

乙醇催化偶合制备 C4 烯烃工艺条件的优化问题

摘要

随着各国二氧化碳排放猛增，世界各国以全球协约的方式减排温室气体，我国因此提出碳达峰和碳中和目标。世界烯烃需求量不断增大，C4 烯烃作为基础化工原料被广泛应用。顺应国家政策，C4 烯烃以清洁能源乙醇作为原料出发制备成为更加符合环境保护和经济实惠的途径。

本文针对乙醇催化偶合制备 C4 烯烃的问题，通过建立 spearman 相关系数模型、多因素方差分析，探索较好的化学工艺条件。

针对问题一，我们研究温度与乙醇转化率和 C4 烯烃转化率的关系。利用附件 1 提供的数据，使用 SPSS 对每组催化剂组合做 spearman 相关系数分析，并且采用多种曲线拟合，得到二次曲线回归模拟效果最好。对给定实验结果进行折线分析并且利用 spearman 相关系数模型进行综合判定。

针对问题二，我们探讨不同催化剂组合及温度对乙醇转化率以及 C4 烯烃选择性大小的影响。作为催化剂载体，HAP 性能优于石英砂。我们选取 Co/SiO₂ 质量与 HAP 质量的比例、乙醇浓度 ml/min、Co 负载量 wt%、Co/SiO₂ 与 HAP 的总质量和温度五个参数作为影响因素，采用多因素方差进行显著性分析，得到温度影响因子数量级最高，影响最显著，Co 负载量 wt%、Co/SiO₂ 与 HAP 的总质量影响显著，略低于温度，Co/SiO₂ 质量与 HAP 质量的比例、乙醇浓度 ml/min 对乙醇转化率影响也较显著，但数量级较低。对 C4 烯烃选择性来说，Co/SiO₂ 质量与 HAP 质量的装料比，Co/SiO₂ 与 HAP 的总质量以及温度都是主要影响因素，Co 负载量也对其产生影响，但影响程度略微小于前三个因素，而比例对其基本没有影响。我们对组合又进行具体分析并用神经网络多层感知器进行佐证。

针对问题三，寻找最优解使得在相同实验条件下 C4 烯烃收率尽可能高，我们通过控制变量和综合问题二的结论初步优化反应条件，采用神经网络，提出了基于多层感知器对工艺条件的优化方法。最终可以得到在 200mg2wt%Co/SiO₂-200mgHAP-乙醇浓度 2.1ml/min、400 摄氏度条件下，C4 烯烃预测收率为 45.86%，为预测所得最优的工艺条件。在 350 摄氏度以下时，得到 200mg2wt%Co/SiO₂-200mgHAP-乙醇浓度 1.68ml/min、325 摄氏度时，即为 A2 收率最高。

针对问题四，增加实验设计得到更好的催化剂组合及温度，对问题三提出的最优解进行验证，具有探究实际结果的现实能力，可根据结果进一步优化工艺条件选择模型。

最后我们对模型的适用范围进行了推广并且对存在的问题进行了改进，在化学实验实际应用中可起到参考作用。

关键字：SPSS、MLP、方差分析、spearman 相关系数分析、神经网络

目录

1 问题背景与重述	4
1.1 问题背景	4
1.2 问题重述	4
2 模型假设	4
3 符号说明	4
4 问题分析	5
4.1 问题一分析	5
4.2 问题二分析	5
4.3 问题三分析	6
4.4 问题四分析	6
5 探究温度与乙醇转化率、C4 烯烃的选择性的关系	6
5.1 模型建立	6
5.2 模型求解	6
5.2.1 数据的预处理	6
5.2.2 计算 spearman 相关系数	7
5.2.3 预判断回归函数	9
5.2.4 回归模拟及确定关系	9
5.2.5 对单次实验结果分析	10
6 探讨催化剂组合及温度对乙醇转化率以及 C4 烯烃选择性的影响	12
6.1 模型建立	12
6.2 模型求解	12
6.2.1 前提条件验证	12
6.2.2 指标影响分析	14
6.2.3 其他因素分析	15
6.2.4 对附件中工艺条件具体分析	15
6.3 问题总结	15
6.4 模型验证	16
7 探究催化剂组合与温度使 C4 收率达到尽可能高	16
7.1 模型建立	16
7.2 模型求解	17
7.2.1 指标数据预处理及补充	17
7.2.2 基于 MLP 神经网络的工艺条件优化模型	19
8 根据模型选择增加实验设计	21

9 模型评价与推广	21
9.1 模型的优点	21
9.2 模型的不足	21
9.3 模型的推广	22
10 参考文献	22
A aa	23
B bb	23

1 问题背景与重述

1.1 问题背景

温室效应是人类亟待需要解决的全球性问题，随着各国二氧化碳排放猛增，世界各国以全球协议的方式减排温室气体，我国由此提出碳达峰和碳中和目标。随国家政策顺势而为，伴随着行业产能的不断提升，烯烃原料多元化呈现波澜壮阔的景象，特别是轻质化原料受到更多的青睐。

乙醇是一种清洁能源，来源丰富，不仅可采用工业制法，也可由粮食及各种植物纤维发酵而得。随着乙醇产量的逐年增加和生产成本的下降，其作为分子转化为其他衍生产物具有广阔的应用前景。随着当前全球烯烃市场的迅猛发展，世界烯烃需求量不断增大。 C_4 烯烃作为轻质化原料的重要方向和基础化工原料被广泛应用。

C_4 烯烃以乙醇作为原料。随着可持续发展战略的部署，选择由乙醇出发制备 C_4 烯烃成为更加符合环境保护和经济实惠的途径。

1.2 问题重述

本文对催化剂组合以及温度对乙醇催化偶合制备 C_4 烯烃的化学反应结果（ C_4 烯烃的选择性和 C_4 烯烃收率）的影响进行了探讨，提出最优化的催化剂组合方式，并对 350 摄氏度时给定的制备条件下在一次实验停留时间不同的测试结果进行了分析，为进一步提高 C_4 烯烃的收率和下一步实验方向提供依据。

问题一：研究每组内催化剂组合与乙醇转化率、 C_4 烯烃的选择性与温度的关系，并对 350 摄氏度时给定的制备条件下在一次实验停留时间不同的测试结果进行分析。

问题二：探讨不同催化剂组合及温度对乙醇转化率以及 C_4 烯烃选择性大小的影响。

问题三：在一定条件限制内，选择较优的催化剂组合与温度在其他条件一致下 C_4 烯烃收率尽可能高。

问题四：在有限实验次数中，选择下一步实验方向，更细致地优化催化剂组合及温度。

2 模型假设

1. 每次实验结果有独立性与可重复性。
2. 实验数据客观真实，实验条件稳定，忽略因测量方法及仪器、周围环境以及人为观察等外部条件所产生的实验误差。
3. 实验各产物间可存在相互转化的关系。
4. 实验转化率、选择性、收率理想状态最大为 1。模型拟合时以上三个指标不得超过 1。
4. 温度可以通过改变催化剂活性和乙醇反应量对反应产生影响。

3 符号说明

符号	意义
M_1	Co/SiO_2 质量
M_2	HAP 质量
W_1	Co/SiO_2 质量与 HAP 质量的比

W_2	乙醇浓度 ml/min
W_3	Co 负载量 $wt\%$
W_4	Co/SiO_2 与 HAP 的总质量
T	温度
P	预测排名
R	实测排名
D	排名差

4 问题分析

4.1 问题一分析

催化是自然界中普遍存在的现象，在现代化学工业中，据统计约有 90% 的化学产品在生产过程中都使用了催化剂。^[1] 在有机化学反应过程中，催化剂的选择性更是起决定性作用。在乙醇催化偶合制备 C_4 烯烃的化学工艺中，催化剂组合与温度对乙醇转化率和 C_4 烯烃选择性会产生影响。

研究附件 1 中每种催化剂组合，分析乙醇转化率、 C_4 烯烃的选择性与温度的关系，采用多种曲线拟合关系，找出最合适的拟合曲线类型，确定其相关关系。根据 350 摄氏度时给定的催化剂组合在一次实验不同时间的测试结果，利用 spearman 相关系数模型进行综合判定，分析乙醇转化率与各生成物选择性之间的相关性。

第一问采用先相关性分析再回归分析的方法。我们使用 SPSS 软件建立对原变量分布不作要求的 spearman 相关系数模型对其进行相关性分析，证明变量间相关性强。为了进一步探究被解释变量和解释变量之间的函数关系，对每种催化剂组合下的乙醇转化率（单位：百分比）——温度（单位：摄氏度）和 C_4 （单位：百分比）——温度图像进行了曲线绘制和曲线拟合。为保证稳健性，使用 SPSS 中曲线拟合工具箱进行多种曲线类型的拟合，并分别对结果进行拟合优度检验，得出准确的数学关系，构造出乙醇转化率和 C_4 选择性与温度间的数学模型。最终，对每种催化剂情况的拟合情况进行总体描述，利用所得数据分类汇总，得到一般情况下最合适的拟合曲线类型，确定其相关关系。

第二问使用 Excel 绘制乙醇转化率和各生成物选择性与时间的折线图进行模糊粗略判断，再使用 spearman 相关系数模型进行综合分析。此外，为保证结果可靠性并考虑 C_4 烯烃收率 = 乙醇转化率 \times C_4 烯烃选择性，将 C_4 烯烃收率也添加为一组变量。

4.2 问题二分析

该问题探讨催化剂组合及温度对乙醇转化率以及 C_4 烯烃选择性的影响，有关催化剂对反应的影响途径进行判断，我们选取 Co/SiO_2 质量与 HAP 质量的比例、乙醇浓度 ml/min 、 Co 负载量 $wt\%$ 、 Co/SiO_2 与 HAP 的总质量和温度五个参数作为影响因素。

若忽略影响因素之间的相互关系，只对某一指标采取单一 t 检验或单因素方差分析，不仅会出现影响因素分析出现差错的几率，还会在遇到分析结果不一致时难以得到一个综合的结论。多因素方差分析正是利用数学模型的可分解性，找出影响结果的主要因素^[2]。

为了研究五种指标对乙醇转化率以及 C_4 烯烃选择性大小的影响，我们采用多因素方差分析探究。将五个影响因素作为方差指标，记作 W_1 、 W_2 、 W_3 、 W_4 、 T 。而该五个影响因素不仅独立影响反应结果，它们还存在交互关系对转化率和选择性产生影响。而多因素方差不仅多个控制变量对观测变量的独立影响，更能够分析多个控制变量的交互作用能否对观测变量产生显著影响，最终找到

利于观测变量的最优组合^[3]。由于多因素方差检验前提条件严格，若指标不符合假设，我们则采取非参数检验。

4.3 问题三分析

针对问题三，寻找最优解使得在相同实验条件下 C4 烯烃收率尽可能高，我们通过控制变量和综合问题二的结论初步优化反应条件，采用神经网络，提出了基于多层感知器对工艺条件的优化方法。

由于实验数量有限且变量条件难以精确控制导致数据量少或部分指标数据量的不匹配直接影响最优工艺条件的构建和预测精度，因而我们基于控制变量和问题二的因素显著性结论对数据进行预处理和补充，在一定程度上提高模型的精确度。

基于 python 搭建神经网络建立工艺条件探究模型。当 MLP 网络应用于催化剂组合与温度工艺条件选择时，输入层神经元是影响 C4 烯烃收率的变量，可直接获取的影响指标，输出层神经元为 C4 烯烃收率。

4.4 问题四分析

针对问题四，基于问题三得到的结果是通过神经网络 MLP 预测所得的较优的五组催化剂组合与温度的选择进行实际检验。由于 MLP 没有一个标准的求最优解的方法，因而我们选择对较优解均进行验证，从而经过实验数据支撑得到最优的催化剂组合与温度。使得制备过程不仅提升催化剂—沸石的利用率，同时提高 C4 烯烃产物的收率。

5 探究温度与乙醇转化率、C4 烯烃的选择性的关系

5.1 模型建立

第一问研究附件 1 中每种催化剂组合，分析乙醇转化率、C4 烯烃的选择性与温度的关系，采用多种曲线拟合关系，找出最合适的拟合曲线类型，确定其相关关系。根据 350 度时给定的催化剂组合在一次实验不同时间的测试结果，利用 spearman 相关系数模型进行综合判定，分析乙醇转化率与各生成物选择性之间的相关性。

步骤如下：

Step1 数据预处理，清除异种数据，保证实验数据的可信性。

Step2 计算 spearman 相关系数，建立相关模型。

Step3 采用多种曲线拟合关系，找出最合适的拟合曲线类型，确定其相关关系。

Step4 对给定实验结果进行折线分析并且利用 spearman 相关系数模型进行综合判定。

5.2 模型求解

5.2.1 数据的预处理

对数据进行特征统计和分析，发现无缺失数据。进行异常数据的处理，由于将其按催化剂组合分为 21 组，考虑不同类型催化剂对最终的转化率和选择性会有较大的影响，因此选择在每一组内进行被解释变量的异常值的处理，而不将数据合并进行异常值处理。

采用 *Excel* 进行异常值处理操作，使用箱型图分析法。选定分析数据集 A ，确定数据集 A 的上四分位数 Q_1 和下四分位数 Q_2 ，并定义四分位距 $IQR = Q_2 - Q_1$ 。给定正常值区间 $U = [Q_1 -$

$1.5 \times IQR, Q_3 + 1.5IQR]$ 。若考察数据 ai 箱型图明确表示了上、下四分位数以及正常值区间 U ，异常值以“o”在图中标示。

A. 组内异常值处理

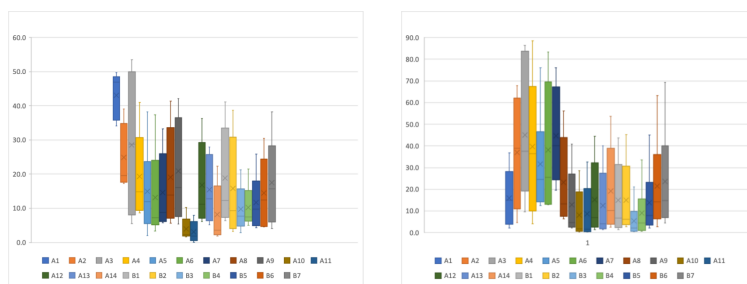


图 1: 单组乙醇转化率和 C_4 烯烃选择性对异常值检验箱型图

观察可知，在所分出的 21 组数据内，并无异常值的出现，符合预期。

B. 总体异常值处理

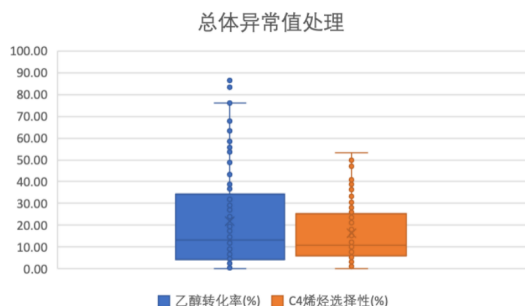


图 2: 总体乙醇转化率和 C_4 烯烃选择性对异常值检验箱型图

经分析可知，在乙醇转化率一项中出现异常值。但是，此异常值存在两个，来自“ $A_3: 200mg1wt\%Co/SiO_2 - 200mgHAP - 乙醇浓度 0.9ml/min$ ”组别和“ $A_2: 200mg2wt\%Co/SiO_2 - 200mgHAP - 乙醇浓度 1.68ml/min$ ”组别。均无法以它组数据作为集合主体，但将某一组的某个数据进行排除，违反了除自变量以外无其他变量的实验数据处理要求。

C. 数据整理

综上所述，所给数据并无异常值和缺失值。可以直接用于 spearman 相关系数模型建立以及后续问题解决。

5.2.2 计算 spearman 相关系数

利用 SPSS 的分析-相关-spearman 系数对每一组的相关系数计算，取 A_1 组合计算结果为例。其中“**”表示在 0.01 级别（双尾）相关性显著；“*”在 0.05 级别（双尾）相关性显著。

整理每一组的乙醇转化率-温度相关系数和 C_4 烯烃选择性-温度相关系数，列表如下。其中转化代表乙醇转化率-温度相关系数、选择代表 C_4 烯烃选择性-温度相关系数。

表 2: A1 相关-spearman 系数计算结果

A1 相关性					
			温度	乙醇转化率 (%)	C4 烯烃选择性 (%)
斯皮尔曼 Rho	温度	相关系数	1	1.000**	.900*
		显著性 (双尾)	.	.	0.037
		N	5	5	5
	乙醇转化率 (%)	相关系数	1.000**	1	.900*
		显著性 (双尾)	.	.	0.037
		N	5	5	5
	C4 烯烃选择性 (%)	相关系数	.900*	.900*	1
		显著性 (双尾)	0.037	0.037	.
	N	5	5	5	

表 3: 乙醇转化率-温度相关系数和 C4 烯烃选择性-温度相关系数

A1		A2		A3		A4		A5		A6	
转化	选择	转化	选择	转化	选择	转化	选择	转化	选择	转化	选择
1**	0.9*	1**	0.9*	1**	0.964**	1**	0.943**	0.943**	1**	0.899*	0.754
A7		A8		A9		A10		A11		A12	
转化	选择	转化	选择	转化	选择	转化	选择	转化	选择	转化	选择
1**	1**	1**	1**	1**	1**	1**	0.7	1**	1**	1**	1**
A13		A14		B1		B2		B3		B4	
转化	选择	转化	选择	转化	选择	转化	选择	转化	选择	转化	选择
1**	1**	1**	1**	1**	1**	1**	1**	1**	1**	1**	0.812*
B5		B6		B7							
转化	选择	转化	选择	转化	选择						
1**	1**	1**	1**	1**	1**						

虽然其中只有 A6 和 A10 的 C4 烯烃选择性-温度相关系数在 95% 的置信水平下没有显著相关, 且相关系数值小。在所有的 42 个相关系数中, A6 和 A10 的 C4 烯烃选择性-温度相关性是最小的。

我们将催化剂的四个参数: Co/SiO₂ 质量、HAP 质量、乙醇浓度、Co 负载量分别以字母 M₁、M₂、W₂、W₃。整理所有催化剂的四类参数如表所示, 以此来探寻 A6 和 A10 催化剂是否有异于它组的特性。

表 4: 催化剂的四类参数数

催化剂编号	M_1 (mg)	W_3 (%)	M_2 (mg)	W_2 (ml/min)	编号	M_1	W_3	M_2	W_2
A1	200	1	200	1.68	A12	50	1	50	1.68
A2	200	2	200	1.68	A13	67	1	33	1.68
A3	200	1	200	0.9	A14	33	1	67	1.68
A4	200	0.5	200	1.68	B1	50	1	50	1.68
A5	200	2	200	0.3	B2	100	1	100	1.68
A6	200	5	200	1.68	B3	10	1	10	1.68
A7	50	1	50	0.3	B4	25	1	25	1.68
A8	50	1	50	0.9	B5	50	1	50	2.1
A9	50	1	50	2.1	B6	75	1	75	1.68
A10	50	5	50	2.1	B7	100	1	100	0.9

通过观察数据, 只有 A6 和 A10 催化剂组合的 C_o 负载量为 5%, 其余所有组合的 C_o 都在 2% 及以下负载量。因此可以得出 C_o 负载量为 5% 会对 C4 烯烃选择性-温度相关性产生影响, 对于此种相关性进行削弱。

经过上述的 spearman 相关系数分析得出: 大部分情况下, 乙醇转化率、C4 烯烃的选择性与温度有很强的相关性。在明确了强相关性的基础上, 接下来将利用回归拟合的方法对乙醇转化率、C4 烯烃的选择性与温度之间建立具体的回归模型。

5.2.3 预判断回归函数

为初步判断选择拟合曲线的类型, 我们分别对 21 组数据进行了散点图绘制。取前五组的散点图为例, 如下所示:

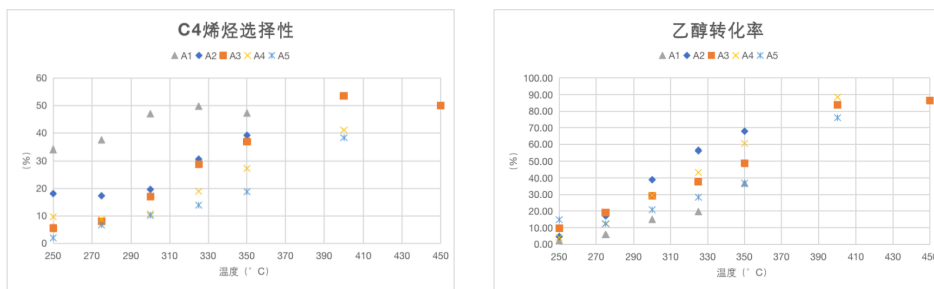


图 3: 乙醇转化率和 C4 烯烃选择性对温度的散点图

在上图中粗略判断关系趋势走向, 乙醇转化率和 C4 烯烃选择性与温度之间有可能成线性、二次方和其他非线性关系。通过题干确定模型中乙醇转化率和 C4 烯烃的选择性为被解释变量, 温度为解释变量。另外题目中的催化剂组合不作为解释变量出现, 仅仅用于标定多次重复实验的序号 (A1 - B7, 共 21 组)。

5.2.4 回归模拟及确定关系

为保证结果的稳健型, 我们利用 SPSS 中的四种模型进行回归曲线模拟。分别是:

- (1) 线性关系: $y = a_1 \times x + b_1$
 (2) 二次方关系: $y = a_2 \times x^2 + b_2x + c_2$
 (3) S 型关系: $\ln y = a_3/x + b_3$
 (4) 指数关系: $\ln y = \ln a_4 x + b_4$

利用 SPSS 分析-回归-曲线拟合工具箱进行拟合, 拟合结果统计如下表。

	乙醇转化率-温度								C4 烯烃选择性-温度								是否无法通过95置信水平显著性检验?
	线性		二次		S		指数		线性		二次		S		指数		
	R方	p	R方	p	R方	p	R方	p	R方	p	R方	p	R方	p	R方	p	
A1	0.9096	0.0077	0.9595	0.0203	0.9779	0.0009	0.9433	0.0038	0.7159	0.0448	0.8319	0.0840	0.7978	0.0263	0.7247	0.0426	√
A2	0.9867	0.0004	0.9822	0.0089	0.9248	0.0058	0.8527	0.0161	0.7812	0.0297	0.9606	0.0197	0.7093	0.0464	0.8022	0.0254	
A3	0.9571	0.0001	0.9495	0.0011	0.9685	0.0000	0.8780	0.0012	0.8954	0.0008	0.9326	0.0020	0.9171	0.0004	0.7952	0.0044	
A4	0.9938	0.0000	0.9931	0.0003	0.9389	0.0009	0.8383	0.0066	0.8966	0.0026	0.9609	0.0036	0.8392	0.0065	0.9037	0.0023	
A5	0.8413	0.0063	0.9901	0.0005	0.8527	0.0054	0.9303	0.0012	0.9251	0.0014	0.9842	0.0009	0.9534	0.0005	0.8915	0.0029	
A6	0.9567	0.0025	0.9718	0.0141	0.9239	0.0059	0.9361	0.0045	0.7119	0.0458	0.8908	0.0546	0.8398	0.0184	0.8873	0.0107	√
A7	0.9983	0.0000	0.9993	0.0003	0.9902	0.0003	0.9386	0.0043	0.9165	0.0068	0.9993	0.0003	0.9389	0.0042	0.9849	0.0005	
A8	0.9395	0.0042	0.9981	0.0009	0.9815	0.0007	0.9922	0.0002	0.9776	0.0009	0.9991	0.0005	0.9980	0.0000	0.9760	0.0010	
A9	0.7967	0.0265	0.9810	0.0095	0.9473	0.0034	0.9938	0.0001	0.9931	0.0002	0.9898	0.0051	0.9844	0.0005	0.9253	0.0057	
A10	0.8025	0.0254	0.9878	0.0061	0.9914	0.0002	0.9835	0.0006	0.6561	0.0606	0.9569	0.0215	0.6121	0.0735	0.7568	0.0351	√
A11	0.7544	0.0356	0.9742	0.0129	0.9945	0.0001	0.9896	0.0003	0.9710	0.0014	0.9989	0.0006	0.8628	0.0145	0.7566	0.0351	
A12	0.9037	0.0084	0.9981	0.0010	0.9992	0.0000	0.9694	0.0015	0.9555	0.0026	0.9985	0.0007	0.9816	0.0007	0.9950	0.0001	
A13	0.8362	0.0190	0.9927	0.0037	0.9817	0.0007	0.9978	0.0000	0.9690	0.0015	0.9633	0.0183	0.9716	0.0013	0.9120	0.0073	
A14	0.9054	0.0082	0.9942	0.0029	0.9985	0.0000	0.9756	0.0011	0.8933	0.0099	0.9989	0.0005	0.9542	0.0027	0.9881	0.0004	
B1	0.9006	0.0088	0.9978	0.0011	0.9996	0.0000	0.9724	0.0013	0.9624	0.0020	0.9944	0.0028	0.9825	0.0006	0.9878	0.0004	
B2	0.8178	0.0224	0.9830	0.0085	0.9560	0.0026	0.9959	0.0001	0.9597	0.0022	0.9955	0.0023	0.9917	0.0002	0.9748	0.0011	
B3	0.7400	0.0175	0.9860	0.0008	0.9521	0.0006	0.9853	0.0001	0.9285	0.0013	0.9605	0.0036	0.9590	0.0004	0.9406	0.0009	
B4	0.7615	0.0146	0.9778	0.0015	0.9935	0.0000	0.9820	0.0001	0.7513	0.0160	0.9561	0.0043	0.5894	0.0459	0.7233	0.0199	
B5	0.7917	0.0111	0.9860	0.0008	0.9720	0.0002	0.9989	0.0000	0.9447	0.0007	0.9970	0.0001	0.9718	0.0002	0.9843	0.0001	
B6	0.8543	0.0053	0.9840	0.0009	0.9802	0.0001	0.9575	0.0004	0.9557	0.0005	0.9515	0.0050	0.9368	0.0010	0.9197	0.0016	
B7	0.8454	0.0060	0.9943	0.0002	0.9797	0.0001	0.9968	0.0000	0.9860	0.0000	0.9952	0.0002	0.9877	0.0000	0.9357	0.0010	
均值	0.8711	0.0110	0.9848	0.0046	0.9669	0.0013	0.9575	0.0021	0.8927	0.0110	0.9674	0.0110	0.8942	0.0116	0.8936	0.0095	

图 4: 对不同因素对乙醇转换率和 C4 烯烃选择性的拟合结果统计

计算出每种拟合曲线的均值, 并将 R 方从最低值到最高值进行色阶标示, 越红说明拟合越好, 越黄说明拟合越差。将 p 按如下规则进行条件突出显示, $p > 0.05$, 显示蓝色; $0.05 > p > 0.01$, 显示绿色。即蓝色标示代表无法通过置信水平为 95% 的显著性检验, 即绿色标示代表无法通过置信水平为 99% 的显著性检验。

21 组催化剂组合的 42 个具体模拟方程在支撑材料中给出。对于具体的催化剂组合, 应依据支撑材料中给出的最佳拟合方程进行拟合。由于附件 1 所给数据最大温度为 450 摄氏度, 我们暂时把温度上限定为 450 摄氏度, 而该温度下乙醇转化率趋于饱和, C4 烯烃选择性也变化较小, 与二次方的趋势符合程度较低。因而回归曲线拟合温度区间 [250 摄氏度, 400 摄氏度]。

整体关系给出最佳的回归曲线拟合结果如下:

a) 在乙醇转化率-温度回归拟合中, 二次 (方) 类型模拟的 R^2 均值最大, P 值最小且 $P < 0.01$ 。因此选用二次 (方) 曲线回归模拟效果最好。

b) 在 C4 烯烃选择性-温度拟合中, 二次 (方) 类型模拟的 R^2 均值最大, P 值最小。因此选二次 (方) 用曲线回归模拟效果最好。

5.2.5 对单次实验结果分析

对附件 2 中 350 度时给定的催化剂组合在一次实验不同时间的测试结果进行分析。我们首先做出了各参数随时间变化的曲线图, 如下:

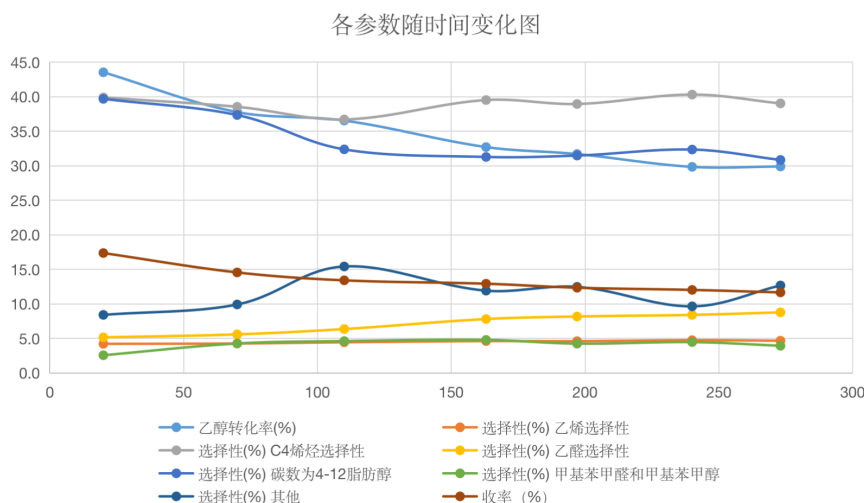


图 5: 各参数随时间变化折线图

分析考察停留时间对转化率以及催化性能的影响，产物的趋势分析如图所示。可以看到，当反应时间从 20min 增加至 273min 时，乙醇转化率缓慢降低后逐渐稳定，各个产物的选择性变化不尽相同。当反应时间不断增长，随着 C4 烯烃选择性先降低再升高。

C4 烯烃收率一直呈下降趋势。为更加精确分析，依旧选择 spearman 相关系数模型进行。单次实验的参与关联性分析的因素更多，我们新添加了 C4 烯烃收率作为因素，检测最终的产物收率与哪项因素关系最大。利用 SPSS 进行 spearman 相关性分析，得到相关系数表，为更有利于直观分析，我们对 R^2 进行了色阶标志，效果图如下。

	时间	乙醇转化率	乙醛选择性	C4 烯烃选择性	乙醛选择性	碳数为4-12脂	甲基苯甲醛和	其他	C4 烯烃收率
时间		-0.964	0.929	0.214	1	-0.857	0.071	0.393	-1
乙醇转化率	-0.964		-0.964	-0.321	-0.964	0.750	-0.179	-0.250	0.964
乙醛选择性	0.929	-0.964		0.393	0.929	-0.786	0.321	0.214	-0.929
C4 烯烃选择性	0.214	-0.321	0.393		0.214	-0.071	-0.179	-0.714	-0.214
乙醛选择性	1	-0.964	0.929	0.214		-0.857	0.071	0.393	-1
碳数为4-12脂	-0.857	0.750	-0.786	-0.071	-0.857		-0.214	-0.536	.857
甲基苯甲醛和	0.071	-0.179	0.321	-0.179	0.071	-0.214		0.286	-0.071
其他	0.393	-0.250	0.214	-0.714	-0.393	-0.536	0.286		-0.393
C4 烯烃收率	-1	0.964	-0.929	-0.214	-1	0.857	-0.071	-0.393	

图 6: 各因素 spearman 相关因素色阶图

分析可得：

1. 我们可以观察到 C4 烯烃收率与时间的相关系数为 1，呈完全负相关，则随着实验时间的拉长，烯烃收率逐渐降低，若想要获得较高收率和经济实惠的 C4 烯烃可以选择控制实验停留时间稍短一些。

2. 乙醛的选择性与 C4 烯烃收率也呈负相关关系，根据假设可知在反应过程中可能存在乙醛向 C4 烯烃发生转化的反应路径。

3. C4 烯烃收率其值为乙醇转化率与 C4 烯烃的选择性的乘积，其中乙醇转化率与 C4 烯烃收率的相关系数为 0.964，基本符合正相关关系，而相比之下 C4 烯烃选择性对收率影响不大，说明物质

收率在更大程度上与反应量具有更大的关系。

4. C_4 烯烃收率与乙烯选择性负相关，由于乙烯与 C_4 烯烃有相似性，均含有碳碳双键，在生成过程中相互竞争与抑制，在整个反应中呈相似趋势。

5. 碳数 4—12 脂肪醇与 C_4 烯烃收率呈正相关，脂肪醇可以断裂脱水生成碳碳双键后形成 C_4 烯烃，也有可能 C_4 烯烃与脂肪醇为同一反应的产物，因而呈正相关。

6. C_4 烯烃收率与甲基苯甲醛以及其他物质关联性较小，可预测其没有转化关系，不会产生较大影响。

6 探讨催化剂组合及温度对乙醇转化率以及 C_4 烯烃选择性的影响

6.1 模型建立

问题二研究五种因素对乙醇转化率以及 C_4 烯烃选择性大小的影响，我们采用多因素方差分析探究。

步骤如下：

Step1 提出假设，对正态分布、同质性、独立性三个前提条件进行验证，保证可以使用多因素方差分析。

Step2 若符合前提条件，则使用多因素方差分析方法；若不符合前提则采用非参数检验。

Step3 对分析结果进行判定，给定显著性水平，得出指标对乙醇转化率以及 C_4 烯烃选择性大小的影响。

6.2 模型求解

6.2.1 前提条件验证

A. 独立性检验

Co/SiO_2 质量与 HAP 质量的比例、乙醇浓度 ml/min 、 Co 负载量 $wt\%$ 、 Co/SiO_2 与 HAP 的总质量和温度五个因素，任取 $200mg$ $1wt\%Co/SiO_2$ - $200mgHAP$ 与乙醇浓度 $0.9ml/min$ 或者 300 摄氏度温度与 $1wt\%Co/SiO_2$ 的积事件的概率，都等于各事件概率之积，则称五个因素相互独立，符合独立性检验。

B. 残差正态检验

指标残差的分布必须是正态分布，否则就会使得得到的回归方程没有任何实际的意义。

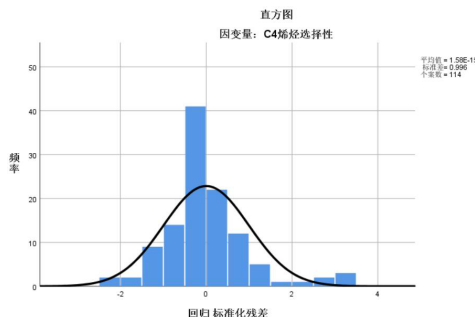


图 7: T 与 C_4 烯烃选择性对应的残差正态检验直方图

从回归标准化残差直方图形态来看，虽有一定的偏差，但基本符合正态分布。 pp 图数据点几乎都围绕在理论直线的周围，成一条直线，如上图所示，就可以认为数据是符合正态分布的。由残差

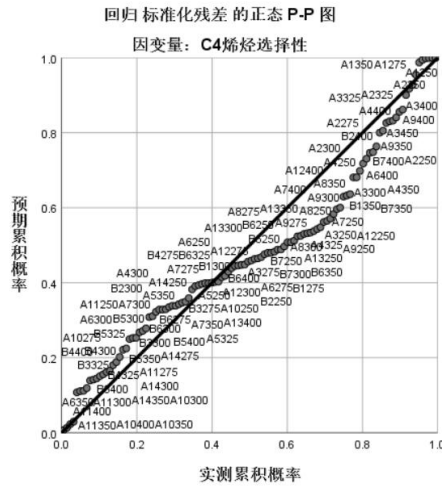


图 8: T 与 C_4 烯烃选择性对应的残差正态检验 pp 图

正态检验结果的支撑材料中可以发现五个因素与乙醇转化率以及 C_4 烯烃选择性的直方图与 pp 图都成上述形状，故不会过于偏态，基本满足正态性要求。

C. 方差齐次性检验

我们采用 *Levene* 检验检验两个或两个以上样本间的方差是否齐性，对五个指标和乙醇转化率以及 C_4 烯烃选择性所在总体的方差是否相等进行验证。*Levene* 检验对于数据的基础分布相当稳健，具有较好的把握度。^[4] 该检验的原假设是各个样本所在总体的方差相等，在显著性水平取 0.05 时，若渐进显著性结果小于显著性水平，则拒绝原假设，认为样本所在的总体方差不相等，则不可以使用多因素方差分析，采用非参数检验正态分布。

		方差齐次性检验			
		莱文统计	自由度 1	自由度 2	显著性
乙醇转化率	基于平均值	1.232	3	110	.302
	基于中位数	.767	3	110	.515
	基于中位数并具有调整后自由度	.767	3	108.905	.515
	基于剪除后平均值	1.000	3	110	.396
C_4 烯烃选择性	基于平均值	3.340	3	110	.022
	基于中位数	1.823	3	110	.147
	基于中位数并具有调整后自由度	1.823	3	104.116	.148
	基于剪除后平均值	2.956	3	110	.036

图 9: W_1 对乙醇转化率和 C_4 烯烃选择性的方差齐次性检验

因为 W_1 与乙转催化和 C_4 烯烃选择性的关系通过残差正态检验，基本符合正态分布，而基于中位数的 *Levene* 统计值主要用于对称分布或正态分布。

$$Z_{ij} = |Y_{ij} - Y_i \cdot|$$

式中: Z_{ij} 为定义数, Y_{ij} 为原始数据, $Y_i \cdot$ 是原始数据中第 i 个样本的算术均数。^[5]

W_1 对乙转催化和 C_4 烯烃选择性的方差同质性检验的验证结果如图 9 所示。其中显著性 $0.515 \geq 0.05$, 则总体方差相等, 符合同质性。

W_1 对乙转催化和 C_4 烯烃选择性的方差同质性检验的验证结果如表 1 所示。由表 1 可知, 其中显著性 $0.000 < 0.05$, 则总体方差不相等, 不符合同质性, 不能采用多因素方差分析。

方差齐性检验^{△,⊥}

		莱文统计	自由度 1	自由度 2	显著性
乙醇转化率	基于平均值	10.305	5	107	.000
	基于中位数	6.265	5	107	.000
	基于中位数并具有调整后自由度	6.265	5	72.072	.000
	基于剔除后平均值	9.971	5	107	.000
C4 烯烃选择性	基于平均值	2.677	5	107	.025
	基于中位数	2.215	5	107	.058
	基于中位数并具有调整后自由度	2.215	5	100.971	.059
	基于剔除后平均值	2.775	5	107	.021

△. 当 温度 = 450.00 时, 乙醇转化率 是常量, 已将其省略。

⊥. 当 温度 = 450.00 时, C4 烯烃选择性 是常量, 已将其省略。

图 10: T 对“乙醇转化率”和 $C4$ 烯烃选择性的方差齐次性检验

D. 前提条件验证小节

据此分析, W_1 、 W_2 、 W_3 、 W_4 满足对乙醇转化率和 $C4$ 烯烃选择性的方差分析条件 (即: 正态分布、同质性、独立性)。可以使用单变量多因素方差分析对 W_1 、 W_2 、 W_3 、 W_4 这四个指标进行对乙醇转化率和 $C4$ 烯烃选择性的影响分析。而对于不满足方差分析条件的 T , 使用非参数检验来分析对乙醇转化率和 $C4$ 烯烃选择性的影响。

6.2.2 指标影响分析

A. 多因素方差分析

使用 *SPSS* 进行多因素方差分析, 分析时因变量为乙醇转化率和 $C4$ 烯烃选择性, 由于前四个指标符合假设检验, 则固定因子为 W_1 、 W_2 、 W_3 、 W_4 , 而由于不符合假设但是对因变量存在影响, 则设其为协变量 T 。

表 5: 对乙醇转化率的主体间效应校验

源	修正模型	截距	温度	比例	速率	WT	总质量	速率 * 总质量
F	39.008	280.342	431.464	11.576	9.379	44.291	35.739	20.127
显著性	0	0	0	0	0	0	0	0

其中显著性数据 ≤ 0.01 说明主效应极显著, 0.05 说明主效应不显著。可以看出五个因素对乙醇转化率的主效应都极显著, 温度影响因子数量级最高, 影响最显著, Co 负载量 $wt\%$ 、 Co/SiO_2 与 HAP 的总质量影响显著, 略低于温度, Co/SiO_2 质量与 HAP 质量的比例、乙醇浓度 ml/min 对乙醇转化率影响也较显著, 但数量级较低。其中乙醇浓度与 Co/SiO_2 与 HAP 的总质量乘积也能对 W_2 、 W_4 两种因素的显著性佐证。

表 6: 对 $C4$ 烯烃选择性的主体间效应校验

源	修正模型	截距	温度	比例	速率	WT	总质量	速率 * 总质量
F	35.805	244.338	444.957	1.31	9.114	3.797	5.057	1.055
显著性	0	0	0	0.276	0	0.013	0	0.352

由图可知, W_2 、 W_3 、 W_4 、 T 的渐进显著性水平 ≤ 0.001 , 远远小于显著性水平 0.05 , 可以看出四个因素对 $C4$ 烯烃选择性的主效应极显著, 但是其中 $W_3 = 0.013$ 虽然小于 0.05 呈显著性水平,

但是比其他三种因素对 C_4 烯烃的影响因子小一些。而 $W_2 = 0.276$ ，远远大于显著性水平， W_2 对 C_4 烯烃选择性几乎不影响。

B. 非参数检验

由于 T 不满足方差分析条件，对 T 进行对乙醇转化率和 C_4 烯烃选择性的非参数检验。该指标采用独立样本 *Kruskal – Wallis* 检验，显著性数据 ≤ 0.05 说明影响极显著。下图显著性为 0.000，说明温度对乙醇转化率与 C_4 烯烃选择性影响极显著。

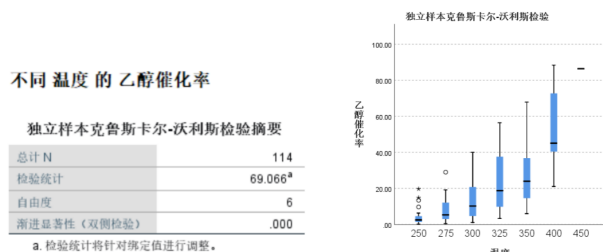


图 11: 温度与乙醇转化率的 *Kruskal – Wallis* 检验结果

6.2.3 其他因素分析

通过 A11 与 A12 两组数据进行对比，我们可以发现其变量只有反应载体——石英砂与 HAP 的不同，其他条件均相同。在温度发生 250 摄氏度——400 摄氏度梯度变化时，乙醇转化率和 C_4 烯烃选择性，以石英砂作催化剂载体均低于 HAP ，可以基本判断 HAP 为更好的催化剂载体，可以形成活化中心，不仅加快乙醇转化率，还可以提高 C_4 烯烃选择性。而 A11 以石英砂作为载体的实验结果与其他 20 组数据有较大出入，对其他物质的选择性相似度较低，因而在后续建模过程中，为保证催化剂活性和效果更优，将 A11 在数据库中剔除。

又根据问题一中 21 组拟合曲线趋势相似，数据相似程度高，可以粗略判断装料方式对反应基本没有影响或者影响较小可以忽略不计。

6.2.4 对附件中工艺条件具体分析

在五种影响因子中温度影响因子数量级最高，先考虑温度该因素，观察数据可知温度越高，乙醇转化率越高，基本可以判定在所给工艺条件中 450 摄氏度下的实验，乙醇转化率最高。在 400 摄氏度下，首要考虑 Co 负载量 $wt\%$ 、 Co/SiO_2 与 HAP 的总质量两个因素，对于 Co 负载量 $wt\%$ ，我们发现 A4、A6、B7 转化率较高，而对于 Co/SiO_2 与 HAP 的总质量，其中转化率较高的为 A3、A4、A5、A6。该几组催化剂组合在乙醇转化前十名存在率为 100%，侧面证明了问题二的分析证明是正确的。采用同样方法可以确定 C_4 烯烃选择性。

6.3 问题总结

题中反应利用了 $SiO_2 - HAP$ 催化剂，并在其表面负载具有脱氢活性的 Co 金属，提高选择性催化乙醇制备 C_4 烯烃。作为催化剂载体， HAP 性能优于石英砂，且催化剂的装料方式不影响反应结果。我们结合多种表征手段，探究催化剂组合及温度对乙醇转化率以及 C_4 烯烃选择性的影响。由上我们可以得到 Co/SiO_2 质量与 HAP 质量的比例、乙醇浓度 ml/min 、 Co 负载量 $wt\%$ 、 Co/SiO_2 与 HAP 的总质量和温度五个因素对乙醇转化率影响都极大，温度影响因子数量级最高，影响最显著， Co 负载量 $wt\%$ 、 Co/SiO_2 与 HAP 的总质量影响显著，略低于温度， Co/SiO_2 质量

与 HAP 质量的比例、乙醇浓度 ml/min 对乙醇转化率影响也较显著，但数量级较低。对 $C4$ 烯烃选择性来说， Co/SiO_2 质量与 HAP 质量的装料比， Co/SiO_2 与 HAP 的总质量以及温度都是主要影响因素， Co 负载量也对其产生影响，但影响程度略微小于前三个指标。

6.4 模型验证

问题二可以采用神经网络中多层感知器（MLP）进行佐证。

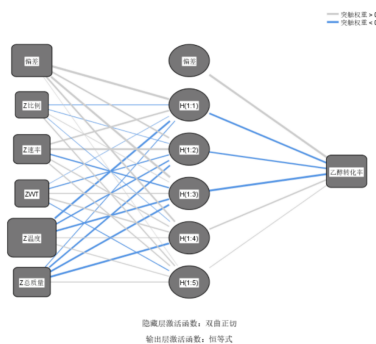


图 12: 神经网络示意图

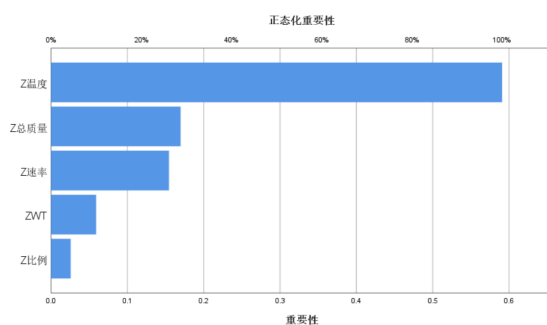


图 13: 对于乙醇转化率神经网络正态化重要性柱状图

图 13 可直观看出，得到温度影响因子数量级最高， Co 负载量 $wt\%$ 、 Co/SiO_2 与 HAP 的总质量影响显著， Co/SiO_2 质量与 HAP 质量的比例、乙醇浓度 ml/min 对乙醇转化率影响显著，但数量级较低。与多元方差分析基本符合，则问题二模型建立正确，佐证成功。

7 探究催化剂组合与温度使 $C4$ 收率达到尽可能高

7.1 模型建立

多层感知器 (Multi-Layer Perceptron, MLP) 是一种前向结构的人工神经网络，目前广泛应用于图像处理、模式识别、优化计算、函数逼近等领域。^[6]

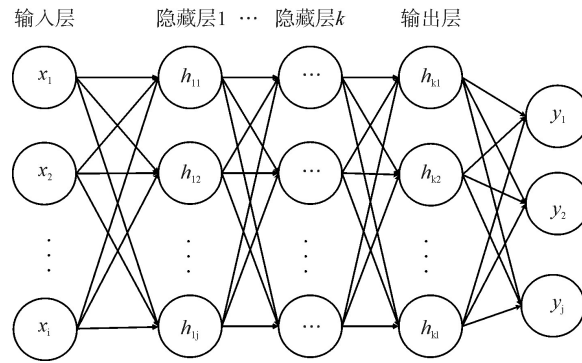


图 14: MLP 结构

步骤如下:

Step1 指标数据预处理

Step2 随机初始化模型权重 W

Step3 将输入值 X 传入模型进行向前传播, 得到结果概率

Step4 使用损失函数计算预测值 (比如 $[0.3, 0.3, 0.4]$) 和实际值 (比如 $[0, 0, 1]$) 的差, 得到损失值 J 常用的逻辑回归损失函数是交叉熵函数。

Step5 根据模型权重计算损失梯度。

Step6 使用梯度下降做反向传播来更新模型权重。更新后的权重将会使得输出不正确的结果 j 的概率降低, 从而提高输出正确类别 y 的概率。

Step7 重复 *Step3—5* 直到模型的表现最好。

7.2 模型求解

7.2.1 指标数据预处理及补充

探究催化剂组合与温度, 使得在相同实验条件下 C_4 烯烃收率尽可能高。由于实验数量有限且变量条件难以精确控制导致数据量少或部分指标数据量的不匹配直接影响最优工艺条件的构建和预测精度, 因而我们基于控制变量和问题二的因素显著性结论对数据进行预处理和补充, 在一定程度上提高模型的精确度。由于温度的影响因子极大, 影响极为显著, 将温度因素滞后处理。为使数据库不至于庞大, 采用控制变量法对影响催化剂的四种因素进行优化处理, 剔除催化剂效果不佳的催化剂组合, 减少神经网络的运算量。

根据问题二中的因素对乙醇转化率和 C_4 选择性的综合影响显著性的顺序进行分析。首先对 Co 的负载量进行控制变量分析。由图可知共存在 $0.5wt\%$ 、 $1wt\%$ 、 $2wt\%$ 、 $5wt\%$ 四个指标, 观察 Co 的负载量为 $0.5wt\%$ 、 $5wt\%$ 时, C_4 收率较大, 则优化 Co 的负载量后剩余变量为 $1wt\%$ 、 $2wt\%$ 。

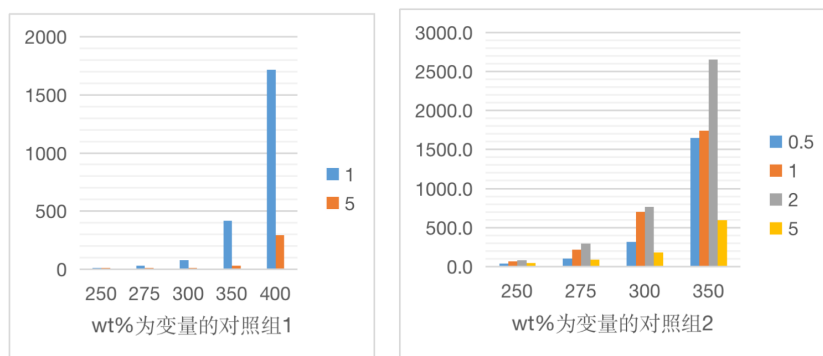


图 15: wt% 为变量的对照组数据直方图

乙醇浓度对乙醇转化率和 C_4 选择性都显著影响，但由于该因素对其他影响存在过多交互效应，在浓度梯度中没有存在较优的浓度，为保证催化剂组合的完整性，我们保留所有乙醇浓度选择。因而乙醇浓度选择性有 $0.3ml/min$ 、 $0.9ml/min$ 、 $1.68ml/min$ 、 $2.1ml/min$ 。

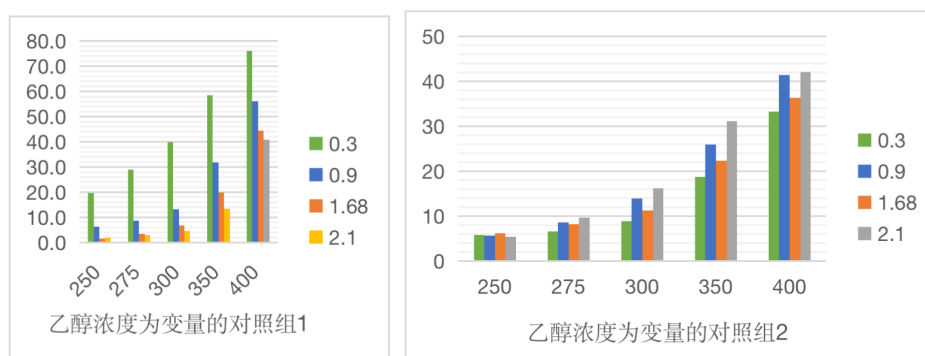


图 16: 乙醇浓度为变量的对照组数据直方图

Co/SiO_2 与 HAP 的总质量有 $20mg$ 、 $50mg$ 、 $100mg$ 、 $140mg$ 、 $150mg$ 、 $200mg$ 、 $400mg$ ，根据控制变量只能寻找到关于 $100mg$ 与 $400mg$ 的对照组， $400mg$ 总质量时收率完全高于 $100mg$ 时。而当其他条件一致，总质量偏小的催化剂组合的 Co 的负载量条件优于总质量偏大的催化剂组合时，依旧是总质量偏大的收率较高，因而该因素优化条件为 Co/SiO_2 与 HAP 的总质量有 $200mg$ 、 $400mg$ 。

根据图直观看出来料比 $1 : 1$ 收率明显优于其他比例，采用相同方法优化优化装料比条件为 $1 : 1$ 。

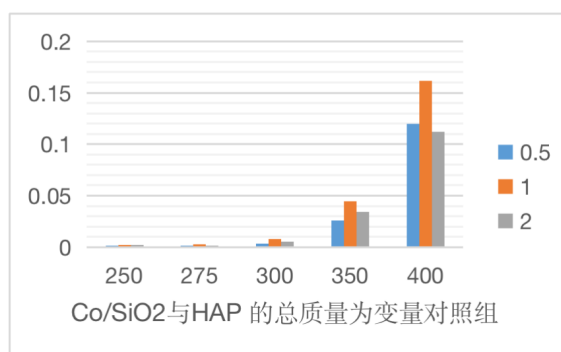


图 17: Co/SiO_2 与 HAP 的总质量为变量对照数据直方图

A、B 装料方式的收率差距不大，装料方式 A、B 同时考虑。

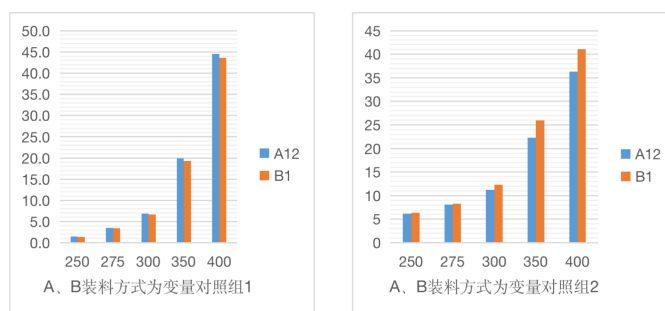


图 18: A、B 装料方式为变量对照组数据直方图

考虑温度因素，在所有给出的实验数据中，只有 A3 组存在 450 摄氏度下的实验数据。计算 A3 组别收率关于温度的分布，折线图表示如下。

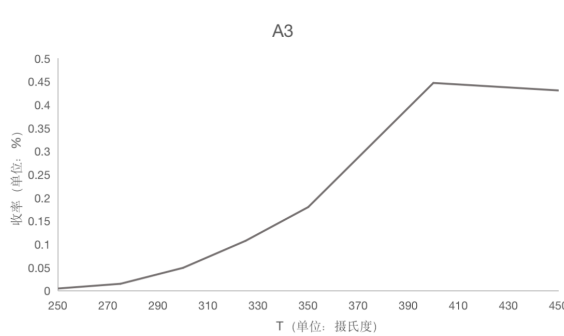


图 19: A3 组有关温度的收率折线图

由此可见，在温度到达 450 摄氏度后，烯烃收率下降。又由于仅给出一组 450 度的数据，不足以代表整体的研究方向，因此，我们的补全数据时并不将 450 度考虑在内。

据上所述，以实验变量分析，共增添 158 种需预测组合，且有 114 种已知组合。至此，将数据补齐共有 A1 – A24 和 B1 – B23，共计 47 种催化剂组合，272 种具体组合方式。

7.2.2 基于 MLP 神经网络的工艺条件优化模型

在 SPSS 作为大环境下，我们建立基于 MLP 神经网络的工艺条件优化模型。

首先设置了一系列名称为是否需要预测的数组，0 代表为新添组合，1 代表已有实验收率测量值的组合。这一步的目的是在神经网络分析的设置中，确定数据的训练集。将补全后的 242 条数据导入 SPSS。由于考虑到装料方式仅有两种：A 和非 A（即 B），因此，我们将装料方式作为一组名义变量测量，A 组装料方式记为 0，B 组装料方式记为 1。其余的负载量、浓度、总质量和温度设定为标度测量。使用 SPSS 的分析-神经网络-多层感知器，设定收率为因变量，装料方式为变量，负载量、浓度、总质量和温度为协变量，进行标准化变量重新标度。分区设定即选择训练集时，选择使用分区变量匹配个体，利用设定好的“是否需要预测”作为分区变量。在保存设置中，选择保存每个因变量的预测值或类别，完成设定，利用 BP 神经网络模型进行对收率的预测，预测值显示后导入 Excel 表格进行分析，将所有的数据进行总体收率降序排名，成为最终排名。将所有已有实测收率值的分别按照实验结果的收率和神经网络预测的收率排名，分别得到预测排名 P 和实测排名 R。

为更加有利于我们的对比和判别，我们又定义了：排名差

$$D = P - R$$

我们将神经网络预测结果取收率最高的十组，列表展示如下。

表 7: 神经网络预测结果

编号	T	W_3	W_2	W_1	W_4	已知收率	预测收率	最终排名	P	R	D
B11	400	2	2.1	1	400		0.459	1			
A18	400	2	2.1	1	400		0.457	2			
A1	400	1	1.68	1	400		0.448	3			
A3	450	1	0.9	1	400	0.431	0.447	4	1	2	-1
A3	400	1	0.9	1	400	0.447	0.432	5	2	1	1
A16	400	1	2.1	1	400		0.422	6			
B14	400	1	0.9	1	400		0.412	7			
B12	400	1	1.68	1	400		0.405	8			
B13	400	2	1.68	1	400		0.393	9			

做出 D 的柱状图，并将 P 和 R 画在同一折线图中。

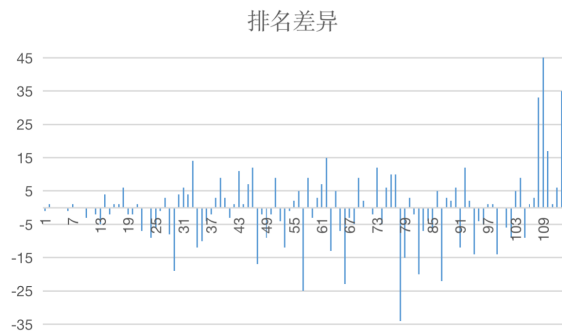


图 20: 排名差柱状图

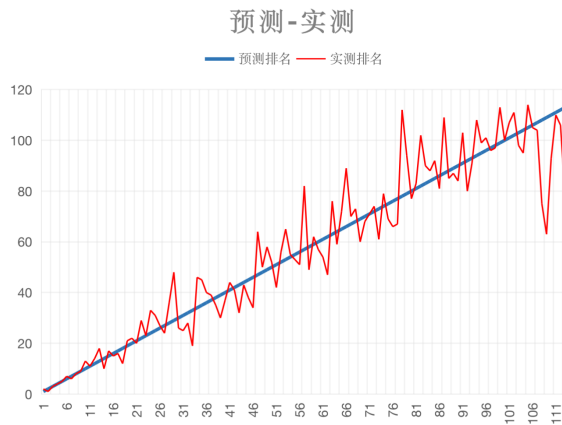


图 21: 排名差折线图

可以发现,排名越高 D 值越小,也就是说明我们所需要的收率最好的组合,在预测值和实测值上,差别很小,并可以非常精确的在所有数据中预测出排名最高的数。这也表明了 MLP 神经网络模型在解决此问题上的优势。最终可以得到在 $200\text{mg}2\text{wt}\%Co/SiO_2-200\text{mg}HAP$ -乙醇浓度 $2.1\text{ml}/\text{min}$ 、 400 摄氏度条件下, $C4$ 烯烃预测收率为 45.86% ,为预测所得最优的工艺条件。即:

$$\begin{pmatrix} W_1 & W_2 & W_3 & W_4 & T \end{pmatrix} = \begin{pmatrix} 1 & 2.1 & 2 & 400 & 400 \end{pmatrix} \quad (1)$$

在 350 摄氏度以下时,得到 $200\text{mg}2\text{wt}\%Co/SiO_2-200\text{mg}HAP$ -乙醇浓度 $1.68\text{ml}/\text{min}$ 、 325 摄氏度时,即为 $A2$ 收率最高。

即:

$$\begin{pmatrix} W_1 & W_2 & W_3 & W_4 & T \end{pmatrix} = \begin{pmatrix} 1 & 1.68 & 2 & 400 & 325 \end{pmatrix} \quad (2)$$

8 根据模型选择增加实验设计

我们选择以下工艺条件时进行进一步实验:

1. $200\text{mg} 2\text{wt}\%Co/SiO_2-200\text{mg} HAP$ -乙醇浓度 $2.1\text{ml}/\text{min}$ 、 400 摄氏度;
2. $200\text{mg} 2\text{wt}\%Co/SiO_2-200\text{mg} HAP$ -乙醇浓度 $2.1\text{ml}/\text{min}$ 、 400 摄氏度;
3. $200\text{mg} 1\text{wt}\%Co/SiO_2-200\text{mg} HAP$ -乙醇浓度 $1.68\text{ml}/\text{min}$ 、 400 摄氏度;
4. $200\text{mg} 1\text{wt}\%Co/SiO_2-200\text{mg} HAP$ -乙醇浓度 $2.1\text{ml}/\text{min}$ 、 450 摄氏度;
5. $200\text{mg} 1\text{wt}\%Co/SiO_2-200\text{mg} HAP$ -乙醇浓度 $0.9\text{ml}/\text{min}$ 、 400 摄氏度。

针对问题四,我们在问题三在寻优时已经增添了较优的工艺条件进行分析,因而可以在问题三中寻找实验。并且问题三给出了所有收率的最终排名,选取最优且未经实验的催化剂组合与温度进行实际检验。问题三中我们也发现预测名次和实际名次之间仍存在差距,这是由于 MLP 没有一个标准的求最优解的方法,因而我们选择对较优解均进行验证,从而经过实验数据支撑得到最优的催化剂组合与温度。使制备过程不仅提升催化剂—沸石的利用率,同时提高 $C4$ 烯烃产物的收率。

9 模型评价与推广

9.1 模型的优点

1. 采用成熟的统计软件 *SPSS* 操作相对简单,可行性较高。
2. 综合考虑影响因素之间的相互关系,采用多因素方差分析,不只是某一指标采取单一 t 检验或单因素方差分析,不会出现影响因素分析出现差错的几率,也不会遇到分析结果不一致时难以得到一个综合的结论。
3. 研究结果从特殊到一般,从优化和预测两个角度出发,建立模型探讨催化剂的最优选择。
4. 该模型根据结果设计实验,有探究实际结果的现实能力,可进一步优化模型。

9.2 模型的不足

1. 问题二中模型应用的限制条件过于苛刻,三个条件同时满足的可能性不大,本文中温度因为不符合假设采用非参数检验。
2. 可用数据较少,可能会对结果产生影响,与实际不完全符合。

9.3 模型的推广

本文构建了有机反应中催化剂—沸石工艺条件的确定方式，采用方差分析和 spearman 相关系数对影响因子进行筛选，确定显著性因素，对化学反应的温度选择以及下一步实验方向确定具有积极意义。

10 参考文献

- [1] Chowdhury, A. D.; Houben, K.; Whiting, G. T.; Chung, S.-H.; Baldus, M.; Weckhuysen, B. M. *Nat. Catal.* 2018, 1, 23.
- [2] 张启贤, 刘新平, 林坤. 海外游客特征对在华停留时间影响的多因素方差分析 [J]. 佳木斯大学学报: 自然科学版, 2009, 5 (27) :727-731.
- [3] 黎毅. 会计实证研究方法: 东北财经大学出版社, 2012.08: 第 143 页
- [4] .Olejnik SF, Algina J.Type I Error Rates and Power Estimates of Selected Parametric and Non-parametric Tests of Scale.Journal of Educational Statistics, 1987, 12, 45-61.
- [5] 程琮, 范华.Levene 方差齐性检验 [J]. 中国卫生统计,2005(06):408+420.
- [6] Li Y, Yuan Y.Convergence Analysis of Two-layer Neural Networks with ReLU Activation[C].neural information processing systems,2017:597-607.

11 附录