Reactive Separation of Gallic Acid: Experimentation and Optimization Using Response Surface Methodology and Artificial Neural Network

K. Rewatkar, D. Z. Shende, and K. L. Wasewar*  
Advanced Separation and Analytical Laboratory,  
Department of Chemical Engineering,  
Visvesvaraya National Institute of Technology (VNIT),  
Nagpur-440010, Maharashtra, INDIA

Gallic acid is a major phenolic pollutant present in the wastewater generated from cork boiling, olive mill, and pharmaceutical industries. Experimental and statistical modelling using response surface methodology (RSM) and artificial neural network (ANN) were carried out for reactive separation of gallic acid from aqueous stream using tri-n-butyl phosphate (TBP) in hexanol. TBP has a more significant effect on extraction efficiency as compared to temperature and pH. The optimum conditions of 2.34 g L\(^{-1}\), 65.65\% v/v, 19 °C, and 1.8 of initial concentration of gallic acid, concentration of TBP, temperature, and pH, respectively, were obtained using RSM. Under optimum conditions, extraction efficiency of 99.45\% was obtained for gallic acid. The ANN and RSM results were compared with experimental unseen data. Error analysis suggested the better performance of ANN for extraction efficiency predictions.  

Key words: gallic acid, reactive extraction, Artificial Neural Network, Response Surface Methodology, optimization

Introduction

Gallic acid (GA), 3,4,5-trihydroxy benzoic acid, is considered a major phenolic pollutant present in the waste streams generated from the cork boiling process,\(^1\) olive mill,\(^2\) agro-industries, pharmaceuticals, and food-processing industries.\(^1\) The wastewater containing GA is harmful to marine life, creates bad odour, and results in dark colour formation of water. The presence of even a trace amount of this phenolic compound in drinking water imparts an objectionable taste and odour. It is reported that the aromatic ring with three hydroxyl groups and –COOH functional group of GA (Figure 1) are responsible for its toxicity.\(^3,5\) The European Community (EC) directive 80/779/EC prescribes a maximum tolerance level of phenolic compounds in drinking water at 0.5 \(\mu\)g L\(^{-1}\).\(^6\)

Biological processes are generally employed to treat the wastewater because of their process efficiency and economic viability. The waste streams containing GA are highly acidic and toxic. This makes the biological process unsuitable by inhibiting the microbial activity.\(^5,7\) Hence, there has been considerable interest among researchers to focus on the removal of GA from the wastewater due to its toxicity, acidity, and unsafe disposal. Methods such as oxidation,\(^8,9\) colloidal surfactant method,\(^10,11\) photocatalytic degradation,\(^12\) solvent extraction\(^13,14\) etc. are being employed in the removal of GA. Among these, the solvent extraction method is found to be the most suitable, as it offers higher efficiency,\(^15\) and GA readily dissolves in most of the organic solvents.\(^16\) Tri-n-butyl phosphate (TBP), an organophosphorus compound, has very low water solubility (0.4 g L\(^{-1}\) at 25 °C). It is considered an effective extractant for most of the organic acids due to its chemical stability and higher distribution coefficient.\(^17\) The electronegative atom in phosphoryl group (=PO) of TBP forms the hydrogen bonding with the acid molecule to produce molecular complexes in the organic phase in order to improve the extraction efficiency.\(^17,18\)

In recent years, response surface methodology (RSM) and artificial neural network (ANN) have been used for modelling and optimization of several physical processes.\(^19-21\) Both the models offer huge advantages over the conventionally followed one-factor-at-a-time approach, and are considered an effective modelling tool for solving complex nonlinear multivariable systems. These models develop a functional relationship between the input variable and the output response using the experimental data. RSM is a statistical technique generally used...
for the design of experiments, model development, and optimization of process conditions to achieve a target response. The ability to optimize a process and interpret the interactive effects of the process variable on the response using a lesser number of experiments is the key attractive feature of the RSM model. It requires good prior knowledge or extra preliminary experiments to fix the search criteria, and works only for a nonlinear quadratic correlation. The ANN model is superior over RSM for processing highly nonlinear complex systems, and is considered a good technique for both data fitting and prediction ability. It does not require a standard experimental design to develop a model, and is highly adaptive towards new process conditions. However, ANN needs a large number of data points to construct a significant model. The development of the ANN model for the prediction of extraction efficiency of gallic acid has not been reported in the literature. Due to the complex mechanism of reactive extraction of gallic acid, it is difficult to analyze the effect of operating parameters on the response and hence the mathematical modelling becomes difficult to accomplish. The ANN model was developed to predict the removal efficiency of gallic acid from the aqueous solution, and the RSM was used to analyse the relative significance and the interactive effect of the operating parameters, and optimization of the process. Various methods like oxidation, photocatalytic degradation, adsorption, Fenton’s method, electrochemical, extraction, microbial processes, are available for the removal of gallic acid from waste streams, but may not be very effective due to various limitations. Solvent extraction is a clean, economical, and energy efficient method, and the solvents can be reused. Hence, solvent extraction was selected for the removal of gallic acid in the present study.

In the present work, an attempt has been made to apply the RSM and ANN for optimizing the reactive separation of GA using TBP in hexanol. The extraction was carried out by varying the concentration of GA in aqueous phase, concentration of TBP in organic phase, temperature, and pH. The RSM model comprising a rotatable orthogonal central composite design (RO-CCD), and an ANN model based on the Levenberg-Marquardt algorithms were developed. The model predictions were compared with the experimental efficiency to evaluate the performance of both models.

**Materials and methods**

**Material**

Gallic acid (99 % pure) was obtained from Sigma-Aldrich, Germany. Tri-n-butyl phosphate, hexanol, and HCl were obtained from S.D. Fine-Chem. Ltd, India with purity of >98 %. The HPLC grade water, acetonitrile, and acetic acid were obtained from Merck, India with purity of >99 %. All the chemicals were used without purification. The pH of aqueous GA solution was adjusted using HCl. A stock solution of GA (5 g L⁻¹) was prepared using double distilled water, and diluted further to obtain the desired concentration. An automatic potentiometric titrator (AT-38C, Spectra-lab instruments Pvt. Ltd.) was used to determine the pH.

**Experimental method**

An equal volume of 20 mL of gallic acid as aqueous phase and TBP with hexanol as organic phase were added in a 150-mL conical flask, and shaken at 250 rpm for 2 h using an orbital shaking incubator (REMI Instruments Ltd., India). The orbital shaker was calibrated for solution temperature in the flask using K-type thermocouple (TC-902). Preliminary experiments were performed and it was observed that 2 h was sufficient to reach equilibrium (refer to supplementary information, Figure S1).

In the experimental runs, the initial concentration of GA in aqueous phase (C₀), TBP concentration in organic phase (Cₜₚ), extraction temperature (T), and initial pH of aqueous GA solution (pH₀) were varied. The experiments were designed with various combinations to study their synergistic interactive effects on the extraction efficiency (Table 1). After extraction, the equilibrium concentration of GA (Cₐₐ) in the aqueous phase was determined using HPLC. Whereas, the equilibrium concentration of organic phase (Cₐₐ) was determined by material balance as water solubility of TBP is negligible. The extraction efficiency E (%) of GA was calculated as:

\[
E (\%) = \frac{C_{G_{aq}}}{C_{G_{0}}} \cdot 100
\]

All the runs, except centre point, were carried out in duplicate to check the reproducibility of extraction efficiency E (%), which were found to be consistent within ±1 % error.
Analytical methods

The equilibrium concentration of GA (C_{\text{GA,aq}}) in the aqueous phase was determined using HPLC (Agilent 1200, USA) equipped with a quaternary pump, diode array detector, 20 µL loop, and a C18 column (4.6 mm ID×250 mm, 5 µm). The mobile phase was a mixture of 10 % acetonitrile and 90 % water, and its flow rate was kept constant at 1 mL min^{-1}. The wavelength of GA was set at 264 nm. The average retention time of GA was found to be 3.92 min. A calibration curve for GA was obtained with the concentration range from 0.005 to 0.3 g L^{-1} (refer to supplementary information, Figure S1). The analysis was repeated twice under identical conditions for each sample, and the average value of C_{\text{GA,aq}} was recorded.

Experimental design

Response Surface Methodology (RSM)

RSM is an effective experimental design technique used to determine the main, quadratic and in-
is the weight trial experiments and literature survey. and their levels, which were decided based on the regression model was evaluated using the Fischer package “Design Expert”. The significance of the tion were performed using the Statistical Software centre points n 0 (Run 25–28: C1–C4). The sum below, developed by Khuri and Cornell. 27

1.607 using the expression below:28

\[ \alpha \rightarrow 1 0 +1 \rightarrow \alpha \]

\[ E = \sum_{i=1}^{4} \beta_{i} + \sum_{j=1}^{4} \beta_{j} + \sum_{i=1}^{4} \beta_{i} x_{i}^{2} + \sum_{j=1}^{4} \beta_{j} x_{j}^{2} \] (4)

where \( E \) is the predicted response (extraction efficiency, %), and \( \beta_{i}, \beta_{j}, \beta_{ij} \) are the regression coefficients representing linear, quadratic, and interactive effects of factors i.e. \( x_{i} (C_{GA0}), x_{j} (C_{TBP}), x_{3} (T), \) and \( x_{4} (pH_{0}) \). Table 2 shows the independent variables and their levels, which were decided based on the trial experiments and literature survey.

The regression analysis and process optimization were performed using the Statistical Software package “Design Expert”. The significance of the regression model was evaluated using the Fischer distribution (F-value) and null-hypothesis test (p-value). A larger F-value indicates a better fit of model to the experimental extraction efficiency. A p-value less than 0.05 (\( p < 0.05 \)) indicates the design variable of a model contributing less than 5 % change in the response. Therefore, the variable with a larger F-value and \( p < 0.05 \) was considered significant.22

Three-dimensional surface plots were drawn to determine the interactive effect of the process variables on extraction efficiency. The numerical optimization method was used to identify the combination of variables that jointly optimize a single response \( E \) (%) in the Design-Expert software.

Artificial Neural Network (ANN)

ANN is considered a powerful modelling tool to study the process of multivariable, complex, non-linear systems. The single hidden layer is a universal approximation for the multilayer feed forward neural network.29 The operation of a single neuron is shown in Figure 2. The input to a neuron is its bias (\( b_{i,j} \)) and the sum of weighted outputs from the previous layer (\( y_{i-1,j} \)), whereas the output depends on the transfer function \( f(x_{i,j}) \) of its input, and can be represented as:

\[ y_{i,j} = f(x_{i,j}) \] (5)

\[ y_{i,j} = x_{i,j} \] for i = 1 (input layer)

where \( x_{i,j} \), \( y_{i,j} \), \( b_{i,j} \) are the input, output, and bias of jth neuron to ith layer, respectively, and \( J_{i-1} \) is the total neurons in the (i-1)th layer. The \( w_{i-1,j} \) is the weight between kth neuron of (i-1)th layer and jth neuron of ith layer, and \( f(x_{i,j}) \) is a suitable transfer function. The sigmoid transfer function (logsig) for hidden neurons and the linear transfer function (purelin) for output neuron were used, as given below:30
The gallic acid is extracted using TBP by forming a complex (GA:TBP) via interfacial reaction (Eq. 13). The phosphoryl group of TBP forms a hydrogen bond with the proton donor gallic acid to form a stable complex in the organic phase. This can be represented as:

\[ pGA_{aq} + mTBP_{org} \rightleftharpoons (GA_pTBP_m)_{org} \]  

It is evident that the maximum GA:TBP complex formation results in maximum removal efficiency of GA that can be obtained only with a particular combination of \( T \), \( \text{pH}_{aq} \), \( C_{GA0} \) and \( C_{TBP} \). However, a large number of experiments may be required for identifying the suitable process conditions. Therefore, the RSM and ANN models were developed to identify the optimal process conditions required for maximum extraction efficiency of GA.

**RSM modelling of GA extraction**

The design of experiments and experimental extraction efficiency \( E(\%) \) are summarized in Table 1. The extraction efficiency for each experiment was calculated using Eq. (1). Runs 25–28 (C1–C4) of centre point were used to determine the experimental error and reproducibility of the data. The experimental data were regressed to obtain the analysis of variance (ANOVA), as shown in Table 3. The model is highly significant and statistically fit with 99.99% confidence level. The empirical relationship between the response and the independent variables is expressed in terms of coded variables as follows:

\[ E = 97.37 + 0.017x_1 + 3.46x_2 - 0.51x_3 - 0.51x_4 - 1.13x_5^2 - 1.51x_6^2 - 0.11x_7^2 - 0.055x_8^2 + 0.24x_9x_2 - 0.25x_9x_3 - 0.011x_9x_4 + 0.42x_2x_3 + 0.38x_2x_4 + 0.24x_3x_4 \]  

The positive and negative sign of the coefficients indicate the linear increasing or decreasing effect of the corresponding variable on the response, respectively. The significance of a particular variable increases with the sum of square (SS). The ANOVA analysis (Table 3) suggests that \( C_{TBP} \) (\( p = 0.0001 \), SS = 253.36, \( F = 882.28 \)) has a most significant effect on the response as compared to \( T \) (\( p = 0.0007 \), SS = 5.58, \( F = 19.43 \)) and \( \text{pH}_{aq} \) (\( p = 0.0007 \), SS = 5.54, \( F = 19.28 \)). The linear coefficients were more significant than the quadratic or interactive coefficients.

The degree of fitness of the model to the experimental data was determined by analysing the coefficient of determination \((R^2)\). The \( R^2 \) obtained was 0.98, which indicates that the developed model...
could well explain 98 % of total variation in the response, and confirms the goodness of fit for the RSM model. The $R^2$ (adj.) and $R^2$ (pred.) values were 0.97 and 0.92, respectively, which are closer to the $R^2$ value of 0.976, suggesting the significance and reasonable prediction ability of the model. Adequate precision (signal to noise ratio) compares the predicted values with the average prediction error. Adequate precision with a ratio greater than 4 is always desirable. The value of 30.82 indicates an adequate signal. The coefficient of variation (CV) was calculated to be 0.56 %, which indicates better precision and reliability of the experiments. The Standard deviation (S) with 0.54 indicates strong compliance with the predicted response. The predicted residual sum of square (PRESS) is a measure of predictive power of the developed model. The PRESS value of 22.14 suggests that the regression model is significant to predict the response for a new experiment. The coded terms of RSM model was then converted into actual variables and represented as follows:

$$E = 90.11 + 1.18(C_{GA0} + 0.29(C_{TBP}) - 0.11(T) - 2.57(pH_0) - 0.26(C_{GA0})^2 - 0.0024(C_{TBP})^2 - 0.0015(T)^2 - 0.11(pH_0)^2 + 0.014(C_{GA0})(C_{TBP}) - 0.043(C_{GA0})T - 0.022(C_{GA0})(pH_0) + 0.002(C_{TBP})(T) + 0.022(C_{TBP})(pH_0) + 0.041(T)(pH_0)$$

Response surface analysis, process optimization and validation

Figures 3a and 3b illustrate the 3D surface plots for the individual and the interactive effects of independent variables on extraction efficiency. In Figure 3a, the concentration of extractant, $C_{TBP}$ ($x_2$) and its interaction with pH ($x_4$) show a positive effect on extraction efficiency (Table 3), whereas pH alone shows a negative effect. A curvilinear effect on extraction efficiency was observed due to the positive linear effect and negative quadratic effect of $C_{TBP}$. This is attributed to the presence of the highly polar phosphoryl group (>P=O) of TBP making it a strong Lewis base, resulting in higher GA:TBP complex formation during the extraction process. Hence, the higher $C_{TBP}$ resulted in higher extraction. The concentration of the undissociated acid always increases at lower pH due to the higher proton concentration. TBP extracts only undissociated form of acid, hence extraction efficiency $E$ (%) increases at lower pH values. A similar effect of pH on the increase in efficiency was also observed in the extraction of other carboxylic acids.

Figure 3b illustrates that $C_{TBP}$ ($x_2$) and $T$ ($x_3$) has a positive and negative effect, respectively, on extraction efficiency $E$ (%), whereas, the combined effect ($x_2x_3$) shows a positive effect (Table 3). With an increase in temperature, the formation of GA:TBP complex decreases. The complex is formed via intermolecular hydrogen bonding through proton transfer between acid and extractant. The process of hydrogen bond formation is exothermic in nature, which shifted the extraction equilibrium in backward direction at higher temperature, thus reducing extraction efficiency. Similar observations have been made by many researchers. Therefore, it is concluded that lower temperature and pH with higher $C_{TBP}$ favour the extraction efficiency.
Optimization of process variables $C_{GA0}$, $C_{TBP}$, $T$, and $pH_0$ was performed using the numerical optimization method to achieve optimal conditions to maximize the extraction. The optimum variables were $C_{GA0} = 2.34$ g L$^{-1}$, $C_{TBP} = 65.65$ % (v/v), $T = 18.4$ °C and $pH_0 = 1.8$ with the predicted optimum extraction of 99.8%. The experiment at optimum conditions was carried out in duplicate to validate the model predictions at the optimum conditions. The experimental extraction efficiency obtained at the optimum conditions was found to be 99.43 %, which is in close agreement with the model predict-
ed value. In Table 1, Runs no. 3, 4, 12, 20 show the extraction efficiency above 99.2 % at the cost of 80–95 % TBP, while the optimum condition gives the same efficiency with 65 % TBP.

**ANN modelling of the GA extraction process**

In ANN, the three layered feed forward error back propagation (FF-EBP) with four neurons in the input layer (input parameters: \( C_{\text{GA0}}, C_{\text{TBP}}, T, \text{pH}_0 \)), ten neurons in the hidden layer (optimum), one neuron in the output layer (extraction efficiency) was selected. Figure 4 illustrates the ANN architecture applied for predicting the extraction efficiency of GA.

A total of 69 experimental data-points (Table 1 and 4) were used to train the ANN model. The input and the output data-points were normalized in the range 0 to 1, and 0.3 to 0.7, respectively. This keeps the back propagation error within the limits and avoids over fitting of the data due to very large and very small values of weights, and speeds up training time. The data were normalized between the two points \( n_i \) and \( n_0 \) using the following generalized equation:

\[
x_n = \left( \frac{n_i - n_0}{x_{\text{max}} - x_{\text{min}}} \right) (x - x_{\text{min}}) + n_0
\]

where \( x_n \) is normalized data-point, and \( x, x_{\text{min}}, x_{\text{max}} \) are the actual, minimum, and maximum of input/output data. The data were split into subsets of training (80 %, 55 data-points), validation (10 %, 7 data-points), and testing (10 %, 7 data-points). Trainlm, a training function that updates weights and bias values based on Levenberg-Marquardt backpropagation (LMB) algorithm, was used for training the network. LMB is an iterative technique which reduces the performance function in each iteration, and makes trainlm the fastest error backpropagation (EBP) algorithm for a moderate sized network. The externally normalized input values were forwarded from the input layer to the hidden layer and then to the output layer to predict the response. The error, MSE was backpropagated from the output to the hidden layer and thereafter to the input layer to modify the weights (IW, LW) and biases (\( b_1, b_2 \)). Figure 5 describes the general concept of FF-EBP algorithm used in ANN training to predict extraction efficiency. During the series of trials for EBP, the training algorithm, connection weights, transfer function, hidden layers, and hidden neurons were varied, and a network topology with an optimal network architecture (4:10:1) was selected based on MSE value.

The training was successfully terminated after 12 epochs (iterations) as the performance function
MSE reached a minimum value of $5 \cdot 10^{-4}$ which is lower than the set goal, $E_0 = 1 \cdot 10^{-3}$ (Figure 6). The similar characteristic curve for the test and the validation were observed, suggesting no significant over-fitting. The estimated MSE values for training ($1 \cdot 10^{-4}$), testing ($3 \cdot 10^{-4}$), and validation data ($5 \cdot 10^{-4}$) were found to be lower than the set goal, which is desirable. Table 5 gives the optimal weights and bias. Wherein $IW = H \times A$, $b_1 = H \times 1$, $LW = O \times H$ matrix, where $H$ is hidden layer neurons (10), $A$ is input layer neurons (4), and $O$ is the output layer neuron (1). The bias $b_2$ is the single value associated with a single neuron at the output.

Figure 7 describes the ANN regression plot for training, validation, testing, and the overall prediction set in the form of network output versus experimental extraction efficiency. It can be observed that the network output values were close to the ex-

![Fig. 5 - Feed forward backpropagation algorithm used in the ANN training for the prediction of GA extraction efficiency](image)
The correlation coefficients $R$ for training, validation, and testing were 0.995, 0.984, 0.989, respectively, whereas the overall prediction set was 0.993, which confirms that the ANN model is satisfactory for interpolating the experimental data.
The ANN model was also validated using statistical analysis, ANOVA. The degree of freedom due to regression \((DF_{\text{regression}})\), and residuals \((DF_{\text{residual}})\) was calculated as follows:\(^{20}\)

\[
DF_{\text{regression}} = M - 1
\]

\[
DF_{\text{residual}} = N - M
\]

\[
M = H(A + O + 1) + O
\]

where, \(M\) is the total number of network connections including weights and biases. \(N\) is the total number of experiments \((N = 69)\) performed to construct and validate the ANN model. Table 6 shows the ANOVA for the developed ANN model. The \(F\)-value and the \(p\)-value were obtained from ANOVA for the ANN model as 10.41 and 0.0004, respectively, which confirms the significance of the ANN model. Table 7 shows the values of \(R^2\) and \(R^2\) (adj.) as 0.986 and 0.88, respectively, indicating the good fit of the ANN model to the experimental extraction efficiency. Therefore, the ANN model for predicting extraction efficiency \(E(\%)\) for the extraction of GA can be written as follows:

\[
y_n = \text{purelin}(LW \times \text{logsig}(IW \times x_n + b_i) + b_j)
\]

\[\text{Fig. 8 – Comparison of RSM and ANN model predictions for the unseen data (data not used in the model development)}\]
Conclusion

The present work suggests that gallic acid can be efficiently removed from aqueous streams using TBP in hexanol. The RSM and ANN models were developed to predict the extraction efficiency of GA by varying the parameters: temperature, concentration of TBP in aqueous phase, pH, and initial concentration of gallic acid in aqueous solution. The importance of each parameter, and its synergistic interactive effects on the extraction of GA was explained with a minimum number of experiments using the RSM model. Both models were efficient in the prediction of GA extraction efficiency, having close agreement with the experimental extraction efficiency. However, ANN showed superiority over RSM in terms of accuracy, but the marginal errors between the models were very low. The optimal conditions to achieve maximum extraction, as predicted by the RSM model were $C_{GA0} = 2.34 \text{ g L}^{-1}$, $C_{TBP} = 65.65 \% \text{ (v/v)}$, $T = 18.4 \, ^\circ \text{C}$, and $\text{pH}_0 = 1.8$, at which an experimental extraction efficiency of 99.5\% was observed. The models also showed better performance with the use of unseen data, which confirmed their generalization capabilities. Thus, the present investigation on experiments, modeling, and optimization of the GA extraction process could be of great significance for the design and evaluation of the performance of similar systems.

Supplementary Information

Supplementary information for confirmation of 2 h shaking time for reactive extraction of gallic acid and retention time of 3.9 min in HPLC analysis.

The trial experiments were conducted for obtaining an optimum shaking period. Equal volumes (20 mL) of aqueous and organic phases were shaken for different time intervals (5–240 min) at 250 rpm. Equilibrium was achieved after half an hour (Figure S1). Hence, further experiments were conducted with 2 h shaking time (sufficient to achieve equilibrium).

Calibration curve and chromatogram for HPLC analysis of gallic acid was obtained by diluting the mother stock of 5 g L$^{-1}$ to prepare different concentrations (5–300 mg L$^{-1}$) of standard sample solutions.

UV absorption spectra for gallic acid was obtained using UV Spectrophotometer Shimadzu UV-1800.

Supplementary Information

Supplementary information for confirmation of 2 h shaking time for reactive extraction of gallic acid and retention time of 3.9 min in HPLC analysis.

The trial experiments were conducted for obtaining an optimum shaking period. Equal volumes (20 mL) of aqueous and organic phases were shaken for different time intervals (5–240 min) at 250 rpm. Equilibrium was achieved after half an hour (Figure S1). Hence, further experiments were conducted with 2 h shaking time (sufficient to achieve equilibrium).

Calibration curve and chromatogram for HPLC analysis of gallic acid was obtained by diluting the mother stock of 5 g L$^{-1}$ to prepare different concentrations (5–300 mg L$^{-1}$) of standard sample solutions.

UV absorption spectra for gallic acid was obtained using UV Spectrophotometer Shimadzu UV-1800.
Nomenclature

- AARE: absolute average relative error
- $C_{GA,aq}$: equilibrium concentration of GA in aqueous phase, g L$^{-1}$
- $C_{GA,org}$: equilibrium concentration of GA in organic phase, g L$^{-1}$
- $C_{GA0}$: initial concentration of GA in aqueous phase, g L$^{-1}$
- $C_{TBP}$: TBP concentration in organic phase, %
- $DF$: degree of freedom in ANOVA analysis
- $E$: extraction efficiency, response, %
- $F$-value: Fischer distribution value (ratio of variances)
- $H, A, O$: number of neurons in hidden layer, input layer, and output layer, respectively
- $IW$: input weight matrix
- $k$: number of input variables or factors in RO-CCD
- logsig: sigmoidal transfer function for neural network
- $LW$: layer weight matrix
- $M$: number of network connections including weights and biases
- $MS$: mean square in ANOVA analysis
- $MSE$: mean squared error
- $N$: total number of experiments to construct ANN model
- $n_0$: centre points in RO-CCD
- $n_s$: star or axial points in RO-CCD
- $pH_0$: initial pH of aqueous GA solution
- PRESS: predicted residual sum of squares
- purelin: linear transfer function for neural network
- $p$-value: null-hypothesis test
- $R$: coefficient of correlation
- $R^2$: coefficient of determination
- $R^2$ (adj.): adjusted R-squared
- $R^2$ (pred.): predicted R-squared
- RMSE: root mean square error
- SEP: standard error of prediction
- SS: sum of squares in ANOVA analysis
- $T$: extraction temperature, °C
- $w$: connection weight between layers
- $x_1, x_2, x_3, x_4$: coded levels of independent variables in RSM
- $x_{ij}, y_{ij}, b_{ij}$: network input, output and bias of $j^{th}$ neuron to $i^{th}$ layer, respectively
- $a$: distance of each star point from centre in RO-CCD
- $\beta_1, \beta_2, \beta_0$: regression coefficients representing linear, quadratic, and interactive effects in RSM

References


