

# 水库工程运行安全监测数据质量控制与异常识别方法研究

胡尔西代姆·库尔班

新疆维吾尔自治区塔里木河流域吉音水利枢纽管理中心（新疆吉音水电开发有限责任公司），中国·新疆 和田 848000

**摘要：**水库大坝作为防洪、灌溉及供水的关键水利基础设施，其运行安全直接关系到流域经济社会的稳定与发展。在新疆维吾尔自治区塔里木河流域，受极端干旱气候、强风沙环境、显著温差及复杂地质条件影响，水库安全监测数据的准确性与可靠性面临严峻挑战。本文系统综述了水库工程运行安全监测数据质量控制与异常识别的理论与方法，并紧密结合塔里木河流域大型平原水库及山区控制性工程的实际特点进行深入分析。文章首先剖析了塔里木河流域监测数据的主要误差来源，包括传感器漂移、传输干扰及环境耦合效应；其次，构建了涵盖数据采集、传输、存储全过程的质量控制体系；再次，重点探讨了基于统计学、机器学习及深度学习的多源异构数据异常识别算法在该区域的适用性与优化策略；最后，提出了适应南疆特殊环境的“空-天-地”一体化智能监测与预警框架。研究表明，通过建立针对性的数据清洗机制与智能化异常诊断模型，可显著提升塔里木河流域水库安全监测的时效性与精准度，为流域水资源优化配置与大坝长治久安提供坚实的技术支撑。

**关键词：**水库安全；监测数据；质量控制；异常识别；塔里木河流域；机器学习；多源融合

## Research on Quality Control and Anomaly Identification Methods for Reservoir Operation Safety Monitoring Data

Hulishidaimu Kurban

Jiyin Water Conservancy Hub Management Center of Tarim River Basin, Xinjiang Uygur Autonomous Region (Xinjiang Jiyin Hydropower Development Co., Ltd.), China Xinjiang Hotan 848000

**Abstract:** Reservoir dams, as key water conservancy infrastructure for flood control, irrigation, and water supply, directly affect the stability and development of the regional economy and society. In the Tarim River Basin of Xinjiang Uygur Autonomous Region, the accuracy and reliability of reservoir safety monitoring data are severely challenged by extreme drought climate, strong wind and sand environment, significant temperature differences, and complex geological conditions. This paper systematically reviews the theories and methods of quality control and anomaly identification for reservoir operation safety monitoring data, and conducts in-depth analysis in combination with the actual characteristics of large plain reservoirs and mountain control projects in the Tarim River Basin. The article first analyzes the main sources of errors in monitoring data in the Tarim River Basin, including sensor drift, transmission interference, and environmental coupling effects; secondly, it constructs a quality control system covering the entire process of data collection, transmission, and storage; thirdly, it focuses on discussing the applicability and optimization strategies of multi-source heterogeneous data anomaly identification algorithms based on statistics, machine learning, and deep learning in this region; finally, it proposes an "air - space - ground" integrated intelligent monitoring and early warning framework suitable for the special environment of southern Xinjiang. The research shows that by establishing a targeted data cleaning mechanism and intelligent anomaly diagnosis model, the timeliness and accuracy of reservoir safety monitoring in the Tarim River Basin can be significantly improved, providing solid technical support for the optimal allocation of water resources and the long-term stability of dams in the basin.

**Keywords:** Reservoir safety; Monitoring data; Quality control; Anomaly identification; Tarim River Basin; Machine learning; Multi-source fusion

## 0 引言

水是新疆经济社会发展的命脉，而水库则是调控水资

源时空分布、保障绿洲生态安全的核心枢纽。塔里木河流域作为中国最大的内陆河流域，横跨南疆五地州，流域内

分布着大石峡、阿尔塔什等控制性骨干工程以及众多服务于农业灌溉的大型平原水库。这些工程大多地处沙漠边缘或高山峡谷，自然环境极其恶劣，夏季高温酷暑、冬季严寒漫长，且常年伴随强风沙天气。这种特殊的地理气候背景，使得水库安全监测系统的硬件设备易受损、信号传输易受扰、观测数据易产生非结构性突变，给大坝安全评估带来了巨大困难。

传统的水库安全监测主要依赖人工巡视与定点仪器观测，数据量少且滞后，难以捕捉突发性险情。随着物联网、大数据及人工智能技术的飞速发展，自动化监测系统已广泛应用于塔里木河流域各大中型水库。然而，海量监测数据中混杂着大量噪声、缺失值及异常点，若不加甄别直接用于安全评估，极易导致误报或漏报，甚至引发错误的决策指令。特别是在塔里木河汛期（融雪型洪水与暴雨型洪水叠加期），数据的实时性与准确性更是关乎下游千万群众的生命财产安全。因此，针对塔里木河流域特殊环境，开展水库运行安全监测数据的质量控制与高效异常识别方法研究，不仅是水利工程管理领域的学术热点，更是保障国家西部水安全的迫切需求。本文旨在梳理相关技术进展，结合流域实际，构建一套科学、robust（鲁棒）的数据处理与分析范式。

## 1 塔里木河流域水库监测数据特征与误差成因分析

### 1.1 复杂环境下监测数据的时空分布特征

塔里木河流域水库监测数据具有鲜明的地域性与季节性特征。从时间维度看，受气温日较差大（可达 20℃以上）的影响，大坝变形、渗流等物理量呈现出显著的周期性波动，且这种波动往往与非线性的温度效应强耦合。例如，平原水库沥青混凝土防渗面板在夏季高温下软化，冬季低温下脆化，其位移监测数据随温度变化的滞后效应明显。从空间维度看，流域内水库类型多样，山区水库（如阿尔塔什）坝高水深，应力应变场复杂；平原水库（如小海子水库）库区广阔，风浪冲刷严重，岸坡稳定性监测点位分散。此外，由于通信基站覆盖不均，部分偏远库区数据传输存在间歇性中断，导致时间序列数据出现不规则缺失。

### 1.2 主要误差来源与数据质量问题诊断

在塔里木河流域，监测数据质量主要受到三类误差的制约。首先是仪器自身误差，长期暴露在强紫外线、高盐碱及风沙环境中，传感器容易发生老化、零点漂移或灵敏度下降，特别是振弦式仪器在极端温差下频率读数易失稳。

其次是传输与环境干扰误差，南疆频发的沙尘暴会导致无线通信信号衰减甚至中断，造成数据包丢失或乱码；同时，强电磁环境（如附近高压输电线路）可能引入脉冲噪声。第三是人为与管理误差，包括校准不及时、安装不规范及运维不到位等。具体表现为数据中的离群点（Outliers）、趋势性漂移、阶梯状突变以及长时段的数据缺失。这些问题若不经严格的质量控制与清洗，将严重扭曲大坝性态的真实反映，使得后续的异常识别模型失效。

## 2 监测数据全流程质量控制体系构建

### 2.1 源头采集端的标准化与冗余设计

提升数据质量的根本在于源头控制。针对塔里木河流域环境特点，应在采集端实施严格的标准化作业与冗余设计。一方面，选用耐候性强、防护等级达到 IP68 以上的工业级传感器，并对关键测点（如坝体内部位移、渗压计）采用“一主一备”或“双套比对”的安装方式，通过硬件冗余剔除单点故障数据。另一方面，制定适应当地气候的巡检与校准规范，特别是在春秋换季及沙尘暴过后，必须增加人工比测频次，利用便携式高精度仪器对自动化数据进行现场校核，及时修正系统偏差。此外，推广智能传感技术，在传感器端嵌入初步的逻辑判断算法，自动过滤明显的物理不可能值（如水位超过坝顶、负向渗压等），实现数据的初级清洗。

### 2.2 传输与存储过程的完整性校验机制

在数据传输环节，需建立多重校验与断点续传机制。针对无线信号不稳定的问题，采用“4G/5G+北斗卫星”双通道传输模式，确保在主网络中断时能通过北斗短报文发送关键报警数据。在协议层，引入 CRC 循环冗余校验与时间戳同步技术，防止数据包篡改与时序错乱。对于接收到的数据，服务器端应部署实时预处理模块，执行格式检查、范围限制检查及变化率检查。一旦发现数据异常（如数值超量程、跳变幅度过大），立即触发重传请求或标记为“可疑”，并通知运维人员介入。在存储层面，构建分布式数据库，实行异地容灾备份，确保历史数据的完整性和可追溯性，为长序列趋势分析奠定基础。

### 2.3 基于多源信息融合的数据清洗算法

针对已入库的含噪数据，需应用先进的数据清洗算法进行深度处理。结合塔里木河水库的实际，可采用“物理模型+数据统计”相结合的混合清洗策略。首先，利用水力学原理和大坝有限元仿真结果，设定各监测量的物理约束边界，剔除违背物理规律的粗差。其次，应用改进的拉依达准则（ $3\sigma$  原则）或小波变换去噪法，识别并平滑随

机噪声。对于缺失数据, 鉴于流域内气象因子(气温、降雨、蒸发)与大坝响应的高度相关性, 可构建基于长短期记忆网络(LSTM)的插补模型, 利用邻近测点数据及周边气象站数据作为输入, 高精度重构缺失的时间序列, 保证数据的连续性。

### 3 基于人工智能的异常识别方法与模型优化

#### 3.1 传统统计方法与阈值判定的局限性

传统的大坝安全监测异常识别多依赖统计过程控制(SPC)图、回归分析及固定阈值法。这些方法在处理线性、平稳序列时效果尚可, 但在塔里木河流域复杂的非线性环境下显得力不从心。例如, 固定的位移报警阈值无法适应季节性温度变化引起的正常胀缩, 容易导致频繁的误报; 而简单的线性回归模型难以捕捉强震后或高水位运行下大坝性态的突变特征。此外, 传统方法对多维变量间的耦合关系(如水位-温度-时效的综合作用)解析能力不足, 难以发现隐蔽的早期异常征兆。

#### 3.2 机器学习与深度学习驱动的异常检测架构

为突破传统方法的瓶颈, 引入机器学习与深度学习技术成为必然选择。针对塔里木河流域水库数据特点, 可构建以下几类模型: 一是基于无监督学习的异常检测。利用孤立森林(Isolation Forest)或局部离群因子(LOF)算法, 无需预先标注故障样本, 即可在高维监测数据空间中快速定位偏离正常模式的离群点, 适用于未知故障类型的早期发现。二是基于时序预测的残差分析。构建基于卷积神经网络(CNN)与LSTM组合的深度学习模型, 学习历史监测数据在水位、温度等环境量驱动下的正常演变规律。将模型预测值与实测值进行比较, 若残差序列超过动态置信区间, 则判定为异常。该方法能有效剥离环境因素影响, 提取出反映大坝结构性损伤的真实异常分量。三是基于知识图谱的因果推理。整合大坝设计资料、地质勘察报告、历史维修记录及专家经验, 构建流域水库安全知识图谱。当监测数据出现异常时, 利用图谱推理引擎关联分析可能的致因(如防渗体破裂、基础渗漏、地震损伤等), 提供可解释的诊断结论, 辅助决策者快速制定处置方案。

#### 3.3 面向小样本与不平衡数据的模型增强策略

在实际工程中, 大坝事故样本极少, 导致训练数据呈现极度的类别不平衡。为提高模型在塔里木河流域的泛化能力, 需采取特定的增强策略。一方面, 利用生成对抗网络(GAN)合成逼真的故障数据样本, 扩充训练集, 平衡正负样本比例。另一方面, 引入迁移学习技术, 将在其他成熟流域或实验室模拟实验中训练好的模型参数迁移至塔

里木河项目, 通过少量本地数据进行微调(Fine-tuning), 解决冷启动问题。此外, 建立“人机回环”机制, 将专家对模型报警结果的反馈纳入再训练过程, 持续迭代优化模型精度, 使其越来越适应当地的具体工况。

### 4 塔里木河流域智能监测预警平台的集成与应用展望

#### 4.1 “空-天-地”一体化监测网络构建

未来的核心发展方向是构建集卫星遥感、无人机低空巡检与地面自动化传感于一体的“空-天-地”立体化协同监测网络, 以突破塔里木河流域地域辽阔、环境恶劣带来的监管盲区。在“天”基层面, 深度融合合成孔径雷达干涉测量(InSAR)技术, 利用其全天候、大范围的优势, 对流域内大型平原水库库岸潜在滑坡体、坝体及附属设施的大范围沉降进行毫米级广域普查, 有效弥补地面测点稀疏的缺陷, 实现宏观形变趋势的早期捕捉。在“空”基层面, 常态化部署搭载激光雷达(LiDAR)、高分辨率可见光及红外热成像载荷的工业级无人机集群, 定期对大坝表面裂缝扩展、渗流逸出点温度异常及护坡冲刷情况进行精细化三维建模与自动巡查, 解决人工巡检难以抵达的死角问题。

#### 4.2 数字孪生驱动的预演预案与智慧决策

依托清洗后的高质量数据底座与先进的深度学习异常识别模型, 全面构建塔里木河流域重点控制性水库的高保真数字孪生体。该系统不仅在虚拟空间实时映射大坝的物理几何状态与力学行为, 更内嵌水动力学、结构动力学及地质灾害演化机理模型, 能够高仿真模拟极端融雪洪水、强降雨叠加、高烈度地震等复杂工况下的灾变演进过程。一旦实时监测数据触发多级预警阈值, 系统立即启动应急响应机制, 在数字孪生体中自动复现故障场景, 快速推演溃坝洪水演进路径、淹没范围及受灾人口分布, 量化评估灾害损失。基于推演结果, 智能算法将结合实时雨情、水情及下游社会经济数据, 自动生成包含泄洪调度方案、人员转移路线及物资调配策略在内的最优应急处置预案, 并支持多方案比选优化。

### 5 结语

水库工程运行安全监测数据的质量控制与异常识别, 是保障塔里木河流域水利设施安全运行的基石。面对南疆特殊的自然环境与复杂的工程条件, 必须摒弃单一的技术手段, 转而构建集标准化采集、智能化清洗、多维度分析及可视化预警于一体的综合技术体系。通过深度融合物理学机制与人工智能算法, 不仅能有效剔除环境噪声干扰,

精准捕捉大坝性态的微小异常,更能挖掘数据背后的演化规律,实现安全隐患的早发现、早诊断、早处置。未来,随着5G、北斗导航及量子传感等前沿技术的进一步应用,塔里木河流域的水库安全监测将更加智慧化、精准化,为新疆社会稳定、长治久安及经济社会高质量发展筑牢坚实的水安全屏障。

#### 参考文献:

[1] 辛秀钰. 考虑多水文要素的峡江水利枢纽近坝水位预测方法研究[D]. 南昌大学, 2025. DOI:10.27232/d.cnki.gnchu.2025.002670.

[2] 徐贤, 孙中晋, 徐洪庆. 棘洪滩水库运行管理矩阵实践与创新[J]. 中国水利, 2025,(09):50-57.

[3] 肖仲凯, 宋世柱, 王吉等. 数字孪生凌塘水库建设

关键技术研究与实践[J]. 水利水电快报, 2025,46(05):123-131. DOI:10.15974/j.cnki.slsdkb.2025.05.020.

[4] 张明. 下坂地水库监测预警与防洪调度方案分析[J]. 水利发展研究, 2024,24(08):38-43. DOI:10.13928/j.cnki.wrd.2024.08.007.

[5] 王保良. 黄河干流生态径流演变及其驱动因素定量评价[D]. 华北水利水电大学, 2024. DOI:10.27144/d.cnki.ghbsc.2024.000232.

[6] 詹树维. 解读大坝安全监测自动化系统应用现状分析及发展趋势[J]. 水上安全, 2023,(09):46-48.

作者简介: 胡尔西代姆·库尔班(1991.10-), 女, 维吾尔族, 新疆阿图什市, 学历: 大学本科。