



Ultrasound assisted green extraction of phenolic components from *Basella alba*

Gulshan Kumar¹ · T. Joshi Jayasree¹ · Rahul Kumar Rout¹ · P. Srinivasa Rao¹ · Padmavati Manchikanti²

Received: 22 January 2024 / Accepted: 8 May 2024 / Published online: 31 May 2024

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2024

Abstract

The investigation was conducted to analyze the effects of ultrasound-assisted extraction (UAE) on the bioactive compounds from *Basella alba* leaves. The experiment used response surface methodology (RSM) to assess the independent variables, which included solvent concentration (50 to 100%), treatment time (5 to 15 min), and amplitude percentage (70 to 90%). The solvent used was ethanol. The extraction yield of response variables was predicted using the RSM and ANN-GA approach. The experimental values of total flavonoid content (TFC), total phenolic content (TPC), and antioxidant activity (AA) ranged from 4.13 to 8.12 mgQEg⁻¹, 5.30–8.23 mgGAEg⁻¹, and 0.29–0.63 mgGAEACg⁻¹ respectively. For the optimized conditions obtained using RSM, the predicted values of TFC, TPC, and AA were 8.31 mgQEg⁻¹, 8.23 mgGAEg⁻¹, and 0.59 mgGAEACg⁻¹, respectively, while the predicted values using ANN-GA were 8.52 mgQEg⁻¹, 7.64 mgGAEg⁻¹, and 0.67 mgGAEACg⁻¹, respectively. ANN-GA exhibited a high R² value and reduced values for other statistical parameters compared to RSM. Moreover, the predicted ANN-GA value closely matched the experimental value. Thus, it can be considered that ANN-GA is the most optimal model for the optimization and modeling of the UAE process pertaining to the bioactive components derived from *B. alba* leaf extracts. The findings of the study will aid in efficiently utilizing sustainable techniques for extracting bioactive compounds from *B. alba* leaves and improving their industrial applications, particularly in the food and therapeutic industries.

Keywords Antioxidant activity · *Basella alba* · TFC · TPC · Ultrasound-assisted extraction

Introduction

Basella alba is an underutilized tropical leafy vegetable commonly found in India and Indonesia. The plant has high nutritional value and medicinal properties [13]. The leaves are high in phytochemicals like polyphenols, flavonoids, and antioxidants [21]. According to Shankul [20], the antioxidant potential of *B. alba* is greater than that of spinach and moringa leaves. Aside from being consumed as a leafy vegetable, the bioactive potential of *B. alba* has

yet to be adequately investigated in terms of extraction, processing, or utilization in food formulations. Extraction is the primary process in analyzing the bioactive compounds present in leaves, and hence, identifying appropriate extraction methods is critical in increasing the yield of bioactive components. The methods of extraction include conventional methods like soxhlet extraction, maceration, and hydro-distillation and novel techniques like ultrasound-assisted extraction (UAE), microwave-assisted extraction (MAE), enzyme-assisted extraction (EAE), pulsed electric field extraction (PEFE), pressurized liquid extraction (PLE), cold plasma assisted extraction and supercritical fluid extraction (SCFE). The primary drawback associated with conventional methods is their lengthy processing time, requirement for high-purity solvents, low rate of solvent recovery, and high cost of extraction [18, 22].

UAE is a highly efficient and promising green technique that can be automated, thereby streamlining and optimizing processes to a great extent. UAE can help in attaining the objective of sustainable green chemistry and extraction

✉ T. Joshi Jayasree
jayasreejoshi48@gmail.com

¹ Agricultural and Food Engineering Department, Indian Institute of Technology Kharagpur, Kharagpur, West Bengal 721302, India

² Faculty of Interdisciplinary Sciences and Engineering, Rajiv Gandhi School of Intellectual Property Law, Indian Institute of Technology Kharagpur, Kharagpur, West Bengal 721302, India

through its environmentally friendly and energy-efficient approaches [3]. UAE has been observed to enhance extraction yields and process kinetics in comparison to conventional extraction methodologies, along with a reduction in extraction time, energy consumption, and environmental impacts. Moreover, the advancement of environmentally friendly technology has been motivated by the increasing public interest in chemical-free substances and the prevention of harmful effluents. Potential challenges associated with implementing and applying these technologies, such as capital investment and operational costs, must be considered [8]. UAE works on the principle of ultrasonication, which applies sound wave at a frequency > 20 kHz, resulting in acoustic cavitation. The cellular matrix of the plant will undergo various changes such as fragmentation, localized erosion, pore creation, shear force, enhanced absorption, and swelling index due to the collapse of cavitation bubbles and the influence of sound waves. This process can lead to the disintegration of the cellular structure into smaller fragments, thereby reducing the particle size and increasing the overall surface area [11]. Consequently, these changes facilitate higher mass transfer rates within the solid matrix boundary layer, enabling efficient solubilization of the bioactive component in the solvent [17]. Ultrasound technique has facilitated the total extraction of phenolic compounds within a short span of time, giving high reproducibility, less solvent consumption, and high purity yield with energy efficiency [3]. In addition to the extraction method used, the bioactive recovery is influenced by the characteristics of the solvent, its concentration during extraction, the feed-to-solvent (F:S) ratio, sonication amplitude, duty cycle, and extraction time. Alves et al. [2] reported that the UAE of bioactive components from *Monteverdia aquifolia* leaves required less time, temperature, and F:S ratio than the conventional method. Kadam et al. [9] also reported that UAE enhanced the recovery and yield of bioactive compounds from *Asco-phylum nodosum* (brown seaweed). They obtained extraction yields of 143.12 mgGAE g_{db}⁻¹, 87.06 mg g_{db}⁻¹, and 128.54 mg g_{db}⁻¹ for total phenolics, fucose, and uronic acid, respectively, at optimized conditions of treatment time (25 min), acid concentration (0.03 M HCl) and amplitude (114 μm).

RSM is a widely employed mathematical and statistical approach used in the modeling and optimization of the bioactive compound extraction process. RSM not only characterizes the influence of independent variables but also captures their interaction effects. Numerous studies have employed RSM technique to optimize the extraction of polyphenolic compounds from various leaf sources, such as oregano leaves [18], mandarin leaves [19], and *B.alba* leaves [21], and have observed its efficacy in process modelling and optimization. AAN is a predictive modeling tool that is becoming increasingly important in recent times. This tool

utilizes a multilayer perceptron, consisting of an input and output layer of processing units called nodes interconnected through one or more hidden layers. The model is developed through a training process, where the input parameters of a set of experimental records are presented to the input layer of the network. These values are then multiplied by a weighting factor and transmitted to the nodes of the initial hidden layer. The combined and transformed values from the initial layer are subsequently integrated and modified through a transfer function. The resulting output is weighted and passed on to the subsequent layer. The outcomes generated by the output layer provide predictions for the dependent variables of the model [15]. GA is an optimization technique rooted in the principle of “survival of the fittest” and facilitates the acquisition of optimum solutions using a well-defined fitness function. Hybrid ANN-GA technique has been reported to be effective in optimizing the process parameters of extraction of bioactive components from plant materials [21]

The potential impact of ultrasound-assisted extraction (UAE) conditions on the bioactive components of *B. alba* leaves has not yet been thoroughly investigated and examined. Consequently, a deliberate effort was undertaken to explore and examine the influence and effect of the UAE conditions on the extraction process and recovery of the bioactive components from the leaves of *B. alba*. This was followed by the implementation of process optimization techniques, specifically the RSM and ANN-GA methodologies, to enhance and optimize the extraction process of the bioactive components from the *B. alba* leaves. These findings will help in using sustainable methods for extracting bioactive compounds from *B. alba* leaves and enhance their applications, especially in the food and therapeutic sectors.

Materials and methods

Raw material

Fresh *B. alba* leaves procured from the local market in Kharagpur, West Bengal, India, was used for the study. The leaves were cleaned and microwave dried at 700 W to a moisture content of 7–8% (wb) in a Microwave-UV-US combined reactor system (Nu tech Analytical Instruments Pvt Ltd, India [13]). The dried leaves were ground into fine powder and sieved through a BS 36 sieve to obtain a uniform particle size, which was used for further analysis.

Chemicals and reagents

The chemicals and reagents used in the investigation were 2, 2-diphenyl-1-picrylhydrazyl (DPPH), 2,4,6-tris(2-pyridyl)-s-triazine (TPTZ), acetic acid, acetone, aluminum chloride, ethanol, ferric chloride, folin–ciocalteau reagent

(FCR), hydrochloric acid, methanol, sodium acetate, sodium carbonate, sodium hydroxide, and sodium nitrite acquired from Himedia Laboratories (Mumbai, India) and Merck specialties private limited (Mumbai, India).

Ultrasound-assisted extraction

The UAE was carried out using a probe sonicator (ATP 500, Athena Technologies Pvt. Ltd., India) with a power output of 500 W and a frequency of 40 kHz. The experiments were carried out at varying solvent concentrations, treatment times, and amplitude percentages. The solvent used was 99% ethanol. These solvents were selected based on their extraction capacity for a wide range of raw materials [5]. In accordance with the regulations of the US Food and Drug Administration, ethanol is classified as a solvent of class 3 and can be used safely within the prescribed limits [18]. Feed to solvent ratio was fixed at 1:15 based on the previous study [21]. After the extraction process, the samples were centrifuged at 6000 rpm for 10 min. This was subsequently followed by the filtration of the samples via Whatman No. 4 filter paper. The resultant extract was collected and stored in a refrigerated environment until further analysis.

Experimental design

A three-level Box—Behnken design with 5 central points was used to optimise the UAE conditions. The independent variables of the study were solvent concentration, extraction duration, and amplitude. The limits (upper and lower) of each independent variable have been identified through preliminary experiments. The independent variables are presented in Table 1, along with their coded representations. The various experimental combinations (17 runs) are listed in Table 2. The impact of the independent factors on the responses, namely TPC, GAEAC, and TFC, was mathematically modelled utilizing the RSM and the ANN-GA methodology.

Table 1 The limits of independent variables used for the study

Independent variable	Terminology	Units	Values		
			(− 1)	0	(+1)
Solvent concentration	A	%	50	75	100
Treatment time	B	Min	5	10	15
Amplitude percentage*	C	%	70	80	90

*Percentage of maximum power (500 W)

Biochemical analysis of extracts

Total phenolic content (TPC)

TPC determination was carried out according to Singleton et al. [23] with minor modifications. The aforementioned procedure facilitates the conversion of the reaction mixture from a yellow hue to a blue shade through the interaction between phenolic compounds and FCR in an alkaline environment. Redox reagent FCR reacts with polyphenols to form phospho-tungstic-phospho-molybdenum complex. In brief, 0.2 mL of the extract was mixed with 2.8 mL of distilled water, 0.5 mL of FCR (10 times diluted), and 0.75 mL Na_2CO_3 (20% w/v) and incubated in the dark for 90 min. A UV–Vis spectrophotometer was used to measure absorbance at 760 nm wavelength. Gallic acid was used as the standard, and TPC was calculated as mg of gallic acid equivalents (mgGAEg^{-1} sample) per dry weight of the sample [18].

Gallic acid equivalent antioxidant capacity (GAEAC)

In a test tube, 0.2 mL of the extract was mixed with 2.85 mL pure methanol and 0.95 mL DPPH solution. Subsequently, the test sample was incubated for 30 min in the absence of light. Employing a UV–Vis spectrophotometer, the absorbance at a specific wavelength of 517 nm was determined. For calibration, pure methanol was used. Gallic acid was used as the reference substance [13].

Total flavonoids content (TFC)

TFC was measured by taking 0.5 mL of diluted extract in a test tube, followed by the addition of 2 mL of distilled water and 0.15 mL of NaNO_2 . To this mixture, 0.15 mL of a 10% AlCl_3 solution was added to it after 4–5 min. Following that, 1 mL of 1 M NaOH was added after waiting 5 min, then 1.2 mL of distilled water. The sample was incubated for 10 min. Using a UV–Vis spectrophotometer, the absorbance was determined at 510 nm wavelength. TFC was calculated as mg quercetin equivalents per g dry weight sample weight (mgQEg^{-1} sample) using quercetin as the standard [18].

RSM modelling and optimisation

Process optimization was carried out in order to maximize the values of the response parameters. In the study, the empirical regression equation (Eq. 1) was used to examine the dependent parameters and predict the optimal values.

Table 2 Actual, RSM-predicted, and ANN-predicted values of the response variables

Run	Solvent conc. (%)	Extraction Time (min)	Amplitude (%)	TPC (mgGAEg ⁻¹)		TFC (mgQEg ⁻¹)		AA (mgGAEACg ⁻¹)				
				Exp. value	RSM pred	ANN pred	Exp. value	RSM pred	ANN pred	Exp. value	RSM pred	ANN pred
1	50	10	70	6.331	6.100	6.319	4.139	1.922	4.216	0.416	0.407	0.417
2	75	15	70	7.564	7.771	7.523	7.487	6.882	7.250	0.564	0.546	0.573
3	75	5	70	6.559	6.558	6.505	6.643	5.570	6.439	0.590	0.623	0.607
4	100	10	70	5.303	5.329	5.300	5.578	3.399	5.499	0.611	0.605	0.618
5	100	5	80	5.546	5.021	5.530	5.578	3.827	5.668	0.627	0.595	0.633
6	75	10	80	8.227	7.037	8.228	7.785	7.710	7.770	0.556	0.549	0.550
7	50	5	80	6.840	7.072	6.884	4.728	3.015	4.709	0.497	0.478	0.493
8	100	15	80	7.844	7.612	7.846	8.124	5.904	8.048	0.585	0.604	0.583
9	75	10	80	8.176	8.037	8.168	7.761	7.710	7.670	0.551	0.549	0.548
10	75	10	80	8.176	8.037	8.168	7.761	7.710	7.670	0.551	0.549	0.560
11	75	10	80	8.176	7.037	8.168	7.761	7.710	7.670	0.551	0.549	0.560
12	50	15	80	6.840	6.865	6.868	4.544	2.363	4.412	0.285	0.317	0.248
13	75	10	80	7.428	8.037	7.428	7.481	7.710	7.670	0.535	0.549	0.520
14	50	10	90	6.345	6.019	6.431	4.323	1.682	4.147	0.387	0.384	0.356
15	100	10	90	6.055	6.286	6.007	7.160	4.558	7.148	0.572	0.591	0.546
16	75	15	90	8.087	8.088	8.139	6.922	6.741	6.912	0.562	0.529	0.589
17	75	5	90	7.625	7.418	7.665	7.279	6.629	7.188	0.586	0.604	0.588

$$y_j = \beta_0 + \sum_{i=1}^3 \beta_i x_i + \sum_{i=1}^3 \beta_{ii} x_i^2 + \sum_{i < j=1}^3 \sum_{i < j=1}^3 \beta_{ij} x_i x_j \quad (1)$$

where y_j is the response variable, β_0 , β_i , β_{ii} and β_{ij} are the coefficients of regression, and x_i & x_j are the coded form of independent parameters. Statistical software, Design Expert 11 (Stat-Ease Inc., Minneapolis, USA) was employed for analysis. The validation of the model was ensured by conducting experimental tests on the optimal conditions for UAE in triplicate.

ANN modelling and GA optimization

The process of ANN modeling was implemented by utilizing the neural network fitting tool provided by MATLAB (version R2020a, The Math Works, Inc., USA). The ANN model composed of three layers, namely the input layer, hidden layer, and output layer (Fig. 1). The number of neurons present in the input layer was taken as three (solvent concentration [X_1], extraction time [X_2], and amplitude percentage [X_3]). Hidden layer was having ten neurons and the output layer had a single neuron representing the response variables. Each response, variable was trained individually. The transfer function for the hidden layer was tansig function and for the output layer was purelin function. The Levenberg–Marquardt (LM) back propagation technique, specifically the trainlm algorithm, was employed to train the data. The training dataset consisted of 70% of the total data, while 15% each was allocated for validation and testing purposes. The training continued until the network achieved the optimal R^2 and MSE metrics values. Later on, the output values

were computed using the weights and bias values obtained from the given equation.

$$y_i = \text{purelin}\{W_{OH} \times \text{tansig}\{U_{IH} \times x_i + TH\} + TO\} \quad (2)$$

where x_i is the input parameter; y_i is the output parameter; W_{OH} and U_{IH} are the weights between different layers; TO and TH are the bias values of neurons.

The process of optimization using the Genetic Algorithm (GA) was conducted by utilizing the GA toolbox of MATLAB (version R2015a, The Math Works, Inc., USA). The GA optimization primarily focused on selecting specific parameters, namely the population type, which was represented as a double vector, the population size set at a value of 200, and the crossover fraction designated as 0.8. Furthermore, several additional factors were taken into consideration during the GA optimization process. These factors included the migration aspect, the creation function, the fitness scaling function, the selection function, the crossover function, and the mutation function. Specifically, forward migration, feasible population, rank, roulette, scattered, and adaptive feasibility were chosen for each respective factor. These selections were determined based on the research conducted by Abdullah et al. [1]. Additionally, a fitness function, denoted as ' f ,' was formulated with the intention of maximizing all output responses. The specific representation of this fitness function is provided below.

$$f = -(Y_1 + Y_2 + Y_3) \quad (3)$$

where Y_1 , Y_2 , and Y_3 indicate the predicted values of responses.

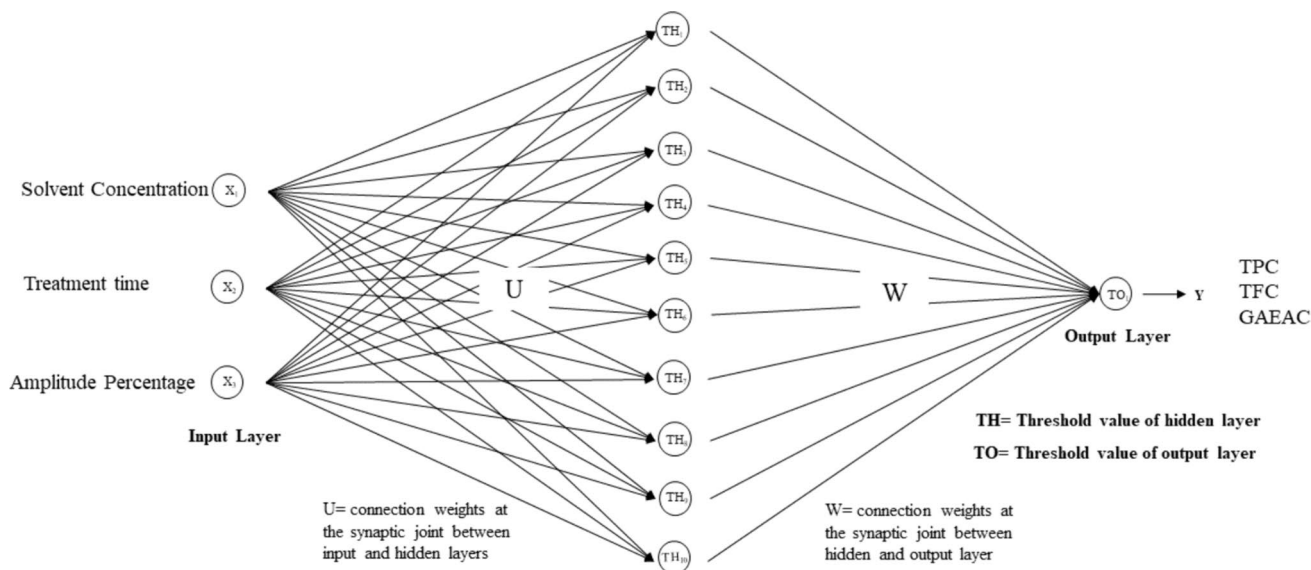


Fig. 1 Proposed ANN model and its various layers

Statistical analysis

The statistical evaluation of the developed models was conducted using various metrics as specified in Eqs. 4, 5, 6, 7, 8, 9, 10 [7, 21]. In order to identify the best model for precisely representing the responses, the model with the lowest values for AAD, MSE, NMSE, RMSE, and NRSME, and the highest value of R^2 was chosen.

$$R^2 = 1 - \frac{\sum (X_p - X_a)}{\sum (X_p - X_a)^2} \tag{4}$$

$$AAD = \frac{\sum |X_p - X_a|}{n} \tag{5}$$

$$MSE = \frac{\sum (X_p - X_a)^2}{n} \tag{6}$$

$$NMSE = \frac{MSE}{X_m} \tag{7}$$

$$MPE = \frac{100}{n} \sum \left| \frac{X_p - X_a}{X_p} \right| \tag{8}$$

$$RMSE = \sqrt{\frac{\sum (X_p - X_a)^2}{n}} \tag{9}$$

$$NMRSE = \frac{RMSE}{X_m} \tag{10}$$

where x_p represents predicted data, x_a represents experimental data, x_m represents mean experimental data, and n is the number of experiments.

Results & discussion

Effect of process parameters on TPC

The changes in TPC during UAE are presented in Table 2. TPC values varied from 5.303 to 8.227 mgGAEg⁻¹ in the extracted samples. The effect of independent variables on UAE of TPC from *B. alba* leaves is shown in RSM plots (Fig. 2). The highest value of TPC (8.227 mgGAEg⁻¹) was obtained at 80% amplitude, 75% concentration, and extraction time of 10 min. The standard curve for the estimation of TPC & TFC was shown in Fig. 4. Regression analysis of the obtained data (Table 3) shows that the linear

term of extraction time and amplitude percentage had a significant effect ($p < 0.05$) on TPC. The multiple linear regression equation (MLR) equation obtained showing the effect of process variables on TPC is presented in Eq. 11. The solvent concentration and extraction time interaction also showed a significant effect on TPC. The increase in TPC with increasing treatment time could be attributed to an increase in microbubble formation, which increased cell damage and resulted in the release of more phenolic compounds [10]. The extraction yield of TPC increased initially with amplitude percentage, after which it reduced. A similar trend has been reported by Alves et al. [2] during the UAE of *Monteverdia aquifolia* leaves. The change in amplitude influences several factors, including the intensification and/or restriction of sonochemical reactions, polymerization and/or depolymerization, chemical bond rupture, and intensification of undesirable interactions between extracted compounds, favouring the production of oxidizing agents, resulting in the degradation of phenolic compounds and limiting their antioxidant capacity [4]. Solvent concentration had significant negative quadratic effects ($p < 0.05$) on TPC. The TPC value increased with an increase in solvent concentration initially (up to 70%), after which it declined. Li et al. [14] noticed a similar trend in the UAE for phenolic chemicals extracted from *Perilla* leaves using ethanol as a solvent. They reported that higher water concentrations in ethanol–water combinations were shown to extract a greater proportion of polar phenolic chemicals. The model exhibited an insignificant lack of fit value [$p > 0.05$] with satisfactory prediction ($R^2 = 0.94$, Adj. $R^2 = 0.88$).

$$\begin{aligned} \text{TPC (mgGAEg}^{-1}\text{)} = & 8.04 - 0.20A + 0.47B \\ & + 0.29C + 0.57AB + 0.18AC \\ & - 0.14BC - 1.36A^2 \\ & + 0.09B^2 - 0.67C^2 \end{aligned} \tag{11}$$

Total flavonoids content (TFC)

The values of TFC ranged from 4.139 to 8.124 mgQEg⁻¹ in extracted samples. The changes in TFC with different processing conditions are presented in Table 2. All linear components (such as time, amplitude, and solvent concentration) had a highly significant ($p < 0.01$) effect on TFC, in accordance with the regression coefficients presented in Table 3. TFC showed a positive correlation with solvent concentration, and the values increased with time at higher concentrations. The solvent concentration and extraction time interaction also showed a significant positive effect ($p < 0.05$) on TFC. Solvent concentration and amplitude percentage showed significant negative quadratic effects ($p < 0.05$) on TFC. Extraction time had an

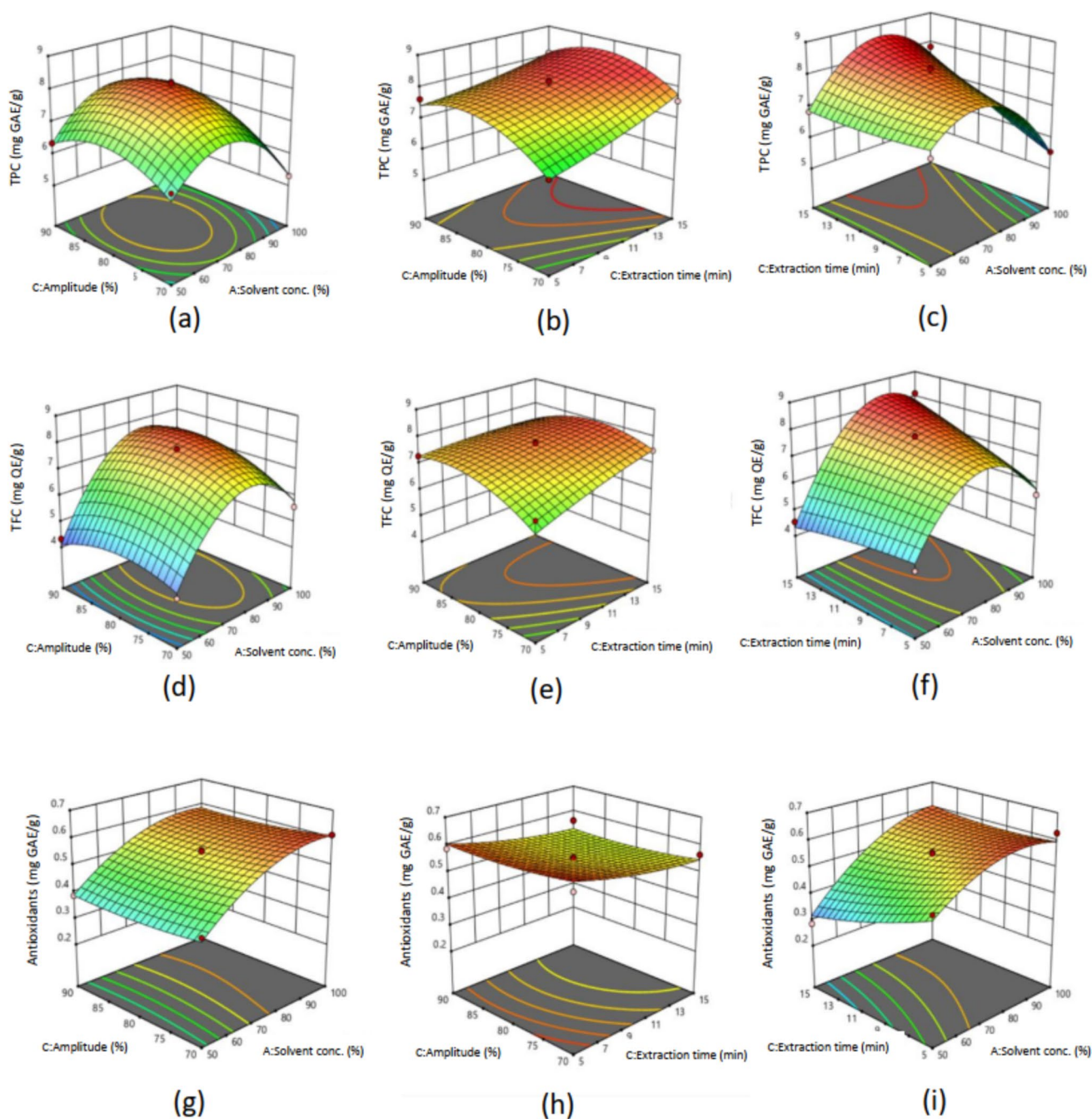


Fig. 2 Response surface plots showing the influence of process variables on **a–c** TPC, **d–f** TFC, and **g–i** AA

insignificant linear effect on the extraction of TFC. The MLR equation obtained showing the effect of process variables on TFC is presented in Eq. 12.

$$\begin{aligned}
 TFC = & 7.71 + 1.09A + 0.36B + 0.23C \\
 & + 0.68AB + 0.35AC - 0.30BC \\
 & - 1.87A^2 - 0.09B^2 - 0.53C^2
 \end{aligned}
 \quad (12)$$

TFC increased with solvent concentration and amplitude percentage and then decreased. Increasing the amplitude of ultrasonic waves frequently leads to an augmentation in the sonochemical impacts, resulting in a greater yield in terms of TFC [3]. The decrease in TFC at 80% amplitude could be attributed to the acceleration of intracellular mass transfer and the degrading effect of ultrasound at elevated output power [6].

Table 3 Regression coefficients and other statistical parameters of the RSM model

Coefficient	TPC	GAEAC	TFC
β_0	8.03	0.54	7.7
β_1	- 0.2	0.10	1.08
β_2	0.47	- 0.03	0.35
β_3	0.29	- 0.01	0.22
$\beta_1 \beta_2$	0.57	0.04	0.68
$\beta_1 \beta_3$	0.18	- 0.002	0.34
$\beta_2 \beta_3$	- 0.13	0.001	- 0.3
β_1^2	- 1.36	- 0.06	- 1.87
β_2^2	0.09	0.01	- 0.09
β_3^2	- 0.66	0.01	- 0.53
Model (F-value)	14.61	15.03	27.23
R ²	0.9495	0.9508	0.9722
Adj. R ²	0.8845	0.8876	0.9365
C.V. (%)	4.65	4.61	4.39
Lack of fit (p-value)	0.520	0.301	0.110

Gallic acid equivalent antioxidant capacity (GAEAC)

GAEAC values ranged from 0.285 to 0.627 mgGAEg⁻¹ in the samples extracted under different treatment combinations. The highest value of GAEAC (0.627 mgGAEg⁻¹) was obtained at a solvent concentration of 100%, a lower extraction time of 5 min, and an amplitude of 80%. Regression analysis demonstrated that the linear component of the solvent concentration and the extraction process's duration had a notable impact (p < 0.05) on GAEAC. Significantly negative quadratic effects (p < 0.05) were observed between solvent concentration and GAEAC. The model exhibited an insignificant lack of fit value (p > 0.05) with satisfactory prediction (R² = 0.95, Adj. R² = 0.88). The multiple linear regression (MLR) equation obtained showing the effect of process variables on GAEAC is presented in Eq. 13. The samples also observed a decrease in GAEAC values with higher treatment time. This is attributed to prolonged extraction time induced by detrimental oxidation reactions [12, 16].

$$\begin{aligned}
 GAEAC = & 0.55 + 0.10A - 0.04 B \\
 & - 0.009C + 0.04AB - 0.003AC \\
 & + 0.0005BC - 0.06465A^2 \\
 & + 0.01B^2 + 0.01C^2
 \end{aligned}
 \tag{13}$$

RSM optimisation

RSM was employed to optimize the extraction of bioactive compounds from *B. alba* leaves. The values of the responses analyzed under different ultrasonic-assisted extraction (UAE) conditions are presented in Table 2. The regression coefficients obtained for each model and other statistical parameters are summarized in Table 3. Additionally, Table 4 provides other relevant statistical parameters. The model's accuracy was confirmed by its high R² value. The adjusted R² values are very similar to the R² values, affirming a profound correlation between the predicted and actual values. Moreover, the coefficient of variation (CV) remained below 5% in every instance, thereby substantiating the reproducibility and precision of the model.

ANN modeling

Model fitting

The study used the LM algorithm to model the dependent variables, specifically TPC, TFC, and AA. In order to achieve this, the ANN framework employed 3, 10, and 1 neuron(s) in the input layer, hidden layer, and output layer, respectively. Figure 1 depicts a visual representation of this architecture. The best model was chosen based on the minimum MSE and maximum R² value criterion. The validation performance for TFC, TPC, and AA, which were the factors under consideration, demonstrated the best results at epochs 4, 5, and 6, respectively. In order to ensure the highest level of rigor in the validation process, the 17 datasets were meticulously divided into three distinct sets during the experiment. This partitioning allowed for an accurate assessment of the models' performance. The training data set for TPC consisted of

Table 4 Summary of the developed model's statistical parameters

Parameter	TPC		TFC		AA	
	RSM	ANN	RSM	ANN	RSM	ANN
AAD	0.164	0.010	1.204	0.100	0.015	0.010
MSE	0.223	0.001	2.419	0.014	0.003	0.001
RMSE	0.473	0.035	1.555	0.119	0.019	0.016
NMSE	0.031	0.005	0.370	0.002	0.001	0.001
NRMSE	0.066	0.001	0.238	0.018	0.036	0.031
MPE	2.588	0.030	3.551	1.630	3.089	2.887
R ²	0.949	0.964	0.970	0.978	0.97	0.987

runs 4, 5, 6, 7, 8, 10, 11, 13, 14, 15, and 17, encompassing a comprehensive range of data points. The R^2 values, representing the coefficient of determination, were calculated for the training, validation, and testing stages, resulting in values of 0.973, 0.931, and 0.964, respectively. Furthermore, the MSE values for these runs were determined to be 0.062, 0.056, and 0.030, providing additional insight into the accuracy of the models. Similarly, runs 2, 3, 4, 7, 8, 9, 11, 12, 14, 16, and 17, comprising 11 observations, were used for training in the case of TFC. A total of 3 observations were allocated for validation, while another three observations were designated for testing. The corresponding R^2 values obtained for the training, validation, and testing sets were recorded as 0.999, 0.945, and 0.975, respectively. Moreover, the MSE values were calculated to be 0.039, 0.0326, and 0.464, respectively. Employing a similar approach, runs 1, 3, 5, 6, 7, 8, 11, 12, 13, 15, and 17 were employed for training in the modeling of AA. Again, three observations were allocated for validation, while three observations were set aside for testing. The R^2 values obtained for the training, validation, and testing of AA were measured to be 0.851, 0.998, and 0.987, respectively. The corresponding MSE values were found to be 0.029, 0.020, and 0.007, respectively. Figure 3 depicts the performance and error histogram of the developed model. The optimization of the ANN model was carried out by comparing the output to the input data. During the optimization procedure, weights were used to minimize the error function. Equations 14, 15, 16, 17, 18, 19, 20, 21 show the values of weights and bias of the ANN model (Fig. 4).

$$U_a = \begin{vmatrix} 0.9152 & 1.8404 & 2.2074 \\ -0.4139 & 2.9872 & 0.0545 \\ -1.9691 & -1.8980 & -1.2719 \\ 0.1437 & -2.9160 & -0.7574 \\ -0.8037 & -0.3515 & -2.8858 \\ -1.8087 & 2.1159 & 1.1615 \\ -0.3770 & 2.7250 & -1.2369 \\ -0.3313 & -0.3246 & 2.9803 \\ 0.3005 & -0.09788 & -2.8371 \\ 1.4639 & -2.1134 & 1.5773 \end{vmatrix} \quad (14)$$

$$U_b = \begin{vmatrix} -1.9170 & 2.3105 & -0.1996 \\ -1.6195 & 2.2493 & -0.1267 \\ -0.1049 & -0.8751 & -2.7861 \\ 0.4369 & -2.2135 & -2.1410 \\ 0.8438 & 0.4724 & -2.8853 \\ -1.9486 & 1.3766 & 1.8252 \\ -2.8187 & -0.7150 & 1.4813 \\ -1.8039 & -0.0612 & -2.5164 \\ 1.1874 & -1.6173 & 2.0010 \\ -0.5843 & 2.7436 & 1.1119 \end{vmatrix} \quad (15)$$

$$U_c = \begin{vmatrix} -0.0194 & -2.6831 & 1.3014 \\ 1.7138 & -2.5459 & -0.4548 \\ -1.6759 & -0.6537 & -2.6249 \\ -2.4119 & 1.3666 & 0.8751 \\ -1.2534 & 2.0736 & 1.7947 \\ 0.7769 & 2.5889 & -1.3248 \\ -0.9260 & 2.0808 & 1.7977 \\ 0.7927 & 1.9759 & -2.1328 \\ -2.3907 & 1.7696 & -0.5702 \\ 1.4252 & 2.0209 & -1.3769 \end{vmatrix} \quad (16)$$

$$W_a = |0.25320.07750.30100.4533 \\ -0.81100.7551 - 0.9713 \\ -0.4114 - 0.64020.8526| \quad (17)$$

$$W_b = |0.16790.3876 - 0.3662 \\ -0.3264 - 0.4711 - 0.4878 \\ -0.81690.16640.0332 - 0.1497| \quad (18)$$

$$W_c = |-0.41680.5613 - 0.51510.3661 \\ -0.10710.3107 - 0.6472 \\ -0.21420.12310.2701| \quad (19)$$

$$TH_a = \begin{vmatrix} -3.0162 \\ 2.3459 \\ 1.6757 \\ -1.0054 \\ 0.3351 \\ -0.3351 \\ -1.0054 \\ -1.6757 \\ 2.3459 \\ 3.0162 \end{vmatrix} \quad TH_b = \begin{vmatrix} 3.0593 \\ 2.6877 \\ 1.7362 \\ -0.9382 \\ -0.2928 \\ -0.3825 \\ -0.5757 \\ -1.5204 \\ 2.4399 \\ -3.0067 \end{vmatrix} \quad TH_c = \begin{vmatrix} -3.0242 \\ -2.2567 \\ 1.4163 \\ 1.1592 \\ 0.3401 \\ 0.0391 \\ -0.9678 \\ 1.6666 \\ -2.3317 \\ 3.1870 \end{vmatrix} \quad (20)$$

$$TO_a = |-0.8636| \quad TO_b = |-0.4281| \quad TO_c = |0.4131| \quad (21)$$

GA optimization

With the aim of increasing the extraction efficacy of TFC, TPC, and AA, the GA optimization process was conducted. The process of optimization was repeated until the MSE and RSME values between the average and individual fitness values were substantially lowered. After the mutation, the optimization cycle continued, and if the intended solution was not achieved, the entire population that was obtained was once again used for reproduction, crossover, and mutation in subsequent cycles. After 96 iterations, the optimal parameters for the UAE were identified to be: solvent concentration of 63.01%, treatment period of 8.92 min, and amplitude

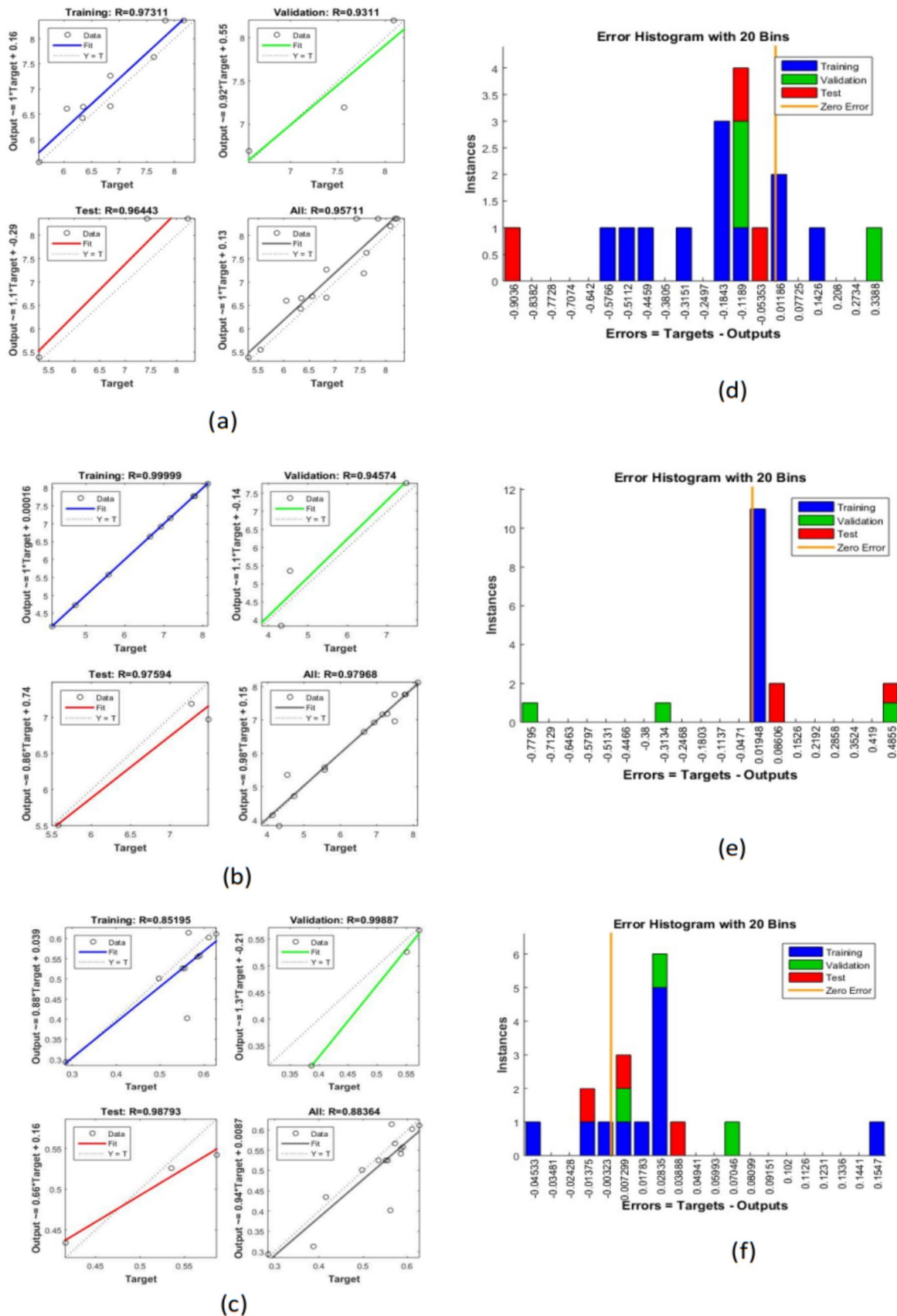


Fig. 3 Performance plot after training and error graph of a, d TPC, b, e TFC, and c, f AA of generated ANN model

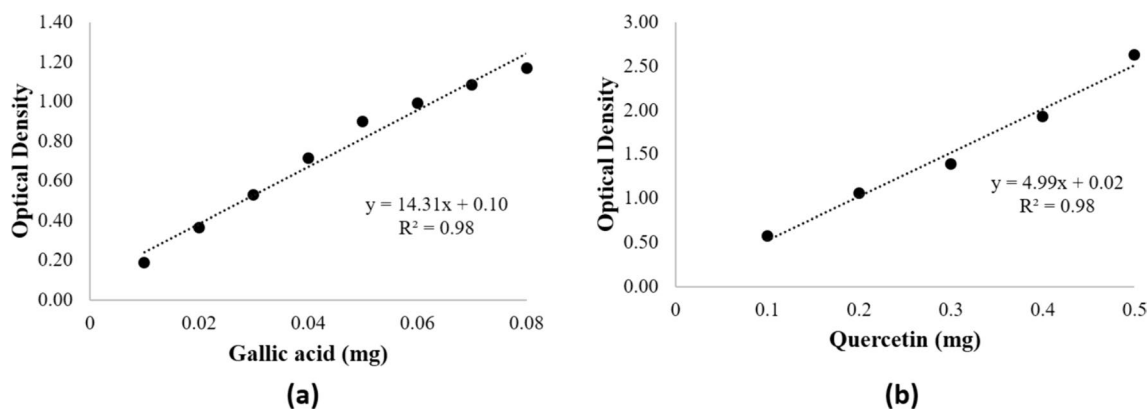


Fig.4 Standard curves for the estimation of **a** TPC and **b** TFC

Table 5 Optimal UAE conditions for maximum yield of bioactives from *B. alba* leaves

Techniques	Optimum conditions			TFC (mgQEg ⁻¹)		TPC (mgGAEg ⁻¹)		AA (mgGAEACg ⁻¹)	
	A	B	C	Pred	Expt	Pred	Expt	Pred	Expt
RSM	91.99	15	80.13	8.227	8.614	8.313	7.96	0.593	0.621
ANN	63.01	8.92	90	7.636	7.601	8.519	8.67	0.674	0.671

percentage of 90. The outcomes predicted by the developed ANN model for the response variables can be observed in Table 5. Subsequently, the experimentally tested optimal process variables were determined. Upon measurement, the obtained responses were determined to be 7.601 mgGAEg⁻¹ of TPC, 8.67 mgQEg⁻¹ of TFC, and 0.671 mgGAEg⁻¹ of AA. It is noteworthy that these experimental findings align well with the expected outcomes, thereby showcasing the precise prediction and optimization capabilities of the ANN-GA model.

Conclusion

UAE is gaining popularity due to its efficacy in phytochemical recovery in less time and low-cost approaches. The technique works through the formation and collapse of cavitation bubbles, resulting in the breakdown and easy release of the cellular matrix to solvent. The influence of UAE conditions on the recovery and quantification of bioactive components from *B. alba* leaves was examined in this work. RSM and the ANN-GA approach were used to investigate, predict, and optimize the influence of UAE conditions on bioactive components. TFC, TPC, and AA, which were extracted using RSM, exhibited predicted values of 8.31 mgQEg⁻¹, 8.23 mgGAEg⁻¹, and 0.59 mg GAEACg⁻¹, respectively. On the other hand, the predicted values for TFC, TPC, and AA obtained through ANN-GA were 8.52 mgQEg⁻¹, 7.64 mgGAEg⁻¹, and 0.67 mgGAEACg⁻¹, respectively. Both models demonstrated a high level of predictability.

However, the ANN model outpaced the RSM model due to its lower values of AAD, NSME, MSE, RSME, NRSME, and MPE, as well as its higher R^2 score. Accordingly, it can be concluded that the ANN-GA is a viable quantitative tool for optimizing the process variables for the UAE of *B. alba*. With continued technological developments in the UAE that involve the development of sophisticated machinery, integration with process modeling, and scaling up capabilities, its commercial applications are expected to expand. This will enable the development of more efficient and environmentally friendly alternatives to traditional extraction methods. Furthermore, the findings will help in widening the use of *B. alba* in various food formulations and pharmaceutical applications.

Data availability Not applicable.

References

1. S. Abdullah, R.C. Pradhan, D. Pradhan, S. Mishra, Modeling and optimization of pectinase-assisted low-temperature extraction of cashew apple juice using artificial neural network coupled with genetic algorithm. *Food Chem.* **339**, 127862 (2021)
2. T.P. Alves, C.C. Triques, P.A. Palsikowski, C. da Silva, M.L. Fiorese, E.A. da Silva, M.R. Fagundes-Klen, Improved extraction of bioactive compounds from *Monteverdia aquifolia* leaves by pressurized-liquid and ultrasound-assisted extraction: yield and chemical composition. *J. Supercrit. Fluids* **181**, 105468 (2022)

3. F. Chemat, N. Rombaut, A.G. Sicaire, A. Meullemiestre, A.S. Fabiano-Tixier, M. Abert-Vian, Ultrasound-assisted extraction of food and natural products: mechanisms, techniques, combinations, protocols, and applications—a review. *Ultrason. Sonochem.* **34**, 540–560 (2017)
4. A.B. Das, V.V. Goud, C. Das, Extraction of phenolic compounds and anthocyanin from black and purple rice bran (*Oryza sativa* L.) using ultrasound: a comparative analysis and phytochemical profiling. *Ind. Crops Prod.* **95**, 332–341 (2017)
5. Q.D. Do, A.E. Angkawijaya, P.L. Tran-Nguyen, L.H. Huynh, F.E. Soetaredjo, S. Ismadji, Y.H. Ju, Effect of extraction solvent on total phenol content, total flavonoid content, and antioxidant activity of *Linnophila aromatica*. *J. Food Drug Anal.* **22**(3), 296–302 (2014). <https://doi.org/10.1016/j.jfda.2013.11.001>
6. B.B. Ismail, M. Guo, Y. Pu, W. Wang, X. Ye, D. Liu, Valorization of baobab (*Adansonia digitata*) seeds by ultrasound-assisted extraction of polyphenolics: optimisation and comparison with conventional methods. *Ultrason. Sonochem.* **52**, 257–267 (2019)
7. J. Joshi, P.S. Rao, Characterization and multivariate analysis of decortication-induced changes in pearl millet. *J. Food Compos. Anal.* **125**, 105788 (2024)
8. T.J. Joshi, S.M. Singh, P.S. Rao, Novel thermal and non-thermal millet processing technologies: advances and research trends. *Eur. Food Res. Technol.* **249**(5), 1149–1160 (2023)
9. S.U. Kadam, B.K. Tiwari, T.J. Smyth, C.P. O'Donnell, Optimization of ultrasound assisted extraction of bioactive components from brown seaweed *Ascophyllum nodosum* using response surface methodology. *Ultrason. Sonochem.* **23**, 308–316 (2015)
10. M. Kazemi, R. Karim, H. Mirhosseini, A.A. Hamid, Optimization of pulsed ultrasound-assisted technique for extraction of phenolics from pomegranate peel of Malas variety: punicalagin and hydroxybenzoic acids. *Food Chem.* **206**, 156–166 (2016)
11. A.C. Khanashyam, M.A. Shanker, P. Thorakkattu, T.J. Joshi, R. Pandiselvam, A. Kothakota, Applications of ultrasound technology for inactivation of food enzymes, in *Enzyme inactivation in food processing*, ed. by M.R. Goyal, J.A. Malik, R. Pandiselvam (Apple Academic Press, New Jersey, 2023), pp.201–237
12. H. Kiziltas, Determination of LC-HRMS profiling, antioxidant activity, cytotoxic effect and enzyme inhibitory properties of *Satureja avromanica* using in vitro and in silico methods. *Process Biochem.* **116**, 157–172 (2022)
13. G. Kumar, J. Joshi, P.S. Rao, P. Manchikanti, Effect of thin layer drying conditions on the retention of bioactive components in Malabar spinach (*Basella alba*) leaves. *Food Chem. Adv.* **3**, 100419 (2023). <https://doi.org/10.1016/j.focha.2023.100419>
14. H.Z. Li, Z.J. Zhang, J. Xue, L.X. Cui, T.Y. Hou, X.J. Li, T. Chen, Optimization of ultrasound-assisted extraction of phenolic compounds, antioxidants and rosmarinic acid from perilla leaves using response surface methodology. *Food Sci. Technol.* **36**, 686–693 (2016)
15. A.P. Plumb, R.C. Rowe, P. York, M. Brown, Optimisation of the predictive ability of artificial neural network (ANN) models: a comparison of three ANN programs and four classes of training algorithm. *Eur. J. Pharm. Sci.* **25**(4–5), 395–405 (2005)
16. M.V. Rao, A.S. Sengar, C.K. Sunil, A. Rawson, Ultrasonication-A green technology extraction technique for spices: a review. *Trends Food Sci. Technol.* **116**, 975–991 (2021)
17. E. Roselló-Soto, M. Koubaa, A. Moubarik, R.P. Lopes, J.A. Saraiva, N. Boussetta, N. Grimi, F.J. Barba, Emerging opportunities for the effective valorization of wastes and by-products generated during olive oil production process: non-conventional methods for the recovery of high-added value compounds. *Trends Food Sci. Technol.* **45**(2), 296–310 (2015)
18. R.K. Rout, A. Kumar, P.S. Rao, A multivariate optimization of bioactive compounds extracted from oregano (*Origanum vulgare*) leaves using pulsed mode sonication. *J. Food Meas. Charact.* **15**(4), 3111–3122 (2021)
19. S. Şahin, Z. İlbağ, Şİ Kırbaslar, Pulsed ultrasound-assisted extraction of natural antioxidants from mandarin (*Citrus deliciosa Tenore*) leaves: experimental and modeling study. *Chem. Eng. Commun.* **205**(6), 717–726 (2018)
20. K. Shankul, Systematic pharmacognostical, phytochemical and pharmacological review on an ethno medicinal plant, *Basella alba* L. *J. Pharmacogn. Phytother.* **5**(4), 53–58 (2013)
21. A.S. Shende, T. Jayasree Joshi, P. Srinivasa Rao, Microwave-assisted aqueous extraction of bioactive components from Malabar spinach (*Basella alba*) leaves and its optimization using ANN-GA and RSM methodology. *J. Food Meas. Charact.* (2023). <https://doi.org/10.1007/s11694-023-02182-2>
22. A.S. Shende, T.J. Joshi, P. Srinivasa Rao, Process optimization of microwave-assisted aqueous extraction of tannins and saponins from Malabar Spinach (*Basella alba*) leaves using ANN-GA and RSM methodology. *Meas. Food* (2024). <https://doi.org/10.1016/j.meaf.2023.100117>
23. V.L. Singleton, R. Orthofer, R.M. Lamuela-Raventós, Analysis of total phenols and other oxidation substrates and antioxidants by means of Folin-Ciocalteu reagent, in *Methods in enzymology*, ed. by D. Jameson, M. Verma, H.V. Westerhoff (Academic press, Cambridge, 1999), pp.152–178

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.