

基于 Python 数据分析的控轧钢板中间坯温降规律研究

张建平^{1, 2}, 徐海健³, 庞宗旭^{1, 2}, 黄健^{1, 2}, 范刘群^{1, 2}

(1.海洋装备用金属材料及其应用国家重点实验室, 辽宁鞍山 114009; 2.鞍钢集团钢铁研究院, 辽宁鞍山 114009; 3.鞍钢股份公司中厚板事业部中厚板厂, 辽宁鞍山 114009)

摘要: 中厚板轧制过程中中间坯厚度对钢板性能和轧制效率具有重要影响。本文以工业生产数据为基础, 以提高轧制效率, 保证性能稳定为目标, 采用 Python 语言的数据分析方法, 开展中间坯厚度对温降速率影响的研究, 建立了基于机器学习的中间坯温降规律预测模型, 为产线生产工艺制定提供指导。模型经测试集的验证表明: 优化后的温降模型对不同钢种温降预测效果较高, 且模型泛化能力较强。

关键词: Python; 控制轧制; 中间坯; 机器学习; 温降预测

近年来, 伴随着国防军工及国家工程的发展需要, 我国中厚板生产水平大大提高, 特别是控制轧制和控制冷却(TMCP)工艺的应用, 使各大钢厂不断开发出具有优良性能的新钢种, TMCP 工艺已成为中厚板生产过程中不可或缺的生产技术^[1]。然而, 控制轧制比传统热轧所需生产时间更长, 特别是两阶段轧制过程中, 完成第一阶段再结晶区轧制后, 钢板需进行待温, 达到规定温度后才能进行第二阶段轧制, 大大降低产线生产效率^[2], 对于单机架中厚板产线而言, 为减少因待温时间造成的产量损失, 往往采用批轧方式进行控轧钢板的生产^[3]。批轧过程中, 批轧块数、辊道空间以及待温时间的匹配关系较为复杂, 会对轧制节奏产生一定影响^[4]。而且, 由于钢板规格及轧制道次分配问题, 同一轧制批次内钢板待温时间不同, 无法保证所有钢板都能在目标温度内完成轧制, 导致性能出现波动。

实际生产过程中, 轧钢操作人员经常根据主观经验决定批轧块数, 这种操作方式既无理论依据也无法实现高效轧制方法的普及。近年来, 机器学习技术在材料研发和工业生产中的应用发展迅速, 利用机器学习进行影响因素相关性分析和相关结果预测, 进而指导生产越来越受到重视^[5], 为通过机器学习建立温降模型进而指导批轧策略制定和高效生产提供了新思路。因此, 本文旨在利用产线生产数据, 建立符合产线生产实际的控轧钢板的中间坯待温模型, 这对提高产线生产效率保证性能稳定具有重要作用。

1 设计思路

鞍钢股份公司中厚板厂产线属于单机架产线, 控轧钢板批轧块数和待温时间因人而异, 无法形成生产技术规程。本文基于 Python 语言的数据分析方法, 结合 Numpy 和 Pandas 库, 整合轧制实绩、轧制道次及天气数据, 通过数据清洗, 过滤非控轧钢板和其他无效数据, 获得控轧钢板粗轧结束道次和精轧开始道次数据; 利用 Matplotlib 库, 实现对所获得数据相关性分析的可视化, 直观确定主要影响因素, 并建立回归模型; 最后利用机器学习方法, 实现控轧钢板中间坯待温时间预测, 指导产线批轧策略制定和工艺调整。

2 数据收集与处理

本文数据集采集于鞍钢股份公司中厚板事业部中厚板厂, 原数据集包含产线 2022 年生产“轧制实绩”和“轧制道次实绩”两个 EXCEL 文件, 其中“轧制实绩”由“物料号”、“生产日期”以及“板坯厚度”等 122 列组成, 总计 13 余万行数据, “轧制道次实绩”由“物料号”、“道次开始时刻”及“道次平均温度”等 33 列组成, 总计 150 余万行记录。

以上述数据为基础, 构建 DataFrame 二维数据表, 调用 pd.read_excel()方法分别读取“轧制实绩”和“轧制道次实绩”两个 EXCEL 表, 通过 Pandas 库的 merge()方法, 将“轧制实绩”与“轧制道次实绩”以“物料号”为连接键进行整合, 赋值给“df_merged”。网络爬虫爬取鞍山市 2022 年全年天气温度, 赋值给“df_weather”,

并分别以“df_weather”中的“天气日期”和“df_merged”中的“生产日期”为连接键，通过 merge() 函数将“df_weather”整合进“df_merged”。利用 Pandas 库的 pop() 方法删除表多余列，仅保留“df_merged”内字段名为“物料号”、“钢牌号”、“出口厚度测量值”、“出口宽度测量值”、“道次结束时刻”、“道次平均温度”以及“天气温度”的数据。利用 Pandas 库中的 sort_values() 方法，以“df_merged”的“道次结束时刻”作为排序字段进行升序排序，利用 Pandas 库的 drop() 方法清洗掉非控轧钢板的无效数据，同时去除控轧钢板无用道次数据，最终只保留仅包含控轧钢板的粗轧结束道次和精轧开始道次的信息。此时，每个“物料号”钢板仅存粗轧结束和精轧开始道次的相关信息，利用索引号将粗轧和精轧道次信息分别赋值给“df_R”及“df_F”两个二维数据表，并将“df_R”内“出口厚度测量值”重命名为“中间坯厚度”，“出口宽度测量值”命名为“中间坯宽度”，“道次结束时刻”重命名为“粗轧结束时刻”，“道次平均温度”重命名为“开始待温温度”。同理，“df_F”内“道次结束时刻”重命名为“精轧开始时刻”，“道次平均温度”重命名为“结束待温温度”。再次调用 merge() 函数，以“物料号”为连接键将“df_R”和“df_F”整合为“df_Data”。在“df_Data”内增加“待温时间”列，其中某物料号钢板的“待温时间”=“精轧开始时刻”-“粗轧结束时刻”。删除“df_Data”内除“钢牌号”、“中间坯厚度”、“中间坯宽度”、“天气温度”、“开始待温温度”、“结束待温温度”、“待温时间”及“天气温度”以外的数据，通过“df_Data.to_excel(“2022 年产线生产数据.xlsx”)”命令，将最终数据存于 EXCEL 表内用以建立温降模型。

3 可视化分析

通过 Pandas 库的 read_excel() 方法读取“2022 年产线生产数据.xlsx”原数据，本文以 Q355B 钢种为例，进行中间坯温降模型建立，因此调用 isin() 方法筛选出“钢牌号”为 Q355B 的钢板轧制信息。导入 matplotlib 和 seaborn 库，值得注意的是，需要通过 plt.rcParams['font.sans-serif']='SimHei' 命令解决可视化工具无法解释中文问题，通过 plt.rcParams['axes.unicode_minus']=False 解决正负号显示问题。调用 matplotlib 库的 title() 方法设置标题，调用 corr() 方法绘制相关系数热力图，本文选择皮尔逊相关系数。最后，对热力图进行颜色、字体及色条阈值等进行设置和调整，结果如图 1。

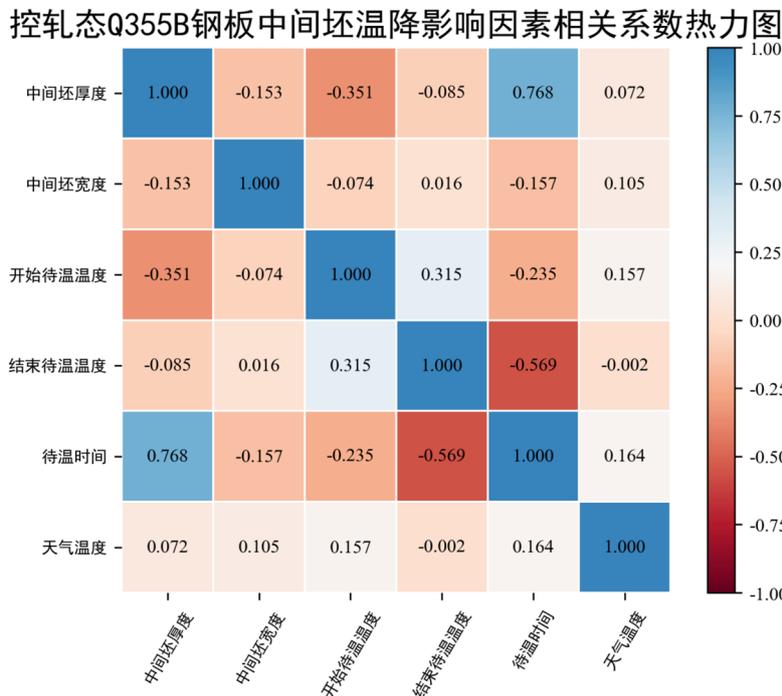


图 1 2022 年中厚板厂控轧态 Q355B 钢板中间坯温降影响因素相关系数热力图

结果表明，待温时间与“中间坯厚度”相关系数为 0.768，有强的正相关关系；与“结束待温温度”相关系数为-0.569，有中等程度的负相关关系；与“天气温度”相关系数为 0.164，表现为极弱的正相关关系，与“中间坯宽度”的相关系数为-0.157，表现为极弱的负相关关系；与“开始待温温度”相关系数为-0.235，存在弱的负相关关系。

5 模型建立及检验

导入 `sklearn.linear_model` 模块,以“待温时间”为因变量 Y,分别以“中间坯厚度”(X1)、“中间坯宽度”(X2)、“开始待温温度”(X3)、“结束待温温度”(X4)、“天气温度”(X5)为自变量建立线性回归模型,调用 `score()`方法对经拟合后的最终回归模型进行 R^2 评分,评分结果为 0.8913,具有较高置信度。

利用该模型对 2023 年生产数据中 20 块控轧态 Q355B 钢板的中间坯待温时间进行预测,并与实际生产待温时间进行对比,结果实例如表 1,结果表明该模型预测待温时间与实际待温时间偏差较小,预测偏差率均介于 $\pm 5\%$ (偏差率=预测偏差量/实际待温时间)。

表 1 2023 年生产数据控轧态 Q355B 钢板中间坯待温时间预测与实际值

序号	中间坯厚度 (mm)	中间坯宽度 (mm)	开始待温温度 (°C)	结束待温温度 (°C)	天气温度 (°C)	预测待温 时间(s)	实际待温 时间(s)	预测偏差 量(s)	偏差率 (%)
1	39.8863	2309.2	1031.9657	899.431	-7.5	98	102	-4	-3.92
2	39.8863	2586.3	990.8379	864.1899	-5	141	138	3	2.17
3	39.8863	2283.4	1000.2688	879.1344	12.5	133	130	3	2.31
4	39.8863	2535.1	1033.0349	856.5919	20.5	188	188	0	0.00
5	39.8863	2283	996.773	817.312	28	221	227	-6	-2.64
6	62.3497	2326.5	1022.2255	890.3568	-1.5	183	180	3	1.67
7	62.3497	2270.7	977.3154	860.4687	4	251	256	-5	-1.95
8	62.3497	2532.3	1008.7623	843.6432	23.5	335	332	3	0.90
9	62.3497	2336.6	1017.4263	934.8465	14.5	142	139	3	2.16
10	79.8703	2611.2	957.9734	807.5013	-7.5	399	395	4	1.01
11	79.8834	2281.2	960.6006	825.95	-5.5	399	391	8	2.05
12	79.8766	2623.4	978.3367	855.0995	5	279	280	-1	-0.36
13	79.8766	2485.1	987.9785	837.9476	16	415	412	3	0.73
14	79.8766	2176.3	1000.181	858.9937	23	336	342	-6	-1.75
15	79.8766	2559.7	978.5552	829.8297	30	460	458	2	0.44
16	99.9082	2640.1	977.1101	892.425	-9	247	242	5	2.07
17	99.9082	2311.4	978.4896	904.6315	-4.5	272	279	-7	-2.51
18	99.9082	2340.8	994.5427	839.8113	6.5	566	561	5	0.89
19	99.9082	2316.7	991.5183	857.3568	16	418	426	-8	-1.88
20	99.9082	2507.7	967.6593	856.0625	24	402	408	-6	-1.47

6 结论

(1) 中间坯厚度对钢板待温时间有强的正相关关系,结束待温温度与待温时间表现为中等程度的负相关关系。

(2) 基于 Python 语言的数据分析方法建立的控轧钢板中间坯温降模型 R^2 评分达到 0.8913,待温时间预测偏差率介于 $\pm 5\%$ 之间,模型具有较高的预测精度。

参考文献:

- [1] 孙震,高朋,刘振宇.待温厚度对一种 Mn 系 HSLA 钢组织及强韧性的影响[J].塑性工程学报,2021,28(8): 161-168.
- [2] 霍文丰.中厚板轧制中间冷却过程控制模型研究与应用[D].东北大学,2009.
- [3] 付颖.中厚板单机架生产效率的优化[J].世界有色金属,2016(1): 118-119.
- [4] 矫志杰,胡贤磊,赵忠等.首钢中厚板轧机轧制节奏的控制[J].钢铁研究学报,2005(1): 72-76.
- [5] 商春磊,王传军,刘文月等.数据驱动的经验辅助管线钢产线落锤撕裂韧性内禀特征关联[J].工程科学学报,2023,45(8): 1390-1399.