

Construction of an optimal model for the preparation of C4 olefins based on multiple regression

Lihong Yang ^a, Leilei Yang ^b

Sichuan normal university, Sichuan, 610000, China

^a ylhfew@163.com, ^b yllyouth@163.com

Abstract. In this paper, based on the data obtained from a series of experiments conducted by a chemical laboratory for different catalysts at different temperatures, the optimization model and regression relationship model are established by using Gibbs free energy law, chemical equation energy analysis, chemical reaction equilibrium, the preparation of C4 alkenes by ethanol coupling was studied. By analyzing the temperature variation law of ethanol conversion and C4 olefin selectivity under different catalyst combinations, a regression relationship model was established, and the functional relationship between the variables was obtained. Then, the effects of different catalyst combinations and temperatures on ethanol conversion and C4 olefin selectivity were discussed. At the same time, in order to establish two optimization models, the first model is the C4 olefin yield model, which takes the minimum C4 olefin yield as the goal, and takes the range of factors such as temperature, and total mass as constraints. The second optimization model is based on the previous one with a constraint that the temperature is below 350°C, and then solved in MATLAB.

Keywords: C4 Olefin Yield; Multiple Regression; Exhaustive method; optimization model.

1. Question restatement

Ethanol is a raw material for the production of C4 olefins and is widely used in pharmaceutical production and chemical products. The combination of temperature and catalyst i.e. Co loading, Co/SiO₂ and HAP charging ratio, ethanol concentration will affect the C4 olefin selectivity and C4 olefin yield during the preparation. Therefore, it is of great significance to design different combinations of catalysts to explore the process conditions for the catalytic coupling of ethanol to prepare C4 olefins.

For a series of experiments performed by a chemical laboratory on different catalysts at different temperatures, a mathematical model was established to solve the following problems:

Question 1: For each catalyst combination in Annex 1, study the relationship between ethanol conversion, C4 olefin selectivity and temperature, and analyze the test results of the given catalyst combination in Annex 2 at 350 degrees at different times in one experiment .

Question 2: Discuss the effects of different catalyst combinations and temperatures on ethanol conversion and C4 olefin selectivity.

Question 3: How should the catalyst combination and temperature be selected to maximize the C4 olefin yield under the same experimental conditions. At temperatures below 350 degrees, how should the catalyst combination and temperature be selected to maximize the C4 olefin yield.

Question 4: If 5 more experiments are allowed, how should it be designed and give detailed reasons.

2. Model assumptions and notation

- 1). The effect of pressure on the reaction process is not considered.
- 2). The concentration change after gas dissolution is not considered.
- 3). Does not consider human measurement error.

Table 1. Symbol and variable description

Symbol	Name	Unit
s_1	Ethanol conversion	
s_2	Selectivity to C4 Olefins	
T	temperature reflex	$^{\circ}\text{C}$
y_1	Ethanol conversion	
y_2	Selectivity to C4 Olefins	
t	Reaction time	min
h	Ethanol conversion	
Δc	change in concentration per unit time	
Δt	Unit of time	min
c_0	concentration	mol/L
v	reaction speed	
x_1	temperature	
x_2	HAP:co/SiO_2	
x_3	Co load	
x_4	Ethanol concentration	
x_5	total mass	

3. Study on the heat transfer performance of winding single tube heat exchanger

3.1 Build the model

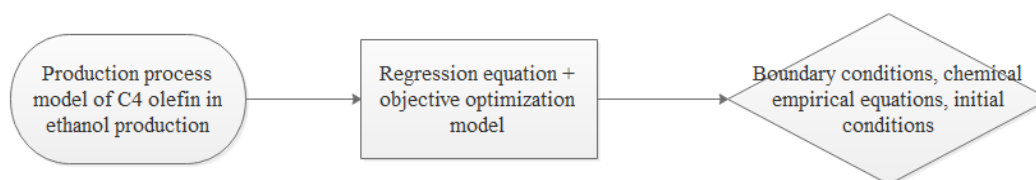


Figure 1. Simplification of the model

(1) Establishment of ethanol conversion rate and temperature regression model

First, we determine the research object as the reactant ethanol, and determine the variable representing the research object, that is, the explained variable of the model is the conversion rate of ethanol, and establish the function of the conversion rate of ethanol with temperature [2] [3]:

$$s_1 = f_1(T) \quad (1)$$

where s_1 is the ethanol conversion rate and T is the reaction temperature.

(2) Establishment of selectivity and temperature regression models for C4 olefins

First, we determine the research object as the product C4 olefin, and determine the variable representing the research object, that is, the explained variable of the model is the selectivity of C4 olefin, and establish the function of the selectivity of C4 olefin as a function of temperature:

$$s_2 = f_2(T) \quad (2)$$

where s_2 is the C4 olefin selectivity and f_2 is the reaction temperature.

3.2 Relationship Analysis

(1) Establishment of time and ethanol conversion rate regression model

For the data given in the title, the conversion rate of the reactant ethanol and the selectivity data of the product C4 olefin and other impurities are measured as a function of time under the same temperature and the same catalyst combination. To analyze the relationship of these test results, we build a time-to-ethanol conversion regression model:

$$y_1 = f_3(T) \quad (3)$$

where y_1 is the ethanol conversion rate and t is the reaction time.

(2) Establishment of time and C4 olefin yield regression model

According to the data, it can be obtained that different reaction conditions have an important impact on the yield of C4 olefins, and it is known that the yield of C4 olefins is the product of ethanol conversion and C4 olefin selectivity. In this experiment, the combination of temperature and catalyst was controlled unchanged, only the reaction time was changed, so the relationship model between time and the yield of C4 olefins was established:

$$h = f_4(T) \quad (4)$$

where h is the ethanol conversion rate, and t is the reaction time.

According to the calculation formula of C4 olefin yield, it can be obtained:

$$h = y_1 \cdot y_2 \quad (5)$$

where y_1 is the ethanol conversion and y_2 is the C4 olefin selectivity.

In summary, the time-C4 olefin yield regression model is established as follows:

$$\begin{cases} h = f_4(t) \\ h = y_1 \cdot y_2 \end{cases} \quad (6)$$

3.3 Solution of the model

(1) Considerations for ethanol conversion

Since the overall conversion rate can better reflect the relationship between the conversion rate and industrial conditions (catalysts, reactors, and reaction conditions), it is beneficial to analyze chemical data to optimize industrial conditions, thereby reducing economic costs. Therefore, this model adopts the one-way conversion rate and the conversion rate per unit time. The calculation formula is as follows:

$$\varphi = \left(\frac{\Delta c}{c_0} \right) = \frac{\frac{\Delta c}{\Delta t}}{\frac{c_0}{\Delta t}} = \frac{v}{c_0} \quad (7)$$

where Δc is the change in concentration per unit time, Δt is unit time, c_0 is concentration, and v is reaction rate.

(2) The relationship between ethanol conversion and temperature

Considering that the conversion rate of ethanol is the conversion rate per unit time, that is, the reaction rate per unit time determines the conversion rate per unit time (the amount of ethanol intake is constant and does not determine the functional relationship), the external factors affecting the reaction rate of this model are mainly external factors. There are catalyst performance, reactant concentration, temperature (ignoring the consideration of pressure, because ethanol has less influence on pressure due to the amount of gas reduced by the reaction, so the degree of influence of pressure on the reaction rate per unit time is compared to temperature, reactant concentration. , the catalyst performance is significantly smaller.)

At the same time, the temperature will affect the activity of the catalyst and the properties of the substance itself, the reaction rate, and whether it can still meet the temperature conditions for the reaction to occur. From the reaction rate formula:

$$\frac{dc}{dt} = k(T) \cdot A^m \cdot B^n \quad (8)$$

Where c is the concentration, t is the reaction time, $k(T)$ is the reaction rate constant and a function related to temperature, A and B refers to the reactants.

(3) Regression of ethanol conversion rate and temperature by chemical kinetics empirical equation

According to the analysis in 2, in order to combine practice and theory, the empirical equation of chemical kinetics is established from empirical equation (8), the kinetic equation is regressed from the data of ethanol conversion rate and temperature (φ, T_1, T_2), and the 21 combinations are made by R language. Point diagram, according to the trend of the scatter diagram to judge whether it is a linear or nonlinear regression analysis, and get the regression equation. And the scatter points are fitted into a function image, and the functional relationship between the ethanol conversion rate and temperature of each combination is obtained, see the appendix.

(4) Test the regression equation using significance tests and goodness-of-fit criteria

In the research of practical problems, we cannot conclude in advance that there is a linear (non-linear) relationship between ethanol conversion rate and temperature. Through the significance test definition and experimental data, the functional relationship of each catalyst combination is obtained ($p < 0.05$), and the fitting degree is good [4].

(5) Conclusions

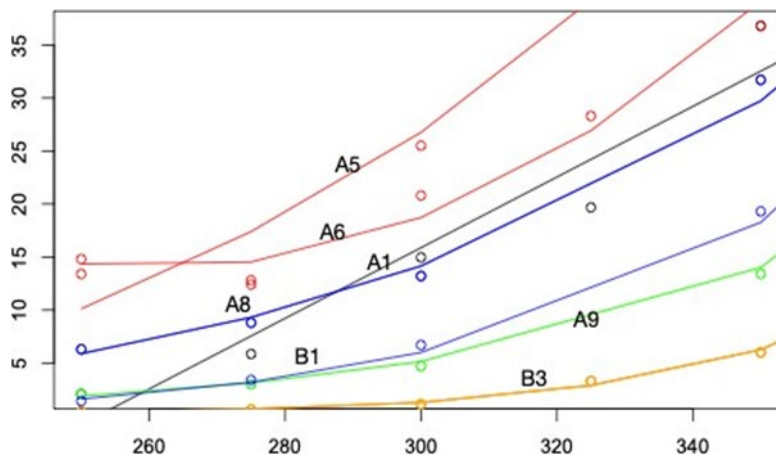


Figure 2. Various regression functions represent images

According to the obtained regression equation function, we can divide it into five categories: linear regression equation, polynomial regression equation (highest order quadratic), polynomial regression equation (highest order cubic), logarithmic regression equation, power function regression equation. At the same time, it is found that the temperature has a positive effect on the ethanol conversion rate. The higher the temperature, the higher the ethanol conversion rate, although the growth rate of different catalyst combinations is different. According to the image extrapolation, it is speculated that different catalyst combinations have different initial reactions. The lower the starting reaction temperature, the earlier the temperature at which the reaction is started is reached, and the earlier the reaction starts.

(6) Ethanol conversion versus time

Because this model is designed for the catalyst, the ethanol conversion rate should first increase and then decrease as the reaction starts and ends. Because the ethanol conversion rate is the highest when the catalyst activity is the highest, the unit reaction rate is the largest, so the ethanol conversion rate also reaches the maximum, and then the unit reaction rate As the rate decreases, the ethanol conversion rate (unit time) decreases. It can be seen from the figure that the ethanol conversion rate

has been in a declining state, which is the part of the curve after reaching the highest catalyst activity. From the downward trend of the curve, it can be found that the decreasing speed is getting slower and slower and tends to a stable value. After equilibrium, the equilibrium conversion rate was 29.9%.

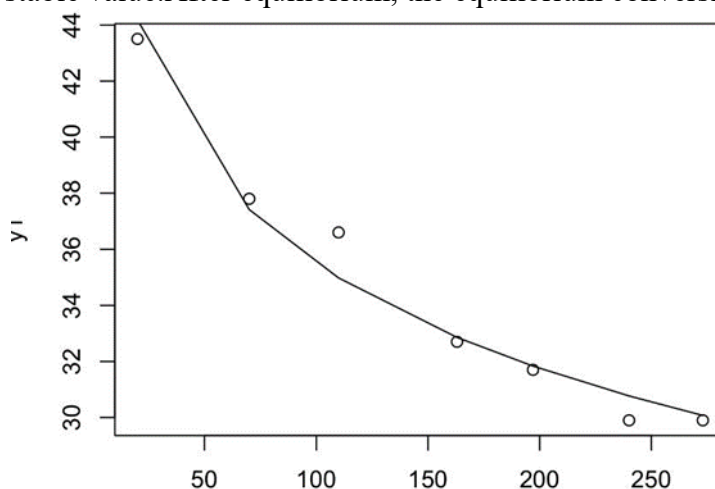


Figure 3. Graph of ethanol conversion versus time

(7) C4 olefin selectivity versus time

It can be roughly seen from the scatter diagram that the C4 selectivity first decreased, then increased and then decreased. Combined with the change of ethanol conversion rate, it can be seen that the decrease in the early stage decreased with the decrease in the ethanol conversion rate, and the decrease in the later period may be due to the continuous addition of ethanol. The resulting other substances and by-products of the reaction lead to poisoning of the catalyst, thereby reducing the selectivity of the catalyst and the yield per unit time.

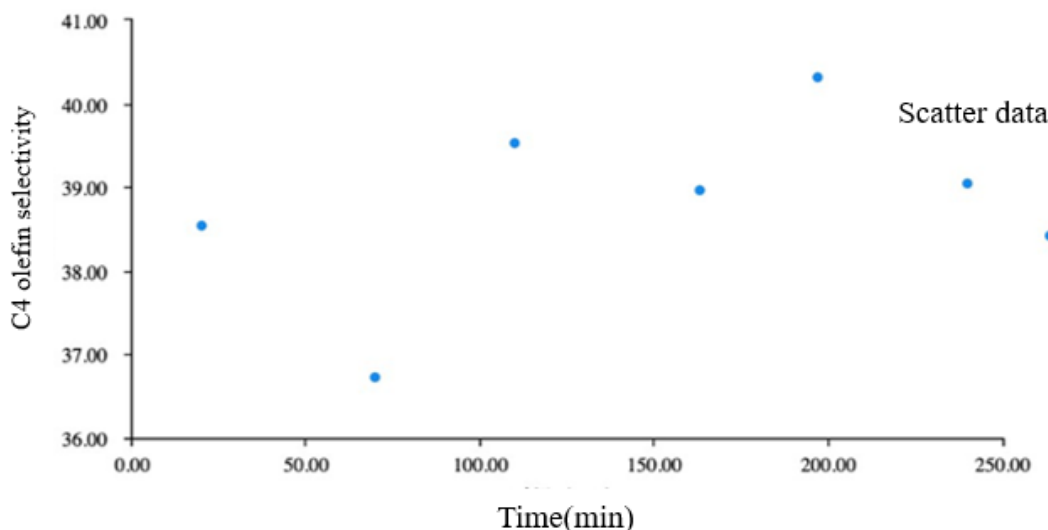


Figure 4. C4 olefin selectivity versus time plot

4. Effects of different catalyst combinations and temperatures on ethanol conversion and C4 olefin selectivity

4.1 Model establishment

(1) Determination of the objective function

It can be seen from the title that in the process of preparing C4 olefins, the influence of temperature, the combination of catalysts and the separation of Co loading, Co/SiO₂ and HAP charging ratio, and ethanol concentration on the selectivity and yield of C4 olefins [8]. The population is subdivided into five variables, namely "temperature", "HAP:co/SiO₂", "Co/SiO₂", "Alcohol concentration",

"Co/SiO₂ total mass". Taking the previous five variables as independent variables, the conversion rate of ethanol is the dependent variable, and the abstract functional relationship that the conversion rate of ethanol is changed due to the five independent variables is expressed. The formula is as follows:

$$Z_1 = g_1(x_1, x_2, x_3, x_4, x_5) \quad (9)$$

Among them, Z_1 is the ethanol conversion rate, x_1 is the temperature, x_2 is HAP:co/SiO₂, x_3 is Co/SiO₂, x_4 is the ethanol concentration, and x_5 is Co/SiO₂ total mass.

In the same way, taking the five variables of "temperature", "HAP:co/SiO₂", "Co/SiO₂", "Alcohol concentration", "Co/SiO₂ total mass" as independent variables, and taking the selectivity of C₄ olefins as the dependent variable, it shows that the selectivity of C₄ olefins is caused by these five variables. The abstract function relation expression that changes with the variable is as follows:

$$Z_2 = g_2(x_1, x_2, x_3, x_4, x_5) \quad (10)$$

Therefore, a general functional relation expression can be established to express the above relation, as follows:

$$Z_i = g_i(x_1, x_2, x_3, x_4, x_5) \quad (i = 1, 2) \quad (11)$$

Considering the error between the estimated value and the true value, by the least square's method, we define a criterion that the sum of squares of the difference between the estimated value and the true value is the smallest, thus transforming the problem into an optimization problem, the formula is as follows:

$$y = \min \sum_{i=1}^N [\hat{Z}_i - Z_i]^2 \quad (12)$$

(2) Limitation of reaction temperature

First, the temperature determines whether the reaction can occur and continue in actual production. Second, the reaction temperature affects the reaction rate and catalyst performance, thereby affecting the ethanol conversion and C₄ olefin selectivity. Considering the activity of the catalyst, there is a temperature range. At the same time, the reaction temperature cannot be increased indefinitely. At the same time, if the temperature is too high, the catalyst will be deactivated or the reaction will stop, and the reaction has a minimum reaction temperature limit. According to the data, the iron-chromium catalyst is a brown cylinder or Plate-like solid particles, the active temperature is 350 ~ 550 °C, so the range can be set as follows according to the experimental data:

$$250 \leq x_1 \leq 500 \quad (13)$$

(3) Limitation of Co/SiO₂ and HAP charging ratio

Because the charging ratio should be selected within a suitable range. On the one hand, when there is too much Co/SiO₂, the catalyst will be wasted; on the other hand, HAP as a carrier can increase the activity of the catalyst, but the activity of the catalyst will also increase to a certain upper limit and remain unchanged, so the range can be set as follows according to the experimental data:

$$0 < x_2 \leq 2 \quad (14)$$

(4) Proportional restrictions of Co/SiO₂

According to the experimental data, the range of the ratio is as follows:

$$0.5 \leq x_3 \leq 2 \quad (15)$$

(5) Limitation of ethanol concentration

The concentration of the reactant will affect the rate of the chemical reaction. If the concentration of the reactant is too low, it will not have a significant positive effect on the reaction, and if the concentration of the reactant is too high, it will inhibit the progress of the reaction. Therefore, ethanol can be set according to the experimental data. The concentration ranges are as follows:

$$0.5 \leq x_4 \leq 1.68 \quad (16)$$

(6) Total mass limit of Co/SiO₂

The reduction of the activation energy of the catalyst in the reaction has a certain limit, so the range of the total mass can be set as follows according to the experimental data:

$$10 \leq x_5 \leq 200 \quad (17)$$

In summary, the optimization model for problem 2 is established as follows:

$$\begin{aligned} \text{Objective function: } y &= \min \sum_{i=1}^N [\hat{Z}_i - Z_i]^2 \\ \text{s.t. } &\begin{cases} 250 \leq x_1 \leq 500 \\ 0 < x_2 \leq 2 \\ 0.5 \leq x_3 \leq 2 \\ 0.4 \leq x \leq 1.68 \\ 10 \leq x_5 \leq 200 \end{cases} \end{aligned} \quad (18)$$

4.2 Solution of the model

(1) Algorithm steps

Step1: Integrate data into tables and position them for easy calculations

Step2: Enter the required data in rows and columns: $x_i, y_i, i=1,2,3,4,5, j=1,2$

Step3: Set the parameters by means of multivariate nonlinear polynomial fitting

Step4: Bring in data to calculate the value of each parameter

Step5: Judging the residual r based on the fitted value and the true value

Step6: Select the appropriate parameters to calculate the fitting equation, if the residual is not within the reference range, go to Step3

Step7: Based on the fitted equations, analyze (different catalyst combinations and temperatures) and (ethanol conversion and olefin selectivity)

Step8: Output result

(2) Calculation results and analysis

Table 2. Goodness of fit and significance for ethanol conversion and C4 olefin selectivity

	R^2	p
Z_1	0.791	$< 2.2e^{-16}$
Z_2	0.7838	$< 2.2e^{-16}$

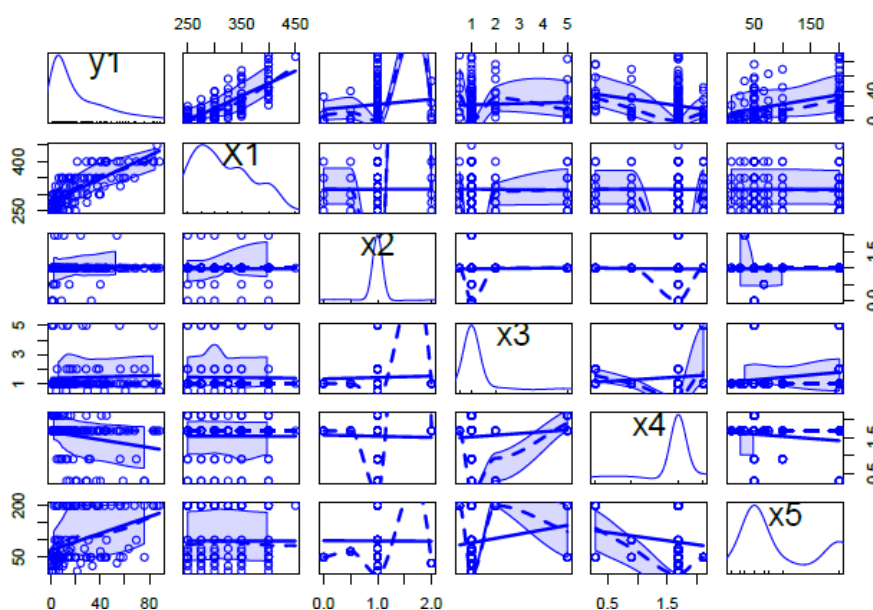


Figure 5. Scatterplot matrix of ethanol conversion versus independent variables

Using the model established above, the set value in the question is brought into R and MATLAB to solve [7], and the functional relationship expression of the ethanol conversion rate with the five factors of "temperature", "HAP:co/SiO2", "Co/SiO2", "Alcohol concentration", "Co/SiO2 total mass" and the C4 olefins are obtained. The selectivity of C4 olefins varies with "temperature", "HAP:co/SiO2", "Co/SiO2", "Alcohol concentration", "Co/SiO2 total mass" and the functional relationship expression of the five factors of ethanol concentration and total mass.

The regression results of different catalyst combinations, temperature and ethanol conversion were obtained by using multiple linear stepwise regression (18), R2 is 0.791, which means 79.1% of the changes in the five variables of "temperature", "HAP:co/SiO2", "Co/SiO2", "Alcohol concentration", "Co/SiO2 total mass". And through the F test ($F=103.422$, $p=0.000<0.05$), this model is valid.

$$Z_1 = -89.445 + 0.336x_1 + 6.590x_2 - 8.322x_3 + 0.119x_4 \quad (19)$$

Similarly, the regression results of different catalyst combinations and temperature and C4 olefin selectivity are as follows:

$$Z_2 = 7.778 = 7.7780.186x_1 + 13.581x_2 + 0.001x_1^2 - 6.489x_2^2 - 0.656x_3^2 + 0.045x_4 \cdot x_5 \quad (20)$$

Table 3. AIC and BIC guidelines

x1	x2	x3	x4	x5	RSS	R2	adjR2	Cp	BIC
*					23924.79	0.5937004	0.5900728	100.454064	-93.20337
*				*	14685.67	0.7506026	0.7461090	21.182257	-144.10409
*	*	*			12765.58	0.7832104	0.7772979	6.292178	-155.34157
*	*		*	*	12279.62	0.7914631	0.7838103	4.017466	-155.02986
*	*	*	*	*	12277.63	0.7914968	0.7818438	6	-150.31210

The AIC and BIC criteria are two important criteria for selecting models in multiple regression. In multiple regression analysis, in order to prevent problems such as overfitting (not only to make the model more explanatory, but also to have a little tension), Akaike (1978) and Schwarz (1978) proposed AIC and BIC respectively as regression models. standard. In regression models, both values are as small as possible.

$$AIC = n \ln \left(\frac{RSS_p}{n} \right) + 2p$$

$$BIC = n \ln \left(\frac{RSS_p}{n} \right) + p \ln(n) \quad (21)$$

Table 4. Global Optimal Criterion

x1	x2	x3	x4	x5	RSS	R2
*					23924.79	0.5937004
*				*	14685.67	0.7506026
*	*	*			12765.58	0.7832104
*	*		*	*	12279.62	0.7914631
*	*	*	*	*	12277.63	0.7914968

Table 5. No multicollinearity results

temperature	HAP/Co: SiO2	Ethanol concentration	Total mass
FALSE	FALSE	FALSE	FALSE

5. Exploring the choice of catalyst combination and temperature

5.1 Establishment of model without temperature constraint

(1) Determination of the objective function

Because we need to determine the different catalyst combinations and temperatures that maximize the yield of C4 olefins under the same experimental conditions, that is, the goal is to maximize the yield of C4 olefins, and accordingly, the objective function can be obtained as follows:

$$Y = \max(Z_1 \cdot Z_2) \quad (22)$$

Where Y is the yield of C4 olefins, y_1 is the ethanol conversion rate, and y_2 is the selectivity of C4 olefins.

(2) Determination of Constraints

After analysis, the objective function will be constrained by the range of five factors: "temperature", "HAP:co/SiO₂", "Co/SiO₂", "Alcohol concentration", "Co/SiO₂ total mass". This constraint is the same as the constraint in model two.

$$\begin{cases} 250 \leq x_1 \leq 500 \\ 0 < x_2 \leq 2 \\ 0.5 \leq x_3 \leq 2 \\ 0.4 \leq x_4 \leq 1.68 \\ 10 \leq x_5 \leq 200 \end{cases} \quad (23)$$

In summary, the optimization model for problem 3 is established as follows:

$$\begin{aligned} Y &= \max(Z_1, Z_2) \\ s.t. \begin{cases} 250 \leq x_1 \leq 500 \\ 0 < x_2 \leq 2 \\ 0.5 \leq x_3 \leq 2 \\ 0.4 \leq x_4 \leq 1.68 \\ 10 \leq x_5 \leq 200 \end{cases} \end{aligned} \quad (24)$$

5.2 Establishment of model with temperature constraint

(1) Determination of the objective function

When the temperature is lower than 350 degrees, in order to determine the different catalyst combinations and temperatures when the yield of C4 olefins is the largest, the target is set as the maximum yield of C4 olefins [9], and the objective function is determined as:

$$W = \max(Z_1 \cdot Z_2) \quad (25)$$

(2) Determination of Constraints

According to the description of the objective function, the difference between this question and the first model is that if the temperature is lower than 350 degrees, it is only necessary to add a constraint on the basis of the previous model:

$$\begin{cases} 250 \leq x_1 \leq 350 \\ 0 < x_2 \leq 2 \\ 0.5 \leq x_3 \leq 2 \\ 0.4 \leq x_4 \leq 1.68 \\ 10 \leq x_5 \leq 200 \end{cases} \quad (26)$$

In summary, the optimization model for problem 3 is established as follows:

$$W = \max(Z_1, Z_2)$$

$$s.t. \begin{cases} 250 \leq x_1 \leq 350 \\ 0 < x_2 \leq 2 \\ 0.5 \leq x_3 \leq 2 \\ 0.4 \leq x_4 \leq 1.68 \\ 10 \leq x_5 \leq 200 \end{cases} \quad (27)$$

5.3 Solution of the model

(1) Algorithm steps

Step1: Integrate data into tables and position them for easy calculations

Step2: Enter the required data in rows and columns: $x_i, y_i, i=1,2,3,4,5, j=1,2$

Step3: Set the parameters by means of multivariate nonlinear polynomial fitting

Step4: Bring in data to calculate the value of each parameter

Step5: Judging the residual r based on the fitted value and the true value

Step6: Select the appropriate parameters to calculate the fitting equation, if the residual is not within the reference range, go to Step3

Step7: find the regression equation

Step8: Optimal C4 Olefin Yield Based on Obtained Functional Equation

Step9: Output result

(2) Calculation results and analysis

The data in Appendix is processed to obtain the yield of C4 alkene, and the chemical empirical equation is used to establish the multiple regression between the yield of C4 alkene and the combination of catalyst and temperature, and the regression equation function of the yield of C4 alkene is obtained, using matlab Solve the solution of the first model: the temperature is 400 °C, the ratio of 1 wt%Co/SiO₂-HAP to HAP is 1:1, and the yield is the largest; the solution of the second model: 350 °C, - The ratio of 1 wt%Co/SiO₂-HAP to HAP is 1:1.

6. Conclusion

Based on the results of a series of experiments on different catalysts at different temperatures in a chemical laboratory, the process conditions for the preparation of C4 alkene by ethanol catalytic coupling were explored. After the above analysis, the main conclusions are as follows: ethanol conversion rate and C4 alkene conversion rate Both are positively correlated with temperature, and different catalyst combinations will affect the growth slope and initial value. The charging method has little effect on the experimental results and will not cause too much deviation in the experimental results. The charging ratio has little effect on the ethanol conversion rate, and the ethanol concentration has little effect on the C4 olefin selectivity. The increase of temperature has a great influence on the ethanol conversion rate and C4 olefin selectivity, and the ethanol concentration will reduce the ethanol conversion rate.

References

- [1] Chinese Society for Industrial and Applied Mathematics. 2021 Higher Education Club Cup National College Students Mathematical Modeling Contest Questions [EB/OL]. 2021-09-09.
- [2] Fan Hongzhen, Sun Xiaowei. Methanol to Olefin Process [O]. Beijing: Chemical Industry Press, 2016.
- [3] Chen Zhongxiu. Chemical Thermodynamics (3rd Edition) [O]. Beijing: Chemical Industry Press, 2012.
- [4] Chen Rongye. Optimization of organic synthesis process [O]. Beijing: Chemical Industry Press, 2006.
- [5] Dou Jinmin. Organic Chemical Technology (Second Edition) [O]. Beijing: Chemical Industry Press, 2012.

- [6] Huang Huajiang / Huang Huajiang. Computer Simulation of Practical Chemical Industry [O]. Beijing: Chemical Industry Press, 2004.
- [7] RobertI.Kabacoff, Kabacoff, Chen Gang, et al. R language in practice[J]. People's Posts and Telecommunications Press, 2013.
- [8] Yiming Xu1 et al. Optimization Analysis of C4 Olefin Preparation by Ethanol Coupling[J]. International Journal of Education and Teaching Research, 2022, 3(2).
- [9] Yifei Wang. Study on the process of preparing C4 olefin from ethanol based on optimal method[J]. Academic Journal of Engineering and Technology Science, 2021, 4.0(7.0).