

# 乙醇制备C4烯烃收率的优化建模

朱瑞依, 张 雯, 邵千一

吉林大学通信工程学院, 吉林 长春

收稿日期: 2022年1月11日; 录用日期: 2022年2月10日; 发布日期: 2022年2月16日

## 摘 要

C4烯烃是重要的基础化工原料, 以乙醇为原料生产C4烯烃具有良好的应用前景及经济效益。在乙醇偶合制备C4烯烃的过程中, 不同的催化剂组合和温度将对C4烯烃的选择性和乙醇转化率产生重大影响, 进一步影响C4烯烃收率。为在减少实验量的基础上探寻多因素下使C4烯烃收率最大的反应条件, 本文采用Lasso回归将不同催化剂条件量化表示为多个变量, 联合温度条件, 分别对乙醇转化率和C4烯烃选择性进行回归, 筛选对二者影响显著的变量, 分别得到回归方程式; 利用回归方程得到C4烯烃收率表达式, 并作为目标函数, 采用模拟退火算法, 解得最优化反应条件; 并在现有实验数据及分析的基础上, 设计探究更优反应条件的正交试验。

## 关键词

C4烯烃收率, Lasso回归, 模拟退火, 正交试验

# Optimization Modeling of Yield of C4 Olefins with Ethanol

Ruiyi Zhu, Wen Zhang, Qianyi Shao

College of Communication Engineering, Jilin University, Changchun Jilin

Received: Jan. 11<sup>th</sup>, 2022; accepted: Feb. 10<sup>th</sup>, 2022; published: Feb. 16<sup>th</sup>, 2022

## Abstract

C4 olefin is a fundamental chemical raw material. The production of C4 olefin with ethanol as raw material enjoys good application prospect and economic benefits. In the process of producing C4 olefins by ethanol coupling, both various catalyst combinations and temperatures will have a significant impact on C4 olefin selectivity and ethanol conversion, further influencing yield of C4 olefins. In order to explore the reaction conditions that maximize the yield of C4 olefins under multiple factors on the basis of reducing experiments, this paper uses Lasso regression to regress the ethanol conversion rate and C4 olefin selectivity respectively by quantifying different catalyst condi-

tions as multiple variables and combining with temperatures, and then the regression equations are obtained respectively through screening the variables that influence ethanol conversion rate and C4 olefin selectivity significantly. Furthermore, the yield expression of C4 olefin is derived from the regression equations, which is taken as the objective function. Consequently, the simulated annealing algorithm is adopted to obtain the optimal reaction condition, and based on existing experimental data and analysis, an orthogonal test for exploring the optimal reaction condition is designed.

## Keywords

Yield of C4 Olefins, Lasso Regression, Simulated Annealing Algorithm, Orthogonal Test

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

C4 烯烃是重要的基础化工原料, 广泛应用于生产医药和化工产品。传统的生产方法以化石能源为原材料, 由碳四馏分分离获得。鉴于化石能源存储量短缺以及对环境的污染, 利用新型清洁能源制备 C4 烯烃越来越重要。乙醇来源广泛, 绿色清洁, 以乙醇为平台分子生产高附加值的 C4 烯烃具有良好的应用前景及经济效益。

在乙醇偶合制备 C4 烯烃的过程中, 不同的温度和催化剂使用将对乙醇转化率和 C4 烯烃的选择性产生重大影响, 而此二者的乘积便是 C4 烯烃的收率。因此通过设计最佳的温度和催化剂组合, 探索乙醇催化偶合制备 C4 烯烃的条件对化工生产具有重要作用。

为探寻多因素下最优反应条件, 需结合反应原理进行大量、多组实验数据。童刘[1]控制反应温度和乙醇浓度, 得到了乙醇制备 C4 烯烃的优化条件; 吕绍沛[2]采用 Co-SiO<sub>2</sub>-HAP 作为催化剂, 通过综合控制温度和催化剂装料方式、装料比、Co 负载量等反应条件, 实现了对乙醇制备 C4 烯烃反应的优化。综上, 反应温度、乙醇浓度、装料方式、Co/SiO<sub>2</sub> 和 HAP 装料比、Co 负载量等自变量的设置都对反应有着重要作用。

利用实验数据进行建模求解, 可以较少的实验结果获得尽可能多的信息量, 获得优化后的反应条件。本文以 2021 年全国大学生数学建模竞赛 B 题[3]所给数据为例进行了分析。为探究温度和催化剂多因素的综合影响, 可作定性和定量分析, 本文采用了 Lasso 回归剔除影响不显著的变量, 分别得到多因素与乙醇转化率和 C4 烯烃选择性的定性关系, 而后重新回归得到定量关系式; 所得定量公式可以用于寻找收率的最大值, 得到优化条件, 在自变量多达 7 个的情况下, 寻优计算量大, 使用模拟退火算法可避免“维数灾难”; 由于所得关系式只对现有数据的取值范围具有普适性, 因而在此基础上做更优反应条件的预测还需设计实验验证和探究, 多因素下设计正交试验做进一步探究可提高效率。

## 2. 分析催化剂与温度对反应影响

为研究不同催化剂组合及温度对乙醇转化率、C4 烯烃选择性的影响, 将催化剂组合拆分成 HAP 质量、乙醇浓度、是否加入石英砂(0 代表不加入石英砂, 1 代表加入石英砂)、Co 负载量、I II 类总量(Co/SiO<sub>2</sub> 和 HAP 总质量)、Co/SiO<sub>2</sub> 和 HAP 装料比、装料方式(0 表示装料方式 II, 1 表示装料方式 I)这些变量。将这些变量联合温度分别对乙醇转化率和 C4 烯烃选择性进行回归, 变量的数量较多, 采用 Lasso 回归。

下文公式中所用符号对应含义见表 1。

**Table 1.** Symbol description  
**表 1.** 符号说明

符号	说明	单位
$T$	反应温度	°C
$\alpha$	单位时间内乙醇的单程转化率	\
$S$	C4 烯烃选择性即在所有产物中的占比	\
$Y$	C4 烯烃收率即乙醇转化率与 C4 烯烃选择性的乘积	\
$P_I$	Co 负载量即 Co 与 SiO <sub>2</sub> 重量比	\
$C_{III}$	反应容器内乙醇浓度	ml/min
$P_{I II}$	Co/SiO <sub>2</sub> 和 HAP 装料比	\
$m_{I II}$	Co 与 SiO <sub>2</sub> 的重量之和	mg
$m_H$	HAP 质量	mg
$K_I$	I 类催化剂种类(Co/SiO <sub>2</sub> )	\
$K_{II}$	是否加入石英砂(0~1 变量)	\
$V$	反应容器的体积	L
$L$	装料方式(0~1 变量)	\

Lasso 回归是一种压缩估计，在最小二乘法的基础上，增加惩罚项以压缩部分回归系数。惩罚项可用于筛选变量，有选择性的把变量纳入模型，起简化模型作用。Lasso 回归还能够通过改变参数控制模型复杂度，避免出现过度拟合。对于线性模型，模型复杂度与变量数之间存在直接关系。具体来说，变量数越多，模型复杂度越高。过多的变量可能导致过度拟合。

Lasso 回归的方程式为：

$$\hat{\beta} = \arg \min \left[ \sum_{i=1}^n (y_i - x_i \hat{\beta})^2 + \lambda \sum_{i=1}^n |\beta_i| \right] \quad (1)$$

Lasso 回归模型通过参数  $\lambda$  调整复杂度， $\lambda$  越大对变量多的模型惩罚力度越大，Lasso 回归模型通过将某些回归系数压缩为零，实现变量的特征选择。Lasso 回归的惩罚项等于各回归系数的绝对值之和乘以参数  $\lambda$ 。

通过最小化 RSS 与惩罚项的和求得变量回归系数。当参数  $\lambda$  等于 0 时，Lasso 回归与最小二乘法没有区别。随着  $\lambda$  逐渐增加，模型开始将回归系数向 0 压缩，当  $\lambda$  足够大时，所有变量的系数将都被压缩到 0。随着参数  $\lambda$  的变化，各变量的回归系数都会跟着改变[4]。

## 2.1. 各因素对乙醇转化率的影响

为探究不同催化剂组合及温度对乙醇转化率的影响，在 Stata 中，以乙醇转化率为因变量，温度以及催化剂组合拆分得出的变量：HAP 质量、乙醇浓度、是否加入石英砂、Co 负载量、I II 类总量、Co/SiO<sub>2</sub> 和 HAP 装料比、装料方式为自变量，归一化后进行 Lasso 回归，设置种子。经 Lasso 回归筛选出有效变量后，得到的较为显著的自变量见表 2。

表 2 是经 Lasso 回归筛选后，保留下的对乙醇转化率影响较大的变量，这种情况下取参数  $\lambda = 127.769$  (lopt)。利用保留下的变量进行多元线性回归(OLS)，得到的结果如最后一列所示。综合分析两种回归，得到一致的结论：温度对乙醇转化率的影响明显高于其他变量。

**Table 2.** Regression results 1**表 2.** 回归结果 1

Selected	Lasso Post-est	OLS
$m_H$	6.8819455	6.9523331
$C_{III}$	-3.9304616	-4.2863523
$K_{II}$	-0.2473751	-0.8484786
$L$	1.5744997	2.1375862
$T$	16.9035327	17.4568015
Partialled-out*		
_cons	21.9892021	21.9892021

为得到上述筛选出每种自变量对乙醇的具体影响情况，我们将每个变量的量纲纳入考虑范围，使用未归一化的数据进行回归，得到表 3：

**Table 3.** Affects results 1**表 3.** 影响结果 1

乙醇转化率	Coef.	Std. Err.	t	P > t	[95% Conf. Interval]
装料方式	4.459328	2.328068	1.92	0.058	-0.15531 9.073963
是否加入石英砂	-4.1251	5.198884	-0.79	0.429	-14.4302 6.179993
乙醇浓度	-8.28019	1.994034	-4.15	0.000	-12.2327 -4.32767
温度	0.333701	0.01871	17.84	0.000	0.296614 0.370788
HAP	0.096099	0.016192	5.93	0.000	0.064003 0.128194
_cons	-82.9088	7.219707	-11.48	0.000	-97.2195 -68.5981

上表第一列显示了自变量每变化一个单位，因变量随变化的程度。如，温度每变化 1℃，乙醇转化率上升约 0.33%。其中，是否加入石英砂对乙醇转化率的影响不显著，应当剔除。其余变量在 90% 的置信概率下，可以认为其对乙醇转化率的影响显著。

重新回归，得到如下表达式：

$$\alpha = 3.903L - 8.325C_{III} + 0.334T + 0.101m_H - 83.99 \quad (2)$$

## 2.2. 各因素对 C4 烯烃选择性的影响

为探究不同催化剂组合及温度对 C4 烯烃选择性大小的影响，以 C4 烯烃转化率为因变量，催化剂组合拆分出的变量联合温度 T 为自变量归一化后进行 Lasso 回归，得到的结果如表 4：

**Table 4.** Regression results 2**表 4.** 回归结果 2

Selected	Lasso Post-est	OLS
$C_{III}$	1.5082040	1.6392717
$K_{II}$	-1.6704803	-1.7402382

Continued

$P_I$	-3.9438316	-4.0725281
$m_{II}$	5.5366024	5.6030215
$P_{II}$	0.5746805	0.6057605
$L$	1.4244204	1.5338241
$T$	9.4273803	9.4938845
Partialled-out*		
_cons	16.4507895	16.4507895

这种情况下, 参数  $\lambda = 14.144$  (lopt)。筛选出 6 个对 C4 烯烃选择性影响显著的自变量。

表 5 显示了考虑量纲后, 6 个自变量对 C4 烯烃选择性的具体影响:

Table 5. Affects results 2

表 5. 影响结果 2

C4 烯烃选择性	Coef.	Std. Err.	t	P > t	[95% Conf. Interval]
温度	0.1809098	0.0153575	11.78	0.000	0.1504654 0.2113542
乙醇浓度	0.8226596	1.643227	0.50	0.618	-2.434846 4.080165
是否加入石英砂	-17.88749	4.370278	-4.09	0.000	-26.55106 -9.223922
Co 负载量	-2.928407	0.7447914	-3.93	0.000	-4.404869 -1.451945
装料比	0.4699531	3.362581	0.14	0.889	-6.195972 7.135878
装料方式	7.439112	1.851654	4.02	0.000	3.768423 11.1098
_cons	-42.21316	6.60067	-6.40	0.000	-55.29822 -29.1281

其中, 乙醇浓度和装料比对 C4 烯烃选择性的影响不显著, 应当剔除或加以改进。其余变量在 95% 的置信概率下, 可以认为其对 C4 烯烃选择性的影响显著。

由于某些自变量间存在约束关系, 如装料比  $P_{II}$  和 HAP 质量  $m_H$  共同影响着 Co 与  $\text{SiO}_2$  的质量之和  $m_{II}$ , 尝试添加相互作用因子  $P_{II} \cdot m_H$ , 同时对乙醇浓度  $C_{III}$  这一参数加以改进, 取对数后再次进行回归, 得到下式。

$$S = 0.181T + 0.339 \ln C_{III} - 9.969K_{II} - 3.471P_I + 0.076P_{II} \cdot m_H + 3.561L - 45.937 \quad (3)$$

### 3. 模拟退火算法预测最佳反应条件

上文得出了不同催化剂组合和温度对乙醇转化率以及 C4 烯烃选择性大小的影响的两个方程, 剔除所得方程中的不显著变量, 再次进行回归分析, 得到新的方程所有变量均显著。将两个方程相乘便得到催化剂组合与温度对 C4 烯烃收率的影响。对得到的方程利用模拟退火算法求解最大值。

确定目标函数为:

$$\min(-Y) = -S \cdot \alpha \quad (4)$$

约束条件:

$$\begin{cases} 250 \leq T \leq 450 \\ 0 \leq m_H \leq 200 \\ 0.3 \leq C_{III} \leq 3 \\ K_{II} \text{ 为 } 0-1 \text{ 变量} \\ 0.5 \leq P_I \leq 5 \\ 0.5 \leq P_{II} \leq 2 \\ L \text{ 为 } 0-1 \text{ 变量} \end{cases} \quad (5)$$

模拟退火算法源于固体退火，遵守 Metropolis 准则，以概率接受新状态。从某一确定初始值开始执行迭代过程，根据目标函数生成新解，以某一概率接受新解，直到求得目标函数的最值[5] [6]。

设函数为  $f(x)$ ，接受新状态的概率为  $p_t$ ， $f(A)$ 为当前解， $f(B)$ 为新解则有：

$$p_t = e^{-|f(B)-f(A)| \times C_t} \quad (6)$$

其中  $C_t$  为关于  $t$  递增的函数。

定义初始温度  $T_0$ ，温度的下降  $T_{t+1} = \alpha^t T_t$ ， $\alpha$  常取 0.95，取

$$C_t = \frac{1}{\alpha^t T_0} \quad (7)$$

得：

$$p_t = e^{-\frac{|f(B)-f(A)|}{\alpha^t T_0}} \quad (8)$$

终止条件通常取为连续若干个新解都没有被接受时终止算法。

模拟退火算法的流程图如图 1 所示：

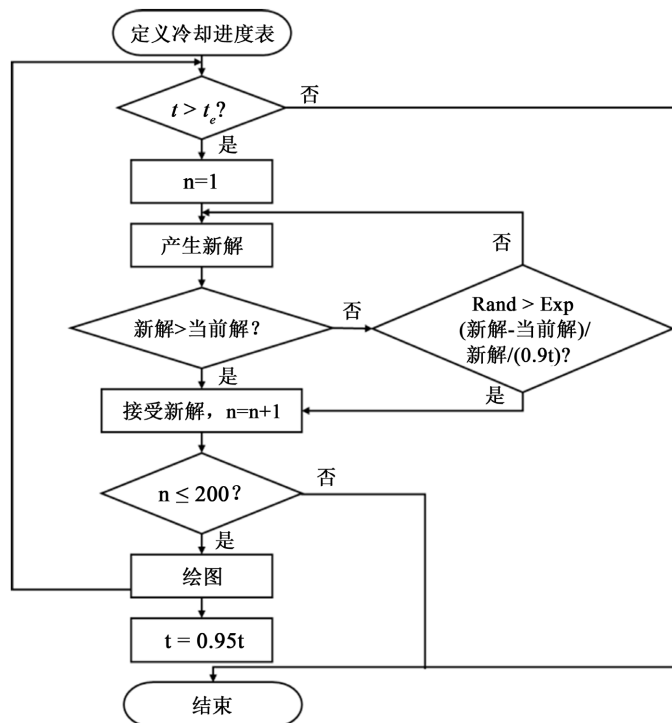


Figure 1. Flowchart of simulated annealing algorithm  
图 1. 模拟退火算法流程图

当求解温度低于 350℃, 使得 C4 烯烃收率尽可能高的催化剂组合与温度时, 将上(9)式中约束条件第一项改为:

$$T \leq 350 \quad (9)$$

其余条件不变。

模拟退火得出预测结果如表 6:

**Table 6.** Predicted results

**表 6.** 预测结果

最优条件	温度/℃	HAP/mg	乙醇浓度/ml/min	Co 负载量/%wt	装料比	装料方式	是否加入石英砂
$T \leq 450^\circ\text{C}$	450	200	0.3	0.5	2	1	0
$T \leq 350^\circ\text{C}$	350	200	1.68	2	1	1	0

上表中分别显示了温度上限分别为 450℃和 350℃时, 使 C4 烯烃选择率最高的温度与催化剂组合预测结果。

#### 4. 在 450℃下设计正交试验

由表 6, 限制 450℃为温度上限时, 最佳温度为 450℃, 但实际数据中 450℃下的实验数据仅有一组, 故将最佳条件与仅有的一组实验的实验条件做对比分析, 发现二者存在三个变量不同, 即乙醇浓度、Co 负载量、装料比(见表 7)。故可以控制温度为 450℃, HAP 质量为 200 mg, 装料方式为 A 类, 针对这三个变量, 设计三因子、两变量的正交试验进一步探究最优解。

**Table 7.** Comparison of experimental data with optimal conditions

**表 7.** 实验数据与最优条件对比

	温度/℃	HAP/mg	乙醇浓度/ml/min	Co 负载量/%wt	装料比	装料方式	是否加入石英砂
预计最大	450	200	0.3	0.5	2	1	0
测量数据	450	200	0.9	1	1	1	0

设计正交试验, 需利用正交表来安排与分析多个试验变量。该方法并不会使用所有变量参数值组合来进行试验, 而是挑选出部分具有代表性的参数组合进行试验, 通过对部分试验结果分析了解全面试验的情况, 从而找出最优的参数组合[7]。通过此种方式设计实验探究多因素对某一变量的影响, 可以减少实验次数, 达到节省时间, 节约材料等效果。试验相应的正交表见表 8:

**Table 8.** Orthogonal table

**表 8.** 正交表

$L_4(2^3)$			
实验号	列号		
	乙醇浓度	Co 负载量	装料比
1	1 (0.3)	1 (0.5)	1 (1)



## Continued

2	1	2 (1)	2 (2)
3	2 (0.9)	1	2
4	2	2	1

以上四组正交试验, 及所预测的最优解组, 构成五组实验来探究 450°C 下的更有利反应条件, 其中加入预测最优解组的目的是验证所预测结果。

## 5. 结论

探究不同温度及催化剂条件下对于反应的影响时, 将催化剂组合拆分成多个自变量并量化表示, 联合温度, 分别以乙醇转化率和 C4 烯烃选择性为因变量, 进行可筛选有效变量的 Lasso 回归。利用筛选后的变量组合再进行多元线性回归, 得出回归方程进行分析。得出自变量每变化一个单位, 因变量的变化量。其中, 温度对乙醇转化率和 C4 烯烃选择性的影响最大。但不足在于因数据量的限制, 进行多元线性回归的拟合优度有待提高。

为预测最优化的反应条件, 将求解得到的回归方程剔除不显著变量进行再次回归, 所得回归方程作为目标函数, 确定约束条件, 进行模拟退火求解最优解, 得到使得 C4 烯烃收率尽可能高的催化剂组合与温度。只需要在约束条件中加入温度小于 350°C 即可得出, 温度低于 350°C 时, 使得 C4 烯烃收率尽可能高的催化剂组合与温度。解决了多变量连续优化问题导致的“维数灾难”, 出现局部最优解的可能较低。

450°C 下的数据量过少, 而最优条件的温度为 450°C, 这值得设计实验进一步探究。选取最优解作为其中一组, 另外控制温度 450°C、装料方式为 A 类、HAP 质量 200 mg, 针对乙醇浓度、Co 负载量、装料比, 设计三因子、两变量的正交实验得到另外 4 组, 设计使 C4 烯烃收率可能达最高的 5 组实验, 以探究进一步优化反应的条件。

本文所建立的基于 Lasso 回归的产物收率模型、基于模拟退火算法的产物收率优化模型及多因素下的正交试验设计方法可推广至多种物质的生产之中, 对解决其他化工生产优化问题, 具有一定的参考意义和经济价值。

## 参考文献

- [1] 童刘, 刘宗章, 张敏华. 乙醇法制备 1,3-丁二烯的研究进展[J]. 化学工业与工程, 2012, 29(4): 38-44.
- [2] 吕绍沛. 乙醇偶合制备丁醇及 C4 烯烃[D]: [硕士学位论文]. 大连: 大连理工大学, 2018.
- [3] 中国工业与应用数学学会. 2021 高教社杯全国大学生数学建模竞赛赛题[EB/OL]. [http://www.mcm.edu.cn/html\\_cn/node/4d73a36cc88b35bd4883c276afe39d89.html](http://www.mcm.edu.cn/html_cn/node/4d73a36cc88b35bd4883c276afe39d89.html), 2021-10-05.
- [4] 荣小辉. Lasso 回归模型的变量选择功能及其在糖尿病发展预测模型中的应用[J]. 中国科技纵横, 2017(24): 26-29.
- [5] 刘娜, 谭亦旻, 莫伟强, 等. 基于模拟退火算法的 Halbach 直线发电机优化设计[J]. 电工技术学报, 2021, 36(6): 1210-1218.
- [6] 司守奎, 孙兆亮. 数学建模算法与应用[M]. 北京: 国防工业出版社, 2015.
- [7] 邓聪颖, 殷国富, 方辉, 肖红. 基于正交试验的机床结合部动刚度优化配置[J]. 机械工程学报, 2015, 51(19): 146-153.