

# 基于机器学习融合的石化企业抢维修业务成本预测与优化研究

谷润泽

中原油田地面工程抢维修中心计划财务办公室

DOI:10.32629/ej.v9i7.3601

**[摘要]** 针对抢维修业务突发、高成本、调度复杂及传统核算粗放、预测不准问题,本文以A油田为例,融合ABC、机器学习与多目标优化,构建“核算—预测—优化”框架。ABC分解四作业中心,建立时间/距离/技能/物料四维模型,MAPE=4.2%。基于9246条工单与20特征,XGBoost预测MAPE=7.8%、 $R^2=0.878$ ,误差较经验降76%。将预测成本嵌入多目标模型,改进遗传算法求解,30天仿真成本降27.7%、响应时间缩27.6%、资源利用率由68%升至82%( $p<0.001$ )。

**[关键词]** 作业成本法; 机器学习; XGBoost; 多目标优化; 遗传算法; 成本预测

中图分类号: TP24 文献标识码: A

## Cost Prediction and Optimization for Emergency Repair Services in Petrochemical Enterprises Based on ABC–Machine Learning Integration: A Case Study of A Oilfield Maintenance Center

Runze Gu

Zhongyuan Oilfield Ground Engineering Maintenance Center Planning and Finance Office

**[Abstract]** Aiming at the prominent problems of emergency maintenance services, including sudden service demands, high operation costs, complex scheduling, extensive traditional cost accounting and inaccurate cost forecasting, this paper takes Oilfield A as the research object. By integrating Activity–Based Costing, machine learning and multi–objective optimization, a three–tier framework of "cost accounting – cost forecasting – resource optimization" is constructed. Four operational centers are decomposed based on ABC theory, and a four–dimensional accounting model covering service time, travel distance, staff skill level and material consumption is established, with the mean absolute percentage error of cost calculation reaching 4.2%. Based on a dataset consisting of 9,246 maintenance work orders and 20 characteristic variables, an XGBoost prediction model is trained. The model achieves a MAPE of 7.8% and a coefficient of determination ( $R^2$ ) of 0.878, reducing forecasting errors by 76% compared with traditional empirical estimation methods. Predicted cost values are embedded into the multi–objective optimization model, and an improved genetic algorithm is adopted to solve the model. The 30–day simulation experiment shows that the overall operation cost is cut by 27.7%, emergency response time is shortened by 27.6%, and the resource utilization rate rises from 68% to 82% ( $p < 0.001$ ), with statistically significant optimization effects.

**[Key words]** Machine Learning; XGBoost; Multi–objective Optimization; Genetic Algorithm; Cost Prediction.

### 1 引言

石油工业是国民经济的战略性支柱产业。随着国际油价波动加剧、能源转型加速以及碳中和目标的提出,石油企业面临前所未有的成本压力。抢维修作为油田生产保障体系的核心环节,具有突发性强、响应要求高、成本波动大的特点,其管理水平直接关系到生产连续性和企业经济效益。行业统计显示,维修成本在运营总成本中占比通常超过20%。以A油田维修中心为例,该单

位年均处理维修工单约3216次、年维修成本约8050万元。

当前成本核算多按部门或设备类型平均分摊,难以反映单次维修的真实成本构成与深层驱动因素;调度决策主要依赖经验判断,缺乏科学的成本预测与优化工具,导致夜间维修、紧急采购等低成本作业频繁发生。作业成本法虽能精确追溯成本动因,但属于事后核算工具,缺乏预测能力;机器学习擅长预测却需高质量训练数据;维修调度优化研究多以历史平均或专家经

验设定目标函数，难以反映成本的动态变化。三条研究脉络在交汇处留有空白：核算不精则预测缺乏可靠标签，预测不准则优化失去动态依据。

基于此，本文将作业成本法、机器学习与多目标优化贯通，构建“核算—预测—优化”一体化框架：以ABC精准核算结果作为机器学习训练标签，再将预测成本嵌入优化模型，形成“预测—优化”闭环决策，推动抢维修管理从经验驱动向数据驱动转变。

2 基于 ABC 的成本现状分析与精细化核算

2.1 成本现状与驱动因素

2022–2024 年数据：年均总成本 8050 万元，人工 41.9%、材料 35.1%、间接 23%，人工+材料=77% 为关键。人工受技能与时段影响：工时单价初级工 150→高级技师 350 元/h；时段系数白天 1.0、晚间 1.5、夜间 2.0–2.5，节假日再乘 1.3–1.5。夜间工单占 35% 但人工成本占 44%。紧急采购占总额 28%，溢价 50%–100%。2024 年 3217 条工单帕累托：前 20% 高成本工单累计占 80%。高成本工单画像：夜间/节假日、远距离、高故障、紧急采购、高技能。距离每增 10km 成本约升 8%；夜间比白天增 85%；紧急采购比库存增 60%。

2.2 ABC 核算模型与多维度成本结构

作业成本法的核心思想是“作业消耗资源，产品消耗作业”，通过两阶段分配建立资源消耗与成本对象之间的因果关系。本文将抢维修流程分解为应急响应、故障诊断、现场维修、验收交付四个作业中心，并对以时间为主要动因的资源采用时间驱动方式核算，工单成本表示为各作业中心单位时间产能成本率与实际耗时乘积之和。

在此基础上构建时间、距离、技能、物料四个维度的成本模型，系统揭示了抢维修成本的驱动机制：时间维度刻画夜间与节假日溢价；距离维度体现交通成本与路途时间成本；技能维度引入技能匹配度系数量化效率与返工影响；物料维度区分库存备件与紧急采购的成本差异。表 1 给出四个作业中心的主要活动及成本动因。

表 1 作业中心划分及主要成本动因

| 作业中心 | 主要活动                | 资源消耗类型   | 主要成本动因          |
|------|---------------------|----------|-----------------|
| 应急响应 | 调度接单、人员召集、车辆调配、路途交通 | 人工、车辆、通讯 | 维修距离、响应次数、时段系数  |
| 故障诊断 | 现场勘查、参数测试、故障分析、方案制定 | 人工、检测工具  | 诊断工时、故障复杂度、技能等级 |
| 现场维修 | 设备拆卸、部件更换、系统调试、功能测试 | 人工、材料、工具 | 维修工时、技能等级、备件数量  |
| 验收交付 | 质量检验、试运行、性能测试、工单结算  | 人工、测试设备  | 验收工时、测试项目数量     |

2.3 模型验证

选取 10 个典型工单作为验证样本，分别用传统平均分摊法和 ABC 方法核算并与财务审计确认的实际成本对比（表 2）。结果

表明，传统方法对夜间、远距离维修普遍低估，对白天简单维修普遍高估，系统性偏差明显，平均绝对百分比误差高达 34.5%；ABC 方法误差在各类工单中分布均匀，MAPE 仅约 2.3%–4.2%，准确性显著提升，为后续机器学习建模提供了高质量训练标签。

表 2 ABC 核算与传统方法对比（部分案例）

| 案例         | 时段  | 实际成本(元) | 传统方法(元) | ABC 方法(元) | ABC 误差 |
|------------|-----|---------|---------|-----------|--------|
| 夜间远距离 A 级  | 夜间  | 15,692  | 8,950   | 15,438    | -1.6%  |
| 白天近距离 C 级  | 白天  | 2,904   | 5,200   | 2,978     | 2.5%   |
| 节假日中距离 B 级 | 节假日 | 17,799  | 6,800   | 17,555    | -1.4%  |
| MAPE       | —   | —       | 34.5%   | 2.3%      | —      |

3 基于机器学习的成本预测模型

3.1 数据与模型

9638 条清洗得 9246 条，按 70%/15%/15% 划分，Z-score 标准化。构建含距离分段、时段溢价、技能匹配度等 27 个特征，筛选保留 23 个。随机森林与 XGBoost 对比：RF：MAE=536，RMSE=768，R<sup>2</sup>=0.854，MAPE=8.5%；XGB：492, 715, 0.878, 7.8%，XGB 全面占优。散点图紧贴 y=x，残差正态。较经验估计误差降 76%，解释力提 35%。

3.2 SHAP 分析

重要性前五：维修距离、派工时段、故障等级、技能等级、备件采购方式，合理解释 75% 变异。依赖图：0–30km 近似线性，30–50km 加速，>50km 陡升（两拐点）；同距离夜间 SHAP 高于白天，呈“乘法效应”。优化靶点：避免夜间远距离叠加，提高备件库存，优化技能匹配。4 融合成本预测的多目标调度优化。

4 融合成本预测的多目标调度优化

4.1 多目标优化模型

将 XGBoost 预测成本嵌入优化模型目标函数，构建包含成本最小化、加权响应时间最短、资源利用率最高三个目标的数学模型。决策变量为任务与技术员的匹配方案；约束条件包括任务分配唯一性、技能匹配、任务开始与完成时间、技术员工作时长上限、任务不重叠及调度时域等。采用加权求和法将三目标统一为单目标，权重设定为成本 0.5、响应时间 0.3、资源利用率 0.2，体现成本优先兼顾效率与资源平衡的管理导向。该模型将机器学习预测成本直接嵌入目标函数，实现了“ABC 精准核算→ML 高精度预测→优化智能调度”的完整闭环。

4.2 改进遗传算法

抢维修调度属典型 NP-hard 组合优化问题，本文设计改进遗传算法求解：采用整数编码表示任务—技术员分配；初始种群采用随机、距离优先与负荷均衡三种策略混合生成；适应度函数采用惩罚函数法处理不可行解；选择算子采用轮盘赌结合精英保留；交叉算子采用顺序交叉并修复技能约束；变异概率采用自适应策略，由初期 0.05 随迭代提升至后期 0.15。算法通常在 60–80 代收敛，单次优化耗时约 8–15 秒，满足实时调度要求。

4. 3 仿真实验与结果

采用Python开发离散事件仿真系统,模拟抢维修中心30天运营。任务到达服从泊松分布( $\lambda=8$ 个/天),技术员配置15人,采用滚动时域优化,每方案运行30次蒙特卡洛实验。设置三种对比方案:方案A为现状;方案B为ABC核算+历史均值+启发式调度;方案C为使用作业成本法+XGBoost+改进遗传算法的完整方法。三方案核心指标对比见表3与图1。

表3 三方案性能对比

| 评价指标       | 方案A   | 方案B   | 方案C   | CvsA    |
|------------|-------|-------|-------|---------|
| 30天总成本(万元) | 206.8 | 175.6 | 149.5 | -27.7%  |
| 平均响应时间(分钟) | 58    | 46    | 42    | -27.6%  |
| 资源利用率(%)   | 68    | 75    | 82    | +14pp   |
| 夜间维修占比(%)  | 35    | 28.5  | 18.3  | -16.7pp |
| 成本波动系数CV   | 0.58  | 0.45  | 0.36  | -37.9%  |

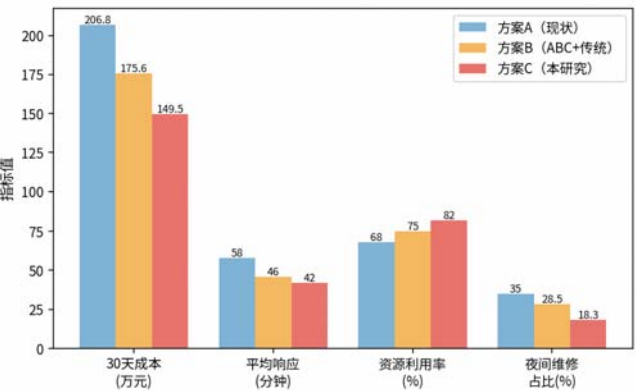


图1 三方案综合性能对比柱状图

仿真结果表明,方案C相比现状方案A,30天成本由206.8万元降至149.5万元,降幅27.7%,按年推算节约约2230万元,达成“年成本降低20%—30%”的预期目标;平均响应时间由58分钟缩短至42分钟,改善27.6%;资源利用率由68%提升至82%;夜间派工次数减少48.6%。方案C相比方案B成本再降14.9%,体现了机器学习预测与遗传算法优化的协同价值。

配对t检验显示,方案C相比方案A与方案B的成本优势均高度显著( $p<0.001$ ),Cohen's d效应量分别为2.18和1.35,均属大

效应。敏感性分析进一步验证了模型鲁棒性:在任务量 $\pm 30\%$ 波动场景下,方案C的成本优势始终保持在25%—30%区间;在技术员减少20%的资源紧张场景下,方案C成本增幅仅11.6%,明显低于方案A的18.5%;在成本参数 $\pm 10\%$ 扰动下优化结果呈线性响应,不会大幅偏离最优解。

5 结论

本文针对成本核算粗放、预测不准、调度非最优、信息化支持不足四大问题,构建了作业成本法、机器学习与多目标优化融合的“核算—预测—优化”一体化成本管理框架,主要结论如下:

(1) 基于作业成本法将业务分解为四个作业中心并建立时间、距离、技能、物料四维成本模型,识别出维修距离、作业时段、技能等级、备件采购方式等关键成本动因,ABC核算MAPE为4.2%,显著优于传统方法的34.5%。

(2) 基于9246条工单与20个特征变量构建XGBoost预测模型,测试集MAPE7.8%、 $R^2$  0.878,较经验估计误差降低76%;SHAP分析揭示了距离—成本的非线性关系与时段—成本的乘法效应。

(3) 将预测成本嵌入多目标优化模型并以改进遗传算法求解,30天仿真显示成本降低27.7%、响应时间缩短27.6%、资源利用率提升至82%,优化效果高度统计显著。

研究的主要创新在于拓展了作业成本法在应急服务领域的应用边界,构建了“预测—优化”闭环决策框架,实现了ABC、机器学习与优化理论的深度融合。研究尚存样本来源单一、未引入深度学习与预防性维护等局限,未来可结合深度学习、数字孪生与强化学习进一步深化,并向其他油田及行业推广。

[参考文献]

[1]徐春茸.基于作业成本法的石化企业成本核算研究[J].市场周刊,2026,39(9):142-145.  
[2]Geurtsen M,Adan J,Atan Z.Production, maintenance and resource scheduling:A review[J]. European Journal of Operational Research,2023,305(2):501-529.  
[3]赵志成,马璐璐,孙萌颖.降本增效视角下油气企业数智化转型的思考与建议[J].中外能源,2026,31(5):1-5.

作者简介:

谷润泽(2001—),男,汉族,河南濮阳人,助理工程师,主要从事财务管理与成本控制研究。