

Optimization of Barbaloin Extraction From *Aloe Vera* Leaf Skin By Adaptive Neuro-Fuzzy Interface System, Support Vector Regression, Response Surface Methodology

Devendra Kumar Pandey*, Babita Pandey¹, R.M. Banik²,

**Department of Biotechnology, School of Biosciences and Biotechnology, Lovely Professional University, Phagwara-144402, Indian, ¹Department of Computer Science and Information Technology, Babasaheb Bhimrao Ambedkar University, ²School of Biochemical Engg. IIT-BHU, * Corresponding author, Tel.: +91-9779683414; E-mail address: dkpandey1974@yahoo.com*

Abstract: *Aloe vera*, an ayurvedic medicinal plants are rich source of barbaloin. The yield of barbaloin from *Aloe vera* leaf is dependent upon various extraction parameters. In this work, we have deployed various optimization methods such as: response surface methodology (RSM), radial basis function (RBF) based support vector regression and adaptive network-based fuzzy inference system (ANFIS) to optimize various extraction parameters to maximize the yield of barbaloin from *Aloe vera* leaf. The work was implemented in two phase: first, identify optimum extraction parameters and then identifying their optimum values. It is observed from the result that the mathematical models i.e. RBF and ANFIS fit well with the experimental data. It was also observed from the result that RSM, RBF and ANFIS performed well and shows stable results in predicting the combined interactions of the independent variables. The assessment of the models was done by coefficient of determination (R^2) showed that ANFIS was superior ($R^2 = 0.971$) to the RSM model ($R^2 = 0.931$) and ANN ($R^2 = 0.970$) in predicting barbaloin yield.

Key words: *Aloe vera*, barbaloin, solid-liquid extraction, central composite design, response surface methodology and artificial neural network.

1. Introduction

Aloe vera (Family Liliaceae) is found in semi wild state and cultivated in all parts of India [1]. Fresh juice of *Aloe vera* is used for various purposes such as: treatment (The wealth of India, 2000), cosmetics and Ayurveda medicine [2]. Barbaloin is a major constituent of *Aloe vera*, which is very important bioactive compound. It is very effective

in various activity on various human organs [3-7]. Therefore, it is urgent need to develop cheap and cost effective extraction methods.

Now days, response surface methodology (RSM) is preferred over classical optimization (CO) methods [9, 10] due to its less time overhead. It is very effective in terms of time as compared to CO. Its study interaction among factors and select optimum conditions of variables for a desirable response [8]. But ANN shows better input–output relationships mapping for complex systems as compared to RSM [11, 12]. Adaptive network-based fuzzy inference system (ANFIS) is a hybrid of ANN and Takagi–Sugeno fuzzy inference system. This hybrid system will take the advantage of self-adaptability and learning characteristics of ANN and the uncertainty and imprecise knowledge dealing of fuzzy logic and built a better system than ANN [13, 14].

Recent studies include the extraction of phenols from mangosteen hull, essential oils from *Diplotaeniachrydifolia* [14], coumarin from *Cuscutareflexa* [15], secoisolariciresinoldiglucoside from flaxseed [16], oils from *Orthosiphon stamineus* [17] and *passiflora* seeds [18], triterpenoids from *Swertiachirata* and the extraction of natural dyes [19, 20] in which RSM, ANN and ANFIS were applied.

In this work, RSM, radial basis function (RBF) based support vector regression (SVR) and ANFIS have been applied to optimize the yield of barbaloin from *Aloe vera* leaf. The proposed extraction method is cost effective and novel. Furthermore, enhancement of barbaloin extraction from the *Aloe vera* leaf by optimization of extraction conditions by reflux method has not been attempted so far.

2. Material and methods:

2.1. Extraction and analysis of barbaloin

Aloe vera plants were vegetative propagated at B.H.U., Varanasi, India. The extraction and analysis of barbaloin was done by heat reflux (Banik and Pandey, 2009) and HPTLC method (Pandey et al. 2016) respectively for barbaloin from *Aloe vera* leaves.

The quantitative estimation of barbaloin by HPTLC was done by on the pre-coated silica gel 60 F₂₅₄ TLC plates (20 cm x 10 cm and mobile phase ethyl acetate: methanol and water 100:16.5:13.5 (v/v) and derivatization with 100 ml of 10% (v/v) Ethanolic sulphuric acid solution, at 366 nm.

2.3. Response Surface Methodology (RSM)

RSM was applied in two stages, first Placket and Burman design (PBD) criterion was deployed to determine the significant parameter affecting barbaloin yield; then followed central composite design (CCD) was deployed to obtained the optimize level of significant parameters. All the implementation was done in Minitab 15 statistical software package.

2.4.1. Placket-Burman design(PBD)

Twelve experiments using PBD criterion at high level and a low level denoted was done to screen the extraction parameter for extraction of barbaloin from *Aloe vera* leaf. The details are shown in **Table 1**. Each experiment was repeated twice and the average value of barbaloin yield was used for statistical analysis at 5% level ($P < 0.05$) from the regression analysis(as given in **Table 2**).

Table 1:Detail of High level and low level of the extraction parameters by PBD criterion

Extraction Code	Extraction condition	High level	Low level
X ₁	temperature	70 °C	35 °C
X ₂	time	60 min	30 min
X ₃	solvent: solid ratio	60:1 ml/g	30:1 ml/g
X ₄	particle size	1.2 mm	0.6 mm
X ₅	extraction steps	3	1

Table 2:Yield of barbaloin at different levels

Treatment variables			Coded levels				in
			Symbol	-1.643	-1	0	
			+1.643				
4	70	30	60:1 ml/g	1.2 mm	1	1.67	
5	70	60	30:1 ml/g	1.2 mm	3	2.53	
6	70	60	60:1 ml/g	0.6 mm	3	1.56	
7	35	60	60:1 ml/g	1.2 mm	1	1.61	
8	35	30	60:1 ml/g	1.2 mm	3	1.41	
9	35	30	30:1 ml/g	1.2 mm	3	1.31	
10	70	30	30:1 ml/g	0.6 mm	3	1.48	
11	35	60	30:1 ml/g	0.6 mm	1	1.54	
12	35	30	30:1 ml/g	0.6 mm	1	0.78	

2.4.2. Central composite design (CCD)

The RSM was used to determine the optimum level defined by maximum barbaloin yield (%) of three significant extraction parameters i.e. screened Temperature, extraction time and solvent:solid ratio from PBD (as given in **Table 3**). The experimental design scheme is given in **Table 4**. barbaloin was analyzed by using a second order polynomial equation and the data were fitted in to the equation by multiple regression procedure.

Table 3: Treatment variables and their coded and actual values used for optimization of barbaloin extraction from *Aloe vera leaf skin* by using CCD

Table 4: CDD criterion of extraction parameters with corresponding experimental and predicted value

Run order	Extraction parameters levels			Barbaloin (% dry weight)			
	inp1	inp2	inp3	EXP	ANFIS	ANN	RSM
1.	52.5	45	50	2.616	2.614	2.48725	2.614
2.	70	60	30	1.747	1.747	1.7725	1.736
3.	35	30	30	1.146	1.146	1.345938	1.153
4.	35	30	70	1.716	1.716	1.7725	1.754

5.	52.5	19.8	50	1.446	1.446	1.645938	1.446
6.	52.5	45	50	2.587	2.335	2.48725	2.614
7.	35	60	30	1.234	1.234	1.433938	1.238
8.	70	60	70	2.459	2.459	2.258938	2.478
9.	23.1	45	50	1.124	1.333	1.323938	1.085
10.	52.5	45	50	2.65	2.335	2.48725	2.614
11.	52.5	70.2	50	1.9	1.900	1.7725	1.883
12.	70	30	70	2.02	2.020	1.819938	2.042
13.	52.5	45	50	2.556	2.635	2.48725	2.614
14.	52.5	45	50	2.687	2.635	2.48725	2.614
15.	35	60	70	1.614	1.314	1.7725	1.652
16.	52.5	45	50	2.58	2.635	2.48725	1.614
17.	52.5	45	16.4	1.439	1.439	1.63925	1.459
18.	52.5	45	83.6	2.645	2.645	2.444938	2.588
19.	70	30	30	1.125	1.325	1.324938	1.113
20.	81.9	45	50	1.745	1.745	1.7725	1.747

2.4.3. Support vector regression (SVR)

The RSM design used in experimental data, were also used in designing the SVR as shown in Table 4. The SVR hyper parameters chosen here cost factor (C=1), epsilon=0.2, kernel= radial basis function (RBF). In current years, RBF based SVR was deployed by many researchers as a modeling tool in bioprocesses [23-24].

The type of kernel was the RBF, due to it fast training and its powerful role in optimization modeling. RBF is radial basis function activation function. In recent years, RBF is deployed as a modeling tool by many researchers in various applications in bioprocesses [23]. The RBF network consisted of input layer, hidden layer of RBF neurons (Fig.1) that is fully interconnected to an output layer [24].

The input can be modeled as a vector of real numbers $x \rightarrow R^n$. The output of the network is then a scalar function of the input vector, $y: R^n \rightarrow R$, and is given by

The relationship between input vector and output is defined by equation 3.

$$y(x) = \sum_{i=1}^n \alpha_i \exp[-\beta \|x - c_i\|^2] \quad (3)$$

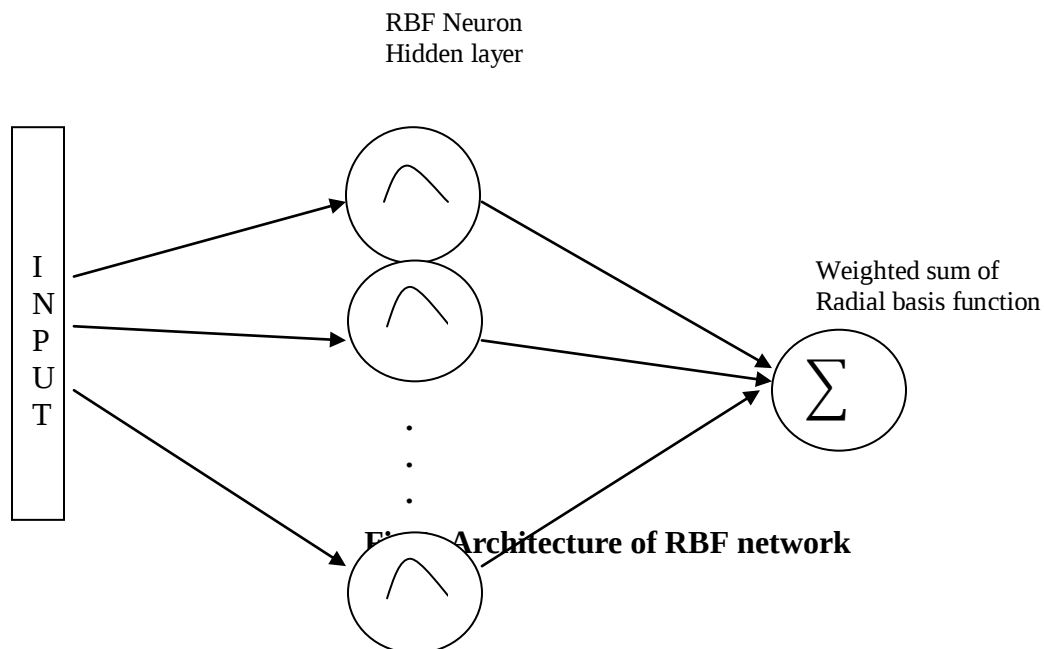
Where w_0 and w_i , represents the weights,

where n is the number of neurons in the hidden layer, c_i is the center vector for neuron, and α_i is the weight of neuron i in the linear output neuron[21]. The SVR is implemented in python and ANFIS was implemented in MATLAB.

The predicted response of barbaloin yield obtained for RSM and RBF network was compared in terms of root mean squared error (RMSE) and coefficient of determination (R^2).

$$RMSE = \sqrt{\left(\frac{1}{n} \sum_{i=1}^n (T_p - T_e)^2\right)} \quad (5)$$

The experimental data given in Table 4 is used as input to RBF. It identify the nonlinear relationship between input variables (temperature, extraction time and solvent-solid ratio) and the output variable (% barbaloin yield).



2.4.5 ANFIS

ANFIS consist of five layers. Each layer contains several nodes described by the node functions. Each layer is cascaded with next layer in such a way that the output of previous

layer is fed as input to the current layer and the output of current layer is fed as input to the next layer. First layer is input and fuzzyfication layer, Second layer is rule layer, third layer is normalization layer, forth layer identify the strength of each rule by assigning the weight and fifth layer performs defuzzification. The rule base of ANFIS consists of fuzzy if then rules. Linguistic variables are assigned to each input. Membership function is associated with each rule. A membership function is a value between 0 and 1. It defines how much a variable belongs to a linguistic set, zero means completely absent and one means completely present in the linguistic set. A membership functions can be trapezoidal, triangular and bell shaped. A membership function is required for each input variable. The parameters used for ANFIs are given in **Table 5**.

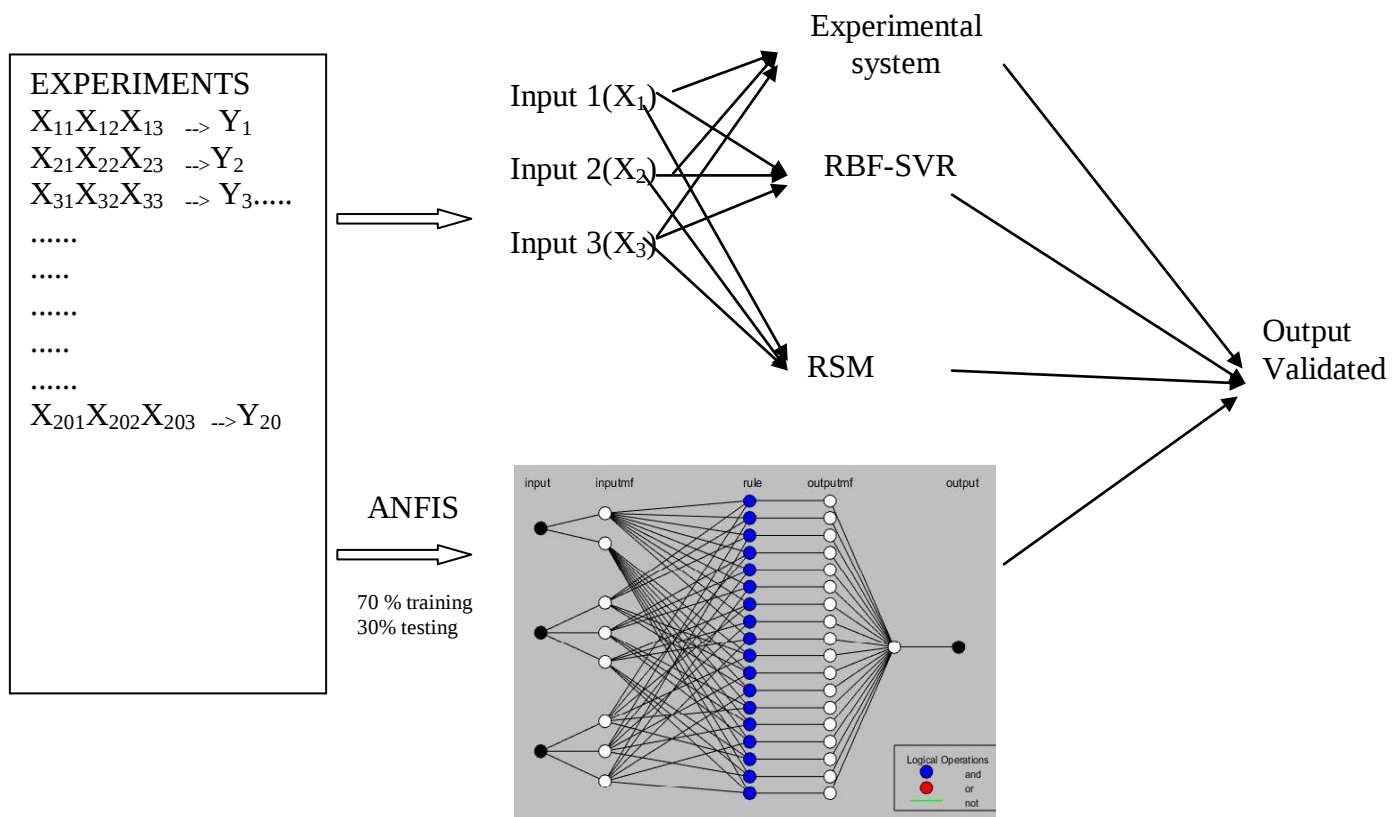


Fig2: Block diagram of proposed system (X_{ij} denotes j^{th} input of i^{th} experiment; Y_i denotes output of i^{th} experiment)

Table 5: ANFIS parameters

ANFIS parameters	Number
nodes	58
linear and non linear parameters	18 and 24
Total parameters	42
training data pairs	14
fuzzy rules	18

The complete block diagram of proposed system is shown in Fig 2. The training and testing results are shown in Fig 3 and Fig 4 respectively. Figure 5a, 5b and 5c shows the surface view of input 2 with respect to 1, surface view of input 3 with respect to 1 and surface view of input 3 with respect to 2 respectively.

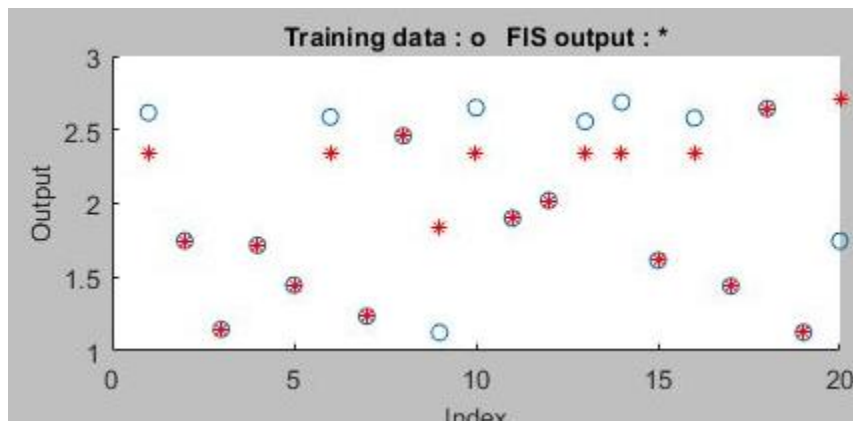


Fig 3: ANFIS training Phase

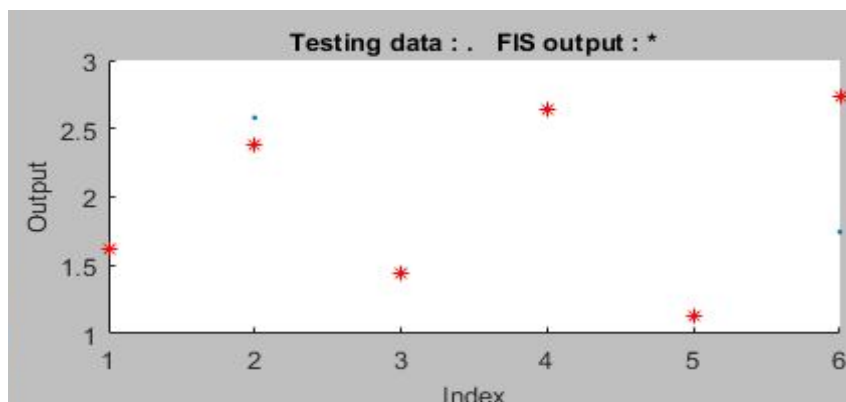


Fig 4: ANFIS testing Phase

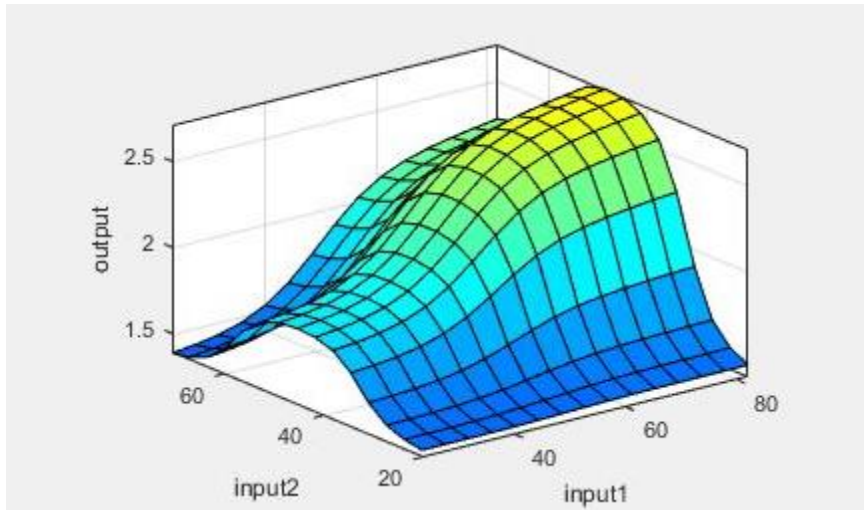


Fig (5a) Surface viewer of input 2 with respect to input 1

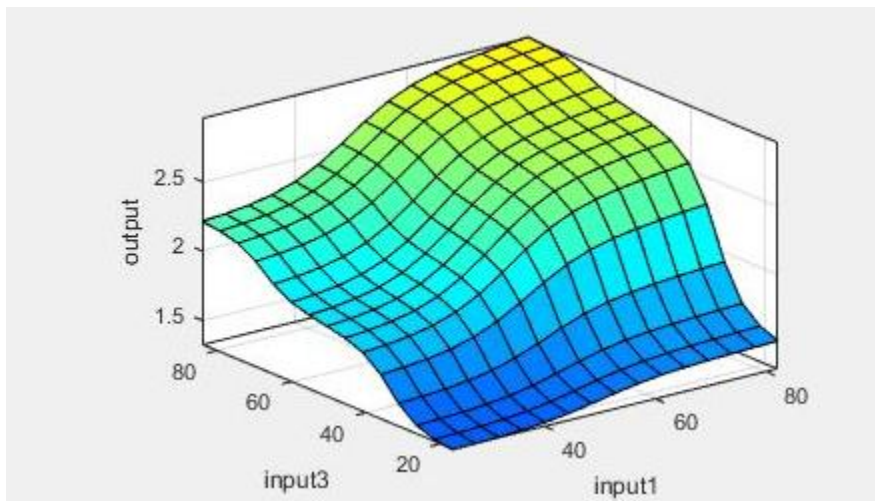


Fig (5b) Surface viewer of input 3 with respect to input 1

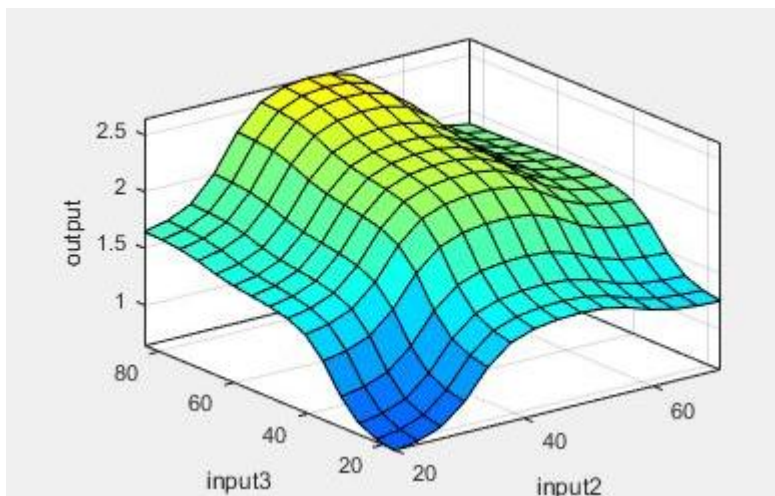


Fig (5c) Surface viewer of input 3 with respect to input 2

3. Results and Discussion

3.1. Screening of extraction parameters using Placket-Burman design criterion

The results of the PBD model indicate that three extraction parameters (temperature, extraction time and solvent: solid ratio) shows significant results with P value <0.05 for extraction of barbaloin (**Table 6**).

Table 6: Result of Regression analysis

Term	Effect	Coef	SE	Coef	T	P
Constant		0.567	0.0529	10.72	0.000	
X1	0.1567	0.078	0.0530	1.48	0.199	
X2	0.3733	0.186	0.0530	3.53	0.017	
X3	-0.6033	-0.302	0.0530	-5.70	0.002	
X4	0.4200	0.2100	0.0530	3.97	0.011	
X5	0.1133	0.057	0.0530	1.07	0.333	

3.2. Effects of Temperature, extraction time and solvent- solid ratio

To further investigate the most significant factors chosen from PBD and to optimize the barbaloin yield, a three-factor three-level CCD was used (**Table. 4**). The effect of following three factors X_1 -temperature, X_2 -extraction time and X_3 - solvent-solid ratio was studied on % barbaloin yield. The fitted models include only the statistically significant coefficients ($P < 0.05$) from **Table 7** are as follows:

$$Y = 2.6132 + 0.3332X_1 + 0.2180X_2 + 0.5639X_3 - 1.1970X_1^2 - 0.9492X_2^2 - 0.5902X_3^2 + 0.3799X_1 X_2 + 0.2322X_1 X_3 - 0.1316X_2 X_3 \dots\dots\dots (4)$$

Table 8 contains the result of analysis of variance of amount of barbaloin extracted from *Aloe vera* leaf.

Table 7: Regression analysis of CCD criterion data for barbaloin extraction from *Aloe vera* leaf skin

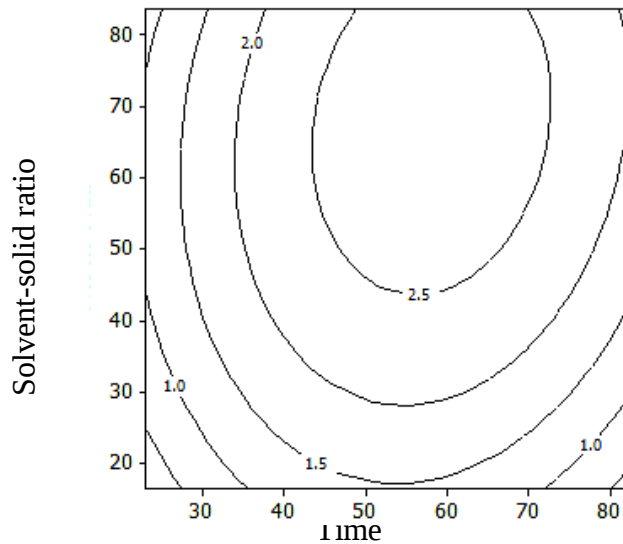
Term	Coef	SE Coef	T	P
Constant	2.6132	0.01920	136.132	0.000
Temperature	0.3332	0.02143	15.549	0.000
Time	0.2180	0.02141	10.186	0.000
Solvent:solid ratio	0.5639	0.02141	26.342	0.000
(Temperature) ²	-1.1970	0.03507	-34.127	0.000
(Time) ²	-0.9492	0.03505	-27.079	0.000
(Solvent:solid ratio) ²	-0.5902	0.03505	-16.837	0.000
Temperature*Time	0.3799	0.04705	8.075	0.000
Temperature*Solvent:solid ratio	0.2322	0.04705	4.935	0.001
Time*Solvent:solid ratio	-0.1316	0.04697	-2.802	0.019

Table8: Analysis of variance for barbaloin extraction from *Aloe vera leaf skin* using CDD criterion

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	9	6.64206	6.64206	0.73801	333.14	0.000
Linear	3	2.29977	2.30265	0.76755	346.47	0.000
Square	3	4.12649	4.12649	1.37550	620.90	0.000
Interaction	3	0.21580	0.21580	0.07193	32.47	0.000
Residual Error	10	0.02215	0.02215	0.00222		
Lack-of-Fit	5	0.01029	0.01029	0.00206	0.87	0.560
Pure Error	5	0.01187	0.01187	0.00237		
Total	19	6.66422				

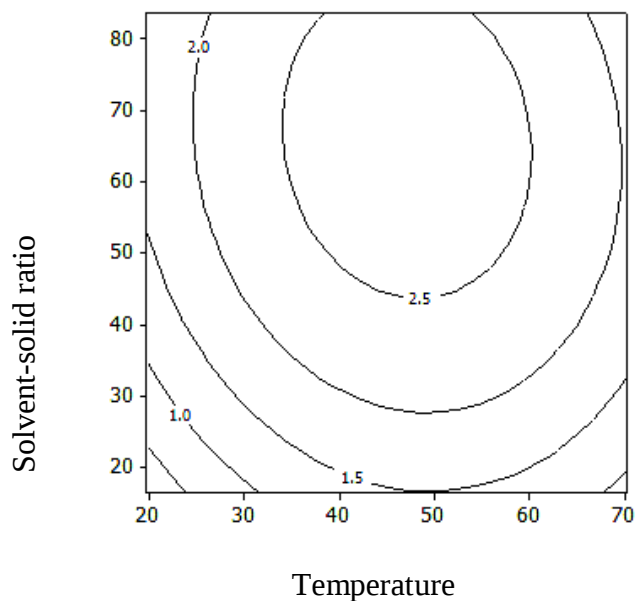
The response surfaces corresponding to the combined effects of three extraction parameters X_1 , X_2 , and X_3 are shown in **Fig.6**. Fig.6a show the contour plots of barbaloin extracted from *Aloe vera leaf skin* at varying level of solvent-solid ratio and time at fixed level of temperature (52.5 °C). Fig.6b show the contour plots of barbaloin extracted from *Aloe vera leaf skin* at varying level of solvent-solid ratio and temperature at fixed level of time (45 min). Fig.6c show the contour plots of barbaloin extracted from *Aloe vera leaf skin* at varying level of time and temperature at fixed level of solvent-solid ratio (50:1). It observed from Fig 6a, 6b, and 6c that maximum barbaloin is extracted at middle level of each pair of extraction parameters. Further increase in these factors above the middle

level showed decrease in barbaloin extraction. The experimental data were fitted in to equation (4) and the optimum values were found to be: temperature (52.5 °C) extraction time (45 min), solvent-solid ratio (56.42:1). At these optimum levels of extraction parameters barbaloin yield in *Aloe vera* leaf was 2.695 %, which is very close to the predicted value of 2.726 %.



Hold values: Temperature (52.5 °C)

Fig 6a. Contour plot for barbaloin yield at varying level of solvent-solid ratio and time



Hold values: Time (45 min)

Fig. 6b. Contour plot for barbaloin yield at varying level of solvent-solid ratio and temperature

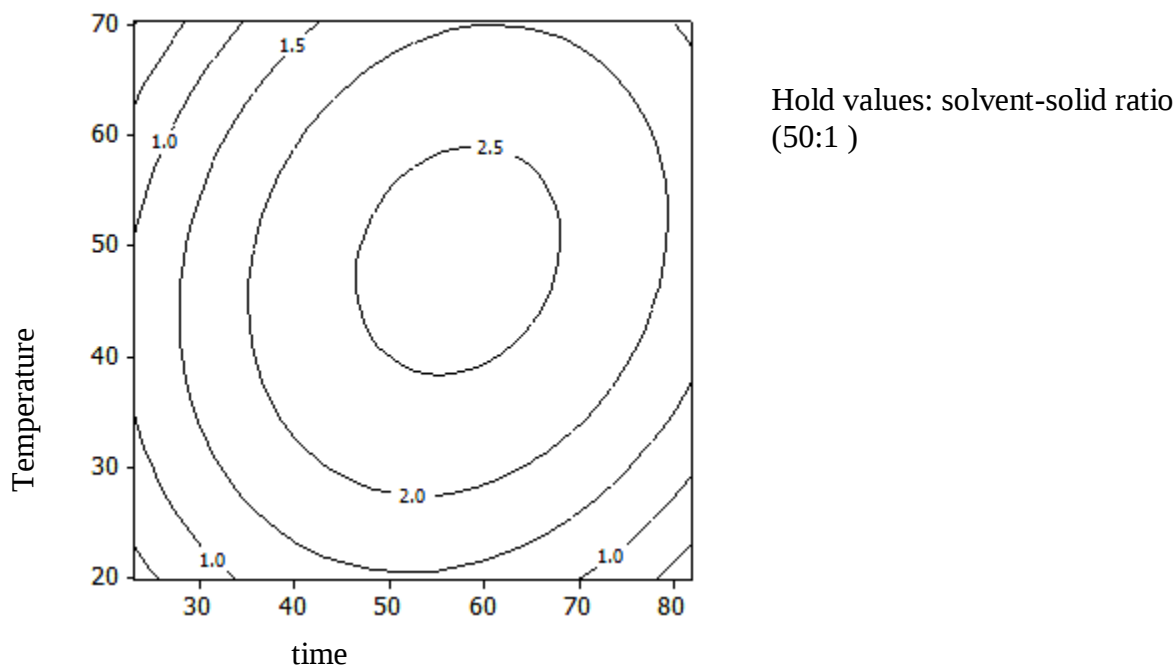


Fig.6c. Contour plot for barbaloin yield at varying level of time and temperature

3.3. Comparison of RSM, ANFIS and RBF-SVR Predicted Values

The RBF network with one hidden layer of RBF neurons and Gaussian basis function of one dimensional input is used. In ANFIS model, triangular membership function was selected for the input variable and output variable. The result predicted from the RBF network and ANFIS model were given in **Table 4**.

All the three models RSM, RBF and ANFIS had been used to optimize the operating conditions. Although all the three models showed good quality in predicting the combined interactions of the independent variables with respect to barbaloin yield, yet the ANFIS results dominated the RSM and RBF-SVR results. The correlation coefficient of RSM, RBF-SVR and ANFIS with experimental data is 0.931, 0.986 and 0.971 respectively. Although RSM used a regression equation approach but SVR-RBF

better than RSM. The RBF approach is flexible and trustable model and more rational to interpret solid-solvent extraction. However, the RBF does not deal with imprecise knowledge. This limitation is overcome by ANNFIS system.

4. Conclusion

RSM was effectively deployed to identify the optimized extraction conditions for barbaloin from *Aloe vera* leaf. It is observed from the result that the second order polynomial model effectively described the experimental data. The five parameters extraction were optimized and then given to prediction model. Among all the five parameters the temperature, time and solvent-solid ratio shows significant effect on extraction of barbaloin.

In this work, experiments were conducted to compare the result of all the models. Among all the models the radial based function network showing more accuracy in estimation over RSM and ANFIS shows slight more accuracy over RBF network. The main drawback of Regression-RSM is that it requires the order of the model to be stated, while RBF network tends to implicitly find the relationship between input and output vector. The major drawback of RBF-SVR is that it does not deal with imprecise knowledge which is overcome with ANFIS.

Acknowledgements

We are thankful to Lovely Professional University, Phagwara.

Reference:

- [1] The Wealth of India, "A Dictionary of Indian Raw Materials and Industrial Products National Institutes of Science Communication", Council of Scientific and Industrial Research, New Delhi, 1:A-Ci (Revised), 2000, pp. 47–49.
- [2] Q.J. Groom, T. Renolds, "Barbaloin in aloespecies", *Planta Medica*, vol. 52, 1986, pp. 345- 348.

- [3] Y. Gutterman, E. Chauser-Volfson, “Peripheral defence strategy: Variation of barbaloin content in the succulent leaf parts of *Aloe arborescens* Miller (Liliaceae)”. Botanical Journal of the Linnean Society, vol. 132, 2000, pp. 385–395.
- [4] K. Nakagomi, S. Oka, N. Tomizuka, M. Jamamoto, T. Masui, H. Nakazawa, “Novel biological activities of aloe components. Effects on mast-cell degranulation and platelet aggregation. Kenkyu Hokoku—Koyo Gijutsu in Bisibutsu Kogyo, vol. 63, 1985, pp. 23–29.
- [5] W.G. Wamer, P. Vath, D.E. Falvey. “In vitro studies on the photobiological properties of aloe emodin and aloin A”. Free Radical Biology and Medicine, vol. 34, 2003, pp. 233–242.
- [6] H. Avila, J. Rivero, F. Herrera, G. Fraile, “Cytotoxicity of a low molecular weight fraction from *Aloe vera* (*Aloe barbadensis* Miller) gel”. Toxicon, vol. 35, 1997, pp. 1423–1430.
- [7] Q.J. Groom, T. Reynolds, “Barbaloin in *Aloe* species”, Planta Medica, vol. 52, 1986, pp. 345–348.
- [8] Y. Gutterman and E. Chauser-Volfson, “Peripheral defence strategy: Variation of barbaloin content in the succulent leaf parts of *Aloe arborescens* Miller (Liliaceae)”, Botanical Journal of the Linnean Society, vol. 132, 2000, pp. 385–395.
- [9] Y. Ishii, H. Tanizawa and Y. Takino, “Studies of *Aloe* V. Mechanism of cathartic effect”, Biological Pharmacology Bulletin, vol. 17, 1994, pp. 651–653.
- [10] D. Wolfe, C. Schmutte, J. Westendorf, “Hydroxyanthraquinones as tumor promoters: Enhancement of malignant transformation of C3H mouse fibroblasts and growth stimulation of primary rat hepatocytes”. Cancer Res, vol. 50, 1990, pp. 6540–6544.
- [11] W.G. Wamer, P. Vath and D.E. Falvey, “In vitro studies on the photobiological properties of aloe emodin and aloin A”, Free Radical Biology and Medicine, vol. 34, 2003, pp. 233–242.
- [12] C.W. Zobel, D.F. Cook, “Evaluation of neural network variable influence measures for process control”. Eng. Appl. Artif. Intell., vol. 24, 2011, pp. 803–812.
- [13] M. Buragohain, “Adaptive Network based Fuzzy Inference System (ANFIS) as a Tool for System Identification with Special Emphasis on Training Data

Minimization”, PhD Thesis submitted to Department of Electronics and Communication Engineering Indian Institute of Technology Guwahati, 2008.

[14] M. Kumara, G. Singh, S. K. Aryab, J. S. Bhatti, P. Sharmad, “Artificial Neuro-Fuzzy Inference System (ANFIS) based validation of laccase production using RSM model”, *Biocatalysis and Agricultural Biotechnology*, vol. 14, 2018, pp. 235–240.

[15] M. Khajeh, M. G. Moghaddam, M. Shakeni, “Application of artificial neural network in predicting the extraction yield of essential oils of *Diplotaeniachrydifolia* by supercritical fluid extraction”. *J. Supercrit. Fluids*, vol. 69, 2012, pp. 91–96.

[16] P. Mitra, P. C. Barman, K. S. Chang, “Coumarin extraction from *Cuscutareflexa* using supercritical fluid carbon dioxide and development of an artificial neural network model to predict the coumarin yield”. *Food Bioprocess Technol.*, vol. 4, 2011, pp. 737–744.

[17] S. M. Nemes, V. Orsat, G. S. V. Raghavan, “Calibration of artificial neural network and partial least squares regression models for the prediction of secoisolariciresinoldiglucoside contents in microwave-assisted extracts of various flaxseed (*Linum usitatissimum* L.) samples”. *Food Chem.*, vol. 133, 2012, pp. 1588–1595.

[18] F. Pouralinalazar, M. A. C. Yunus, G. Zahedi, “Pressurized liquid extraction of *Orthosiphon stamineus* oil: experimental and modeling studies”. *J. Supercrit. Fluids*, vol. 62, 2012, pp. 88–95.

[19] G. Zahedi, A. Azarpour. “Optimization of supercritical carbon dioxide extraction of Passiflora seed oil. *J. Supercrit. Fluids*, vol. 58, 2011, pp. 40–48.

[20] K. Sinha, S. Chowdhury, P. Das Saha, S. Datta. “Modeling of microwave-assisted extraction of natural dye from seeds of *Bixa orellana* (Annatto) using response surface methodology (RSM) and artificial neural network (ANN)”. *Ind. Crop. Prod.*, vol. 41, 2013, pp. 165–171.

- [21]K. Sinha, P. Das Saha, S. Datta. “Response surface optimization and artificial neural network modeling of microwave assisted natural dye extraction from pomegranate rind”. Ind. Crop. Prod., vol. 37, 2012, pp. 408–414.
- [22]H. Liu, K. Wang, H. Xu,X. Chen, Z.Hu, “Application of Experimental Design and Artificial Neural Networks to Separation and Determination of Active Components in Traditional Chinese Medicinal Preparations by Capillary Electrophoresis”. Chromatographia, vol. 55, 2002, pp.579-583.
- [23]M.Rajasimman, S.Subathra,“Optimization of Gentamicin Production: Comparison of ANN and RSM Techniques”, World Academy of Science, Engineering and Technology, vol. 27, 2009, pp. 533-539.
- [24] D.K.Pandey, S. Paridaand A. Dey,. Comparative HPTLC analysis ofbioactive marker barbaloin from in vitro and naturally grown Aloe vera (L.) Burm. f. RevistaBrasileira de Farmacognosia - 26(2), 2016,161-16.