

选煤数智化系统精准对口建模与优化构建研究

刘左权, 张 驰, 王 倩, 郝明辉

陕西新能选煤技术有限公司 陕西西安

【摘要】针对选煤数智化系统中算法模型与业务场景匹配不足、统一模型泛化套用导致控制效果不稳定的问题, 提出“场景任务-数据画像-模型对口-指标验模-部署闭环”的精准对口建模方法。按设备诊断、过程优化、质量预测、故障分析和工艺控制任务, 分别匹配神经网络、多元线性回归、支持向量机、聚类分析及专家规则/比例-积分-微分 (PID) 等模型, 建立离线验证、工业验证和在线迭代机制。工程应用结果表明, 系统故障诊断准确率达 93%, 精煤回收率提升 1.1%, 介耗降低 16%, 日均停机时间减少 0.55 h, 年直接经济效益 341.9 万元。研究表明, 按场景约束选择模型并持续优化, 可提高模型有效性、控制适配性和工程可复制性。

【关键词】智慧选煤; 精准建模; 算法适配; 专家知识库; 模型优化; 智能控制

【收稿日期】2026 年 3 月 5 日 **【出刊日期】**2026 年 8 月 20 日 **【DOI】**10.12208/j.merd.20260004

Precision scenario-matched modeling and optimized construction of digital-intelligent coal preparation systems

Zuoquan Liu, Chi Zhang, Qian Wang, Minghui Hao

Shaanxi Xinneng Coal Preparation Technology Co., Ltd. Xi'an, Shaanxi

【Abstract】To address the mismatch between algorithmic models and business scenarios in digital-intelligent coal preparation systems, a precision scenario-matched modeling method consisting of scenario-task definition, data profiling, model matching, metric-based validation and closed-loop deployment is proposed. For an engineering smart coal-preparation project, neural networks, multiple linear regression, support vector machines, clustering, expert rules and Proportional-Integral-Derivative (PID) control are matched respectively to equipment diagnosis, process optimization, quality prediction, fault analysis and process control tasks. Offline validation, industrial verification and online iterative optimization are integrated into the model lifecycle. Engineering application achieved 93% fault-diagnosis accuracy, a 1.1% increase in clean-coal recovery, a 16% reduction in medium consumption, a 0.55 h/day reduction in downtime and annual direct economic benefits of CNY 3.419 million. The results indicate that scenario-constrained model selection and continuous optimization can improve model effectiveness, control adaptability and engineering replicability.

【Keywords】Smart coal preparation; Precision modeling; Algorithm matching; Expert knowledge base; Model optimization; Intelligent control

引言

智能化选煤厂建设正在由设备自动化、数据可视化向全流程智能分析和自主调控发展。现有研究已在选煤大数据治理、数字孪生、智能感知、设备全生命周期管理和专家系统等方面形成较为完整的技

术框架^[1-2]。但是, 选煤生产包含设备故障诊断、产品质量预测、工艺参数优化、液位控制、压滤卸料判定等差异明显的任务, 其输入数据形态、输出目标、时滞特征、可解释性要求和控制风险并不相同。若以单一算法或统一模型处理全部场景, 容易出现模

第一作者简介: 刘左权 (1985-) 男, 本科, 助理工程师, 研究方向: 机电设备维护管理。

型精度高但难以部署、局部有效但跨工况失效、预测有效但不能闭环控制等问题。

近年来,面向煤泥浮选的 WOA-GRU 系统辨识模型与模型预测控制、无模型深度强化学习等研究表明,不同工艺过程应根据时滞、多变量耦合和在线学习需求选择相应模型^[3-4]。这说明“算法先进性”并不等于“工程适配性”,模型选择需要与具体任务、数据条件和控制目标进行对口匹配。

智慧选煤数智管控一体化项目已构建神经网络、多元线性回归、支持向量机、聚类分析以及专家知识规则等算法库,并采用问题分析、算法筛选、参数优化、验证测试和算法集成的适配流程。在此基础上,本文提出“精准对口建模”概念:不追求单一通用算法,而是按照场景任务、数据特征、输出目标、实时约束和控制权限选择模型族,建立统一验证标准,并通过工业反馈持续迭代。研究重点包括精准对口建模方法、典型场景模型配置、生产过程模型优化及工程效果验证。(如图 1 所示)

核心原则:一场景一任务,一任务一数据画像,一数据画像对应模型族与验收指标

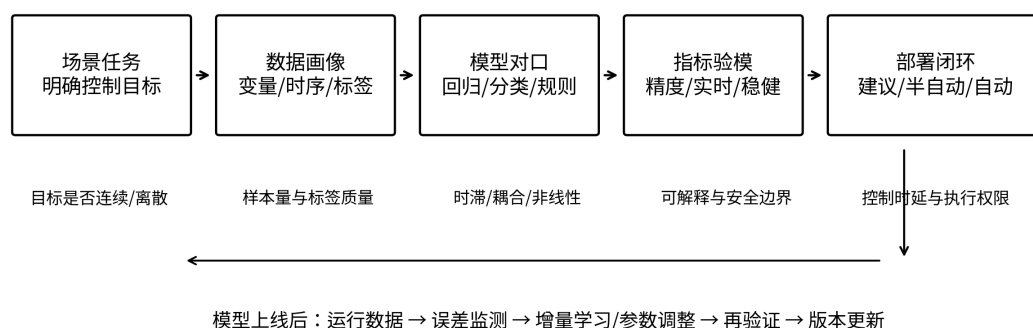


图 1 选煤数智化系统精准对口建模方法框架

1 精准对口建模方法

1.1 方法定义与基本原则

本文所称“精准对口建模”是指以业务问题为起点,而不是以某一种算法为起点,将场景任务、数据条件、模型类型、评价指标和部署方式进行一一映射。对任一业务场景 i , 可将其建模对象表示为:

$$M_i = \langle O_i, D_i, T_i, A_i, E_i, C_i \rangle$$

式中: O_i 为业务目标,如故障诊断、质量预测或参数优化; D_i 为数据画像,包括数据类型、样本量、时序性和标签质量; T_i 为任务类型,如回归、分类、聚类、规则推理或闭环控制; A_i 为候选算法族; E_i 为验模指标; C_i 为部署控制方式。该表达式为本文提出的工程化建模框架,用于统一不同场景的模型选择逻辑,并非项目原报告已有数学公式。

精准对口建模遵循 3 项原则:一是任务优先,先明确要解决的是预测、分类、优化还是控制,再选择模型;二是数据约束,算法复杂度与可获得数据量、标签质量和实时性相匹配;三是控制边界约束,模型输出是提示、推荐还是自动执行,应由安全风

险和可验证程度决定。对于安全等级高、难以完整建模的场景,应保留专家规则和确定性控制逻辑,而不是强行采用复杂数据模型。

1.2 “硬约束筛选+性能验模+工业复核”三级适配

第一层为硬约束筛选。根据输出类型、样本情况、时滞与耦合特征、实时响应要求、解释性和控制风险,排除不适合的模型。例如连续质量/能耗预测可优先考虑回归模型;设备状态分类可采用神经网络或支持向量机;缺乏完整数学模型但具有明确经验规则的场景可采用专家规则;液位等连续控制问题则需考虑 PID 或其他确定性控制器。

第二层为性能验模。项目研究报告对分类/诊断模型提出准确率 $\geq 92\%$ 、精确率 $\geq 90\%$ 、召回率 $\geq 88\%$ 、 $F1 \geq 89\%$ 、响应时间 $< 1\text{ s}$ 等指标;对生产过程多元线性回归模型要求 $R^2 \geq 0.85$, F 检验显著、回归系数 t 检验显著且残差满足统计要求,并在工业验证阶段要求预测误差 $\leq 5\%$ 。这些指标使模型选择从经验判断转化为可验收过程。

第三层为工业复核。离线模型达标后仍需进入

真实工况验证，重点检查控制稳定性、异常工况、传感器漂移、网络延时和人工操作影响。模型只有在生产质量、能耗、设备稳定性或维护效率等工程指标改善后，才能作为正式上线模型。

1.3 场景—算法对口关系

项目针对不同业务任务采用了差异化模型：设备健康管理采用神经网络，依托实时运行数据和历史故障记录建立诊断模型；生产过程调控采用多元

线性回归，量化工艺参数与产品质量、综合能耗的关系；产品质量管控采用支持向量机预测关键质量指标；故障分析采用聚类算法识别故障类型、部位和诱因；磁选机液位采用 PID 与专家规则；板框压滤机采用滤液流量阈值和压力关系进行状态触发控制；原煤分选采用灰分-密度-产率关系模型动态给出分选密度设定值。上述配置体现出“不同问题采用不同模型”的基本工程逻辑。

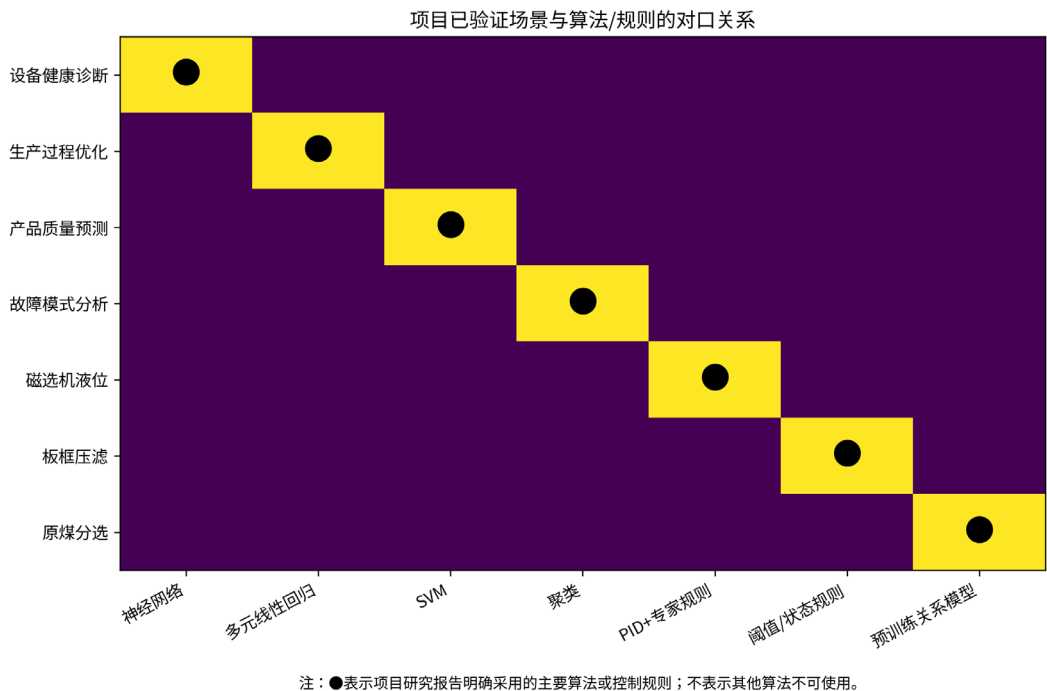


图 2 项目已验证场景与主要算法/规则对口关系

表 1 选煤典型任务的数据特征与对口模型

业务任务	主要数据特征	任务类型	对口模型/规则	核心评价指标	部署方式
设备健康诊断	振动、温度、压力、历史故障； 时序+标签	分类/诊断	神经网络	准确率、召回率、F1、响应时间	预警+维护决策
生产过程优化	工艺、质量、能耗； 连续变量	多变量回归/优化	多元线性回归	R ² 、显著性、残差、预测误差	参数推荐/闭环调节
产品质量预测	原煤性质+工艺参数； 标签明确	分类/预测	支持向量机	分类准确率、稳定性	质量预判+控制依据
故障模式分析	故障记录； 标签不完整	无监督模式识别	聚类分析	类间分离、案例可解释性	知识库/维修策略
磁选液位控制	液位、底流状态； 实时连续	反馈控制 +规则推理	PID+专家规则	液位稳定性、介质回收	自动调节
压滤卸料判定	流量、压力； 状态边界明确	状态识别	阈值/流量-压力规则	周期、含水稳定性	自动停料/卸料
原煤分选	灰分、粒度、压力、密度	关系预测+优化	灰分-密度-产率模型	产率、介耗、质量稳定	设定值计算+阀门联动

2 数据基础与模型构建

2.1 标准化数据底座

精准建模的前提是不同场景共享可信数据。项目基础信息模型接入设备信息、生产过程、安全环保、管理文档、视频监控及传感器数据,对数据进行清洗、去重、标准化、向量化和特征提取,并建立设备-工艺、生产-质量、设备-维护等关系。统一数据、接口和语义标准后,形成可供算法调用的标准化数据底座。

项目资料显示,基础信息模型数据覆盖率达到98%以上、准确率达到99%以上;通过模型压缩、层次化建模和动态加载,数据查询响应时间缩短至1.5 s以内,无故障运行时间达到99.9%以上。数据标准化的意义在于减少不同算法重复进行数据清洗和字段映射,使模型能够围绕业务目标直接训练与部署。

2.2 训练集构建与模型训练

项目将设备运行、生产工艺、故障记录和维护保养等数据经清洗、去重和标准化后,按7:3划分训练集和验证集。模型训练包括参数初始化、迭代训练、验证评估和参数优化4个步骤,并根据算法特点采用梯度下降、遗传算法等方法调整参数。

针对选煤场景,训练数据不宜简单按随机样本

划分,还应尽量覆盖不同煤质、负荷、设备状态和异常工况。对于时间序列任务,应避免未来数据泄漏到训练集;对于故障诊断,应关注少数类故障样本不平衡;对于工艺优化,应确保不同工况窗口具有足够代表性。上述扩展属于本文基于工程建模原则提出的优化建议,用于增强项目现有训练流程的泛化能力。

2.3 模型优化与生命周期管理

项目采用定期更新、增量学习、多模型融合和反馈调整进行模型优化。本文进一步将其归纳为“离线训练-验模-部署-性能监测-再训练-版本更新”的生命周期闭环。当出现新煤种、新设备、传感器更换或模型预测误差持续升高时,应触发模型更新;更新后的模型需重新通过离线指标和工业验证后才能替换线上版本。

对于关键控制模型,应同时保留原模型或规则控制作为回退路径。当实时数据缺失、模型置信度下降或输出超过工艺约束时,系统自动切换至规则/人工控制。这样可避免“模型越复杂风险越高”,实现智能算法与工业安全控制的分层协同。

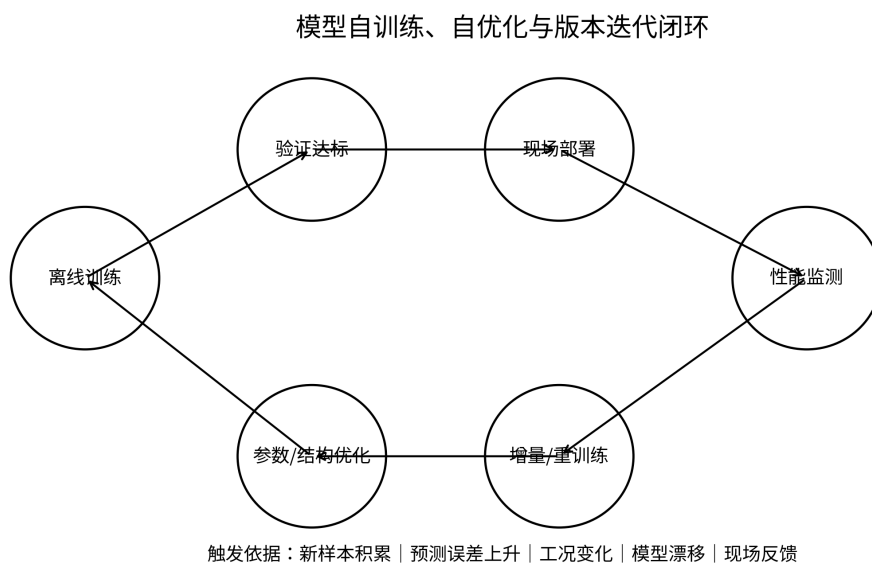


图3 模型自训练、自优化与版本迭代闭环

3 生产过程精准对口建模与优化构建

3.1 多元线性回归模型构建

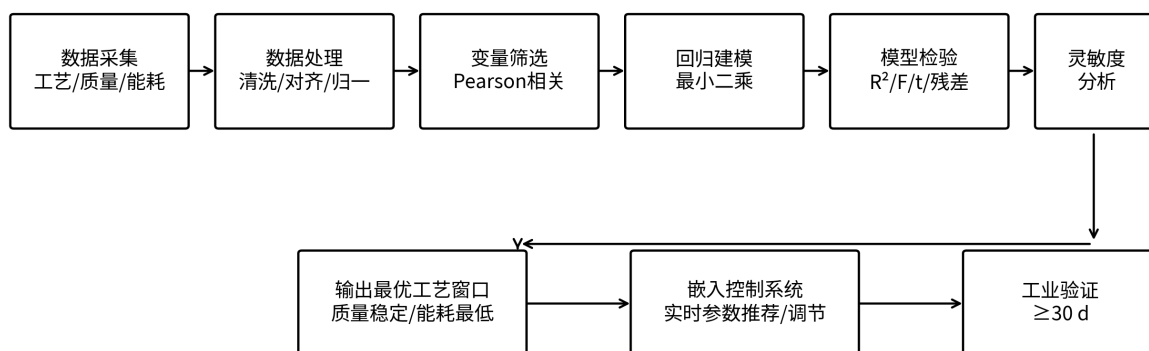
生产过程优化需要同时描述工艺参数、产品质量和能耗之间的关系。项目选择多元线性回归而非直接采用高复杂度黑箱模型,主要原因在于工艺变

量连续、样本数量较大且需要给出参数影响方向。以质量或能耗指标 y 为因变量,以入料浓度、分选密度、循环水量、筛机参数、介质流量和给料量等关键工艺参数 x_j 为自变量,可表示为:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \varepsilon$$

通过 Pearson 相关分析筛选显著变量后,采用最小二乘法求解回归系数。模型验收不仅考察 R^2 , 还包括 F 检验、各回归系数 t 检验和残差检验, 以保

证模型相关关系具有统计意义。项目规定 $R^2 \geq 0.85$ 作为拟合精度要求, 并在连续不少于 30 d 的工业验证中要求预测误差 $\leq 5\%$ 。



验收约束： $R^2 \geq 0.85$ ；F检验显著；回归系数t检验显著；残差满足要求；工业预测误差 $\leq 5\%$

图4 生产过程精准对口回归建模与闭环优化流程

3.2 从预测模型到工艺参数窗口

对口建模的目标不是得到一个预测值, 而是形成可执行工艺窗口。项目利用合格回归模型开展灵敏度分析, 分别寻找能耗最低、质量稳定合格和生产效率较高的参数组合, 并形成最优工艺区间。在系统运行中实时采集当前工艺参数, 模型计算质量与能耗趋势, 再给出分选密度、筛机频率、介质流量等调节建议, 由系统自动或人工执行。

这种“模型-窗口-执行”路径比直接让黑箱模型输出控制量更容易解释和管理。对关键变量可叠加上下限、变化率和设备能力约束, 确保推荐参数不突破工艺安全边界; 当不同目标发生冲突时, 可采用分层优先级, 例如先满足产品质量约束, 再优化介耗和能耗。

3.3 对高时滞、强耦合场景的模型升级路径

多元线性回归适合解释性较强、关系相对稳定的生产优化任务, 但对强时滞、显著非线性和多变量耦合过程存在局限。煤泥浮选研究表明, WOA-GRU 模型可利用 GRU 表征时序特性, 并结合模型预测控制处理多输入多输出耦合; 相关工业试验实现灰分波动降低^[3]。进一步的无模型深度强化学习也可通过在线交互学习控制策略^[4]。

因此, 精准对口建模应形成分级升级机制: 当线性模型已满足精度和稳定性要求时优先采用可解释模型; 当残差呈明显非线性或时滞效应导致预

测偏差时, 再升级至时序神经网络、模型预测控制 (MPC) 或强化学习, 并同步提高验证强度和安全约束。模型复杂度应由业务问题驱动, 而非以算法新颖性为唯一目标。

4 多场景模型集成与系统构建

4.1 专家知识库与算法模型协同

选煤现场存在大量难以用完整数学模型表达的经验知识。项目专家知识来源包括领域专家经验、历史生产数据、技术文献和现场实践, 采用规则、框架、语义网络和案例等形式组织。专家知识与数据模型并非替代关系: 神经网络、回归或支持向量机 (SVM) 负责从数据中提取模式, 专家规则负责异常工况解释、逻辑校核和控制边界^[5-6]。

例如磁选机液位控制采用 PID 维持连续稳定, 同时用“液位持续下降且底流浓度异常”等规则识别堵塞风险; 设备健康模型给出故障概率后, 可通过故障知识库匹配部位、原因和维修方法。混合模式能够提升模型输出的可解释性和可执行性。

4.2 数据—模型—控制一体化集成

项目最终形成基础设施层、数据源层、基础平台层、应用层和业务三维展示层的一体化系统。基础平台负责数据存储、处理、算法引擎和知识服务, 应用层部署设备健康、生产控制、安全与决策场景, 三维展示层实现设备状态和业务交互。精准对口模型并非独立算法文件, 而是以服务方式接入统一平

台, 由标准数据对象触发计算, 并将结果回写至控制、工单或可视化系统。

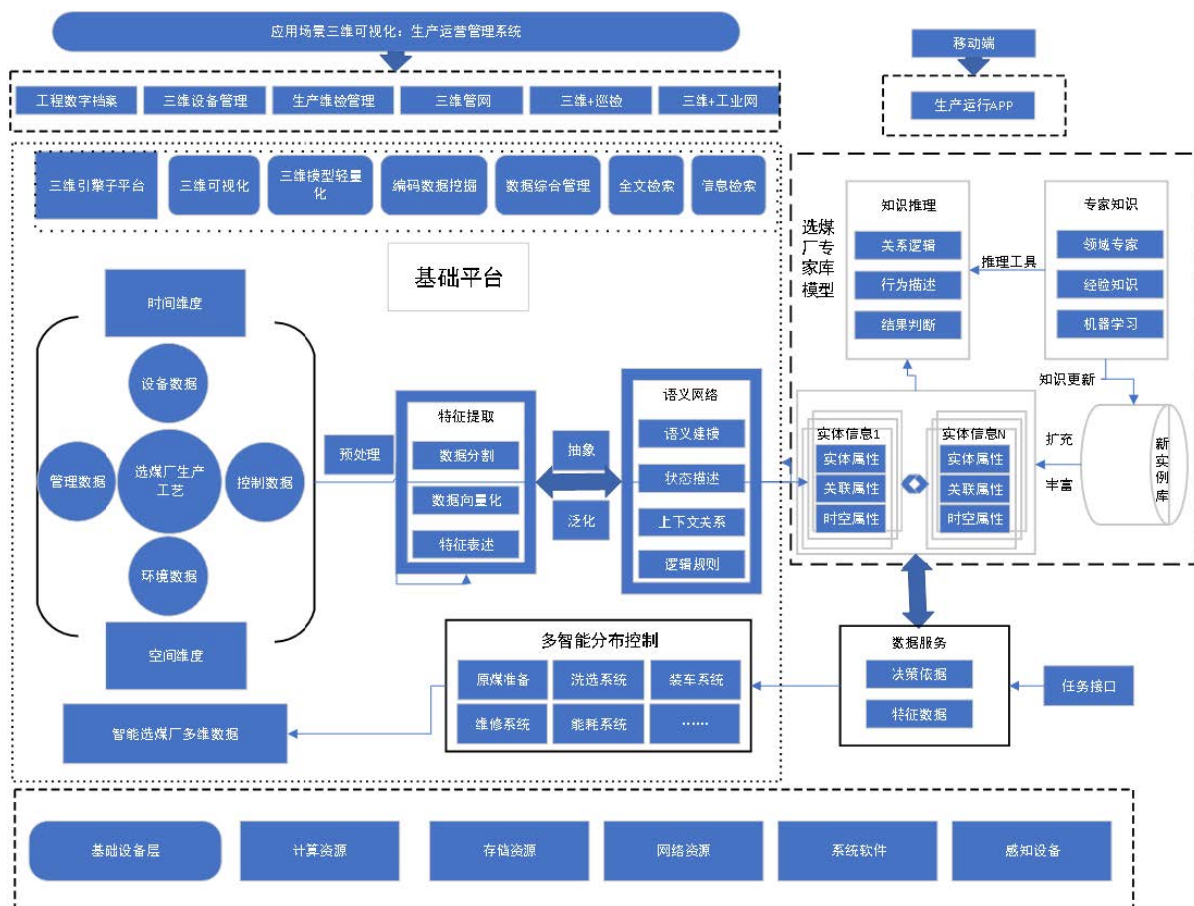


图5 智慧选煤数智管控一体化系统总体架构（源自项目总结报告）

4.3 场景模型的分级部署

依据模型输出风险, 可将部署方式划分为三级: 一级为诊断与提示, 仅输出状态、风险和建议, 适用于设备健康、故障知识检索等; 二级为辅助决策, 由模型输出参数窗口或工艺建议, 经人工确认后执行; 三级为自动闭环, 在输入可靠、控制边界清晰且经过工业验证后, 由系统直接调节执行机构, 如磁选机液位和压滤机状态控制。

分级部署可以使模型能力与安全责任相匹配, 也是精准对口建模从“算法适配”向“系统适配”的延伸。对于停送电等高风险场景, 即使具备数字化流程和五防逻辑, 也应保留必要的确定性联锁和人工安全程序。

5 工业应用与效果分析

5.1 典型场景效果

项目在文家坡选煤厂完成设备健康、磁选机液位、板框压滤、原煤分选、阀门远控、智能润滑和停

送电等场景应用。其中, 板框压滤机依据滤液流量阈值自动判断卸料时机, 单循环时间平均缩短约 15%; 原煤分选根据灰分、粒度、旋流器压力和介质密度动态计算分选密度, 场景试点精煤产率平均提升约 1.2 个百分点, 介质消耗降低约 16%; 智能润滑使润滑油消耗降低约 18%; 智能停送电平均操作时间缩短约 40%以上。

这些场景并未采用统一算法, 而是根据数据结构和控制任务选择不同模型或规则, 验证了精准对口思路的工程合理性。尤其是状态边界明确的压滤机采用阈值规则, 而不是额外引入复杂机器学习模型, 说明对口建模强调“适合”而非“复杂”。

5.2 综合验收指标

项目最终验收结果显示, 精煤回收率提升 1.1%, 介耗降低 16%, 每天平均减少停机时间 0.55 h, 运维作业时间减少 6%, 设备故障诊断准确率达到 93%, 年直接经济效益 341.9 万元, 投资回收期

约 4 年。系统同时在线人数超过 1000 人，核心页面响应时间小于 2 s。上述结果表明，模型性能提升需要与数据治理、系统集成和控制执行共同作用，单独比较算法精度不能完全反映数智化系统价值。

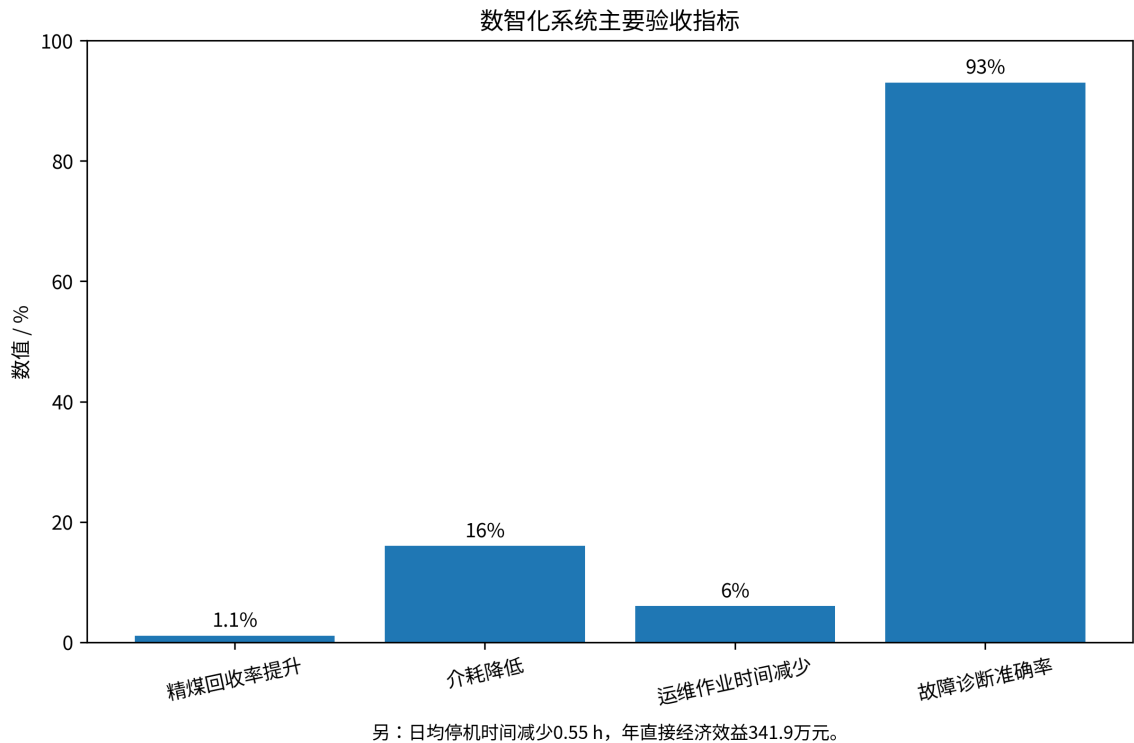


图 6 数智化系统主要验收指标

表 2 精准对口建模的模型指标与工程指标对应关系

评价层级	主要指标	项目/方法要求	工程意义
数据层	覆盖率、准确率、响应	覆盖率>98%；准确率>99%；查询≤1.5 s	保证模型输入可信和实时
分类诊断层	准确率、精确率、召回率、F1	≥92%、≥90%、≥88%、≥89%	保证故障识别有效性
回归预测层	R ² 、显著性、残差、预测误差	R ² ≥0.85；P<0.05；工业误差≤5%	保证过程模型可解释、可应用
控制层	稳定性、周期、能耗、产率	按场景工业验证	验证模型能转化为控制效果
系统层	停机、运维、介耗、回收率	0.55 h/d；-6%；-16%；+1.1%	评价整体数智化收益
经济层	直接效益、投资回报	341.9 万元/年；约 4 年	评价工程推广可行性

6 讨论

6.1 精准对口建模的工程优势

与“一个平台配置一个统一算法”的建设方式相比，精准对口建模有 3 方面优势。首先，算法可解释性和任务要求一致，能够降低模型部署难度；其次，评价指标与业务效果直接对应，避免仅以离线精度评价模型；再次，模型可按场景独立升级，不需要因单个工艺变化重构整个平台。对于工况复杂、设备多样的选煤厂，这种模块化模型体系更利于复制推广。

6.2 现有模型体系的进一步优化方向

项目模型体系仍存在进一步完善空间。一是建立统一模型注册与版本管理，记录训练数据、参数、适用工况、上线时间和回退版本；二是引入概念漂移和数据质量监测，发现煤质或设备状态变化导致的性能衰减；三是对复杂时序和强耦合场景增加 GRU、MPC、强化学习等模型候选，但必须通过对比试验确认其相对于简单模型的增益；四是加强模型不确定性评估，当置信度不足时自动转人工或规则控制。

此外，项目报告中不同阶段对精煤回收率存在约 0.8%与最终 1.1%的统计口径差异。本文综合效果

采用验收总结反复出现的最终口径, 而阶段性数据仅作为试运行过程信息。后续工程研究应统一基准期、统计周期和工况范围, 避免模型效果与系统整体效益在不同口径下混用。

6.3 从精准建模向系统级协同优化发展

当前精准对口建模主要解决“一个业务场景应选什么模型”的问题, 下一阶段需要解决“多个场景如何协同优化”。例如原煤性质变化不仅影响分选密度, 也会影响介质循环、脱水负荷和能耗; 设备健康下降后, 系统可在生产计划允许范围内降低负荷并自动生成维护窗口^[7-8]。

因此, 应在场景模型之上增加系统级协调层, 将质量、产率、能耗、设备健康和安全约束纳入统一多目标优化, 同时保留局部模型的专业性。这样才能实现从单点精准建模到全厂数智协同控制的进一步跃迁。

7 结论

1) 提出了面向选煤数智化系统的“场景任务-数据画像-模型对口-指标验模-部署闭环”精准对口建模方法, 以任务类型、数据条件、实时约束和控制风险为依据选择算法, 避免复杂模型的泛化套用。

2) 项目形成设备健康-神经网络、生产优化-多元线性回归、质量预测-支持向量机、故障分析-聚类、磁选液位-PID 与专家规则、压滤机-状态阈值规则等差异化模型体系。对分类模型设置准确率、精确率、召回率和 F1 指标, 对生产回归模型设置 R^2 、显著性、残差和工业预测误差约束, 实现模型可量化验收。

3) 构建离线训练、工业验证、在线监测、增量学习和版本更新的模型优化闭环, 并通过统一基础信息模型和专家知识库实现数据-模型-控制集成。工业应用最终实现故障诊断准确率 93%、精煤回收率提升 1.1%、介耗降低 16%、日均停机时间减少 0.55 h 和年直接经济效益 341.9 万元。

4) 精准对口建模的下一步应由单场景模型匹配向系统级协同优化发展, 重点加强模型版本管理、数据漂移监测、复杂时序模型升级、安全回退以及质量-能耗-设备健康多目标协同, 为选煤厂无人化和自优化运行提供模型基础。

参考文献

- [1] 母长春, 杨国峰. 我国智能化选煤厂技术与展望[J]. 选煤技术, 2024, 52(3): 1-8.
- [2] 申斌学. 基于矿山工业互联网的选煤厂智能管控体系设计与关键技术[J]. 煤炭科学技术, 2024, 52(9): 210-219.
- [3] 窦治衡, 王然风, 秦新凯, 等. 基于 WOA-GRU 模型的煤泥浮选智能控制研究[J]. 工矿自动化, 2025, 51(4): 153-159, 168.
- [4] 秦新凯, 王然风, 付翔, 等. 基于无模型深度强化学习的煤泥浮选智能控制研究[J]. 工矿自动化, 2025, 51(8): 25-33, 58.
- [5] 张帆, 葛世荣, 李闯. 智慧矿山数字孪生技术研究综述[J]. 煤炭科学技术, 2020, 48(7): 168-176.
- [6] 邢震. 面向智能矿山的数字孪生技术研究进展[J]. 工矿自动化, 2024, 50(3): 22-34, 41.
- [7] 胡亚辉, 赵国瑞, 吴群英. 面向煤矿智能化的 5G 关键技术研究[J]. 煤炭科学技术, 2022, 50(2): 223-230.
- [8] 孙继平. 矿山人工智能大模型标准研究制定[J]. 工矿自动化, 2025, 51(10): 1-6.

版权声明: ©2026 作者与开放获取期刊研究中心(OAJRC)所有。本文章按照知识共享署名许可条款发表。

<https://creativecommons.org/licenses/by/4.0/>



OPEN ACCESS