Study on the reaction conditions for preparing C4 olefins from ethanol coupling based on multi-factor optimization

Zhong Fang[1] Yufeng Wang[1] Yunchang Zhu[1]

[1]Wuhan University

Brief Introduction

China Undergraduate Mathematical Contest in Modeling (CUMCM), the highest level MCM in China. The contest requires us to submit a technical paper including the model establishment, simulation based on the given data set within 3 days. This poster is made since the original paper is overall 40 pages, where some sections are discussing details related to the specific problem itself as opposed to techniques.
Idea goes: The chemical background of the problem is unfamiliar to us at first glance. It is troublesome to figure out the correlation in these intricate reactions, and only rely on experiment data is obviously not a clever choice. We tried to get first intuition by illustrating the sketch figure with reactants and products according to the experimental data. The oscillation curves depicting intermediates acetaldehyde and C4-12 aliphatic alcohols aroused our interest. It took us a long time to find an oldish interdisciplinary book focusing on chemical multi factors reaction process after retrieving many chemical references trying to seize any potential ODEs analysts. Finally we clarified each reactant and product goes by testing possible mutual reactions.

Problem Statement

In the organic chemical industry, it is of great significance to use relatively cheap C4 olefins to produce low-carbon olefins and other chemical and pharmaceutical products rather than the traditional steam cracking method [Lv 2018]. Adjusting the selectivity and yield of C4 olefins by controlling the reaction conditions is the key to chemical production. It is valuable to establish models through data and design different combinations of reaction conditions to improve C4 olefins yiled (multiplication of ethanol conversion rate and C4 olefins selectivity) and explore its reaction mechanism.
A laboratory conducted a series of experiments at different temperatures for different catalyst combinations, Results are recorded in the attachment. The following questions are raised.

1. For each catalyst combination in Attachment 1, study the relationship between ethanol conversion rate and C4 olefins selectivity versus temperature respectively. Analyse the test results of the catalyst combination given at 350 degrees in Attachment 2 at different times in an experiment.
2. Explore the effects of different catalyst combinations and temperatures on ethanol conversion rate and C4 olefins selectivity.
3. How to select catalyst combinations and temperatures to make the C4 olefins yield as high as possible under the same experimental conditions, what if temperature is lower than 350 degrees?
4. Design 5 more experiments and give detailed reasons. - not math related, omitted

Analysis of Problem 1

Since data is from a lab report, there is a few noise need to be filtered first, data preprocessing is essential according to the variable consistency. It can be seen that the problem is not under a big-data context by our preprocessed data in Fig. 1, thus, the classic fitting method will create a considerable misfitting in this case.

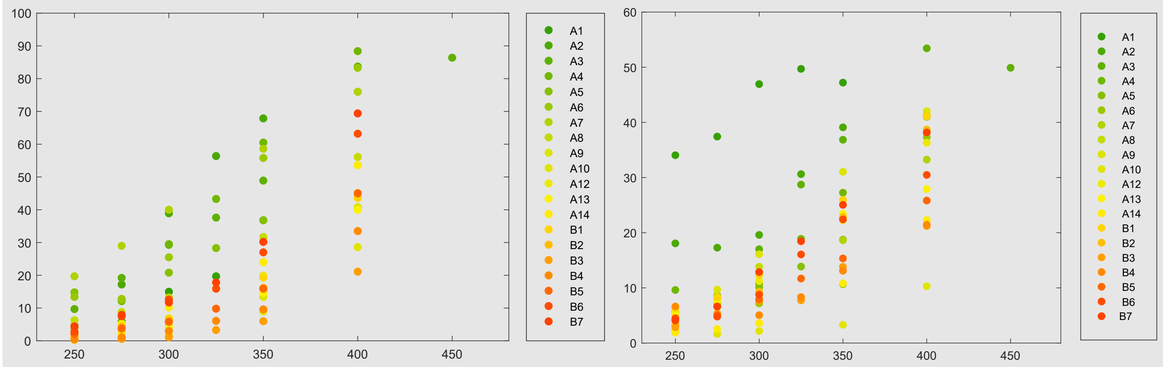

Figure 1. Scatter plot of preprocessed data (left: ethanol conversion versus temperature, right: C4 olefins selectivity versus temperature

We consider them separately. In chemical equilibrium, ethanol is used as a reactant, and its conversion rate is strongly correlated with the direction of equilibrium movement and the degree of reaction. Since the equilibrium constant is only related to temperature, the relationship between ethanol conversion rate and temperature can be explored based on the chemical reaction isotherm.

$$\Delta_r G_m^\Theta = -RT \ln K^\Theta \Leftrightarrow K^\Theta = e^{-\frac{\Delta_r G_m^\Theta}{RT}}$$

where $\Delta_r G_m^\Theta$ is the standard molar reaction Gibbs free energy change, R is the gas constant and K^Θ the standard equilibrium constant. To uniform the unit of temprature with the data, we use $f(t) = \alpha e^{-\frac{\beta}{t+273}}$ as fitting function.
Current organic chemistry research has not yet clarified the quantitative relationship between temperature and reactant selectivity, so we will consider the establishment and analysis of the model from a statistical perspective. Gaussian process regression is considered by the widely use in chemistry, especially computational chemistry. $f_i(t) = \alpha e^{-(\frac{t_i-\beta}{\gamma})^2}$ is used for above 20 data sets. To simplify computation and decrease error, we enforce logarithm for both sides, and a linear equation is obtained by redefine variables. Implement Moore–Penrose inverse we can easily solve α, β, γ.
The other one takes more work. By analyzing the relative ratio of each reactant in Attachment 2, combined with the organic chemical reaction mechanism, it can be seen that there is an obvious oscillation phenomenon, and the reaction mechanism conforms to the nonlinear differential equation model of multimolecular reaction. The speculation about the possible pathways for the formation of C4 olefins is omitted here. Multimolecular biochemical reaction model [Wang 2008] is established reasonably by the analysis. The reaction can be represented as

$$A \xrightarrow{k_1} X,\ nX + mY \xrightarrow{k_2} (m+n)Y,\ Y \xrightarrow{k_3} P$$

where the initial substance A represents ethanol, and its concentration is $[A]_0$. P is the final product, and the intermediates X and Y represent acetaldehyde and carbon number 4-12 fatty alcohols, respectively. Their concentrations are $[X] = x(\tau)$, $[Y] = y(\tau)$, τ is the time, and k_i represents the rate constant of the reaction in the i-th step. Assume that the concentration change of A and the reversibility of the chemical reaction are ignored, the concentrations of X and Y have oscillatory behavior. The differential reaction rate equations describing the concentration change with time τ are further given:

$$\begin{cases} \dfrac{\mathrm{d}[X]}{\mathrm{d}\tau} = k_1[A]_0 - k_2[X]^n[Y]^m \\ \dfrac{\mathrm{d}[Y]}{\mathrm{d}\tau} = k_2[X]^n[Y]^m - k_3[Y]_0, \quad m, n \in \mathbb{N} \end{cases}$$

To obtain the dimensionless form of the above equations, consider the following transformation

$$\begin{cases} [X] \to x, [Y] \to y, \tau k_2 \to \tau \\ \alpha = \dfrac{[A]_0 k_1}{k_2}, k_2 = k_3 \end{cases} \Rightarrow \begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}\tau} = \alpha - x^n y^m \\ \dfrac{\mathrm{d}y}{\mathrm{d}\tau} = y(x^n y^{m-1} - 1) \end{cases} \quad (1)$$

Algorithm and Results for Problem 1

Fitting parameters can be obtained by using `Curve Fitting` toolbox in MATLAB. R^2 is an important evaluation criteria for fitting performance.

Symbol	Description	Unit
t	Celsius temperature	$^\circ$C
T	Kelvin temperature	K
τ	Time	s
A	Initial material (ethanol)	ml/min
η	Ethanol conversion	\
ξ	C4 olefins selectivity	\

Table 1. List of notation

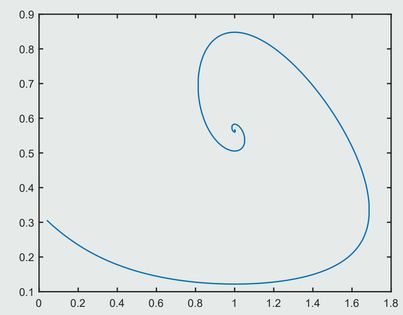

Figure 2. Oscillation trajectory of relative content of Y v.s. X

In this paper, relative content is used instead of concentration when establishing the model (Note: due to the characteristics of organic chemical cracking reactions, the relative content can be greater than 1). Based on model (1), a group of differential equations is established and the data in Attachment 2 is implemented as (x, y) in the model, that is, acetaldehyde and fatty alcohols with carbon numbers of 4-12. According to the given reaction data, the coefficient α can be solved, trajectory of (x, y) is illustrated in Fig. 2, from the starting point $(0.0398, 0.3060)$, and the oscillation trajectory is asymptotically stable at the point $(1, 0.565)$. Therefore, it can be considered that data follows the Lyapunov asymptotic stability, which proves that the catalyst stability is good at this time scale.

Potential extension

I focused on some details about chemistry reaction process in the paper, where I didn't analyse about equilibrium point heavily. It's also interesting to explore it if in a different context.

Analysis of Problem 2

- The **Grey Relational Analysis** (GRA) model [Si and Sun 2015] is considered for manifesting the impact from each factors.
  - As a multi-factor statistical analysis method, determines whether the relationship is close based on the similarity of the sequence curve shape, thereby establishing an evaluation model between different chemical reaction independent variable factors and dependent variables.
  - After determining the evaluation objects and evaluation indicators, the weight for each indicator ω_k is given by **entropy weight method** which based on data as opposed to subjective choice.
  - Grey correlation coefficient $\zeta_i(k)$ can be computed by its definition, and then size the grey weighted correlation $r_i = \sum_{k=1}^n \omega_k \zeta_i(k)$, the evaluation objects are sorted to establish the correlation sequence of them.
- **Control Variable** is a crucial idea about analysing the influence of different factors due to heavily coupled experimental data.

Analysis of Problem 3

There are 6 chemical reaction factors influencing the C4 olefins yield, which is reasonably problematic to find independent functions to performance. Multi-factor model combined with Backpropagation Neural Network(BPNN) is proposed for solving the optimization problem. BPNN is a trade-off decision in this situation.
Let 6 factors be input, and 2 output is ethanol conversion rate output$_1$ and C4 olefins selectivity output$_2$. Our objective is to maximize their multiplication, i.e., output$_1$ $\cdot$ output$_2$ Algorithm is designed as follows:

1. Write a Matlab program based on model 3 to generate a prediction data set.
2. Initially set a larger step size, traverse the values of the six independent variables and find a rough solution (local optimal solution) through the BPNN model objective function.
3. Further narrow the search accuracy and find the optimal solution near the rough solution of step 3. Repeat step 3 until the optimal solution is found.
4. Backtrack the data set according to the maximum index of the objective function to find the values of the relevant independent variables. The results show that when each variable meets the data in the table below, the C4 olefins yield is the highest, and the yield is 58.99%.

References

Lv, S. (2018). "Coupling of ethanol to butanol and C4 olefins". MA thesis. Dalian University of Technology.

Si, S. and X. Sun (2015). *Mathematical modeling algorithms and applications*. Beijing: Industrial Press.

Wang, S. (2008). *Selected Lectures on Mathematical Models*. Science Press.

Summary

Competition	Level	Involvement(teams)	Proportion
CUMCM	National 1st	45,075	0.65%
HZMCM	Regional 1st Innovative paper	1,927	4.7% 3(1.6‰)
EEMCM	National 2nd	6,389	15%

Table 2. Corresponding competition data