

水利水电工程中大坝渗流监测数据异常值识别与修正方法

李汉伟

宁夏太阳山水务有限责任公司

DOI:10.12238/jpm.v6i10.8469

[摘要] 大坝渗流监测数据的准确性对水利水电工程的安全管理至关重要。然而，监测过程中常常会遇到数据异常值，这些异常值会影响对大坝渗流状态的评估与预警，甚至可能导致工程决策错误。识别并修正这些异常值成为确保数据准确性的关键任务。本文提出了一种基于多维度数据分析和算法修正的渗流监测数据异常值识别方法，并结合实际案例对修正效果进行了验证。通过合理的异常值识别与修正方法，能够有效提高监测数据的准确性，为大坝安全运行提供可靠的数据支持。

[关键词] 大坝渗流；数据异常值；识别与修正；水利水电工程；数据分析

Identification and correction methods for outliers in dam seepage monitoring data in water conservancy and hydropower engineering

Li Hanwei

Ningxia Taiyangshan Water Co., Ltd.

[Abstract] The accuracy of dam seepage monitoring data is crucial for the safety management of water conservancy and hydropower projects. However, data outliers are often encountered during the monitoring process, which can affect the assessment and warning of dam seepage status, and may even lead to errors in engineering decision-making. Identifying and correcting these outliers has become a key task in ensuring data accuracy. This article proposes a method for identifying outliers in seepage monitoring data based on multidimensional data analysis and algorithm correction, and verifies the correction effect through practical cases. By using reasonable methods for identifying and correcting outliers, the accuracy of monitoring data can be effectively improved, providing reliable data support for the safe operation of dams.

[Key words] dam seepage; Data Outliers; Identification and correction; Water conservancy and hydropower engineering; data analysis

引言

水利水电工程的大坝渗流监测是保障工程安全的重要环节。随着技术的发展，大坝渗流监测系统已经能够提供大量实时数据，但在数据采集、传输和处理过程中，往往会产生一些异常值。这些异常值如果未能及时识别和修正，可能导致对大坝渗流状态的错误判断，从而影响工程的安全性评估和决策执行。当前，异常值的识别与修正方法虽然已有一定研究，但在实际应用中仍存在诸多挑战。特别是对于大坝渗流这一复杂的多变量监测体系，如何有效地识别和修正数据异常，保证监测系统的稳定性和数据的可靠性，成为亟待解决的重要问题。

一、大坝渗流监测数据异常值问题的识别与定义

（一）渗流监测数据的特点分析

大坝渗流监测数据具有复杂性和多样性。其数据采集过程中，受到气候变化、地质结构、施工环境等多方面因素的影响，往往呈现出一定的时空变异性。渗流数据通常是通过多个监测点在不同时间进行实时采集，因此数据之间存在时间和空间的相关性。这种相关性使得渗流数据不仅具有局部的变化，还受到外部环境因素的干扰，导致一些数据在特定时段或条件下出现波动。监测设备的精度、位置误差以及数据传输的稳定性也可能影响渗流数据的准确性。在数据分析时需要特别注意这些特点，以避免由于数据采集过程中产生的误差导致错误的结果。

（二）异常值的识别标准与分类

渗流监测数据中的异常值，指的是那些在特定条件下显著偏离正常数据范围的值。这些异常值可能来源于仪器故障、数据传输错误、环境变化或人为操作失误。根据异常值的来源和表现，可以将其分为两类：一类是基于统计学标准识别的异常值，通常表现为在数据集中的极端值；另一类是基于工程实际情况判断的异常值，这类异常值可能没有明显的统计学偏差，但由于不符合工程常规规律或现场条件，依然被视为异常。常见的异常值包括由于设备故障引起的极大或极小读数，以及受外部环境（如暴雨、地震等）影响的突发异常数据。对异常值的分类与识别，需要根据工程背景和历史数据进行综合判断，从而确保修正工作的针对性和有效性。

（三）常见异常值的来源与影响

渗流监测数据的异常值来源可以归结为以下几类：一是仪器故障或传感器失效，导致数据出现极大或极小值，这类异常值常见于设备老化或安装不当的情况下；二是数据传输过程中发生错误，可能由于网络不稳定或传输协议的问题，造成数据丢失或不一致；三是外部环境变化，如暴雨、地震等自然灾害，可能导致瞬时的大幅度变化，从而出现离群数据；四是人为操作错误，例如数据录入时的失误或测量点选择不当，也可能引发数据异常。

二、数据异常值的分析与检测方法

（一）统计方法在异常值检测中的应用

统计方法在大坝渗流数据的异常值检测中具有重要作用。常见的统计方法包括 Z-Score、箱线图法和标准差法等。这些方法基于数据的分布特征，通过分析数据与其均值或中位数的偏差，来判断是否存在异常值。Z-Score 方法通过计算数据的标准化值（即与均值的偏差与标准差的比值），若 Z-Score 值超出预设阈值（如 ± 3 ），则认为数据是异常的。箱线图法则通过数据的四分位数来定义异常值，任何超出“盒子”范围的数据都被视为异常。标准差法同样基于数据的离散程度，若数据点与均值的差距大于设定的标准差倍数，则视为异常。

（二）机器学习算法在数据异常识别中的优势

机器学习算法能够处理复杂、多维的渗流监测数据，识别其中隐藏的异常模式。常见的机器学习方法包括支持向量机（SVM）、决策树、聚类算法等，这些方法能通过训练模型自动识别数据中的异常特征。支持向量机可以通过构建一个最优超平面，将正常数据和异常数据分开；决策树则通过分支判断，将数据划分为不同的类别，异常值通常被划分到偏离大多数的分支上。聚类算法，如 K-means，可以将数据分为多个组，并通过分析每个组的内部一致性来发现异常数据。机器学习方法不仅能处理大规模数据，而且能够捕捉到数据中的非线性关系，特别适用于渗流数据这种受多因素影响、变化复杂的情况。通过这些算法，可以更高效、更准确地识别

渗流监测中的异常值。

（三）多维度数据融合分析方法

渗流监测数据往往受到多种因素的影响，单一的数据源可能难以全面反映出问题。多维度数据融合分析方法则通过结合不同来源的数据，提供更全面的异常值识别结果。将渗流数据与气象数据、土壤湿度数据、地震数据等进行融合，可以综合考虑多种因素对渗流状态的影响，从而更准确地识别异常情况。数据融合不仅可以提高数据的准确性，还能帮助排除由单一传感器或单一环境因素引发的异常值。数据融合还能优化监测方案，提高系统的整体鲁棒性。在实际应用中，通过构建基于多维数据的分析模型，可以更好地掌握大坝渗流变化的全貌，增强对异常值的识别能力。

三、渗流监测数据异常值的修正技术

（一）基于统计学的修正方法

统计学方法在渗流监测数据修正中起着基础性作用。对于识别出的异常值，统计学方法常通过插值法、均值法和加权平均法等方式进行修正。插值法根据已有的正常数据点，通过数学模型推测出异常数据点的合理值。线性插值法或多项式插值法可以根据邻近数据点推算出异常值的修正值。均值法则简单直接，通过计算异常值前后数据的均值或中位数来替代异常值。加权平均法在考虑异常数据影响的基础上，对周围数据赋予不同权重，以减少异常值对修正结果的影响。这些方法通常用于处理那些相对简单且线性的数据异常，适用于渗流监测数据中的一些基础性修正需求。

（二）基于回归模型的修正方法

回归模型是渗流数据修正中的另一种重要技术，特别是在处理数据呈现出一定趋势性或规律性的情况下。回归模型通过建立输入与输出变量之间的数学关系，能够预测并修正异常数据。利用线性回归或非线性回归模型，可以根据其他正常数据点的变化趋势，预测出异常数据应该具备的合理值。回归模型不仅能够处理单一变量的数据，还能处理多变量数据之间的关系，通过多元回归模型进一步提高数据修正的精确度。在大坝渗流监测中，这种方法能够结合多个监测点的数据信息，从而进行更加准确的修正，避免单一数据点异常对结果的影响。

（三）基于人工智能的修正策略

人工智能（AI）技术在数据修正中逐渐展现出强大的优势，尤其在处理复杂、非线性的大坝渗流监测数据时，AI 算法能提供高效、精确的修正策略。常见的 AI 算法包括神经网络、深度学习和强化学习等，这些算法能够从大量历史数据中自动学习规律，并对新的异常数据进行修正。神经网络通过模拟人脑的结构和功能，能够对数据中的潜在非线性关系进行建模，从而预测并修正异常值。深度学习则能够通过多层次的神经网络，更加精准地捕捉复杂数据中的模式。强化学习则可以在修正过程中进行动态优化，逐步提高修正精度。AI 方法在大坝渗

流监测中的应用，能够结合传感器数据、环境数据等多维度信息，提出个性化的修正策略，提升修正结果的准确性和可靠性。

四、渗流监测数据修正效果评估与应用

（一）修正效果的评价指标

修正效果的评价是衡量渗流监测数据修正技术有效性的关键。常用的修正效果评价指标包括均方误差（MSE）、平均绝对误差（MAE）和相关系数等。均方误差用于量化修正前后数据之间的偏差，数值越小表明修正效果越优；平均绝对误差则关注修正后的数据与实际值之间的平均误差，数值较低表明修正效果较好；相关系数用于评估修正数据与实际渗流状态的一致性，数值越高表示修正后的数据与实际情况越接近。这些评价指标能清晰地展示修正方法的实际表现，为后续技术优化提供指导。在评价时，还需综合考虑修正方法的实时性、计算复杂度和适用场景等因素。

（二）实际案例中修正效果的验证

通过实际案例对修正效果进行验证是评估修正方法有效性的一个重要环节。在某水利水电工程的大坝渗流监测中，采用了统计方法、机器学习算法以及人工智能修正策略对异常值进行修正。通过与历史正常数据的对比分析，修正后的数据显示出较高的准确性和一致性。在这一案例中，修正后的渗流数据较大程度上恢复了正常的趋势，并有效地反映了大坝渗流情况，避免了错误的安全评估。这一验证结果表明，结合多种修正方法，可以有效提高渗流监测数据的准确性，进而为大坝的安全运行提供更加可靠的数据支持。

（三）修正技术在大坝监测中的应用实例

在实际应用中，渗流监测数据修正技术已经成功地应用于多个大坝监测项目。在某大型水电大坝的渗流监测中，系统通过结合统计修正和机器学习修正策略，有效识别并修正了由设备故障和环境变化引发的异常数据。修正后的数据成功地反映了大坝的渗流趋势，有效提高了渗流监测系统的可靠性。这一应用实例证明了数据修正技术在大坝监测中的重要作用，也为后续相关项目提供了宝贵的经验。

五、大坝渗流监测系统的优化与未来发展

（一）异常值修正技术的持续优化

异常值修正技术的持续优化是适应大坝渗流监测需求日益复杂化的必然趋势。传统的统计方法，如均值法和标准差法，虽然简便易行，但无法有效处理复杂的非线性关系和大规模数据。近年来，机器学习和人工智能技术逐渐取代了这些传统方法，使得数据修正过程更加智能化和高效。采用深度学习、神经网络等先进算法，能够通过自学习和自优化机制，识别和修正数据中的异常值。利用实时数据分析平台，这些技术能够在数据采集的同时进行自动校正，保证数据的高精度和高可靠性。随着计算能力和数据处理技术的飞跃，未来的渗流监测系

统将实现更大规模的自动化修正，显著提升数据分析的准确性和响应速度。

（二）智能化监测系统的未来应用

未来的大坝渗流监测将朝着全自动化和智能化的方向发展。集成物联网技术、人工智能算法和大数据分析工具，监测系统能够实时自动获取数据，并对数据进行智能处理和分析，及时识别异常值并进行修正。通过智能化的监测系统，大坝的渗流状态将被持续、实时地监控，并通过数据模型预测可能发生的安全风险。这种智能化系统不仅能提高数据准确性，还能减少人工干预，降低人为错误带来的风险。随着云计算和边缘计算的结合，数据处理将更加灵活和高效，为大坝监测提供更为精确的支持，确保工程安全稳定运行。

（三）多系统融合与大坝监测的长远展望

多系统融合将成为未来渗流监测技术发展的核心。未来的大坝监测系统将不再仅仅依赖渗流数据，而是将气象数据、土壤湿度数据、地震数据等多维度数据进行有效融合。通过整合来自不同传感器的数据，系统可以从多个角度对大坝的渗流状态进行全面分析。这种多维度的融合能更好地识别潜在的风险，及时调整监测策略，确保大坝安全。5G技术的应用将大大提升数据的实时传输和处理能力，进一步缩短反应时间，并增强监测系统的响应速度和灵活性，为大坝的长期安全运行提供更加可靠的数据支持。

结语

本文详细探讨了大坝渗流监测数据异常值识别与修正方法的研究进展。随着监测技术的不断发展，传统的异常值处理方法逐渐向更加智能化的方向转变，机器学习、人工智能等技术的应用为修正方法提供了更高的精度与可靠性。智能化监测系统的普及以及多系统数据融合的不断推进，必将大大提高大坝渗流监测的效率和准确性，进一步保障大坝的安全运行。在未来的工作中，结合多源数据和智能算法的修正技术将成为渗流监测领域的重要发展方向。

[参考文献]

- [1]王晓梅, 赵志强.基于机器学习的大坝渗流监测数据异常值修正研究[J].水利科技, 2023, 45(6): 128-135.
- [2]张娜, 高明.大坝渗流监测系统的数据修正技术及应用探讨[J].水利工程学报, 2024, 38(3): 92-98.
- [3]李鹏飞, 吴强.大坝渗流数据异常值修正算法的研究与应用[J].岩土工程技术, 2023, 42(2): 210-217.
- [4]朱国强, 王建波.基于深度学习的渗流监测数据异常值修正方法[J].水电工程技术, 2023, 39(1): 76-82.
- [5]刘瑾, 孙浩.多维数据融合技术在渗流监测中的应用与展望[J].水资源保护, 2024, 40(7): 168-173.