变电设备数据质量评价模型研究

王峰

（中国电力科学研究院有限公司）

摘 要：变电设备数据质量和数据共享利用水平不高，没有形成一个完整性的数据质量管控体系和全面有效的数据质量保障机制，制约数据价值的深度挖掘。为此针对变电设备相关信息系统数据质量现状,构建涵盖数据接入情况、准确性、完备性、一致性和及时性等多维度目标的变电设备数据质量评价体系，提出适用于变电设备智能运维管控需求的数据质量评价指标计算方法，建立融合主客观权重的数据质量综合评价模型，解决目前方法评价维度偏少以及自动化和实用化程度不高的问题，使电力公司设备运维人员能够及时了解变电设备数据质量水平及存在问题，为变电设备监测预警、状态评价、故障诊断等高级应用提供必要的数据质量支撑。

基金项目：国家电网有限公司总部科技项目

作者简介：

引文信息

★王峰. 变电设备数据质量评价模型研究［J］．广东电力，2020，30(3)：17-25.

WANG Feng. Research on Data Quality Evaluation Model of Substation Equipment［J］．Guangdong Electric Power，2020，30(3)：17-25.

研 究 背 景

特高压及智能电网建设的持续推进对电网的安全稳定运行提出严峻挑战。一方面，电网线路路径长、覆盖广，沿途所经地域自然环境恶劣，极端天气多发，线路安全运行风险较大；另一方面，电网交直流混联，结构复杂，运行控制难度较大。

存 在 问 题

电网变电设备数据具有量大、面广、多源异构的特点，包括出厂试验报告、监造报告、设备台账、试验报告、状态监测、带电检测、气象、雷电等设备、电网、环境数据，为设备状态管控奠定了坚实的数据基础。但由于录入人员失误、系统设计缺陷、通信异常、装置故障、融合过程导致的数据不一致等原因，使得数据存在错误、缺失、重复等质量问题，迫切需要建立一个科学合理的数据质量管理体系，促进变电设备数据质量的提升和有效应用，为设备状态评价和故障诊断提供准确、全面的数据基础，有效支撑电网安全稳定运行。

计算机系统处理的数据和信息量在全球范围内不断增长，数据质量是从这些数据导出信息的质量和有用性的关键要素，大多数业务过程都依赖于数据质量。数据质量问题的研究，分别在统计领域始于20世纪60年代末，管理领域始于20世纪80年代初，计算机领域始于20世纪90年代初。国内电力信息化进程较晚，针对电力数据质量管理研究大约起始于21世纪初，早期侧重于电网运行数据异常值的辨识与处理，以及数据抽取过程中简单的数据清洗规则建立，主要采用插值、拟合等数值分析手段对异常数据进行平滑处理。2006年以来，国家电网有限公司（以下简称“国网公司”）高度重视信息化系统建设，相继提出SG186、SG-ERP等建设规划，极大提高了国网公司的信息化程度。同时电力数据质量问题逐渐引起相关部门关注，电力数据质量研究范围逐步扩展到营销、配网、售电、计量等领域，研究方法也从简单的数值分析延伸至数据挖掘、统计分析等计算机科学的方法。总体来说，国内电力数据质量管理研究尚处于起步阶段，侧重于对异常值的检测与处理，缺乏一套完善的理论体系与实践方法。

数据质量评价是数据质量管理的重要组成部分，承担着发现数据问题的重任，是提升数据质量的驱动力和风向标。目前数据质量评价多集中于电网安全可靠、电能质量、配网等领域。文献[5]讨论电力企业运营监测中心数据质量，评价指标不够全面，组合评价模型偏于简单。文献[6]设计了电网调度数据质量的综合分析与评价系统，针对每类数据给出了评价指标，通过主客观权重配比确定其综合评分；但评价指标及量化方法针对性较强，难以用于其他类型数据质量评价。文献[7]构建智能配电网数据质量评价体系，通过准确度、完整度、一致度、及时度和冗余度5个维度对智能配电网数据进行分析和评价；但指标计算方法过于复杂，可操作性不强。综上所述，目前缺乏针对设备运检领域的数据质量特性研究和应用，没有对这些质量特性的等级或评分进行取值范围的定义。

本文研究内容

本文基于中国电力科学研究院拥有的国网公司每年设备评价统计数据和与之相关的重大缺陷或故障历史数据等丰富的变电设备海量数据资源，以及常年的特高压设备状态监控、变电设备状态评价、事故分析等工作经验，从变电设备运检业务需求出发，借鉴数据与软件工程领域中数据质量模型和数据质量测量的最新标准GB/T 25000.12《系统与软件工程 系统与软件质量要求和评价（SQuaRE）第12部分：数据质量模型》和GB/T 25000.24《系统与软件工程 系统与软件质量要求和评价（SQuaRE）第24部分：数据质量测量》，从固有的和依赖系统的2种视点，给出了变电设备数据的质量特性，构建了涵盖数据接入情况、准确性、完备性、一致性和及时性等多维度目标的变电设备数据质量评价体系，提出了实用、有效的多维度评价指标及量化方法。同时着重研究了指标权重的计算方法，建立了融合主客观权重的数据质量综合评价模型，使电力企业运检人员能够全面、系统地掌握变电设备数据质量的真实水平和存在的问题，并及时进行改进和修复，有力支撑变电设备的智能运维管控。