Inversion.* This operator is active through the head of chromosomes and enables the user to reverse the length of 1 to 3 in the head.

III. *Transposition.* This operator has three types: (i) insertion sequence transposition that is responsible for transposing a fragment to the head of its own or other genes; (ii) root insertion sequence transposition that is responsible for transposing a segment with a function in the first position of the root of the genes; and (iii) gene transposition that is responsible for transposing all genes of the first chromosomes [22,30,35].

3.2. Training GEP as an intelligent approach to modeling and obtaining the mathematical equation

The main objective of this study was to present a mathematical equation for GEP model (GEP-1 to GEP-10) to predict the size of AgNPs. Table 4 summarizes the derived equations of GEP-1 to GEP-10 with better performance. The number of genes is supposed to be eight. The multiplication and addition randomly selected as correlating function. In the testing and training phases of GEP models, 10 different

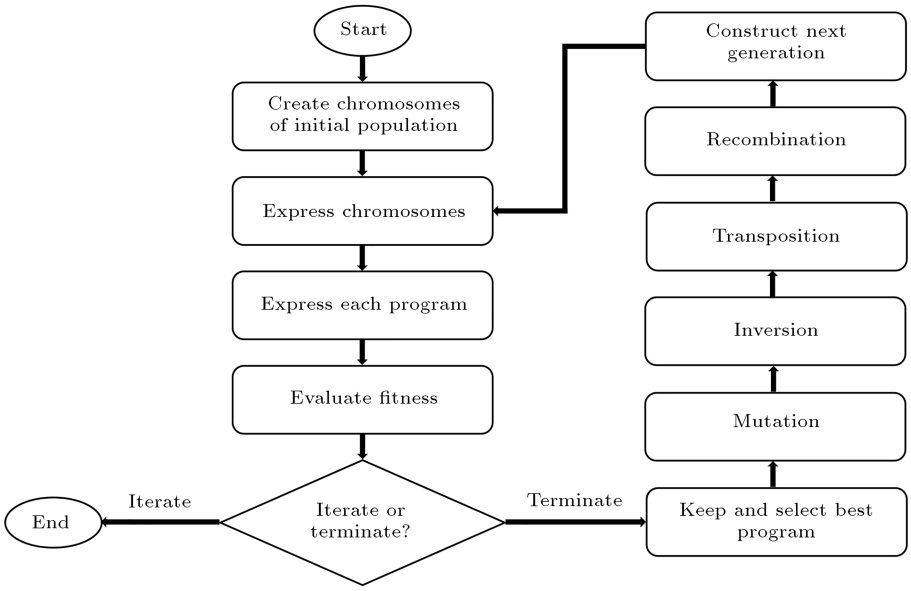


Figure 3. Flowchart of the Gene Expression Programming (GEP) strategy.

Table 4. Representation of the inferring equations for different Gene Expression Programming (GEP) models.

Models	Predicted equation
GEP-1	$Y = ((((W_{MMT}{}^{*}M_{AgNO3})^{*}T)^{*}(0.034^{*}0.034)) + ((0.034^{*}W_{MMT})^{*}$ $(M_{AgNO3} + Cts))) + ((((Cts - W_{MMT})^{*}(W_{MMT}{}^{*}0.074))^{*}((W_{MMT}{}^{*}W_{MMT}) -$ $M_{AgNO3})) - T) + ((M_{AgNO3}/(((-2.33) + Cts)^{*}T)) - (((-2.33) - W_{MMT}) - T))$
GEP-2	$Y = ((((M_{AgNO3} - Cts)^{*}(0.006^{*}T)) + ((0.006/Cts) + W_{MMT}))/W_{MMT})^{*}$ $((((W_{MMT} + W_{MMT})^{*}Cts)/((-0.446) - W_{MMT})) + ((W_{MMT} + M_{AgNO3})/(W_{MMT}{}^{*}$ $(-0.446))))^{*}((((0.587 - Cts)^{*}(0.587/T))/((M_{AgNO3}/0.587) - W_{MMT})) + 0.587)^{*}$ $((W_{MMT} + W_{MMT})/(((W_{MMT}/T) - (-0.220)) - (M_{AgNO3} + W_{MMT})))$
GEP-3	$Y = ((((((((((-0.126) - Cts) + W_{MMT})^{*}M_{AgNO3})^{*}$ $M_{AgNO3})))\hat{}(1/3))))\hat{}(1/3)) + (1.0/((((W_{MMT} + (-0.830)) -$ $tanh(M_{AgNO3})) + (W_{MMT} + Cts)/(((T))\hat{}(1/3)))))) + (d(4) + ((((-0.834)^{*}Cts) -$ $tanh((-0.834)))^{*}(tanh(M_{AgNO3}) - W_{MMT})))$
GEP-4	$Y = ((Atan((T^{*}(-0.432))) + Atan(W_{MMT}))^{*}((M^{*}_{AgNO3}(-0.432)) +$ $(-0.432))) + exp(Atan(((exp(M_{AgNO3}) + (Cts/W_{MMT})) + (0.150/Cts)))) +$ $(((-0.007) + M_{AgNO3})/(((T + W_{MMT}) * (-0.007)) * (W_{MMT} + T)))$
GEP-5	$Y = (Log((\sin(\sin(\sin(M_{AgNO3})))) + (Log(((M_{AgNO3}{}^{*}$ $(-1.020)) + T))/Log(10))))/Log(10)) + \sin((((W_{MMT} * Cts) -$ $((-0.304)/W_{MMT}))^{*}(((-0.304) + W_{MMT})^{*}\sin((-0.304))))) +$ $(T + ((\sin(W_{MMT}) + (W_{MMT} - T)) + ((0.891^{*}Cts) + 0.891)))$
GEP-6	$Y = ((\cos((W_{MMT} + W_{MMT}))^{2}) - ((Cts^{*}W_{MMT}) - (0.006 + 0.006))) +$ $((W^{2}_{MMT}) + (((M_{AgNO3}/0.321)^{*}(Cts - M_{AgNO3}))/(T^{*}Cts))) + (\cos((\cos(W_{MMT}) -$ $((-0.530)^{2}))) - ((Cts^{2})^{*}((-0.530) + (-0.530))))$ $+(((\cos(W_{MMT})/Cts) + (Cts/T))^{*}((W_{MMT} - M_{AgNO3})^{*}(0.200^{*}W_{MMT})))$
GEP-7	$Y = ((((((((T))\hat{}(1/3)) + (1.176 - W_{MMT})) * ((1.176 + 1.176) + (Cts^{*}$ $M_{AgNO3})))))\hat{}(1/3)) + ((((((0.607 + Cts) + (M_{AgNO3} + M_{AgNO3})) + ((0.607 +$ $M_{AgNO3}) + (0.607 + M_{AgNO3})))))\hat{}(1/3)) + ((((M_{AgNO3} -$ $W_{MMT}) * Cts) * M_{AgNO3})/(((((-0.682)))\hat{}(1/3))^{*}(T + M_{AgNO3})))$
GEP-8	$Y = ((0.520^{*}Cts) + ((((0.520 - M_{AgNO3}{}^{*}Cts))))\hat{}(1/3))) +$ $((((((((-1.110)))\hat{}(1/3))^{3}) + (((M_{AgNO3}))\hat{}(1/3)))^{*}((((((T))\hat{}(1/3))))\hat{}(1/3)))^{3}) +$ $((((((Cts))\hat{}(1/3))/T) + (0.333^{*}Cts)) + ((((0.333))\hat{}(1/3)) + (Cts + W_{MMT)})) + (0.391 -$ $(((M^{3}_{AgNO3})/(W^{*}_{MMT}T))/((Cts^{3}) - M_{AgNO3})))$
GEP-9	$Y = ((1.0/((((Cts - W_{MMT})^{*}(T/(-0.333))) + ((-0.333) - T)))) +$ $W_{MMT}) + (1.0/((W_{MMT} + (((W_{MMT} + (-0.970))/$ $((-0.970) + Cts))^{*}(W_{MMT}/T))))) + ((tanh((W_{MMT}{}^{*}M_{AgNO3})) -$ $((Cts^{*}(-0.317))/(M_{AgNO3} + (-0.317)))) - (-0.317)) + (((-0.185)/((T/M_{AgNO3}) - W_{MMT})) -$ $((M_{AgNO3}{}^{*}(-0.185)) - (T - T)))$
GEP-10	$Y = ((((((Cts^{*}W_{MMT})/(((0.369))\hat{}(1/3))) +$ $(M_{AgNO3}{}^{*}Ln(T)))))\hat{}(1/3)) + Ln(Ln((T + (((1.685^{2})^{2}) -$ $(M^{2}_{AgNO3})))))) + ((((((((Cts^{2}) - (W_{MMT} - (-0.209))) + (M_{AgNO3}/Cts))))\hat{}(1/3))))\hat{}(1/3))$

Table 5. The most appropriate characteristics of Gene Expression Programming (GEP) algorithm in the present study.

GEP parameters	GEP-1 to GEP-10
Number of runs	100
Parameters	Value
Used functions	$*, /, +, -, 3Rt, Inv, Tanh, Arctan, Cos, Sin, Log, Ln, Exp, X2$
Number of chromosomes	34 (GEP-2), 32 (GEP-8), 30 for other models
Size of head	8
Number of genes	4 (GEP-2), (GEP-8) and (GEP-9), 3 for other models
Linking function	Addition, Multiplication
Fitness function error type	Root relative squared error (RRSE)
Constant per gene	1
Mutation rate	0.0044
Inversion rate	0.1
One-point recombination rate	0.2
Two-point recombination rate	0.2
Gene recombination rate	0.1
Gene transportation rate	0.1

sub-ETs including M_{AgNO3}, T, Cts, and W_{MMT} as the input parameters as well as AgNPs as the output value were used. Of the 30 collected datasets, 20 trails were selected as the training dataset and 10 trails were chosen for the testing phase. The fitness function, f_i, of an individual program can be estimated through Eq. (4):

$$f_i = \frac{1}{N}\sum_{i=1}^{i=N}\frac{|t_i - p_i|}{t_i} \times 100. \tag{4}$$

In the case of precision $|C_{(ij)} - T_j| \leq 0.01$, the precision would be $= 0$ and $i = f_{\max} = MC_t$. However, for $f_{\max} = 1000$, $M = 100$. This approach provides the possibility of finding an optimum condition [36,37].

Then, the sets of terminals (T) and functions (F) for generating the chromosomes were suggested. While the former contains M_{AgNO3}, T, Cts, and W_{MMT}, the latter includes four basic arithmetic operators (−, +, / , *). In addition, some basic mathematical functions (−, +, /, *, *3Rt, Inv, Tanh, Arctan, Cos, Sin, Log, Ln, Exp*, and X^2) were utilized. At the same time, the performance of the proposed GEP models was monitored during the testing and training phases. Selection of the chromosome tree is another step of GEP. First, this study considered the single gene and two lengths of heads as the first approximation and then, each length increased in the course of every single run. Table 5 presents the abbreviations of the investigated characteristics during the GEP modeling in the current study. Of note, the size of AgNPs is a function of M_{AgNO3}, T, Cts, and W_{MMT} [36].

3.3. Evaluating the training performance of the model

To evaluate the approximation performance of GEP equations, the real and predicted outputs were compared in terms of statistical performance indices including Mean Absolute Percentage Error (MAPE), Root Relative Squared Error (RRSE), Root Mean Square Error (RMSE), and correlation coefficient (R^2) (Eqs. (5)–(8)) [38,39].

$$MAPE = \frac{1}{N}\sum_{i=1}^{i=N}\frac{|t_i - p_i|}{t_i} \times 100, \tag{5}$$

$$RRSE = \sqrt{\frac{\sum_{i=1}^{i=N}(t_i - p_i)^2}{\sum_{i=1}^{i=N}\left(t_i - (\frac{1}{N})\sum_{i=1}^{i=N}t_i\right)^2}}, \tag{6}$$

$$RMSE = \sqrt{\sum_{i=1}^{i=N}\frac{(t_i - p_i)^2}{N}}, \tag{7}$$

Table 6. Different statistical criteria for the 10 most appropriated Gene Expression Programming (GEP) models.

Models	Performance indices	R^2	RMSE	RRSE	MAPE
GEP-1	Training set	0.992	0.046	0.086	0.989
	Testing set	0.954	0.153	0.222	2.134
GEP-2	Training set	0.995	0.039	0.073	0.628
	Testing set	0.945	0.160	0.233	2.215
GEP-3	Training set	0.985	0.067	0.125	1.238
	Testing set	0.987	0.100	0.146	0.221
GEP-4	Training set	0.980	0.076	0.141	1.212
	Testing set	0.953	0.168	0.244	2.092
GEP-5	Training set	0.995	0.040	0.074	0.564
	Testing set	0.985	0.102	0.149	1.274
GEP-6	Training set	0.991	0.050	0.094	1.034
	Testing set	0.969	0.129	0.188	2.089
GEP-7	Training set	0.990	0.052	0.097	0.916
	Testing set	0.963	0.143	0.208	1.827
GEP-8	Training set	0.988	0.058	0.107	1.095
	Testing set	0.982	0.113	0.164	1.554
GEP-9	Training set	0.987	0.062	0.114	1.090
	Testing set	0.973	0.219	0.319	3.039
GEP-10	Training set	0.986	0.064	0.119	0.957
	Testing set	0.980	0.129	0.188	1.714

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (t_i - p_i)^2}{\sum_{i=1}^{N} {p_i}^2}. \tag{8}$$

Table 6 summarizes the values for these criteria in the GEP models.

3.4. Sensitivity analysis

The significance of practical parameters and their relative effects on the size of AgNPs size were investigated through sensitivity analysis (Eq. (9)). Accordingly, the value of each parameter oscillates between maximum and minimum, while the other practical variables are supposed to be constant in their mean values. Then, the output changes are measured and employed as the threshold for comparing the effect of each parameter. Typically, to determine the effect of W_{MMT} changes on the size of AgNPs, all practical variables (except W_{MMT}) are supposed to be fixed in their mean values, while W_{MMT} varies [39]. To carry out sensitivity analysis, a step-by-step methodology was applied to the most appropriate GEP model by changing the input parameters one at a time at constant rates [15,40].

$$S_i(\%) = \frac{1}{N} \sum_{j=1}^{N} \left(\frac{\%\ change\ in\ output}{\%\ change\ in\ input} \right)_j. \tag{9}$$

4. Results and discussion

In the present study, R^2 (the closer to 1, the better), $RMSE$ (the closer to 0, the better), $RRSE$ (the closer to 0, the better), and $MAPE$ (the closer to 0, the better) criteria were utilized to evaluate the accuracy of the proposed GEP models. According to Table 6, $0.945 < R^2 < 0.995$, $0.073 < RMSE < 0.319$; $0.039 < RRSE < 0.219$; and $0.221 < MAPE < 3.039$. Figure 4 shows the changes of the selected statistical indices for the 10 most appropriate models. Accordingly, the highest values for R^2 are attributed to the GEP-5 model equal to 0.995 and 0.985 for the training and testing phases, respectively. In case of

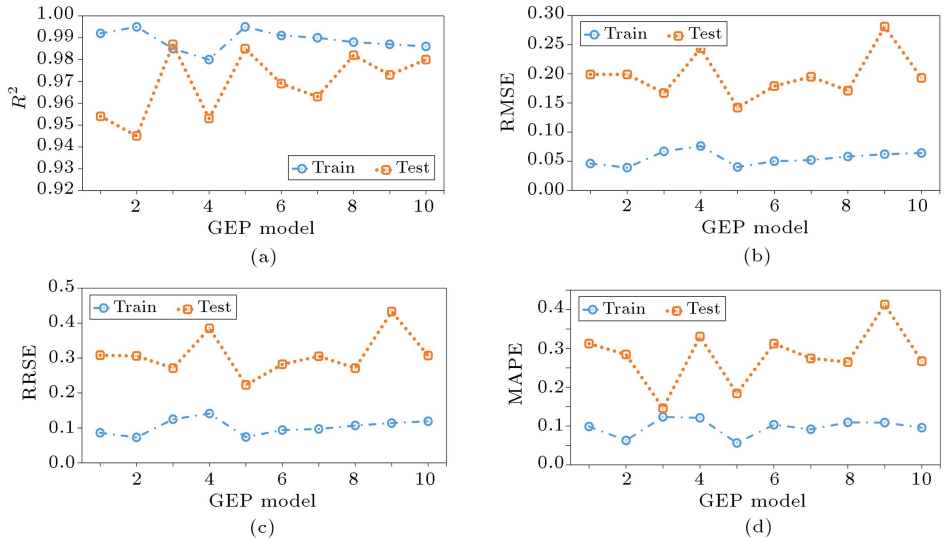


Figure 4. Comparison of validation criteria for Gene Expression Programming (GEP) models: (a) R^2, (b) Root Mean Square Error (RMSE), (c) Root Relative Squared Error (RRSE), and (d) Mean Absolute Percentage Error (MAPE) for the GEP-3 structure.

using the RMSE as a metric for GEP validation, the lower value for the RMSE belongs to GEP-3 due to its closer value to zero, equal to 0.067 and 0.100 in the training and testing phase, respectively. Similar to the GEP model, based on the RRSE, it can be found that the GEP-3 with RRSE equal to 0.125 for training and 0.146 for testing phases had the best performance; in terms of MAPE as a criterion, GEP-3 was selected as the most appropriate model. Consequently, using one statistical approach cannot simultaneously compensate for the lower values of errors (RMSE, RRSE, and MAPE) and higher values of R^2. For all the statistical indices employed in selecting the most appropriate GEP models, the fitness value was defined through Eq. (10):

$$Fitness - value = 1/R^2 + RMSE + RRSE + MAPE. \quad (10)$$

In this approach, the most appropriate model has the lowest fitness value (Eq. (10)). Figure 5 shows the changes in the fitness value for different GEP models. Accordingly, the best structure belongs to GEP-3 with three genes, eight heads, and 30 chromosomes where the genes are linked to each other using the addition function.

Figure 6 shows the ET of GEP-3. The inferring equation between the practical parameters and size of the Ag particles in the GEP-3 model in Table 6 reveals the complexity of the practical parameters of the output.

Figure 7 compares the practical and predicted values of the size of AgNPs, where the experiment number is defined by the number in the circle and hexagon, and the size of Ag nanoparticles is determined

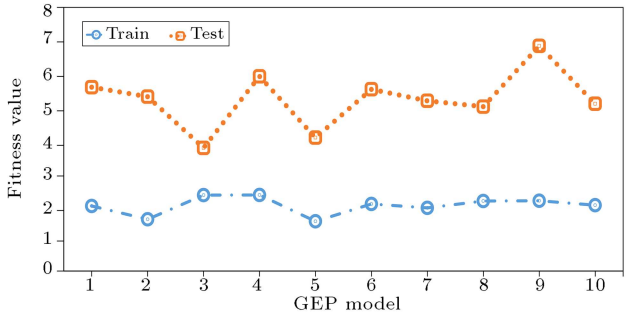


Figure 5. The changes of fitness value for 10 Gene Expression Programming (GEP) models.

by the number of the vertical axis. The values of the statistical indices as well as fitness value are added to Figure 7. According to this figure, GEP-3 provides acceptable accuracy during the testing and training sets, respectively.

Figure 8 shows the effect of experimental parameters on the outputs in 3D diagram. Figure 8(a) illustrates the size variation of AgNPs as a function of M_{AgNO3} and Cts and confirms the significant effect of these parameters. In addition to M_{AgNO3} and T as the selected variables, Figure 8(b) shows the relatively greater effect of M_{AgNO3} with respect to T on the output. The yellow and red points in Figure 8(c) indicate that the effect of M_{AgNO3} is more notable than W_{MMT}. Furthermore, the output value changes directly with an increase in both M_{AgNO3} and W_{MMT}. Analysis of Figure 8(d) reveals that the effect of Cts on the output is more considerable than T. In addition, W_{MMT} factor can affect the size of AgNPs more considerably than T (Figure 8(e)). At the same time, increasing W_{MMT} and Cts could increase the size of the output (Figure 8(f)).

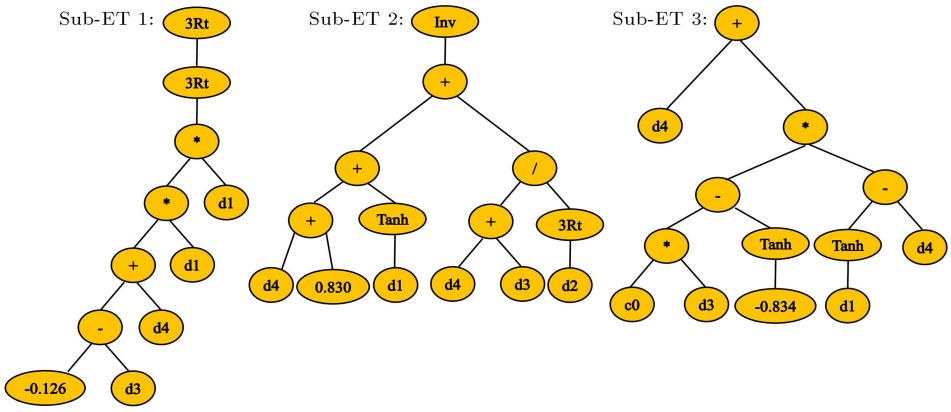
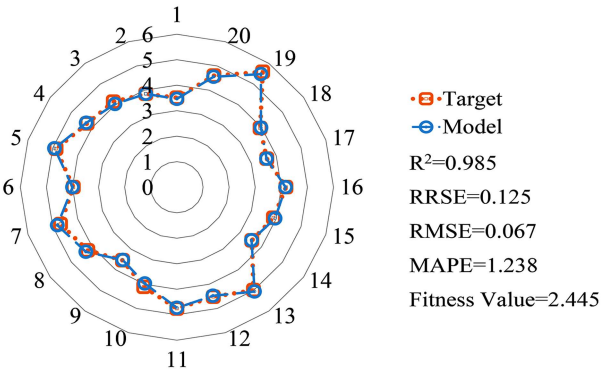
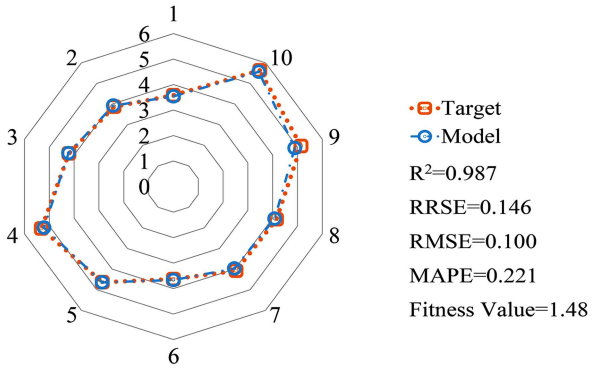


Figure 6. Expression tree of GEP-3 with three genes used for predicting the average particle size of AgNPs.



(a) GEP-training data



(b) GEP-testing data

Figure 7. Comparison of the predicted values obtained from GEP-3 and actual data: (a) Training datasets and (b) testing datasets.

Generally, this study aimed to predict the size of AgNPs with emphasis on the effect of four variables including Cts, M_{AgNO3}, W_{MMT}, and T. As expected, the size of AgNPs increased upon increasing the $AgNO_3$ concentration, reaction temperature, and value of MMT. However, the effect of reaction temperature on the size enhancement of AgNPs is quite low; moreover, upon increasing the chitosan percentages, the size of AgNPs would decrease. However, with simultaneous increase in the chitosan with other parameters, the size of the AgNPs would increase. According to Figure 9, the important factors in the size of AgNPs include Cts, M_{AgNO3}, W_{MMT}, and T of the reaction.

5. Conclusion

The present study employed the Gene Expression Programming (GEP) to predict the size of AgNPs and consider the effect of M_{AgNO3}, T, Cts, and W_{MMT} as the input variables. The results confirmed that the GEP structure consisting of three genes, eight heads, and 30 chromosomes exhibited better behavior during the training algorithm with an acceptable performance in the GEP molding of the prepared bionanocomposites. As a result, GEP-3 model with $R^2 = 0.977$, $RRSE = 0.146$, $RMSE = 0.100$, and $MAPE = 0.221$ was recommended for estimating the size of AgNPs prepared with different practical parameters. The results confirmed that upon increasing the value of Ag^+ in terms of both concentration and temperature, the size of AgNPs increased. However, at a higher chitosan percentage, the size of the outputs decreased. Further, GEP models could provide a suitable approach to predicting the size of AgNPs.

Nomenclature

M_{AgNO3}	Concentration of silver nitrate (M)
T	Reaction temperature (°C)
Cts	Percentage of chitosan (w/w%)
MMT	montmorillonite
R^2	Correlation coefficient
RMSE	Root Mean Square Error
RRSE	Root Relative Squared Error
MAPE	Mean Absolute Percentage Error
$C_{(i,j)}$	Value returned by the individual chromosome i for fitness case j

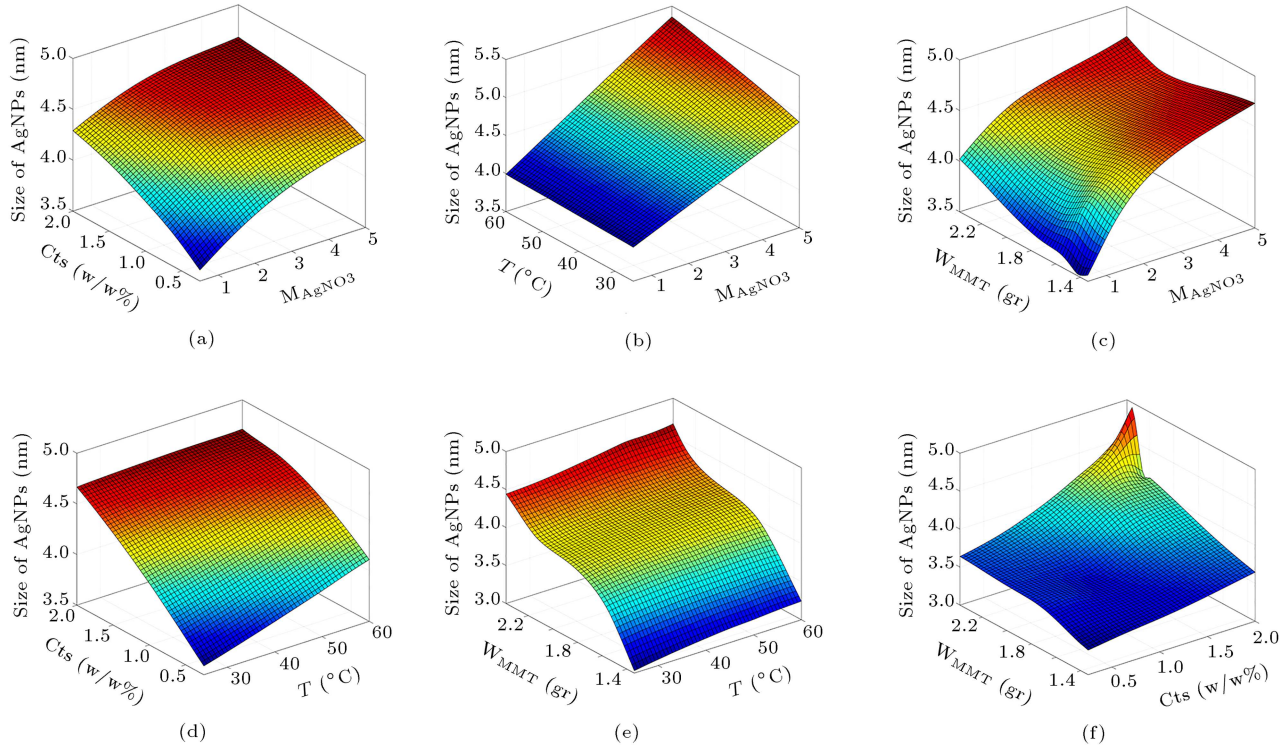


Figure 8. Three-dimensional surfaces plots: (a) The effects of M_{AgNO3} and Cts (w/w%), (b) the effects of M_{AgNO3} and T, (c) the effects of M_{AgNO3} and W_{MMT}, (d) the effects of Cts (w/w%) and T, (e) the effects of T and W_{MMT}, and (f) the effects of Cts (w/w%) and W_{MMT} on the size of AgNPs.

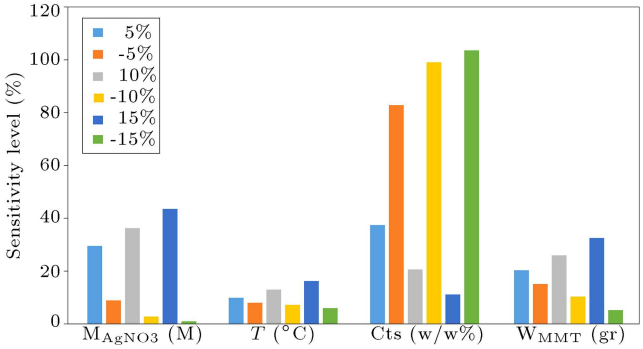


Figure 9. Sensitivity analysis of the practical parameter during the preparation of AgNPs.

S_i	Sensitivity level of an input parameter (%)
$N(=30)$	Number of datasets used for sensitivity test
t_i	The measured values for models
p_i	The predicted values for models

References

1. Shabanzadeh, P., Yusof, R., and Shameli, K. "Artificial neural network for modeling the size of silver nanoparticles' prepared in montmorillonite/starch bionanocomposites", *Journal of Industrial and Engineering Chemistry*, **24**, pp. 42–50 (2013).
2. Veisi, H., Azizi, S., and Mohammadi, P. "Green synthesis of the silver nanoparticles mediated by Thymbra spicata extract and its application as a heterogeneous and recyclable nanocatalyst for catalytic reduction of a variety of dyes in water", *Journal of Cleaner Production*, **170**, pp. 1536–1543 (2016).
3. Rai, M., Yadav, A., and Gade, A. "Silver nanoparticles as a new generation of antimicrobials", *Biotechnology Advances*, **27**(1), pp. 76–83 (2009).
4. Shameli, K., Ahmad, M.B., Al-Mulla, E.A.J., et al. "Green biosynthesis of silver nanoparticles using Callicarpa maingayi stem bark extraction", *Molecules*, **17**(7), pp. 8506–8517 (2012).
5. Shameli, K., Ahmad, M.B., Yunus, W.M.Z.W., et al. "Synth2esis of silver/montmorillonite nanocomposites using γ-irradiation", *International Journal of Nanomedicine*, **5**, p. 1067 (2010).
6. Makwana, D., Castaño, J., Somani, R.S., et al. "Characterization of Agar-CMC/Ag-MMT nanocomposite and evaluation of antibacterial and mechanical properties for packaging applications", *Arabian Journal of Chemistry*, **13**(1), pp. 3092–3099 (2018).
7. Lavorgna, M., Attianese, I., Buonocore, G.G., et al. "MMT-supported Ag nanoparticles for chitosan nanocomposites: structural properties and antibacterial activity", *Carbohydrate Polymers*, **102**, pp. 385–392 (2014).
8. Usman, M.S., Ibrahim, N.A., Shameli, K., et al. "Copper nanoparticles mediated by chitosan: synthesis and

characterization via chemical methods", *Molecules*, **17**(12), pp. 14928–14936 (2012).

9. Shameli, K., Ahmad, M.B., Zargar, M., et al. "Synthesis and characterization of silver/montmorillonite/chitosan bionanocomposites by chemical reduction method and their antibacterial activity", *International Journal of Nanomedicine*, **6**, p. 271 (2011).
10. Ahmad, M.B., Shameli, K., Tay, M.Y., et al. "Antibacterial effect of silver nanoparticles prepared in bipolymers at moderate temperature", *Research on Chemical Intermediates*, **40**(2), pp. 817–832 (2014).
11. Shameli, K., Ahmad, M.B., Yunus, W.M.Z.W., et al. "Green synthesis of silver/montmorillonite/chitosan bionanocomposites using the UV irradiation method and evaluation of antibacterial activity", *International Journal of Nanomedicine*, **5**, pp. 5:875–887 (2010).
12. Jafari, M.M. and Khayati, G.R. "Prediction of hydroxyapatite crystallite size prepared by sol-gel route: gene expression programming approach", *Journal of Sol-Gel Science and Technology*, **86**(1), pp. 112–125 (2018).
13. Ebrahimzade, H., Khayati, G.R., and Schaffie, M. "A novel predictive model for estimation of cobalt leaching from waste Li-ion batteries: Application of genetic programming for design", *Journal of Environmental Chemical Engineering*, **6**(4), pp. 3999–4007 (2018).
14. Shabanzadeh, P., Senu, N., Kamyar, S., et al. "Prediction of silver nanoparticles' diameter in montmorillonite/chitosan bionanocomposites by using artificial neural networks", *Research on Chemical Intermediates*, **41**(5), pp. 3275–3287 (2015).
15. Khayati, G.R. "Adaptive neuro-fuzzy inference system and neural network in predicting the size of monodisperse silica and process optimization via simulated annealing algorithm", *Journal of Ultrafine Grained and Nanostructured Materials*, **51**(1), pp. 43–52 (2018).
16. Faradonbeh, R.S. and Monjezi, M. "Prediction and minimization of blast-induced ground vibration using two robust meta-heuristic algorithms", *Engineering with Computers*, **33**(4), pp. 835–851 (2017).
17. Sun, M. "Prediction of viscosity of branched alkanes using gene expression programing", *Petroleum Science and Technology*, **36**(23), pp. 2049–2056 (2018).
18. Tiryaki, B. "Predicting intact rock strength for mechanical excavation using multivariate statistics, artificial neural networks, and regression trees", *Engineering Geology*, **99**(1–2), pp. 51–60 (2008).
19. Sayadi, A.R., Khalesi, M.R., and Borji, M.K. "A parametric cost model for mineral grinding mills", *Minerals Engineering*, **55**, pp. 96–102 (2014).
20. Sayadi, A.R., Lashgari, A., and Paraszczak, J.J. "Hard-rock LHD cost estimation using single and multiple regressions based on principal component analysis", *Tunnelling and Underground Space Technology*, **27**(1), pp. 133–141 (2012).
21. Kaiser, H.F. "An index of factorial simplicity", *Psychometrika*, **39**(1), pp. 31–36 (1974).
22. Al-Anni, R., Hou, J., Azzawi, H., et al. "New gene selection method using gene expression programing approach on microarray data sets", *International Conference on Computer and Information Science*, Springer, Cham (2018).
23. Dikmen, E. "Gene expression programming strategy for estimation performance of LiBr-H_2O absorption cooling system", *Neural Computing and Applications*, **26**(2), pp. 409–415 (2015).
24. Steeb, W.-H., *The Nonlinear Workbook: Chaos, Fractals, Cellular Automata, Neural Networks, Genetic Algorithms, Fuzzy Logic with C++, Java, Symbolicc++ and Reduce Programs*, World Scientific (2001).
25. Brownlee, J., *Clever Algorithms: Nature-Inspired Programming Recipes*, Jason Brownlee (2011).
26. Monjezi, M., Baghestani, M., Shirani Faradonbeh, R., et al. "Modification and prediction of blast-induced ground vibrations based on both empirical and computational techniques", *Engineering with Computers*, **32**(4), pp. 717–728 (2016).
27. Faradonbeh, R.S., Jahed Armaghani, D., Abd Majid, M.Z., et al. "Prediction of ground vibration due to quarry blasting based on gene expression programming: a new model for peak particle velocity prediction", *International Journal of Environmental Science and Technology*, **16**(6), pp. 1453–1464 (2016).
28. Khandelwal, M., Armaghani, D.J., Faradonbeh, R.S., et al. "A new model based on gene expression programming to estimate air flow in a single rock joint", *Environmental Earth Sciences*, **75**(9), p. 739 (2016). https://doi.org/10.1007/s12665-016-5524-6
29. Aval, S.B., Ketabdari, H., and Gharebaghi, S.A. "Estimating shear strength of short rectangular reinforced concrete columns using nonlinear regression and gene expression programming", *Structures*, Elsevier, **12**, p. 13–29 (2017).
30. Ferreira, C., *Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence*, Springer, **21** (2006).
31. Jafari, M.M. and Khayati, G.R. "Prediction of hydroxyapatite crystallite size prepared by sol-gel route: gene expression programming approach", *Journal of Sol-Gel Science and Technology*, **86**(1), pp. 112–125 (2018).
32. Zhong, J., Feng, L., and Ong, Y.S. "Gene expression programming: A survey", *IEEE Computational Intelligence Magazine*, **12**(3), pp. 54–72 (2017).
33. Faradonbeh, R.S., Monjezi, M., and Armaghani, D.J. "Genetic programing and non-linear multiple regression techniques to predict backbreak in blasting operation", *Engineering with Computers*, **32**(1), pp. 123–133 (2016).
34. Baykasoğlu, A., Güllüb, H., Çanakçi, H., et al. "Prediction of compressive and tensile strength of limestone via genetic programming", *Expert Systems with Applications*, **35**(1–2), pp. 111–123 (2008).
35. Ferreira, C. and Gepsoft, U. "What is gene expression programming" (2008).

36. Sarıdemir, M. "Effect of specimen size and shape on compressive strength of concrete containing fly ash: Application of genetic programming for design", *Materials & Design*, **56**, pp. 297–304 (2014).
37. Jafari, S. and Mahini, S.S. "Lightweight concrete design using gene expression programing", *Construction and Building Materials*, **139**, pp. 93–100 (2017).
38. Fallahpour, A., et al. "An integrated model for green supplier selection under fuzzy environment: application of data envelopment analysis and genetic programming approach", *Neural Computing and Applications*, **27**(3), pp. 707–725 (2016).
39. Khozani, Z.S., Bonakdari, S., and Ebtehaj, I. "An analysis of shear stress distribution in circular channels with sediment deposition based on Gene Expression Programming", *International Journal of Sediment Research*, **32**(4), pp. 575–584 (2017).
40. Hoseinian, F.S., Faradonbeh, R.S., Abdollahzadeh, A. et al. "Semi-autogenous mill power model development using gene expression programming", *Powder Technology*, **308**, pp. 61–69 (2017).

Biographies

Elham Sarvestani received his MSc degree in Materials Science Engineering from Shahid Bahonar University of Kerman (Iran) in 2020 and BSc degree in Materials Science Engineering from Shahid Bahonar University of Kerman (Iran) in 2017.

Gholam Reza Khayati is currently a faculty member at the Department of Materials Science and Engineering at Shahid Bahonar University of Kerman. He received his PhD in Materials Science from Shiraz University of Shiraz (Iran) and MSc (honor student) in Materials Science and Engineering at the Department of Materials Science & Engineering, Tehran University (Iran). He received his BSc in Materials Science from Shahid Bahonar University of Kerman, Iran in 2004. He has published more than 40 scientific papers in the field of synthesis and characterization of nanostructures materials.