British Journal of Applied Science & Technology
4(21): 3039-3052, 2014

SCIENCEDOMAIN *international*
www.sciencedomain.org

Optimization of Process Variables for C-massecuite Exhaustion in a Nigerian Sugar Refinery

M. O. Aremu[1], D. O. Araromi[2*], J. A. Adeniran[2] and O. S. Alamu[1]

[1]Biochemical Engineering Laboratory, Department of Chemical Engineering, Ladoke Akintola University of Technology, Ogbomoso, Oyo State, Nigeria.
[2]Optimization and Control Unit, Department of Chemical Engineering, Ladoke Akintola University of Technology, Ogbomoso, Nigeria.

Authors' contributions

This work was carried out in collaboration between all authors. All authors write, read and approved the final manuscript.

Original Research Article

Received 20th February 2014
Accepted 3rd April 2014
Published 29th May 2014

ABSTRACT

Sucrose in the final molasses continues to be a source of major financial loss to sugar refineries worldwide. This study therefore aims at rectifying this anomaly. In this study, the final molasses exhaustibility was predicted using Adaptive Neuro Fuzzy Inference System (ANFIS) and Response Surface Methodology (RSM) based ondata generated from molasses sample collected from the recovery end of refining processes. The results show that both models are able to predict the final molasses exhaustibility with sufficient accuracy. The optimum sucrose recovery of 49.18% was achieved at the point when Brix^0 is 96.00%, Purity of 65.00% and pH of 4.50. Also, both models agree on the combination of purity and pH as the two factors interaction that have optimal effect on the sucrose recovery. The correlation coefficient (R^2) value obtained for ANFIS was 0.96 while that of RSM was 0.99. Thus, the RSM model has better prediction performance than ANFIS.

Keywords: Adaptive Neuro Fuzzy Inference System (ANFIS); Response Surface Methodology (RSM); design expert; molasses exhaustion; purity; sucrose.

**Corresponding author: E-mail: doararomi@lautech.edu.ng;*

1. INTRODUCTION

Sugar production from sugarcane and sugar beets has been improved upon today by research studies carried out by intellectuals in minimizing sucrose loss at the recovery end of refining processes. The main goal of sugar cane factory is to have an efficient and profitable operation with the required sugar quality and maximum sugar recovery. The purity of molasses (sucrose % in total solid molasses) is a very important process variable. It is a measure of sugar content in molasses and therefore of the factories sugar loss; the higher the purity, the higher this loss is[1].Currently, final molasses is usually produced at about 3.50 to 4.00% of cane by weight. In order to minimize the loss of sugar in the final molasses it is essential to consider the factors responsible for the varying degrees of molasses exhaustibility [2].

The problem of molasses exhaustion and the factors influencing the same can be considered from two angles namely; compositional and operational factors. The compositional factors depend on the quality of cane and juice clarification efficiency. The most important factor determining the exhaustion potential of molasses is the reducing sugar/ ash ratio. The higher this ratio, lower should be loss of sucrose in molasses. This ratio should be 1.0 or more for good exhaustion of final molasses [2]. The operational factors, also called controllable factors, affecting exhaustibility of final molasses are the practices and operations mainly at the low grade massecuite processing.

In order to assess the degree of exhaustion achieved, a benchmark is necessary. Target purity of final molasses is an accepted concept designed to provide a practical benchmark of factory molasses exhaustion [3]. Over the years, several researchers had proposed target purity formula and equations for minimizing loss of sucrose in final molasses [4,5,6,7,8,9]. In 2013, [10] applied the target purity formula proposed by ASI, 1993 to quantify molasses exhaustion at Dangote Sugar refinery, Nigeria. Target purity is a reference or equilibrium purity of molasses taking into account the effect of the non-sucrose present on its exhaustion and represents the minimum value that can be achieved according to the process thermodynamics.[11].

In conventional sugar technology, target purity has been modeled with equations which only take into account the contents of monosaccharide (glucose, fructose) and ash, hence not providing a good correlation.Goza-Leon et al. [11], developed a model for final molasses purity of Cuban sugar factory using an artificial neural network (ANN) rather than the conventional approach. From his work, he was able to take into account a greater number of non-linear variables and dealt with glucose, fructose and ash separately as compared with the conventional modeling of reference purity of the molasses. In 2011, [1] researched on process optimization to reduce sucrose loss with final molasses at Metahara Sugar factory, Ethiopia using response surface methodology (RSM). This paper aims at comparing model analysis for final molasses exhaustion developed using RSM and ANFIS.

2. METHODOLOGY

2.1 Description of Study Area

This research work covers activities taken place at a foremost refinery in Nigeria. Raw sugar importation is the chief source of refined sugar in Nigeria. At the plant, the sucrose available in the liquor is crystallized in several stages, conducted at descending purities. The molasses gotten from each massecuite (Refined massecuite, special A, A, B or C massecuite) were boiled depending on purity gotten. This process was continuous until the C molasses purity was low, therefore sent out as final molasses This is because the liquor becomes more and more exhausted of sugar and crystallization becomes more difficult, since the non- sugars in the liquor inhibit sugar crystallization. The sugar in the raw sugar typically makes up 98% to 99% of the dissolved solid. The final molasses in the plant has sucrose content of between 40% and 45%, which always translate to between 15% and 20% of the entering raw sugar. The true purity (sucrose/dissolved solids) of the final molasses in the refinery is also between 45% and 60% as against the 35% and 45% range that are obtained in places like South Africa and Louisiana.

2.2 Data Collection

The molasses sample used in this work was collected at recovery end (low grade C-massecuite), of the refinery which is the most critical operational stage in sugarcane refining. The sample was analysed in the laboratory for Brix, Purity of C-massecuite, pH of C-massecuite and sucrose recoveryand the data obtained were used for the optimization studies. The input variables are Brix of C-massecuites, Purity of C-massecuite and pH of C-massecuite, while the sucrose recovery is the output variable. These data were collected over a period of eight months from January, 2013 to August, 2013. The total data points were 181.

2.3 Adaptive Neuro Fuzzy Inference System (ANFIS)

ANFIS is a hybrid integrated neuro-fuzzy model which makes use of the complementary strength of Artificial Neural Network (ANN) and Fuzzy Inference System (FIS). ANFIS proposed by Jang [12] is based on the first-order Sugeno fuzzy model that uses either a back propagation algorithm alone or a hybrid learning algorithm. ANFIS networks have been successfully applied to classification tasks, rule-based process controls, pattern recognition problems and the like.

ANFIS have been widely applied to predict process parameters or effluent parameters for aerobic biological treatment processes [13], for predicting suspended solids (SS) and COD in the effluent from hospital WWTP [14]. Some models have been developed by using ANFIS for effluents only and for single parameter output for a particular treatment unit. Models considering the main treatment units separately and estimating multiple parameters have not been sufficiently developed yet. Studies are generally based on two or three parameters input and single output.

An ANFIS can be viewed as a special three-layer feed forward neural network. The first layer represents input variables, the hidden layer represents fuzzy rules and the third layer is for output. Fig. 1 represents a typical ANFIS architecture. For simplicity, a fuzzy inference system has two inputs x and y and one output is assumed. For a first-order Sugenofuzzy model, a common rule set with two fuzzy if–then rules is defined as:

Rule 1 If x is $x1$ and y is $y1$, then $f1 = p1\,x + q1\,y + r1$

Rule 2 If x is $x2$ and y is $y2$, then $f2 = p2\,x + q2\,y + r2$

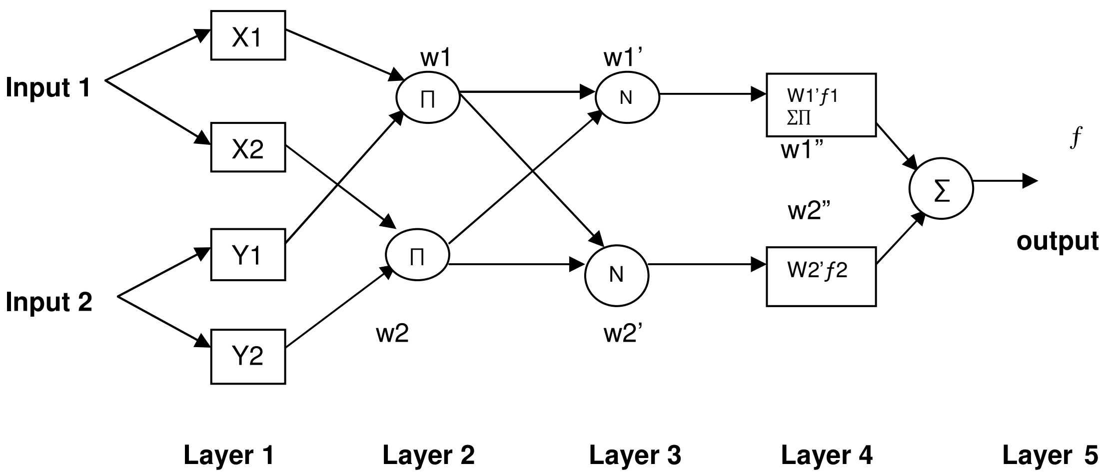

Fig. 1. Adaptive Neuro Fuzzy Inference System structure
(Source: [15])

2.4 Modeling

The collected data sets were treated to remove outliers. The data pre-processing was accomplished by removing measurements that were not within the range of$\pm 2\sigma$, standard deviations around the group or design cell mean (Statistical package version 6). Further pre-processing of the data was done using smoothing technique. The pre-processed data were partitioned into training and checking data of equal parts. These data were loaded into the ANFIS graphical user interface (GUI) of Fuzzy logic toolbox [16] for training of Fuzzy Inference System (FIS) using grid-partitioning and validation of the FIS generated. Exhaustive search was performed on the data. Exhaustive search is a combinatorial function which selects the required number of inputs combination that has optimal effect on the output performance. The selected model was then optimised using hybrid optimization method because it combines the least square and back propagation gradient descent method. Regressors were formulated from the selected input combinations to build ANFIS structure. The ANFIS structure was trained using 2 MF, different input MF type and output MF type 'constants' as benchmark. The selected model was then validated using RMSE and correlation coefficient. The validation was done using MF number of 2, different input MF type and output MF type 'constant.'

2.5 Response Surface Methodology (RSM)

RSM is used to generate mathematical model by considering input parameters to the response generated through experiments.

$$Y = f(X_1, X_2, X_3, \ldots X_n) \pm \varepsilon \tag{1}$$

Where Y is the response, *f* is the unknown function of response, X_1, X_2, X_3,…, X_4 are the input variables which can affect the response, n is the number of the independent variables and ε is the statistical error that represents other sources of variability not accounted for by *f*. The least square technique was used to fit a model equation containing the input variables by minimizing the residual error measured by the sum of square deviations between the actual and the estimated responses.

In this study, a three-variable Box Behnken experimental design was employed. The chosen variables were obtained from the data collected at DSR plant. The input variables are Brix, purity and pH of C-massecuite while the response is purity of final molasses. The Brix has a minimum value of 94 and high value of 98. The purity of C-massecuite ranges from 55 to 65 and low pH value was 4.5 while 5.5 was the high value.

Table 1 show the factor design of the three input variables and response (Sucrose recovery) obtained from the experimental work done in the refinery. This table was obtained by inserting minimum and maximum values into Design expert response surface Box BehnkenProgram.

Table 1. The Box Behnken experimental design for the three variables with three levels employed for C-Massecuite optimization

Run	Factor 1 A:brix	Factor 2 B:purity	Factor 3 C:pH	Response 1 Sucrose recovery
1	96	65	5.50	43.47
2	96	60	5.00	46.83
3	96	60	5.00	46.83
4	98	60	4.50	48.66
5	96	60	5.00	47.50
6	98	65	5.00	46.21
7	96	65	4.50	49.18
8	96	60	5.00	46.83
9	98	55	5.00	46.40
10	96	60	5.00	46.83
11	98	60	5.50	43.96
12	96	55	5.50	47.68
13	94	60	5.50	47.19
14	94	60	4.50	47.50
15	94	65	5.00	46.44
16	96	55	4.50	46.97
17	94	55	5.00	48.25

2.6 Model Estimation

The model performance predicted by ANFIS was evaluated using correlation coefficient (R) and Root Mean Square Error (RMSE).

$$R = \frac{\sum(obs-obs^{'})(pre-pre^{'})}{\sqrt{\sum(obs-obs^{'})^2\sum(pre-pre^{'})^2}} \tag{2}$$

$$RMSE= \sqrt{\frac{1}{n}}\sum(obs\text{-}pre)^2 \tag{3}$$

(Source: [15])

Where obs = observed values; Pre=predicted values; obs^{i}=average value of observed values and Pre^{i} = average value of predicted values.

In the RSM model developed, the analysis of variance (ANOVA) including sequential F-test and lack of fit test were used to assess the performance of the model. Also, residual analysis and diagnostics case statistics were checked to ensure adequacy of the model. The quality of fit of the polynomial model was expressed by the coefficient of determination (R^2) ,adjusted coefficient (R^2_{Adj}) and predicted coefficient (R^2_{Pred}).

$$R^2 = 1 - \frac{SS_{residual}}{SS_{model} + SS_{residual}} \tag{4}$$

$$R^2_{Adj} = 1 - \frac{SS_{residual}/DF_{residual}}{(SS_{model} + SS_{residual})/(DF_{model} + DF_{residual})} \tag{5}$$

$$R^2_{Pred} = 1 - \frac{PRESS}{(SS_{model} + SS_{residual})} \tag{6}$$

Where SS means sum of square, DF means degree of freedom and PRESS means Predicted residual sum of squares.

3. RESULTS AND DISCUSSION

3.1 ANFIS Modelling Analysis

Exhaustive search of 1 input was performed on the collected data to know the input that has optimal effect on the purity of final molasses.

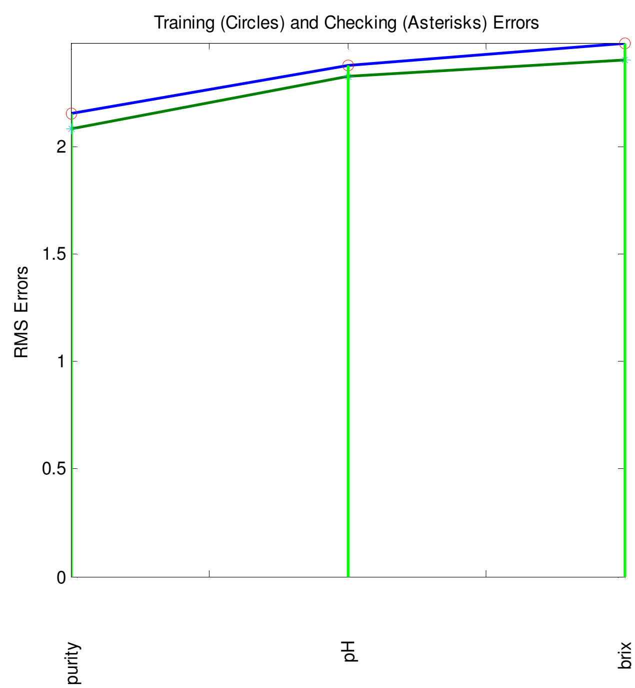

Fig. 2. Plot of RMS Errors of input parameters

From Fig. 2, the input to the left-most part of the plot is the most significant. The purity of C-massecuite was found to have the least training error of 2.15 and checking error of 2.08.

Therefore it is the input variable that is most influential on the output performance. The second most important variable is pH of C-massecuite with a training error of 2.38 and checking error of 2.33 while the less important variable is Brix having training error of 2.48 and checking error of 2.41.

Two input variables were combined to determine which input combination has optimal effect on the output performance.

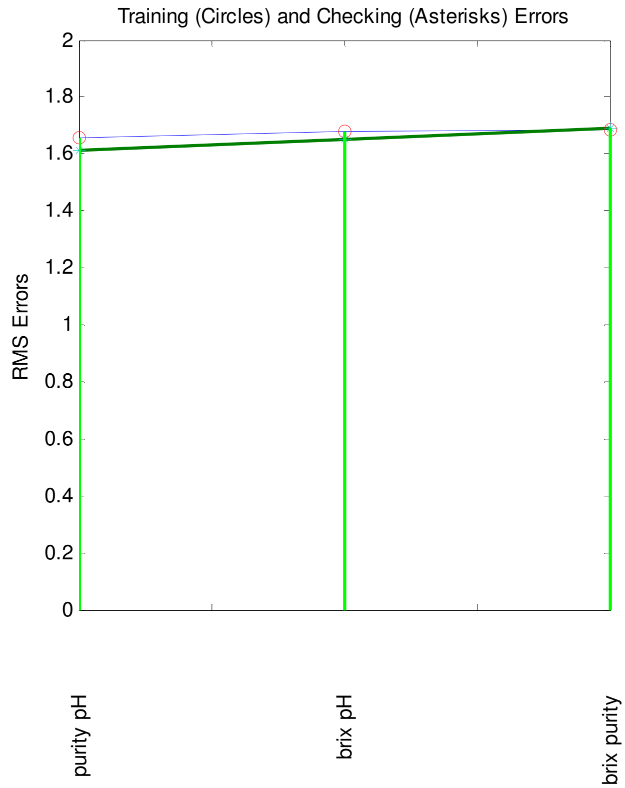

Fig. 3. Plot of RMS errors against the input parameters combination

Fig. 3 shows the combinatorial effect of two inputs on the output performance. The left-most input combination which is the most significant is combination of purity and pH.

Table 2. ANFIS training showing result of two inputs combination

Input combination	Training error	Checking error
Brix and purity	1.69	1.69
Brix and pH	1.68	1.65
Purity and pH	1.66	1.62

From Table 2, the two inputs combination which is most significant to the output performance, that is, purity of final molasses is purity and pH because it has the least training error of 1.65 and checking error of 1.61.

Table 3. Model estimation result showing RMSE and R-values

Inputmf type	MF no: 2 , Epoch = 1000		
	RMSE		(R-value)
	Training	Checking	
Trimf	0.74	0.84	0.96
Trapmf	Nil	nil	Nil
Gbellmf	0.80	0.83	0.96
Gaussmf	0.80	0.85	0.96
Gauss2mf	0.80	0.86	0.95
Pimf	0.74	0.90	0.95
Dsigmf	0.79	0.84	0.95
Psigmf	0.79	0.84	0.95

Table 3 shows the model validation at different input Mf type. The highest R-value was gotten at ‘Trimf’ and low RMSE value for the training data. ‘Gbellmf’ gave the least RMSE difference but with high training error and low R-value when compared with ‘Trimf.’ Therefore, ‘Trimf’was the selected input MF type for the prediction showing R-value of 0.96.

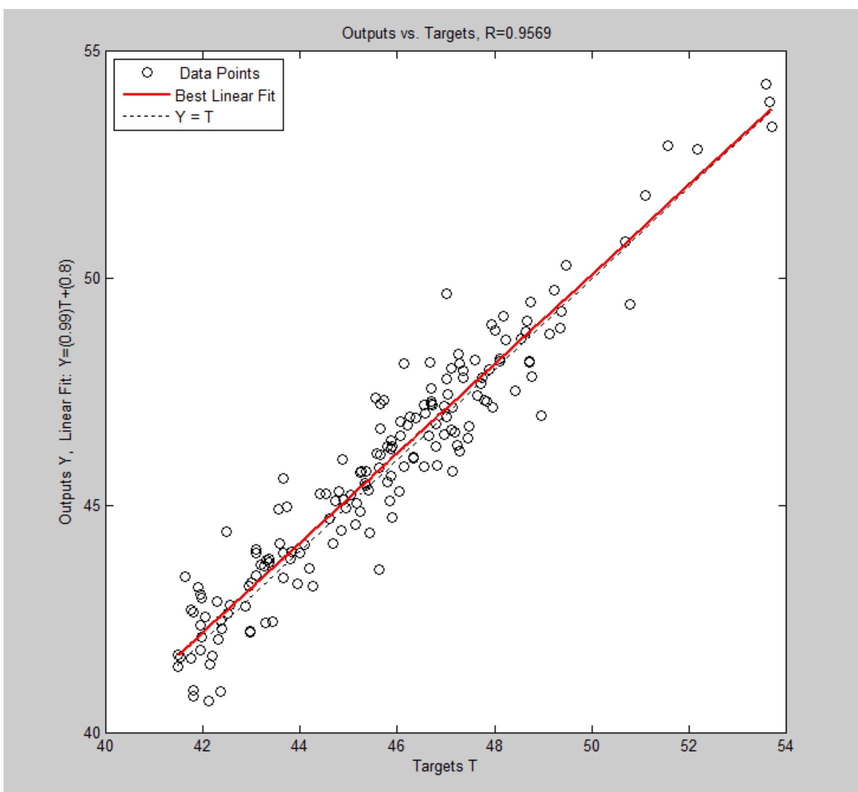

Fig. 4. A plot of predicted against actual values of the output parameter

Fig. 4 shows the plot of actual output parameter against the predicted value having a correlation coefficient value of 0.96. The data points cluster round the straight line showing less disparity and a good model prediction.

3.2 RSM Modelling Analysis

3.2.1 Model fitting

Table 4a. ANOVA for response surface quadratic model

Source	Sum of squares	Mean square	F value	p-value Prob > F	
Model	32.52	3.61	69.91	<0.0001	Significant
X_1-brix	2.15	2.15	41.65	0.0003	
X_2-purity	2.00	2.00	38.69	0.0004	
X_3-Ph	12.53	12.53	242.33	<0.0001	
X_1X_2	0.66	0.66	12.69	0.0092	
X_1X_3	4.82	4.82	93.22	<0.0001	
X_2X_3	10.30	10.30	199.36	<0.0001	
X_1^2	0.02	0.02	0.37	0.5630	
X_2^2	0.02	0.02	0.37	0.5630	
X_3^2	0.02	0.02	0.37	0.5630	
Residual	0.36	0.05			
Lack of Fit	0.00	0.00	0.00	1.0000	not significant
Pure Error	0.36	0.09			
Cor Total	32.88				

From Table 4a, the input with highest F-value and minimum prob>F value is input C which is the pH of C-massecuite. This means that pH is the most influential factor on the RS 2FI model. The best two input combination is the one with highest F-value which is 199.36 and that is input BC which is purity and pH of C-massecuite. The model F-value of 69.91 implies the model is significant. Values of "Prob > F" less than 0.0500 indicate model terms are significant.In this case X_1, X_2, X_3, X_1X_2, X_1X_3, X_2X_3 are significant model terms. Values greater than 0.1000 indicate the model terms are not significant.

Table 4b. Model estimation result

Std. Dev.	0.23	R-Squared	0.99
Mean	46.87	Adj R-Squared	0.98
C.V. %	0.485	Pred R-Squared	0.98
PRESS	0.57	Adeq Precision	32.76

From Table 4b, the "Pred R-Squared" of 0.98 is in reasonable agreement with the "Adj R-Squared" of 0.98."Adeq Precision" measures the signal to noise ratio. A ratio greater than 4 is desirable. The model ratio of 32.76 indicates an adequate signal. This model can be used to navigate the design space. The model equation in terms of actual factor is given as;

Sucrose recovery $= -567.099 + 6.026125 * X_1 - 0.4552 * X_2 + 144.0675 * X_3 + 0.0405 * X_1 * X_2 - 1.0975 * X_1 * X_3 - 0.642 * X_2 * X_3 - 0.01681 * X_1^2 - 0.00269 * X_2^2 - 0.269 * X_3^2.$

Where X_1= Brix, X_2= Purity and X_3=pH of C-massecuite.

Table 5. Model analysis performance

Response	1	Susrose Recovery			Transform:	None			
Diagnostics Case Statistics									
					Internally	**Externally**	**Influence on**		
Standard order	**Actual value**	**Predicted value**	**Residual**	**Leverage**	**Studentized residual**	**Studentized residual**	**Fitted value dffits**	**Cook's distance**	**Run order**
1	48.25	48.25	0	0.75	0	0	0	0	17
2	46.40	46.40	0	0.75	0	0	0	0	9
3	46.44	46.44	0	0.75	0	0	0	0	15
4	46.21	46.21	0	0.75	0	0	0	0	6
5	47.50	47.50	0	0.75	0	0	0	0	14
6	48.66	48.66	0	0.75	0	0	0	0	4
7	47.19	47.19	0	0.75	0	0	0	0	13
8	43.96	43.96	0	0.75	0	0	0	0	11
9	46.97	46.97	0	0.75	0	0	0	0	16
10	49.18	49.18	0	0.75	0	0	0	0	7
11	47.68	47.68	0	0.75	0	0	0	0	12
12	43.47	43.47	0	0.75	0	0	0	0	1
13	46.83	46.96	-0.1345	0.2	-0.66	-0.63	-0.32	0.01	10
14	47.50	46.96	0.538	0.2	2.65	0.00	0	0.18	5
15	46.83	46.96	-0.1345	0.2	-0.66	-0.63	-0.32	0.01	3
16	46.83	46.96	-0.1345	0.2	-0.66	-0.64	-0.32	0.01	8
17	46.83	46.96	-0.1345	0.2	-0.66	-0.63	-0.32	0.01	2

Table 5 gave experimental data values as against the actual values.

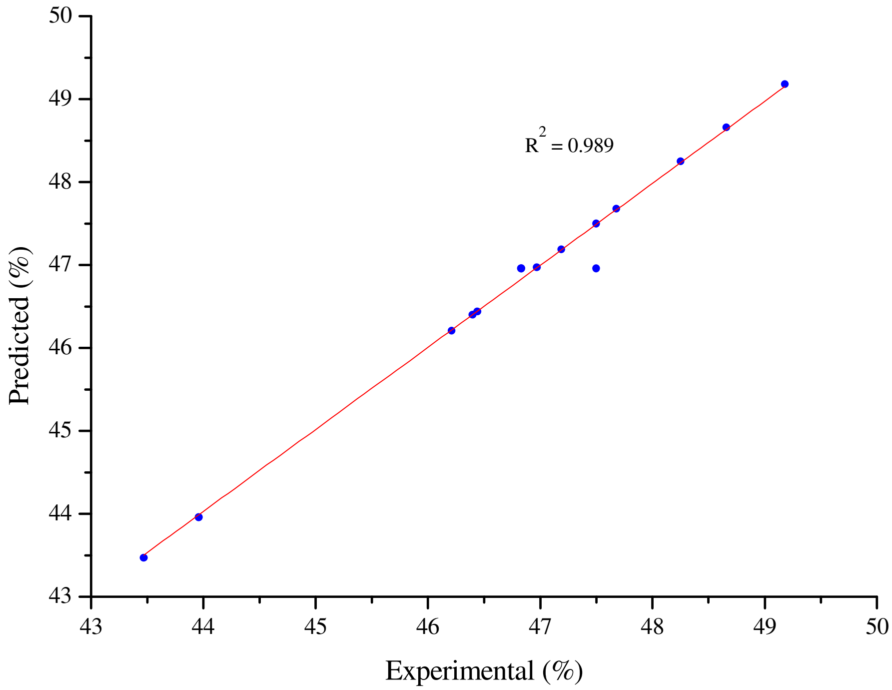

Fig. 5. Plot of predicted against actual responses

Fig. 5 shows the plot of the predicted versus actual responses and a close agreement between experimental and predicted values was indicated. The R^2 value is 0.989.

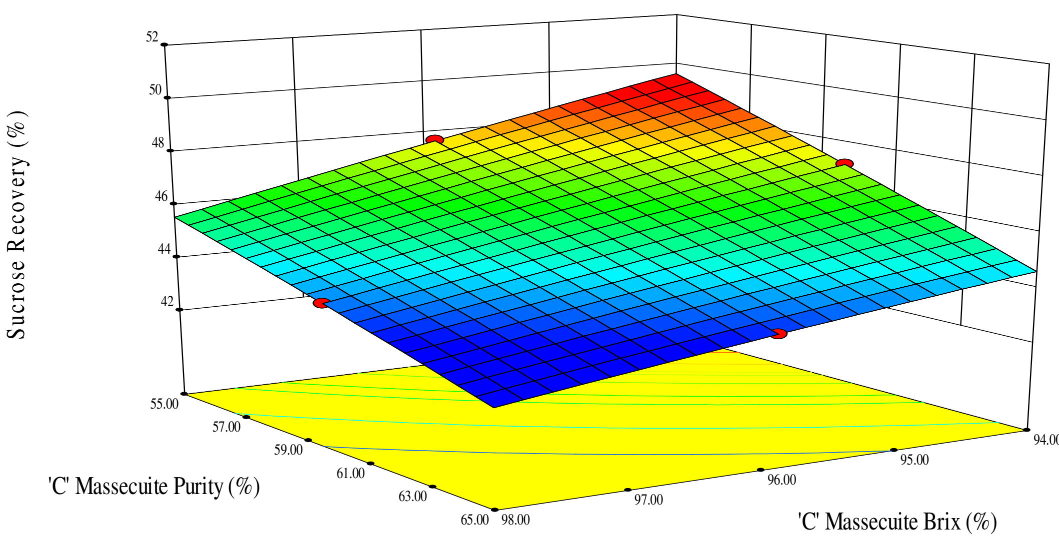

Fig. 6. 3D plot of two factors interaction against sucrose recovery

Fig. 6 gave the interactions of two factors (Brix and purity) as against the sucrose recovery. At high pH, high purity value of 65, the sucrose recovery decreases around 46 as brix value increases. Also, at low pH, low purity value of 55, sucrose recovery decreases from around 50.5 to 48 region as the brix increases. Furthermore, at high pH, high purity value of 65, the sucrose recovery increases from 48 upward as brix value increases. At low purity value of 55, sucrose recovery tends to be linear as brix increases.

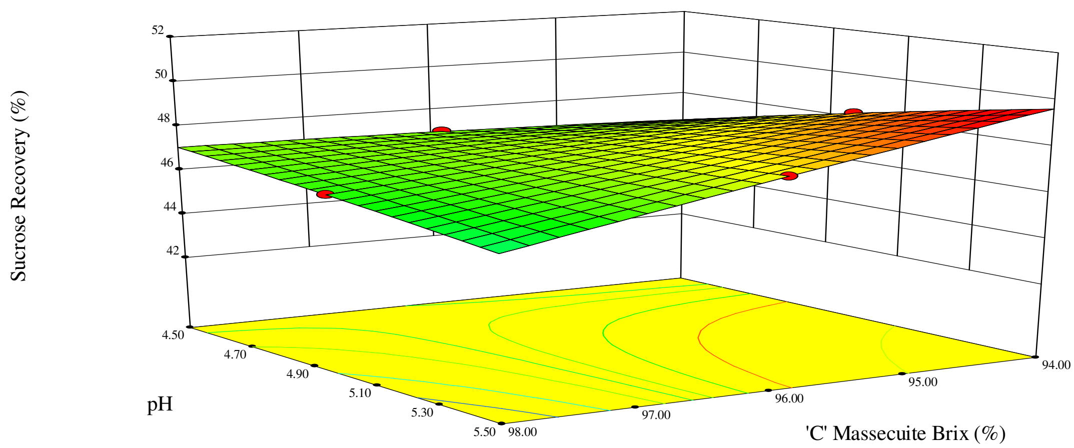

Fig. 7. 3D plot of two factors interaction (pH and Massecuite Brix)

Fig. 7 show the interaction of two factors (pH and massecuite brix) and their behavior with sucrose recovery. At low purity and high pH value, sucrose recovery decreases around 50-48.5 region as the brix increases. At low purity, low pH value, sucrose recovery increases within 46-48 region as the brix increases. Also, at high purity value, low pH value, sucrose recovery increases at 48% and above at increasing brix. At high purity, high pH value, sucrose recovery decreases below 46% as brix increases.

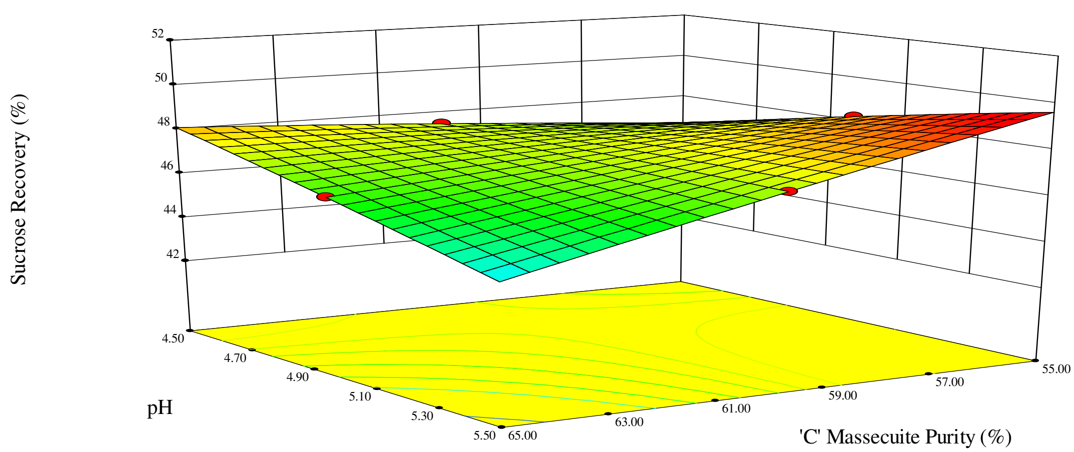

Fig. 8. 3D plot of two factors interaction (pH and purity)

Fig. 8 show the plots of two factors interaction (pH and purity) as against the sucrose recovery. At low brix, high pH value (5.50), sucrose recovery decreases as massecuite purity increases. And at low pH value (4.50), sucrose recovery increases as purity increases. At high brix, high pH value (5.50), sucrose recovery increases from 46 – 51% as massecuite purity increases.

3.3 Comparison between ANFIS and RSM Model

ANFIS and RSM are both data driving model. RSM is an optimization method of process variables gotten from experimental works. ANFIS is a black box modeling approach which requires sufficient data for a better model development. RSM is more statistical in analyzing and optimizing process variables. In this study, ANFIS model developed gave a correlation coefficient R-value of 0.96 while RSM gave a R^2 –value of 0.99.

4. CONCLUSION

Final molasses of the refinery were not well exhausted with reference to the high purity values obtained for the data collected. This study focused on the comparative study of final molasses exhaustibility using Adaptive Neuro Fuzzy Inference System (ANFIS) model and Response surface methodology (RSM).

The input parameters were examined by the two models to determine which input parameter was the most significant to give better output performance. The input parameter that has the optimal effect on purity of final molasses as determined by ANFIS is purity of C-massecuite while pH was the most influential parameter as determined by RSM. Considering two input interaction on the output performance, RSM and ANFIS predicted the combination of purity and pH of C-massecuite as the best input combination. The optimal conditions determined by RSM were 97.51 brix, 64.61 purity value and 5.42 pH value. The results obtained showed that the predicted and experimental values were not significantly different. For the ANFIS prediction, the predicted values on average gave a value of 45.66 and the actual average value is 45.82. The correlation coefficient R^2-value for validating ANFIS model prediction was 0.96 while the R^2 value for validating RSM model prediction was 0.99. RSM gave a better prediction of final molasses purity than the ANFIS model.The optimum process parameters suggested in this study for the recovery of sucrose from C-massecuite should be put in practice and strictly controlled. The models developed can be used for process behaviors prediction for performance measure, for process optimization and for training tools for operators

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Kassahun M. ‘Process optimization to reduce sucrose loss with final molasses’,Published Master of Science thesis. School of Graduate Studies, Addis Ababa University, Ethiopia; 2011.
2. Kulkarni DP. Cane sugar manufacture in India. The Sugar Technologists' Association of India, New Delhi, India. 2005;292-305.

3. Saska M, Goudeau S, Andrews L. 'Molasses Exhaustion and Target Purity Formulas'. Sugar Journal. 1999;20-24.
4. Rein PW, Bento L, Cortez R. The Direct Production of White Sugar in a Cane Sugar Mill. Int Sugar J. 2007;109(1301):286-299.
5. Bruijn J, Fitzergerald JR, Koenig S, Macgillivary AW. 'Proceedings of the South African Sugar Technologists' Association'. 1972;103–109.
6. Matthesius GA, Mellet P. 'Proceedings of the South African Sugar Technologists' Association. South Africa. 1976;206-207.
7. Rein PW, Smith IA. 'Proceedings of the South African Sugar Technologists' Association. 1981;55:85-91.
8. Smith IA. 'Proceedings of The South African Sugar Technologists' Association. 1995;69:163–165.
9. Matthesius GA, Mellet P. 'Proceedings of The South African Sugar Technologists' Association. 1976;50:206–207.
10. Nwalor JU, Adeniran JA. 'The application of target purity formula to quantify molasses exhaustion at Dangote sugar refinery, Apapa, Nigeria'. Adv Appl Sci Res. 2013;4(1):337-344.
11. Goza –León CO, Kazemian CH, Hormaza CJ, Rojas LYM, Santana CR. 'The International Journal of Cuban Studies'. 2008;2:96–103.
12. Jang JSR. ANFIS: Adaptive network based fuzzy inference system, IEEE Trans. On System, Man and Cybernetics. 1993;23(3):665-685.
13. Murnleitner E, Becker TM, Degado A. State detection and control of overloads in the anaerobic wastewater treatment using fuzzy logic. Water Research. 2002;36(1):201–211.
14. Pai TY, Chiang CF, Tsai YS, Chu HH, Lio WC. Using fuzzy inference system to improve neural network for predicting hospital wastewater treatment plant effluent. Computers & Chemical Engineering. 2009;33(7):1272-1278.
15. Amuthar R, Pochelvan P. 'Seasonal prediction of groundwater levels using ANFIS and radial basis neural network' Vellore. 2011;98-108.
16. MathworksAdaptive neuro fuzzy inference system toolbox; 2009b.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history.php?iid=541&id=5&aid=4737