Analysis Methodology of C4 olefins by ethanol coupling based on correlation analysis

Yiran Wang[1, *, #], Zhuer Lu[2, a, #], Yishan Xie[2, b, #]

[1]Beijing Forestry University, College of Science, Beijing, 100083

[2]Beijing Forestry University, College of Biological Sciences and Technology, Beijing, 100083

*Corresponding Author Email: yrjjlili@gmail.com, [a]luzhuer0506@126.com, [b]mayxie_2000@163.com

#These authors contributed equally.

Abstract. C4 olefins are widely used to produce various chemical products and pharmaceuticals. Besides, Ethanol is a good raw material for the production of C4 olefins instead of petroleum cracking and ethylene cracking. First, the optimal choice of catalyst combination and temperature for C4 olefins preparation has been studied, and the process conditions forC4 olefins preparation by catalytic coupling of Ethanol are of great significance and value for promoting industrial production. Besides, a simple scoring method is also proposed to evaluate the differences in the catalytic effects of different experimental groups. We first analyzed the correlation with the R language and found that the temperature was positively correlated with the C4 olefin selectivity and ethanol conversion rate. Then, we fitted 21 curves sets of ethanol conversion rate and C4 olefin selectivity with temperature. Second, R language is used to perform correlation analysis to screen out variables that are significantly related to ethanol conversion and C4 olefin selectivity, then obtain a stepwise regression equation to quantitatively describe the effects of different catalyst combinations and temperatures on ethanol conversion and C4 olefin selectivity. Thirdly, we used the stepwise regression method to obtain the optimal local solution at each temperature, and the optimal combination of temperature and catalyst was 0.5wt% Co loading at 400℃, 200mgCo/SiO2-200mgHAP, ethanol concentration is 0.3ml/min. If the temperature is required to be controlled below 350 degrees, the optimal combination of temperature and catalyst is 2.145wt% of Co loading, 200mgCo/SiO2-200mgHAP, and ethanol concentration of 2.1ml/min at 350℃.

Keywords: Production of C4 olefins from Ethanol, correlation analysis, stepwise regression, R language.

1. Introduction

C4 olefins (i.e., 1,3-butadiene) are widely utilized to produce various chemical products and medicines. [1] Besides, Ethanol is a kind of clean energy and good raw material to replace petroleum cracking and ethylene cracking to produce C4 olefins. In the reaction

$$2CH_3 2CH_3OH \frac{catalyst}{temperature} C_4H_6,$$

Catalyst, temperature, catalyst usage, and heating temperature play an essential role in the conversion of Ethanol and the selectivity of C4 olefins [1, 2]. Therefore, finding the best catalyst combination and optimizing the heating temperature is the key to improving final production efficiency.

HAP (Hydroxyapatite) is a catalyst carrier, and the SiO2-HAP catalyst loaded with a certain wt% of Co has both acid active sites and active base sites and has good catalytic efficiency. During the preparation, the catalyst combination (i.e., Co loading, Co/SiO2, HAP charge ratio, and ethanol concentration) and temperature will affect ethanol conversion, C4 olefin selectivity, and C4 olefin final absorb rate. [3] Thus, it is of great practical significance and economic value to obtain the optimal process conditions for preparing C4 olefins by catalytic coupling of Ethanol through the study of catalyst combinations. We selected 21 catalyst combinations, designed cross-experiments, and obtained accurate experimental data. The analysis of experimental data is needed.[4]

Correlation analysis refers to analyzing two or more correlated variable elements to measure the degree of correlation between two factors. We correlated data within each catalyst combination to study ethanol conversion, C4 olefin selectivity versus temperature, and their correlation with other by-product selectivities [1].

2. Correlation Analysis

2.1. Methodology

Correlation analysis [5, 6] is performed on each experimental data of the 21 catalyst combinations. Through the R language [7], 21 groups of correlation coefficient tables, correlation analysis diagrams, and P-value tables were obtained. The range of the correlation coefficient is [-1,1]. The closer the absolute value of the correlation coefficient is to 1, the greater the correlation between variables. In the correlation analysis graph, the darker the color red, the more evident in negative correlation, and the darker the color blue, the more obvious in positive correlation. In the P-value table, where $P>0.05$ indicates a significant correlation.

From the correlation analysis of each experimental group, the temperature has a relatively obvious positive correlation with etha_rate and C4_rate in all experimental groups, but the correlation coefficients are different to some extent.

The following takes catalyst combinations A5 and B3 as examples to explain the correlation analysis results in detail.

1) *A5* combination

The analysis shows that under the action of the A5 catalyst combination, the temperature is significantly positively correlated with the C4 olefin selectivity and the ethanol conversion rate, and the C4 olefin selectivity is significantly positively correlated with the ethanol conversion rate. The temperature was also negatively correlated with selectivity to fatty alcohols with carbon numbers 4-12 and positively correlated with selectivity to ethylene.

Among the by-products, the selectivity of fatty alcohols with carbon numbers 4-12 was significantly negatively correlated with the selectivity of C4 olefins and ethylene. Ethylene selectivity is positively correlated with C4 olefin selectivity and ethanol conversion. Acetaldehyde selectivity, methylbenzaldehyde, and methyl benzyl alcohol selectivity were not significantly correlated with other variables. The correlation coefficient and P-value are shown in TABLE I and FIGURE I.

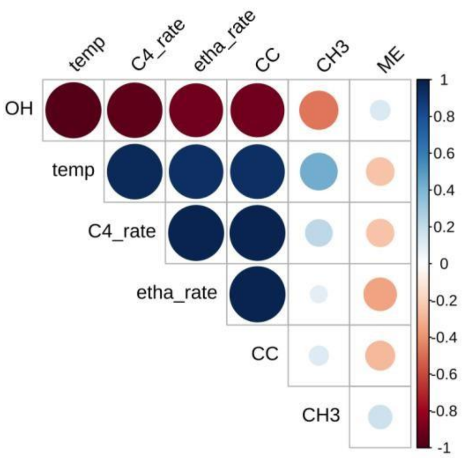

FIGURE I Correlation Analysis Chart for B3 Combined

TABLE I B3 Combined Correlation Coefficient Table

	OH	temp	C4_ rate	Etha _rate	CC	CH3	ME
OH	1	-0.99	-0.96	-0.91	-0.92	-0.47	0.13
temp	-0.99	1	0.97	0.93	0.94	0.44	-0.25
C4_ rate	-0.96	0.97	1	0.99	0.98	0.23	-0.24
Etha _rate	-0.91	0.93	0.99	1	0.99	0.09	-0.34
CC	-0.92	0.94	0.98	0.99	1	0.12	-0.27
CH3	-0.47	0.44	0.23	0.09	0.12	1	0.18
ME	0.13	-0.25	-0.24	-0.34	-0.27	0.18	1

2) *B3* combination

The analysis shows that under the action of the B3 catalyst combination, the temperature is significantly positively correlated with the C4 olefin selectivity and the ethanol conversion rate, and the C4 olefin selectivity is significantly positively correlated with the ethanol conversion rate. In addition, the temperature was also negatively correlated with selectivity to fatty alcohols with carbon numbers 4-12 and positively correlated with selectivity for ethylene, methyl benzaldehyde, and methyl benzyl alcohol.

Among the by-products, fatty alcohols with carbon numbers 4-12 were significantly negatively correlated with ethanol conversion, C4 olefin selectivity, and ethylene selectivity. Methyl benzaldehyde and methyl benzyl alcohol have a specific positive correlation with ethanol conversion, C4 olefin selectivity, and ethylene selectivity. Ethylene selectivity is positively correlated with C4 olefin selectivity and ethanol conversion. On the opposite, Acetaldehyde selectivity was not significantly correlated with other variables. The correlation coefficient and P-value are shown in TABLE II and FIGURE II.

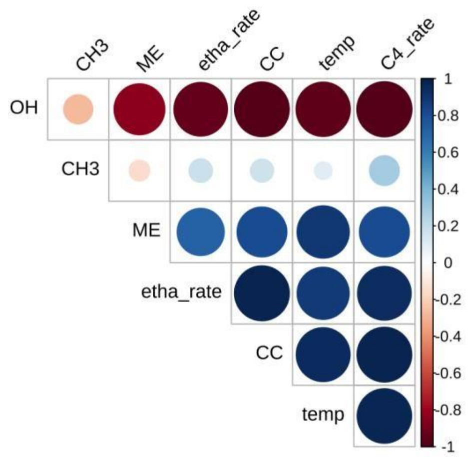

FIGURE II Correlation Analysis Chart for A5 Combined

TABLE II A5 Combined Correlation Coefficient Table

	OH	**CH3**	**ME**	**Etha _rate**	**CC**	**temp**	**C4_ rate**
OH	1	−0.28	0.85	−0.94	−0.98	−0.97	-1
CH3	−0.28	1	0.14	0.18	0.18	0.1	0.29
ME	−0.85	−0.14	1	0.73	0.81	0.89	0.82
Etha _rate	−0.94	0.18	0.73	1	0.99	0.89	0.94
CC	−0.98	0.18	0.81	0.99	1	0.95	0.98
temp	0.97	0.1	0.89	0.89	0.95	1	0.97
C4_ rate	-1	0.29	0.82	0.94	0.98	0.97	1

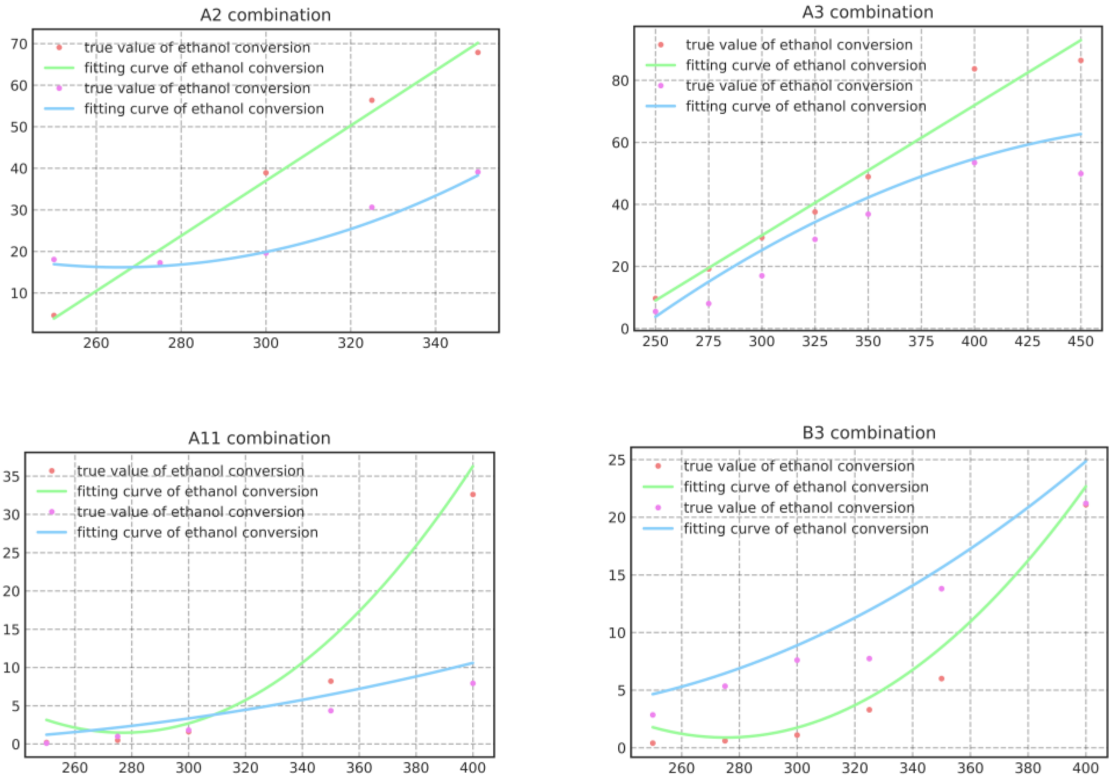

FIGURE III 18 sets of ethanol conversion and C4 olefin selectivity-temperature graphs

2.2. Evaluation of Effects

After calculating the correlation, we roughly estimated the difference in the effect of each catalyst combination and gave a new evaluation system.

According to the 21 groups of ethanol conversion and C4 olefin selectivity-temperature relation, it is doubtless to observe that most of the ethanol conversion and C4 olefin selectivity are positively correlated with temperature. There are exceptions: A2, A4, A5, A6, A10, B4 first decreased and then increased, A1, A3 increased first and then decreased, but the overall trend is still in line with the upward trend. The rate of increase in ethanol conversion and C4 olefin selectivity with increasing temperature varied among groups.

In order to quantitatively compare the differences between different catalyst combinations, four key index values, A_{250}, B_{250}, Ka, Kb, were selected. Based on the principle of control variables, the temperature is controlled the same, and the ethanol conversion rate A_{250} and the C4 olefin selectivity

B_{250} at 250 °C are picked out for each group. Defining the slope of the line connecting the maximum and minimum influence ratios, representing ethanol conversion velocity changes with temperature and C4 olefin selectivity changes with temperature, and can be expressed as:

$$Ka = \frac{A\max - A\min}{TA\max - TA\min} \tag{1}$$

$$Kb = \frac{B\max - B\min}{TB\max - TB\min} \tag{2}$$

TABLE III A250, B250, Ka, Kb are used for the value classification of N and M

	A_{250}	Ka	B_{250}	Kb
N	2.49	0.33	5.19	0.21
M	7.46	0.37	13.11	0.19

After calculating and organizing the A250, B250, Ka, and Kb of each group, it is used to make the difference between the experiments of each group more comprehensible. It classified the data and calculated the median N and the self-defined values M(average values of median, maximum and minimum values). The results are shown in TABLE III. According to the actual meaning of A250, B250, Ka, Kb, it is known that the experimental group with all four values that are larger has achieved good ethanol conversion. C4 olefin selectivity is 250 ℃, and the ethanol conversion and C4 olefin selectivity performance increase significantly with increasing temperature. That is, the experimental group has a good catalytic effect. If the 4 indicators have the same weight, a simple scoring method can be established to grade the data (those between N and M are level 1, those above N and M but not the highest are level 2, and the highest are level 3). It is found that the higher the level, the higher the score obtained. A rough estimation of the catalytic effect according to the total score of the levels is shown that the best catalyst combinations are A2 and A3, the catalyst combinations of A4, A5, A6, and A7 are excellent, the catalyst combinations of A1, A8 and B7 are good, and the catalyst combinations of the other experimental groups are in normal.

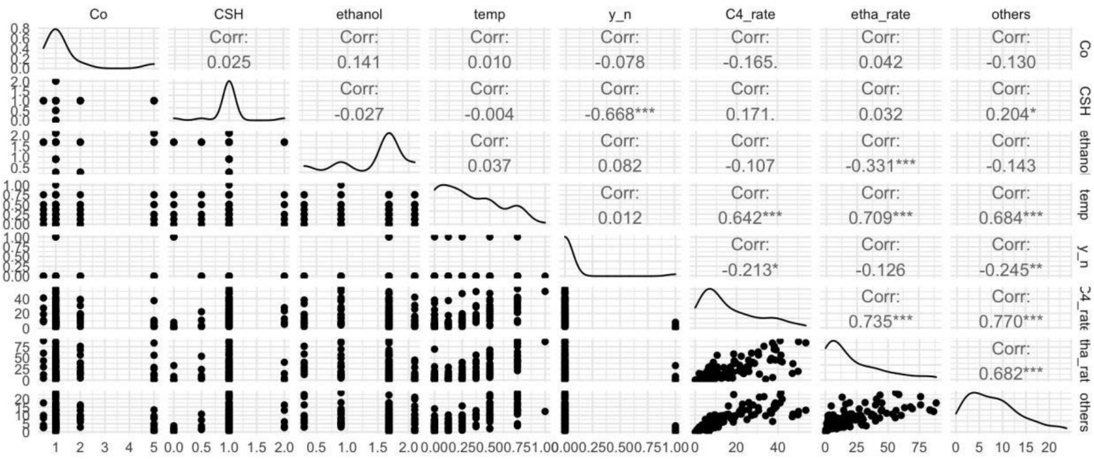

FIGURE IV Correlation analysis results

2.3. Visualization

Through the R language, **we fitted 21 curves sets of ethanol conversion rate and C4 olefin selectivity with temperature.** We selected the four most representative data groups: A2 has the fastest ethanol conversion rate with temperature, A3 has the fastest C4 olefin selectivity with temperature, and A11 and B3 have the slower ethanol conversion rate C4 olefin selectivity with temperature. The four sets of data are shown in FIGURE VI.

3. Stepwise Regression

3.1. Data processing

Twenty-one catalyst combinations from A1 to B7 were re-stated, considering Co loading, Co/SiO2-HAP mass ratio, and ethanol concentration. According to the chemical principle, the amount of catalyst is positively correlated with the catalytic effect, and the amount of Co/SiO2 and HAP has a clear relationship with ethanol conversion and C4 olefin selectivity, so it is no longer considered a factor. According to statistics, among the 21 combinations of questions, the Co loading levels are 0.5wt%, 1wt%, 2wt%, and 5wt%; the Co/SiO2-HAP mass ratios are 1:1, 2:1, and 1:2. There are four levels of ethanol concentration: 1.68ml/min, 0.9ml/min, 0.3ml/min and 2.1ml/min. The selectivity of ethylene, acetaldehyde, methyl benzaldehyde, C4~C12 aliphatic alcohol, and other products was studied, and their values were summed, defined as the by-product selectivity of the reaction. Descriptive statistics were performed on the preprocessed data, and TABLE IV was obtained.

TABLE IV Descriptive Statistics

	Co	**CSH**	**ethanol**	**temp**	**y_n**	**C4_ rate**	**Etha _rate**	**others**
Min Value	0.500	0.000	0.300	0.000	0.000	0.100	0.200	0.000
Middle Value	1.000	1.000	1.680	0.250	0.000	10.775	13.300	7.425
Mean Value	1.421	0.978	1.483	0.310	0.044	16.451	21.985	8.115
Max Value	5.000	2.000	2.100	1.000	1.000	53.430	88.400	23.710

After analyzing the correlation between variables, the results are shown in FIGURE VI. It can be seen from Figure 8 that the ethanol conversion rate, C4 olefin selectivity, and selectivity of other products are significantly related to temperature, and the correlation coefficients (Pearson correlation coefficient) are 0.709, 0.642, and 0.684, respectively. There is also a strong correlation among them. The correlation coefficient between whether to use quartz sand and the mass ratio of Co/SiO2-HAP reaches 0.668 because of the HAP mass in the A11 combination. It is used quartz sand is 0, and the correlation between whether to use quartz sand and the mass ratio of Co/SiO2-HAP is sexually significant.

3.2. Stepwise regression method

Stepwise regression is to gradually input independent variables into the model and then test whether the model has statistical significance. Finally, a regression model with the best fitting effect is obtained. It is essentially a linear regression.

Introduce all explanatory variables, perform significance analysis, filter out CSH and Ethanol's insignificant explanatory variables, and get the results in TABLE V.

TABLE V Results after screening out insignificant variables (ethanol conversion rate)

	Estimate	**Std.Error**	**t value**	**Pr(>|t|)**
Intercept	10.0599	1.8109	5.555	1.95e–07 ***
Co	–2.2140	0.7911	–2.799	0.006064 **
temp	32.9261	3.4466	9.553	4.14e–16 ***
y_n	–15.5069	4.4652	–3.473	0.000737 ***

Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Residual standard error: 9.732 on 110 degrees of freedom12
Multiple R-squared: 0.4964, Adjusted R-squared: 0.4827
statistic: 36.15 on 3 and 110 DF, p-value: 2.458e-16

A regression equation can be obtained:

$$etha\_rate = -2.2140*Co + 32.9261*temp - 15.5069*y\_n + 10.0599 \quad (3)$$

There is no doubt that the loading of Co and whether quartz sand has a significant negative correlation with the ethanol conversion, the temperature has a significant positive correlation with the ethanol conversion, and the ratio of Co/SiO2-HAP has a more negligible effect on it. Positive correlation, ethanol concentration produces a more negligible negative correlation with it.

Introduce all explanatory variables, conduct significant analysis, filter out the variables with little significance, and get the results in TABLE VI.

TABLE VI Results after screening out insignificant variables (C4 olefin selectivity)

	Estimate	**Std.Error**	**t value**	**Pr(>|t|)**
Intercept	26.053	4.183	6.229	8.81e–09 ***
ethanol	–15.408	2.525	–6.103	1.59e–08 ***
temp	62.169	4.903	12.679	<2e-16 ***
y_n	–11.752	6.350	–1.851	0.0669

Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Residual standard error: 13.84 on 110 degrees of freedom
Multiple R-squared: 0.6423, Adjusted R-squared: 0.6325
F-statistic: 65.84 on 3 and 110 DF, p-value: < 2.2e-16

Intercept, Ethanol and temp can be significantly different from 0 at the level of P=0.05 and pass the significance test. The overall equation passed the significance test and got the regression equation:

$$C4\_rate = -15.408*ethanol + 62.169*temp - 11.752*y\_n + 26.053 \quad (4)$$

Ethanol concentration produces a significant negative correlation with C4 olefin selectivity, while temperature produces a significant positive correlation with C4 selectivity.

4. Conclusion

This paper correlates variables such as temperature, C4 olefin selectivity, and ethanol conversion and draws reliable conclusions. In addition, we also innovatively established a simple classification and scoring method to evaluate the differences in the catalytic effects of different experimental groups.A2, A3, A4, A5, A6, and A7 were estimated to be the catalyst combination with better catalytic effect by using four indicators. It is worth mentioning that we considered the correlation of each by-product with other variables separately. These data can be used to judge further which by-products may affect the C4 olefin selectivity and ethanol conversion of the reaction. After that, we fitted the curves of C4 olefin selectivity and ethanol conversion with temperature, and the fitting effect was satisfactory. This proves that in the chemical process of preparation of olefins from ethanol, ethanol concentration produces a significant negative correlation with C4 olefin selectivity, while temperature produces a significant positive correlation with C4 olefin selectivity. Our research has certain reference significance for improving the efficiency of producing C4 olefins by ethanol coupling in the future.

References

[1] Wang, S., et al., Design and research on preparation of C4 olefins by ethanol coupling based on logistic. 2022. 8: p. 370-376.

[2] Zacharopoulou, V. and A.A.J.C. Lemonidou, Olefins from biomass intermediates: a review. 2017. 8(1): p. 2.

[3] Gao, X., et al., Research on Preparation of C4 Olefins by Ethanol Coupling.

[4] Zhang, S.-M., et al., Research on Ethanol Coupling to Prepare C4 Olefins Based on BP Neural Network and Cluster Analysis. 2022. 2022.

[5] Alpaydin, E., Machine learning. 2021: MIT Press.

[6] Thearling, K.J.D.M.M., An introduction to data mining. 1999: p. 28-31.

[7] Keka, I. and B. Çiço. Data Visualization as Helping Technique for Data Analysis, Trend Detection and Correlation of Variables Using R Programming Language. in 2019 8th Mediterranean Conference on Embedded Computing (MECO). 2019. IEEE.