

大数据领域指标数据管理方法及系统的研究

彭华养

珠海华发集团科技研究院有限公司, 广东 珠海 519000

DOI: 10.61369/TACS.2025120008

摘 要 : 新能源电池储能与现代农业领域数据来源多样、结构复杂, 现有指标数据管理模式在数据一致性、实时性与风险预警方面仍存在不足。针对上述问题, 构建了一套基于大数据技术的指标数据管理系统, 系统设计遵循可扩展性、可靠性与通用性原则, 采用分层架构实现数据采集、处理、存储与应用协同运行。重点研究多源数据融合、动态质量指标提取、多维度风险评估、实时预警决策树及数据标准化处理等关键技术。通过系统实现与行业应用验证, 结果表明该系统能够有效提升新能源储能运行监测与现代农业生产管理中的数据治理水平和风险识别能力。尽管在复杂工况数据采集和跨系统协同方面仍存在一定局限, 但结合数字孪生技术可进一步增强系统的智能分析与决策支持能力。

关 键 词 : 新能源电池储能; 现代农业; 指标数据管理系统; 大数据技术

Research on Management Methods and Systems for Indicator Data in the Field of Big Data

Peng Huayang

Zhuhai Huafa Group Technology Research Institute Co., Ltd., Zhuhai, Guangdong 519000

Abstract : The data sources in the fields of new energy battery storage and modern agriculture are diverse and structurally complex. Existing indicator data management systems still exhibit shortcomings in data consistency, real-time performance, and risk early warning capabilities. To address these challenges, a big data-based indicator data management system has been developed. The system design adheres to principles of scalability, reliability, and versatility, employing a hierarchical architecture to achieve coordinated operations in data collection, processing, storage, and application. Key research focuses include multi-source data fusion, dynamic quality indicator extraction, multi-dimensional risk assessment, real-time early warning decision trees, and data standardization. Through system implementation and industry application validation, the results demonstrate that this system effectively enhances data governance levels and risk identification capabilities in new energy storage operation monitoring and modern agricultural production management. Although limitations persist in complex condition data acquisition and cross-system collaboration, integration with digital twin technology can further strengthen the system's intelligent analysis and decision support capabilities.

Keywords : new energy battery storage; modern agriculture; indicator data management system; big data technology

引言

随着数字化与智能化进程的加快, 大数据技术在实体产业中的应用不断深化。2023 年发布的《关于推进城市基础设施数字化转型的指导意见》明确提出, 要以数据治理为抓手, 提升重点行业的运行监测与风险防控能力。新能源电池储能与现代农业作为典型的数据密集行业, 在生产运行和管理过程中产生了大量多源异构指标数据, 对指标数据的统一管理 with 高效利用提出了更高要求。新能源电池储能系统涉及电压、电流、温度及电池健康状态等多维度运行指标, 数据更新频率高、实时性要求强; 现代农业依托传感器和信息化平台采集环境参数、作物生长及生产管理数据, 数据类型复杂、标准不统一。现有指标数据管理模式普遍存在数据分散、质量评估静态化、风险预警依赖经验等问题, 难以支撑精细化管理与科学决策。在此背景下, 本文围绕大数据领域指标数据管理方法及系统展开研究, 遵循可扩展性、实时性与安全性原则, 采用分层架构实现多源数据融合、动态质量指标提取、多维度风险评估与实时预警。通过系统实现与应用验证, 表明该方法能够有效提升新能源储能与现代农业领域的指标数据管理与风险识别能力, 为相关行业的数据治理与智能决策提供技术支撑。

一、新能源电池储能与现代农业指标数据管理现状分析

（一）行业指标数据特征

新能源电池储能与现代农业在运行与生产过程中均呈现出显著的多源异构数据特征。新能源电池储能系统运行数据主要包括电压、电流、温度、荷电状态（SOC）、健康状态（SOH）等关键指标，数据采集依赖大量传感器和监测设备，具有采样频率高、实时性要求强的特点，对设备稳定性和数据传输可靠性提出了较高要求。现代农业领域的的数据则涵盖土壤墒情、气象环境、作物生长状态、农业投入与作业记录等信息，数据来源包括传感器、遥感影像及管理系统平台，既包含结构化数据，也包含文本和图像等非结构化数据。上述两类行业数据在采集过程中普遍面临数据标准不统一、采集环境复杂、数据质量易受外界因素影响等问题。例如，储能系统传感器易受工况波动影响产生噪声数据，农业生产记录存在主观性强、格式多样等情况，增加了数据治理与管理难度^[1]。

（二）现行管理模式缺陷

当前新能源电池储能与现代农业领域的指标数据管理模式仍存在一定不足。在实时处理方面，部分企业仍采用分散采集、定期汇总的方式，数据更新滞后，难以及时反映系统运行状态或农业生产过程变化^[2]。在数据关联分析方面，不同业务系统之间数据相互割裂，缺乏统一的数据治理与指标管理框架，导致难以从整体视角分析储能运行安全与农业生产效率之间的潜在关系。在风险预警环节，现行管理方式多依赖经验判断或简单阈值规则，缺乏基于大数据分析的动态评估模型，难以及时识别电池运行异常或农业生产风险。这些问题在一定程度上制约了新能源储能系统安全运行和现代农业精细化管理水平的提升，亟需引入更加系统化、智能化的指标数据管理方法加以解决。

二、指标数据管理系统架构设计

（一）系统设计原则

面向新能源电池储能与现代农业的指标数据管理系统设计，需要遵循一系列基本原则，以满足系统运行监测、质量追溯与风险动态评估等应用需求。在可扩展性方面，考虑到新能源储能设备规模扩大及农业生产场景不断丰富，系统架构应具备良好的扩展能力，能够灵活接入新的数据源与指标类型，并支持存储与计算资源的弹性扩展，以适应业务持续增长^[3]。实时性原则要求系统能够对高频运行数据和生产过程数据进行及时处理与分析，为储能系统运行状态评估和农业生产过程调控提供快速、有效的数据支撑，从而辅助管理人员及时作出决策。安全性原则同样至关重要，由于指标数据涉及设备运行状态、生产管理信息等关键内容，系统需通过数据加密、访问控制和身份认证等技术手段，保障数据在采集、传输与存储过程中的安全性，防止数据泄露和非法篡改，确保数据的保密性、完整性与可用性。

（二）分层体系架构

指标数据管理系统采用分层体系架构，以提升系统的可维护性与运行效率。数据采集层负责从新能源电池储能系统监测设备、农业传感器及业务管理平台等多类数据源获取指标数据，需具备对不同通信协议和数据格式的高效适配能力^[4]。存储计算层对采集的数据进行集中存储与计算处理，采用 Hadoop 分布式文件系统等技术实现海量数据的可靠存储，并结合 Spark 等大数据计算框架完成高效的数据清洗、统计分析与特征提取。模型分析层基于存储计算层的数据，构建指标分析与风险评估模型，包括统计分析模型和机器学习模型，用于挖掘新能源储能运行状态和农业生产过程中的潜在规律与风险特征。应用服务层则将分析结果以可视化界面、报表或 API 接口的形式对外提供，支持运行监控、风险预警与管理决策等多种应用场景。各层之间通过标准化接口和消息机制实现解耦与协同运行，构建稳定、高效的指标数据管理系统。

三、核心管理方法研究

（一）数据处理关键技术

1. 多源数据融合技术

在新能源电池储能与现代农业领域的指标数据管理中，多源数据融合技术是提升数据治理能力的关键环节。新能源储能系统运行过程中会产生来自电池管理系统（BMS）、环境监测设备及能量管理平台的多源数据；现代农业则依托传感器、遥感及生产管理系统采集环境参数、作物生长和生产作业数据。这些数据在时间粒度、空间分布和数据格式方面存在明显差异，需通过有效的时空配准方法实现统一。时空配准的核心在于对不同数据源的时间戳和空间位置进行对齐，保证各类指标数据在同一时间和空间基准下进行分析。在此基础上，采用特征级或指标级数据融合方法，将经配准的数据进行整合，形成更加完整、可靠的指标数据集。通过多源数据融合，可有效减少单一数据源带来的偏差，提升指标数据的完整性与一致性，为新能源储能运行状态评估和现代农业生产分析提供可靠数据支撑^[5]。

2. 动态质量指标提取

动态质量指标提取是大数据环境下实现精细化管理的重要手段。在新能源电池储能与现代农业场景中，系统通过对运行与生产过程中产生的海量数据进行分析，利用机器学习方法实现质量指标的动态提取。具体而言，从多源异构数据中挖掘与设备运行安全、生产稳定性密切相关的潜在特征，如电池性能波动特征、环境参数变化趋势等。通过数据挖掘与特征选择算法，对原始指标进行筛选与组合，提取能够实时反映系统运行状态和生产质量变化的关键指标。相较于传统静态指标体系，该方法能够根据运行工况和生产条件变化动态调整指标内容，提高指标的适应性和时效性。同时，结合实际应用场景对指标进行持续优化，为后续异常检测与风险评估模型提供高质量输入数据，从而提升模型分析的准确性与可靠性^[6]。

（二）风险预警模型构建

1. 多维度风险评估模型

多维度风险评估模型从多个角度对指标数据管理过程中的潜在风险进行综合分析。在新能源电池储能与现代农业应用中，风险评估需重点关注数据准确性、完整性、一致性以及运行环境不确定性等因素。通过引入蒙特卡洛模拟方法，对指标数据在不同随机扰动条件下的变化情况进行多次模拟，获取风险发生的概率分布，实现对不确定性的量化分析。同时，利用深度神经网络对历史数据进行训练，自动学习指标之间的非线性关系和潜在模式，识别影响系统运行安全和生产稳定性的关键风险特征。蒙特卡洛模拟与深度神经网络的融合能够实现优势互补，前者用于刻画随机性和不确定性，后者用于挖掘复杂特征关系，共同构建多维度风险评估模型，为指标数据管理提供更加全面、精准的风险识别能力^[7]。

2. 实时预警决策树

实时预警决策树是实现风险快速响应的重要工具。该方法结合规则引擎与流式计算技术，对新能源储能系统运行数据和农业生产过程数据进行实时分析。规则引擎依据业务需求和管理经验设定关键指标阈值和判断规则，当数据触发特定条件时自动执行相应逻辑判断。流式计算框架对持续输入的数据进行实时处理，保证预警分析的时效性。在此基础上构建决策树结构，以不同节点表示指标判断条件，分支表示风险等级或处理策略。当指标数据满足某一节点条件时，系统沿对应分支快速给出预警结果，实现对运行异常或生产风险的即时识别与响应，为新能源电池储能和现代农业领域的指标数据管理提供高效的风险预警手段^[8]。

四、系统实现与工程验证

（一）工程数据集构建

1. 数据标准化处理

在新能源电池储能与现代农业指标数据管理系统的实现过程中，数据标准化处理是工程数据集构建的核心环节。针对储能系统运行数据和农业生产数据来源多样、格式不统一的问题，依据统一的数据分类与指标编码规范^[9]，对不同来源的数据进行标准化编码，确保指标口径一致。对于数值型数据，如电压、电流、温度、环境参数等，采用归一化或标准化方法消除量纲差异，提升不同指标之间的可比性；对于文本型数据，如农业生产记录、运维日志等，利用自然语言处理技术进行清洗与规范化处理，去除冗余信息和无效字符。通过构建质量校验规则体系，对标准化后的数据从准确性、完整性和一致性等多个维度进行校验，为后续模型分析和系统稳定运行提供高质量数据基础。

2. 分布式存储方案

在系统实现与工程验证中，需构建覆盖多个项目的数据集，采集新能源储能运行指标和现代农业生产指标，确保数据的全面性与代表性。针对数据规模大、并发访问频繁的特点，系统采用基于 Hadoop 的分布式文件系统（HDFS）作为底层存储方案^[10]。HDFS 通过数据分块与多副本机制提升数据存储的可靠性和容错

能力，并结合合理的数据块大小和副本策略优化存储性能。同时，依托分布式架构实现数据的并行读写，为后续跨项目数据仓库构建、指标查询与分析提供高效支撑。

（二）系统功能模块开发

1. 可视化监控模块

可视化监控模块面向新能源电池储能与现代农业应用场景，构建基于时间序列的指标数据可视化平台。模块对采集的运行与生产数据进行实时处理和分析，并以图表形式直观展示关键指标变化趋势，如储能系统运行状态、农业环境参数变化等。系统支持多维度指标联动展示和交互式查询，用户可按时间、项目或指标类型灵活筛选数据，实现对运行状态和生产过程的动态监控，为管理人员快速掌握系统运行情况提供直观支持。

2. 智能决策模块

智能决策模块通过整合多源指标数据，利用大数据分析与机器学习模型，为管理决策提供智能支持。系统对历史运行数据和实时监测数据进行分析，挖掘指标之间的关联关系和变化规律。在新能源储能场景下，模块可辅助判断运行异常和潜在风险；在现代农业场景下，可为生产调控和风险防范提供决策建议。模块注重与其他功能模块的协同运行，确保数据流转顺畅，实现指标数据管理的智能化与精细化。

（三）典型工程验证

1. 新能源电池储能系统运行监测案例

在新能源电池储能系统运行监测案例中，基于所构建的指标数据管理系统，对某储能项目进行持续运行监测。系统采集电压、电流、温度及电池状态等关键指标，并对数据进行实时分析。在运行过程中，系统成功识别出多次异常波动情况，并及时触发预警。经运维人员核查，预警结果与实际运行状态高度一致，有效降低了运行风险，验证了指标数据管理系统在储能系统运行监测中的实用性和可靠性。

2. 现代农业生产风险管控案例

在现代农业生产风险管控案例中，系统对农业环境与生产过程数据进行持续监测。应用系统前，生产异常主要依赖人工经验判断，问题发现滞后。系统投入使用后，通过对环境参数和生产指标的实时分析，异常识别效率显著提升。系统能够及时发现环境异常或生产偏差，并自动推送预警信息，辅助管理人员快速采取应对措施。实践表明，该指标数据管理系统在提升农业生产风险识别能力和管理效率方面具有良好的应用效果，验证了其在现代农业领域的工程应用价值。

五、总结

大数据领域指标数据管理方法及系统在新能源电池储能与现代农业等行业的数据治理实践中展现出显著应用价值。通过对多源异构数据的标准化处理、动态质量指标提取及风险预警模型构建，有效提升了储能系统运行监测、农业生产过程管理的数据一致性、准确性与实时性，为企业运营决策和精细化管理提供了可靠的数据支撑。在新能源电池储能场景中，该体系有助于提升运

行状态感知与风险预警能力；在现代农业应用中，则可支撑环境监测、生产调度与质量评估等关键环节的数据治理需求。但现阶段研究仍存在一定局限。一方面，复杂工况和极端环境下的数据采集稳定性不足，影响指标数据的完整性；另一方面，不同业务系统之间的数据标准与接口差异较大，跨系统协同治理仍有提升

空间。未来，可结合数字孪生与智能分析技术，构建面向新能源与现代农业的行业级数据治理与决策支撑体系，实现指标数据的实时映射、预测分析与智能优化，进一步提升企业数据治理能力和行业应用水平。

参考文献

[1] 王佳兰. 制约大数据在建筑领域中应用的关键影响因素研究及对策探讨 [D]. 重庆交通大学, 2023.

[2] 康雨萌. 大数据技术在汉语国际教育领域的应用研究 [D]. 西安石油大学, 2022.

[3] 轩振原. 基于大数据的动车组数据处理及故障诊断研究 [D]. 中国铁道科学研究院, 2022.

[4] 季书涵. 基于大数据技术的混凝土桩基质量评估系统研究 [D]. 东南大学, 2022.

[5] 马林. 大数据时代电网企业多维指标预算管理方法研究 [D]. 山东财经大学, 2021.

[6] 赵丹. 大数据技术在审计领域的应用研究 [J]. 现代商业, 2021, (26): 163-165.

[7] 杨晓倩, 周和益. 大数据技术在企业审计领域的应用及对策研究 [J]. 现代商业, 2022, (15): 181-183.

[8] 吴玉芳. 大数据技术在企业审计领域的应用及优化建议 [J]. 营销界, 2022, (21): 90-92.

[9] 张缪亮, 陆挺. 大数据技术在通讯领域的运用研究 [J]. 中国设备工程, 2021, (08): 25-26.

[10] 韩钰婷, 陶冶. 大数据技术在教育领域的应用研究 [J]. 电大理工, 2024, (01): 30-33.