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



doi: 10.15255/CABEQ.2016.931

Original scientific paper
Received: February 9, 2016
Accepted: March 22, 2017

**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.[3] 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.[4,5] The European Community (EC) directive 80/779/EC prescribes a maximum tolerance level of phenolic compounds in drinking water at 0.5 μg $L^{-1}$.[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.[1,7] 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

*Corresponding author; Email: k_wasewar@rediffmail.com;
Phone: +91-712-2801561, Fax: +91-712-2223969

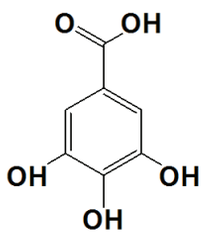



Fig. 1 – *Chemical structure of GA*

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.[22,23] 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.[21,24,25] 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.[26] 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 ($C_{GA0}$), TBP concentration in organic phase ($C_{TBP}$), extraction temperature ($T$), and initial pH of aqueous GA solution ($pH_0$) 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 ($C_{GA,aq}$) in the aqueous phase was determined using HPLC. Whereas, the equilibrium concentration of organic phase ($C_{GA,org}$) was determined by material balance as water solubility of TBP is negligible. The extraction efficiency $E$ (%) of GA was calculated as:

$$E\ (\%) = \frac{C_{GA,org}}{C_{GA0}} \cdot 100 \qquad (1)$$

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.

Table 1 – *Rotatable orthogonal central composite design (RO-CCD) with experimental and predicted extraction efficiency of GA*

| Run | Type | Independent variables (factors) | | | | | | | | Response, $E$ (%) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $C_{GA0}$ (g $L^{-1}$) | $x_1$ | $C_{TBP}$ (% v/v) | $x_2$ | $T$ (°C) | $x_3$ | $pH_0$ | $x_4$ | Experimental | RSM | ANN | Remark |
| 1 | O1 | 0.7 | –1 | 30 | –1 | 18 | –1 | 1.2 | –1 | 94.27 | 94.14 | 94.55 | Training |
| 2 | O2 | 2.1 | +1 | 30 | –1 | 18 | –1 | 1.2 | –1 | 94.47 | 94.23 | 94.43 | Training |
| 3 | O3 | 0.7 | –1 | 80 | +1 | 18 | –1 | 1.2 | –1 | 99.37 | 98.96 | 99.37 | Training |
| 4 | O4 | 2.1 | +1 | 80 | +1 | 18 | –1 | 1.2 | –1 | 99.30 | 99.99 | 99.27 | Training |
| 5 | O5 | 0.7 | –1 | 30 | –1 | 35 | +1 | 1.2 | –1 | 92.04 | 92.28 | 92.02 | Training |
| 6 | O6 | 2.1 | +1 | 30 | –1 | 35 | +1 | 1.2 | –1 | 91.71 | 91.35 | 91.43 | Training |
| 7 | O7 | 0.7 | –1 | 80 | +1 | 35 | +1 | 1.2 | –1 | 98.92 | 98.81 | 98.84 | Training |
| 8 | O8 | 2.1 | +1 | 80 | +1 | 35 | +1 | 1.2 | –1 | 98.99 | 98.83 | 98.73 | Testing |
| 9 | O9 | 0.7 | –1 | 30 | –1 | 18 | –1 | 2.6 | +1 | 91.53 | 91.88 | 92.18 | Training |
| 10 | O10 | 2.1 | +1 | 30 | –1 | 18 | –1 | 2.6 | +1 | 91.76 | 91.92 | 92.39 | Training |
| 11 | O11 | 0.7 | –1 | 80 | +1 | 18 | –1 | 2.6 | +1 | 97.84 | 98.24 | 97.03 | Testing |
| 12 | O12 | 2.1 | +1 | 80 | +1 | 18 | –1 | 2.6 | +1 | 99.29 | 99.24 | 99.91 | Training |
| 13 | O13 | 0.7 | –1 | 30 | –1 | 35 | +1 | 2.6 | +1 | 91.66 | 91.03 | 91.62 | Training |
| 14 | O14 | 2.1 | +1 | 30 | –1 | 35 | +1 | 2.6 | +1 | 89.43 | 90.03 | 90.03 | Training |
| 15 | O15 | 0.7 | –1 | 80 | +1 | 35 | +1 | 2.6 | +1 | 98.62 | 99.06 | 98.88 | Training |
| 16 | O16 | 2.1 | +1 | 80 | +1 | 35 | +1 | 2.6 | +1 | 98.87 | 99.05 | 98.11 | Validation |
| 17 | S1 | 0.275 | –$\alpha$ | 55 | 0 | 26.5 | 0 | 1.9 | 0 | 96.98 | 97.01 | 96.98 | Training |
| 18 | S2 | 2.525 | +$\alpha$ | 55 | 0 | 26.5 | 0 | 1.9 | 0 | 97.47 | 97.06 | 97.26 | Training |
| 19 | S3 | 1.4 | 0 | 14.83 | –$\alpha$ | 26.5 | 0 | 1.9 | 0 | 87.79 | 87.91 | 87.58 | Training |
| 20 | S4 | 1.4 | 0 | 95.18 | +$\alpha$ | 26.5 | 0 | 1.9 | 0 | 99.55 | 99.04 | 99.28 | Training |
| 21 | S5 | 1.4 | 0 | 55 | 0 | 12.8 | –$\alpha$ | 1.9 | 0 | 98.29 | 97.90 | 98.26 | Training |
| 22 | S6 | 1.4 | 0 | 55 | 0 | 40.2 | +$\alpha$ | 1.9 | 0 | 96.25 | 96.25 | 96.43 | Validation |
| 23 | S7 | 1.4 | 0 | 55 | 0 | 26.5 | 0 | 0.775 | –$\alpha$ | 97.65 | 98.04 | 97.72 | Training |
| 24 | S8 | 1.4 | 0 | 55 | 0 | 26.5 | 0 | 3.025 | +$\alpha$ | 97.18 | 96.40 | 96.73 | Training |
| 25 | C1 | 1.4 | 0 | 55 | 0 | 26.5 | 0 | 1.9 | 0 | 97.26 | 97.36 | 97.38 | Training |
| 26 | C2 | 1.4 | 0 | 55 | 0 | 26.5 | 0 | 1.9 | 0 | 97.28 | 97.36 | 97.38 | Testing |
| 27 | C3 | 1.4 | 0 | 55 | 0 | 26.5 | 0 | 1.9 | 0 | 97.22 | 97.36 | 97.38 | Training |
| 28 | C4 | 1.4 | 0 | 55 | 0 | 26.5 | 0 | 1.9 | 0 | 97.17 | 97.36 | 97.38 | Training |

O: orthogonal design points, C: centre points, S: star points, +/–α: star point value, –1: low value, 0: centre value, +1: high value

## Analytical methods

The equilibrium concentration of GA ($C_{GA,aq}$) 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 $C_{GA,aq}$ was recorded.

## Experimental design

### *Response Surface Methodology (RSM)*

RSM is an effective experimental design technique used to determine the main, quadratic and in-

Table 2 – *Levels of independent variables used for rotatable orthogonal central composite design*

| Independent variables (factors) | Levels of factors | | | | |
|---|---|---|---|---|---|
| | $-\alpha$ | –1 | 0 | +1 | $+\alpha^*$ |
| $C_{GA0}$ (g $L^{-1}$) | 0.275 | 0.7 | 1.4 | 2.1 | 2.525 |
| $C_{TBP}$ (% v/v) | 14.83 | 30 | 55 | 80 | 95.18 |
| $T$ (°C) | 12.84 | 18 | 26.5 | 35 | 40.16 |
| $pH_0$ | 0.775 | 1.2 | 1.9 | 2.6 | 3.025 |

* $\alpha = 1.607$

teractive effects of operating variables on the response. The rotatable orthogonal central composite design (RO-CCD) was used to develop a quadratic experimental model. In the rotatable design, the prediction variance of a point depends only on its distance from the centre of the design irrespective of its direction. The orthogonality allows all the factors to be estimated independently. RO-CCD consists of the corner points $n_c = 2^k$, the star points $n_s = 2k$, and the centre points $n_0$.

To make CCD both orthogonal and rotatable, the centre points, $n_0$, were calculated as 4 by substituting k = 4 factors ($C_{GA0}$, $C_{TBP}$, $T$, $pH_0$), in the Eq. (2) below, developed by Khuri and Cornell.[27]

$$n_0 = (4n_c)^{1/2} + 4 - 2k \tag{2}$$

The axial distance, $\alpha$ was calculated to be 1.607 using the expression below:

$$\alpha = \left\{ \left[ \left( n_c + n_s + n_0 \right)^{1/2} - n_c^{1/2} \right]^2 . \frac{n_c}{4} \right\}^{1/4} \tag{3}$$

Table 1 shows the corner points $n_c$ (Run 1–16: O1–O16), star points $n_s$ (Run 17–24: S1–S8), and centre points $n_0$ (Run 25–28: C1–C4). The sum $n_c + n_0 + n_s = 28$. Hence, 28 batch experiments were carried out to construct the RSM model. The relationship between independent variables (factors) and the extraction efficiency (response) was established using a second-order polynomial equation as follows:[28]

$$E = \beta_0 + \sum_{i=1}^{4} \beta_i x_i + \sum_{i=1}^{4} \beta_{ii} x_i^2 + \Sigma \sum_{i=1}^{4} \beta_{ij} x_i x_j \tag{4}$$

where $E$ is the predicted response (extraction efficiency, %), and $b_i$, $b_j$, $b_{ij}$, are the regression coefficients representing linear, quadratic, and interactive effects of factors i.e. $x_1$ ($C_{GA0}$), $x_2$ ($C_{TBP}$), $x_3$ ($T$), 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

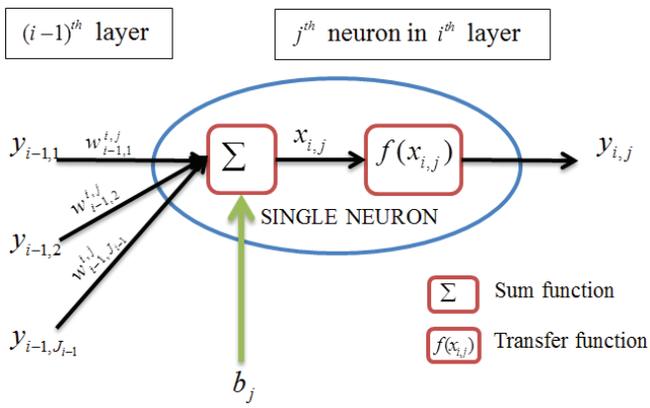



Fig. 2 – *Operation of single neuron in ANN*

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, nonlinear 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 ($b_j$) and the sum of weighted outputs from the previous layer ($y_{i-1}$), whereas the output depends on the transfer function $f(x_{i,j})$ of its input, and can be represented as:

$$\begin{aligned} &\text{Input} \quad x_{i,j} = \sum_{k=1}^{J_{i-1}} w_{i-1,k}^{i,j} y_{i-1,k} + b_{i,j} \\ &\text{Output} \quad y_{i,j} = f(x_{i,j}) \quad \text{for } i > 1 \\ &\qquad\qquad y_{i,j} = x_{i,j} \qquad \text{for } i = 1 \quad \text{(input layer)} \end{aligned} \tag{5}$$

where $x_{i,j}$, $y_{i,j}$, $b_{i,j}$, are the input, output, and bias of $j^{th}$ neuron to $i^{th}$ layer, respectively, and $J_{i-1}$ is the total neurons in the $(i-1)^{th}$ layer. The $w_{i-1,k}^{i,j}$ is the weight between $k^{th}$ neuron of $(i-1)^{th}$ layer and $j^{th}$ neuron of $i^{th}$ 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]

$$logsig\ (x) = \frac{1}{1+e^{-x}} \tag{6}$$

$$purelin\ (x) = x \tag{7}$$

Matlab software R2012a was used for ANN modelling. The connections between the input layer and the hidden layer, and the hidden layer and the output layer consist of weights and biases. The weights which express the interactions between the neurons were used in two types of matrix, i.e., input weight matrix (IW) connecting the input layer to the hidden layer, and the layer weight matrix (LW) connecting the hidden layer to the output layer. Each neuron in the first layer was connected to all the neurons in the next layer, with a total of 40 connections between the input and hidden layers. The bias determines the threshold to fire a neuron. Each neuron has an individual bias value except for the input layer neurons.[30] The optimal number of neurons in the hidden layer was decided by the trial and error approach by varying it, until a minimum mean squared error (MSE) was achieved. MSE is an error function which measures the performance of the ANN model, and can be calculated as follows:[21]

$$\text{MSE} \quad \frac{\sum_{i\ 1}^{i\ N}(E_{i,pred} \quad E_{i,exp})}{} \tag{8}$$

where, N is the number of experimental data points, $E_{i,pred}$ and $E_{i,exp}$ are the ANN predicted and experimental responses, respectively.

To compare both constructed models, an error analysis was carried out with experimental and the predicted response using coefficient of determination ($R^2$), root mean square error (RMSE), absolute average relative error (AARE), and standard error of prediction (SEP), calculated as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(E_{i,pred} - E_{i,exp})^2}{\sum_{i=1}^{N}(E_{i,pred} - E_{avg})^2} \tag{9}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(E_{i,pred} - E_{i,exp})^2}{N}} \tag{10}$$

$$\text{AARE} = \frac{100}{N}\sum_{i=1}^{N}\frac{\left|E_{i,pred} - E_{i,exp}\right|}{\left|E_{i,exp}\right|} \tag{11}$$

$$\text{SEP} = \frac{\text{RMSE}}{E_{avg}}.100 \tag{12}$$

Where $E_{avg}$ is the average value of the experimental response.

## Results and discussion

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:

$$pGA_{aq} + mTBP_{org} \rightleftharpoons (GA_p{:}TBP_m)_{org} \tag{13}$$

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 $T$, $pH_0$, $C_{GA0}$ 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:

$$\begin{aligned} E = {} & 97.37 + 0.017x_1 + 3.46x_2 - 0.51x_3 - \\ & - 0.51x_4 - 1.13x_1^2 - 1.51x_2^2 - 0.11x_3^2 - \\ & - 0.055x_4^2 + 0.24x_1x_2 - 0.25x_1x_3 - 0.011x_1x_4 - \\ & + 0.42x_2x_3 + 0.38x_2x_4 + 0.24x_3x_4 \end{aligned} \tag{14}$$

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 $pH_0$ ($p$ = 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

Table 3 – *Analysis of variance (ANOVA) for the experimental results of the rotatable orthogonal central composite design*

| Source | Coded terms | Coefficient | Sum of Squares | *F*-value | *p*-value | Remarks |
|---|---|---|---|---|---|---|
| *Model* | | | 303.3 | 75.44 | 0.0001 | Significant |
| Constant | | 97.37 | | | 0.0001 | |
| ***Linear*** | | | | | | |
| $C_{GA0}$ | $x_1$ | 0.017 | 0.00604 | 0.021 | 0.887 | |
| $C_{TBP}$ | $x_2$ | 3.46 | 253.36 | 882.28 | 0.0001 | Significant |
| $T$ | $x_3$ | –0.51 | 5.58 | 19.43 | 0.0007 | Significant |
| $pH_0$ | $x_4$ | –0.51 | 5.54 | 19.28 | 0.0007 | Significant |
| ***Square*** | | | | | | |
| $C_{GA0} \cdot C_{GA0}$ | $x_1^2$ | –0.13 | 0.22 | 0.77 | 0.395 | |
| $C_{TBP} \cdot C_{TBP}$ | $x_2^2$ | –1.51 | 30.24 | 105.3 | 0.0001 | Significant |
| $T \cdot T$ | $x_3^2$ | –0.11 | 0.17 | 0.58 | 0.461 | |
| $pH_0 \cdot pH_0$ | $x_4^2$ | –0.055 | 0.041 | 0.14 | 0.712 | |
| ***Interaction*** | | | | | | |
| $C_{GA0} \cdot C_{TBP}$ | $x_1 x_2$ | 0.24 | 0.92 | 3.19 | 0.097 | |
| $C_{GA0} \cdot T$ | $x_1 x_3$ | –0.25 | 1.03 | 3.51 | 0.081 | |
| $C_{GA0} \cdot pH_0$ | $x_1 x_4$ | –0.011 | 0.00181 | 0.00629 | 0.940 | |
| $C_{TBP} \cdot T$ | $x_2 x_3$ | 0.42 | 2.88 | 10.03 | 0.007 | Significant |
| $C_{TBP} \cdot pH_0$ | $x_2 x_4$ | 0.38 | 2.36 | 8.23 | 0.013 | Significant |
| $T \cdot pH_0$ | $x_3 x_4$ | 0.24 | 0.96 | 3.33 | 0.091 | |
| ***Statistics*** | | | | | | |
| Adequate precision | | 30.82 | | | | |
| PRESS | | 22.14 | | | | |
| S | | 0.54 | | | | |
| CV (%) | | 0.56 | | | | |
| $R^2$ | | 0.98 | | | | |
| $R^2$(adj.) | | 0.97 | | | | |
| $R^2$(pred.) | | 0.92 | | | | |

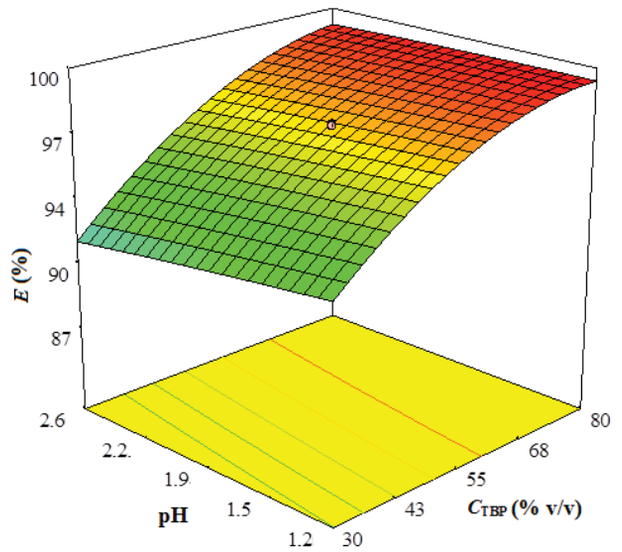
could well explain 98 % of total variation in the response, and confirms the goodness of fit for the RSM model. The $R^2$ (adj.) and $R^2$ (pred.) values were 0.97 and 0.92, respectively, which are closer to the $R^2$ value of 0.976, suggesting the significance and reasonable prediction ability of the 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:

$$\begin{aligned} E \;=\; & 90.11 + 1.18(C_{GA0}) + 0.29(C_{TBP}) - \\ & - 0.11(T) - 2.57(\mathrm{pH}_0) - 0.26(C_{GA0})^2 - \\ & - 0.0024(C_{TBP})^2 - 0.0015(T)^2 - 0.11(\mathrm{pH}_0)^2 + \\ & + 0.014(C_{GA0})(C_{TBP}) - 0.043(C_{GA0})(T) - \\ & - 0.022(C_{GA0})(\mathrm{pH}_0) + 0.002(C_{TBP})(T) + \\ & + 0.022(C_{TBP})(\mathrm{pH}_0) + 0.041(T)(\mathrm{pH}_0) \end{aligned} \qquad (15)$$

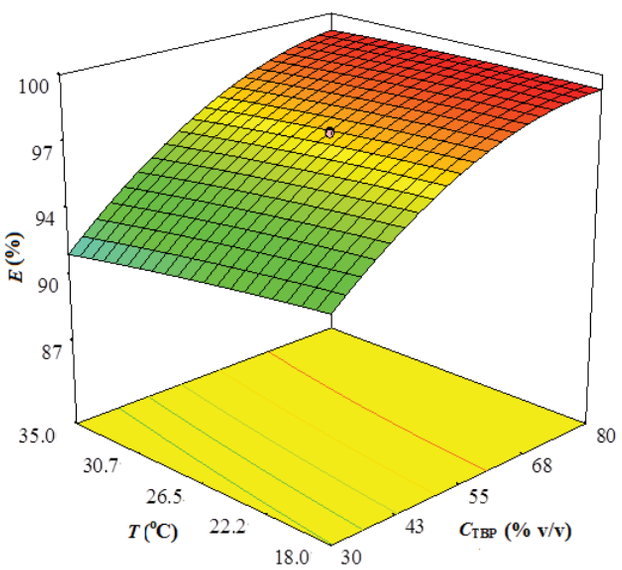
*Response surface analysis, process optimization and validation*

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 ($x_2 x_4$) 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.[18] 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,[31] 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.[32]

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 ($x_2 x_3$) 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.[33,34] Therefore, it is concluded that lower temperature and pH with higher $C_{TBP}$ favour the extraction efficiency.




(a)




(b)

Fig. 3 – *Three dimensional surface plot showing (a) the effect of $C_{TBP}$ and $pH_0$ on the extraction of GA at the centre level ($T$ = 26.5 °C, $C_{GA0}$ = 1.4 g L$^{-1}$), and (b) the effect of $C_{TBP}$ and $T$ on the extraction of GA at the centre level ($pH_0$ = 1.9, $C_{GA0}$ = 1.4 g L$^{-1}$).*

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-

Table 4 – *Additional experimental dataset used for the construction of ANN model*

| Run | Independent variables | | | | Response, $E$ (%) | | |
|---|---|---|---|---|---|---|---|
| | $C_{GA0}$ (g L$^{-1}$) | $C_{TBP}$ (% v/v) | $T$ (°C) | $pH_0$ | Experimental | ANN | Remark |
| 29 | 0.5 | 65 | 11 | 1.96 | 98.95 | 98.68 | Training |
| 30 | 2.5 | 30 | 11 | 3.03 | 94.88 | 95.08 | Training |
| 31 | 1.4 | 30 | 11 | 3.02 | 93.04 | 93.25 | Training |
| 32 | 1.8 | 53 | 11 | 2.13 | 97.28 | 97.00 | Training |
| 33 | 2.5 | 75 | 11 | 3.05 | 98.86 | 98.51 | Training |
| 34 | 1.2 | 22 | 11 | 2.98 | 90.13 | 90.99 | Validation |
| 35 | 0.4 | 82 | 11 | 1.67 | 97.23 | 97.66 | Training |
| 36 | 0.28 | 15 | 11 | 2.25 | 85.72 | 87.04 | Validation |
| 37 | 1.0 | 30 | 15 | 3.08 | 92.77 | 92.43 | Training |
| 38 | 2.5 | 30 | 15 | 3.03 | 94.73 | 94.29 | Validation |
| 39 | 0.28 | 95 | 15 | 0.78 | 97.30 | 97.47 | Training |
| 40 | 1.4 | 30 | 16 | 3.01 | 92.93 | 93.05 | Training |
| 41 | 2.5 | 30 | 16 | 3.02 | 94.49 | 94.02 | Training |
| 42 | 2.5 | 30 | 18 | 3.03 | 93.59 | 93.47 | Training |
| 43 | 0.5 | 25 | 20 | 1.97 | 90.76 | 90.36 | Training |
| 44 | 1.4 | 30 | 20 | 3.02 | 92.74 | 92.12 | Training |
| 45 | 2.5 | 30 | 25 | 3.01 | 91.24 | 91.79 | Training |
| 46 | 0.28 | 15 | 25 | 0.78 | 84.85 | 84.82 | Training |
| 47 | 2.5 | 30 | 27 | 3.03 | 92.63 | 91.50 | Validation |
| 48 | 1.4 | 30 | 27 | 3.02 | 87.68 | 89.34 | Training |
| 49 | 0.4 | 85 | 27 | 1.39 | 98.08 | 97.96 | Training |
| 50 | 2.4 | 72 | 27 | 2.55 | 98.66 | 98.43 | Training |
| 51 | 1.2 | 24 | 27 | 2.2 | 89.31 | 90.44 | Testing |
| 52 | 1.0 | 62 | 28 | 1.72 | 97.58 | 98.26 | Testing |
| 53 | 1.5 | 72 | 28 | 1.65 | 98.52 | 98.31 | Testing |
| 54 | 2.5 | 30 | 28 | 3.03 | 91.71 | 91.36 | Training |
| 55 | 1.6 | 32 | 30 | 2.46 | 92.24 | 92.10 | Training |
| 56 | 1.8 | 78 | 30 | 2.61 | 98.95 | 98.52 | Testing |
| 57 | 1.8 | 27 | 30 | 2.61 | 90.39 | 89.30 | Training |
| 58 | 1.2 | 65 | 30 | 2.54 | 98.06 | 97.92 | Training |
| 59 | 2.5 | 95 | 30 | 0.78 | 98.64 | 98.79 | Training |
| 60 | 1.4 | 30 | 33 | 3.02 | 87.98 | 87.66 | Training |
| 61 | 2.5 | 30 | 35 | 3.03 | 91.14 | 90.62 | Validation |
| 62 | 2.5 | 30 | 35 | 3.012 | 90.38 | 90.63 | Training |
| 63 | 1.4 | 30 | 37 | 3.02 | 87.33 | 87.29 | Training |
| 64 | 2.4 | 80 | 37 | 2.22 | 98.59 | 98.57 | Training |
| 65 | 0.8 | 60 | 37 | 2.18 | 97.18 | 96.94 | Training |
| 66 | 2.4 | 16 | 37 | 2.09 | 85.15 | 85.36 | Training |
| 67 | 1.6 | 32 | 37 | 1.78 | 91.04 | 91.02 | Training |
| 68 | 2.5 | 30 | 41 | 3.03 | 90.99 | 90.32 | Training |
| 69 | 1.0 | 30 | 41 | 1.50 | 88.83 | 88.90 | Training |

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: $C_{GA0}$, $C_{TBP}$, $T$, $pH_0$), 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 $n_1$ and $n_0$ using the following generalized equation:

$$x_n = \left( (n_1 - n_0) . \frac{(x - x_{min})}{(x_{max} - x_{min})} \right) + n_0 \qquad (16)$$

where $x_n$ is normalized data-point, and $x$, $x_{min}$, $x_{max}$ 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.[35] 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

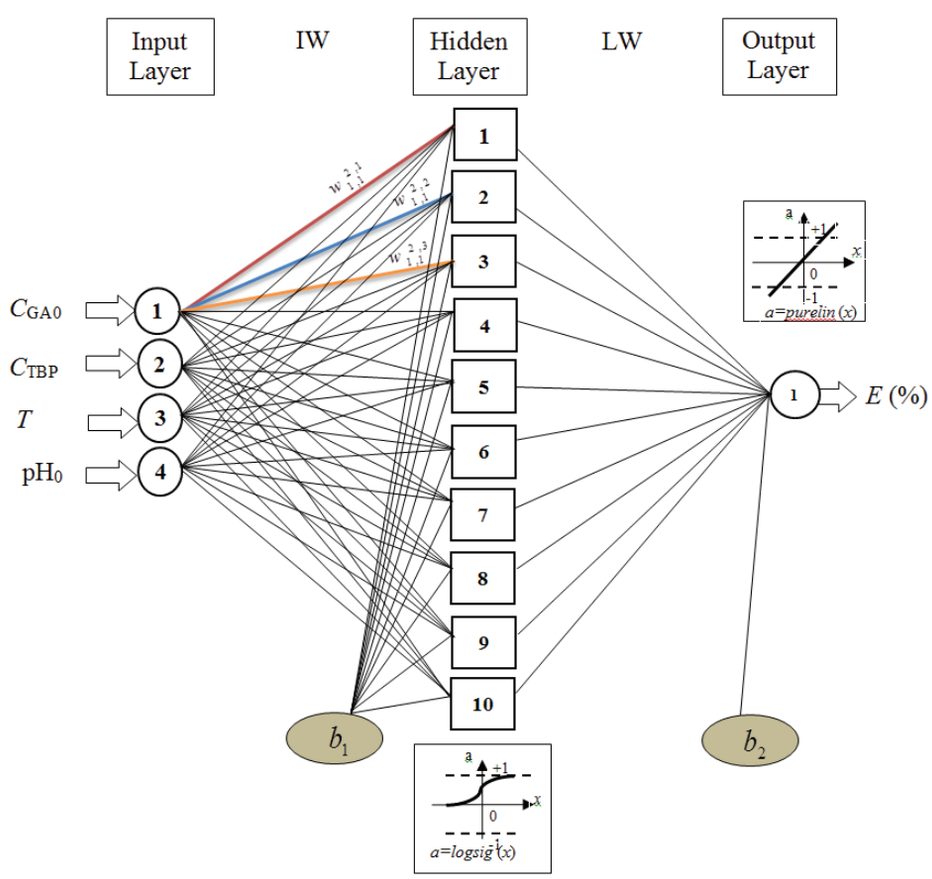



Fig. 4 – *ANN architecture used for the prediction of GA extraction efficiency*

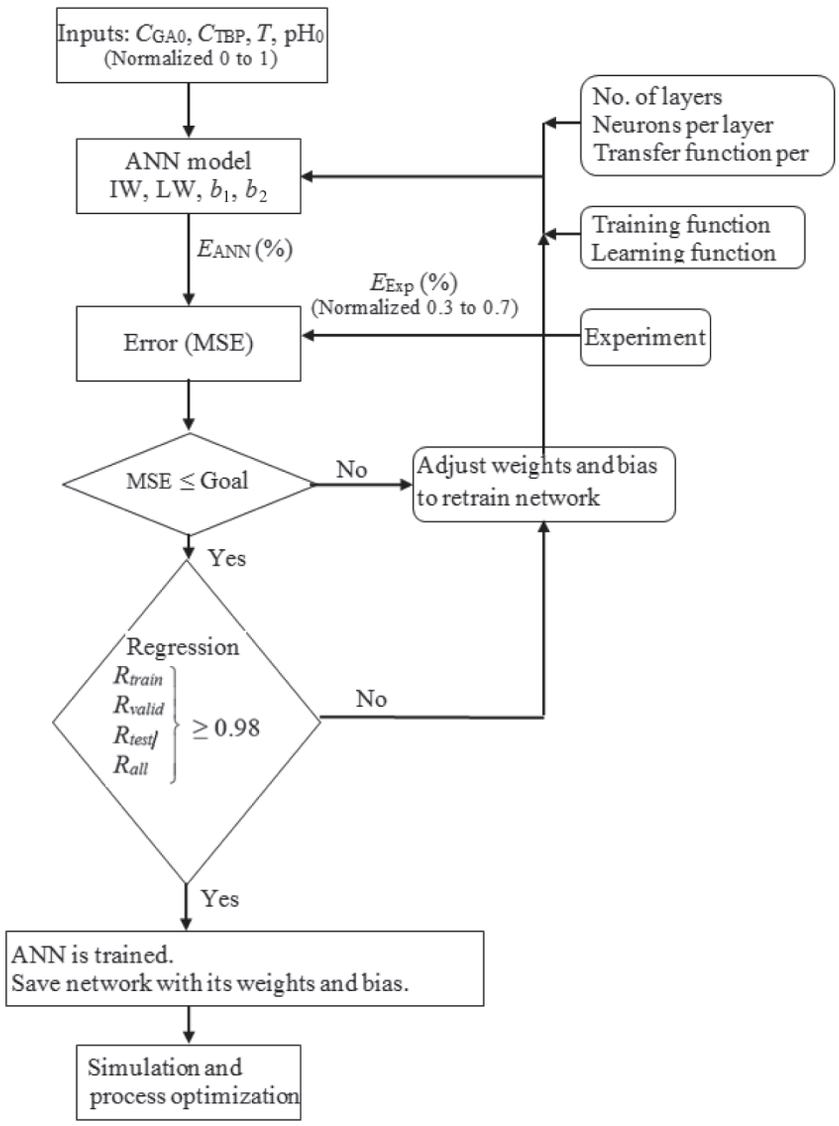



Fig. 5 – *Feed forward backpropagation algorithm used in the ANN training for the prediction of GA extraction efficiency*

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×A, $b_1$ = H×1, LW =O×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-

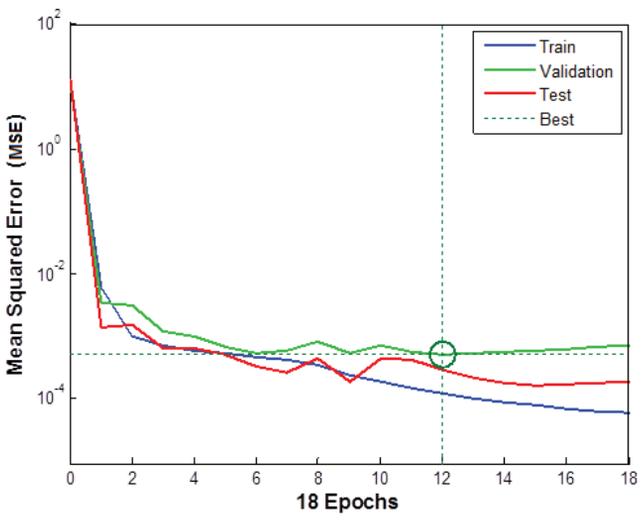



Fig. 6 – *MSE for training, validation, and test dataset computed using LMB (with performance of $5 \cdot 10^{-4}$ at 12 epochs and goal of $1 \cdot 10^{-3}$)*

Table 5 – *Optimal values of weights and biases obtained from ANN training using Levenberg-Marquardt backpropagation algorithm*

| | |
|---|---|
| Input weight matrix, IW (Weights to Hidden layer from inputs) | $\begin{vmatrix} 5.4485 & 7.4654 & -1.4374 & -2.0099 \\ -3.4801 & 1.0038 & -7.1242 & 0.6678 \\ -9.9763 & 4.6322 & 4.1336 & -7.9662 \\ 1.6333 & -4.4955 & -0.545 & -2.1418 \\ -4.5359 & -4.2326 & 5.6151 & 1.6171 \\ 8.043 & 2.7753 & 2.0607 & 0.6206 \\ 3.7705 & 3.0939 & 6.0152 & 7.2166 \\ 8.1829 & 11.8002 & -6.2271 & -3.8066 \\ 4.5814 & -0.8091 & 6.2607 & 2.5373 \\ 3.3117 & -4.7426 & 7.2301 & 0.8914 \end{vmatrix}$ |
| Bias vector (Hidden layer), $b_1$ | $\begin{vmatrix} -10.2912 & 11.4372 & 9.4153 & 0.4321 & 2.6838 & -5.0968 & -7.3116 & 0.1728 & -4.231 & 3.7916 \end{vmatrix}^T$ |
| Layer weight vector, LW (Weight to output layer from hidden layer) | $\begin{vmatrix} -0.1575 & -0.4295 & 0.0956 & -0.3633 & -0.2071 & -0.2006 & -0.1154 & 0.2645 & 0.2656 & 0.313 \end{vmatrix}$ |
| Bias scalar (output layer), $b_2$ | 0.6545 |

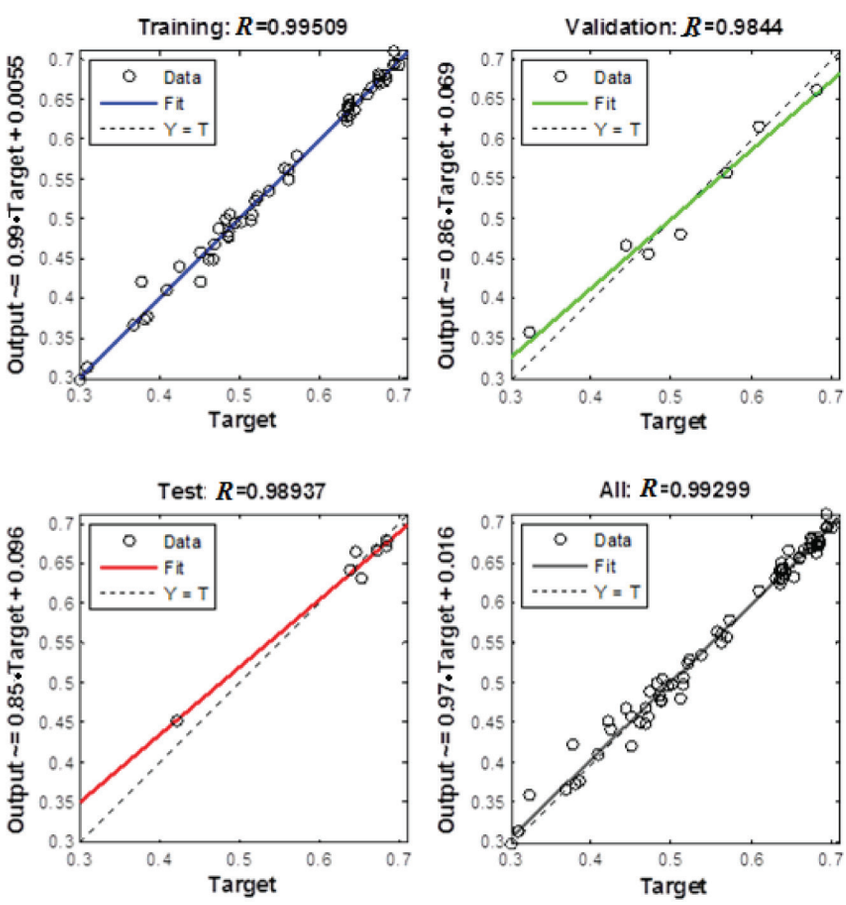



Fig. 7 – *ANN model showing the regression plots for training, validation, testing, and overall prediction set*

perimental extraction efficiency in all the cases. 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.

Table 6 – *Analysis of variance (ANOVA) analysis for ANN model*

| Source | DF[a] | SS[b] | MS[c] | *F*-value | *p*-value | $R^2$ | $R^2$(adj.) |
|---|---|---|---|---|---|---|---|
| Model or Regression | 60 | 1142.1 | 19.04 | 10.41 | 0.0004 | 0.986 | 0.88 |
| Residual Error | 9 | 16.46 | 1.83 | | | | |
| Total | 69 | 1158.7 | | | | | |

[a]Degree of Freedom, [b]Sum of Squares, [c]Mean Square

Table 7 – *Unseen experimental dataset used in the developed RSM and ANN models for the prediction of extraction efficiency of GA*

| Run | Independent variables | | | | Response, *E* (%) | | |
|---|---|---|---|---|---|---|---|
| | $C_{GA0}$ (g $L^{-1}$) | $C_{TBP}$ (% v/v) | *T*(°C) | $pH_0$ | Experimental | RSM | ANN |
| 1 | 1.6 | 15 | 11 | 1.74 | 90.57 | 90.15 | 91.40 |
| 2 | 2.5 | 30 | 40 | 3.0 | 90.05 | 88.36 | 90.37 |
| 3 | 2.5 | 23 | 37 | 2.34 | 88.88 | 87.23 | 88.53 |
| 4 | 2.5 | 37 | 27 | 2.11 | 92.26 | 93.23 | 92.27 |
| 5 | 2.5 | 45 | 27 | 1.85 | 95.25 | 95.31 | 94.78 |
| 6 | 2.5 | 23 | 37 | 2.34 | 87.98 | 87.23 | 88.53 |
| 7 | 1.4 | 30 | 40 | 3.01 | 85.33 | 89.67 | 87.36 |
| 8 | 1.8 | 20 | 32 | 2.21 | 87.45 | 88.05 | 86.63 |
| 9 | 1.0 | 30 | 24 | 2.6 | 92.56 | 91.69 | 92.57 |
| 10 | 1.8 | 36 | 32 | 2.1 | 93.02 | 92.85 | 93.89 |
| 11 | 1.0 | 30 | 40 | 1.5 | 88.83 | 91.18 | 89.08 |

Table 8 – *Comparison of RSM and ANN models for its generalization capability (using unseen data)*

| Parameters | RSM | ANN |
|---|---|---|
| $R^2$ | 0.61 | 0.915 |
| RMSE | 1.73 | 0.80 |
| % AARE | 1.43 | 0.66 |
| % SEP | 1.91 | 0.88 |

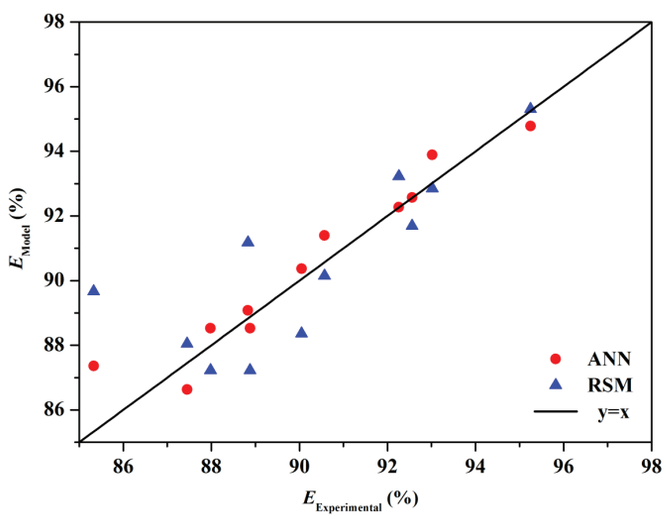



Fig. 8 – *Comparison of RSM and ANN model predictions for the unseen data (data not used in the model development)*

The ANN model was also validated using statistical analysis, ANOVA. The degree of freedom due to regression ($DF_{regression}$), and residuals ($DF_{residual}$) was calculated as follows:[20]

$$DF_{regression} = M - 1 \quad (17)$$

$$DF_{residual} = N - M \quad (18)$$

$$M = H(A + O + 1) + O \quad (19)$$

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 10.4 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:

$$y_n = purelin\ (LW \times logsig\ (IW \times x_n + b_1) + b_2) \quad (20)$$

## Comparison of ANN and RSM models

The prediction and generalization capabilities of RSM and ANN models were evaluated based on its $R^2$, RMSE, AARE, and SEP values. Table 1 and 4 indicate that both the models predict well the experimental extraction efficiency, and their $R^2$ values are closer to 1. Therefore, both models can be considered good for predicting the extraction efficiency of GA. The generalization capabilities of both the models were compared only with unseen dataset. Table 7 shows the unseen data set which were not used for the model development. Figure 8 shows the comparative parity plot for RSM and ANN model predictability for the unseen data set, and suggests that both the model predictions fit the experimental data very well. The comparative values of $R^2$, RMSE, AARE, and SEP values for both models using unseen data are shown in Table 8. Thus, higher $R^2$ and lower RMSE, AARE, SEP values for ANN indicate that the ANN model is superior over the RSM model for its generalization capability of the GA extraction process.

## 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 g $L^{-1}$, $C_{TBP}$ = 65.65 % (v/v), $T$ = 18.4 °C, and $pH_0$ = 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, modelling, 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.

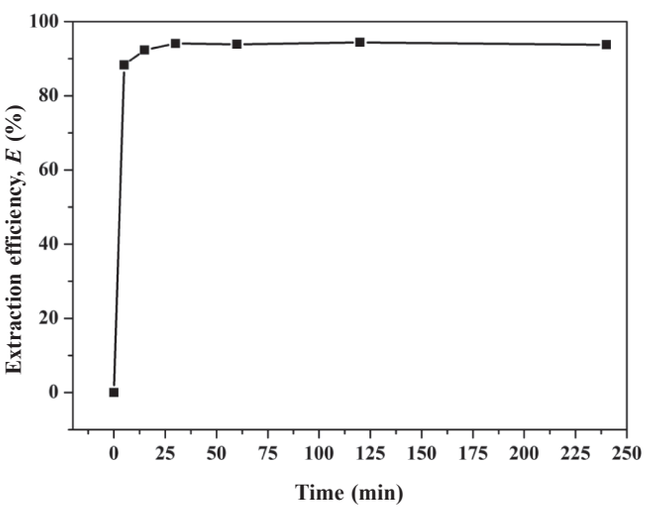



Fig. S1 – *Experimental extraction efficiency (E%) of gallic acid versus time for the operating conditions:* $C_{GA0}$ = 2 g $L^{-1}$, $C_{TBP}$ = 40 % (v/v), $T$ = 30 °C, $pH_0$ = 3

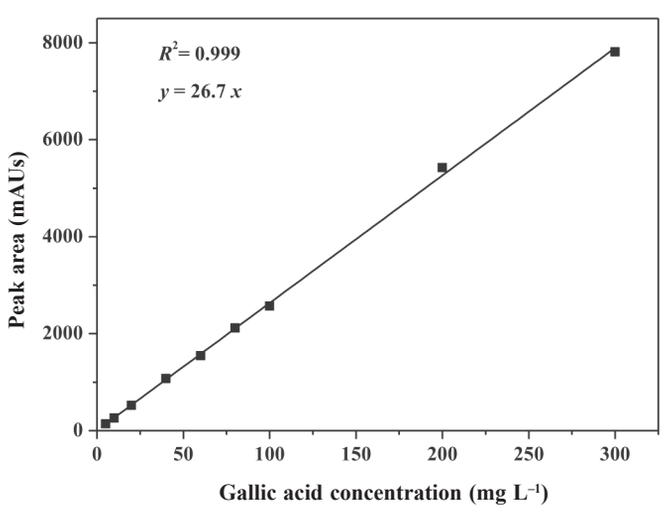



Fig. S2 – *Calibration curve for gallic acid obtained using HPLC*

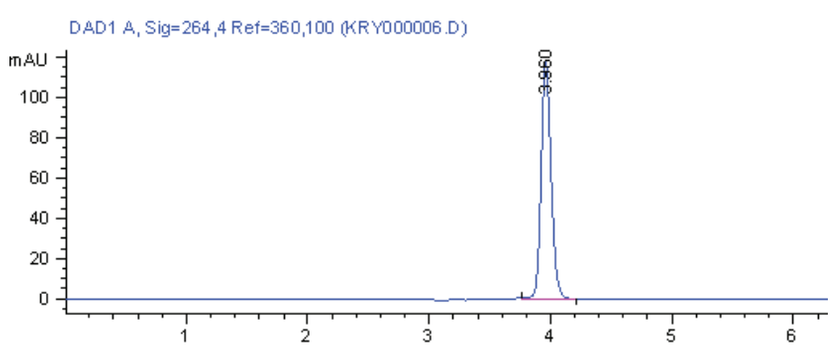



Fig. S3 – *HPLC Chromatogram of gallic acid at 264 nm and 3.96 min*

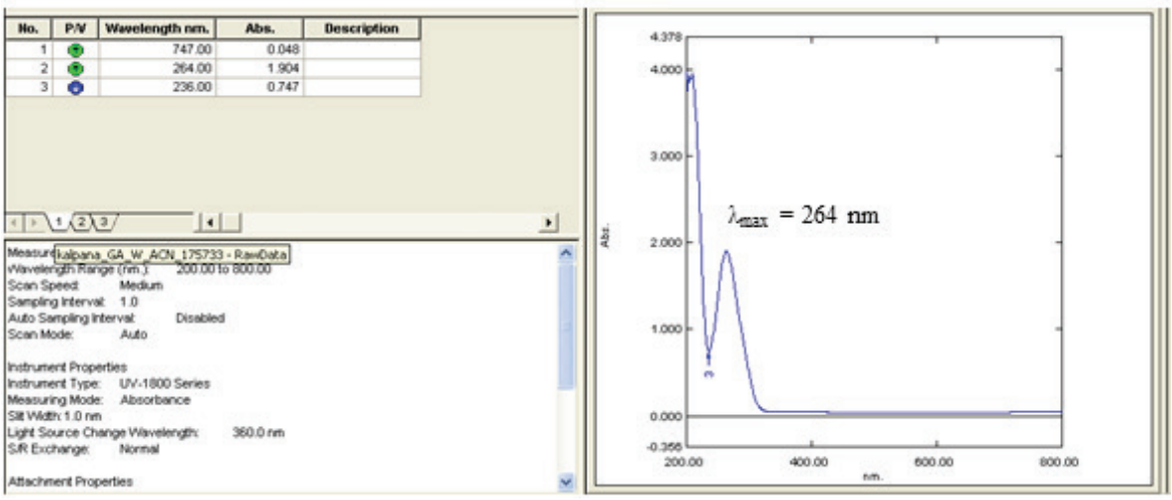



Fig. S4 – *Absorption spectra of gallic acid showing* $l_{max}$ *= 264 nm*

## 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
$n_c$ – corner points or factorial 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 vaiables in RSM
$x_{i,j}$, $y_{i,j}$, $b_{i,j}$ – network input, output and bias of $j^{th}$ neuron to $i^{th}$ layer, respectively
$\alpha$ – distance of each star point from centre in RO-CCD
$\beta_i$, $\beta_j$, $\beta_{ij}$ – regression coefficients representing linear, quadratic, and interactive effects in RSM

## References

1. *Benitez, F. J., Real, F. J., Acero, J. L., Leal, A. I., Garcia, C.,* Gallic acid degradation in aqueous solutions by UV/$H_2O_2$ treatment, Fenton's reagent and the photo-Fenton system, J. Hazard. Mater. **126** (2005) 31. doi: https://doi.org/10.1016/j.jhazmat.2005.04.040
2. *El-Gohary, F. A., Badawy, M. I., El-Khateeb, M. A., El-Kalliny, A. S.,* Integrated treatment of olive mill wastewater (OMW) by the combination of fenton's reaction and anaerobic treatment, J. Hazard. Mater. **162** (2009) 1536. doi: https://doi.org/10.1016/j.jhazmat.2008.06.098
3. *Herrero, M., Cifuentes, A., Ibanez, E.,* Sub- and supercritical fluid extraction of functional ingredients from different natural sources: plants, food-by-products, algae and microalgae A Review, Food Chem. **98** (2006) 136. doi: https://doi.org/10.1016/j.foodchem.2005.05.058
4. *Kougan, G. B., Tabopda, T., Kuete, V., Verpoorte, R.,* Simple phenols, phenolic acids, and related esters from the medicinal plants of africa, in: *Kuet, V.* (First Ed.), Medicinal plant research in Africa, Pharmacology and chemistry, Elsevier Inc., London, 2013, 225–249.

5. *Kapellakis, I. E., Tsagarakis, K. P., Crowther, J. C.*, Olive oil history, production and by-product management, Rev. Environ. Sci. Biotechnol. **7** (2008) 1. doi: https://doi.org/10.1007/s11157-007-9120-9
6. *Calace, N.*, *Nardi, E.*, *Petronio, B.*, *Pietroletti, M.*, Adsorption of phenols by papermill sludges, Environ. Pollut. **118** (2002) 315. doi: https://doi.org/10.1016/S0269-7491(01)00303-7
7. *Fki, I., Allouche, N., Sayadi, S.,* The use of polyphenolic extract, purified hydroxytyrosol and 3,4-dihydroxyphenyl acetic acid from olive mill wastewater for the stabilization of refined oils: A potential alternative to synthetic antioxidants, Food Chem. **93** (2005) 197. doi: https://doi.org/10.1016/j.foodchem.2004.09.014
8. *Ollis, D. F., Pelizzetti, E., Serpone, N.,* Photocatalyzed destruction of water contaminants, Environ. Sci. Technol. **25** (1991) 1522. doi: https://doi.org/10.1021/es00021a001
9. *Masten, S. J., Davies, S. H.,* The use of ozonation to degrade organic contaminants in wastewaters, Environ. Sci. Technol. **28** (1994) 180A. doi: https://doi.org/10.1021/es00053a001
10. *Spigno, G., Jauregi, P.,* Recovery of gallic acid with colloidal gas aphrons (CGA), Int. J. Food Eng. **1** (2005) 1. doi: https://doi.org/10.2202/1556-3758.1038
11. *Spigno, G., Dermiki, M., Pastori, C., Casanova, F., Jauregi, P.,* Recovery of gallic acid with colloidal gas aphrons generated from a cationic surfactant, Sep. Purif. Technol. **71** (2010) 56. doi: https://doi.org/10.1016/j.seppur.2009.11.002
12. *Quici, N., Litter, M. I.,* Heterogeneous photocatalytic degradation of gallic acid under different experimental conditions, Photochem. Photobiol. Sci. **8** (2009) 975. doi: https://doi.org/10.1039/b901904a
13. *Singh, M., Jha, A., Kumar, A., Hettiarachchy, N., Rai, A. K., Sharma, D.*, Influence of the solvents on the extraction of major phenolic compounds (punicalagin, ellagic acid and gallic acid) and their antioxidant activities in pomegranate aril, J. Food Sci. Technol. **51** (2014) 2070. doi: https://doi.org/10.1007/s13197-014-1267-0
14. *Claudio, A. F. M., Ferreira, A. M., Freire, C. S. R., Silvestre, A. J. D., Freire, M. G., Coutinho, J. A. P.*, Optimization of the gallic acid extraction using ionic-liquid-based aqueous two-phase systems, Sep. Purif. Technol. **97** (2012) 142. doi: https://doi.org/10.1016/j.seppur.2012.02.036
15. *Li, Y., Wang, Y., Li, Y., Dai, Y.*, Extraction of glyoxylic acid, glycolic acid, acrylic acid, and benzoic acid with trialkylphosphine oxide, J. Chem. Eng. Data **48** (2003) 621. doi: https://doi.org/10.1021/je020174r
16. *Daneshfar, A., Ghaziaskar, H. S., Homayoun, N.,* Solubility of gallic acid in methanol, ethanol, water, and ethyl acetate, J. Chem. Eng. Data **53** (2008) 776. doi: https://doi.org/10.1021/je700633w
17. *Wasewar, K. L., Shende, D. Z.*, Reactive extraction of caproic acid using tri- n -butyl phosphate in hexanol, octanol, and decanol, J. Chem. Eng. Data **56** (2011) 288. doi: https://doi.org/10.1021/je100974f
18. *Keshav, A., Chand, S., Wasewar, K. L.,* Reactive extraction of acrylic acid using tri- n -butyl phosphate in different diluents, J. Chem. Eng. Data **54** (2009) 1782. doi: https://doi.org/10.1021/je800856e
19. *Zhong, K., Wang, Q.*, Optimization of ultrasonic extraction of polysaccharides from dried longan pulp using response surface methodology, Carbohydr. Polym. **80** (2010) 19. doi: https://doi.org/10.1016/j.carbpol.2009.10.066
20. *Marchitan, N., Cojocaru, C., Mereuta, A., Duca, G., Cretescu, I., Gonta, M.*, Modeling and optimization of tartaric acid reactive extraction from aqueous solutions: A comparison between response surface methodology and artificial neural network, Sep. Purif. Technol. **75** (2010) 273. doi: https://doi.org/10.1016/j.seppur.2010.08.016
21. *Messikh, N., Samar, M. H., Messikh, L.*, Neural network analysis of liquid–liquid extraction of phenol from wastewater using tbp solvent, Desalination **208** (2007) 42. doi: https://doi.org/10.1016/j.desal.2006.04.073
22. *Montgomery, D.*, Design and analysis of experiments, (Fifth. Ed.), John Wiley & Sons Inc., New Jersey, 2001.
23. *Box, G. E. P., Hunter, J. S., H. W. G.*, Statistics for experimenters: design, innovation and discovery, (second ed.), Vol. 48, John Wiley &Sons, New Jersey, 2006.
24. *Dornier, M., Decloux, M., Trystram, G., Lebert, A.,* Dynamic modeling of crossflow microfiltration using neural networks, J. Memb. Sci. **98** (1995) 263. doi: https://doi.org/10.1016/0376-7388(94)00195-5
25. *Niemi, H., Bulsari, A., Palosaari, S.*, Simulation of membrane separation by neural networks, J. Memb. Sci. **102** (1995) 185. doi: https://doi.org/10.1016/0376-7388(94)00314-O
26. *Bourquin, J., Schmidli, H., van Hoogevest, P., Leuenberger, H.*, Pitfalls of artificial neural networks (ANN) modelling technique for data sets containing outlier measurements using a study on mixture properties of a direct compressed dosage form, Eur. J. Pharm. Sci. **7** (1998) 17. doi: https://doi.org/10.1016/S0928-0987(97)10027-6
27. *Cornell, A. I., Khuri, J. A.*, Response surfaces: Designs and analyses, (second ed.), Marcel Dekker, Inc., New York, 1996.
28. *Bezerra, M. A., Santelli, R. E., Oliveira, E. P., Villar, L. S., Escaleira, L. A.*, Response surface methodology (RSM) as a tool for optimization in analytical chemistry, Talanta **76** (2008) 965. doi: https://doi.org/10.1016/j.talanta.2008.05.019
29. *Hornik, K., Stinchcombe, M., White, H.*, Multilayer feedforward networks are universal approximators, Neural Networks **2** (1989) 359. doi: https://doi.org/10.1016/0893-6080(89)90020-8
30. *Hagan, M. T., Demuth, H. B., Beale, M. H.,* Neural network design; PWS Publishing Company: Boston, 1995.
31. *Keshav, A., Wasewar, K. L., Chand, S.,* Extraction of propionic acid using different extractants (tri-n-butylphosphate, tri-n-octylamine, and Aliquat 336), Ind. Eng. Chem. Res. **47** (2008) 6192. doi: https://doi.org/10.1021/ie800006r
32. *Yang, S. T., White, S. A., Hsu, S. T.,* Extraction of carboxylic acids with tertiary and quaternary amines: Effect of pH, Ind. Eng. Chem. Res. **30** (1991) 1335. doi: https://doi.org/10.1021/ie00054a040
33. *Keshav, A., Wasewar, K. L., Chand, S.,* Extraction of Acrylic, Propionic, and butyric acid using Aliquat 336 in oleyl alcohol: Equilibria and effect of temperature, Ind. Eng. Chem. Res. **48** (2009) 888. doi: https://doi.org/10.1021/ie8010337
34. *Uslu, H., Kırbaslar, S. I.*, Effect of temperature and initial acid concentration on the reactive extraction of carboxylic acids, J. Chem. Eng. Data **58** (2013) 1822. doi: https://doi.org/10.1021/je4002202
35. *Pham, D. T., Sagiroglu, S.,* Training multilayered perceptrons for pattern recognition: A comparative study of four training algorithms, Int. J. Mach. Tools Manuf. **41** (2001) 419. doi: https://doi.org/10.1016/S0890-6955(00)00073-0