



# Optimization of distillation column in phenol production process for increasing the isopropyl benzene concentration using response surface methodology and radial basis function (RBF) coupled with leave-one-out validation method

Hani Vaziri<sup>1</sup> · Amin Hedayati Moghaddam<sup>1</sup> · Seyed Amin Mirmohammadi<sup>1</sup>

Received: 6 January 2020 / Accepted: 11 April 2020 / Published online: 21 April 2020  
© Institute of Chemistry, Slovak Academy of Sciences 2020

## Abstract

In this study, phenol production process was simulated. Further, the performance of distillation column was optimized through maximizing the mole fraction of cumene in upstream flow. Response surface methodology was applied for design of experiment, modelling, and optimizing the cumene mole fraction in upstream flow of separation column. The analysis of variance was performed for finding the important operative parameters as well as their effects. In this experiment, the effects of three parameters on separation performance were investigated, including number of tray (*A*), column temperature (*B*), and reflux ratio (*C*). Further, radial basis function (RBF) was applied to model the separation column. To develop the neural network model, leave-one-out method was used. This robust model was used for optimizing the performance of separation column. The statistical and artificial intelligence system were capable of predicting mole fraction in upstream flow of distillation column in different conditions with  $R^2$  of 0.99 and 0.93, respectively. According to statistical and RBF models, the optimized values of cumene mole fraction are 0.45 and 0.44, respectively.

**Keywords** Design of experiment · Cumene · Radial basis function · Leave-one-out · Phenol production

## Introduction

Phenol is a simple organic substance comprised of cyclohexene and a hydroxyl group used as an intermediate substance in production of phenolic resins, base phenol A, caprolactam, adipic acid, alkyl phenols (cresols, xylenes, nonylphenols), aniline, and chlorinated phenols (such as penta-chlorophenol). Further, this substance is used in production of disinfecting materials and lotions, ointment, pain killers, soap, etc.

There are several phenol productions based on cumene route such as Sunoco/UOP, KBR, and GE/Lummus. Since about 90% of phenol production around the world is based on cumene route (Schmidt 2005), this study focuses on assessing and optimization of this process. It is important to note that there are some other routes to produce phenol,

but the selectivity of these methods is low and they are not economical (Yadav and Asthana 2003).

The computer simulation programs such as HYSYS and ASPEN PLUS have been widely used in chemical industry for design and optimization (Davis 2002; Munoz et al. 2006; Aspelund et al. 2010; Bassyouni et al. 2014; Sunny et al. 2016; Zolfaghari et al. 2017). Smejkal and Soos made a comparison between the capability of HYSYS and ASPEN PLUS in simulation of reactive distillation column (Smejkal and Šoóš 2002). They reported good agreement between the results of these two simulation programs. There are only few types of researches for simulation of phenol production process. Chudinova et al. developed a mathematical model to describe the benzene alkylation with propylene using Borland Delphi 7. The error of the model was less than 7.5%. Through optimization of process by this model, it is found that the catalyst consumption could be reduced to 10–15% and cumene concentration in product mixture could increase to 25wt% (Chudinova et al. 2015).

In phenol production process, separation of excess cumene from cumene hydroperoxide is critical. This process

✉ Amin Hedayati Moghaddam  
ami.hedayati\_moghaddam@iauctb.ac.ir

<sup>1</sup> Department of Chemical Engineering, Central Tehran Branch, Islamic Azad University, Tehran, Iran

takes place in a distillation column. The performance of distillation column as a process unit can be affected by different parameters during the separation. It is clear that controlling the process and finding the optimized parameters are challenging for process engineers. The lack of knowledge about the transfer phenomena in column and thermodynamic equations as well as relations between operation parameters leads to low efficiency. In separation process, the performance of column depends on operating parameters and their interactions with different degree of sensitivity. So, using a systematic approach to find out the influences of different parameters on the yield and consequently determining the optimum conditions is of interest. In traditional approach of assessing the effects of operative parameters, one parameter is varying while the other factors are kept constant. This method is time-consuming and cost consuming. Further, the interactions between parameters are not considered. To face this issue, a set of design of experiment (DoE) methods could be used. Recently, DoE techniques have been used in several fields of science and engineering (Wu et al. 2014; Boudjema et al. 2018; Ferdosh et al. 2012; Moghaddam et al. 2016; Zeynali et al. 2016; Heydari and Pirouzfard 2016; Bagheri et al. 2018). DOE consists of various methods. Fractional factorial design (FFD) and response surface methodology (RSM) are the two most applicable techniques to find out the interactions between operative parameters and their effects on output response.

The modelling methods based on artificial neural networks (ANNs) offer good tools to simulate the separation processes. These methods were used in several fields of science and engineering to describe the input–output relations (Tashaouie et al. 2012; Rahmanian et al. 2012; Dolatabadi et al. 2018; Hazrati et al. 2017). Motlaghi et al. applied ANN to model and consequent optimization of a crude oil distillation column (Motlaghi et al. 2008). Osuolale and Zhang developed a strategy based on ANN to model exergy efficiency in distillation columns using process operational data. After developing the ANN model, this model was used in finding the optimal condition of distillation operation to increase the energy performance of distillation process (Osuolale and Zhang 2016).

The aim of the present study is consisted of two parts. First goal is simulation of the phenol production process using ASPEN-HYSYS. Second goal is to investigate the influences of several operative parameters including number of trays, column temperature, and reflux ratio on the performance of distillation column for cumene separation as an important unit in phenol production process. In phenol production process, separation of excess cumene from cumene hydroperoxide is critical. The cumene mole fraction in upstream flow of distillation column was selected as the output response. The major novelties of this study are applying the DoE method, simulation of the process using

Aspen-Hysys, and developing a model based on ANN as well as process optimization. RSM was selected for DoE, modelling, and optimizing the conditions of production process. For simulating the process by ANN, a robust model was developed through applying radial basis function (RBF) network. The model allows predicting the composition of cumene mole fraction in upstream flow by changing the operative parameters of column without running conventional simulation. This robust model was used for optimization the separation process.

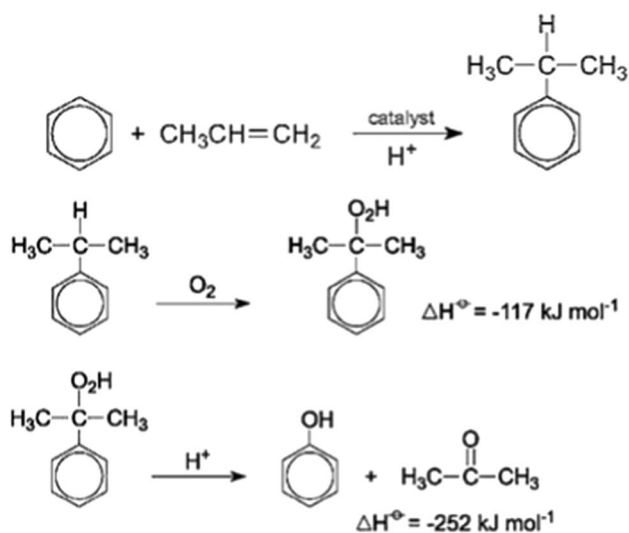
This effort could be resulted in improving the yield, environmental aspects of phenol production as a strategic substance and decreasing the production cost. In other words, the optimized operative conditions of distillation column were proposed as the best candidates for the phenol process on the basis of the Hysys simulations, optimization results obtained by DoE method as well as RBF. According to the author's knowledge, simultaneous using of Aspen-Hysys simulation, RSM, and RBF for modelling of separation process has not applied, yet.

## Methodology and procedure

### Process description

Phenol and acetone are produced in a three-step process from benzene and propylene called Hock process (Yadav and Asthana 2003). Cumene (isopropyl benzene) is an intermediate substance. In this process, two rather cheap substances (benzene and poly propylene) could be converted to two rather costly products (phenol and acetone). In first step, cumene is produced through alkylation of benzene with propylene. At the moment, all the cumene is produced commercially through using zeolite-based processes (Schmidt 2005). In the second step, cumene is oxidized by oxygen which results in cumene hydroperoxide (CHP) production. This reaction is auto-catalyzed by CHP. Finally, CHP is degraded into phenol and acetone in the presence of a mineral acid catalyst (Fig. 1). The main by-product of side reaction of cumene oxidation is dimethyl phenyl carbonyl (DMPC).

In Fig. 2, the process flow diagram of phenol production process (Sunoco process) has been presented. As is shown, feed cumene is mixed with recycled cumene from the next steps and the mixture is oxidized to CHP in two bubble columns (oxidizers). These two oxidizer columns are in series with respect to cumene mixture stream. It is important to note that more reactors could be used according to the capacity of unit (Schmidt 2005). The excess air is released into atmosphere from the top of the bubble columns after passing through an absorber to remove hydrocarbons. In distillation column, excess cumene is separated from CHP and is recycled to the feed column. Then, CHP is sent to



**Fig. 1** Reaction of phenol & acetone production through cumene route

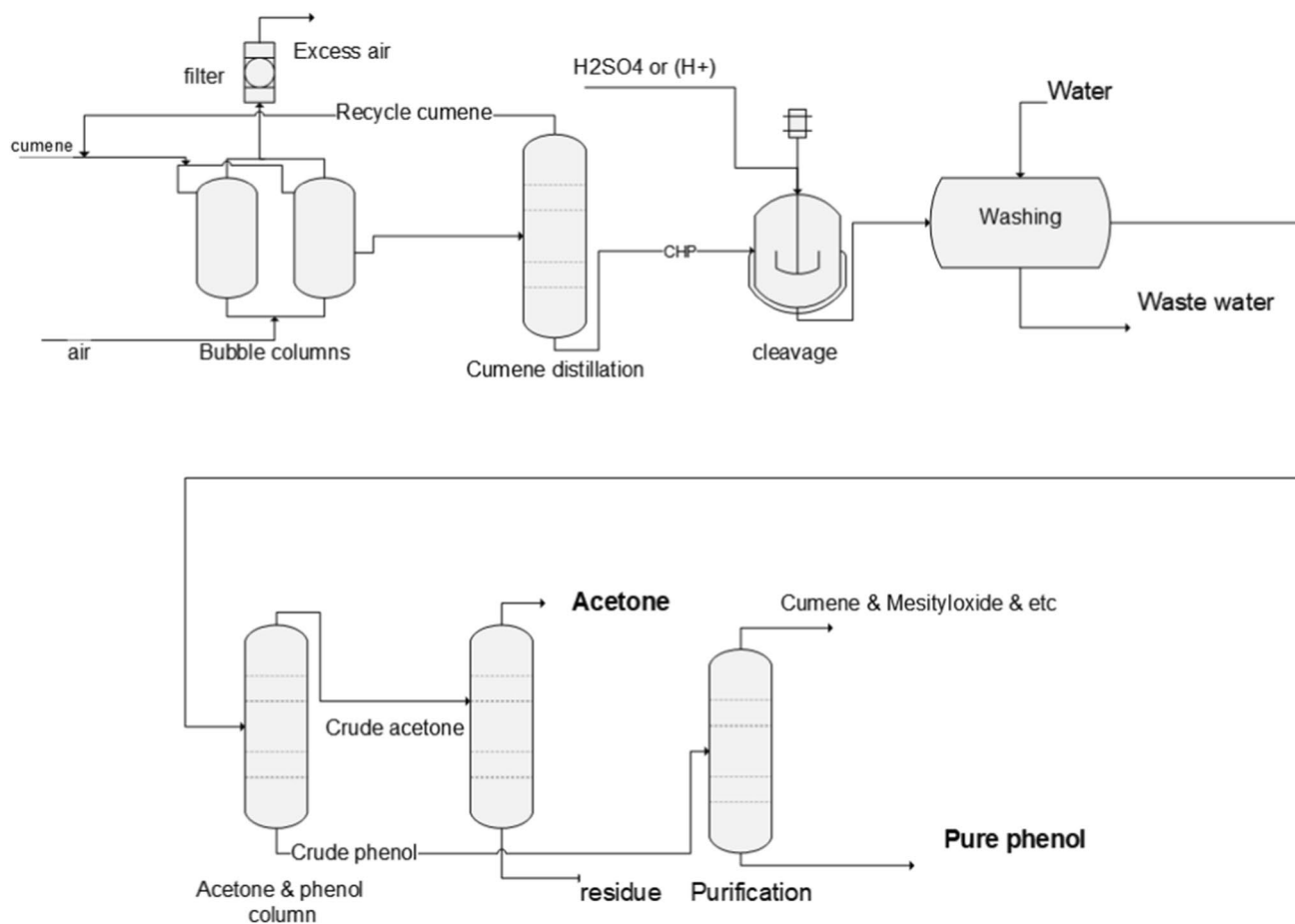
a cleavage column and will be decomposed to phenol and acetic acid in the presence of a mineral acid catalyst. In the

next step, the output stream of cleavage unit is washed by water to remove the sulfuric acid. Then, the mixture of phenol and acetone is sent to a separation column. In final step, pure acetone and phenol will be produced.

## Process simulation

### Oxidation process simulation

In Fig. 3, the oxidation process has been shown. In this section of process, cumene feed, air, and recycle stream of cumene enter into mixer 1, and consequently, CHP is produced in three series reactors. In this study, aerated reactors were replaced by conversion reactors according to simulation restrictions. The reactions of oxidation section and conversion values are presented in Table 1. The inlet feed is pure cumene with flow rate and temperature of 100 kg mol/h and 25 °C, respectively. Further, the flow rate and pressure of air stream are 100 kg mol/h and 101.3 kPa, respectively. Since the oxidation reactions are exothermic, the temperature in first reactor is fixed at 110 °C; and the temperature of next reactors was found to be at 115 °C and 119 °C, respectively.



**Fig. 2** Phenol production process

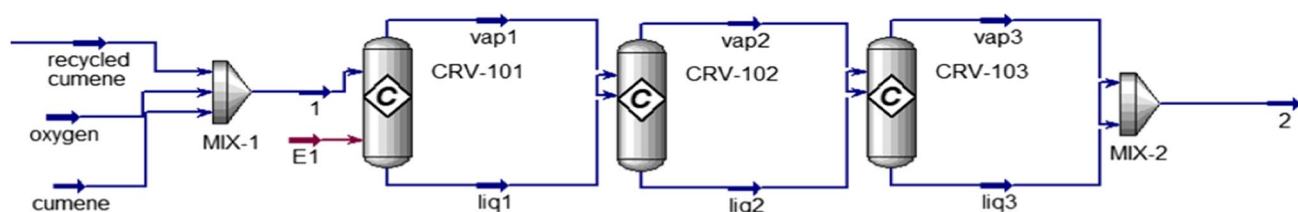


Fig. 3 Oxidation process section

**Table 1** The reactions of oxidation section and their conversion percentage

Reactions	Conversion percentage		
	First reactor	Second reactor	Third reactor
Cumene + O <sub>2</sub> → CHP	12	5	3.67
Cumene + 0.5O <sub>2</sub> → DMPH carbonyl	0	0.6	0.16
Cumene + 0.5O <sub>2</sub> → AMS + H <sub>2</sub> O	0	0.01	0
Cumene + O <sub>2</sub> → acetophenone + CH <sub>3</sub> OH	0.5	0.05	0.4
CH <sub>3</sub> OH + O <sub>2</sub> → formic acid + H <sub>2</sub> O	100	100	100

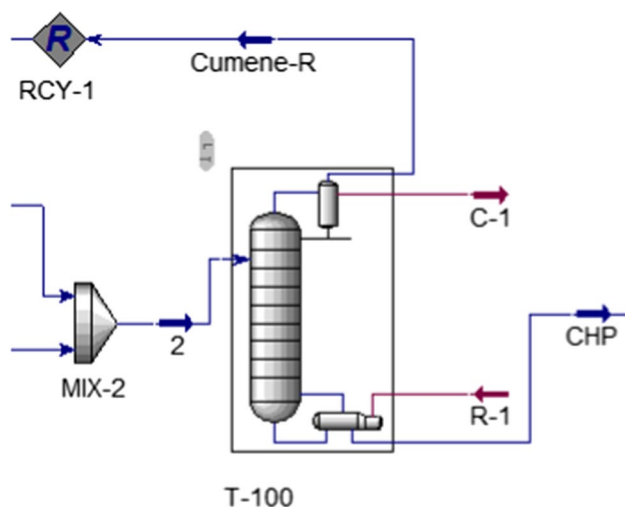


Fig. 4 Distillation column (cumene column)

The output stream of mixer 2 consisted of 18% cumene, 37% H<sub>2</sub>O, 39% formic acid, 4.5% CHP, and a few amount of DMPH carbonyl. The flow rate and temperature of output stream were 119 °C and 1908 kg mol/h, respectively.

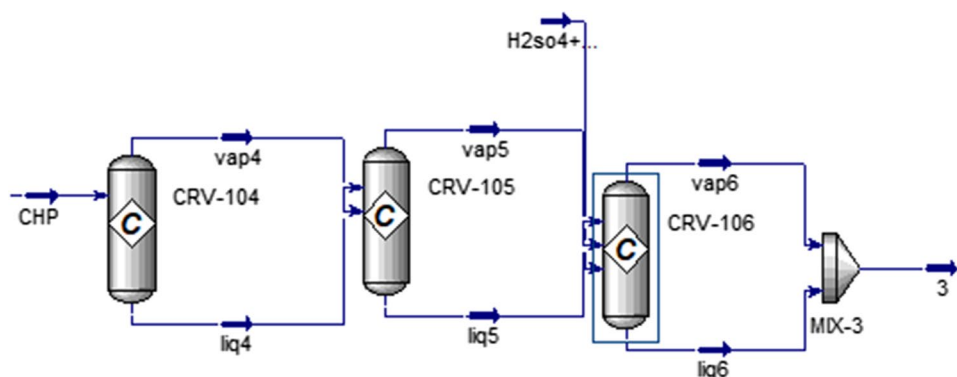
#### Concentration process simulation

In the next step, stream enters into the concentration section. As shown in Fig. 2, in a distillation column, CHP will be separated from unreacted cumene. CHP will be transferred into the next section and cumene will be recycled to the bubble columns. The thermodynamic model used to simulate the distillation column is SPRV. The schematic view of this distillation column is shown in Fig. 4.

#### Conversion process simulation

In this section, CHP and other reactants enter into three conversion reactors in series, as shown in Fig. 5. In these

Fig. 5 Conversion process section



reactors, CHP converts to acetone and phenol. The reaction kinetics is presented in Table 2. The catalyst of main reaction is low levels of  $\text{H}_2\text{SO}_4$ . To improve the preciseness of simulation, a stream comprising of 20 kg mol/h phenolate sodium, 40 kg mol/h sulphuric acid, and 40 kg mol/h sodium hydroxide enters into third reactor. The thermodynamic model that was used to simulate the distillation column is SPRV.

## Response surface methodology

There are two methods for investigating the influences of operative parameters on the process performance (conventional and DoE method). In this study, 3 operative parameters are involved in distillation column. In conventional method, one factor is changing and all the other factors are kept constant. And this procedure is repeated 3 times. So, it is clear that this method including a lot of experiments and doing that is a time-consuming and cost-consuming work. But in design of experiment methodology, all the factors are changed simultaneously. So, the number of experiments will decrease notably. Further, in DoE possible interactions between operative parameters could be considered. RSM is a hybrid statistical-mathematical methodology that could be used in design of experiment as well as analyzing the experimental data using analysis of variance. Further, statistical model may be generated. This developed model is a useful tool to optimize the process. In this approach, design points consisted of factorial, axial, and center points (Montgomery 2017). The number of experiments could be calculated by the following equation:

$$N = 2^n + 2 \times n + N_c \quad (1)$$

where  $n$  is the number of operative parameters. The terms  $2^n$ ,  $2 \times n$ , and  $N_c$  are the number of factorial, axial, and center points, respectively. Since the number of operative parameters is 3, the numbers of factorial and axial points are 8 and 6, respectively. Further, the number of center points is 5.

**Table 2** The reactions of conversion section and their conversion percentage

Reactions	Conversion percentage		
	Fourth reactor	Fifth reactor	Sixth reactor
CHP → acetone + phenol	60	100	–
DMPH carbinol → AMS + $\text{H}_2\text{O}$	20	99	–
Phenol + AMS → cumylphenol	2.1	1.99	–
Phenol + 2 acetone + $3\text{H}_2\text{O}$ → 6 acetaldehyde	0.07	0.06	–
2 acetone → mesityloxide + $\text{H}_2\text{O}$	2	1.09	–
2 sodium phenoxide + sulfuric acid → sodium sulfate + 2 phenol	–	–	86.4
Sulfuric acid + 2 sodium hydroxide → sodium sulfate + $2\text{H}_2\text{O}$	–	–	13.6

The experimental strategies in statistical section are (I) performing the design of experimental layout, (II) model development, and (III) optimization.

For modelling and optimizing the performance of distillation column, a central composite design (CCD) with 5 replicates at the center point was performed. The upper and lower limits of the operative parameters are shown in Table 3. The final goal of the design of experiments is seeking for optimal conditions to operate the distillation column with highest cumene mole fraction and lowest column temperature, lowest number of trays, and lowest reflux ratio.

There are several software for performing the design of experiment. In this study, Design Expert software is used to design experimental layout, data analysis, and process optimization. In this method, data processing is performed in coded forms. Operating parameters were coded by the following equation:

$$X_i = \frac{x_i - \frac{x_{i,\text{high}} - x_{i,\text{low}}}{2}}{\frac{x_{i,\text{high}} - x_{i,\text{low}}}{2}} \quad (2)$$

where  $X_i$  and  $x_i$  are the coded form and real value of operative parameter. And  $x_{i,\text{high}}$  and  $x_{i,\text{low}}$  are the maximum and minimum values of the operative parameters, respectively.

## Radial basis function

ANN is a simple model of human neural network that could perform data processing. There are several different types

**Table 3** Operative ranges of process parameters

Parameter	Code	Range
Number of tray	A	3–35
Column temperature ( $^{\circ}\text{C}$ )	B	175–180
Reflux ratio	C	0.7–5

of ANN, but all of them consist of two main component: (I) neurons that perform data processing on input data and (II) connections that determine the transferring of information between neurons. The type of connection between neurons, neuron layout (network topology), and how to train the network are critical issues in development of ANN-based model.

RBF neural network is a nonlinear layered feed forward network that consists of 3 layers (input, hidden, and output layers) and uses radial basis function as activation function. In this research, MATLAB 2010a was used to devise and train the networks. The architecture of RBF is shown in Fig. 6.

The role of first layer (input layer) is just receiving the input data and transferring it to the next layer (hidden layer). Main portion of data processing is performing in hidden radial basis layer. The distance box produces several elements that are the differences between input and weight matrix. Then, the bias is added and the result elements are introduced to Gaussian function (f1) which in MATLAB environment is termed radbas. And third layer is an output linear layer. The function of f2 in output layer is a linear transfer function (in MATLAB termed purelin).

Model development and data processing were performed for normalized operative parameters. data normalization was done as follow:

$$X_i = 0.05 + 0.9 \times \frac{(x_i - x_{i,low})}{x_{i,high} - x_{i,low}} \quad (3)$$

where  $X_i$  and  $x_i$  are normalized and real values, respectively. And  $x_{i,high}$  and  $x_{i,low}$  are high and low levels of parameters, respectively.

In this study, for data mining, RBF was selected to generate a robust model. The properties of RBF are as follows:

- Mean squared error goal.
- Spread of radial basis function.
- Maximum number of neurons.
- Number of neurons to add between displays.

Further, trail-and-error method was used to develop the optimized neural network. Mean squared error (MSE) could be calculated as follows:

$$MSE = \frac{\sum_{i=1}^n (y_{i,pred.} - y_{i,real})^2}{n} \quad (4)$$

where  $y_{pred.}$  and  $y_{real}$  are predicted and real parameters, respectively. And  $n$  is the number of data.

### Cross-validation—leave one out method

In this method, since there are  $N$  data, the dataset is divided into two partitions with  $(N - 1)$  data and one data. The goal of performing this partitioning is to set the section with one data as validation dataset and the other section with  $(N - 1)$  data as training dataset. This procedure will be repeated  $N$  times. So, this validation method termed one-leave-out cross-validation.

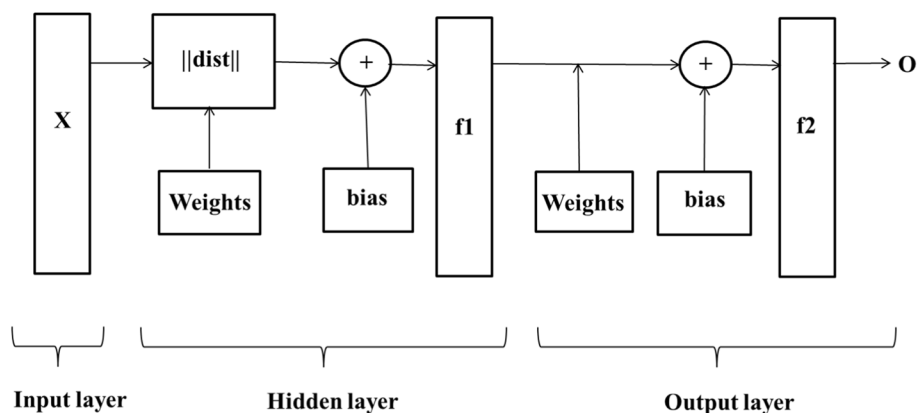
In other words, for network training by this method, in each step,  $(N - 1)$  data are used to train a network with optimized properties and 1 data of dataset will be out of consideration. Then, this trained network is validated by this one data that would be out of consideration in previous step. Since this procedure is repeated  $N$  times, finally, there would be  $N$  predicted data (validation data) that could be compared with goal (experimental data) to calculate the model accuracy.

The accuracy and preciseness of network prediction were assessed by determination coefficient ( $R^2$ ). This index could be calculated as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_{i,real} - y_{i,pred.})^2}{\sum_{i=1}^n (y_{i,real} - \bar{y})^2} \quad (5)$$

$$\bar{y} = \frac{\sum_{i=1}^n y_{i,real}}{n} \quad (6)$$

Fig. 6 RBF structure





where  $y_{\text{pred}}$ ,  $y_{\text{real}}$ , and  $\bar{y}$  are predicted, real, and average values, respectively. And  $n$  is the number of data. In this study, to develop a model, several RBF networks with different properties were constructed and trained. Each generated net was validated through one-leave-out method as mentioned above. Then, the values of  $R^2$  for training and validation phase were calculated for each network. Next, the average values of validation- $R^2$  were compared to select the optimum network. Finally, the optimized properties of network were set according to the value of the validation- $R^2$  (maximum value). It is important to note that in this manner, for data generating for comparing with real data a distinct neural network is not used, and a hybrid of several networks ( $N$  nets) as a prediction tool is developed. But for optimization and plotting the graphs, a unique optimized network is used.

## Result and discussion

In this section, the distillation column that has been used in concentration unit is assessed and simulated by applying statistical analysis and data mining strategy. So, the mole fraction of cumene in upstream flow could be predicted according the operative parameters using the developed models.

### Statistical analysis

#### Cumene mole fraction in upstream flow

DoE is performed using CCD. Accordingly, 19 experiments have been ran and the results are assessed by applying analysis of variance (ANOVA). ANOVA table of cumene mole fraction in upstream flow, which has been obtained from RSM methodology, is shown in Table 4.

Values of "Prob >  $F$ " less than 0.0500 indicate that the model terms are effective. Values greater than 0.1000 indicate the model terms are not important. From Table 4, it is

clear that reflux ratio ( $C$ ) could not influence the upstream cumene mole fraction, but  $C^2$  is a considerable factor. So, to support the hierarchy of model, reflux ratio must be considered as a model term. According to the  $F$ -value results, the ranking of operative parameter on the basis of importance is as follows:  $B > B^2 > A > AB > C^2$ . Adeq Precision of 50.375 indicates that signal to noise is in acceptable range. The model  $F$ -value of 322.09 shows that this model is significant. Accordingly, there is a little chance that this model  $F$ -value is due to noise. The precision of a model can be checked by determination of an  $R^2$  coefficient and adjusted  $R^2$ . In this case,  $R^2$  is 0.9938. In other words, 99.38% of response variability is achieved by the obtained regression model. The adjusted  $R^2$  is 0.9907 reasonably close to 1, which is in acceptable agreement with  $R^2$ . The final equation for cumene mole fraction in terms of coded factors obtained from regression of values is as follows:

$$\begin{aligned} \text{Up - stream cumene mole fraction} &= 0.13 + 0.032 \\ &\times T - 0.16 \times B + 5.36 \times 10^{-3} \times C - 0.018 \times AB \\ &+ 0.12 \times B^2 - 0.023 \times C^2 \end{aligned} \quad (7)$$

For review the predicted data and plotting the related graphs, introducing the regression model in actual values is preferred. The extracted model in actual values of operative parameters for cumene mole fraction is as follow:

$$\begin{aligned} \text{Up - stream Cumene mole fraction} &= +601.97033 \\ &+ 0.082763 \times \text{Tray} - 6.72548 \times \text{Temperature} + 0.030599 \\ &\times \text{Reflux ratio} - 4.55000 \times 10^{-4} \times \text{Tray} \times \text{Temperature} \\ &+ 0.018785 \times \text{Temperature}^2 - 4.93087 \times 10^{-3} \times \text{Reflux ratio}^2 \end{aligned} \quad (8)$$

where  $3 < \text{number of tray} < 35$ ,  $175 < \text{column temperature} < 180$  ( $^{\circ}\text{C}$ ),  $0.7 < \text{reflux ratio} < 5$

This model can be used to predict the mole fraction only in the limits of operative parameters. To check the adequacy of final model, the normal probability plot vs. standardized

**Table 4** ANOVA table for cumene concentration in upstream flow in RSM methodology

Source	Sum of squares	<i>Df</i>	Mean squares	<i>F</i> value	Prob > <i>F</i>	Comment
Model	0.33	6	0.056	322.09	<0.0001	Significant
<i>A</i> (tray)	0.010	1	0.010	59.35	<0.0001	Significant
<i>B</i> (temperature)	0.27	1	0.27	1548.53	<0.0001	Significant
<i>C</i> (reflux ratio)	2.873E-004	1	2.873E-004	1.66	0.2213	
<i>AB</i>	2.650E-003	1	2.650E-003	15.35	0.0020	Significant
<i>B</i> <sup>2</sup>	0.043	1	0.043	251.97	<0.0001	Significant
<i>C</i> <sup>2</sup>	1.639E-003	1	1.639E-003	9.50	0.0095	Significant
Residual	2.072E-003	12	1.726E-004			
Lack of fit	2.072E-003	8	2.589E-004			
Pure error	0.000	4	0.000			
Cor total	0.34	18				

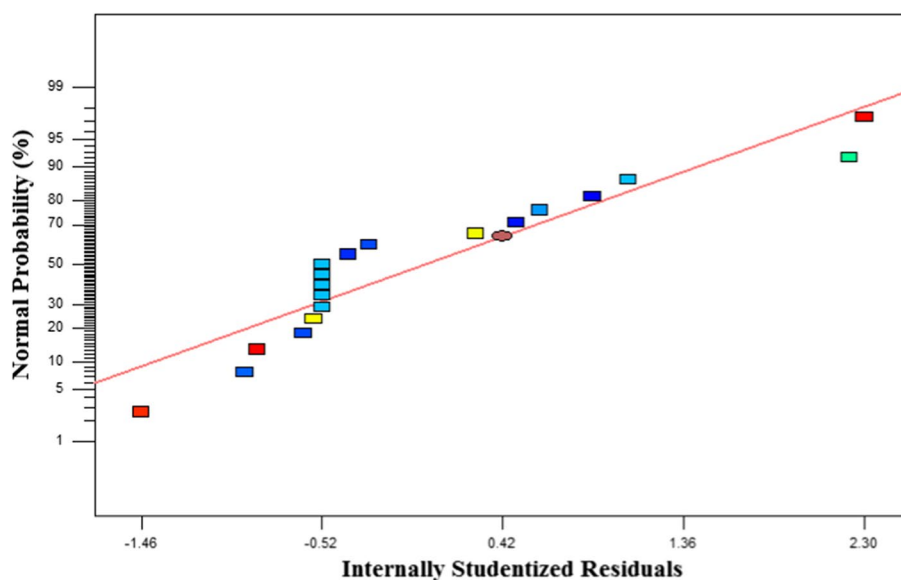
residuals is presented in Fig. 7. It is obvious that the points follow a straight line. So, there is no need for transformation.

The residual versus run number is shown in Fig. 8. This is a useful tool to check the lurking variables that influenced the response. In this plot, random scattering is of interest. The validation tools such as coefficient determination and related plots figured out that the response values determined by the obtained statistical model were in good agreement with experimental data for mole fraction of cumene in upstream flow of distillation column. So, this model is found to be useful for data prediction, plotting the related graphs, graphically assessment of the effect of operative parameters as well as finding the optimized conditions.

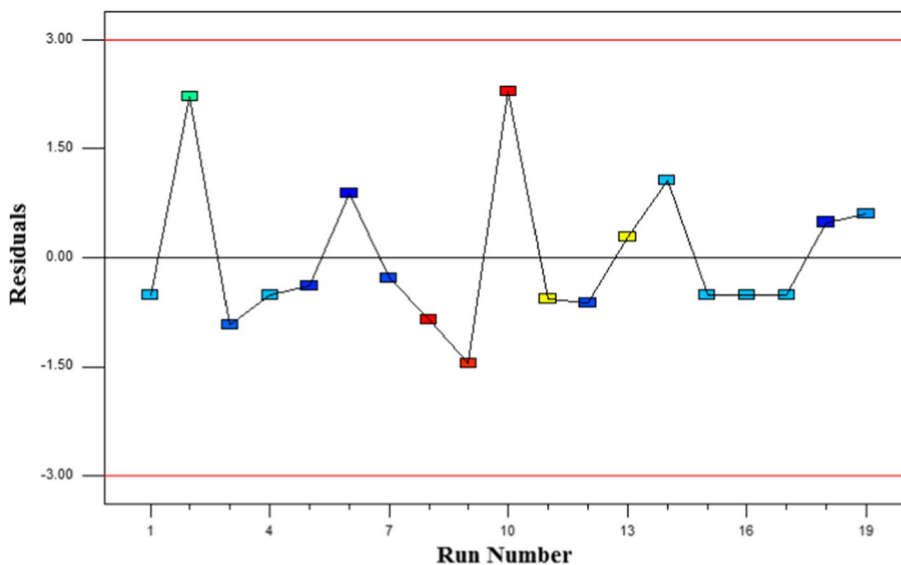
In Fig. 9a, the interaction effect of the number of trays and column temperature on cumene mole fraction is

displayed while reflux ratio is fixed at 2.85. As can be seen, when the temperature reduces and number of trays grows simultaneously, cumene mole fraction would be peaked. It can be concluded that increasing the column temperature and decreasing the number of trays have obviously negative effect on response. It is important to note that a decrease in the number of trays causes lower cumene mole fraction moderately in comparison with increasing the column temperature. In Fig. 9b, cumene mole fraction has been plotted vs. number of tray and reflux ratio. As shown, the influence of number of tray on response is greater than the influences of reflux ratio. This observation is in agreement with the result of ANOVA table that marked reflux ratio as an unimportant operative factor. Figure 9c reveals that the effect of

**Fig. 7** Normal probability plot of residuals for cumene concentration in upstream flow

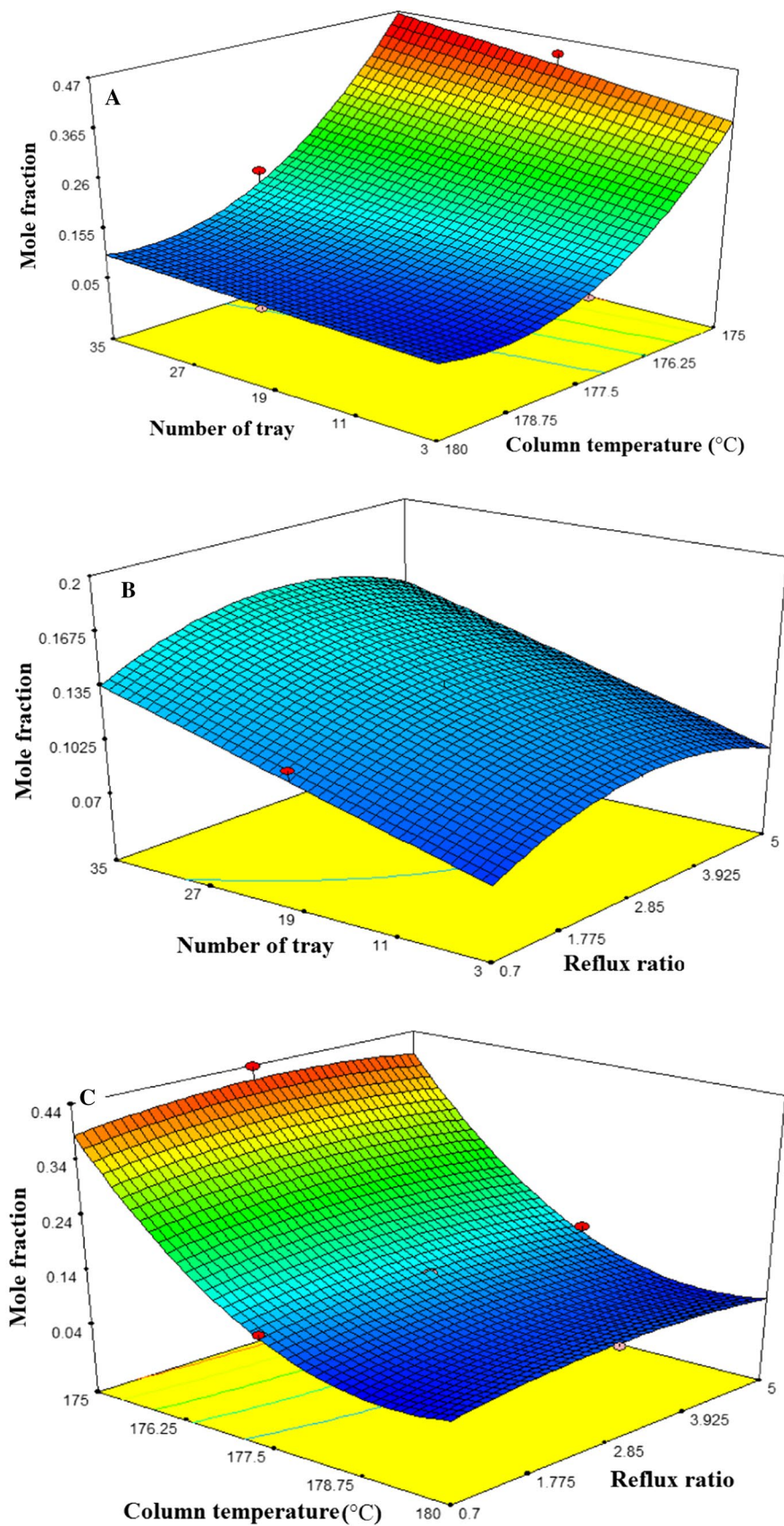


**Fig. 8** Residuals versus run number for cumene concentration in upstream flow





**Fig. 9** 3D and contour plot of predicted cumene mole fraction as **a** function of a number of trays and column temperature, **b** number of tray and reflux ratio, and **c** column temperature and reflux ratio



**Table 5** Optimized condition-statistical method

Goal	Number of trays <sup>a</sup>	Column temperature (C) <sup>a</sup>	Reflux ratio <sup>a</sup>	Cumene mole fraction <sup>a</sup>	Desirability
I	31	175	1.60	0.44	1.000
II	3	175	3.12	0.36	0.912
III	3	175	3.10	0.36	0.878
IV	30	175	1.87	0.44	1.000
V	3	175	0.70	0.34	0.911

<sup>a</sup>Values have been rounded

reflux ratio on response is negligible in comparison with column temperature.

### Optimization

In this study, minimum value of operative parameters and maximum value of response is of interest. To optimize the process, after developing a statistical model, desirability function approach is used for finding optimized conditions.

In this approach, five different goals have been set as follows:

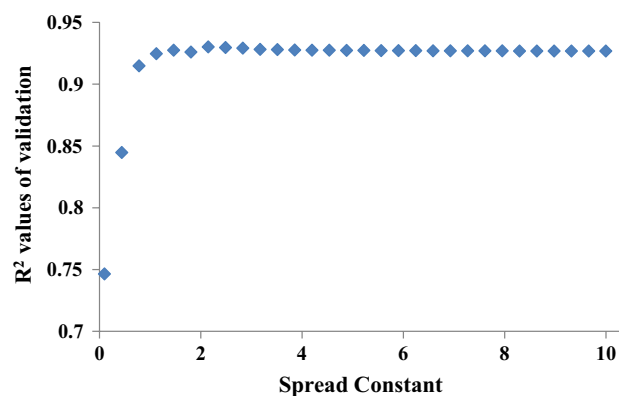
- I: maximizing the cumene mole fraction.
- II: minimizing the number of trays/minimizing the column temperature/maximizing the cumene mole fraction.
- III: minimizing the number of trays/maximizing the cumene mole fraction.
- IV: Minimizing the column temperature /maximizing the cumene mole fraction.
- V: minimizing the column temperature/minimizing the number of trays/minimizing the reflux ratio/maximizing the cumene mole fraction.

In above strategies, the degree of importance of response is set 3 and the degrees of importance of operative parameters were set at 2. The reason of minimizing the column temperature is to decrease the energy consumption or minimizing the number of trays leads to decrease in construction cost. In Table 5, the optimized conditions for each goal have been shown.

### Data mining

#### Model development

In this work, for networks deal with cumene mole fraction in upstream flow, maximum number of neurons is set at 20 and number of neurons to add between displays is set at 1. Accordingly, several networks with different structures were investigated. Each network was trained through leave-one-out method as explained previously. The preciseness



**Fig. 10** The average values of validation- $R^2$  versus mean squared error goal

of artificial networks was assessed by determination coefficient ( $R^2$ ) of validation. In Fig. 10, the average values of  $R^2$  of validation for different structures of RBF with different spread constant and mean squared error goal of 0.01 have been shown where maximum number of neurons and number of neurons to add between displays were 20 and 1, respectively. It is found that the maximum  $R^2$  is 0.93 that belongs to a network with spread constant of 2.148.

Assessing the influence of the variation of mean squared error on the values of validation- $R^2$  revealed that this parameter has no considerable effect on the prediction performance of developed models.

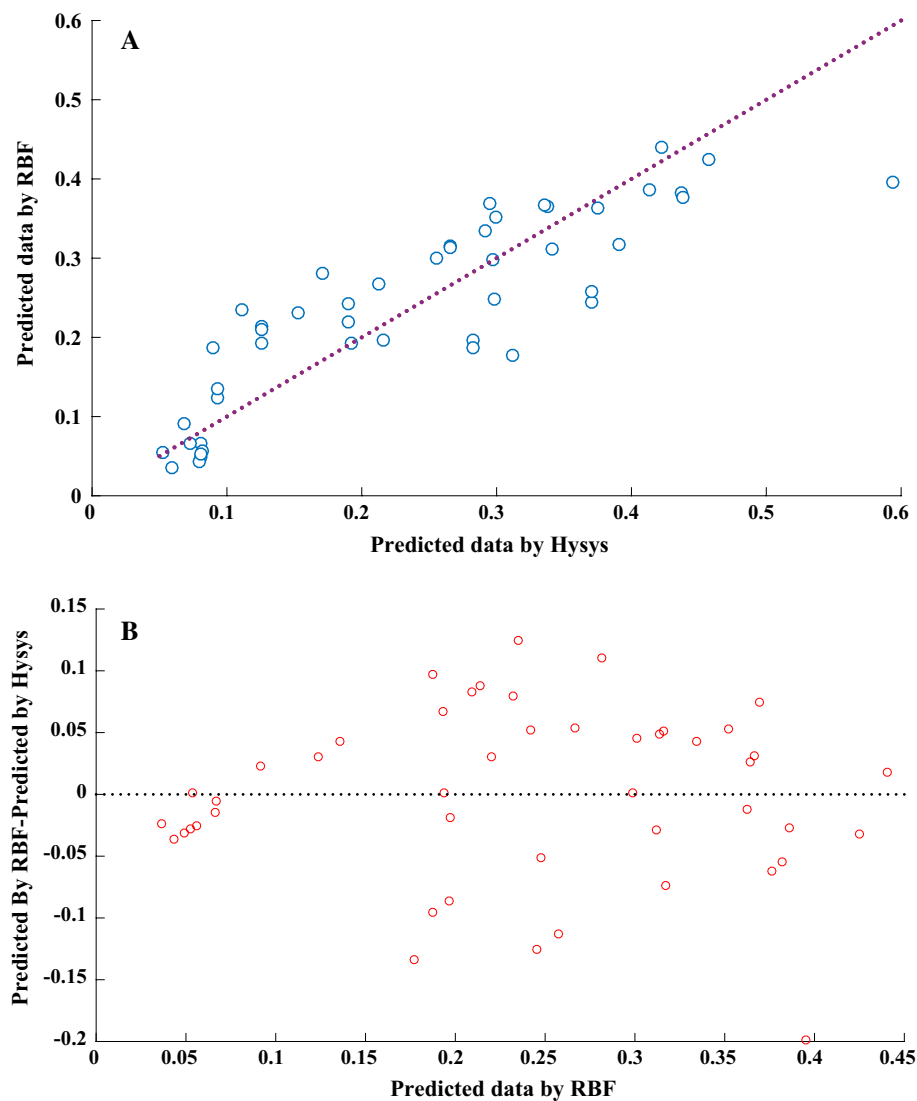
In Fig. 11a, the predicted cumene mole fraction by optimized RBF has been scattered versus data generated by Hysys simulation. As is shown, acceptable distribution around the line  $y=x$  proves the agreement of between data generated by developed neural network model and Hysys simulation. This developed model can be used to predict the cumene mole fraction only in the limits of operative parameters. In Fig. 11b, the difference between data predicted by RBF and Hysys simulation termed residuals has been scattered versus predicted data by RBF. Random distribution of residuals is of interest.

#### Data assessment

According to the prediction result of developed neural network model, the maximum value of cumene mole fraction is 0.4407. This value could be achieved when the number of tray, reflux ratio, and column temperature are 35, 0.7, and 175 °C, respectively. Similarly, the minimum value of cumene mole fraction is 0.0361. The number of tray, column temperature, and reflux ratio corresponded to this minimum value are 3, 4.73, and 180, respectively.

In Fig. 12a, mole fraction of cumene in upstream flow of distillation column that was predicted by optimized RBF model has been plotted versus number of tray and column

**Fig. 11** **a** Predicted data by RBF versus predicted data by Hysys and **b** residuals versus predicted data by Hysys



temperature while the reflux ratio was fixed at 2.85. By comparing with Fig. 9, it is found that the result of RBF is in agreement with the result of Hysys simulation, but RBF proposed linear relations. Figure 12b, c proves that the effect of reflux ratio on cumene mole fraction is less than column temperature and number of trays. Further, Fig. 12a, b shows that the effect of column temperature on response is severe than the effects of number of tray and reflux ratio. These observations are in agreement with the results that were obtained from ANOVA table (Table 4).

### Optimization

After developing the robust RBF model, this model could be used to find the optimized conditions. Similar to optimization through statistical analysis, five goals have been defined. Each criterion is reevaluated by assigning weights according to their importance. For example in goal II, 3 criteria are

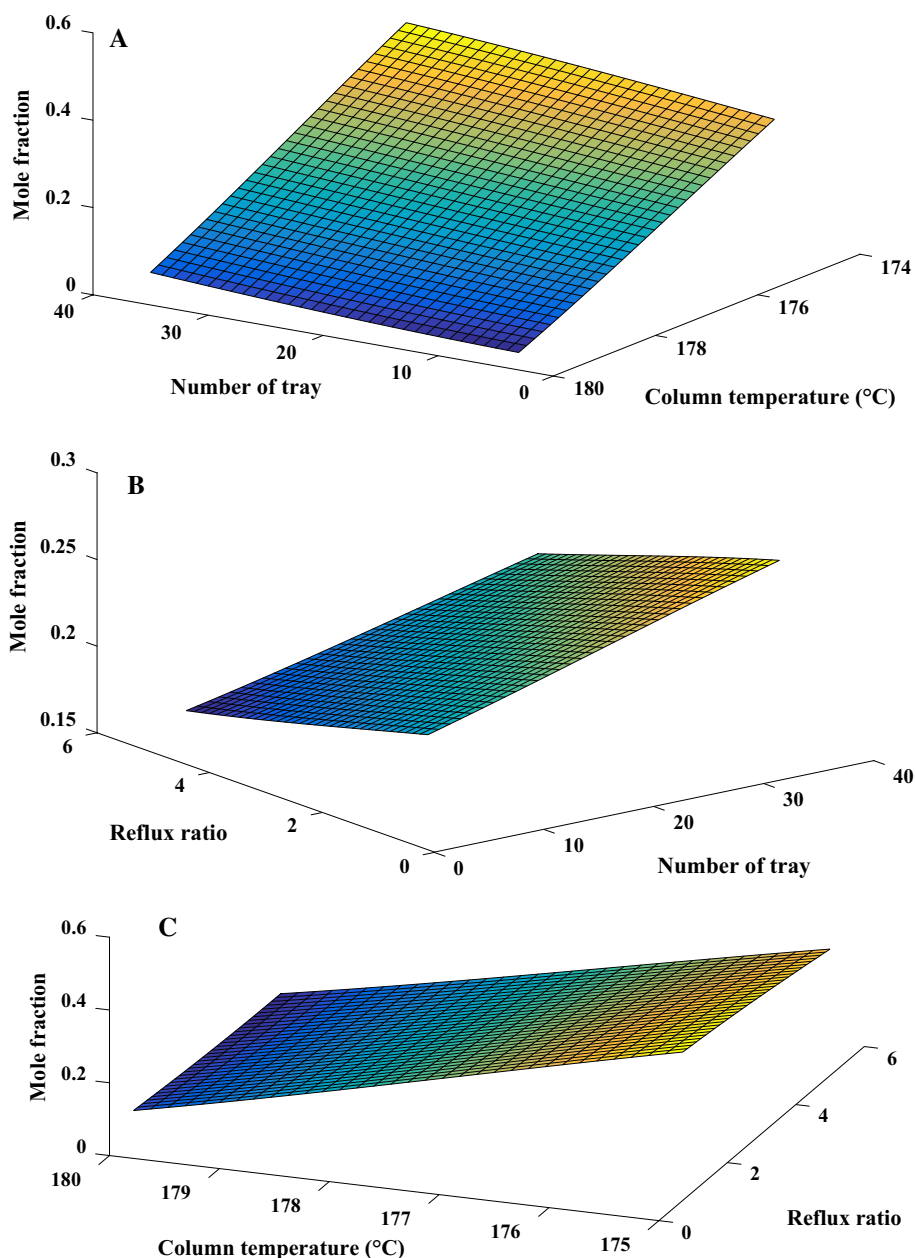
defined. Each of these criteria has different degree of importance. Accordingly, the importance of each criterion could be determined by assigning the weights. In other words, the weight could determine the degree of importance of each criterion. The optimization process is performed in a way to decrease the optimization function. Optimization function is defined as follows:

$$\text{Optimization function} = \sum_{i=1}^n |y_i - y_{ig}| \times w_i \quad (9)$$

where  $w_i$  is the weight of each criterion. And  $y_i$  and  $y_{ig}$  are the criterion and its maximum or minimum value.

In optimization process, the value of optimization function was checked and the optimized condition according to the each goal and assigned weights could be found. The reason of minimizing the column temperature is to decrease the energy consumption or minimizing the number of trays leads

**Fig. 12** Mole fraction of cumene in upstream flow as a function of **a** number of tray and column temperature, **b** number of tray and reflux ratio, and **c** column temperature and reflux ratio



to decrease in construction cost. In Table 6, the optimized conditions for each goal have been shown.

## Conclusion

In this research, an investigation was made on the simulation of phenol production process. The data of cumene mole fraction in upstream flow of distillation column (an important unit in phenol production process) obtained from ASPEN-HYSYS simulation were applied to develop statistical and artificial neural network models. Statistical analysis revealed that number of trays and column temperature have

significant effect on cumene mole fraction. The predicted results of the optimized neural network model proved that the effect of reflux ratio on cumene mole fraction is less than column temperature and number of trays. Further, the effect of column temperature on cumene mole fraction is severe than the effects of number of trays. Optimization of process based on neural network model revealed that the maximum cumene mole fraction in upstream flow is 0.44 that will be obtained if the number of trays, reflux ratio, and column temperature are 35, 0.7, and 175 °C, respectively. So, use of these values in distillation column design leads to an improved process performance in distillation column.



**Table 6** Optimized condition-neural network method

Strategy	Weights (response-A–B–C)	A (number of trays)	B (reflux ratio)	C [column temperature (°C)]	Cumene mole fraction
I	–	35	0.7	175	0.4407
II	0.33–0.33–0.33	3	0.7	175	0.36422
	0.7–0.15–0.15	11	0.7	175	0.38662
	0.5–0.1–0.4	17	0.7	175	0.40241
III	0.5–0.5	3	0.7	175	0.36422
	0.9–0.1	35	0.7	175	0.4407
IV	0.5–0.5	35	0.7	175	0.4407
V	0.25–0.25–0.25–0.25	3	0.7	175	0.36422
	0.7–0.1–0.1–0.1	32	0.7	175	0.43551

## Compliance with ethical standards

**Conflict of interest** On behalf of all authors, the corresponding author states that there is no conflict of interest.

## References

- Aspelund A, Gundersen T, Myklebust J, Nowak M, Tomasgard A (2010) An optimization-simulation model for a simple LNG process. *Comput Chem Eng* 34(10):1606–1617. <https://doi.org/10.1016/j.compchemeng.2009.10.018>
- Bagheri S, Aghaei H, Ghaedi M, Asfaram A, Monajemi M, Bazrafshan AA (2018) Synthesis of nanocomposites of iron oxide/gold (Fe<sub>3</sub>O<sub>4</sub>/Au) loaded on activated carbon and their application in water treatment by using sonochemistry: optimization study. *Ultrason Sonochem* 41:279–287. <https://doi.org/10.1016/j.ultsonch.2017.09.031>
- Bassyouni M, ul Hasan SW, Abdel-Aziz M, Abdel-hamid S-S, Naveed S, Hussain A, Ani FN (2014) Date palm waste gasification in downdraft gasifier and simulation using ASPEN HYSYS. *Energy Convers Manag* 88:693–699. <https://doi.org/10.1016/j.enconman.2014.08.061>
- Boudjema S, Zerrouki M, Choukhou-Braham A (2018) Experimental design for modeling and multi-response optimization of catalytic cyclohexene epoxidation over polyoxometalates. *J Chin Chem Soc* 65(4):435–444. <https://doi.org/10.1002/jccs.201700291>
- Chudinova A, Salischeva A, Ivashkina E, Moizes O, Gavrikov A (2015) Application of cumene technology mathematical model. *Procedia Chem* 15:326–334. <https://doi.org/10.1016/j.proche.2015.10.052>
- Davis RA (2002) Simple gas permeation and pervaporation membrane unit operation models for process simulators. *Chem Eng Technol Ind Chem Plant Equip Process Eng. Biotechnol* 25(7):717–722. [https://doi.org/10.1002/1521-4125\(20020709\)25:7%3c717::AID-CEAT717%3e3.0.CO;2-N](https://doi.org/10.1002/1521-4125(20020709)25:7%3c717::AID-CEAT717%3e3.0.CO;2-N)
- Dolatabadi M, Mehrabpour M, Esfandyari M, Alidadi H, Davoudi M (2018) Modeling of simultaneous adsorption of dye and metal ion by sawdust from aqueous solution using of ANN and ANFIS. *Chemometr Intell Lab Syst* 181:72–78. <https://doi.org/10.1016/j.chemolab.2018.07.012>
- Ferdosh S, Sarker MZI, Abd Rahman NNN, Selamat J, Karim MR, Razak TA, Abd Kadir MO (2012) Fish oil recovery from viscera of Indian mackerel (*Rastrelliger kanagurta*) by supercritical fluid: an optimization approach. *J Chin Chem Soc* 59(11):1421–1429. <https://doi.org/10.1002/jccs.201100235>
- Hazrati H, Moghaddam AH, Rostamizadeh M (2017) The influence of hydraulic retention time on cake layer specifications in the membrane bioreactor: experimental and artificial neural network modeling. *J Environ Mental Chem Eng* 5(3):3005–3013. <https://doi.org/10.1016/j.jece.2017.05.050>
- Heydari S, Pirouzfard V (2016) The influence of synthesis parameters on the gas selectivity and permeability of carbon membranes: empirical modeling and process optimization using surface methodology. *RSC Adv* 6(17):14149–14163. <https://doi.org/10.1039/C5RA27772H>
- Moghaddam AH, Shayegan J, Sargolzaei J (2016) Investigating and modeling the cleaning-in-place process for retrieving the membrane permeate flux: Case study of hydrophilic polyethersulfone (PES). *J Taiwan Inst Chem Eng* 62:150–157. <https://doi.org/10.1016/j.jtice.2016.01.024>
- Montgomery DC (2017) Design and analysis of experiments. Wiley, Hoboken
- Motlaghi S, Jalali F, Ahmadi MN (2008) An expert system design for a crude oil distillation column with the neural networks model and the process optimization using genetic algorithm framework. *Expert Syst Appl* 35(4):1540–1545. <https://doi.org/10.1016/j.eswa.2007.08.105>
- Munoz R, Monton J, Burguet M, De la Torre J (2006) Separation of isobutyl alcohol and isobutyl acetate by extractive distillation and pressure-swing distillation: simulation and optimization. *Sep Purif Technol* 50(2):175–183. <https://doi.org/10.1016/j.seppur.2005.11.022>
- Osuolale FN, Zhang J (2016) Energy efficiency optimisation for distillation column using artificial neural network models. *Energy* 106:562–578. <https://doi.org/10.1016/j.energy.2016.03.051>
- Rahmanian B, Pakizeh M, Mansoori SAA, Esfandyari M, Jafari D, Maddah H, Maskooki A (2012) Prediction of MEUF process performance using artificial neural networks and ANFIS approaches. *J Taiwan Inst Chem Eng* 43(4):558–565. <https://doi.org/10.1016/j.jtice.2012.01.002>
- Schmidt RJ (2005) Industrial catalytic processes—phenol production. *Appl Catal A* 280(1):89–103. <https://doi.org/10.1016/j.apcata.2004.08.030>
- Smejkal Q, Šošč M (2002) Comparison of computer simulation of reactive distillation using ASPEN PLUS and HYSYS software. *Chem Eng Process* 41(5):413–418. [https://doi.org/10.1016/S0255-2701\(01\)00160-X](https://doi.org/10.1016/S0255-2701(01)00160-X)
- Sunny A, Solomon P, Aparna K (2016) Syngas production from regasified liquefied natural gas and its simulation using Aspen HYSYS. *J Nat Gas Sci Eng* 30:176–181. <https://doi.org/10.1016/j.jngse.2016.02.013>

- Tashaouie HR, Gholikandi GB, Hazrati H (2012) Artificial neural network modeling for predict performance of pressure filters in a water treatment plant. *Desalination Water Treat* 39(1–3):192–198. <https://doi.org/10.1080/19443994.2012.669175>
- Wu YT, Cham TM, Tsai TR (2014) Development of HPLC with photodiode array method for the determination of Ramipril in tablets using factorial design. *J Chin Chem Soc* 61(12):1388–1394. <https://doi.org/10.1002/jccs.201400130>
- Yadav GD, Asthana NS (2003) Selective decomposition of cumene hydroperoxide into phenol and acetone by a novel cesium substituted heteropolyacid on clay. *Appl Catal A* 244(2):341–357. [https://doi.org/10.1016/S0926-860X\(02\)00605-1](https://doi.org/10.1016/S0926-860X(02)00605-1)
- Zeynali V, Sargolzaei J, Hedayati Moghaddam A (2016) Optimization of several hydrodynamic and non-hydrodynamic operating parameters in treatment of synthetic wastewater containing wheat starch in a sequencing batch reactor (SBR) using response surface methodology. *Desalination Water Treat* 57(51):24240–24256. <https://doi.org/10.1080/19443994.2016.1141372>
- Zolfaghari M, Pirouzfard V, Sakhaeinia H (2017) Technical characterization and economic evaluation of recovery of flare gas in various gas-processing plants. *Energy* 124:481–491. <https://doi.org/10.1016/j.energy.2017.02.084>

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