Chapter-4

To Comparison of response surface methodology (RSM) and artificial neural network (ANN) modelling for supercritical fluid extraction of phytochemicals from Terminalia chebula pulp and optimization using RSM coupled with desirability function (DF) and genetic algorithm (GA) and ANN with GA

4.1. Introduction

In developing nations, access to health care originating from the significant expenses of medications, health services and diagnostics, has become a significant matter of concern. *Terminalia chebula Retz*. *(T. chebula)* often called as haritaki in Sanskrit native to South Asia, is considered as a source of nutrients having a medical advantage and is available in the tropic territories in the world referred to as *Chebulae Fructus* (Hezi) in China [21]. It has traditionally been utilized for its therapeutic properties in various bioactive compounds since ancient times to fix geriatric diseases and to improve memory [3]. It has been accounted that haritaki organic products are highly rich in phytochemicals. To get a high return of phenolic components from haritaki, it is important to build up an efficient extraction technique. Currently, natural herbal antioxidants from various fruits are gaining attention for their potential helpfulness. For prospective usage in functional foods or nutraceuticals, the bioactive components from haritaki may be extracted. For the extraction of phenolic mixtures from haritaki, many extraction techniques have been used, including subcritical water extraction and reflux framework coupled with water-ethanol and water-propylene glycol [23].

In any case, because of high temperature and long treatment periods, the antioxidant activities of natural products usually reduce in traditionally extraction techniques. Many authors have featured the advantages of utilizing supercritical fluid extraction to diminish process energy, reduce manpower, and increase shelf life. Supercritical CO_2 extraction is a green, new approach that successfully employed for selective extraction of comparable chemicals and offers strong selectivity for lipophilic/non-polar or mildly polar compounds [16, 28]. For the extraction of bioactive components from herbs and other products, organic solvent extraction and steam distillation are traditionally used. These procedures are characterised by a labor-intensive process, a lengthy extraction duration, a low yield, toxic solvent residue, and the degradation of chemicals that are sensitive to temperature. The use of the supercritical fluid extraction (SFE) technique can eliminate these drawbacks [10, 26]. It is a fast evolving technique to make bioactive chemicals using only technology and benign settings. The most common solvent in SFE is carbon dioxide since it is medically inert, environmentally safe, non-explosive, and easily available [29].

Many researchers have successfully extracted phenolics from fruits using supercritical fluid extraction, demonstrating the technology's ecological suitability and usefulness as a complementary safe solvent for extracting food-grade bioactive components from agricultural products. One important thermodynamic benefit of using supercritical liquid is the simple isolation from extracted solutes that can be achieved by simply changing the temperature. As compared to organic solvents, supercritical fluids have liquid-like densities and provide improved mass exchange characteristics. They are classified as low-viscosity and high diffusivity fluids. The lower surface tension allows supercritical fluids to penetrate the porous biological matrix while removing solutes [25]. Various approaches like response surface methodology (RSM) or artificial neural network (ANN) can be used for modelling of the process and efficient optimization of the process parameters of extraction for maximum recovery of the phytochemicals from plant materials. RSM is a collection of statistical methods for planning experiments, constructing models, assessing the impact of processing parameters on responses, and optimizing processes [27]. Artificial neutral (ANN) is the complex mathematical modelling commonly used to mimic the biological neural networks and processes information. ANN is used to optimize and model complex biological process and highly non-linear data and overcome the problem uneasy for human or statistical methods. Its acceptance in food processing is growing rapidly day by day for data modelling due to its advanced design to manage complex biological data and non-linear data. It has properties such as noise resistance, multi-nonlinear variable accommodation, parallel processing power, Capability for approximation of universal function, and strong generalization efficiency [8, 20, 22]. For data fitting and prediction, a well-trained ANN model can be used. For the clamorous data, ANNs have been used for optimization and prediction and are often favoured over regression models. The basic behaviour of neural computations comes from connecting neurons in a network. ANN is showing superiority over the RSM, unpredicted nonlinear data, fuzzy inputs, and subtle patterns under certain conditions [1, 33]. Genetic algorithm (GA) is an optimization method that can be used even if a full model of the process is not available. GA is based on Darwin's genetic evolution theory and employs genetic operators such as selection, mutation, and crossover to find the best solution to problems. In many practical applications, the combination of ANN and GA (ANN-GA) has been used for optimization [1, 12, 30].

Therefore, the present study was taken up with the aim to model the process of supercritical fluid extraction by different methods viz. RSM and ANN, to optimize the process parameters of supercritical fluid extraction on the extractability of phytochemicals from dried *T. chebula* pulp using various approaches viz. RSM coupled with desirability function, RSM coupled with GA and ANN coupled with GA, and compare them.

4.2. Materials and methods

4.2.1. Preparation of dried haritaki pulp powder

The mature *T. chebula* fruits were collected from the Horticulture section of Tezpur University. The fruit was cleaned, and the pulp and seed were separated. A laboratory tray dryer (Labotech, BDI-51, B. D. Instrumentation, Ambala, India) was used to dry the pulp at a temperature of 40 °C. The dry pulp was crushed and put through a 100 mesh screen before being sealed in polythene bags with an aluminium laminate until further usage. All the reagents used in the present study were of analytical grade.

4.2.2. Extraction of phytochemicals by supercritical CO_2

The supercritical fluid extractor vessel (Applied Separations, USA) was filled with the powdered haritaki pulp (1g), and the extracts were then collected in glass tubes in a separator. According to the experimental design, the extraction was carried out under varied extraction circumstances. The extraction co-solvent utilised was a 1:1 mixture of ethanol and water, and the constant mass flow rate of CO_2 during all tests was 5 kg/h.

Carbon dioxide was pressurised using a high-pressure pump and then charged into the extraction vessel at the necessary pressure. According to the experimental plan, the extraction vessel containing the sample was allowed to heat in the oven while a thermocouple recorded the temperature. The extract was placed in a glass vial after the dissolved chemicals in supercritical CO_2 were forced through a heated micrometre valve and expanded at room temperature and pressure. For each extraction experiment, the extractor was charged with a steady supply of CO_2 at a rate of 5 mL/min. The internal flow metering mechanism of the SFE maintained a constant flowrate.

4.2.3. Experimental design

The experiments were designed using a central composite rotatable design (CCRD) with four numerical factors (independent and dependent parameters of CCRD in

Table 4.1). The numerical variables included temperature (X_1), pressure (X_2), time (X_3), and co-solvent flow rate (X_4). Temperature, pressure, time, and flow rate were all adjusted to range from 40 to 60 °C, 100 to 200 bar, 40 to 80 min, and 1 to 5 mL/min, respectively. 30 tests in total were conducted (**Table 4.2**). To minimize the impact of external variables, all experiments were conducted in a randomized order [31].

Table 4.1. Independent and Dependent parameters of CCRD

Sl. No.	Independent parameters	Dependent parameters
1.	Time (40, 50, 60, 70 and 80 min)	Total phenolic content
2.	Temperature (40, 45, 50, 55 and 60 °C)	Total flavonoid content
3.	Pressure (100, 125, 150, 175 and 200 bar)	DPPH radical scavenging activity
4.	Flow rate (1, 2, 3, 4 and 5 mL/min)	

Dependent variables were TPC, TFC and DPPH free radical scavenging activity.

4.2.4. Determination of bioactive compounds

The total phenolic content was calculated as discussed in section 3.2.8.2, the total flavonoid content was calculated as discussed in section 3.2.8.3, and the total antioxidant activity was calculated as discussed in section 3.2.8.4.

4.2.5. RSM modelling

For response surface methodology, Design Expert version 8 was used to analyze and model data for responses. The experimental data was fitted to a second order polynomial model:

$$Y = \beta_o + \sum_{i=1}^{4} \beta_i X_i + \sum_{i=1}^{4} \beta_{ii} X_i^2 + \sum \sum_{i<j=1}^{4} \beta_{ij} X_i X_j \quad (4.1)$$

Where Y represents the responses TPC, TFC and DPPH radical scavenging activity, $\beta_{o,}$ is the constant, β_i, β_{ii} and β_{ij} are the regression coefficients and X_i and X_j are the independent variables in coded values. An analysis of variance was used for model analysis (ANOVA). Lack-of-fit test and R^2 (coefficient of determination) was calculated for determining the adequacy of the model. To determine the relative dispersion of the experimental points from the model's prediction, the coefficient of variation (CV) was

calculated. Response surfaces were generated to study the effect of interactions on the responses.

4.2.6. ANN modelling

The neural network fitting tool of Matlab (Matlab 9.6 – R2019a, The Mathworks Inc., MA, USA) was employed for modelling of experimental data by ANN that was produced during extraction of phytochemicals of haritaki pulp. In ANN, Multilayer perceptron (MLP) is the broadly used technique for data modelling. MLP comprises of feed-forward (FF) backpropagation (BP) of algorithm input layer in neural network, hidden layers, and output layer. There can be one or more neurons contained in each layer. The number of hidden layer neurons was varied from 8 to 12 for the training of data for development of various neural network models. The number of neurons in the hidden layers determine the performance of the ANN model as a very small number of neurons in the hidden layers may limit the ability of the ANN to properly model the process and the network may not train well, whereas too many neurons might make the network memorizing the data rather than training it [2]. Sigmoid transfer function was used between the input layer and hidden layer as well as between the hidden layer and output layer. Trial and error method were used to obtain the best combination of hidden layer and transfer function (minimum error condition) for the provided data [11, 22, 6]. Training of ANN models was done until the error reaches the minimum between experimental and predicted values of responses. For training of data set, the Levenberg-Marquardt training algorithm was used. The weights and bias are all together known as neural network parameters. The trained network model was validated using validation data set (experimental data which was not used for training). The development of the ANN model was carried out by dividing the data set in three groups: 70% for training, 15% for testing and 15% for validation. The weight values of the synaptic joints between the input and hidden layer and that between the hidden and output layer were calculated by well versed ANN model for optimization of the parameters.

4.2.7. Analysis of the developed models

The performances of the models developed by RSM and ANN were compared statistically by calculating coefficient of determination (R^2) when intercept is zero, mean absolute error (MAE), root mean square error (RMSE) and Chi-square (χ^2) values [27] as follows:

$$R^2 = \frac{\sum_{i=1}^{n}(y_{pi})^2}{\sum_{i=1}^{n}(y_{oi})^2} \quad (4.2)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_{pi} - y_{oi}\right| \quad (4.3)$$

$$RMSE = \sqrt[2]{\frac{\sum_{i=1}^{n}(y_{pi}-y_{oi})^2}{n}} \quad (4.4)$$

$$\chi^2 = \sum_{i=1}^{n}\frac{(y_{pi}-y_{oi})^2}{y_{pi}} \quad (4.5)$$

where, n is the number of data points, y_{pi} is the predicted value and y_{oi} is the observed value.

4.2.8. Optimization of the process

Three different approaches were used for optimization of the optimization using desirability function (RSM-DF), RSM couples with GA (RSM-GA) and ANN coupled with GA (ANN-GA). The numerical optimization technique of the "Design-Expert" software was used for RSM-DF approach. For RSM-GA and ANN-GA approaches, the models generated through RSM and ANN were optimized by employing genetic algorithm toolbox in MATLAB. For the ANN-GA and RSM-GA optimization, the population size used was 200. The criteria selected for optimization was maximization of all the responses i.e., TPC, TFC and DPPH.

4.3. Results and discussion

4.3.1. RSM modelling and Effect of process parameters and their interaction on the responses

Extraction of phytochemicals from haritaki was investigated with the application of CCRD based response surface methodology with four independent variables, viz. flow rate (X_1), pressure (X_2), temperature (X_3), and time (X_4) along with TPC, TFC, and DPPH scavenging activity as dependent variables. **Table 4.2** represents the experimental and predicted values of runs along with dependent variables while **Table 4.3** displayed the response of ANOVA for all the dependent variables. ANOVA was conducted to evaluate the model significance and the result showed that all the models were significant with a level of $p< 0.05$. In the case of TPC, X_1, X_2, X_3, X_4, X_1^2, X_2^2 and X_3^2 model parameters were found significant at a level of $p<0.05$ with values of 0.958 coefficient of determination (R^2) and 5.15% coefficient of variance (C.V). In the case of TFC, X_1, X_2,

X_3, X_4, X_1X_2, X_1^2, X_2^2, X_3^2 and X_4^2 while in the case of DPPH radical scavenging activity X_1, X_2, X_3, X_4, X_1^2, X_2^2, X_3^2 and X_4^2 model parameters appeared to be significant. The R^2 value for TFC and DPPH scavenging effect obtained was 0.918 and 0.919 along with 6.95 and 1.88% of C.V respectively.

Table 4.2. Experimental and predicted values of responses for different combinations of experimental conditions

Sl. No.	Co-solvent flow rate (mL/min)	Pressure (bar)	Temperature (°C)	Time (min)	Experimental values			ANN predicted values			RSM predicted values		
					TPC (mg GAE/mL)	TFC (mg QE/mL)	DPPH (%)	TPC (mg GAE/mL)	TFC (mg QE/mL)	DPPH (%)	TPC (mg GAE/mL)	TFC (mg QE/mL)	DPPH (%)
1	3	150	50	60	425.17	131.58	92.28	411.67	134.48	90.45	416.98	133.55	90.79
2	2	175	55	70	324.09	117.77	88.27	386.12	130.72	89.01	339.49	116.83	89.24
3	4	125	45	70	370.3	121.26	84.46	393.45	127.83	86.58	395.11	127.08	86.44
4	2	175	45	70	290.68	112.01	84.45	300.04	123.72	84.49	296.62	114.93	84.73
5	3	150	50	60	421.95	139.18	91.46	411.67	134.48	90.45	416.98	133.55	90.79
6	3	150	50	60	401.28	131.28	93.84	411.67	134.48	90.45	416.98	133.55	90.79
7	3	200	50	60	394.27	119.84	89.48	397.71	137.76	89.54	386.20	129.77	90.41
8	4	175	45	70	396.11	119.33	89.37	379.64	110.23	88.95	397.48	116.89	88.57
9	4	175	55	70	394.27	122.94	90.28	395.04	121.42	90.09	405.30	122.24	90.46
10	5	150	50	60	342.34	101.4	85.26	343.51	100.83	84.62	325.17	101.62	85.36
11	3	150	50	60	411.16	135.82	89.46	411.67	134.48	90.45	416.98	133.55	90.79
12	1	150	50	60	158.49	69.59	77.24	172.87	74.62	79.08	145.50	63.43	77.22
13	2	125	55	50	280.01	76.28	79.24	285.47	77.15	80.22	285.28	82.54	79.90
14	2	175	45	50	271.25	106.24	81.45	277.59	134.39	82.94	272.26	104.18	81.21
15	4	175	45	50	366.99	101.49	88.26	338.56	96.30	86.07	377.32	99.43	88.29

16	3	150	50	60	404.95	127.63	89.34	411.67	134.48	90.45	416.98	133.55	90.79
17	2	175	55	50	322.17	117.31	86.47	358.53	130.03	83.36	320.88	113.62	84.56
18	2	125	45	50	241.12	71.21	78.44	248.84	85.40	81.43	253.61	74.04	78.33
19	4	125	55	70	380.35	125.61	86.45	376.99	120.40	87.69	385.98	131.49	86.55
20	2	125	45	70	281.31	78.82	82.46	280.49	84.00	82.57	285.45	87.59	82.75
21	3	150	60	60	411.16	116.23	86.47	409.44	127.39	86.47	391.58	116.58	86.73
22	3	150	50	80	436.57	138.32	92.43	430.16	131.28	90.78	410.89	129.55	90.88
23	3	150	50	60	437.35	135.79	88.37	411.67	134.48	90.45	416.98	133.55	90.79
24	2	125	55	70	298.17	84.35	85.44	300.59	85.62	85.85	311.36	88.54	85.48
25	4	175	55	50	388.38	117.27	89.45	378.71	107.02	84.33	390.88	112.32	89.02
26	3	150	50	40	369.11	103.25	83.38	373.55	92.19	81.60	364.63	106.08	85.01
27	4	125	55	50	346.51	119.57	84.42	324.00	87.02	80.91	364.09	118.78	84.21
28	4	125	45	50	376.24	102.08	86.37	350.73	103.70	81.29	367.48	106.83	85.25
29	3	150	40	60	362.68	109.03	83.45	339.90	100.50	84.58	352.10	102.74	83.27
30	3	100	50	60	370.32	124.76	84.46	350.03	120.16	82.40	348.23	108.88	83.61

Based on results on responses, a second-order polynomial quadratic regression equation was established in terms of coded values (**Eq. 4.6-4.8**) for the parameters of extraction:

$$TPC = 416.98 + 44.92X_1 + 9.49X_2 + 9.87X_3 + 11.56X_4 - 2.20X_1X_2 - 8.76X_1X_3 - 1.05X_1X_4 + 4.24X_2X_3 - 1.87X_2X_4 - 45.41X_1^2 - 12.44X_2^2 - 11.28X_3^2 - 7.30X_4^2 \quad (4.6)$$

$$TFC = 133.55 + 9.55X_1 + 5.22X_2 + 3.46X_3 + 5.86X_4 - 9.38X_1X_2 + 0.86X_1X_3 + 1.68X_1X_4 + 0.24X_2X_3 - 0.70X_2X_4 - 1.88X_3X_4 - 12.75X_1^2 - 3.55X_2^2 - 5.97X_3^2 - 3.93X_4^2 \quad (4.7)$$

$$DPPH = 90.79 + 2.07X_1 + 1.53X_2 + 0.87X_3 + 1.46X_4 + 0.04X_1X_2 - 0.65X_1X_3 - 0.81X_1X_4 + 0.44X_2X_3 - 0.22X_2X_4 + 0.29X_3X_4 - 2.42X_1^2 - 0.74X_2^2 - 1.49X_3^2 - 0.75X_4^2 \quad (4.8)$$

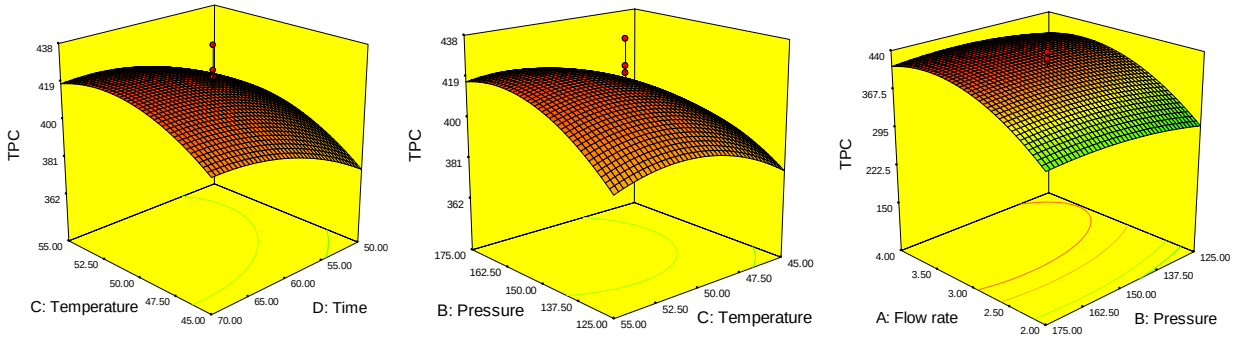
Table 4.3. Analysis of variance (ANOVA) of the RSM models for the responses

Source	TPC		TFC		DPPH radical scavenging activity	
	Sum of Squares	p-value	Sum of Squares	p-value	Sum of Squares	p-value
Model	115860	< 0.0001	10341.55	< 0.0001	457.4706	< 0.0001
X_1-Flow rate	48424.66	< 0.0001	2112.751	< 0.0001	99.55227	< 0.0001
X_2-Pressure	2162.771	0.0228	654.5882	0.0052	69.22407	0.0001
X_3-Temperature	2338.598	0.0187	287.4568	0.0471	18.02667	0.0198
X_4-Time	3209.288	0.0075	825.792	0.0023	51.56802	0.0005
X_1X_2	77.57206	0.6379	1409.252	0.0002	0.0225	0.9278
X_1X_3	1228.678	0.0752	11.9025	0.6661	6.8644	0.1286
X_1X_4	17.61901	0.8220	44.95703	0.4057	10.4976	0.0652
X_2X_3	287.3873	0.3699	0.893025	0.9056	3.1684	0.2917
X_2X_4	55.83826	0.6894	7.7841	0.7268	0.81	0.5887
X_3X_4	33.03376	0.7583	56.8516	0.3513	1.3456	0.4873
X_1^2	56560.88	< 0.0001	4318.204	< 0.0001	154.7957	< 0.0001
X_2^2	4245.026	0.0029	354.693	0.0297	24.52681	0.0083

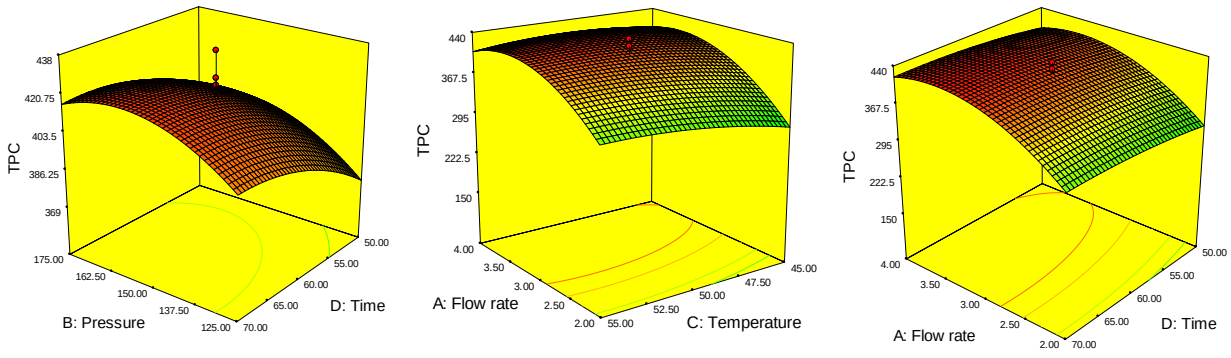
X_3^2	3492.611	0.0057	991.8907	0.0011	57.51953	0.0003
X_4^2	1463.379	0.0544	433.3431	0.0180	13.89987	0.0370
Residual	5043.436		921.4223		39.79563	
Lack of Fit	4111.639	0.1979	835.4748	0.0474	18.09634	0.8879
R^2	0.958		0.918		0.919	
Adj R^2	0.919		0.841		0.845	
%C.V.	5.15		6.95		1.88	
Adequate Precision	20.94		12.84		11.86	

The adjusted R^2 was found to be near to R^2 of the models and adequate precision recorded was in the range of 11.86 – 20.94 that depicts an adequate signal with less noise. Also, the non-significant lack of fit displayed that the selected model is a better fit for the predicted response.

Fig. 4.1-4.3 depicts the response surface plots for extraction parameters which were generated using second-order quadratic polynomial regression equation. The general trend of all the investigated compounds was similar, their values increased with the increasing trend of all variables to a certain extent followed by a regular decline. It is evident from (**Fig. 4.1**) that the extraction of phenols from haritaki was significantly influenced by independent variables X_1,X_2, X_3, and X_4 as well as the quadratic model terms X_1^2, X_2^2 and X_3^2.



(a) (b) (c)



(d) (d) (d)

Fig. 4.1. Response surface plot showing the effects of time, temperature, pressure and flow rate on TPC

TPC of haritaki extract was found to be gradually increasing with the increase of extraction time, flow rate, pressure and temperature. In case of the time of extraction, a steady increase in phenol content (252.45–431.56 mg GAE) until 70 min of extraction was observed with a maximum content of 431.56 mg GAE thereafter gradual but the drop was noticed up to 80 min. The flow rate of extraction ranging from 1-5 mL/min was used, an increasing trend was recorded with an increase of 158.38–430.41 mg GAE that subsequently decreases with increasing flow rate. In contrary, changing the level of pressure from 150 to 200 bar significantly decreased the phenol content. A similar pattern was seen in the temperature instance, where phenol content was significantly changed by a temperature increase of 50°C at a significant level of 0.05 with a significant p-value of 0.0187. The ANOVA findings indicated that the interaction between the models had no discernible impact ($p>0.05$) on the phenol content of haritaki.

In the case of flavonoid content, the interaction between X_1X_2 and quadratic models was found significant at a significant level of $p<0.05$ which was also supported by ANOVA results with F-value of 12.03 representing the significance of the model (**Fig. 4.2**). The trend of change in flavonoid extraction was quite similar in all independent variables; however, it was significantly influenced by the flow rate that produced the maximum flavonoid content with the level of 132.85 mg QE/mL.

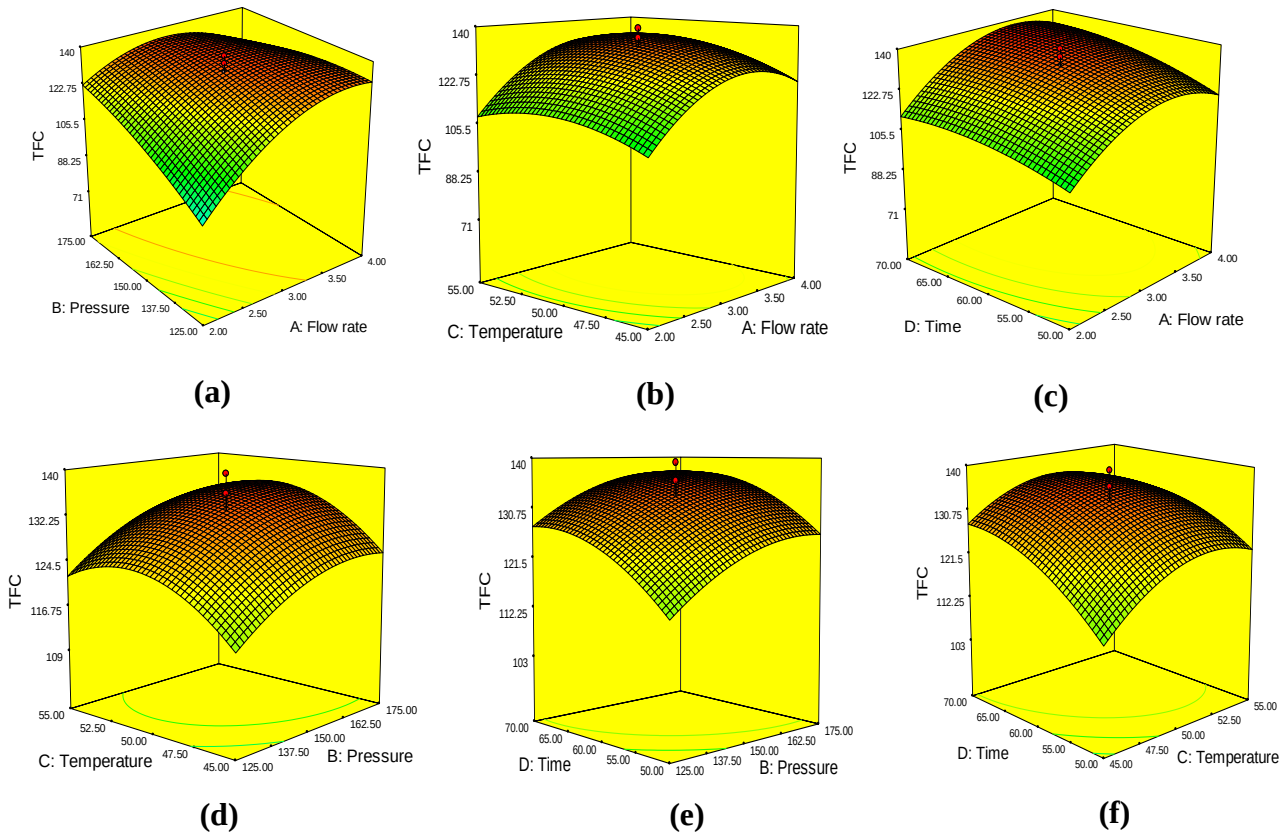


Fig. 4.2. Response surface plot showing the effects of time, temperature, pressure and flow rate on TFC

The variables (extraction time and flow rate) had less of an impact on the change in flavonoid content even when they tended to increase to their uppermost permissible limits. However, in the case of pressure and temperature, flavonoid extraction displayed a bell-shaped curve, and the variables mostly affected the extraction of flavonoid content, with the highest content noted at 150 bar and 50 °C and a notable tendency of drop after that. The present study showed similar result as reported by Yin et al. [32] that depicted highest flavonoid content extracted at 50 °C from *Pueraria lobate.*

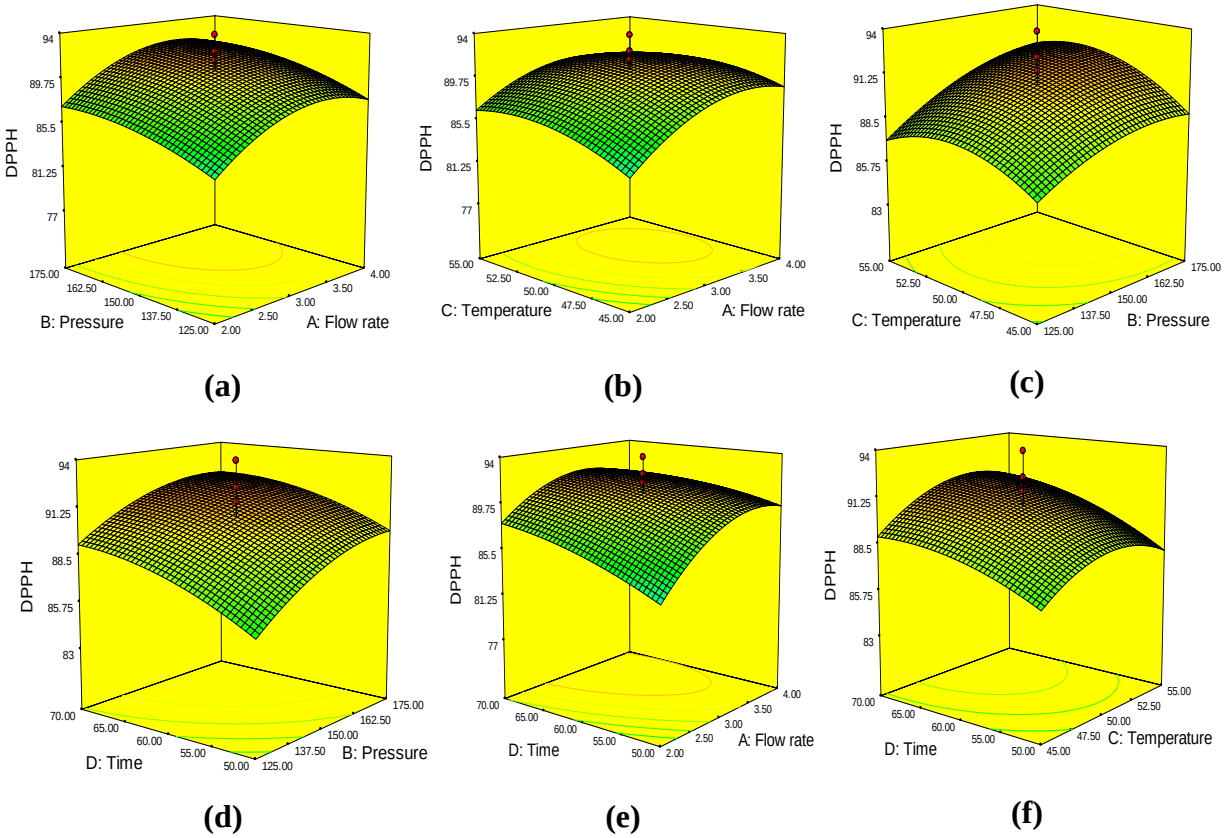


Fig. 4.3. Response surface plot showing the effects of time, temperature, pressure and flow rate on DPPH

The phytochemicals present in the extract contributes to the total antioxidant activity as measured by DPPH radical scavenging activity. DPPH radical scavenging activity of haritaki extract was drastically affected by the extraction parameters which was also depicted by ANOVA results where all the independent variables and quadratic models were found significant with non-significant lack of fit. With an increase in extraction time, DPPH radical scavenging activity of haritaki extract was seen to increase and maximum activity recorded was at 80 min where flow rate depicted a systematic trend in the DPPH radical scavenging activity of haritaki extract that tends to decrease significantly after 3 mL/min of flow rate (**Fig. 4.3**). The 3-D surface plot also shows that as the extraction pressure is increased, the activity of the haritaki extract likewise rises until a pressure of 150 bar, after which it gradually decreases. When it came to temperature, DPPH radical scavenging activity grew systematically as extraction temperature rose, then it began to fall. The peak activity was found to be at 50 °C. The results obtained for the effect of extraction time on phytochemicals showed higher recovery with the advancement of time. The same was supported by the linear effect in

3-D surface plots of TPC, TFC and DPPH activity. On the other hand, ANOVA results also showed statistical significance of the model terms for every response. The flow rate of the solvent influenced the extraction of the phytochemicals. The recovery of the phytochemicals i.e., concentration in the extract increased with increase in the flow rate up to certain value (approximately 3 mL/min) and then became constant. The reason for increased recovery of haritaki phytochemicals at higher flow rate might be due to decreased mass transfer resistance, saturation of CO_2 and fulfilment of equilibrium condition because of higher yield of bioactive compounds. However, with further advancement in extraction flow rate caused decrease in residence time i.e., fluid bed interaction time [18]. Hence, determination of the optimum flow rate of the extraction of a particular compound in a semi-continuous manner using SFE becomes an important factor that depends on temperature and pressure of extraction.

It has been reported that higher pressure favours the extraction of chemicals [4], however the contrary tendency was seen in the present investigation. This is because the pressure utilised in the recovery of bioactive compounds greatly influenced the content of haritaki. Initial increases in extraction pressure improved CO_2 density and solubility, leading to higher compound recovery (**Figs. 4.1 and 4.2**), but further increases in pressure decreased the diffusivity of haritaki phytochemicals, decreased convective mass transfer, and negatively affected compound recovery [18].

The extraction solvent (ethanol) improved CO_2 solvation effectiveness at lower temperatures, however the extracted bioactive components of haritaki were dramatically impacted by an increase in extraction temperature. A declining trend was observed in all compounds, which might be attributed to the decrease in density of CO_2 and resulted in decreased the components extraction, however interaction of pressure and temperature was opposite [13]. According to a study by Kassama et al. [14], immediate heating has a considerable negative impact on the effect that an increase in temperature and solubility of the solute has on the compound of interest. A similar condition was reported by Kuś et al. [17] in the extraction of bio actives of black poplar buds using SFE-CO_2.

4.3.2. ANN modelling

For training, testing, and validation of the ANN, the feed forward back propagation model with three layers input, hidden, and output and 30 data sets was utilised. While the input layer contained four neurons and the output layer contained three

neurons as decided by the number of independent variables and responses respectively, the number of neurons were varied from 8 to 12 for the hidden layer.

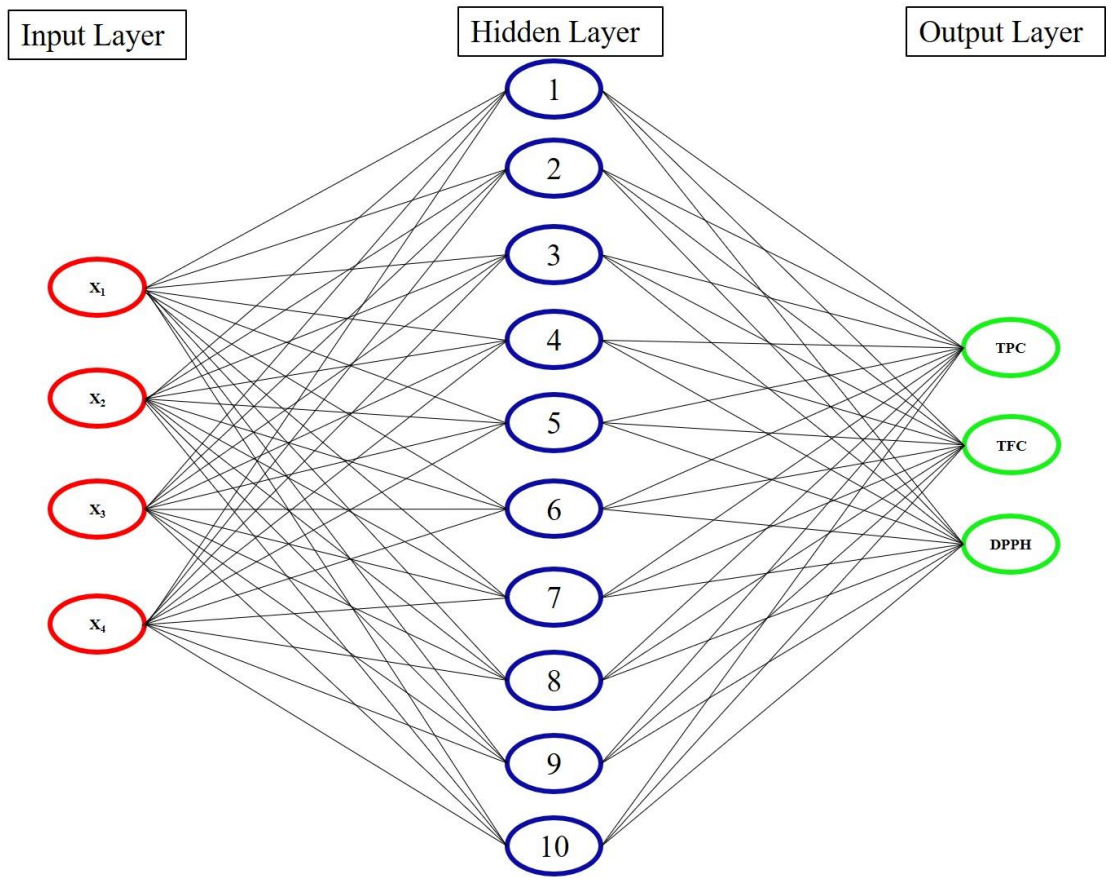


Fig. 4.4. General architecture of the feed forward back propagation multilayer perceptron (MLP) neural network consisting of 4 neurons in the input layer, 10 neurons in the hidden layer and 3 neurons in the output layer.

The optimal topology was discovered to be attained when the hidden layer included 10 neurons. Therefore, the final topology for the developed ANN model contained four, ten and three neurons for input, hidden and output layers respectively (**Fig. 4.4**).

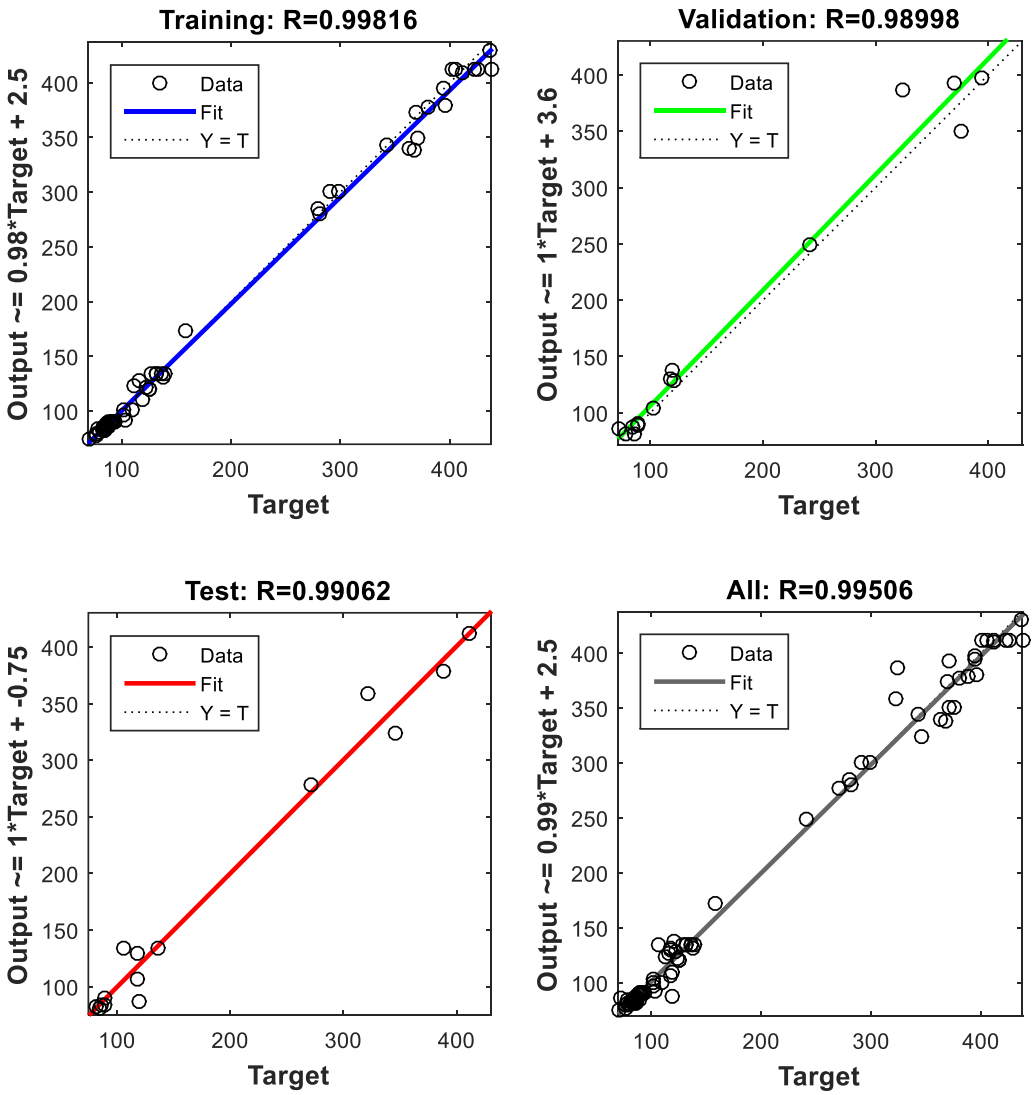


Fig. 4.5. Correlation coefficients (R) for training, validation, testing and overall datasets for the developed ANN model

As noted in **Fig. 4.5**, the correlation coefficients (R) for training, testing, and validation are 0.998, 0.991, and 0.989, respectively, while R for the entire dataset is 0.995. This data shows that the constructed model is sufficiently reliable to predict the outputs for various sets of input values. The sizes for the matrices for the weights joining the neurons input layer to hidden layer is 10×4 and for the weights joining the hidden layer to output layer is 3×10, while the size of the matrices of the biases for the hidden layer is 10×1 and for the output layer is 3×1.

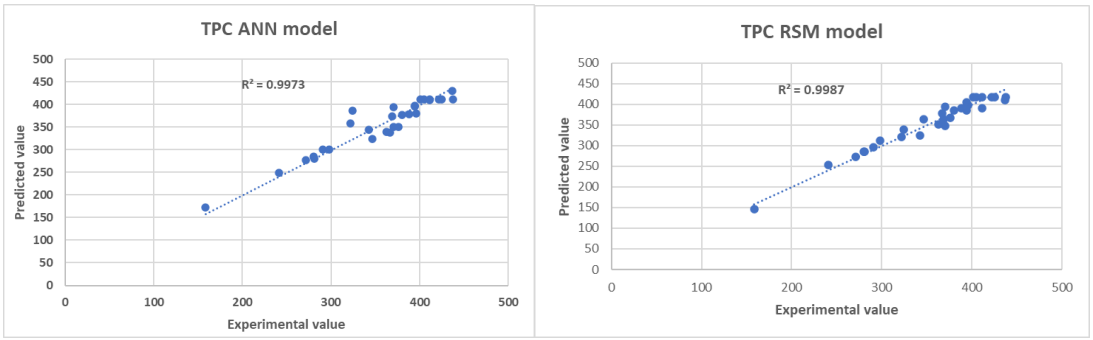
4.3.3. Comparison of the developed RSM and ANN models

The experimental values and corresponding predicted values for the developed RSM and ANN models are presented in **Table 4.4**. The performances of the RSM and the ANN models were compared using the statistical parameters such as coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE) and Chi square (χ^2) values and are given in **Table 4.4**.

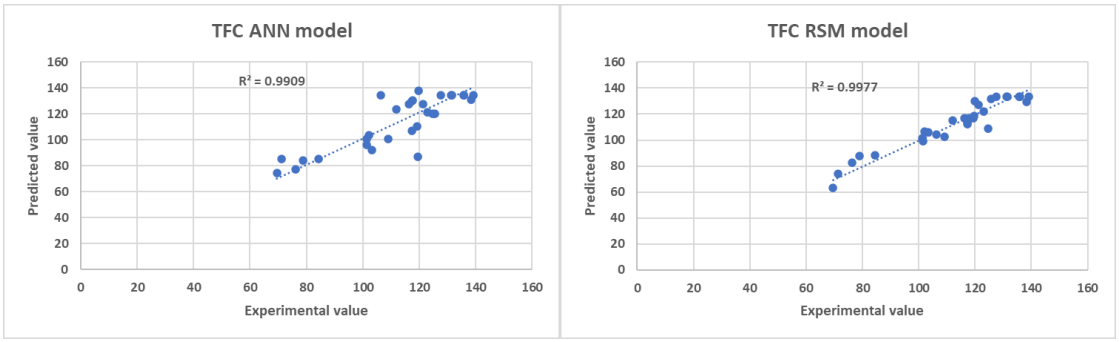
Table 4.4. Comparison of the statistical parameters of ANN and RSM models for the various responses

	TPC	**TFC**	**DPPH**
R^2 value			
ANN Model	0.9973	0.9909	0.9994
RSM Model	0.9987	0.9977	0.9998
MAE			
ANN Model	13.40311	8.175279	1.643748
RSM Model	10.94936	4.326667	0.843028
RMSE			
ANN Model	18.78651	11.0563	2.125115
RSM Model	12.9659	5.462292	1.151747
χ^2			
ANN Model	29.38151	33.96452	1.611373
RSM Model	14.28889	8.336857	0.448451

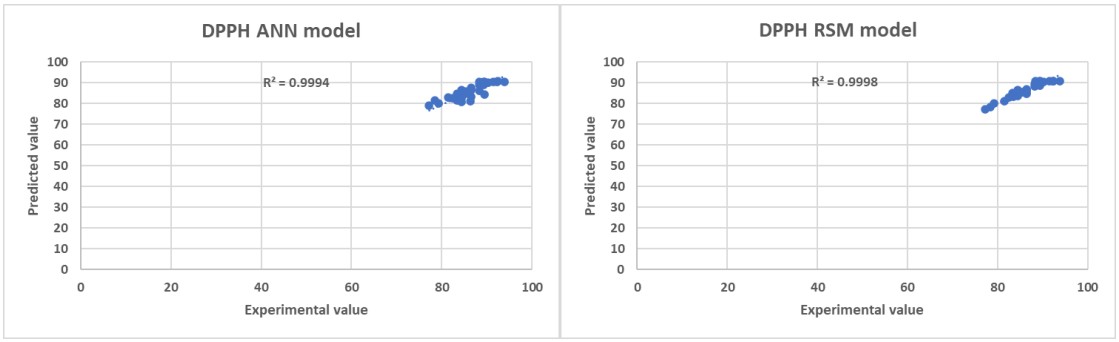
Table 4.4 shows that the R^2 values for the TPC, TFC, and DPPH RSM models are greater than the comparable R^2 values for the created ANN models. Additionally, the RSM models had lower RMSE, MAE, and χ^2 values than the ANN models. In the current study, RSM models exhibit greater predictive capability than ANN models, as evidenced by higher R^2 values and lower MAE, RMSE, and χ^2 values for the RSM models [7].



(a)



(b)



(c)

Fig. 4.6. Comparison of the performances of the ANN and RSM models for **(a)** TPC; **(b)** TFC; and **(c)** DPPH

For all of the responses, the experimental values were plotted against the ANN and RSM model predictions, and the results were compared (**Fig. 4.6**). Additionally, it is evident from (**Fig. 4.6**) that the RSM model outperformed the ANN model in the current study. The effectiveness of a generated model depends on a variety of variables, including the quantity of experiments, the kind of procedure, etc. While ANN performs well with both smaller and bigger numbers of experimental data, RSM, which is simpler, occasionally predicts well when the number of experiments is limited. RSM was determined to be better in the current analysis because there were only 30 experiments,

but the ANN model could also accurately predict the answers. Similar results were reported by Zeng et al. [34], Rizalman and Lee [24] and Naderloo [19] while comparing performances of RSM and ANN models.

4.3.4. Optimization of the process parameters by different approaches

The extraction method was optimised using three alternative strategies. The methods include ANN combined with a genetic algorithm (ANN-GA), RSM combined with a genetic algorithm (RSM-GA), and RSM combined with a numerical method employing a desirability function (RSM-DF). For ANN-GA optimization, the final trained ANN model was used as the fitness function and for RSM-GA optimization, the fitness function used was as follows:

$$F = - (y1 + y2 + y3) \quad (4.9)$$

Where, y1, y2 and y3 are the predicted values of TPC, TFC and DPPH respectively for the RSM model. In a genetic algorithm, response maximisation was achieved by using the negative sign. The population size employed for the ANN-GA and RSM-GA optimizations was 200, as previously noted, and the optimal value was attained after 63 and 69 generations, respectively. Using Design Expert software, optimization was also carried out numerically for the RSM models using the desirability function. A comparison of the optimum values of the responses obtained by the three optimization methods viz. ANN-GA, RSM-GA and RSM-DF is shown in **Table 4.5**.

Table 4.5. Comparison of the optimization results obtained from the three different approaches

	Flow rate (mL/min)	**Pressure (bar)**	**Temperature (°C)**	**Time (min)**	**TPC (mg GAE/ml)**	**TFC (mg QE/ml)**	**DPPH (%)**	
Approach								
RSM-GA	3.23	172.79	52.37	68.53	428.03	136.58	92.63	
ANN-GA	3.30	174.07	51.18	65.23	414.25	135.55	91.32	
RSM-DF	3.34	166.94	51.97	67.47	432.28	137.36	92.54	Desirability (0.951)

The results demonstrate that, although the results are extremely similar, optimization using RSM-GA and RSM-DF generated slightly higher values of the responses, i.e. TPC, TFC, and DPPH. This shows that all three techniques are effective

and can be employed to maximise the process of phytochemical extraction from *T. chebula* pulp by supercritical fluid extraction.

4.4. Conclusion

In order to extract phytochemicals from the pulp of haritaki, the supercritical extraction process was modelled using statistical and artificial intelligence techniques (RSM and ANN), and several methods were used to optimise the extraction conditions. In this chapter, it was found that the RSM model performed better in response prediction than the ANN model. By contrasting the statistical parameters of the two models, this was proven. As a result, it can be said that both RSM and ANN approaches are reasonably effective at predicting responses, albeit the performances may vary based on a variety of aspects, such as the specific process, the quantity of tests conducted, the independent and dependent parameters used, etc. Additionally, the RSM and ANN models were used to optimise the SFE process using the desirability function and genetic algorithm as RSM-DF, RSM-GA, and ANN-GA. Additionally, in this instance, it was found that the RSM-DF and RSM-GA models performed better than the ANN-GA model, which was the goal of optimization. The optimum conditions obtained for RSM-DF were 3.34 mL/min, 166.94 bar, 51.97 ℃, 67.47 min, for RSM-GA were 3.23 mL/min, 172.79 bar, 52.37 ℃, 68.53 min, while that for ANN-GA were 3.30 mL/min, 174.07 bar, 51.18 ℃, 65.23 min. The optimized values of responses i.e., TPC (mg GAE/mL), TFC (mg QE/mL) and DPPH (%) for RSM-DF approach were 432.28, 137.36, 92.54 respectively, for RSM-GA approach were 428.03, 136.58, 92.63 respectively, and for ANN-GA approach were 414.25, 135.55, 91.32 respectively. Therefore, finally it can be concluded that both RSM and ANN models can be used for predicting the responses with good accuracy, but one must select the specific model depending on the situation to get the best results. Similarly, for optimization also, all the approaches gave fairly good results and again one must choose a particular approach for fulfilling the specific criteria for optimization.

4.5. References

1. Akintunde, A. M., Ajala, S. O., and Betiku, E. Optimization of Bauhinia monandra seed oil extraction via artificial neural network and response surface methodology: a potential biofuel candidate. *Industrial Crops* and *Products*, 67, 387–394, 2015.
2. Arce-Medina, E., and Paz-Paredes, J. I. Artificial neural network modeling techniques applied to the hydrodesulfurization process. *Mathematical* and *Computer Modelling*, 49(1–2), 207–214, 2009.
3. Avula, B., Wang, Y. H., Wang, M., Shen, Y. H., and Khan, I. A. Simultaneous determination and characterization of tannins and triterpene saponins from the fruits of various species of *Terminalia* and *Phyllantus emblica* using a UHPLC-UV-MS method: application to triphala. *Planta Medica*, 29(02), 181-188, 2013.
4. Bogdanovic, A., Tadic, V., Stamenic, M., Petrovic, S., and Skala, D. Supercritical carbon dioxide extraction of *Trigonella foenum-graecum* L. seeds: process optimization using response surface methodology. *Journal* of *Supercritical Fluid*, 107, 44–50, 2016.
5. Brand-Williams, W., Cuvelier, M. E., and Berset, C. L. W. T. Use of a free radical method to evaluate antioxidant activity. *LWT-Food science and Technology*, 28(1), 25-30, 1995.
6. Chegini, G. R., Khazaei, J., Ghobadian, B., and Goudarzi, A. M. Prediction of process and product parameters in an orange juice spray dryer using artificial neural networks. *International journal of food engineering*, 84(4), 534–543, 2008.
7. Chouaibi, M., Daoued, K.B., Riguane, K., Rouissi, T., and Ferrari, G. Production of bioethanol from pumpkin peel wastes: comparison between response surface methodology (RSM) and artificial neural networks (ANN). *Industrial Crops and Products*, 155, 112822, 2020.
8. Coit, D. W., Jackson, B. T., and Smith, A. E. Static neural network process models: considerations and case studies. *International Journal* of *Production Research*, 36(11), 2953–2967, 1998.
9. Dewanto, V., Wu, X., Adom, K. K., and Liu, R. H. Thermal processing enhances the nutritional value of tomatoes by increasing total antioxidant activity. *Journal of Agricultural and Food Chemistry*, 50(10), 3010-3014, 2002.

10. Guan, W., Li, S., Yan, R., Tang, S., and Quan, C. Comparison of essential oils of clove buds extracted with supercritical carbon dioxide and other three traditional extraction methods. *Food Chemistry*, 101(4), 1558–1564, 2007.
11. Hornik, K., Stinchcombe, M., and White, H. Multilayer feedforward networks are universal approximators. Neural Netw. 2(5), 359–366, 1989.
12. Hugget, A., Sebastian, P., and Nadeau, J.P. Global optimization of a dryer by using neural networks and genetic algorithms. *AIChE Journal,* 45(6), 1227–1238, 1999.
13. Joki´c, S., Nagy, B., Zekovi´c, Z., Vidovi´c, S., Bili´c, M., Veli´c, D., and Sim´andi, B. Effects of supercritical CO_2 extraction parameters on soybean oil yield. Food Bioprod. Process. 90(4), 693–699, 2012.
14. Kassama, L.S., Shi, J., and Mittal, G.S. Optimization of supercritical fluid extraction of lycopene from tomato skin with central composite rotatable design model. Sep. Purif. Technol. 60(3), 278–284, 2008.
15. Kim, D. O., Jeong, S. W., and Lee, C. Y. Antioxidant capacity of phenolic phytochemicals from various cultivars of plums. *Food Chemistry*, 81(3), 321-326, 2003.
16. Kong, Y., Fu, Y. J., Zu, Y. G., Liu, W., Wang, W., Hua, X., and Yang, M. Ethanol modified supercritical fluid extraction and antioxidant activity of cajaninstilbene acid and pinostrobin from pigeonpea *Cajanus cajan* (L.) Millsp. leaves. *Food Chemistry*, 117(1), 152-159, 2009.
17. Ku´s, P., Jerkovi´c, I., Jakovljevi´c, M., and Joki´c, S. Extraction of bioactive phenolics from black poplar (*Populus nigra* L.) buds by supercritical CO_2 and its optimization by response surface methodology. J. Pharm. Biomed. Anal. 152, 128–136, 2018.
18. Masghati, S., and Ghoreishi, S. M. Supercritical CO_2 extraction of cinnamaldehyde and eugenol from cinnamon bark: optimization of operating conditions via response surface methodology. *Journal* of *Supercritical Fluids*, 140, 62–71, 2018.
19. Naderloo, L. Prediction of solar radiation on the horizon using neural network methods, ANFIS and RSM (case study: Sarpol-e-Zahab Township, Iran). *Journal of Earth System Science,* 129(1), 1–11, 2020.
20. Pradhan, P., Tingsanchali, T., and Shrestha, S. Evaluation of Soil and Water Assessment Tool and Artificial Neural Network models for hydrologic simulation in different climatic regions of Asia. Sci. Total Environ. 701, 134308, 2020.

21. Prado, P., and Thibaut, T. Differences between Epiphytic Assemblages on Introduced Caulerpa taxifolia and Coexisting Eelgrass (*Zostera Capricorni*) in Botany Bay (NSW, Australia), 2008.
22. Rai, P., Majumdar, G.C., Das, G. S., and De, S. Prediction of the viscosity of clarified fruit juice using artificial neural network: a combined effect of concentration and temperature. *Journal* of *Food Engineering*, 68(4), 527–533, 2005.
23. Rangsriwong, P., Rangkadilok, N., Satayavivad, J., Goto, M., and Shotipruk, A. Subcritical water extraction of polyphenolic compounds from *Terminalia chebula* Retz. fruits. *Separation and Purification Technology*, 66(1), 51-56, 2009.
24. Rizalman, A. N., and Lee, C. C. Comparison of artificial neural network (ANN) and response surface methodology (RSM) in predicting the compressive strength of POFA concrete. *Applications of Modelling and Simulation*, 4, 210–216, 2020.
25. Roy, B. C., Goto, M., and Hirose, T. Temperature and pressure effects on supercritical CO_2 extraction of tomato seed oil. *International Journal* of *Food Science* & *Technology*, 31 (2), 137–141, 1996.
26. Scalia, S., Giuffreda, L., and Pallado, P. Analytical and preparative supercritical fluid extraction of chamomile flowers and its comparison with conventional methods. J. Pharm. Biomed. Anal. 21 (3), 549–558, 1999.
27. Shafi, J., Sun, Z., Ji, M., Gu, Z., and Ahmad, W. ANN and RSM based modelling for optimization of cell dry mass of Bacillus sp. strain B67 and its antifungal activity against Botrytis cinerea. *Biotechnology* & Biotechnological *Equipment*, 32 (1), 58–68, 2018.
28. Shilpi, A., Shivhare, U. S., and Basu, S. Supercritical CO_2 extraction of compounds with antioxidant activity from fruits and vegetables waste-a review. *Focusing on Modern Food Industry*, 2(1), 43-62, 2013.
29. Sim´andi, B., Sass-Kiss, ´A., Czukor, B., De´ak, A., Prechl, A., Csord´as, A., and Sawinsky, J. Pilot-scale extraction and fractional separation of onion oleoresin using supercritical carbon dioxide. *Journal* of *Food Engineering*, 46 (3), 183–188, 2000.
30. Srinivas, M., and Patnaik, L. M. Adaptive probabilities of crossover and mutation in genetic algorithms. *IEEE Transactions* on *Systems, Man,* and *Cybernetics*, 24 (4), 656–667, 1994.
31. Wanasundara, U.N., and Shahidi, F. Concentration of ω-3 polyunsaturated fatty acids of marine oils using *Candida cylindracea* lipase: optimization of reaction conditions. *Journal* of the *American Oil* Chemists' *Society*, 75, 1767–1774, 1998.

32. Yin, W., and Wang, S. The character of pueraria lobata starch, chemical composition and pharmacological function of Pueraria lobata. Grain Process. 1, 84–86, 2008.
33. Zbici´nski, I., Kami´nski, W., Ciesielski, K., and Strumiłł, P. Dynamic and hybrid neural model of thermal drying in a flutdized bed. Dry. Technol. 15 (6–8), 1743–1752, 1997.
34. Zeng, Z., Chen, M., Wang, X., Wu, W., Zheng, Z., Hu, Z., and Ma, B. Modeling and optimization for konjac vacuum drying based on response surface methodology (RSM) and artificial neural network (ANN). *Processes* 8 (11), 1430, 2020.