Study on Preparation of Olefins by Coupling ethanol based on partial correlation Analysis and Linear stepwise regression

Shuo Pang [1,#,*], Jingyi Zhu [1,#], Xinshuang Zhang [2,#]

[1] Central South University Business School Changsha, Hunan, 410083 China

[2] Central South University School of Mathematics and Statistics Changsha, Hunan, 410083 China

*Corresponding author. Email: pangshuo220322@163.com

#These authors contributed equally.

Abstract. A correlation analysis was constructed according to the effect of catalyst and temperature on the chemical reaction in the reaction of ethanol coupling to C4 olefin. Firstly, the relationship between the reaction product, temperature, and the reaction product was explored. The chemical mechanism analysis and qualitative image analysis inferred that temperature is related to ethanol conversion and C4 olefin selectivity. Then the Pearson correlation coefficients of each group of temperatures and the corresponding experimental data are calculated, and 90% of the correlation coefficients are greater than 0.9. it is confirmed that they have a significant positive correlation.

Keywords: Ethanol, Preparation of olefin, Correlation analysis, Positive correlation.

1. Introduction

With the improvement of chemical production technology, C4 olefins are used more and more widely, covering but not limited to biology, medicine, chemistry, and the chemical industry. The coupling reaction using ethanol is an essential way of Preparation [4]. Therefore, exploring the influence of catalyst combination (mainly composed of CO loading, Co / SiO2, HAP loading ratio, Co / SiO2 and HAP total dosage, ethanol concentration, and other factors) and temperature on C4 olefin selectivity and C4 olefin yield is explored. The causes and trends and designing appropriate schemes are essential for further improving C4 olefin yield and improving raw material utilization, which has important value and significance.

2. Analysis of reaction mechanism

The Preparation of olefins from ethanol can be divided into two steps: the Preparation of butanol from ethanol and the conversion of butanol to olefins. Among them, according to the literature [1], it can be seen that the direct coupling reaction for preparing butanol from ethanol is an exothermic reaction, which can be roughly divided into:

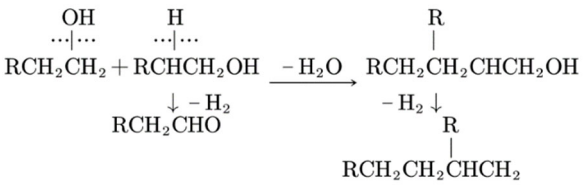


Fig.1 The direct coupling mechanism of ethanol to olefins

Therefore, it can be reasonably inferred that the Preparation of olefins from ethanol is also an exothermic reaction. Next, the preliminary observation data can be found when the temperature increases during the reaction process. The conversion of ethanol will increase under any condition[5-7], and the olefin selection rate will increase, which means that the proportion of olefins in the reaction products increases. The increase of temperature promotes the forward progress of the reaction, which is consistent with the inference.

Arrhenius formula combined with chemical kinetics:

$$k = Ae^{-Ea/RT} \tag{1}$$

Where k is the rate constant, R is the molar gas constant, T is thermodynamic temperature, Ea is the apparent activation energy, and A is a pre-exponential factor (also known as frequency factor)

On this basis, we can use the existing data to preliminarily observe the quantitative relationship between temperature change and ethanol conversion and olefin selectivity, as shown in the Figure:

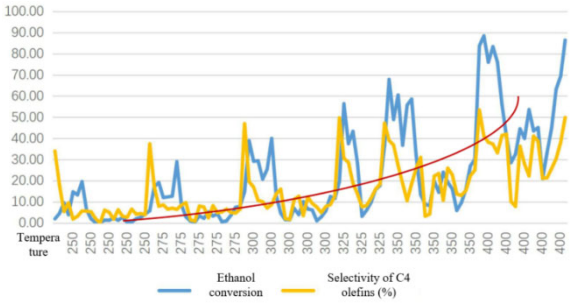


Fig.2 Diagram of correlation between temperature change and olefin selectivity and ethanol conversion

It can be seen from the Figure that with increasing temperature, the overall trend of both shows an upward state, roughly as shown by the red line. Moreover, the rising stage and amplitude are similar. It is shown that olefins and ethanol react in the same direction when the temperature increases, which is consistent with the previous conclusion of the reaction mechanism, indicating that there is a strong correlation between the temperature change and the two.

3. Correlation analysis between ethanol conversion, C4 olefin selectivity, and temperature

For each group of experimental data, it can be regarded as the sample under the experimental conditions of this group. Set the temperature as the variable, the ethanol conversion rate as, and the olefin selectivity as, which represents the number of tests at different temperatures of each group of experiments, so the mean value is:

$$\overline{T} = \frac{\sum_{i=1}^{n} T_i}{n}, \overline{X_1} = \frac{\sum_{i=1}^{n} X_{1i}}{n}, \overline{X_2} = \frac{\sum_{i=1}^{n} X_{2i}}{n} \tag{2}$$

The sample covariance between any two variables is:

$$\begin{cases} \operatorname{cov}(T, X_1) = \dfrac{\sum_{i=1}^{n} (X_{1i} - \overline{X}_1)(Y_i - \overline{Y})}{n-1} \\ \operatorname{cov}(T, X_2) = \dfrac{\sum_{i=1}^{n} (X_{2i} - \overline{X}_2)(Y_i - \overline{Y})}{n-1} \\ \operatorname{cov}(X_1, X_2) = \dfrac{\sum_{i=1}^{n} (X_{1i} - \overline{X}_1)(X_{2i} - \overline{X}_2)}{n-1} \end{cases} \tag{3}$$

The sample correlation coefficient is [2]:

$$\begin{cases} r_{X_1T} = \dfrac{\operatorname{cov}(X_1, T)}{S_{X_1} S_T} \\ r_{X_2T} = \dfrac{\operatorname{cov}(X_2, T)}{S_{X_2} S_T} \\ r_{X_1T} = \dfrac{\operatorname{cov}(X_1, X_2)}{S_{X_1} S_{X_2}} \end{cases} \tag{4}$$

S_T, S_{x1}, S_{x2} are the standard deviation of samples, respectively. This method calculates the correlation coefficient between temperature and the two in each group of experiments. The p-value is used as the judgment basis for the reliability of the calculation results. If the p-value meets the

significance requirements, it shows that the calculation results have high reliability and can well explain the specific relationship and correlation degree between variables to analyze the influence of temperature on them quantitatively.

Tab.1 Calculation results of correlation coefficient with temperature

Experimental group number	Ethanol conversion	p	Olefin selectivity	p
A1	0.965**	0.007655	0.887*	0.044769
A2	0.995**	0.000425	0.914*	0.029731
A3	0.982**	0.000083	0.977**	0.000153
A4	0.998**	0.000009	0.958**	0.002639
A5	0.934**	0.006321	0.970**	0.001374
A6	0.984**	0.002507	0.885*	0.045754
A7	0.999**	0.000019	0.968**	0.006781
A8	0.977**	0.004155	0.992**	0.000926
A9	0.921*	0.026539	0.997**	0.000159
A10	0.923*	0.025371	0.861*	0.0606
A11	0.968**	0.006759	0.989**	0.001374
A12	0.963**	0.008419	0.983**	0.002616
A13	0.937*	0.018991	0.988**	0.001513
A14	0.964**	0.008198	0.959**	0.00985
B1	0.962**	0.008846	0.986**	0.002031
B2	0.929*	0.022395	0.985**	0.002248
B3	0.890*	0.017503	0.971**	0.001252
B4	0.900*	0.014623	0.895*	0.015958
B5	0.913*	0.011056	0.978**	0.000746
B6	0.940**	0.005306	0.982**	0.000477
B7	0.936**	0.005991	0.994**	0.000047

From the calculation results of each group of experiments in the chart, it can be observed that the maximum correlation coefficient is (0.999,0.997), which has been underlined in red; The minimum correlation coefficient is (0.890,0.861), which is underlined in blue. All correlation coefficients are greater than 0.85, and 90% are more excellent than 0.05 and 9. Almost all ethanol conversion and olefin selectivity are significantly positively correlated with temperature.

The only non-significant correlation is the olefin selectivity and temperature of a10. When calculating the correlation coefficient of olefin selectivity, its p-value is 0.0606, greater than 0.05, indicating that the calculated correlation coefficient is not significant (marked with red *). All other p-values are significant, so it is likely to be caused by some accidental factor in the experiment, which can be ignored. In addition, the error will continue to be explored, verified, modified, and improved in the subsequent experimental process.

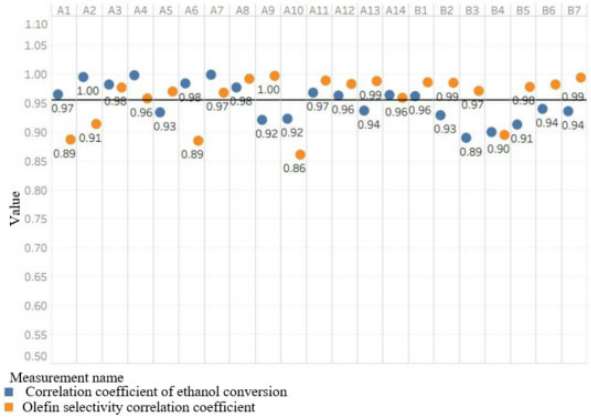


Fig.3 Distribution of correlation coefficient between ethanol conversion and temperature, olefin selectivity, and temperature

In addition, we plotted the scatter plots of all correlation coefficients and calculated that the mean value of their correlation coefficients was 0.955, where the black horizontal line is the position of the mean value of the correlation coefficient. In conclusion, we can conclude that both ethanol conversion and olefin selectivity are significantly positively correlated with temperature.

4. The influence of different factors on the experiment was analyzed

4.1. Total Co / SiO2 and HAP charge

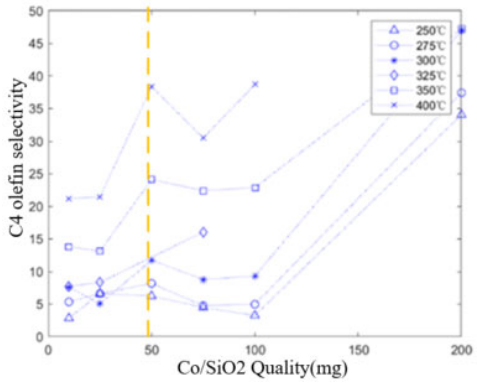


Fig.4 Change of ethanol conversion with CO / SiO2 content

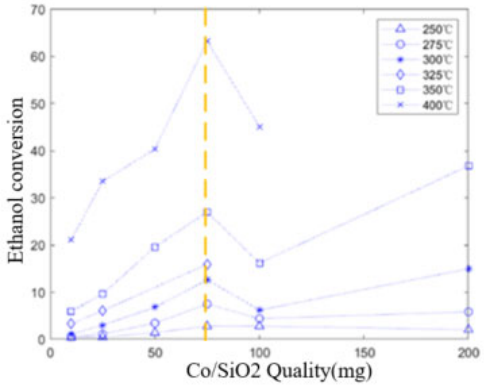


Fig.5 Olefin selectivity with CO / SiO2 content

For ethanol conversion rate, although the experimental data are missing at 400 ℃, it can be inferred from the combination of 350 ℃ and 325 ℃. When the total amount of catalyst is small initially, the degree of reaction is low. With the increase of Catalyst Co / SiO2, the reaction continues to move forward. The conversion rate peaks at 75mg, then gradually decrease and reach the minimum at 100mg. The ethanol conversion first increases and then decreases, then continues to rise; The dosage of catalytic materials has the same trend for C4 olefin selectivity, but it reaches the highest and lowest points in advance.

4.2. Load Co

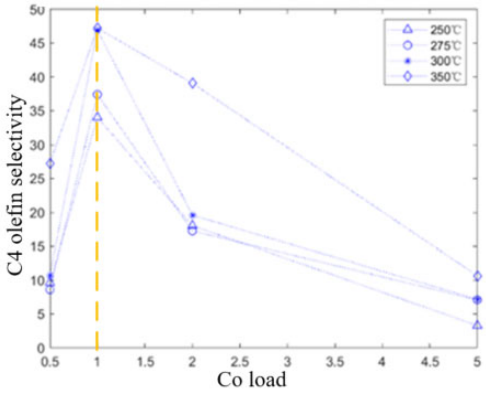


Fig.6 Change of ethanol conversion with CO loading

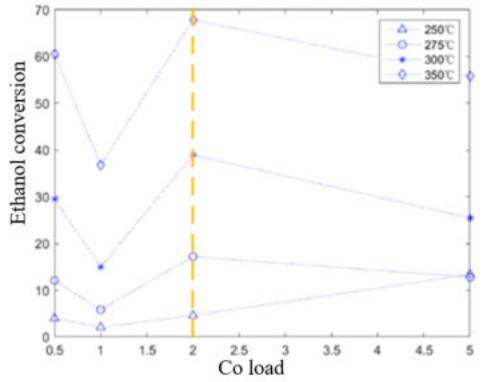


Fig.7 Olefin selectivity with CO loading

It can be seen that the effects of CO loading on ethanol conversion and olefin selectivity are consistent at different temperatures. The ethanol conversion rate increases first and then decreases with the co-load, and the inflection point is 1. The selectivity increases slightly with the increase of CO loading and then decreases gradually with CO loading.

4.3. Ethanol concentration

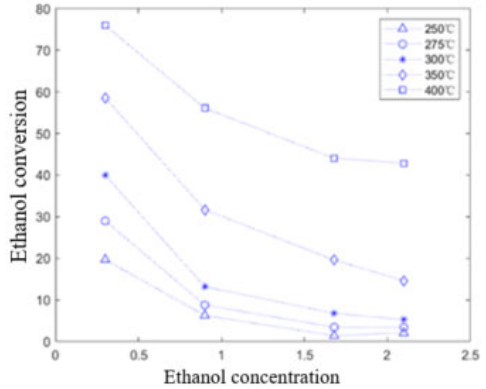


Fig.8 Change of ethanol conversion rate with ethanol concentration

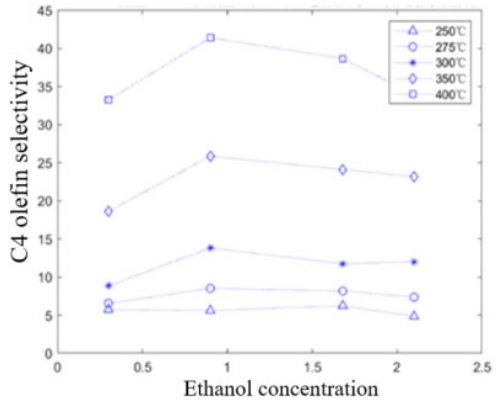


Fig.9 Variation of olefin selectivity with ethanol concentration

It can be seen that with the increase of ethanol concentration, the conversion rate of ethanol decreases gradually. By consulting the data [1], we know that the reaction is reversible, and the reactant will not be completely converted. With the gradual increase of substrate concentration, the relative content of catalyst can not fully meet the demand of catalytic reaction, so the conversion rate of ethanol decreases gradually. Because the reaction is endothermic, the increase in temperature promotes the positive progress of the reaction. Therefore, the experimental data with high temperatures are in a higher position. However, even if the concentration of reactants continues to increase, the selectivity of olefins in the products has not changed much, indicating that the reaction has been carried out thoroughly, and even increasing the concentration of ethanol will not have a great impact on the selectivity.

5. Catalytic factor model based on multiple linear stepwise regression

We choose the most basic, multiple linear regression model [8]. It is a form of linear regression used when two or more predictors are used. Its formula is:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + \varepsilon \quad (5)$$

Among them, we set the independent variable Co / SiO2 dosage as, CO loading as, ethanol concentration as, temperature as 14 ℃, dependent variable ethanol conversion as, C4 olefin selectivity as; Since the ratio of CO / SiO2 to HAP is 1:1 and HAP does not have any characteristics different from CO / SiO2 for this reaction, we only choose Co / SiO2 as the representative catalyst. Considering the interaction between independent variables, the most basic multiple regression form cannot be used, so we introduce the interactive terms, x_1x_2、x_1x_3, x_1x_4, x_2x_3、x_2x_4, x_3x_4 and through the combination of any two different variables, and finally build a formula containing ten independent variables.

The variables are introduced into the model one by one. The selected independent variables are tested one by one for each independent variable. They are deleted when the initially introduced explanatory variables become no longer significant due to the introduction of later explanatory variables. Introduce or eliminate an independent variable from the regression equation. As a step of stepwise regression, F test shall be conducted to ensure that only significant variables are included in the regression equation before introducing new variables each time

First regression result:

Tab.2 First regression results of ethanol conversion

Variable	Standardization coefficient	Significance	Partial correlation	Variance expansion factor
x_4	0.726	3.59E-18	0.722	3.934
x_1x_4	1.226	2.15E-07	0.48	40.771
x_3x_4	-0.528	1.32E-11	-600	4.024
x_1	-1.382	4.02E-08	-505	45.431

x_1x_3	0.569	1.17E-07	489	8.367

There may be multiple collinearities among the model variables obtained by the first iteration. In order to eliminate this effect, we introduce the variance expansion factor VIF:

Suppose there are k independent variables (k is 5), then the of the M independent variable is:

$$VIF_m = \frac{1}{1 - R^2_{1\sim k/m}} \tag{6}$$

Where $R^2_{1\sim k/m}$ does regression of the m independent variable obtain the goodness-of-fit as the dependent variable and the remaining 1 independent variable as a dependent variable. The larger the VIF_m, the greater the correlation between the m-th variable and other variables. Therefore, it can define a regression model:

$$VIF = \max\left\{VIF_1, VIF_2, \cdots, VIF_k\right\} \tag{7}$$

Where the adjusted R square of the sample is 0.871, and the significance test of each variable is established, it is evident that the fitting degree is good, and the obtained model can well explain the dependent variables. However, according to the calculation results, it can be judged that there are multiple collinearities in the first regression, and the model is greater than 10, and the variable appears more frequently, so we propose variables and enter the second regression.

Combined with the previous analysis, ethanol is used as a reactant, and the reaction can be carried out in reverse, increasing the concentration of ethanol at the same temperature. The same type and number of catalysts will decrease ethanol conversion, and the variable parameter is negative. When the concentration of ethanol increases, it is likely that the reaction can not be carried out entirely due to the lack of catalysts, so that the number of catalysts can be increased at the same time, and the two can meet each other's needs without considering the influence of other factors. in order to improve the conversion of ethanol. The quantitative ethanol conversion equation we obtained accords with the actual analysis and has high accuracy and credibility.

6. Conclusions

With the improvement of chemical production technology, C4 olefins are used more and more widely, including but not limited to biology, medicine, chemistry, and chemical engineering. The coupling reaction of ethanol is an essential way of Preparation, so this paper mainly explores the factors affecting the Preparation of C4 olefins by ethanol coupling. Firstly, through chemical mechanism analysis and qualitative image analysis, it is speculated that there is a correlation between temperature and ethanol conversion and C4 olefin selectivity. Then, we calculate the Pearson correlation coefficients of each temperature group and the corresponding experimental data and get that 90% of the correlation coefficients are more than 0.9. it is confirmed that they have a significant positive correlation, which provides a relevant research basis for the Preparation of olefins by ethanol coupling.

References

[1] Yang C, Meng ZY. Bimolecular Condensation of Ethanol to 1-Butanol Catalyzed by Alkali Cation Zeolites[J]. Journal of Catalysis, 1993, 142(1):37-44.

[2] Shi Haina, Liu Yutian, Li Shien, Liu Gang, Zhang Jinxia, Duan Rui. Correlation and regression analysis of body mass and body size index of Duper sheep [J]. Feed Research, 2021 and 44 (13): 85-89.

[3] Zeng Jipeng, Lu Yongfeng, Hou Kepeng. Application of two-factor analysis method in the analysis of the influence of fault on slope stability [J]. Chemical Minerals and processing, 2021, 50 (04): 6-8, 13.

[4] Huang Boqiang, Li Qicai. Linear regression modeling of two-factor analysis of variance with interaction [J]. Statistics and decision-making, 2021 Magi 37 (01): 10-15.

[5] Ge Peiyun, Xu Hengying, Bai Chenglin, Yang Lishan, Sun Weibin, Yu Xinkuo, Zhao Ruqing, Li Baokun, Zhou Tanglei, Qin Peng, Bi Yanfeng, Cao Lingguo. A Low-complexity OSNR monitoring scheme based on amplitude variance analysis [J]. Optik,2021,246:

[6] Tour soldiers, Yan Yan. Stepwise regression analysis and its application [J]. Statistics and decision-making, 2017 (14): 31-35.

[7] Shen Huayu, Wang Zhaoxia, Gao Chengyao, Qin Juan, Yao Fubin, Xu Wei.Definition of units in hidden layer of BP neural network [J]. Journal of Tianjin University of Technology, 2008 (05): 13-15.

[8] Zhao Yuxin. Study on multicollinearity in multivariate linear regression [J]. Industry and Science and Technology Forum, 20119, 18 (03): 71-72.