摘要

一体化压铸，英文文献中常称 Integrated Die Casting，是新能源汽车大型结构件，如后地板、前舱一体化构件、电池包壳体等，实现轻量化与集成化的重要制造路径。该工艺在超大型压铸装备上完成充型、增压补缩与凝固成形，具有工艺窗口窄、过程变量强耦合、质量波动传播快等特点；传统以离线抽检与经验调参为主的质控方式，难以及时覆盖“模次级”的过程风险。面向工程现场对过程可视化推演与质量量化预测的需求，本文构建“数字孪生推演—工艺窗口治理—数据驱动质量预测”一体化软件架构：以熔体温度、模具温度、慢/快压射速度、增压与型腔真空度等关键工艺参数为主线，在孪生侧实现参数微调、派生指标计算与多视图可视化；在模型侧基于监督学习建立由过程特征至缺陷风险等质量输出的映射，并以 JSON 配置的工艺窗口实现参数越界提示与条形图相对位置编码，将规则边界与数据驱动结果集成于统一交互界面。系统基于 Python 3.10 实现，可视化采用 matplotlib 绘制曲线与统计图；机器学习采用 scikit-learn，模型工件与元数据以 joblib/json 管理，形成自数据管理、模型训练、批量预测、结果导出至孪生监控的闭环。在示例数据集样本量有限、泛化能力不足的情形下，如验证集决定系数为负、风险分布近似常数等，本文从统计学习视角分析其成因，并讨论工业部署阶段在数据采集、标签体系、特征工程、模型迭代与在线更新等方面的改进路径。研究成果可为新能源汽车一体化压铸场景的数字化质量管理提供系统设计参考。文中对工艺窗口归一化、孪生派生指标与回归评价指标等给出了相应的数学表达式。

从工程边界看，本文着重阐明模块职责划分、数据与配置文件契约、训练推理与可视化闭环，以及小样本条件下评价指标的合理解释方式，便于在课程设计、企业预研或产线试点中按需裁剪与替换实现。摘要后文各章对上述要点分模块展开，并在第 7 章讨论与工业现场对接时的数据、标签与在线集成约束。全文在保持公式与字段说明一致的前提下，力图使“界面操作—数据文件—模型工件”形成可追溯链条，以支撑论文评审与软件验收中对可复现性的要求。

关键词：新能源汽车；一体化压铸；数字孪生；质量预测；工艺窗口；机器学习；过程数据

Abstract

Integrated die casting has become a key route for manufacturing large aluminum structural parts in new energy vehicles (NEVs), yet its quality is sensitive to tightly coupled process variables and narrow operating windows. This paper presents a unified software design that combines digital-twin what-if deduction, process-window governance, and data-driven quality prediction. Critical parameters such as melt/mold temperatures, slow/fast shot speeds, intensification pressure, and cavity vacuum are mapped into a twin layer for interactive simulation and visualization, while a supervised learning module predicts quality-related risk indicators from engineered features. Process limits are externalized as JSON configuration to enable consistent boundary checks and interpretable bar encodings. The implementation targets Python 3.10; charts rely on matplotlib. Together with scikit-learn and standard artifact formats (joblib/json), the system closes the loop from CSV data ingestion to model training, batch inference, exportable reports, and twin monitoring. Limitations under extremely small demonstration datasets are analyzed, and industrialization paths involving richer data, calibrated labels, and online integration (e.g., OPC) are outlined. The proposed architecture serves as a reference for digital quality management in NEV integrated die casting.

Keywords: new energy vehicle; integrated die casting; digital twin; quality prediction; process window; machine learning; process data

目录

摘要 1
Abstract 2
第 1 章 绪论 7
1.1 研究背景与意义 7
1.1.1 新能源汽车结构集成与一体化压铸的产业动因 7
1.1.2 质量管控从离线抽检走向数据驱动的内在需求 8
1.1.3 研究目的与工程价值概述 9
1.2 国内外研究现状与评述 9
1.2.1 压铸过程建模与仿真研究现状 9
1.2.2 数字孪生在制造过程的应用模式 10
1.2.3 机器学习在铸造缺陷预测中的研究现状 10
1.2.4 现有研究的不足与本文定位 11
1.3 研究内容与技术路线 11
1.3.1 研究内容 11
1.3.2 技术路线 12
1.4 论文结构安排 12
第 2 章 一体化压铸过程机理与数字孪生质量管控相关研究 13
2.1 高压压铸与一体化压铸的工艺特点 13
2.1.1 压铸工艺流程概述 13
2.1.2 典型缺陷类型与过程诱因 13
2.2 数字孪生用于压铸质量管控的关键要素 14
2.2.1 孪生对象与边界 14
2.2.2 同步、推演与可视化 14
2.3 数据驱动质量预测的基本范式 15
2.3.1 监督学习问题表述 15
2.3.2 特征工程与可解释性 15
2.4 本章小结 16
2.5 面向工程落地的关键约束与伦理边界 16

第 3 章 系统需求分析、设计原则与总体架构 18
3.1 需求分析 18
3.1.1 功能需求 18
3.1.2 非功能需求 18
3.2 设计原则 18
3.2.1 分层与解耦 18
3.2.2 规则与学习互补 19
3.3 总体架构 19
3.3.1 逻辑架构 19
3.3.2 模块映射与实现目录的对应关系 19
3.3.3 系统分层与功能模块一览 19
3.4 本章小结 20
第 4 章 工艺窗口约束与数据管理层设计 22
4.1 工艺窗口的概念与作用 22
4.1.1 从经验区间到可执行规则 22
4.1.2 配置外置与可追溯 22
4.1.3 工艺窗口相对位置与越界判定的计算 22
4.2 工艺窗口参数项与示例取值 23
4.2.1 参数项与物理含义 23
4.2.2 示例窗口 23
4.3 数据管理模块设计 24
4.3.1 数据字段与单位一致性 24
4.3.2 数据质量与预处理策略 25
4.4 本章小结 25
第 5 章 数字孪生推演与可视化模块设计 26
5.1 交互式推演与人机界面组织 26
5.1.1 模次选择与参数来源 26
5.1.2 滑条微调与实时反馈 26
5.2 压铸单元孪生监控与设备信息展示 27

5.2.1 设备型号与能力参数 27
5.2.2 状态量与滑条一致性 27
5.3 孪生派生指标设计 27
5.3.1 充型指数 Filling Index 27
5.3.2 气孔代理 Porosity Proxy 27
5.3.3 孪生派生指标的量化计算示例 27
5.4 多视图可视化设计 28
5.4.1 孪生可视化子视图汇总 29
5.5 本章小结 31
第 6 章 铸件质量预测模型与评价体系设计 32
6.1 预测任务定义与输出语义 32
6.1.1 回归型风险输出的适用场景 32
6.1.2 分类与异常检测扩展 32
6.2 特征集合设计 32
6.2.1 原始工艺特征 32
6.2.2 派生特征：孪生输出进入学习闭环 33
6.2.3 特征标准化与数据集划分 33
6.2.4 特征标准化与预测模型的矩阵形式 33
6.3 模型训练、评估与工件管理 34
6.3.1 训练流程 34
6.3.2 指标解释：MAE 与 R^2 34
6.3.3 特征重要性 34
6.3.4 回归评价指标的计算公式 34
6.3.5 模型选择与偏差—方差权衡简述 35
6.4 批量预测与结果导出 36
6.4.1 批次表与风险分布图 36
6.4.2 导出 CSV 的工程意义 36
6.5 本章小结 36
第 7 章 软件实现、运行环境、功能验证与讨论 37

7.1 运行环境与依赖 37
7.1.1 Python 版本与库 37
7.1.2 数据与工件格式 37
7.1.3 主要软件与库版本建议 37
7.2 典型操作流程 38
7.3 能验证与演示现象解释 38
7.3.1 演示数据规模与指标现象 38
7.3.2 验证方法建议 39
7.4 工业落地扩展讨论 39
7.4.1 标签体系建设 39
7.4.2 在线数据接入 39
7.4.3 与 CAE 的协同 39
7.5 本章小结 39
第 8 章 结论与展望 41
致谢 42
参考文献 43

第 1 章 绪论

1.1 研究背景与意义

1.1.1 新能源汽车结构集成与一体化压铸的产业动因

在全球碳减排与能源结构转型背景下，新能源汽车产业快速扩张。为提升续航里程、降低整备质量并缩短总装节拍，车身制造呈现出“以铝代钢、以集成代分件”的趋势。一体化压铸通过单次充型成形大型薄壁复杂构件，可显著减少焊点数量与连接误差累积，对后地板、前舱、电池包上盖等部件尤为适用。该路线同时带来制造挑战：铸件尺寸大、壁厚差异显著，局部凝固顺序与补缩条件对工艺参数高度敏感；一旦出现气孔、缩松、冷隔、欠铸等缺陷，返修成本与报废损失显著上升。因此，面向单模次乃至射出阶段的过程可视化与风险预测，成为连接“装备能力”与“质量稳定性”的关键环节。后文所称模次与英文术语 shot 含义一致。

产业侧对一体化压铸的关注不仅在于减重，还在于供应链简化与尺寸链收敛：大型单件减少总成公差累积，同时对模具寿命、温控与喷涂维护提出更高要求。若缺乏过程层面的数字化记录，质量问题往往在总装或路试阶段才暴露，追溯成本高。故以工艺参数为主线构建可观测、可推演、可预测的软件闭环，与车身开发节奏相匹配，具有明确的工程动机。

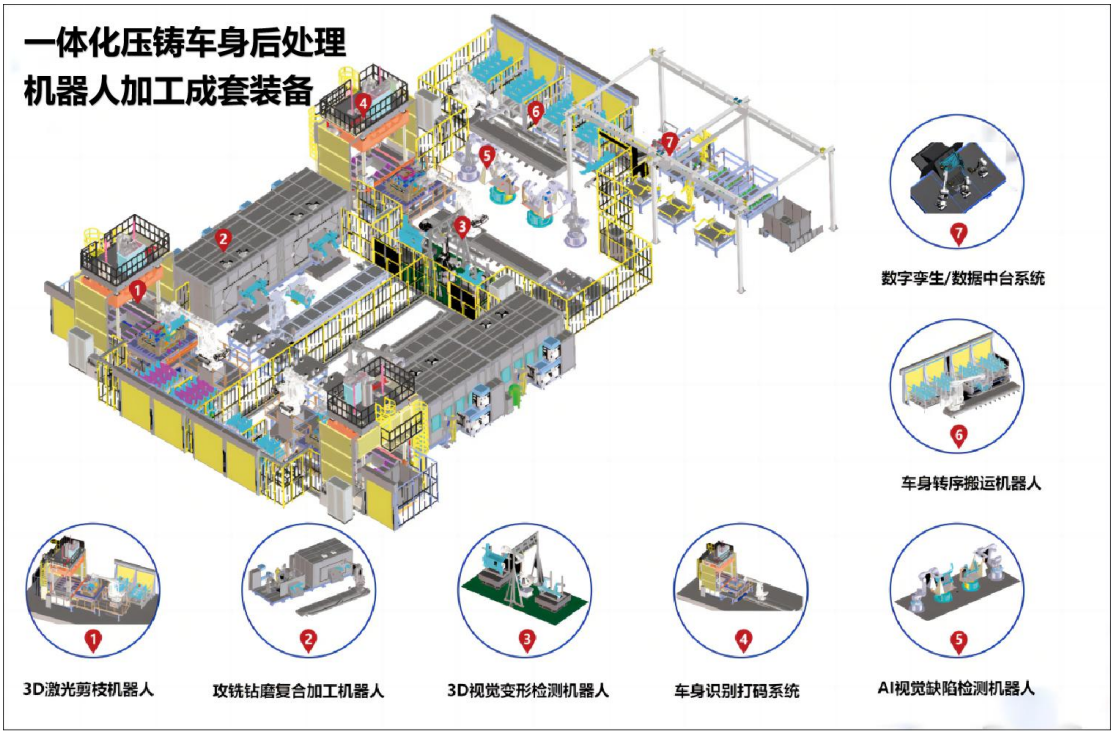


图 1-1　新能源汽车一体化压铸典型应用场景示意

1.1.2 质量管控从离线抽检走向数据驱动的内在需求

传统质量管理往往依赖抽检、金相、射线检测与事后追溯。对一体化压铸而言，抽检比例受限、检测周期长，且缺陷形成与射出曲线、真空建立、增压阶段等动态过程密切相关，仅依赖静态标准难以覆盖过程波动。工业物联网与设备数据采集能力提升，使温度、压力、速度、真空等信号能够以更高频率记录；若可将上述信号与质量结果建立关联，则有望在更早阶段实现风险预警。数字孪生侧重物理实体与运行过程的数字化映射与推演；机器学习侧重从数据中学习输入输出关系。二者结合，可在工程实践中形成可观测的过程状态与可量化的质量风险表征。

数据驱动质控并不否定标准与规范，而是将标准从“事后符合性检验”前移为“过程符合性监控与风险排序”。其实施前提包括：数据可采、模次可关联、时钟可对齐、异常可解释。若仅采集总量而缺乏模次粒度，则学习问题退化为粗粒度统计，难以支撑工位级调参。本文后续章节以 CSV 模次表与 JSON 工艺窗口为最小契约，正是为了在有限复杂度下满足上述前提。

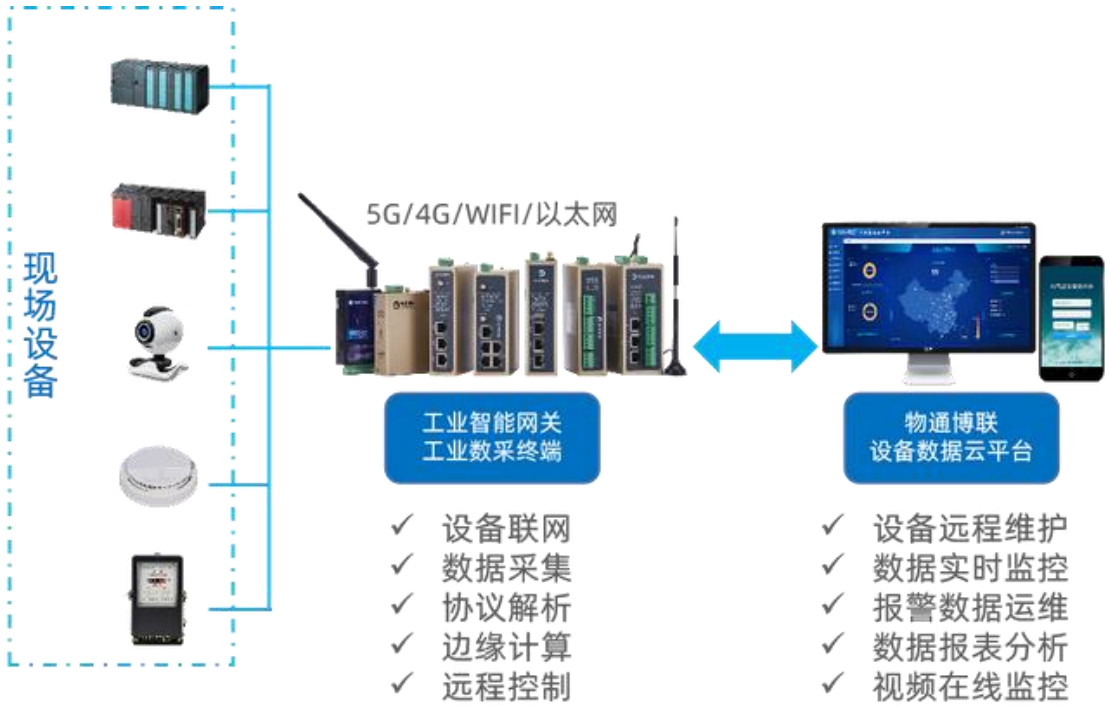


图 1-2 过程数据驱动的质量管控

1.1.3 研究目的与工程价值概述

本研究不以替代材料机理与工艺规范为目标，而着眼于可实施的软件系统设计：以可配置工艺窗口刻画工程边界，以孪生可视化支撑参数微调与情景推演，以模型预测支撑批次对比与报表导出，以服务于工艺人员的决策需求。在实验教学与课程设计场景中，该系统亦可作为算法、数据与界面协同的参考实例，用于说明模块化架构与可扩展接口。

从交付形态看，论文将同时给出概念结构、关键计算公式、数据字段与软件栈说明，使读者能够区分“可复现的工程设计”与“依赖特定演示数据的偶然指标”。当企业已有 MES、QMS 或专用检测设备时，本设计强调导出与配置外置，以便在不大改既有流程的前提下嵌入风险预测与窗口治理模块。

1.2 国内外研究现状与评述

1.2.1 压铸过程建模与仿真研究现状

压铸过程涉及流体力学、传热、相变与应力演化；数值仿真如充型与凝固模拟可揭示卷气、冷隔等机理，但计算成本较高，标定依赖材料与边界条件，且与产线节拍并行实施存在困难。另一类研究关注射出曲线、压力峰值、真空度等过程信号与缺陷统计关联，以统计模型或机器学习建立预测关系。其优势是更贴近

产线数据形态；挑战在于标签噪声、分布漂移与小样本问题。

两类研究路线在工厂中常呈互补：仿真用于工艺窗口初设与缺陷机理解释，数据模型用于节拍内的风险预警与批次对比。本文不追求替代全场仿真，而关注如何将窗口规则、孪生可视化与监督学习组织为统一软件架构，使工程人员能在同一交互语境下理解“规则越界”与“模型风险”两类信息。

1.2.2 数字孪生在制造过程的应用模式

数字孪生在制造领域常见三类形态：几何孪生，刻画设备与模具几何及装配关系；过程孪生，刻画信号与状态随时间的演化；性能孪生，刻画质量、寿命与可靠性等指标。对压铸单元而言，过程孪生与性能孪生往往更直接服务质控。工程实现上，孪生并非仅等同于几何三维展示，而强调与数据同步的状态更新及可交互推演。本系统采用参数滑条与模次选择实现轻量级推演，并以曲线与三维示意进行直观表达，属于面向工位监控的轻量孪生形态。

孪生与业务系统的衔接方式影响落地效果：若孪生仅停留在展示层，难以形成闭环；若与数据采集、工艺授权与质量追溯打通，则推演结果可回写为试验计划或参数变更单。本文在架构层预留 JSON 配置与 CSV 导出，以便与外部系统对接时减少定制开发。

1.2.3 机器学习在铸造缺陷预测中的研究现状

常见做法包括：回归，用于预测孔隙率或力学性能等代理指标；分类，用于判别合格与不合格；排序与异常检测，多用于无标签或弱标签场景。特征方面，除原始传感器特征外，常构造统计特征、曲线特征如峰值、斜率、能量以及工艺派生指标。评价指标包括平均绝对误差 MAE、均方误差 MSE、决定系数 R^2、查准率与查全率以及 AUC 等。当样本量较小时，复杂模型易出现过拟合，R^2 亦可能为负，即模型预测劣于简单均值基线；该现象主要源于数据规模与噪声结构，须在论述中予以充分说明。

跨批次泛化是工业场景的核心难点：材料批次、模具磨损与环境温湿度变化会导致特征分布漂移。研究上可引入迁移学习、域自适应或在线更新，但前提是标签与过程数据长期可追溯。本文以可替换估计器与版本化元数据为设计要点，使数据条件改善后能够升级模型策略，而不推翻数据管道与界面结构。

1.2.4 现有研究的不足与本文定位

综合上述，本文定位在“系统级设计、可复现实现与面向落地的接口”，重点描述三方面内容：其一，工艺窗口作为可治理规则；其二，孪生层与模型层的职责边界；其三，数据文件、模型工件与配置文件的工程化组织。对机理仿真与在线 OPC 的深度耦合，本文在第 7 章作为扩展路线讨论。

与纯算法论文不同，本文刻意将“能跑通的闭环”置于论证中心：从 CSV 载入到模型导出，再到孪生侧条形图与派生指标刷新，形成可演示、可写进说明书的操作路径。对机理细节与先进网络结构，仅在必要处点到为止，避免喧宾夺主。

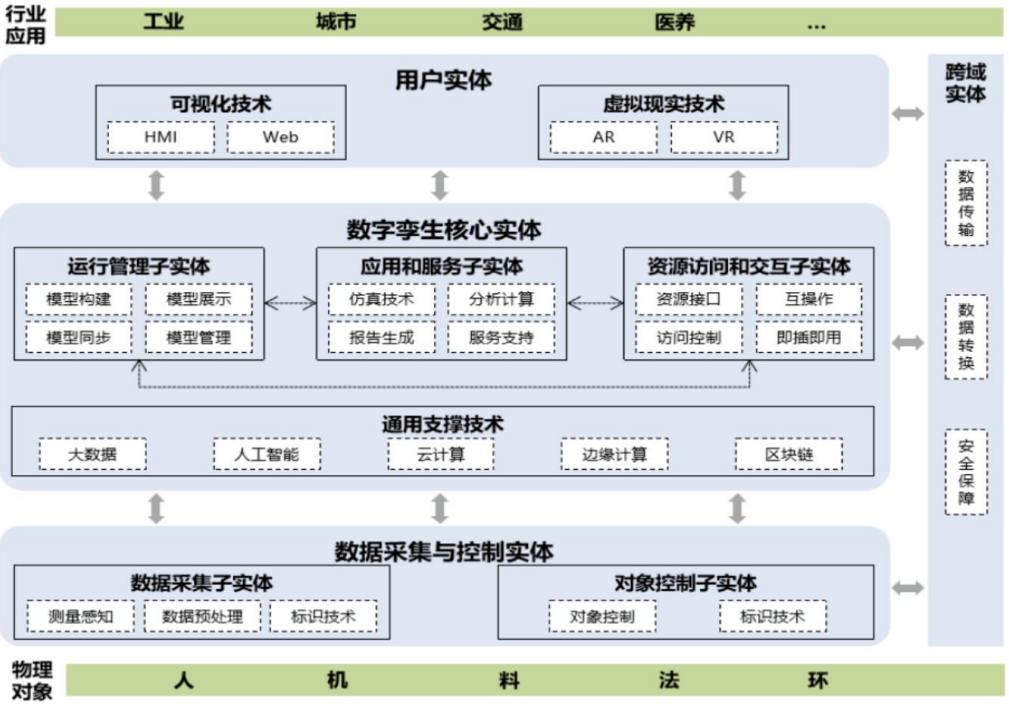


图 1-3 数字孪生架构图

1.3 研究内容与技术路线

1.3.1 研究内容

其一，阐述压铸质量管控场景下的功能需求与非功能需求，形成模块化总体架构。其二，设计工艺窗口配置方案与数据管理流程，明确字段含义与单位。其三，设计数字孪生模块的人机交互与可视化内容，包括派生指标、条形图编码与射出时间轴。其四，设计质量预测模型的特征集合、训练评估流程与报告输出。其五，结合示例数据讨论指标解释与向工业场景扩展的要点。

五项内容在逻辑上形成递进：需求与架构决定模块边界；数据与窗口契约是孪生与预测的公共基础；孪生派生特征又可进入预测特征集，构成“孪生—学习”

闭环。论文后续章节按数据流顺序展开，便于读者从界面操作反查到公式与字段定义。

1.3.2 技术路线

技术路线可概括为“数据—规则—模型—可视化”四段闭环：数据层导入 CSV；规则层读取工艺窗口 JSON；模型层训练并输出风险；可视化层统一呈现。该路线强调模块低耦合，以利于替换数据源与模型算法，并支持接入实时数据流。

实施顺序上建议先冻结字段与配置 schema，再实现校验与持久化，最后挂载图表与交互，以减少返工。若未来引入实时流，可在数据层外侧增加采集与对齐模块，而不必修改窗口归一化与回归指标定义，从而保护已验证的数学层与实验可复现性。

1.4 论文结构安排

第 2 章阐述一体化压铸过程要点与数字孪生质控相关概念；第 3 章阐述需求与总体架构；第 4 章阐述工艺窗口与数据管理；第 5 章阐述孪生模块；第 6 章阐述质量预测模型；第 7 章阐述实现、运行环境与讨论；第 8 章总结全文并展望。



图 1-4 系统首页总览界面

第 2 章 一体化压铸过程机理与数字孪生质量管控相关研究

2.1 高压压铸与一体化压铸的工艺特点

2.1.1 压铸工艺流程概述

高压压铸，英文缩写 HPDC，通常包括模具闭合、熔体输送、慢速封射、快速充型、增压补缩、开模取件等阶段。慢速阶段用于排除空气与稳定流态，快速阶段决定充型速度与前沿形态，增压阶段影响凝固补缩与微观组织。真空工艺用于降低型腔气体含量，减少卷气类缺陷风险。一体化压铸面向超大模具与超大投影面积，设备合模力高、热容量大，工艺参数与模具温度场、喷涂润滑、冷却回路的耦合更强。

从控制角度看，慢压与快压的切换时刻、快压峰值与建压斜率往往由机台程序与模具热状态共同决定；真空建立时序若与充型不匹配，易在局部形成低压卷气区。故仅记录单点设定值不足以刻画风险，理想情况下应保存射出曲线与报警事件；本文原型以模次级聚合量与示意时间轴为折中，以兼顾教学可实现性与工程可解释性。

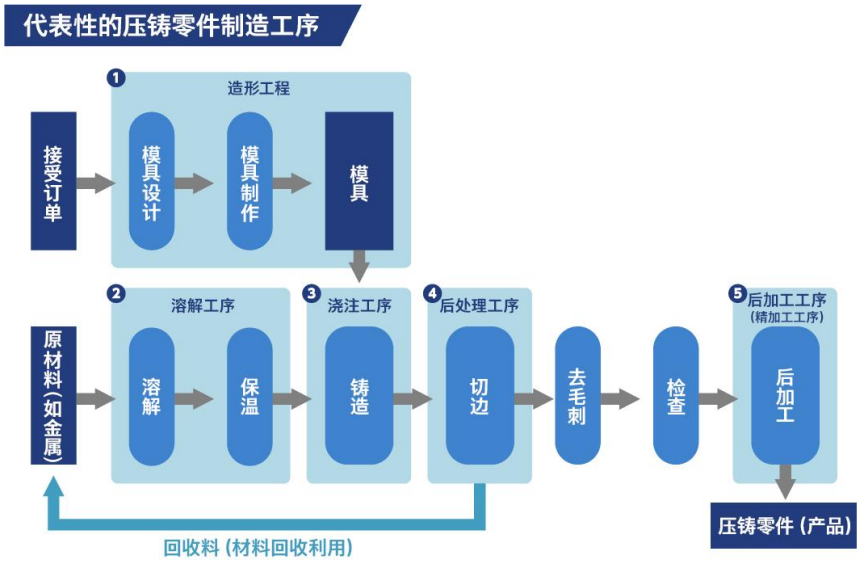


图 2-1 代表性高压压铸典型

2.1.2 典型缺陷类型与过程诱因

气孔类缺陷常与卷气、脱模剂分解气体、真空不足等相关；欠铸与冷隔常与

充型末端温度过低或流动前沿失稳相关；缩松与补缩不足相关，增压与局部壁厚差异是重要因素。工程上常用无损检测与解剖验证，但成本高，促使过程侧预警成为研究热点。

缺陷在铸件上的位置信息对根因分析至关重要：浇口附近与远端的凝固顺序不同，同一组工艺参数在不同壁厚处的表现亦不同。数字孪生若仅给出全局代理指标，应明确其空间平均意义；若未来引入局部传感或射线分层标签，模型输入可扩展为空间相关特征，本节为后续升级预留概念接口。

2.2 数字孪生用于压铸质量管控的关键要素

2.2.1 孪生对象与边界

孪生对象可以是压铸设备整机、模具型腔、机器人辅具或整线单元。边界清晰有助于定义输入输出：输入为可测或可设定的工艺参数；输出为质量指标或风险代理。本系统将“压铸单元、关键工艺参数与派生质量指标”作为孪生对象边界。

边界之外的因素，如铝液含气量、脱模剂喷涂均匀性、冷却水道结垢等，可能显著影响质量却未进入本原型特征表。论文在讨论部分提醒：当残差系统性偏大时，应回到边界外因素排查，而非盲目提高模型复杂度。

2.2.2 同步、推演与可视化

同步强调孪生状态与物理状态一致；推演强调在数字空间进行参数变化并观察结果；可视化强调工程可理解性。本研究采用滑条推演与图表刷新实现轻量推演，适用于实验演示与工艺探索；若接入实时数据，同步频率可提升至秒级或毫秒级。

推演与优化的区别在于：前者回答“若参数如此，指标如何变化”，后者回答“参数应取何值以满足目标”。本系统停留在前者，并借助工艺窗口提供可行域；若需后者，应引入多目标优化与约束求解，并明确目标函数与成本权重，已超出本文原型范围。

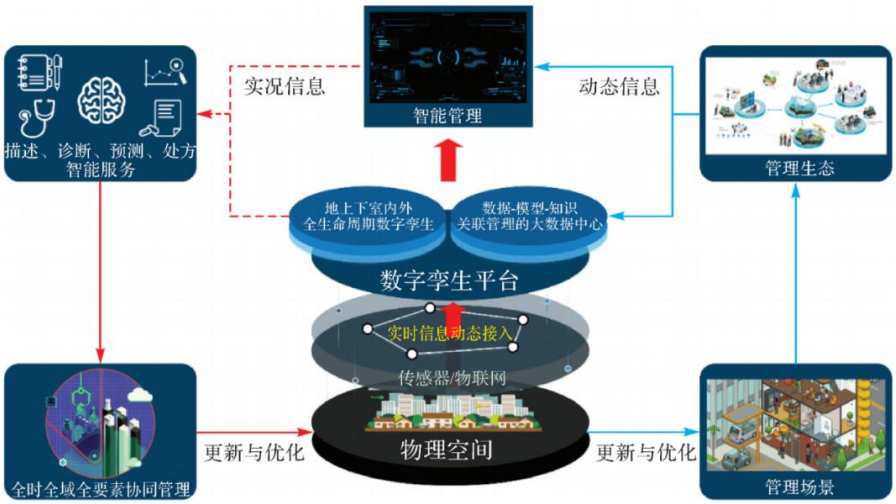


图 2-2 制造场景下数字孪生要素关系

2.3 数据驱动质量预测的基本范式

2.3.1 监督学习问题表述

设输入特征向量为 $\mathbf{x}\in\mathbb{R}^{p}$，包含工艺参数与派生特征；输出 y 为连续型风险或离散型标签。学习目标是得到映射 $\hat{y}=f(\mathbf{x})$，训练阶段通过最小化经验损失 $\frac{1}{n}\sum_{i=1}^{n}\mathcal{L}\big(f(\mathbf{x}_i),y_i\big)$ 估计模型参数。对连续输出常用 MAE、RMSE 与 R^2；对分类输出常用准确率、F1 与 AUC。下文第 6 章给出 $\mathcal{L}$ 取平方损失或绝对值损失时评价指标的具体计算式。

监督学习在压铸场景中的标签往往滞后：质检完成时刻晚于射出时刻，关联模次需依赖追溯体系。若标签与特征时间错位，会引入标签噪声甚至虚假相关。论文在实现章节强调元数据记录与样本量报告，其目的之一正是使读者能判断“可学性”是否被数据链路破坏。

2.3.2 特征工程与可解释性

特征工程包括缺失值处理、异常值裁剪、标准化、派生指标构造与特征选择。可解释性方面，树模型可输出特征重要性；线性模型可看系数符号与幅度。工程上需警惕“虚假相关”，应以工艺知识复核。

对一体化压铸，批次效应常见：同一套参数在不同模具寿命阶段表现不同。可在特征中加入累计模次、修模后计数等慢变量，或采用分层建模。本文示例数据未必包含这些字段，但字段扩展路径应在架构层面保持向后兼容。

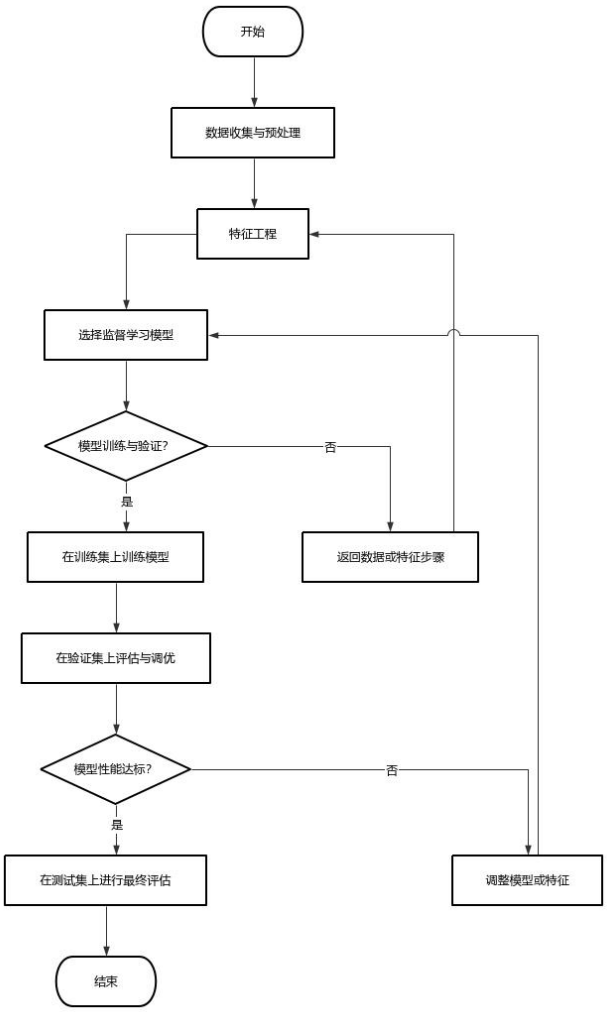


图 2-3 监督学习用于质量预测的基本流程图

2.4 本章小结

本章从工艺机理、孪生要素与学习范式三方面为后续系统设计奠定概念基础，并明确本研究侧重系统架构与工程接口，而非取代高保真 CAE。

后续章节将把“窗口—孪生—预测”串为一致的数据语义：第 4 章给出规则与 CSV 契约，第 5 章给出交互与派生指标，第 6 章给出学习闭环与评价指标，第 7 章讨论从演示到产线的差距与扩展。

2.5 面向工程落地的关键约束与伦理边界

在将“预测风险”用于生产决策时，应明确模型适用范围与不确定性来源：其一，训练数据若来自单一设备、单一模具或短周期，模型可能无法跨批次泛化；其二，当标签存在主观判定或检测阈值变化时，预测输出更宜作为辅助信息而非

唯一判据；其三，对关键安全件或法规件，应保留人工复核与工艺授权流程。本系统强调工艺窗口规则与风险预测的并列呈现，使可解释规则与数据驱动结果同步呈现于工程人员，以降低误用风险。此外，数据采集与模型部署需遵守企业与个人信息保护要求；若数据包含产线节拍、班次与人员信息，应做脱敏与权限控制。上述约束不构成软件实现细节，但应在论文“讨论”章节予以明确，以体现工程严谨性。

组织层面还应约定模型输出与工艺放行的关系：预测风险高是否触发停线或仅提示，须由质量管理体系定义。软件设计提供信息并列与导出能力，不替代质量授权人的签字责任。

第 3 章 系统需求分析、设计原则与总体架构

3.1 需求分析

3.1.1 功能需求

数据管理模块应支持 CSV 载入、路径浏览、表格预览与模次列表展示。工艺窗口模块应呈现关键参数上下限及说明文本，作为越界判定依据。数字孪生模块应支持模次切换、参数微调、派生指标呈现、工艺条形图、型腔三维示意与射出时间轴。质量预测模块应支持模型训练、批量预测、指标呈现、风险分布图与结果导出。模型与报告模块应呈现模型文件、元数据 JSON 与训练指标。帮助模块应提供环境依赖、操作步骤与文档索引。

功能需求还应满足“最小可复现”原则：同一 CSV 与同一随机种子下，训练与指标应可重复；若引入随机搜索或自动调参，则应在界面或日志中显式记录随机状态与参数范围，否则验收与论文实验难以对齐。

3.1.2 非功能需求

可维护性方面，配置外置，模块目录清晰。可扩展性方面，可替换数据源与模型算法。可复现性方面，模型与元数据版本化保存。易用性方面，界面分区明确，操作流程可依据帮助文档完成。

非功能需求还包括可测试性：关键函数如窗口归一化、越界判定、指标计算应可单元测试；界面层可依赖少量集成测试或人工检查表。若部署到生产环境，还需考虑日志、监控与备份策略，本文原型仅作提示。

3.2 设计原则

3.2.1 分层与解耦

将数据、规则、算法与展示分层，避免界面逻辑与训练逻辑纠缠。孪生模块侧重交互与可视化；质量模块侧重训练推理；公共配置由 JSON 管理。

分层带来的代价是跨层调用与数据拷贝，需在性能上接受桌面端小规模数据的权衡。若未来迁移至 Web 或微服务，可将算法层独立为服务，数据层对接对象存储，而规则与孪生仍可复用同一 JSON schema。

3.2.2 规则与学习互补

工艺窗口提供硬边界与可解释约束；机器学习提供从数据中学习到的趋势与风险排序。两者同时存在时，应明确优先级：通常规则用于报警与着色，模型用于风险量化与报表。

当规则与模型结论冲突时，界面应呈现冲突本身而非隐藏其一：例如“在窗但模型高风险”可能提示窗口需更新或存在未观测扰动；“越界但模型低风险”可能提示标签噪声或特征不足。此类并列呈现是工程可解释性的重要组成部分。

3.3 总体架构

3.3.1 逻辑架构

逻辑架构包含数据层、孪生推演层、算法层与应用层。数据层负责 CSV、配置与工件存储；孪生层负责参数映射、派生指标与图形；算法层负责训练评估与推理；应用层负责导航与交互。

数据层在孪生与算法之间充当单一事实来源：若两侧各自缓存 CSV，易出现“孪生所见与训练所用不一致”。故应在设计层面约定：训练读取路径与孪生载入路径指向同一文件或通过同一加载函数。

3.3.2 模块映射与实现目录的对应关系

参考实现可将孪生相关代码置于 src/twin/，质量预测相关代码置于 src/quality/，并在 docs/ARCHITECTURE_AND_DESIGN.md 中记录设计决策，以利于代码评审与二次开发。

公共工具如配置解析、日志与数据校验可置于 src/common/，避免复制粘贴式重复。课程设计多人协作时，目录边界与接口契约应写入仓库说明，减少合并冲突。

3.3.3 系统分层与功能模块一览

为与界面导航及实现目录对照，将逻辑分层与主要功能归纳为表 3-1；表中“主要职责”概括各层在闭环中的输入输出边界。

表 3-1　系统逻辑分层与功能模块一览

层级	界面或实现对应	主要职责

层级	界面或实现对应	主要职责
数据层	数据管理、CSV/JSON	过程数据与配置载入、路径管理与预览
规则层	工艺窗口·限值	读取工艺上下限、支撑越界判定与条形图编码
孪生推演层	数字孪生·推演与窗口	参数微调、派生指标、曲线与三维示意
算法层	质量预测、模型与报告	训练、批量预测、指标与工件持久化
应用层	首页、帮助与说明	导航聚合、文档索引与操作流程提示



图 3-1 系统功能模块导航



图 3-2 模型与报告界面

3.4 本章小结

本章归纳需求、设计原则与分层架构，为后续各模块的细化界定约束与边界。下一章将工艺窗口与 CSV 字段落实为可计算、可校验的契约，使规则可视

化与监督学习共享同一输入语义。

第 4 章 工艺窗口约束与数据管理层设计

4.1 工艺窗口的概念与作用

4.1.1 从经验区间到可执行规则

工艺窗口将工程师经验与试验验证所得可接受参数区间予以结构化。系统在界面呈现上下限，并在孪生条形图中将当前值映射至 $[0,1]$ 相对位置：0 对应下限，1 对应上限。该映射使不同量纲参数得以同屏比较，以利于识别接近窗口边界的参数。相对位置与越界判定的数学表达见第 4.1.3 节。

工程上还需区分“设备允许范围”与“质量合格窗口”：前者可能更宽，后者由质量部门发布。若混用，会导致条形图语义混乱。本文默认 JSON 中存储的是质量窗口；若需同时展示设备极限，应增加第二套上下限或独立视图。

4.1.2 配置外置与可追溯

将窗口写入 config/default.json 一类配置文件，可实现版本管理与变更追溯。界面提示“修改需重启”属于典型桌面应用行为；若未来服务化，可改为热更新与权限审批。

配置变更应与工艺变更单关联：至少记录变更人、时间与适用模具批次。论文级描述强调可追溯，便于在质量争议中回答“当时采用的窗口版本为何”。

4.1.3 工艺窗口相对位置与越界判定的计算

对第 i 个工艺参数，记其工艺下限为 L_i、上限为 U_i，当前设定或实测值为 x_i，且满足 $U_i > L_i$。条形图横轴采用的相对窗位 s_i 可取线性归一化

$$s_i = \frac{x_i - L_i}{U_i - L_i}。$$

当 $x_i < L_i$ 时 $s_i < 0$，当 $x_i > U_i$ 时 $s_i > 1$，故界面可用 $s_i \in [0,1]$ 表示在窗、超出上端或下端用颜色区分。若仅展示“窗内相对位置”，可采用截断映射

$$\tilde{s}_i = \min\left(\max(s_i, 0), 1\right)，$$

此时 $\tilde{s}_i$ 恒落于 $[0,1]$，以利于多参数同屏对比。越界判定可定义示性变量 $I_i^{\text{out}} \in \{0,1\}$：当 $x_i < L_i$ 或 $x_i > U_i$ 时取 $I_i^{\text{out}} = 1$，否则取 0。对 K 个参数统计

在窗个数时，有 $K_{\text{in}}=\sum_{i=1}^{K}\left(1-I_i^{\text{out}}\right)$，与界面中“$K/K$ 在窗”类汇总信息一致。真空度等变量虽常为负值，只要 L_i,U_i 与 x_i 处于同一坐标系，上式仍适用。

若某参数在工艺上存在“单侧约束”或分段窗口，线性归一化需分段定义或改用分段仿射映射；本原型采用统一上下限以简化实现，扩展时应在配置 schema 中增加分段节点并同步更新条形图绘制逻辑。

4.2 工艺窗口参数项与示例取值

4.2.1 参数项与物理含义

熔体温度影响流动性与卷气倾向；模具温度影响凝固壳层与粘模风险；慢压射速度影响排气与流态稳定性；快压射速度影响充型时间与前沿形态；增压影响补缩与致密性；真空度影响型腔气体残留。各参数相互耦合，窗口应通过工艺试验与质量数据校准，而非仅凭单一指标设定。

试验设计可采用正交表或空间填充设计，在有限试验次数下覆盖多维窗口内部；标定完成后，再将保守边界写入 JSON。论文中的示例数值不承担标定职能，仅用于界面与公式对齐说明。

4.2.2 示例窗口

下列区间为与原型系统一致的界面表格示例。为与配置文件字段及第 4.1.3 节符号一致，将关键工艺参数上下限整理为表 4-1。实际生产应以企业工艺卡与设备能力为准，表中数值仅作设计说明。

表 4-1　原型系统工艺窗口参数及示例上下限

序号	参数项	符号约定	单位	下限 Li	上限 Ui
1	熔体温度	—	℃	640	720
2	模具温度	—	℃	120	200
3	慢压射速度	—	mm/s	50	200
4	快压射速度	—	mm/s	2000	6000
5	增压	—	MPa	60	120
6	型腔真空度	—	kPa	−95	−50



图 4-1 工艺窗口限值配置界面

4.3 数据管理模块设计

4.3.1 数据字段与单位一致性

数据表字段建议保持命名与单位一致，避免同一物理量出现多种列名。示例字段含义与表 4-2 一致，便于与第 4 章工艺窗口及第 6 章特征列对齐。

表 4-2 示例过程数据 CSV 主要字段说明

字段名	含义	单位或格式	备注
shot_id	模次 / 批次编号	文本，如 S001	主键
melt_temp_c	熔体温度	℃	与表 4-1 对应
mold_temp_c	模具温度	℃	
slow_shot_mm_s	慢压射速度	mm/s	
fast_shot_mm_s	快压射速度	mm/s	
intensification_mpa	增压	MPa	
vacuum_kpa	型腔真空度	kPa	多为负值

字段名	含义	单位或格式	备注
cycle_time_s	单周期时间	s	可作辅助特征

4.3.2 数据质量与预处理策略

建议记录数据来源、采集时间、设备编号与模具批次；对缺失值应区分“未采集”与“无效零值”。预处理可包括范围校验并与工艺窗口联动标记异常、平滑与离群点处理。对示例数据集，预处理过程应以可复现脚本形式记录，以利于实验与复现。

对时间序列型原始曲线，若下采样为模次级特征，应记录采样窗口与聚合规则，例如峰值取最大、能量取积分，以免不同批次特征不可比。本文表 4-2 以标量字段为主，曲线特征作为第 7 章扩展讨论内容。

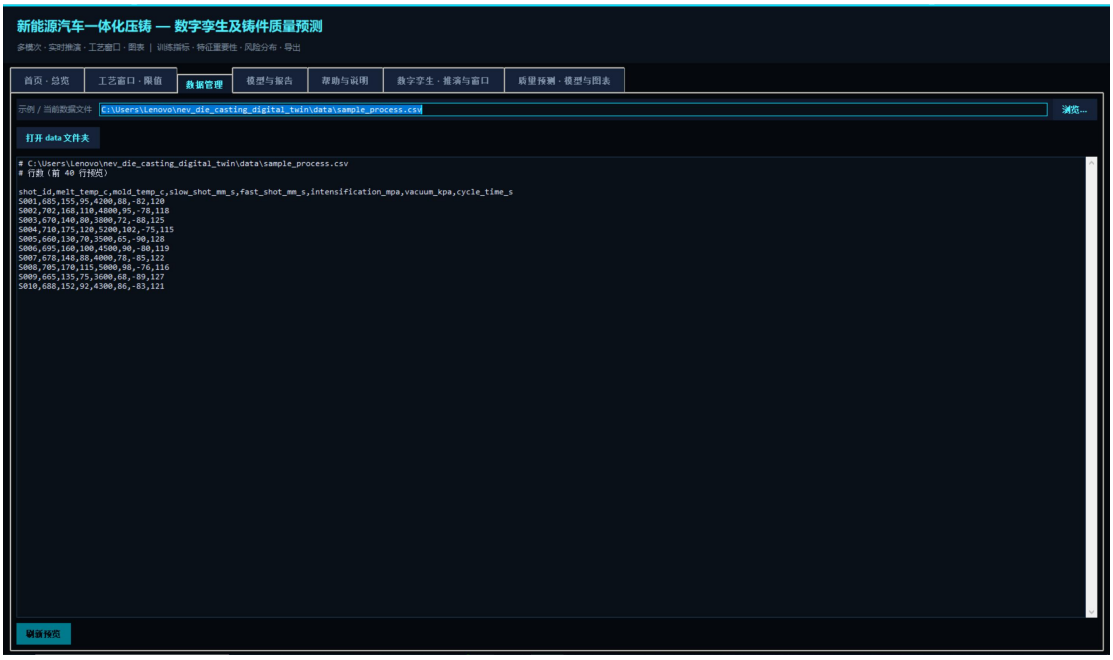


图 4-2 过程数据预览与管理界面

4.4 本章小结

本章归纳工艺窗口规则化方法与数据字段设计要点，为孪生与预测模块提供统一的输入语义。

第 5 章在同一输入上叠加交互推演与多视图表达，使工程人员能够用视觉语言理解规则与代理指标，为第 6 章特征设计提供来源说明。

第 5 章　数字孪生推演与可视化模块设计

5.1　交互式推演与人机界面组织

5.1.1　模次选择与参数来源

操作人员可在历史模次列表中选取一条记录作为初始参数，并在此基础上进行微调。该模式符合由已知工况出发探索邻域的习惯，亦便于对比调整前后模次差异。

列表若支持按时间、风险或越界数排序，可加速定位异常批次；教学演示可预置“贴近窗口边界”的样本，使条形图与汇总状态变化更明显。

5.1.2　滑条微调与实时反馈

滑条覆盖熔体温度、模具温度、慢/快压射、增压、真空等关键量。调整后触发刷新：更新孪生读数、派生指标与图表。界面可显示“6/6 在窗”等汇总状态，将多参数合规性压缩为单条信息，以减轻信息呈现负荷。

交互响应方面，若重绘成本较高，可采用防抖：滑条拖动结束后再统一刷新，或降低三维刷新频率、仅更新读数与条形图。实现细节依框架而定，设计原则是避免操作迟滞削弱推演体验。



图 5-1　数字孪生推演与工艺窗口可视化主界面

5.2 压铸单元孪生监控与设备信息展示

5.2.1 设备型号与能力参数

界面可呈现设备型号，如 LK DCC16000 IMPRESS-PLUS，以及合模力与关键状态量。上述信息用于增强场景可辨识性；若数据并非来源于真实设备实时采集，应在正文中明确数据来源与性质。

展示信息应与法律责任边界一致：设备铭牌参数可公开引用，若涉及产能或客户项目代号，应脱敏处理。

5.2.2 状态量与滑条一致性

合模力、末端增压、快压射速度、真空等显示应与滑条设定一致，体现“孪生映射”的基本一致性原则。若未来接入 OPC，则应显示实测值与设定值，并标注时间戳。

接入实测后，建议增加“数据新鲜度”提示：若 OPC 中断超过阈值，孪生应提示使用最后有效值或切换为离线推演模式，以免现场误判为实时状态。

5.3 孪生派生指标设计

5.3.1 充型指数 Filling Index

充型指数用于表达充型充分性相关的综合指标，常以 0–1 标度展示。工程上可由经验公式、代理模型或学习模型输出。论述中应明确：该指标在原型实现中可为简化代理；工程应用须与射线检测、CT 或试切等验证手段对齐标定。

标定策略可分阶段推进：先保证与报废率或人工分级的单调趋势一致，再在小样本金相或 CT 上精化尺度。任何系数调整应写入配置或元数据，避免界面与训练侧使用不同权重。

5.3.2 气孔代理 Porosity Proxy

气孔代理用于表征气孔类风险的趋势，不等同于真实孔隙率测量。其价值在于横向对比不同参数组合与不同模次之间的相对风险，而不是给出绝对物理量。

若将代理同时作为预测特征，应注意与真空、增压等原始量的信息重叠；可通过消融实验比较“含代理 / 不含代理”的验证指标，判断其边际贡献。

5.3.3 孪生派生指标的量化计算示例

为与软件实现及符号表述一致，可将充型指数 I_{fill} 与气孔代理 P_{poro} 视为

由工艺参数经归一化特征 $\tilde{\mathbf{s}}=(\tilde{s}_1,\dots,\tilde{s}_K)^{\top}$ 映射至 $[0,1]$ 的标量。原型实现可采用加权平均型充型指数

$I_{\mathrm{fill}}=\sum_{i=1}^{K}\omega_i\tilde{s}_i$，其中 $\omega_i\ge 0$ 且 $\sum_i\omega_i=1$；

若强调“短板效应”，可改为 $I_{\mathrm{fill}}=\min_i \tilde{s}_i$。气孔代理可用单调变换构造，例如令 g 为对真空与增压敏感的可调函数，写

$P_{\mathrm{poro}}=\sigma\big(\alpha_1(1-\tilde{s}_{\mathrm{vac}})+\alpha_2(1-\tilde{s}_{\mathrm{press}})+\beta\big)$，

其中 $\sigma(z)=\dfrac{1}{1+e^{-z}}$ 为 logistic 函数，将实数压缩到 $(0,1)$，系数 α_1,α_2,β 由标定或训练确定。三维示意中液面高度可与 I_{fill} 成比例，颜色通道可与 P_{poro} 成比例，即高度 $\propto I_{\mathrm{fill}}$、色调权重 $\propto P_{\mathrm{poro}}$。实际部署时应说明代理式与物理孔隙率之间的对应关系。

5.4 多视图可视化设计

1）工艺条形图：窗口相对位置编码

条形图以相对窗位映射参数，并以颜色区分在窗与越界，如绿色与橙色。该设计将规则约束可视化，适用于现场快速判读。

配色应考虑色弱友好与打印灰度可读性，必要时附加文字标签“在窗/越界”。截图用于论文时，须在图注中解释颜色含义与对应窗口版本。

2）型腔三维示意：几何直觉辅助决策

三维示意可用液面高度与充型指数关联、颜色与气孔代理关联，用于增强空间直觉。界面提示可替换为 STL/STEP 或 CAE 网格，以对接更真实几何。

替换几何时需统一坐标系与单位，否则液面高度映射失去可比性；大体量网格还应考虑渲染性能与内存占用。

3）射出时间轴：速度—压力过程曲线

射出时间轴呈现慢压—快压—增压的典型曲线形态，用于分析参数调整对过程历程的影响。若接入 OPC，可用实测曲线替代示意曲线。若以时间 t 为自变量，慢压射结束时刻记为 t_1，快压射阶段射速可抽象为自 v_{slow} 向 v_{fast} 过渡的

分段函数，增压 $p(t)$ 在充型后段上升，示意模型可写为 $v(t)=v_{\text{slow}}+(v_{\text{fast}}-v_{\text{slow}})\cdot h(t-t_1)$，其中 h 为平滑阶跃，例如 $h(\tau)=\dfrac{1}{1+e^{-\kappa\tau}}$，$\kappa$ 控制上升陡度；该式仅用于界面曲线形态说明，不代表具体机台 PLC 内部控制律。

5.4.1 孪生可视化子视图汇总

数字孪生模块中多视图的分工可归纳为表 5-1，与系统界面各功能分区及图相对应。

表 5-1　数字孪生模块主要可视化子视图及用途

子视图名称	核心内容	工程用途
工艺条形图	各参数在工艺窗口内的相对位置	快速判断在窗 / 越界与贴近边界程度
型腔三维示意	液面高度、颜色与代理指标映射	增强空间直觉，可替换为 CAD/CAE 网格
射出时间轴	射速、增压随时间变化曲线	理解慢压 — 快压 — 增压历程与调参影响
孪生监控区	设备型号、合模力与关键状态量	与现场机台信息对齐，用于展示与培训



图 5-2 S001 型腔三维示意与质量代理映射

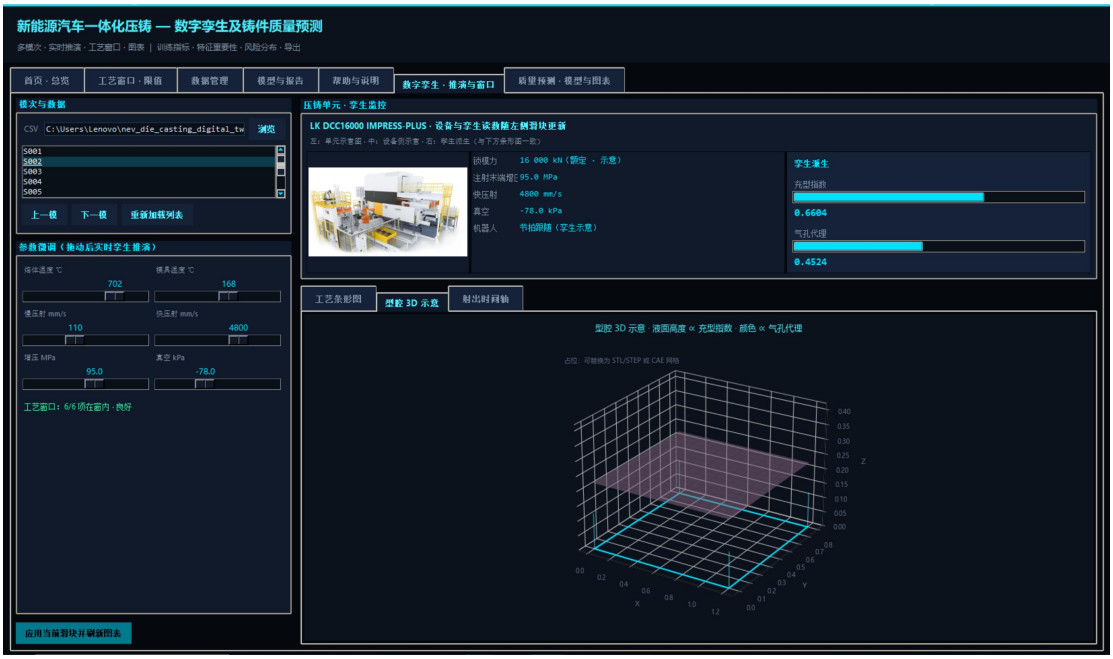


图 5-3 S002 型腔三维示意与质量代理映射



图 5-4 S001 射出过程速度—压力时间轴



图 5-5 S002 射出过程速度—压力时间轴

5.5 本章小结

本章从交互组织、孪生监控、派生指标与多视图可视化四方面阐述孪生模块设计，强调可解释规则可视化与代理指标趋势化相结合的思路。

第 6 章将孪生输出与工艺参数一并纳入特征集合，并给出标准化、训练流程与回归评价指标，使预测模块与孪生模块在数据语义上闭环。

第 6 章 铸件质量预测模型与评价体系设计

6.1 预测任务定义与输出语义

6.1.1 回归型风险输出的适用场景

当标签为连续型缺陷风险或连续型质量指标时，可采用回归模型。输出可在 0–1 区间解释为风险强度，也可映射为等级。需要明确：风险是否具有单调可解释性，是否与工艺方向一致。

若标签经非线性变换，如取对数或分箱，应在论文中写明变换与反变换，避免跨系统比较时量纲混乱。对安全相关件，风险阈值应由质量部门授权，而非仅由模型默认决定。

6.1.2 分类与异常检测扩展

若标签为合格/不合格，可采用分类模型；若缺乏标签，可采用无监督异常检测。本研究以回归与连续风险为主展开论述，架构上可替换 scikit-learn 中的其他估计器。

分类场景下需关注类别不平衡：不合格样本稀少时，准确率失真，宜报告查准率、查全率与 F1，并采用分层抽样或代价敏感学习。异常检测可作为冷启动手段，但阈值仍需工艺与质量共同确认。

6.2 特征集合设计

6.2.1 原始工艺特征

原始特征与数据表一致：熔体温度、模具温度、慢/快压射速度、增压、真空等。此类特征具有明确物理意义，便于与工艺窗口联合解释。与模型元数据 feature_columns 的对应关系可参照表 6-1。

表 6-1　质量预测模型输入特征与数据来源说明

特征名	类型	来源
melt_temp_c	工艺参数	CSV 列，与表 4-2 一致
mold_temp_c	工艺参数	CSV 列
slow_shot_mm_s	工艺参数	CSV 列

特征名	类型	来源
fast_shot_mm_s	工艺参数	CSV 列
intensification_mpa	工艺参数	CSV 列
vacuum_kpa	工艺参数	CSV 列
derived_fill_index	派生特征	孪生侧计算或规则输出，见第 5.3 节
derived_porosity_proxy	派生特征	孪生侧计算或规则输出，见第 5.3 节

6.2.2 派生特征：孪生输出进入学习闭环

将 derived_fill_index、derived_porosity_proxy 等派生量纳入 feature_columns，形成孪生计算与机器学习相衔接的闭环。须注意：若派生特征由与输入参数相同的规则直接计算，可能与原始特征强相关并引发多重共线性；实践中应开展相关性分析与特征选择。

若孪生侧与训练侧使用不同实现或浮点顺序，派生特征可能出现细微差异，建议在流水线中单点计算并序列化到 CSV，或共享同一函数库，保证“所见即所训”。

6.2.3 特征标准化与数据集划分

对梯度敏感或距离度量敏感的算法，常对特征做标准化。数据集划分可采用 hold-out 或交叉验证。样本极少时，交叉验证更稳妥，但计算成本与方差需要权衡。

划分时应避免信息泄漏：同一模具连续模次不宜同时落入训练与验证集时可采用分层或按模次分组；若标签随时间漂移，可按时间切分验证集以更贴近部署。

6.2.4 特征标准化与预测模型的矩阵形式

设训练集特征矩阵为 $\mathbf{X}\in\mathbb{R}^{n\times p}$，第 j 列均值为 $\mu_j=\frac{1}{n}\sum_{i=1}^{n}X_{ij}$，标准差 $\sigma_j=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_{ij}-\mu_j)^2}$，标准化后

$Z_{ij}=\frac{X_{ij}-\mu_j}{\sigma_j}$，当 $\sigma_j=0$ 时该列可剔除或置零处理。

线性回归可写为 $\hat{y}_i = \mathbf{w}^\top \mathbf{z}_i + b$，其中 $\mathbf{z}_i$ 为第 i 条样本标准化后的特征向量；含 L2 正则时目标为 $\min_{\mathbf{w},b} \frac{1}{n}\sum_i (y_i - \hat{y}_i)^2 + \lambda \| \mathbf{w} \|_2^2$。随机森林等树模型可不标准化，但为与线性基线公平比较，可在实验中同时报告“原始尺度 / 标准化尺度”两套结果。

6.3 模型训练、评估与工件管理

6.3.1 训练流程

训练流程包括读取 CSV、清洗、特征对齐、划分训练/验证集、拟合模型、计算指标、保存 joblib 与 json 元数据。元数据建议记录 feature_columns、指标、样本量、时间戳与随机种子，以支持复现。

对极小样本，宜优先简单模型与留一法等评估策略，避免网格搜索与复杂模型叠加导致过拟合。若引入随机搜索，应将搜索空间与最优超参写入元数据。

6.3.2 指标解释：MAE 与 R^2

MAE 衡量平均绝对误差，直观易懂。R^2 衡量模型相对均值基线的改进程度。当 R^2 为负，说明模型在验证集上比简单均值预测更差，常见于样本极少、噪声大或特征与标签关系不稳定等情况。应同时报告样本量与划分方式，避免仅凭单一指标作出判断。各指标定义与计算式见第 6.3.4 节。

6.3.3 特征重要性

树模型如随机森林与梯度提升树可输出特征重要性；线性模型可看系数。若界面未呈现特征重要性，多与所选模型类型或训练实现方式有关，须在技术说明中予以解释。

解释特征重要性时应结合量纲与标准化：树模型基于不纯度的重要性受类别与分箱影响；线性系数在标准化前后尺度不同。更稳妥的做法是结合置换重要性或部分依赖图，并与工艺窗口关键参数对照。

6.3.4 回归评价指标的计算公式

设验证集或测试集样本量为 n，真实值为 y_i，预测值为 $\hat{y}_i$，均值 $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$。平均绝对误差定义为 $\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} \left| y_i - \hat{y}_i \right|$；均方误差为

$\mathrm{MSE}=\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$；均方根误差为 $\mathrm{RMSE}=\sqrt{\mathrm{MSE}}$。决定系数采用 $R^2=1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2}$，分母为总平方和，分子为残差平方和；当残差平方和大于总平方和时 $R^2<0$。若需对异常值更鲁棒，可报告 MAE 为主、RMSE 为辅，并在附录给出各模次残差 $r_i=y_i-\hat{y}_i$ 表。上述指标的名称、含义与取值特点汇总见表 6-2。

表 6-2　回归任务常用评价指标一览

符号	名称	取值特点或解释要点
MAE	平均绝对误差	同量纲于 y，对离群点相对鲁棒
MSE	均方误差	放大较大误差，敏感于离群点
RMSE	均方根误差	与 MSE 等价，量纲同 y
R2	决定系数	不大于 1；越接近 1 越好；可为负

注：当验证集样本量极小时，R^2 为负的情形较易出现，须结合 MAE 与残差表一并报告。

6.3.5 模型选择与偏差—方差权衡简述

在样本量有限时，复杂模型往往降低训练误差但升高验证方差。实践中可先建立线性回归或岭回归作为强基线，再尝试随机森林与梯度提升；若验证集提升不显著，应优先保留简单模型以利于解释与部署。偏差—方差权衡并非本文理论重点，但应在论文讨论中体现工程判断：避免仅因训练集拟合更好而选择复杂结构。



图 6-1 质量预测训练指标与批次风险分布

6.4 批量预测与结果导出

6.4.1 批次表与风险分布图

批量预测生成模次—风险表，并以柱状图展示分布，以利于识别异常模次。若风险值近似常数，则可能源于模型表达能力不足、特征区分度有限或标签近似常数，须从数据与建模两方面加以诊断。

柱状图纵轴与刻度应随数据自适应，并在样本极少时提示“统计意义有限”。导出表可附加各参数越界标记，便于与工艺根因分析会同一张表。

6.4.2 导出 CSV 的工程意义

导出预测结果可支撑与 MES/QMS 的对接，并便于离线复核与审计。

导出文件宜包含预测时间、模型版本与特征列快照，使历史结果可复现；若企业有留存期限与脱敏要求，应在系统说明中列明字段范围。

6.5 本章小结

本章归纳预测任务界定、特征设计、训练评估与导出流程，并强调指标解释须结合样本规模与数据质量。

第 7 章将给出运行栈、操作流程与演示现象解释，使本章公式与工件管理与可执行软件对应，并讨论工业扩展路径。

第 7 章 软件实现、运行环境、功能验证与讨论

7.1 运行环境与依赖

7.1.1 Python 版本与库

系统实现采用 Python 3.10 作为运行环境；可视化基于 matplotlib，用于绘制射出曲线、工艺条形图、风险分布柱状图等；机器学习采用 scikit-learn；模型持久化采用 joblib。上述技术栈在工业研究与教学场景中应用广泛，生态成熟。

依赖版本建议用 requirements.txt 或 conda environment.yml 固定，避免 sklearn 小版本差异导致默认参数变化。无界面服务器环境需将 matplotlib 后端设为 Agg，以免绘图报错。

7.1.2 数据与工件格式

CSV 为通用表格交换格式，可与 Excel、数据库导出结果对接；JSON 适用于配置与元数据；joblib 适用于 scikit-learn 模型对象的序列化。

读写 CSV 时建议显式指定 UTF-8 编码；列名含中文或路径含中文时，Windows 默认编码易引发乱码。JSON 中浮点数与 null 的解析应与保存端一致，避免指标被读成字符串。

7.1.3 主要软件与库版本建议

实现环境与第三方库以表 7-1 为准：运行环境为 Python 3.10，可视化依赖 matplotlib；复现实验时宜与表 7-1 保持一致，以避免解释器版本差异导致结果不一致。

表 7-1 系统实现主要软件与库

类别	名称	用途说明
运行环境	Python 3.10	开发与运行所采用版本
数值与模型	scikit-learn	回归模型训练、评估与推理
可视化	matplotlib	图表绘制依赖该库
持久化	joblib	序列化模型对象

类别	名称	用途说明
配置与元数据	JSON	工艺窗口、模型指标与 feature_columns
数据交换	CSV	过程数据与预测结果导出



图 7-1 帮助与系统说明界面

7.2 典型操作流程

典型操作顺序为：在数据管理页核对 CSV 字段；在质量预测页完成模型训练并查看指标；执行批量预测并导出结果；在孪生页选取模次并微调参数以观察派生指标；在工艺窗口页核对限值是否需更新。上述顺序与帮助文档中“训练—预测—推演”的表述一致。

首次运行建议用极小样本做冒烟测试，确认列名与 feature_columns 一致后再扩大数据。答辩演示可预置已训练 joblib，缩短现场等待。

7.3 能验证与演示现象解释

7.3.1 演示数据规模与指标现象

在样本量极小的情形下，例如训练集 7 条、验证集 3 条，验证集 R^2 为负

较易出现；若风险分布近似常数，则表明模型可能未有效刻画结构或标签变化不足。上述现象应归因于数据条件不足，而不宜据此否定系统架构本身。

撰写论文或答辩时应同时给出样本量、标签范围与特征范围，使听众建立与指标相符的预期，避免“界面华丽但指标不可解释”的印象。

7.3.2 验证方法建议

更严谨的验证应包括：扩大样本、分层抽样、交叉验证、与无损检测结果对齐、对比多种模型如线性回归、随机森林、梯度提升以及正则化策略。

交叉验证折数不宜过高以免每折样本过少；可重复随机划分多次报告均值与方差。与射线或 CT 对齐时须核对模次与时间戳，防止标签滞后错位。

7.4 工业落地扩展讨论

7.4.1 标签体系建设

建议建立缺陷编码标准，将射线、CT、报废原因与模次关联；对连续标签可采用缺陷面积、孔隙率估计等。

编码体系应版本化：标准变更时保留旧码映射，否则历史训练数据与现行质检口径不可比。主观分级标签宜做一致性抽检，降低标注漂移。

7.4.2 在线数据接入

通过 OPC UA 接入实时曲线与设备报警，孪生页可用实测曲线替换示意曲线；同时引入时间序列特征，如峰值、上升斜率与积分能量。

接入后需处理时钟同步与丢包：时间轴错位会使积分类特征失真。可在边缘侧缓冲对齐模次后再上传，云端刷新频率与产线节拍匹配。

7.4.3 与 CAE 的协同

可将 CAE 输出的温度场/卷气风险指数作为额外特征，形成“机理仿真 + 数据学习”的混合范式。

全场仿真与模次级预测在时效上不对等，可采用离线预计算场特征或降阶模型，在线仅更新边界敏感量，以平衡精度与延迟。

7.5 本章小结

本章从环境、流程、验证与扩展讨论四方面补充实现细节，并强调演示系统与工业系统之间的关键差距主要在数据规模与在线集成。

第 8 章对全文论点作收束，并对应第 1 章工程价值，提出数据、在线集成与系统协同三方面可延续工作。

第 8 章 结论与展望

本文面向新能源汽车一体化压铸场景，提出并阐述了一套融合数字孪生推演、工艺窗口治理与数据驱动质量预测的软件系统设计。系统以关键工艺参数为主线，完成数据管理、模型训练、批量预测、结果导出与孪生可视化闭环；通过 JSON 配置实现工艺窗口可治理与可视化一致性；通过 joblib/json 工件实现模型可复现。对演示小样本条件下的负 R^2 与风险近似常数等现象，本文从统计学习角度给出解释，并指出改进方向。

在创新点表述上，本文侧重可落地的模块划分与数据契约，而非提出新的压铸材料机理或全新深度学习结构；在应用价值上，可作为企业数字化项目的需求说明蓝本或院校课程设计的结构参考，按需替换数据源与估计器即可演进。

未来工作可在三方面深入：一是构建高质量标签与大规模过程数据；二是引入在线实时数据与更丰富的曲线特征；三是与 CAE、MES、QMS 深度集成，实现从预警到工艺优化闭环。

此外可研究模型监控与漂移告警：当特征或残差分布随季节、模具寿命缓慢变化时，提示再训练或重标定工艺窗口；在合规前提下亦可探索跨工厂迁移学习，缓解单产线数据不足。上述方向需质量与工艺部门共同定义指标与责任边界，技术实现应服务于可审计的质量体系。持续运营阶段还应建立模型与数据版本台账，使任意历史预测结果可追溯到当时生效的窗口、特征列与估计器配置。

致谢

本论文的顺利完成，首先要向我的指导教师致以最由衷的感谢与最崇高的敬意。从论文最初的选题构思、研究方向的精准定位，到整体框架的搭建、研究方案的细化完善，再到系统设计的落地实施、实验数据的论证分析，以及最终文稿的反复修改、逻辑梳理与格式规范，导师始终给予我细致入微的指导与耐心细致的点拨。导师以其严谨务实的治学态度、深厚扎实的专业学识和丰富的科研经验，在我遇到困惑迷茫时为我指明研究方向，在课题推进遭遇瓶颈、技术难题难以突破时，悉心为我分析问题、提供思路，帮助我逐一攻克难关。无论是系统五层架构的优化设计、压铸工艺参数的窗口分析，还是质量预测模型的构建与实验调试，每一个关键环节都离不开导师的悉心指引与无私帮助，这份言传身教不仅让我顺利完成了本次课题研究，更培养了我严谨的科研思维和求真务实的学习态度，让我受益终身。

同时，我要感谢朝夕相伴的同窗好友们。在课题研究的漫长过程中，我们相互交流学习心得、分享研究思路，在遇到技术难点时彼此鼓励、共同探讨，一起查阅文献、分析数据、调试模型，在一次次的讨论与尝试中攻克了诸多难题，不仅提高了研究效率，更收获了珍贵的友谊与并肩前行的力量。感谢我的家人，一直以来给予我无条件的理解、支持与包容，默默承担起生活的重任，为我创造了安心钻研、专心完成学业的良好环境，他们的牵挂与鼓励，是我克服困难、奋勇前行的坚实后盾。此外，还要感谢院校为我们提供的优质学习平台、完善的实验资源以及良好的科研氛围，为本次课题的顺利开展提供了有力保障。这段求学路上的每一份收获与成长，都离不开身边人的帮助与支持，这些温暖与力量，都将成为我未来人生道路上的宝贵财富，激励我在专业领域不断深耕细作、精进自我，稳步前行。

参考文献

[1]刘峻作,李俊林,程超锋,等. 基于数字孪生的液氨电阻推进系统故障诊断研究[J/OL].推进技术,1-16[2026-04-14].

[2]张群,郭园,杨文栋,等. 数字孪生赋能城市防洪排涝“四预”决策体系构建分析[J].陕西水利,2026,(04):68-70.

[3]赵英豪,李勇坚,李佳艺. 数字孪生多模态耦合：驱动工业智能设备创新的使能机理与优化策略[J/OL].价格理论与实践,1-6[2026-04-14].[4]魏炜,何彬,王彦军.电机压铸模具再制造工艺优化与应用分析[J].南方农机,2026,57(03):122-125.

[5]乔子龙,周扬,毛高峰,等. 高压压铸铝合金零件多工序加工工艺集成优化[J].汽车知识,2026,26(02):89-91.

[6]谢峰,唐磊,雷超. 基于 GABP 优化算法的压铸件质量预测研究[J].制造技术与机床,2025,(10):214-220.

[7]龚欢.铝合金复杂压铸件数字孪生系统研究[D].大连交通大学,2025.

[8]师博人.铝合金压铸制造质量追溯及数字孪生监控[D].大连理工大学,2025.

[9]刘双勇,刘维洲,赵永刚,等. 工业元宇宙技术在高压铸造岛全生命周期中的应用研究[J].特种铸造及有色合金,2024,44(11):1489-1497.

[10]徐超. 数字孪生技术在新能源汽车一体压铸生产中的创新应用[J].金属加工(冷加工),2024,(08):8-12.

[11]雷超.大型压铸机数字孪生模型构建及铸件质量预测研究[D].安徽大学,2024.

[12]许宝奇.面向压铸过程的数字孪生监控及铸件质量预测方法[D].大连理工大学,2024.

[13]杜宇,柴文杰,刘冬. 基于数字孪生的压铸过程监控和产品质量预测[J].组合机床与自动化加工技术,2023,(12):167-173.

[14]李悦.基于 DVSM 的 S 公司变速箱铝壳体生产线精益改善[D].河北科技大学,2023.

[15]黄胜,张国富,赵子龙. 基于数字孪生的电机定转子车间设计[J].时代汽车,2023,(20):10-12.

[16]侯远韶. 基于数字孪生的压铸机器人模型构建与演化[J].安阳工学院学

报,2023,22(04):68-72.

[17]王艳,洪海玻,张宇,等. 一体化压铸在白车身中的应用浅析[J].汽车工艺师,2023,(07):58-60.

[18]舒吉平.压铸工艺过程与模具设计虚拟仿真实验及数字孪生平台搭建[D].重庆交通大学,2023.

[19]柴文杰.基于数字孪生的压铸单元监控和产品质量预测方法[D].大连理工大学,2022.

[20] 日本芝浦机械推出电动合模大型压铸机[J].铸造工程,2022,46(02):72.