

智能制造过程大数据技术

Big Data Technology in Intelligent Manufacturing Process

第八讲：应用实例

Lecture 8: Application example

丁敏 dingmin@cug.edu.cn



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

- 高炉炼铁的基本原理
- 高炉参数的相关性分析
- 高炉数据预处理
- 高炉操作参数状态聚类
- 高炉炉况分类
- 高炉煤气利用率拟合与预测
- 本章小结



- 高炉炼铁的基本原理
- 高炉参数的相关性分析
- 高炉数据预处理
- 高炉操作参数状态聚类
- 高炉炉况分类
- 高炉煤气利用率拟合与预测
- 本章小结



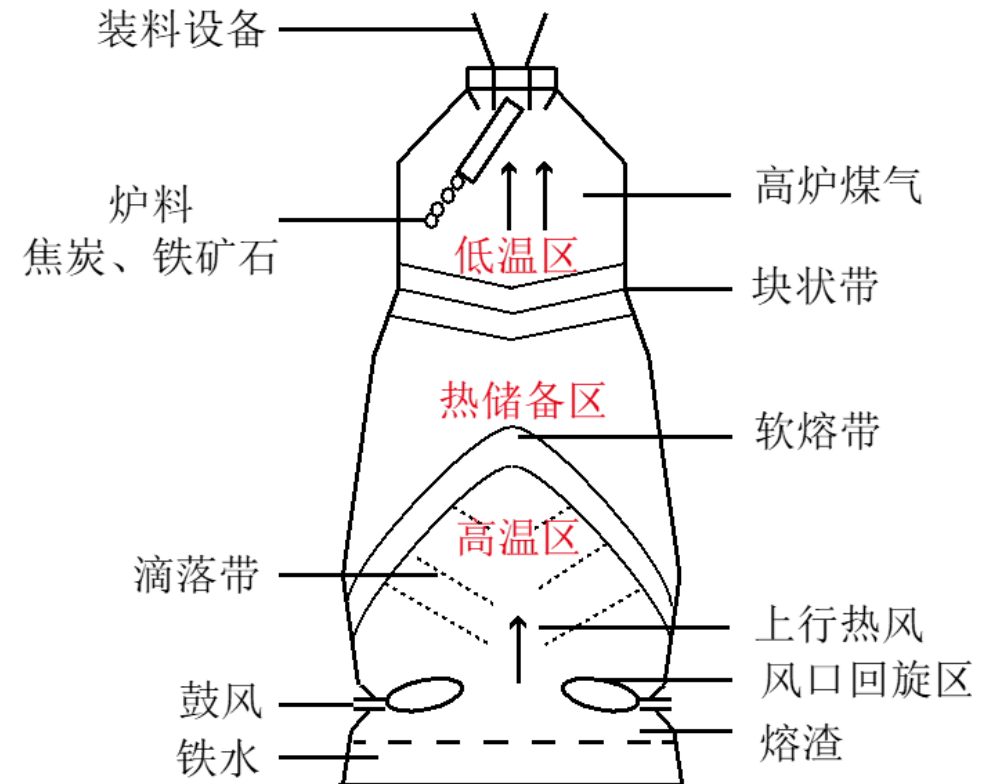
1 高炉炼铁基本原理

4

➤ 高炉炼铁的生产机理

□ 高炉炼铁是钢铁生产过程中的上游工序，是通过高炉中的一系列复杂反应，将铁矿石冶炼，还原成生铁的过程。

- ✓ 主体设备为高炉，是一个密闭的逆流式热交换竖炉，用高强度钢板作炉壳，壳内砌耐火砖内衬
- ✓ 在炉内堆积成料柱状的炉料，受逆流而上的高温还原气流的作用，不断地被加热、分解、还原、软化、熔融、滴落，并最终形成渣铁融体而分离



➤ 高炉炼铁的生产机理

- ✓ 在生产过程中，炉长需要根据不同的高炉冶炼指标去判断冶炼过程的好坏，并通过**调节不同的高炉操作**调节炉内的氧化还原反应，将高炉控制在期望的煤气流状态与热状态下，从而保证炉况保持在一定范围内的稳定状态下。——**主要依托人工经验**
- ✓ 高炉炉况运行情况的好坏决定了铁水的出炉质量，因此需要对高炉炉顶**布料**与炉底**送风**配合控制，以保持炉况的稳定。

必要操作

明确高炉操作参数与高炉状态参数之间的关系，建立合理的模型对煤气流分布进行调整与控制，并对不同高炉炉况进行相应的分类分析，针对炉况进行特定的调整。

➤ 高炉操作对高炉状态的影响

□ 高炉主要操作：布料操作与送风操作。两种操作从不同角度对高炉状态产生影响。

布料操作

- ✓ 布料操作通过改变高炉内固体原料分布、改变高炉内部煤气流发展与分布、调控高炉内部燃料比、调整热状态和物理化学反应来影响煤气利用率。
- ✓ 布料操作的调节主要是改变布料矩阵，其中包括改变布料溜槽的角度、布料圈数、料批重量等参数，主要是影响软融带和块状带的煤气流分布。

送风操作

- ✓ 送风操作通过在炉底风口鼓入热风并吹动焦炭，使其回旋燃烧形成初始煤气流。
- ✓ 送风主要对高炉内部煤气流初始分布、高炉底部热能的调节、高炉内部压力环境、煤气上升动能和高炉冶炼强度等进行调。常见的送风操作参数有风量、风压、富氧等。

➤ 高炉操作对高炉状态的影响

□ **高炉煤气利用率**能够实时反映煤气利用程度、煤气流分布状态、能源消耗程度以及铁水的质量和产量，是反映高炉整体状态的重要参数。

□ 高炉煤气利用率 $\eta_{CO} = \frac{V_{CO_2}}{V_{CO_2} + V_{CO}}$ V_{CO} : 一氧化碳含量
 V_{CO_2} : 二氧化碳含量

- ✓ 在一定范围内，炉煤气利用率越高，代表还原效率越高、气体分布越合理、高质量和低的高炉消耗
- ✓ 煤气利用率越低，则表示能量利用率低，冶炼状态越差。

- 高炉炼铁的基本原理
- 高炉参数的相关性分析
- 高炉数据预处理
- 高炉操作参数状态聚类
- 高炉炉况分类
- 高炉煤气利用率拟合与预测
- 本章小结



2 高炉参数的相关性分析

➤ 高炉操作参数间的相关性分析

- ❑ 问题：众多的高炉操作参数均作为模型输入会造成数据冗余
- ❑ 解决：利用皮尔逊相关性分析法进行参数选择

	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀
X ₁	1.00	-0.31	-0.33	0.13	-0.27	-0.14	0.25	0.25	0.01	-0.14
X ₂	-0.31	1.00	0.99	-0.19	0.97	0.03	-0.002	-0.008	-0.008	0.09
X ₃	-0.33	0.99	1.00	-0.02	0.97	0.03	-0.02	-0.03	-0.08	0.11
X ₄	0.13	-0.19	-0.21	1.00	-0.08	0.04	-0.11	-0.12	0.08	-0.14
X ₅	-0.28	0.97	0.97	-0.08	1.00	0.04	-0.01	-0.02	-0.08	0.08
X ₆	-0.14	0.03	0.03	0.04	-0.04	1.00	-0.04	-0.04	0.03	0.03
X ₇	0.23	0.01	-0.005	-0.1	0.003	-0.06	1.00	0.93	-0.18	0.20
X ₈	0.23	0.01	-0.01	-0.12	-0.001	-0.04	0.93	1.00	-0.21	0.22
X ₉	0.02	-0.08	-0.08	0.09	-0.08	0.03	-0.18	-0.21	1.00	-0.69
X ₁₀	-0.15	0.11	0.12	-0.15	0.10	0.05	0.220	0.22	-0.69	1.00

注：冷风流量（X₁）、冷风压力（X₂）、热风压力（X₃）、富氧流量（X₄）、富氧压力（X₅）、喷煤量（X₆）、边缘矿焦比3（X₇）、边缘矿焦比4（X₈）、中心焦比7（X₉）、中心焦比11（X₁₀）

➤ 高炉操作参数间的相关性分析

- 皮尔逊相关性： C_p （皮尔逊相关系数）的绝对值越大，两个时间序列之间的相关性越大。 C_p 为负值时，两个时间序列为负相关，其为正值时，两个时间序列为正相关。
- 相关性分析结果：
 - ✓ 冷风压力（ X_2 ）、热风压力（ X_3 ）以及富氧压力（ X_5 ）相关性高
 - ✓ 边缘矿焦比3（ X_7 ）与边缘矿焦比4（ X_8 ）相关性高
 - ✓ 中心焦比7（ X_9 ）与中心焦比11（ X_{10} ）相关性高
- 参数选择：热风压力（ X_3 ）、边缘矿焦比3（ X_7 ）、中心焦比11（ X_{10} ）、冷风流量（ X_1 ）、富氧流量（ X_4 ）、喷煤量（ X_6 ）

➤ 高炉操作参数与状态参数间的相关性分析

□ 选取影响高炉煤气利用率大小的关键操作参数，进一步做相关性分析。

高炉操作参数与状态参数皮尔逊相关性分析结果

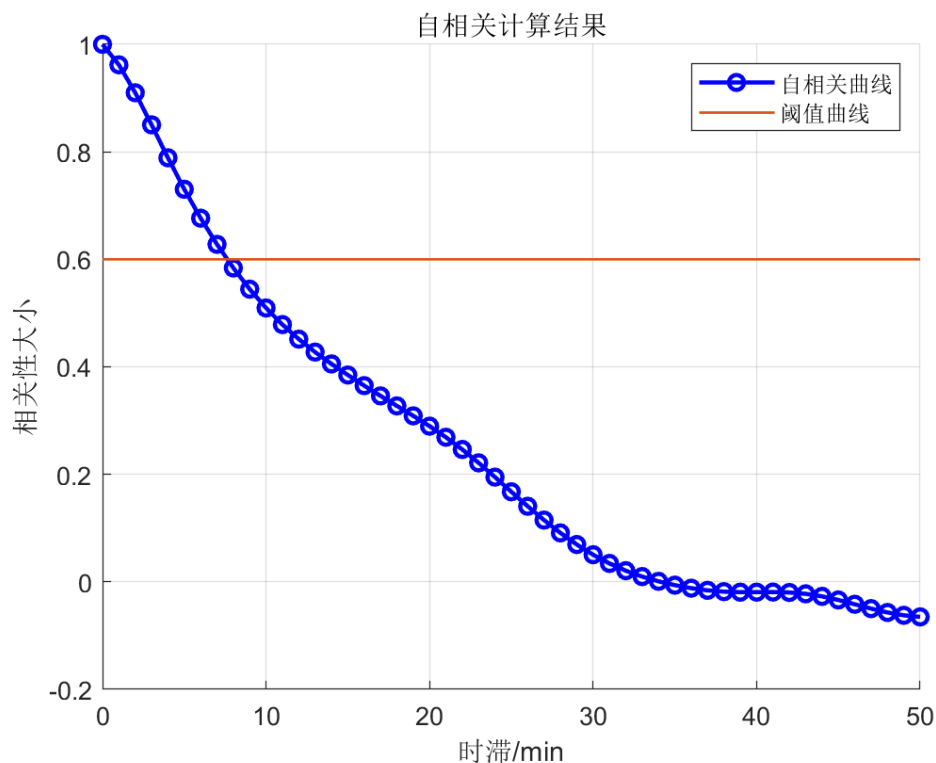
	X_1	X_3	X_4	X_6	X_7	X_{10}
煤气利用率	0.42	0.16	0.04	-0.05	0.12	-0.12

注：冷风流量（ X_1 ）、热风压力（ X_3 ）、富氧流量（ X_4 ）、喷煤量（ X_6 ）、边缘矿焦比3（ X_7 ）、中心焦比11（ X_{10} ）

□ 最终参数选择：冷风流量（ X_1 ）、热风压力（ X_3 ）、边缘矿焦比3（ X_7 ）、中心焦比11（ X_{10} ）

➤ 高炉多参数的自相关性分析

- ❑ 高炉炼铁生产状态参数具有一定的**自相关性**，即前一段时间的系统状态会对之后的状态产生明显的影响。



高炉煤气利用率自相关分析结果

- ✓ 随着时滞时间的延长，煤气利用率的ACF（自相关系数）值从1开始逐步下降
- ✓ 取**阈值水平为0.6**所对应的**前7个历史煤气利用率历史数据**作为预测模型的部分输入

- 高炉炼铁的基本原理
- 高炉参数的相关性分析
- **高炉数据预处理**
- 高炉操作参数状态聚类
- 高炉炉况分类
- 高炉煤气利用率拟合与预测
- 本章小结



➤ 高炉数据预处理

- **必要性**：高炉冶炼过程是一个非线性、大耦合、多变量的高难度冶金过程，参数测量和数据采集容易受到过程状态和环境因素影响而产生质量差的现场数据

➤ 高炉数据问题分析

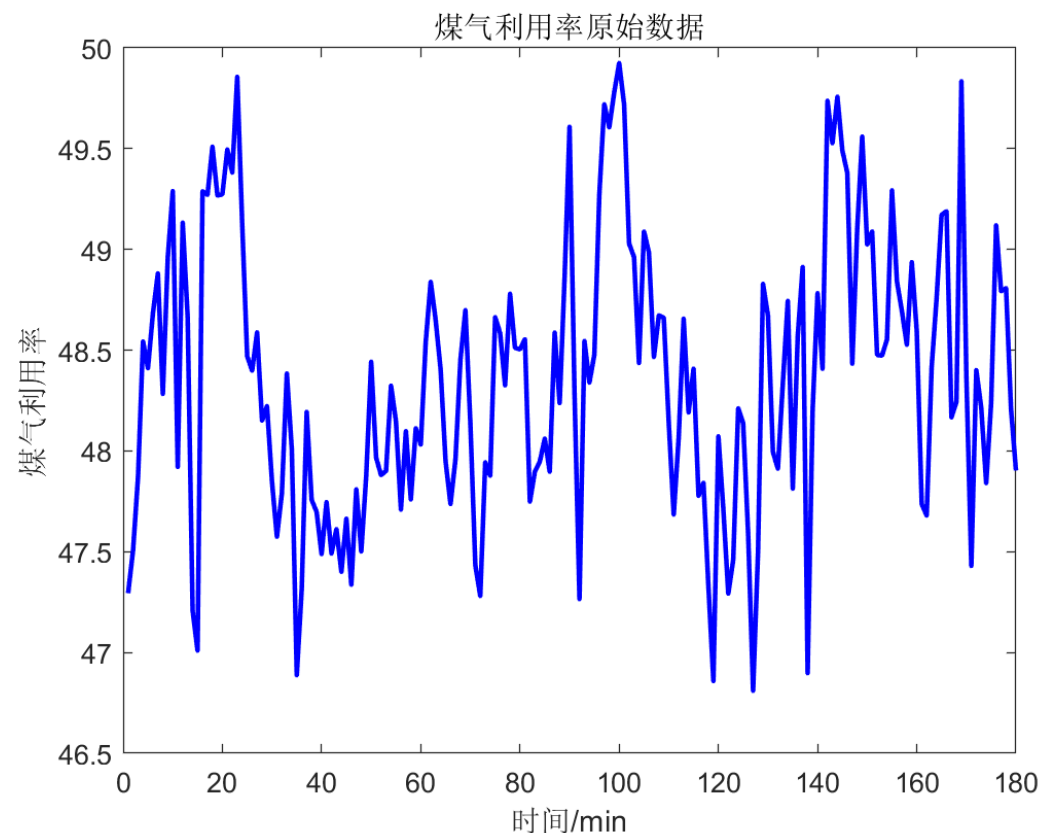
- 高炉炼铁连续性生产过程的属性导致数据具有一定的**时序性**
- 传感装置本身的敏感性易受外部条件干扰，可能产生**数据缺失**、**测量不准**等问题
- 多种物理化学反应同时进行，交叉影响，加剧了数据的**复杂性**和**耦合性**



数据去噪

➤ 高炉数据去噪处理

□ 高炉煤气利用率的原始数据



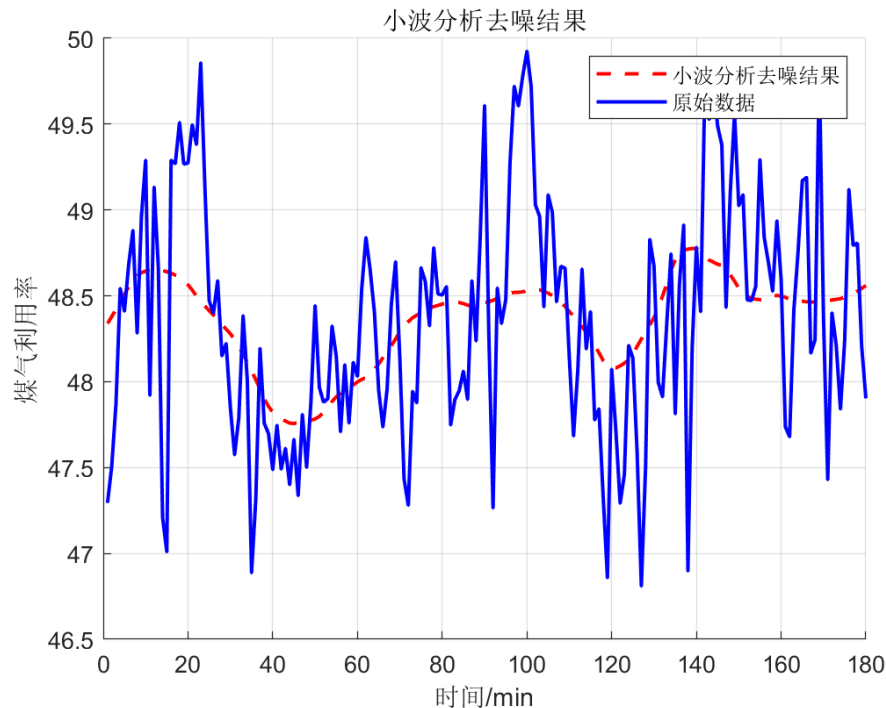
- ✓ 高炉煤气利用率的原始数据中具有明显**噪声波动特征**
- ✓ 高炉信号中的噪声不是平稳的白噪声，存在**多尖峰或突变**



小波去噪

➤ 高炉数据去噪处理

- ❑ **小波分析**能够将信号中各种不同的频率成分分解到互不重叠的频带上，可以保留具有重要意义的信号奇异点
- ❑ 对高炉现场采集多个参数的数据样本序列（冷风流量（ x_1 ）、热风压力（ x_3 ）、边缘矿焦比3（ x_7 ）、中心焦比11（ x_{10} ）、煤气利用率）分别进行**去噪处理**。以煤气利用率为例：



- ✓ 小波分析能够很有效地**去除**高炉现场数据的**尖峰和毛刺**，**圆滑曲线**的同时**保留**数据的**总体变化趋势**

➤ 高炉数据规范化

- 必要性：消除指标之间的量纲影响，以解决数据指标之间的可比性
- 方案：采用最小最大（Min-Max）规范化方法对原始数据进行线性变换，使结果值映射到 $[0, 1]$ 区间

➤ 送风操作相关参数规范化结果

序号	冷风流量	热风压力	煤气利用率	序号	冷风流量	热风压力	煤气利用率
1	0.9046	0.7267	0.6294	14	0.5435	0.9319	0.0412
2	0.6604	0.8495	0.2833	15	0.6736	0.6971	0.8034
3	0.6992	0.7357	0.1815	16	1	0.6030	0.8785
4	0.5471	0.7527	0.1371	17	0.8625	0.8898	0.8398
5	0.5039	0.6613	0.2091	18	0.6797	0.9615	1
6	0.4025	0.4991	0.3318	19	0.4768	0.8602	0.7535
7	0.3590	0.7052	0.5577	20	0.3964	1	0.5334
8	0.9154	0.8674	0.6981	21	0.2892	0.8396	0.5090
9	0.7698	0.8764	0.8091	22	0.7829	0.6694	0.5044
10	0.8409	1	0.3479	23	0.8730	0.8271	0.3773
11	0.8050	0.8656	0.7565	24	0.7723	0.9758	0
12	0.8035	0.7303	0.5988	25	0	0.2975	0.2895
13	0.3364	0	0.1088				

➤ 布料操作相关参数的规范化结果

序号	矿焦比3	中心焦比11	煤气利用率	序号	矿焦比3	中心焦比11	煤气利用率
1	0.7690	0.4122	0.6294	14	0.5491	0.4274	0.8785
2	0.6519	0.3916	0.5577	15	0.5026	0.4675	0.8398
3	0.5163	0.4006	0.5130	16	0.2853	0.4213	1
4	0.6578	0.3626	0.6047	17	0.3364	0.4642	0.7535
5	0.7906	0.4005	0.6716	18	0	0.4169	0.5334
6	0.5968	0.4027	0.4699	19	0.5010	0.5862	0.5090
7	0.6614	0.3815	0.6981	20	0.5125	0.9444	0.5731
8	0.5257	0.3796	0.8091	21	1	0.9049	0.3044
9	0.8802	0	0.3479	22	0.9953	0.9405	0.5044
10	0.6448	0.3925	0.7565	23	0.7602	1	0.3773
11	0.6423	0.3713	0.5988	24	0.5458	0.5344	0
12	0.8192	0.4484	0.1088	25	0.4039	0.4667	0.1439
13	0.6612	0.4559	0.8034				

- 高炉炼铁的基本原理
- 高炉参数的相关性分析
- 高炉数据预处理
- **高炉操作参数状态聚类**
- 高炉炉况分类
- 高炉煤气利用率拟合与预测
- 本章小结



➤ 高炉操作参数状态聚类

□ 必要性

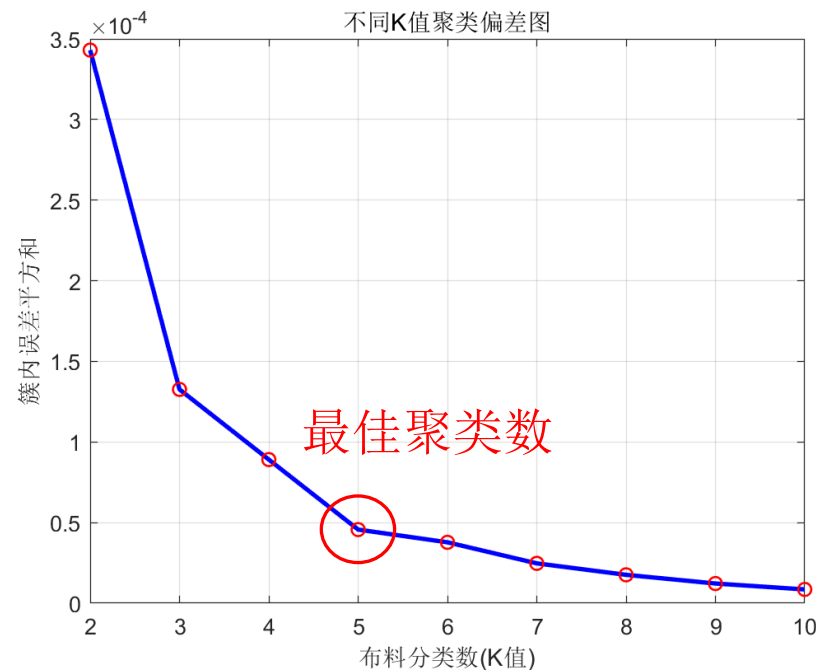
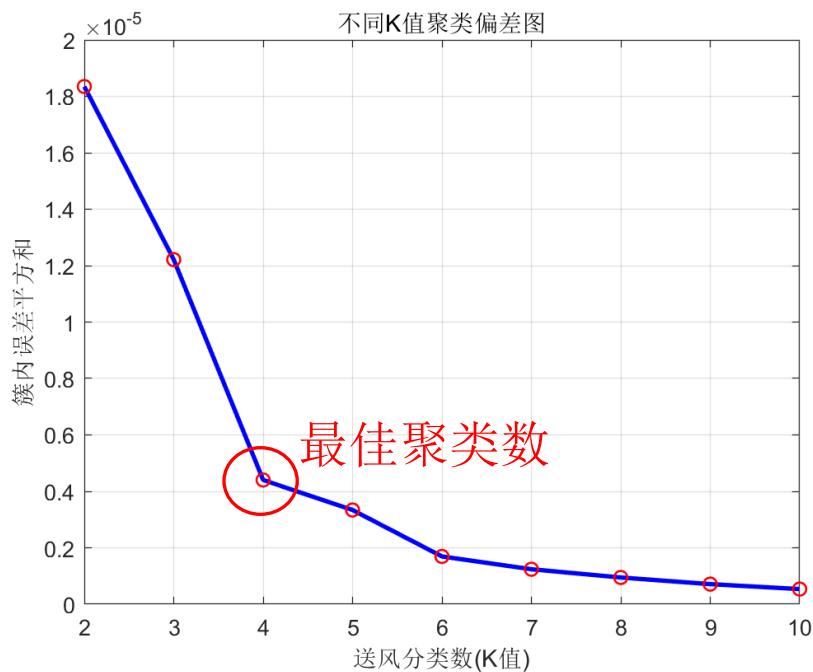
- ✓ 布料操作对高炉炉况控制作用大，对高炉生产状态的调节具有长时间尺度的滞后作用
- ✓ 送风操作对高炉煤气流状态和热状态的控制速度快，对高炉生产状态的调节具有短时间尺度的滞后作用
- ✓ 基于操作参数预测煤气利用率时，如果不区分时间尺度很难得到较好的预测结果



利用聚类方法分别处理布料与送风操作参数，为后续高炉煤气利用率的多尺度预测做建模准备

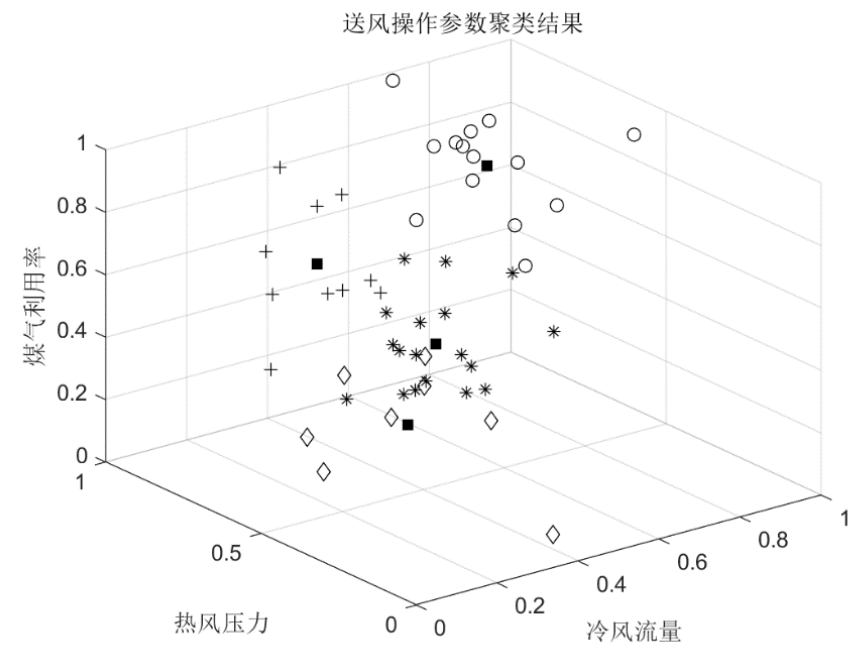
➤ 高炉操作参数状态聚类

- ❑ 输入参数选择：冷风流量 (x_1)、热风压力 (x_3)、边缘矿焦比3 (x_7)、中心焦比11 (x_{10}) 以及煤气利用率
- ❑ K-means聚类算法对送风与布料操作聚类
 - 肘部法确定最佳聚类中心



➤ 送风操作聚类

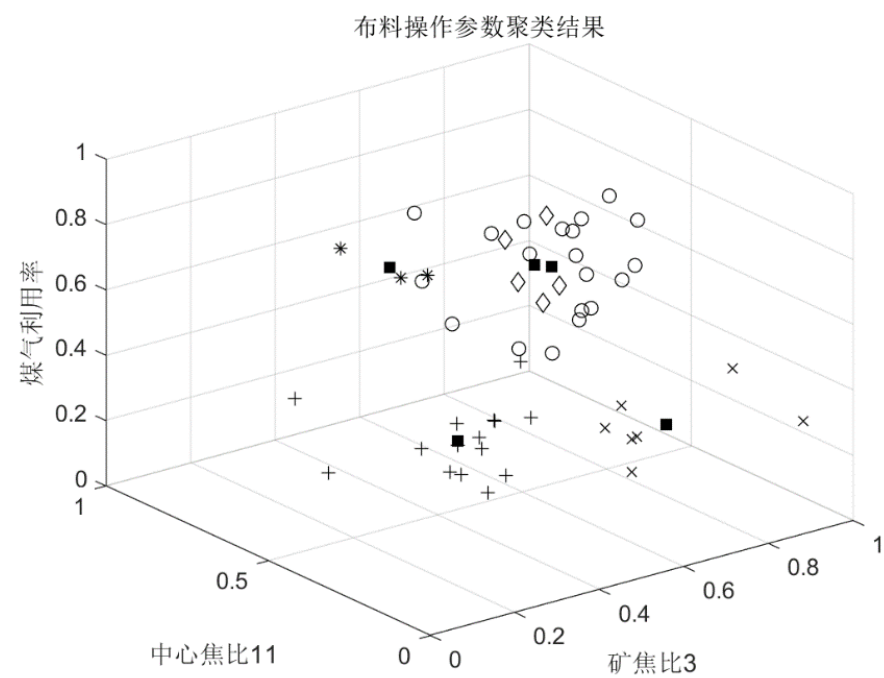
□ 采用K=4作为高炉送风与煤气利用率的类别数，进行K-means聚类。



簇	聚类中心	送风操作参数聚类结果
1	(0.7941,0.2987,0.6616)	{1,9,12,13,15,16,20,21,22,23,24,25,26,36}
2	(0.4056,0.8080,0.4178)	{6,17,31,32,43,44,47,48}
3	(0.7982,0.8483,0.7579)	{2,3,4,5,11,14,18,33,34,35,37,38,39,41,42,45,49,50}
4	(0.2803,0.2379,0.5604)	{7,8,10,19,27,28,29,30,40,46}

布料操作聚类

采用K=4作为高炉布料与煤气利用率的类别数，进行K-means聚类。



簇	聚类中心	送风操作参数聚类结果
1	(0.8275,0.6443,0.6118)	{ 3,4,5,6,14,17,18 }
2	(0.4019,0.9676,0.3532)	{30,31,37}
3	(0.9655,0.4232,0.4398)	{1,7,8,9,10,11,12,13,15,16,19,20,21,22,23,24,25,26,27,29}
4	(0.9405,0.1939,0.4586)	{2,28,38,39,40,41,42,43,44,45,46,47,48,49,50}

通过聚类将高炉送风与布料操作分别与煤气利用率聚为多个不同类别，可减少由于炉况波动而导致的类间数据变化过大对炉况调节造成的影响

- 高炉炼铁的基本原理
- 高炉参数的相关性分析
- 高炉数据预处理
- 高炉操作参数状态聚类
- **高炉炉况分类**
- 高炉煤气利用率拟合与预测
- 本章小结



➤ 高炉炉况分类

□ 炉况二分类：正常状态1与非正常状态0

□ 设置训练集为20、测试集为6

高炉分类训练数据集

训练集	x_7	x_{10}	x_1	x_3	高炉炉况
1	0	1	1	0	1
2	1	1	1	0	1
3	0	1	1	0	1
4	0	1	0	1	1
5	1	1	0	0	1
6	0	1	1	0	1
7	1	1	0	0	1
8	0	1	1	0	1
9	1	1	1	0	1
10	1	0	0	1	1

训练集	x_7	x_{10}	x_1	x_3	高炉炉况
1	0	0	0	0	0
2	0	0	0	1	0
3	0	1	0	0	0
4	1	1	0	0	0
5	1	0	0	1	0
6	0	1	0	0	0
7	1	0	0	0	0
8	0	1	0	1	0
9	0	0	0	1	0
10	1	0	1	0	0

➤ 高炉炉况分类

□ 朴素贝叶斯分类器测试结果

测试集	x_7	x_{10}	x_1	x_3	高炉炉况 (真实值)	高炉炉况 (预测值)
1	0	0	0	0	0	0
2	1	1	0	0	1	1
3	0	1	0	0	0	0
4	1	1	1	0	1	1
5	1	0	0	1	0	0
6	0	1	0	1	1	0

□ 分类整体精度

准确率	精准率	召回率	F度量
5/6	2/3	2/2	4/5

✓ 准确率均较高，反映了朴素贝叶斯分类器能满足高炉煤气利用率的基本分类要求

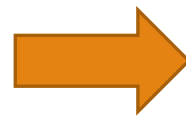
- 高炉炼铁的基本原理
- 高炉参数的相关性分析
- 高炉数据预处理
- 高炉操作参数状态聚类
- 高炉炉况分类
- 高炉煤气利用率拟合与预测
- 本章小结



➤ 高炉煤气利用率拟合与预测

□ 基于高炉煤气利用率预测模型主要分为机理模型和数据驱动模型

- 机理模型：研究高炉内部的化学反应，结合传热学、流体力学等学科来建立可以描述高炉内部反应的偏微分方程组并进行求解
- 数据驱动模型：由于检测方法的提高，为数据驱动模型提供了良好的数据基础



✓ 理想化的模型，假设条件过多，高炉的异常情况难以满足，且需要大量的计算，难以实现在线求解



采用最小二乘法和支持向量回归法，建立高炉煤气利用率的线性和非线性两种预测模型

➤ 高炉煤气利用率线性回归分析

□ 数据：某钢铁厂内体积为2800 m³的高炉连续五个月的现场数据

□ 数据描述：

✓ 归一化后的数据集 (x_i, y_i) $x_i \in$ 冷风流量 X_1 $y_i \in$ 煤气利用率 Y

□ 冷风流量 (X_1) 与煤气利用率之间的回归模型：
$$f(x) = a_1x^2 + a_2x + a_3$$

□ 原始数据集

x_i	0.4896	0.4994	0.5142	0.5341	0.5592	0.5886	0.6216	0.6571	0.6930	0.7283
y_i	0.7994	0.7952	0.7896	0.7868	0.7888	0.7902	0.7887	0.7809	0.7708	0.7606

➤ 最小二乘法估计

□ 二项式方程拟合结果

参数	a_1	a_2	a_3
结果	-0.5551	0.5448	0.6601

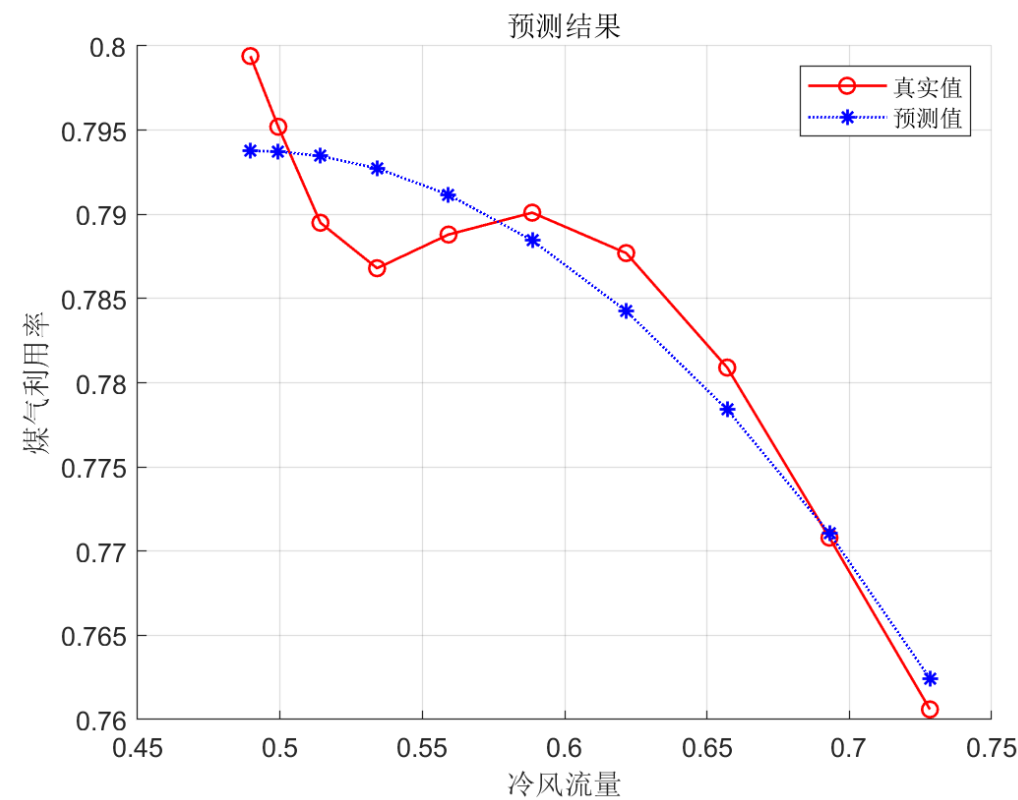
□ 高炉煤气利用率的预测模型为 $f(x) = -0.5551x^2 + 0.5448x + 0.6601$

□ 最小二乘拟合结果

x_i	0.4896	0.4994	0.5142	0.5341	0.5592	0.5886	0.6216	0.6571	0.6930	0.7283
y_i	0.7994	0.7952	0.7896	0.7868	0.7888	0.7902	0.7887	0.7809	0.7708	0.7606
拟合值	0.7938	0.7937	0.7935	0.7927	0.7912	0.7885	0.7843	0.7784	0.7711	0.7624

➤ 最小二乘法估计

□ 最小二乘拟合图



□ 模型性能指标

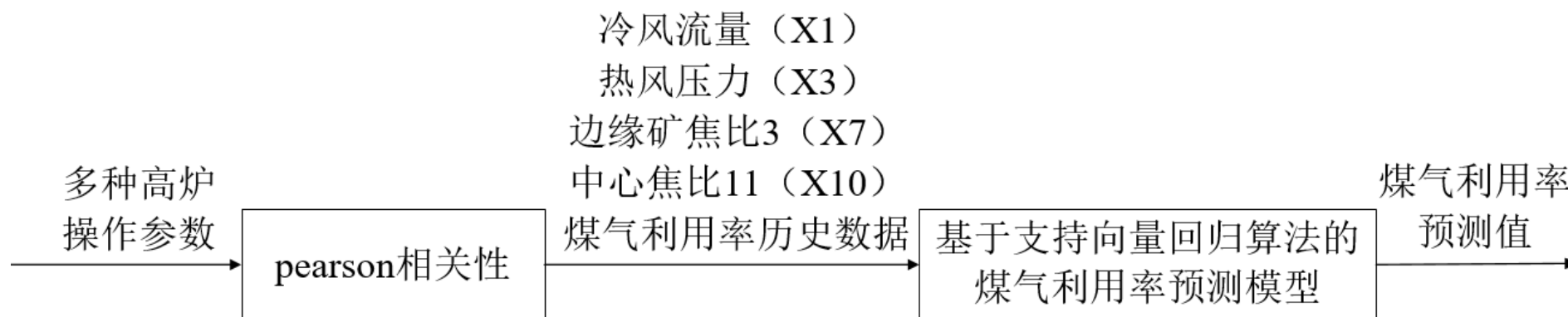
决定系数	标准差	均方根误差
89.41%	0.0034	0.0034

□ 结果分析

- ✓ 该回归拟合模型大体上能够对一段时间内的煤气利用率进行拟合，基本能够达到预期效果

➤ 高炉煤气利用率非线性回归分析

- ❑ 线性拟合问题：高炉冶炼过程中参数众多，反应复杂，多种参数相互影响，使得高炉炼铁过程呈现出**大非线性特性**。**线性回归方法不适用**
- ❑ 解决方案：非线性拟合模型——支持向量机
- ❑ 基本思路



6 高炉煤气利用率拟合与预测

34

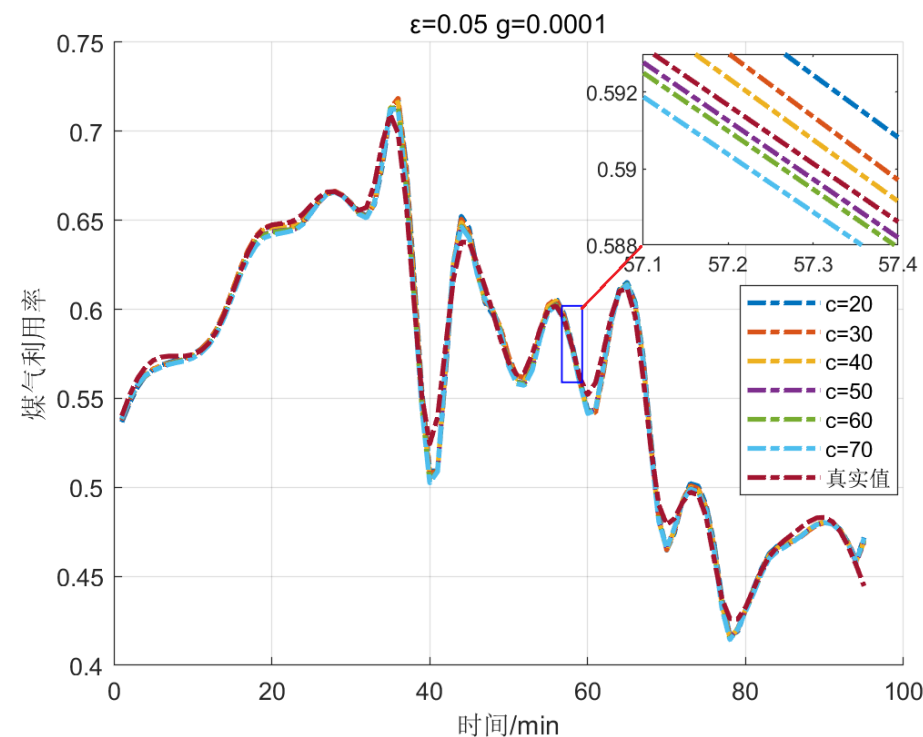
➤ 高炉煤气利用率非线性回归分析

□ 超参数选择： c ， ϵ 和 g

□ c 的测试结果

c	20	30	40	50	60	70
标准差	0.0077	0.0073	0.0069	0.0066	0.0066	0.0067
均方根误差	0.0078	0.0074	0.0071	0.0069	0.0070	0.0072
决定系数	98.89%	98.99%	99.08%	99.12%	99.11%	99.05%

✓ 综合考虑： $c=50$ 是较为合适的



6 高炉煤气利用率拟合与预测

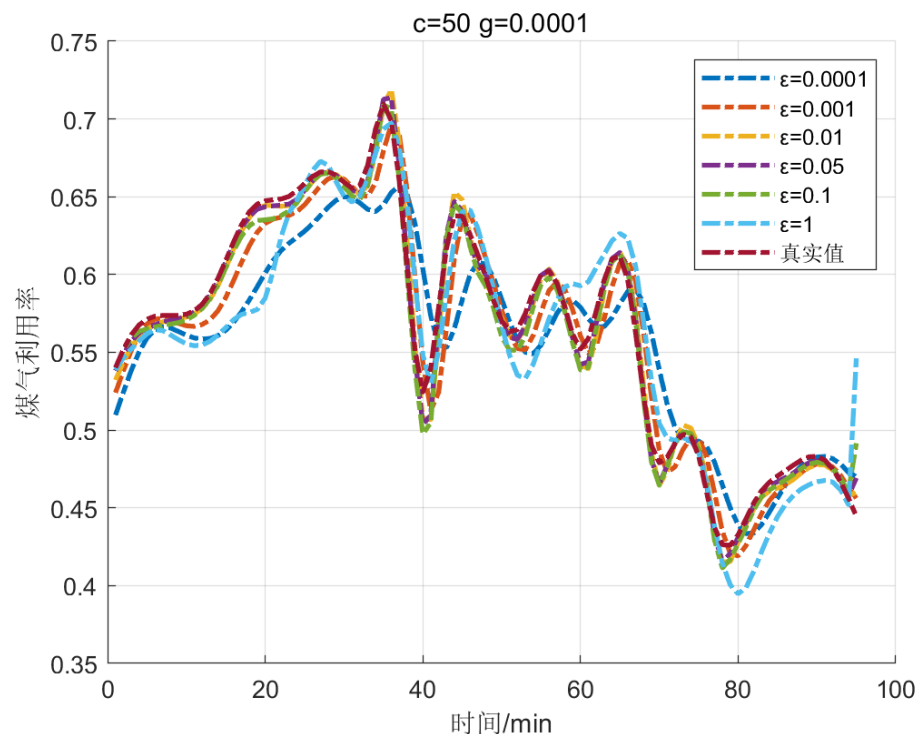
35

➤ 高炉煤气利用率非线性回归分析

□ ϵ 的测试结果

ϵ	0.0001	0.001	0.01	0.05	0.1	1
标准差	0.0317	0.00180	0.0083	0.0066	0.0085	0.0279
均方根误差	0.0330	0.00187	0.0088	0.0069	0.0098	0.0289
决定系数	80.30%	93.58%	98.58%	99.12%	98.23%	84.82%

✓ 综合考虑: $\epsilon=0.05$ 是较为合适的



6 高炉煤气利用率拟合与预测

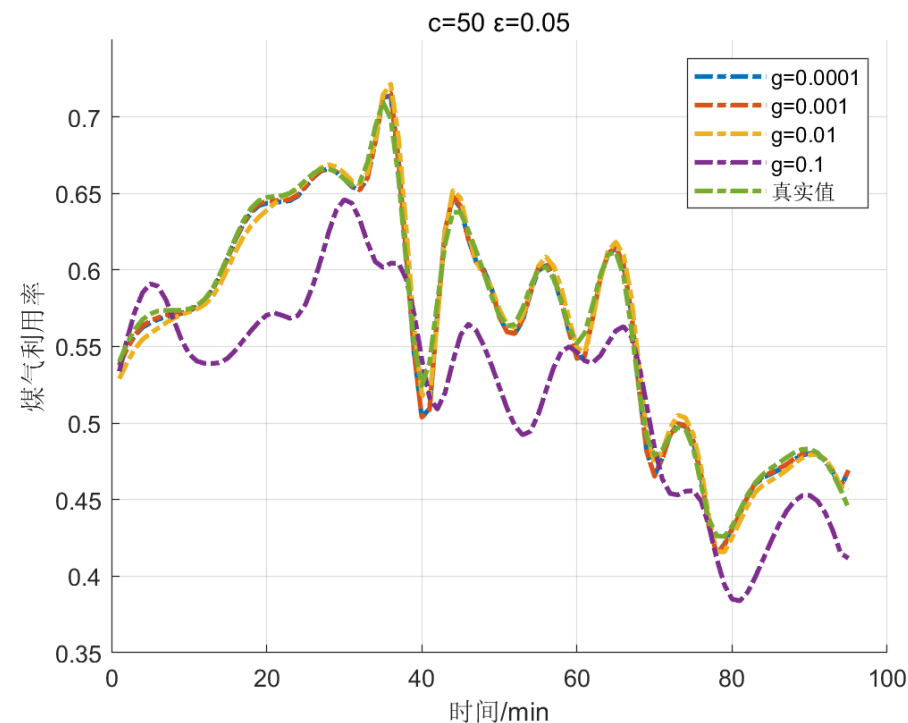
36

➤ 高炉煤气利用率非线性回归分析

□ g 的测试结果

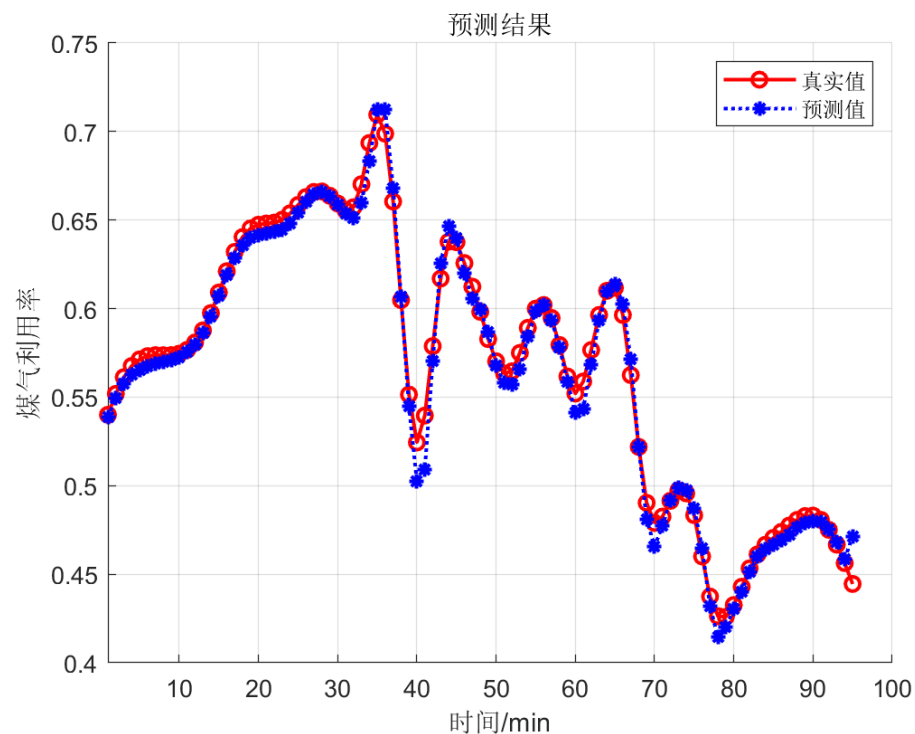
g	0.0001	0.001	0.01	0.05
标准差	0.0066	0.0068	0.0089	0.0319
均方根误差	0.0069	99.10%	0.0089	0.0526
决定系数	99.12%	99.10%	98.54%	61.54%

✓ 综合考虑: $g=0.001$ 是较为合适的



➤ 高炉煤气利用率非线性回归分析

□ 支持向量回归预测结果



决定系数	标准差	均方根误差
99.12%	0.0066	0.0069

✓ 评价指标都具有较好的数值，故该回归模型能够较好地对未来一段时间的煤气利用率进行预测

- 高炉炼铁的基本原理
- 高炉参数的相关性分析
- 高炉数据预处理
- 高炉操作参数状态聚类
- 高炉炉况分类
- 高炉煤气利用率拟合与预测
- **本章小结**



□ 聚类分析

- 利用K-means聚类算法对高炉不同布料操作参数进行距离聚类

□ 数据分类

- 利用贝叶斯分类模型完成的高炉煤气利用率的二分类问题

□ 线性回归

- 利用最小二乘线性回归模型完成相关参数估计
- 检验模型的适用性以及拟合性能

□ 非线性回归

- 利用支持向量非线性模型完成超参数求解
- 检验模型的适用性以及拟合性能

- 杨天钧, 张建良, 刘征建, 等. 低碳炼铁 势在必行[J]. 炼铁, 2021, 40(04):1-11.
- 周传典. 高炉炼铁生产技术手册. 北京: 冶金工业出版社, 2002.
- 吴敏, 曹卫华, 陈鑫. 复杂冶金过程智能控制, 科学出版社, 2016.
- 吴敏, 王昌军, 安剑奇, 等. 基于料面温度场的高炉煤气流分布识别方法[J]. 信息与控制, 2011, 40(01): 78-82.
- 吴敏, 聂卓赞, 曹卫华, 等. 面向高炉布料操作优化的在线信息检测方法及其应用(上)[J]. 冶金自动化, 2008(03): 6-9.
- 吴敏, 聂卓赞, 曹卫华, 等. 面向高炉布料操作优化的在线信息检测方法及其应用(下)[J]. 冶金自动化, 2008(04): 5-8+56.
- 曲飞, 吴敏, 曹卫华, 等. 基于支持向量机的高炉炉况诊断方法[J]. 钢铁, 2007(10): 17-19.
- 安剑奇, 陈易斐, 吴敏. 基于改进支持向量机的高炉一氧化碳利用率预测方法[J]. 化工学报, 2015, 66(01): 206-214.
- 陈少飞, 刘小杰, 李宏扬, 等. 高炉炼铁数据缺失处理研究初探[J]. 中国冶金, 2021, 31(02): 17-23.
- 李浩然, 邱彤. 基于因果分析的烧结生产状态预测模型[J]. 化工学报, 2021, 72(03): 1438-1446.
- 周志华. 机器学习. 北京: 清华大学出版社, 2016.
- An J Q, Shen X L, Wu M, et al. A multi-time-scale fusion prediction model for the gas utilization rate in a blast furnace[J]. Control Engineering Practice, 2019, 92: 104120.