

智能制造大数据技术实践

Practices of Big Data Technology in Intelligent Manufacturing

课程实践：TEP仿真实验
Practice Part：TEP Simulation

胡文凯

wenkaihu@cug.edu.cn



中国地质大学(武汉) 自动化学院
School of Automation, China University of Geosciences

- 了解工业过程数据特点和典型质量问题，练习并掌握工业数据分析常用方法
- 掌握数据预处理方法，包括数据平滑、离群点检测和处理、数据变换、数据缩减等
- 熟悉常用的数据可视化工具，能够运用不同方法对数据进行可视化分析，并观察数据特征
- 掌握数据基本统计和关联度分析方法
- 掌握其他数据处理与分析方法，如回归分析、分类分析等

➤ 线上讲解

- 总体任务安排介绍和TEP模型学习

➤ 线下实验

- 数据预处理和可视化仿真实验
- 参数预测仿真实验
- 故障诊断仿真实验

实验模型—TE模型

4

➤ 模型介绍

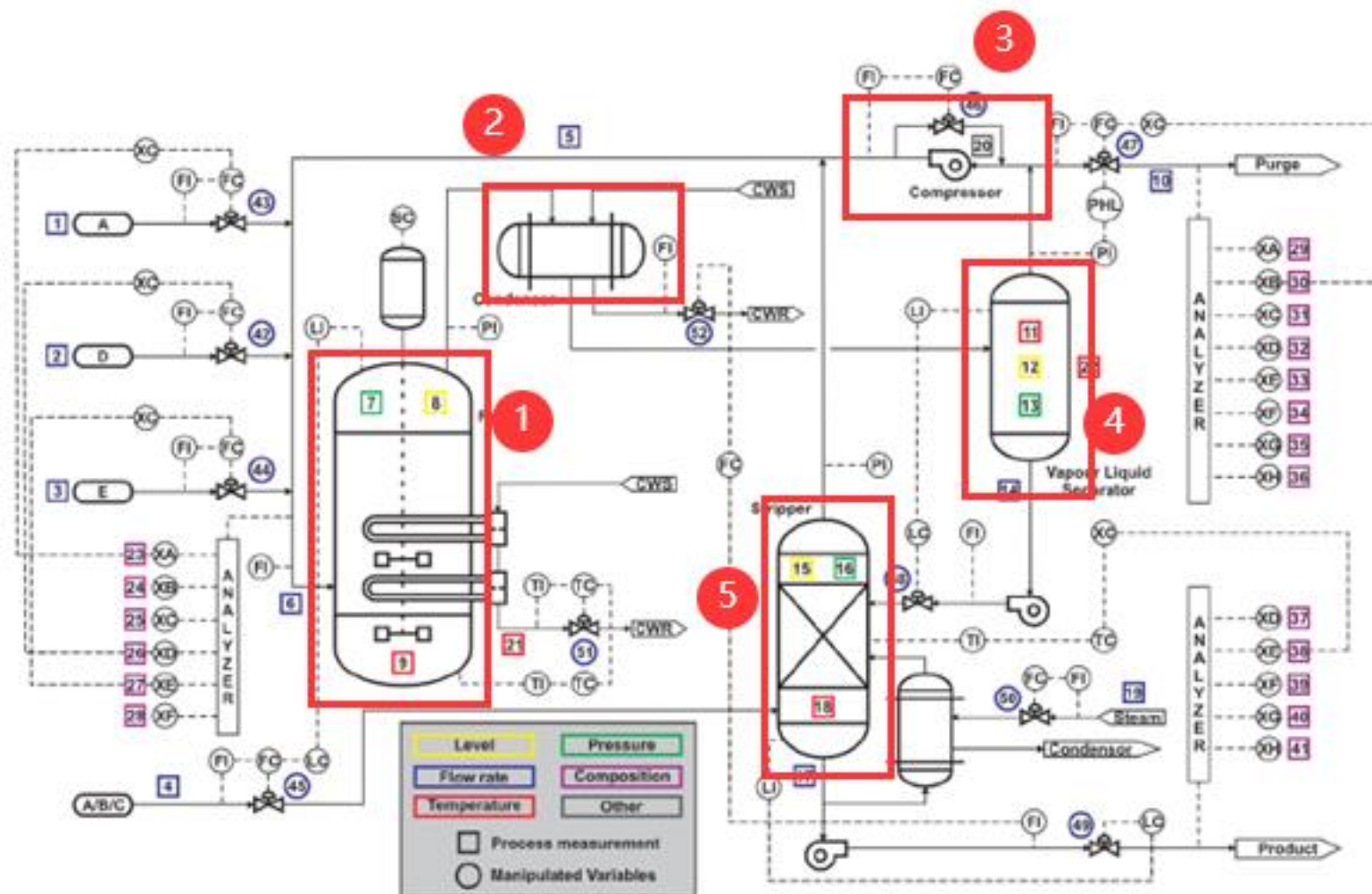
根据实际化学反应过程，美国 Eastman 公司开发了开放的化学模型模拟平台—TEP仿真平台。其生成数据具有时变性、强耦合、非线性等特点，被广泛应用于复杂系统的控制和故障诊断测试。

TE过程包括五个部分：

- 1) 反应器
- 2) 冷凝器
- 3) 压缩机
- 4) 分离器
- 5) 解吸塔

[depts.washington.edu/control/LARRY/TE/download.html#Basic TE Code](https://depts.washington.edu/control/LARRY/TE/download.html#Basic%20TE%20Code)

[temexd_mod.zip](#)



➤ 模型介绍

利用 Matlab 的 TEP_Model 生成不同故障的过程信号。该过程共有 11 个操纵变量（XMV1~11），41 个测量变量，有 6 种运行模式。

反应起始物包括四种：A、C、D、E
目标生产物：G、H
副产物：F

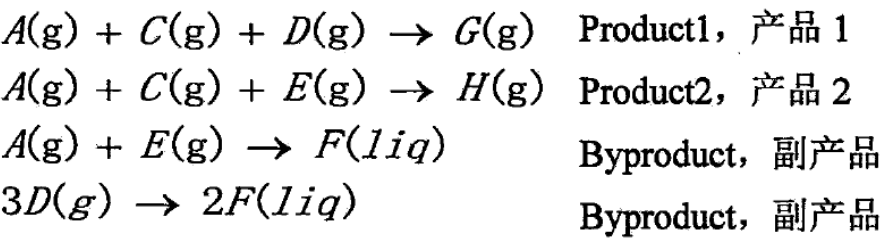


表1 TE化工过程操作模式

操作模式	产品中 G 和 H 比值	G/H 的生产率(流 11)
1	1:1	7038kg/h 和 7038kg/h
2	1:9	1408kg/h 和 12669kg/h
3	9:1	1000kg/h 和 1111kg/h
4	1:1	最大值
5	1:9	最大值
6	9:1	最大值

➤ 模型介绍

利用Matlab的TEP_Model生成不同故障的过程信号。该过程共有11个操纵变量（XMV1~11），41个测量变量，有6种运行模式。

表3 22个过程变量的信息

变量名	变量含义	基础值	单位	低报警限	高报警限
F1	A 物料流量	0.25052	kscmh	0.15758	0.34291
F2	D 物料流量	3664.0	kg/h	3557.1	3765.9
F3	E 物料流量	4509.3	kg/h	4394.7	4627.3
F4	A、C 混合物料流量	9.3477	kscmh	9.1050	9.5833
F5	回收流量	26.902	kscmh	26.287	27.517
F6	反应釜进料流量	42.339	kscmh	41.666	42.989
P7	反应釜压力	2705.0	kPa gauge	2685.0	2726.3
L8	反应釜液位	75.00	%	73.43	76.60
T9	反应釜温度	120.40	℃	120.34	120.45
Pr10	放空率	0.3372	kscmh	0.3734	0.3003
T11	产品分离器温度	80.109	℃	80.771	79.401
L12	产品分离器液位	50.000	%	46.834	53.025
P13	产品分离器压力	2633.7	kPa gauge	2612.9	2655.9
F14	产品分离器出口流量	25.160	m ³ /h	22.068	28.304
L15	汽提塔液位	50.000	%	46.836	53.081
P16	汽提塔压力	3102.2	kPa gauge	3084.3	312.4
F17	汽提塔出口流量	22.949	m ³ /h	21.104	24.710
T18	汽提塔温度	65.731	℃	64.462	67.245
F19	汽提塔蒸汽流量	230.31	kg/h	201.64	266.39
T21	反应釜冷却水出口温度	94.599	℃	94.220	94.990
T22	分凝器冷却水出口温度	77.297	℃	76.493	78.066

表4 19个成分变量的信息

变量号	描述	流号	基本模式值	单位
XMEAS(23)	成分 A	6	32.188	mol%
XMEAS(24)	成分 B	6	8.8933	mol%
XMEAS(25)	成分 C	6	26.383	mol%
XMEAS(26)	成分 D	6	6.8820	mol%
XMEAS(27)	成分 E	6	18.776	mol%
XMEAS(28)	成分 F	6	1.6567	mol%
XMEAS(29)	成分 A	9	32.958	mol%
XMEAS(30)	成分 B	9	13.823	mol%
XMEAS(31)	成分 C	9	23.978	mol%
XMEAS(32)	成分 D	9	1.2565	mol%
XMEAS(33)	成分 E	9	18.579	mol%
XMEAS(34)	成分 F	9	2.2633	mol%
XMEAS(35)	成分 G	9	4.8436	mol%
XMEAS(36)	成分 H	9	2.2986	mol%
XMEAS(37)	成分 D	11	0.01787	mol%
XMEAS(38)	成分 E	11	0.83570	mol%
XMEAS(39)	成分 F	11	0.09858	mol%
XMEAS(40)	成分 G	11	53.724	mol%
XMEAS(41)	成分 H	11	43.828	mol%

表2 52个过程信号的信息

	代号
1	控制变量 (XMV1~11)
2	过程变量 (XMEAS1~22)
3	成分变量 (XMEAS23~41)

➤ 模型介绍

TEP模型中有20个预设故障，包括15种已知故障和5种未知的故障。

表5 故障类型

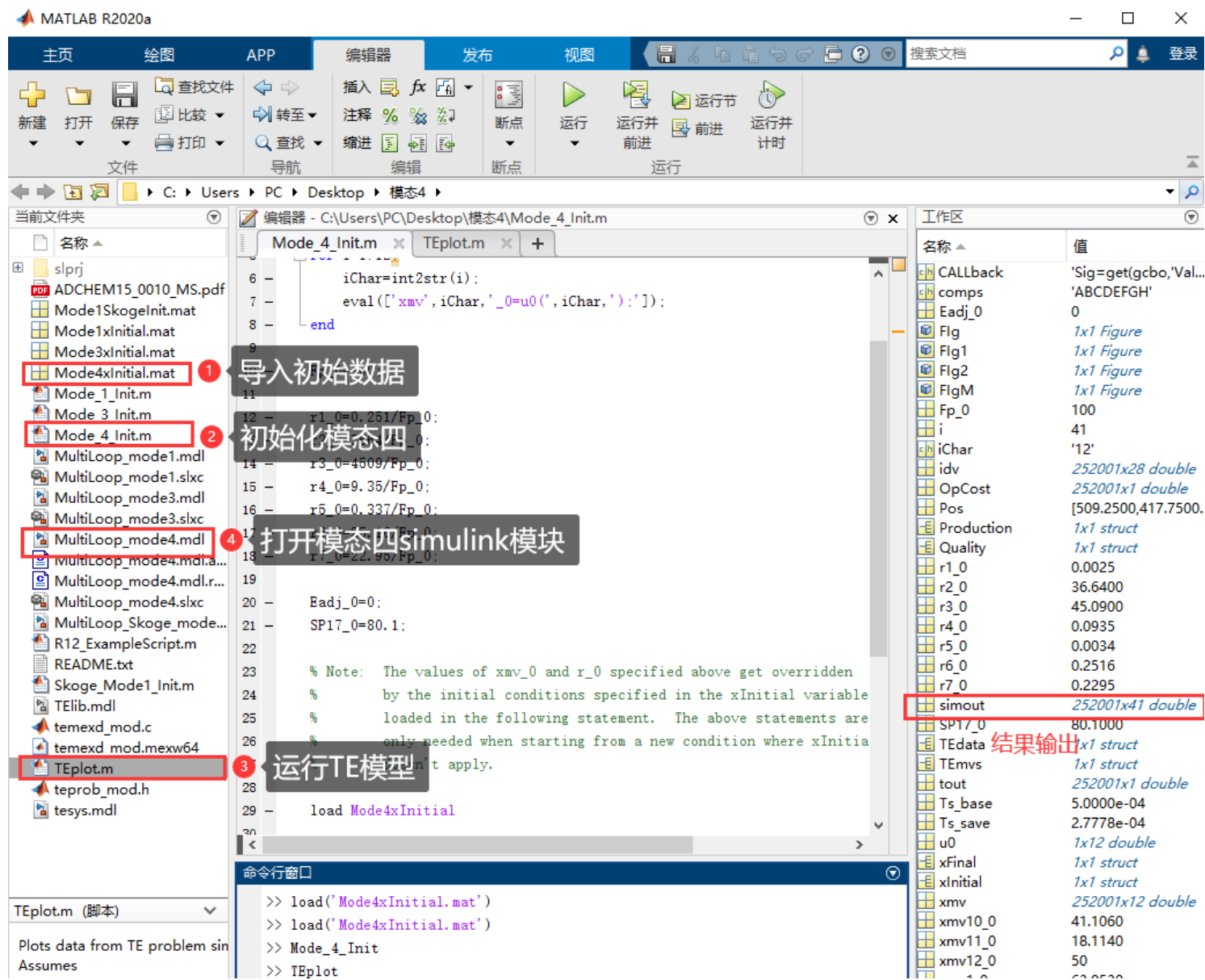
故障序号	故障原因
1~ 7	物料阶跃型变化
8 ~12	随机变换
13	慢偏移
14, 15	粘住
16 ~ 20	未知故障

表6 故障具体原因

变量号	变量名	类型
1	A/C 进料比率, B 成分不变 (流 4)	阶跃
2	B 成分, A/C 进料比率不变 (流 4)	阶跃
3	D 的进料温度 (流 2)	阶跃
4	反应器中冷却水的入口温度	阶跃
5	冷凝器中冷却水的入口温度	阶跃
6	A 进料损失 (流 1)	阶跃
7	C 存在压力损失-可用性降低 (流 4)	阶跃
8	A、B、C 进料成分 (流 4)	随机变量
9	D 的进料温度 (流 2)	随机变量
10	C 的进料温度 (流 2)	随机变量
11	反应器中冷却水的入口温度	随机变量
12	冷凝器中冷却水的入口温度	随机变量
13	反应动态	慢偏移
14	反应器冷却水阀门	粘住
15	冷凝器冷却水阀门	粘住
16	未知	未知
17	未知	未知
18	未知	未知
19	未知	未知
20	未知	未知

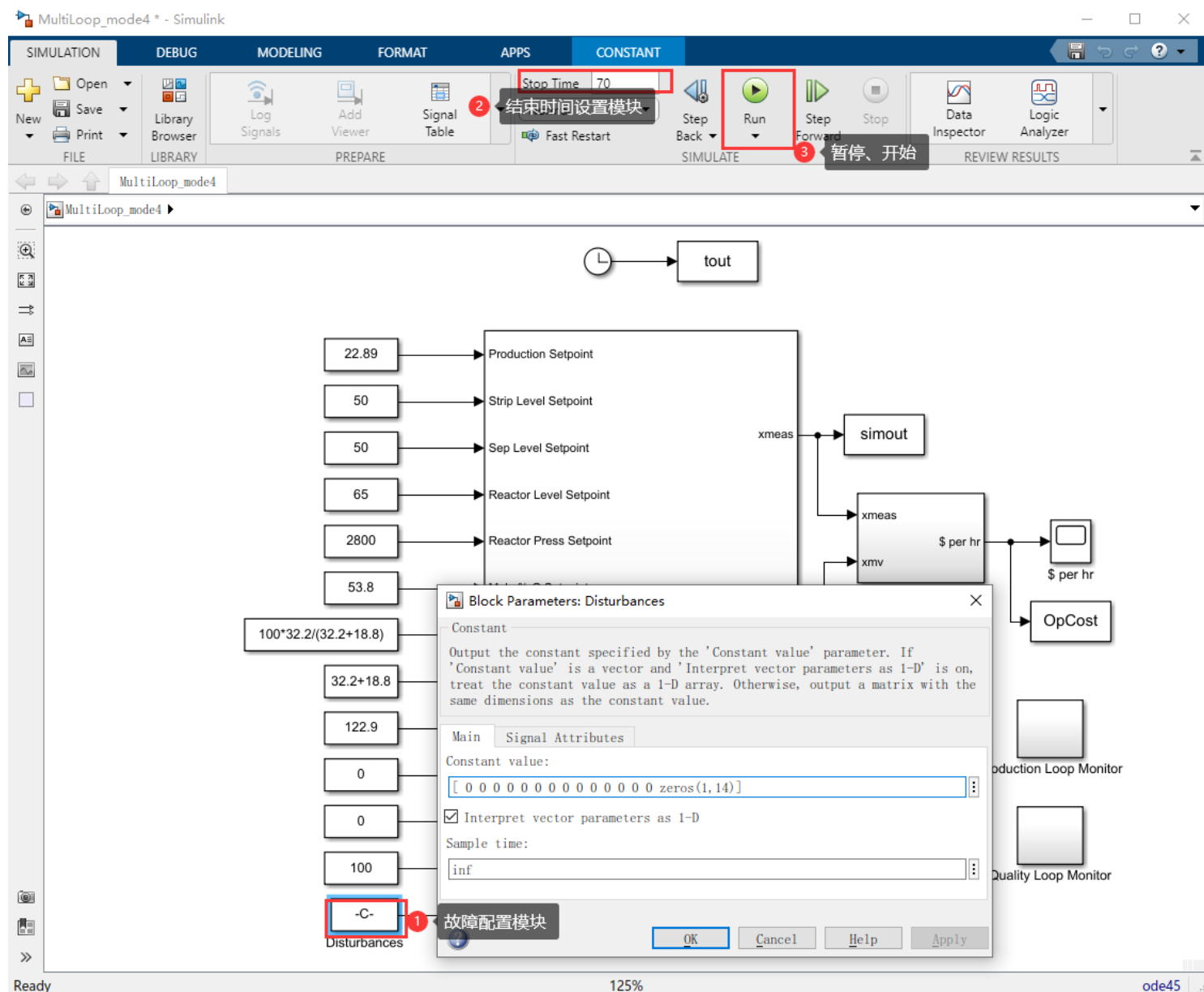
➤ 原始数据生成

- ❑ 导入模态初始数据
- ❑ 模态初始化
- ❑ 运行TE模型
- ❑ 打开对应模态simulink模块
 - 设置仿真总时间
 - 设置故障类型
 - 设置故障加入时间



➤ 原始数据生成

- ❑ 导入模态初始数据
- ❑ 模态初始化
- ❑ 运行TE模型
- ❑ 打开对应模态simulink模块
 - 设置仿真总时间
 - 设置故障类型
 - 设置故障加入时间



➤ 主要任务安排如下：

□ 数据基本统计分析

- 对原始数据进行可视化分析，分析典型数据质量问题（噪声、离群点等），比较不同变量间的统计值差异

□ 数据清洗

- 对比不同的方法在数据预处理的差异，并根据数据特征进行对应操作（如噪声问题，根据数据分布选择合适的方法去噪声）

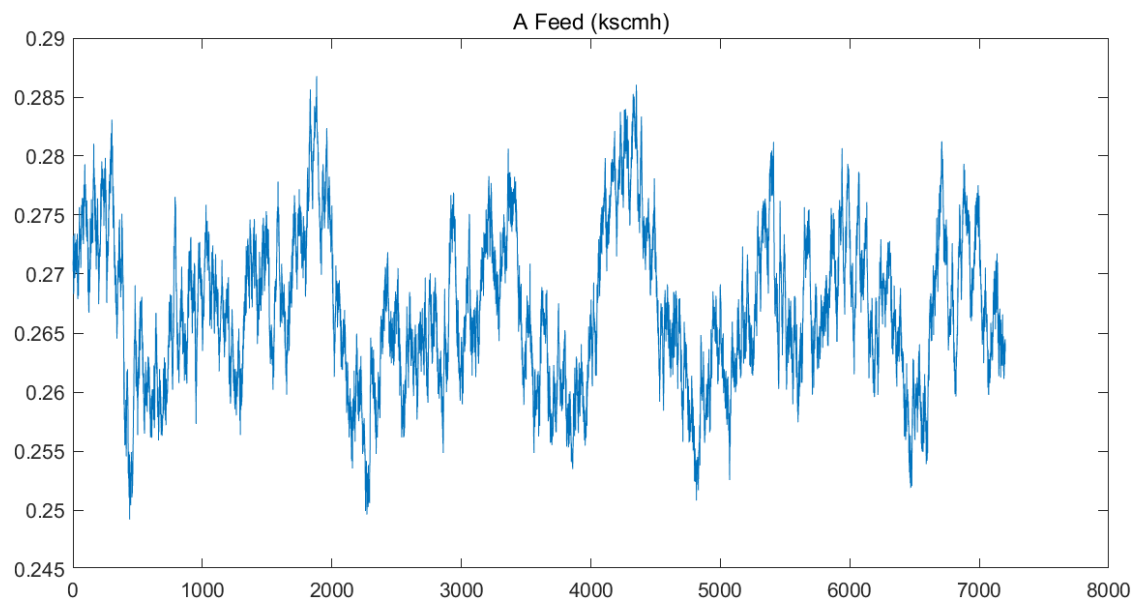
□ 数据变换

- 对数据采用不同的归一化方法，并通过可视化方法对比其差异性
- 采用主元分析，对数据进行降维处理

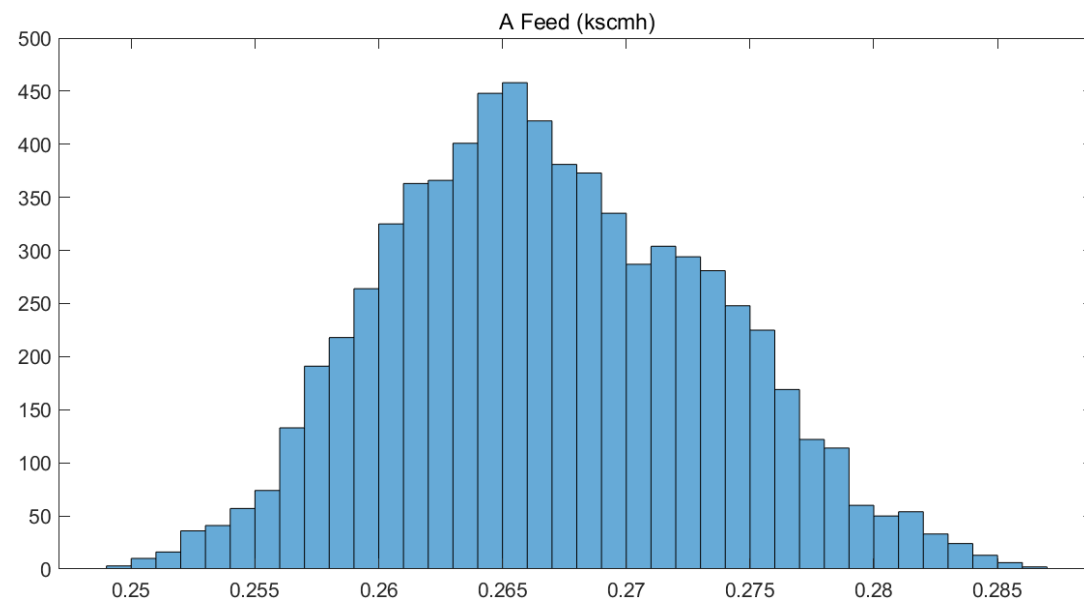
➤ 可视化图标示例

□ 通过统计分析和数据可视化，观察数据特征以及分布情况

plot()画出时序特征



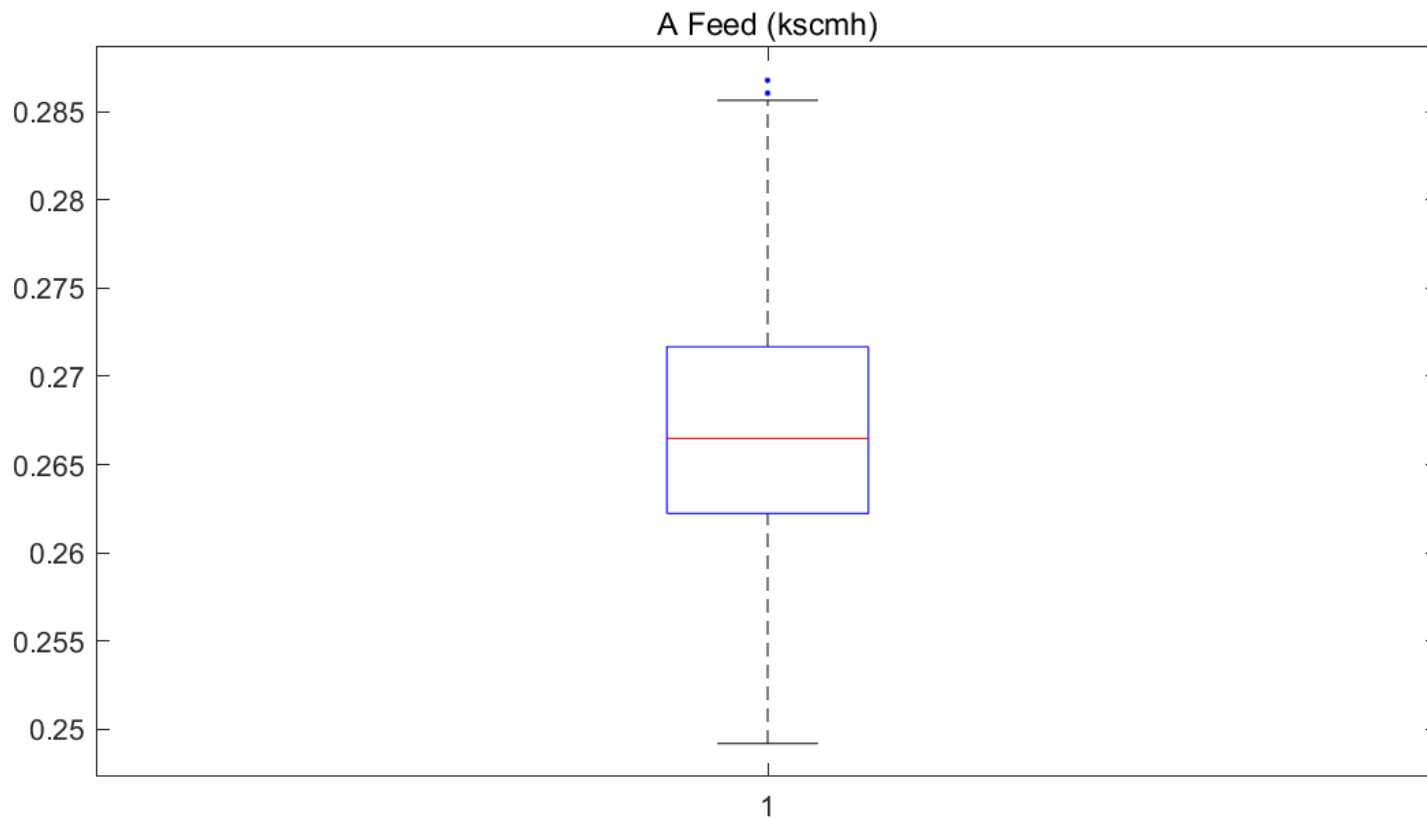
histogram()画出数据分布



➤ 可视化图标示例

- 通过统计分析和数据可视化，观察数据特征以及分布情况

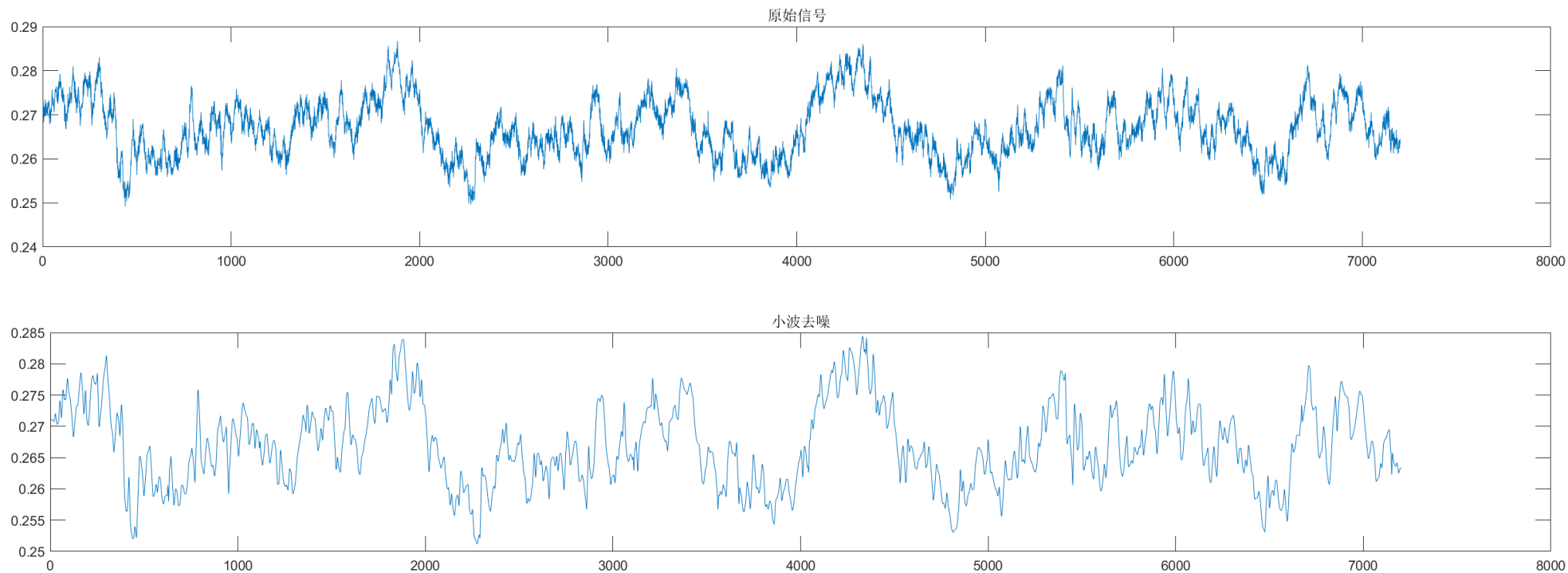
boxplot()画出箱体图



➤ 数据去噪示例

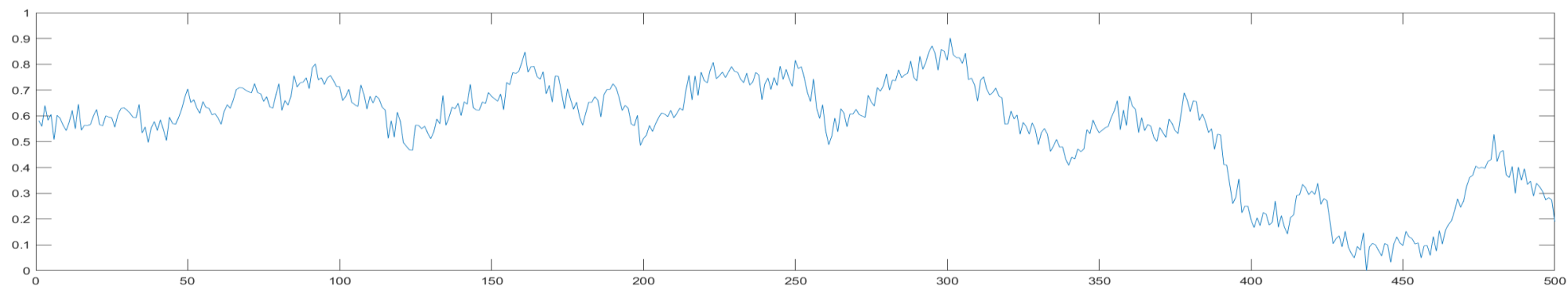
□ 根据数据特征，选用不同去噪声方法

wdenoise()小波变换去噪声

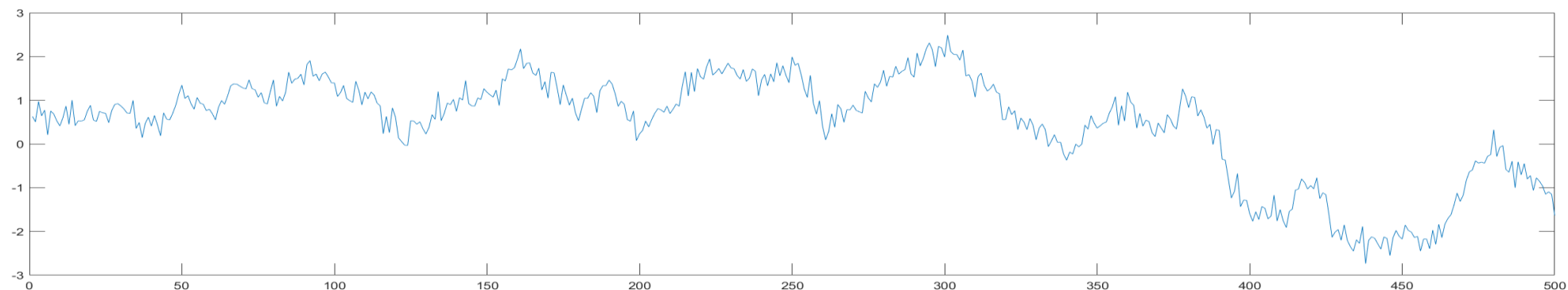


➤ 数据归一化示例

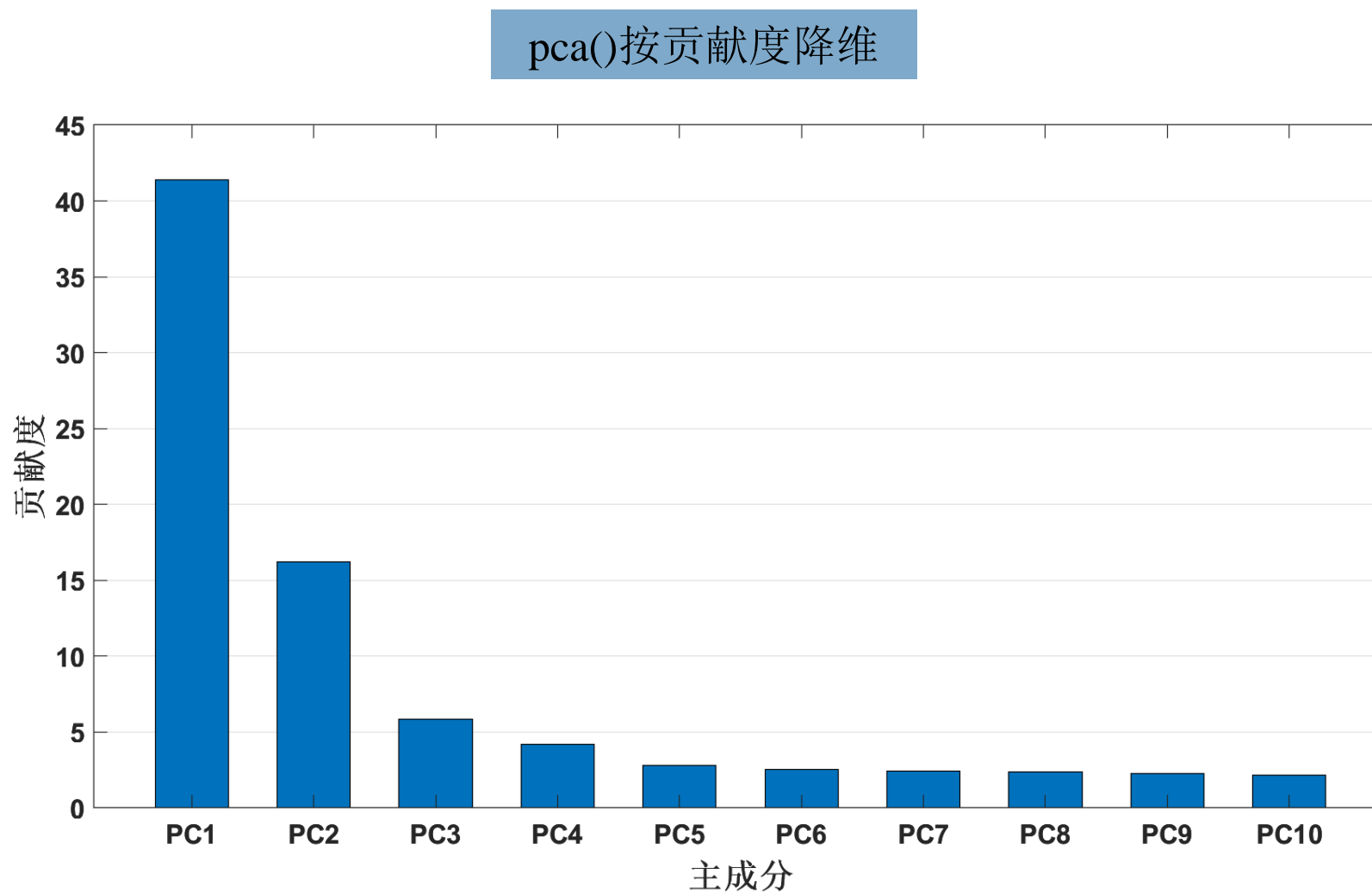
mapminmax()归一化



zscore()归一化



➤ PCA降维示例



➤ 主要任务安排如下：

□ 关联关系分析

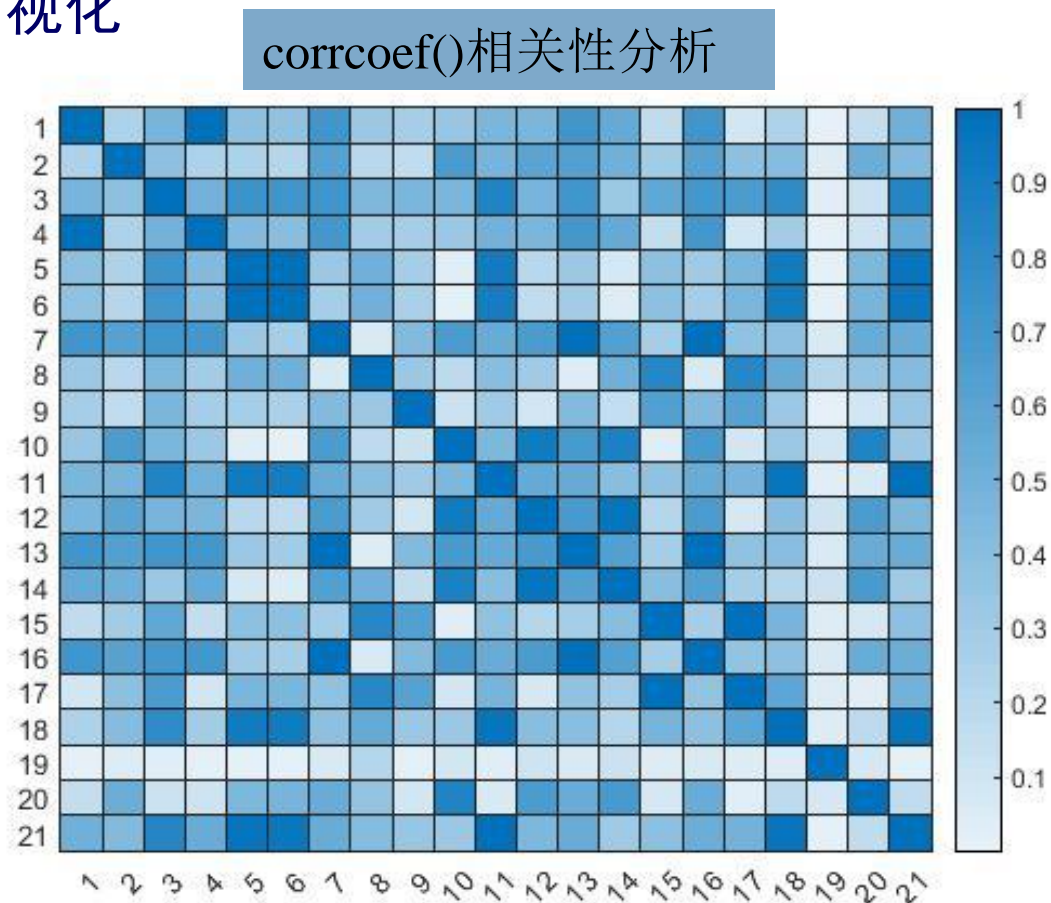
- 通过计算协方差、相关系数等相关性指标，计算变量的关联度，并通过可视化方法展示出来，找出**强关联变量**

□ 回归分析

- 1) 选择某种**成分变量**（如A物料、C物料等），根据关联关系分析，确定其相关**过程测量变量**
- 2) 采用回归分析方法，建立其他解释变量关于该**成分变量**的预测模型
- 3) 采用不同性能指标对分析模型进行评价，分析并对比不同回归模型性能

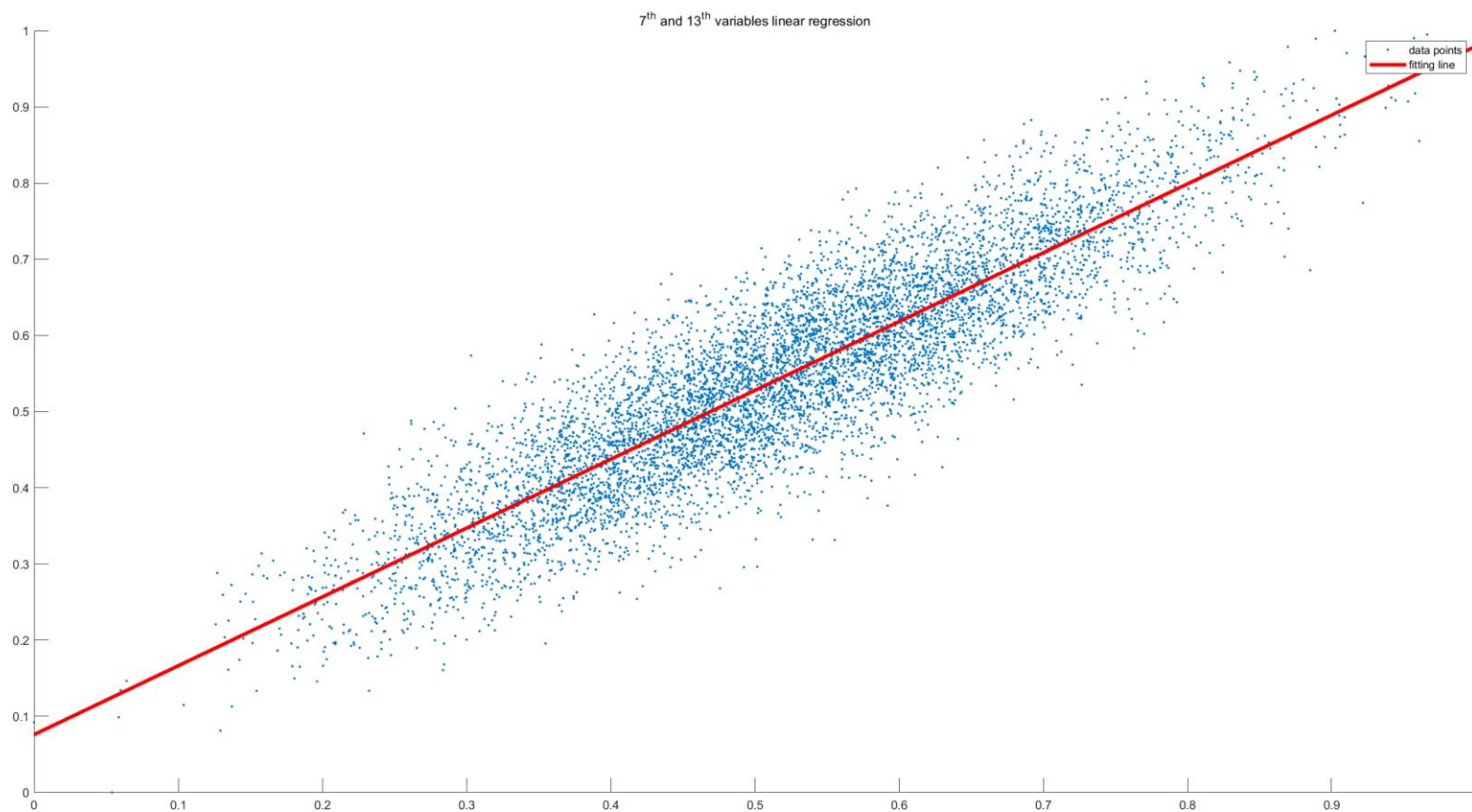
➤ 变量相关性分析示例

- ❑ 计算协方差(Covariance)、相关系数(Correlation Coefficient)等相关性指标，用图将结果可视化



➤ 变量散点图及回归建模示例

□ 采用最小二乘方法，建立回归模型



➤ 主要任务安排如下：

- 设定不同故障类型，故障类型选择不少于5个，获得带有故障标签的数据集
- 设计仿真实验
 - 选择合适的分类算法
 - 对数据进行划分，包括训练集和测试集
 - 在训练集上，训练分类器
 - 在测试集上，测试分析已训练好的分类器的性能
 - 考虑交叉验证等方法，对比不同分类算法用于故障诊断的效果

➤ 故障诊断性能评价可视化图表示例

□ 利用分类算法，完成TEP模型的故障诊断，并用尝试以下图表分析诊断性能

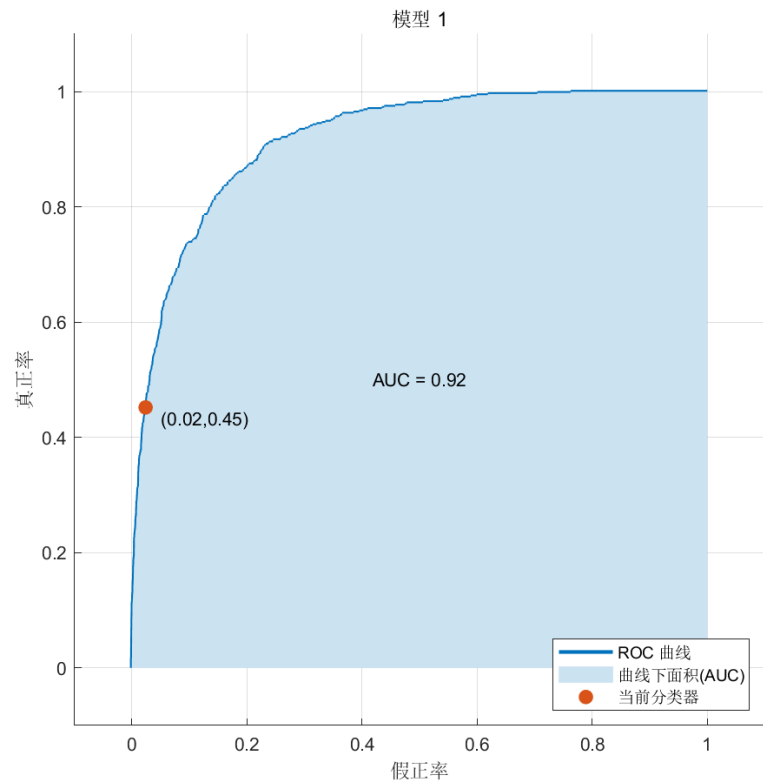


图 ROC曲线

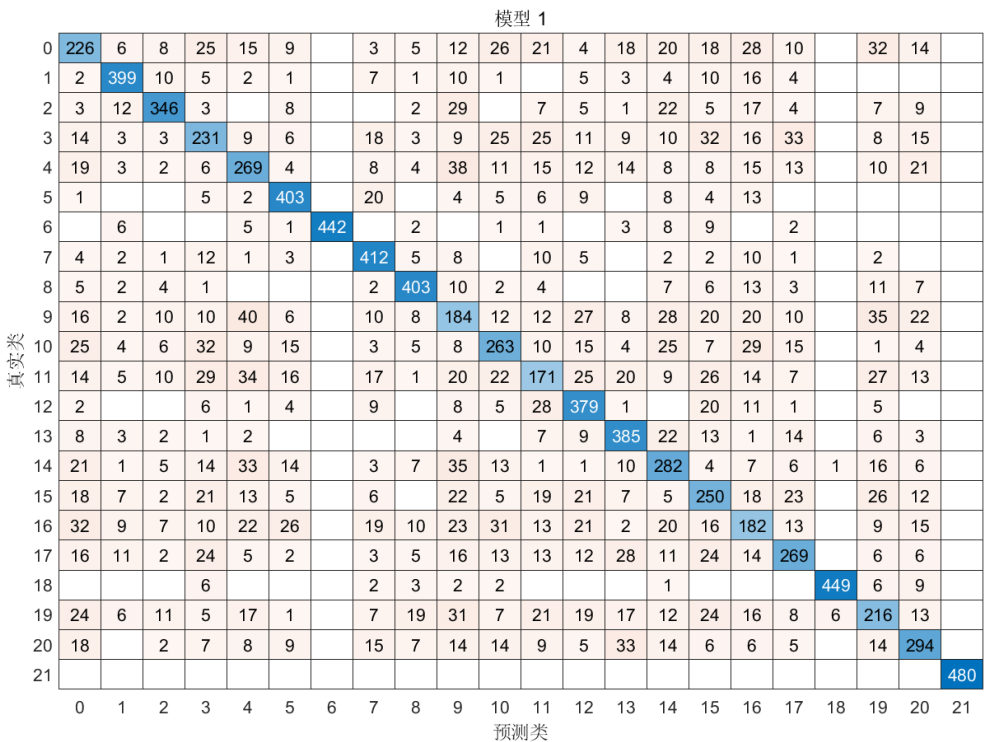


图 混淆矩阵

- ❑ 实验报告要图文并茂，对于实验结果可以用图和表格表示，同时要采用文字来解释和说明，注意对图和表格要进行编号和引用
- ❑ 实验报告排版规范美观，注意字体、字号、字距等的调整，避免出现前后不一致的情况，注意调整图表位置，使排版紧凑。
- ❑ 在实验报告最后要附上代码，且代码中要有一定注解。
- ❑ 实验分析和报告要独立完成