

Research on preparation process of ethanol and C4 olefin based on linear and nonlinear regression fitting algorithm

Ziyuan Zhao*, Shougan Ye

Beijing University of Chemical Technology, Beijing, 100029

*Corresponding author

Keywords: Linear regression, Nonlinear regression, Fitting algorithm

Abstract: First of all, aiming at "studying the relationship between ethanol conversion, C4 olefin selectivity and temperature for each catalyst combination", using the knowledge of fitting arbitrary data and adopting the basic idea of control variable method, linear regression equation and nonlinear regression equation model are established. Aiming at the solution of linear regression and non-linear regression model, this paper uses the fitting algorithm to calculate the functional relationship between the two variables, and uses Matlab to solve the specific function expression and draw the function relation diagram to further solve the relationship between catalyst combination, temperature and ethanol conversion and C4 olefin selectivity.

1. Introduction

With the development of society and the changes of the times, the production of chemical products and pharmaceuticals depends more and more on the wide application of C4 olefins. Ethanol is the raw material for the production of C4 olefins [1]. Therefore, it is of great significance and value to explore the process conditions for the preparation of C4 olefins by catalytic coupling of ethanol [2]. Therefore, the effects of different catalyst combinations and temperature on the selectivity and yield of C4 olefins were studied in this paper [3].

2. Establishment of regression equation

Person correlation coefficient should satisfy continuous data, normal distribution and linear relationship. If these conditions are met, person correlation coefficient should be used first, otherwise Spearman correlation coefficient will be used. Spearman can be used for non-linear, non-normal distribution and so on. Because the data in this question is less and does not necessarily satisfy the person correlation coefficient, so the spearman test is used.

In order to describe the relationship between ethanol conversion, selectivity of C4 olefins and temperature, the internal relationship is explored and solved by regression equation. Thus, the trend of ethanol conversion and C4 olefin selectivity changing with temperature under the same catalyst combination was predicted.

Here, the Spearman coefficient is selected for test.

$$r = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} \quad (1)$$

Calculate the spearman correlation coefficient with corr:

$$\text{corr}(X, Y, \text{'type'}, \text{'Spearman'}) \quad (2)$$

$$\text{corr}(X, \text{'type'}, \text{'Spearman'}) \quad (3)$$

$$\rho = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}} \quad (4)$$

Then carry on the stepwise regression analysis.

$$Y = -3616.928 + X_1 \times 12.895 + X_2 \times 4.64 + X_3 \times (-72.422) + X_4 \times (-137.919) \quad (5)$$

3. Model solving

In this paper, the above formulas are modeled and analyzed by Matlab and SPSS software.

(1) Relationship between temperature and ethanol conversion

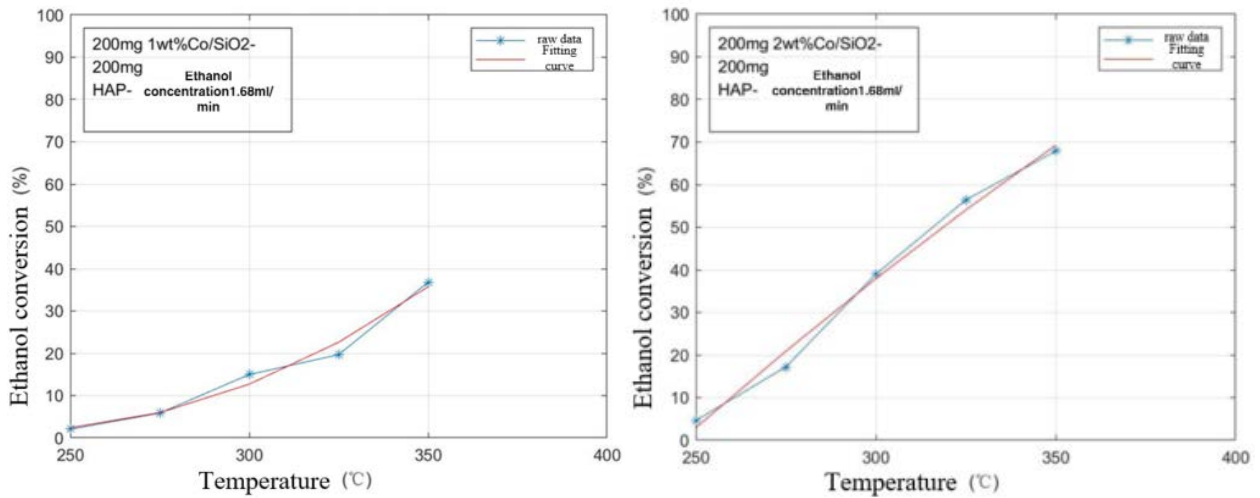


Figure 1: Relationship between temperature and ethanol conversion

(2) Selectivity of C4 olefins

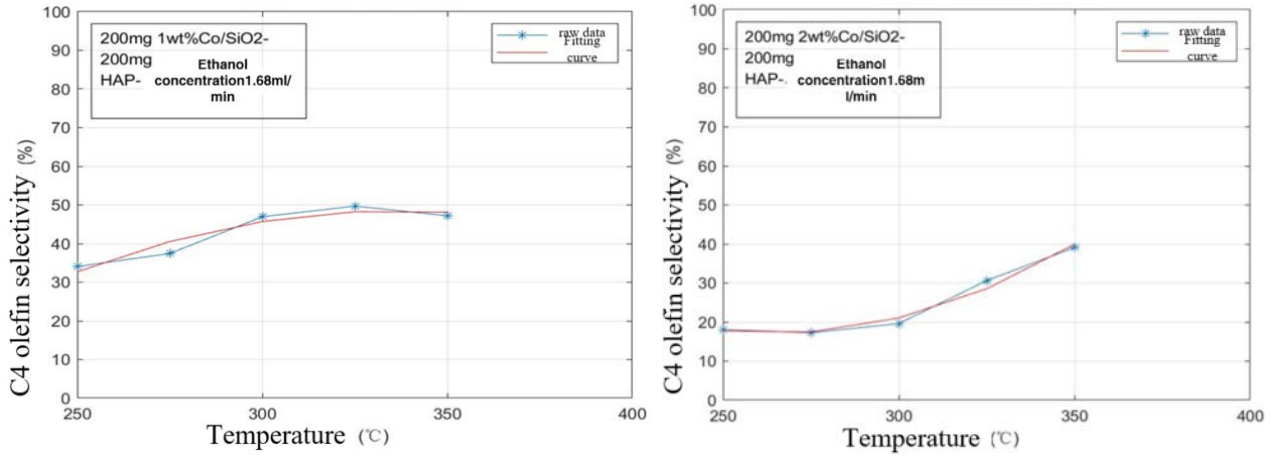


Figure 2: C4 olefin selectivity

4. Model checking

Mean square error (MSE):

$$MSE = \frac{1}{m} \sum_{i=1}^m \left(y_i - \hat{y}_i \right)^2 \quad (6)$$

Mean absolute error (MAE): $\frac{1}{m} \sum_{i=1}^m \left| y_i - \hat{y}_i \right|$

Sum of squares of errors (SSE):

$$SSE = \sum (Y_{actual} - Y_{predict})^2 \quad (7)$$

In the case of the same data set, the smaller the SSE is, the smaller the error is, and the better the model is.

Determination coefficient (R-square):

$$R^2 = 1 - \frac{(Y_{actual} - Y_{predict})^2}{(Y_{actual} - Y_{mean})^2} \quad (8)$$

Theoretically take the value range $(-\infty, 1)$, the normal value range is $[0, 1]$ -in practice, we usually choose a well-fitted curve to calculate R^2 , so the occurrence of $-\infty$ is rare. The closer it is to 1, the stronger the explanatory ability of the variables of the equation to y , and the better the model fits the data. The closer to 0, the worse the model fitting. Empirical value: > 0.4 , the fitting effect is good.

5. Conclusion

First of all, according to different catalyst combinations, this paper studies the relationship between ethanol conversion, selectivity of C4 olefins and temperature respectively. For the solution of linear regression and nonlinear regression model, this paper uses the fitting algorithm to calculate the function relationship between the two variables, and uses Matlab to solve the specific function relation expression and draw the function relation diagram. Furthermore, the relationship between catalyst combination, temperature and ethanol conversion and C4 olefin selectivity was solved. Using the basic idea of the control variable method, the linear regression equation and the nonlinear regression equation model are established. The model can vividly express the specific function

expression and function image between the two variables, through the function expression can roughly predict: when the catalyst combination is constant, the ethanol conversion and C4 olefin selectivity corresponding to unknown temperature. The changing trend of the relationship between the two variables can be vividly expressed by the function image. It can not only be used to explore the process of catalytic coupling of ethanol to prepare C4 olefins. It can also be used in all reaction processes that require catalyst catalysis in a wider range of chemical production fields. Its modeling and problem-solving ideas are more or less the same as this problem.

References

- [1] Zar, J. H. (1972). *Significance testing of the Spearman rank correlation coefficient*. *Journal of the American Statistical Association*, 67, 578. Re- printed with permission from the *Journal of the American Statistical Association*. Copyright © 1972 by the American Statistical Association. All rights reserved.
- [2] https://blog.csdn.net/qq_40783326/article/details/103137788?utm_medium=distribute.pc_relevant.none-task-blog-2%7Edefault%7EBlogCommendFromBaidu%7Edefault-7.no_search_link&depth_1-utm_source=distribute.pc_relevant.none-task-blog-2%7Edefault%7EBlogCommendFromBaidu%7Edefault-7.no_search_link
- [3] <https://www.iikx.com/news/statistics/459.html>.