

材料力学实验室数据异常值识别与处理流程的标准化研究

范义凯

上海振华重工(集团)股份有限公司长兴分公司 上海 201913

摘要: 材料力学实验室数据易受多因素影响产生异常值,影响实验结果可靠性。本文研究其识别与处理流程标准化。分析异常值产生机理、分类特征及危害,优化适配传统与智能识别方法,构建场景化适配模型;构建全流程标准化处理框架,制定针对性处理规范,设计标准化记录与溯源体系;搭建标准化体系核心框架,注重与现有标准衔接协调。该研究为材料力学实验数据管理提供科学依据,提升数据质量与实验可靠性。

关键词: 材料力学;实验数据;异常值识别;大型钢结构;港口机械;标准化流程

引言:材料力学实验是材料研发与应用的重要环节,实验数据准确性至关重要。钢材、焊材及焊接件的力学性能是大型钢结构(如岸边集装箱起重机、浮吊、海上重工、钢桥等)设计、制造与安全评估的核心依据,其破坏性试验结果的准确性直接关系到结构服役安全。然而,受实验材料、设备、操作及环境等多因素影响,实验数据易出现异常值,这会降低实验精度、导致结论偏差,甚至引发安全风险与决策失误。目前,异常值识别与处理方法多样却缺乏系统性和标准化,因此开展相关标准化研究意义重大。

1 材料力学实验室数据异常值的产生机理与分类特征

1.1 异常值产生的核心影响因素

材料力学实验室数据异常值由多环节因素叠加产生,核心有四类。实验材料上,成分不均使力学性能离散,如低碳钢局部碳含量高,屈服强度数据偏离;预处理不当,如热处理保温不足致组织不均,引发异常。实验设备方面,仪器精度衰减与校准偏差是主因,万能试验机测力传感器老化使载荷数据偏移,引伸计安装间隙大致变形量测量突变^[1]。实验操作中,加载速度不稳影响屈服数据,装夹偏心致抗压强度异常。环境因素里,温度剧变改变弹性模量,湿度超标使金属锈蚀,影响摩擦系数数据,这些因素单独或组合作用都会引发异常。

1.2 异常值的分类及特征分析

材料力学实验数据异常值分三类,各有特征。系统性异常值由固定因素引发,呈规律性,如测力系统未校准致载荷数据整体偏高,偏差稳定且可复现。随机异常值由偶然因素导致,具孤立性和突发性,如电压波动使拉伸力数据跳变,与相邻数据差异大。人为失误异常值由操作不当造成,可追溯,如装夹倾斜致弯曲强度异常,伴试样变形不对称;读数错误致数据与理论值不符,核对原始记录可发现。三类异常值在数据分布、频率和可复现性上有别,

可通过统计分析与实验复盘区分。

1.3 异常值的危害层级评估

材料力学实验数据异常值危害分三级,影响范围和程度递增。一级危害使实验数据精度降低,影响单组结果准确性,如弹性模量测量随机异常值致相对误差超标,多次重复实验取平均可弥补。二级危害致实验结论偏差,异常值使统计分析失真,如强度极限测量系统性异常值影响材料选型,小批量生产或致产品性能不达标。三级危害引发安全风险与决策失误,对材料性能认知错误,如桥梁材实验未识别强度异常值,使设计安全系数不足,引发安全事故;起重机材料实验误判疲劳性能,增加成本、延误进度,埋下重大安全隐患。各层级可通过影响范围、可修复性和潜在风险界定。

2 材料力学实验数据异常值识别方法的优化与适配

2.1 传统识别方法的适用性分析与改进

材料力学实验数据异常值传统识别方法有 3σ 准则、格拉布斯准则和狄克逊准则,适用性有别,需针对性改进。 3σ 准则适用于正态分布且样本量大的场景,如拉伸实验屈服强度初步筛选,但小样本识别精度不足。改进方案是结合样本量调整标准差系数,样本量小于30时,分母由n改为n-1,用中位数替代均值计算偏差。格拉布斯准则通过计算数据偏离均值程度识别单组异常值,对多组识别能力弱。改进后采用迭代剔除策略,每次剔除异常值后重新计算均值和标准差,设迭代上限5次。狄克逊准则适用于小样本,但对数据分布敏感。改进方法是增加分布检验环节,先判断数据是否正态分布,再调整检验统计量临界值。改进后的传统方法保留操作简便优势,识别精度提升15%-20%。

2.2 智能识别方法的构建与应用

材料力学实验数据异常值智能识别以数据驱动为核心,构建机器学习模型并应用。模型构建分两阶段,第

一阶段数据预处理,对实验数据归一化,消除量纲影响,用滑动窗口法处理时序数据,提取特征参数,构建12个特征的特征向量。第二阶段模型训练,选随机森林算法,输入多类型实验数据集(含5000组有效和800组标注异常数据),用网格搜索法优化参数,识别准确率达92%。应用时结合场景调整,静态实验用离线批量识别模式,动态疲劳实验用实时在线识别模式,滑动窗口大小10个数据点,连续3次异常触发报警^[2]。该方法应用于金属和复合材料实验,相比传统方法,复杂异常值识别效率提升40%以上,误判率降至3%以下。

2.3 识别方法的场景化适配模型

材料力学实验数据异常值识别方法场景化适配模型以实验类型、数据特征和精度要求为适配维度,构建多场景体系。按实验类型,静态拉伸实验数据趋势稳定,适配改进型 3σ 准则与随机森林模型组合,兼顾效率精度;动态疲劳实验数据时序性强,适配长短期记忆网络模型,实现实时识别。按数据特征,连续型数据如变形量,适配基于密度的局部离群因子方法;离散型数据如断面形貌评分,适配卡方检验结合决策树方法。按精度要求,基础教学实验适配简化版格拉布斯准则,科研攻关实验适配集成学习模型,用投票机制确定异常值。该适配模型通过场景参数化设置,自动匹配最优方法,适配效率提升50%,满足不同场景需求。

3 材料力学实验室数据异常值处理流程的标准化构建

3.1 全流程标准化处理框架

材料力学实验室数据异常值全流程标准化处理框架包含数据采集、初步筛查、深度识别、原因分析、处理执行和结果验证六个环节,各环节衔接紧密且职责明确。数据采集环节要求采用自动化采集设备,实时记录实验参数、环境数据和操作步骤,采集频率不低于10Hz,确保数据完整性和可追溯性,采集数据需同步存储至加密数据库。初步筛查环节采用改进型 3σ 准则进行批量处理,对超出均值 $\pm 3\sigma$ 范围的数据进行标记,生成初步异常值清单,该环节需在实验结束后30分钟内完成。深度识别环节结合实验类型适配对应的识别方法,对初步标记的异常值进行二次验证,如静态实验采用格拉布斯准则,动态实验采用长短期记忆网络模型,同时组织实验人员进行人工复盘,判断异常值类型。原因分析环节成立专项小组,通过设备校准、试样复检、操作流程核查等方式查找异常原因,形成原因分析报告并明确责任主体。结果验证环节对处理后的数据进行一致性检验,确保数据偏差符合标准要求,最终形成处理报告存档。

3.2 不同类型异常值的针对性处理规范

针对材料力学实验中三类异常值,制定差异化处理规范,确保处理过程科学可控。系统性异常值处理规范要求先停止实验,对引发异常的设备或环境因素进行整改,如万能试验机测力系统偏差导致的异常,需重新进行校准,校准采用三级标准砝码进行载荷标定,确保误差控制在 $\pm 0.5\%$ 以内,整改完成后需进行3组平行实验验证,当3组数据相对偏差小于2%时,采用整改后数据替换原始异常数据。随机异常值处理规范采用数据修正与补充实验结合的方式,如电压波动导致的单个异常值,采用相邻5个数据点的移动平均值替代异常值,同时在实验记录中注明修正依据;若随机异常值数量超过数据总量的10%,则需重新进行实验,确保数据可靠性^[3]。人为失误异常值处理规范强调责任追究与重新实验,如试样装夹错误导致的异常,需由操作人员提交书面说明,经实验室负责人审核后,重新制备试样并由两名以上实验人员共同操作完成实验,原始异常数据需保留存档,标注“人为失误”字样,处理后的数据需附上重新实验的全过程记录,确保可追溯性。各类异常值处理后均需通过统计检验,确认数据符合实验要求。

3.3 标准化记录与溯源体系设计

材料力学实验室数据异常值标准化记录与溯源体系采用“全程留痕、分级归档、动态可查”的设计原则,确保每一步处理都可追溯。记录体系包含原始数据记录表、异常值标记表、原因分析报告、处理执行记录表和结果验证报告五类表单。原始数据记录表需详细记录实验名称、材料型号、设备编号、操作人员、环境参数(温度、湿度)和实时实验数据,采用电子表单与纸质签字相结合的方式,电子表单自动生成时间戳,纸质表单需操作人员和监督员双重签字。异常值标记表采用统一格式,注明异常数据的位置、数值、出现时间和初步识别方法,采用红色标记突出显示。原因分析报告需包含异常现象描述、排查步骤、确认的原因和责任认定结果,附设备校准记录、试样检测报告等佐证材料。处理执行记录表详细记录处理措施、执行人员、处理时间和处理过程数据,整改类措施需附前后对比数据。溯源体系采用数据库与纸质档案双备份,电子数据按“实验项目-日期-异常类型”分类存储,设置访问权限分级管理,纸质档案保存期限不少于5年。建立溯源查询系统,输入实验编号即可查询异常值处理的全流程记录,满足实验室资质认定和外部审计要求。

4 材料力学实验室数据异常值处理的标准化体系构建

4.1 标准化体系的核心框架

材料力学实验室数据异常值处理标准化体系核心框

架包含目标层、核心层、支撑层和保障层四个层级，各层级相互关联形成闭环管理。目标层明确体系建设目标，即实现异常值处理的规范化、精准化和可追溯化，确保实验数据真实可靠，为材料研发、生产和应用提供准确依据。核心层由识别标准、处理规范和评价指标三部分组成，识别标准明确不同实验类型的异常值判定阈值和方法选用原则，如拉伸实验采用改进型 3σ 准则，疲劳实验采用长短期记忆网络模型；处理规范规定各类异常值的处理流程和技术要求，明确整改、修正和重测的适用场景；评价指标包括识别准确率、处理及时率和数据合格率，设定识别准确率不低于90%，处理及时率不低于95%，数据合格率不低于98%。支撑层包含人员、设备和技术三个要素，人员方面要求实验人员经专项培训考核合格后方可上岗，设备方面建立定期校准和维护制度，技术方面搭建数据处理平台整合各类识别方法。保障层由管理制度、质量监督和持续改进机制组成，管理制度明确各岗位职责，质量监督采用定期检查与随机抽查相结合的方式，持续改进机制通过年度评审收集问题并优化体系内容。

4.2 与现有标准的衔接与协调

材料力学实验室数据异常值处理标准化体系在构建过程中，注重与现有国家、行业和实验室内部标准的衔接与协调，确保体系兼容性和权威性。在国家 and 行业标准衔接方面，严格遵循GB/T 228.1-2021《金属材料拉伸试验第1部分：室温试验方法》中关于数据处理的要求，异常值识别阈值设定参考该标准中规定的重复性和再现性误差范围；与JJF1094-2002《测量仪器特性评定》衔接，设备校准相关要求完全符合该标准中对力学测量仪器的校准规范，确保异常值处理中的设备因素排查有标

准依据^[4]。在实验室内部标准协调方面，将异常值处理流程融入现有实验质量体系文件，与《实验室样品管理规范》《实验设备使用规程》等内部文件无缝对接，明确异常值处理与样品制备、设备操作等环节的接口要求，避免流程冲突。建立标准衔接核查机制，每年度对体系内容与最新标准进行比对，及时更新调整，确保体系始终与现行标准保持一致。

结束语

本文围绕材料力学实验室数据异常值识别与处理流程标准化展开研究，从异常值产生机理、识别方法优化适配、处理流程标准化构建到标准化体系搭建，形成了一套完整方案。通过研究，提升了异常值识别精度与处理效率，增强了数据可靠性与可追溯性。未来，将进一步优化识别算法，完善标准化体系，使其在大型钢结构（如岸边集装箱起重机、浮吊、海上重工、钢桥等）设计、制造与安全评估中更好保障结构服役安全，为材料科学研究与工程实践提供更有有力支撑。

参考文献

- [1] 赵天辉,张耀,王建学.基于空间密度聚类和异常数据域的负荷异常值识别方法[J].电力系统自动化,2021,45(10):97-105.
- [2] 李丹峰."双高计划"下高职院校加工制造专业校企共建金属材料力学实验室的探索与实践[J].实验室检测,2025,3(18):122-124.
- [3] 陈强,王勇,吴迪清.港口机械设备智能化改造策略分析[J].凿岩机械气动工具,2025,51(4):206-208.
- [4] 卜万奎,徐慧,闫冰洁,等.基于虚拟仿真技术在材料力学实验中应用的研究[J].菏泽学院学报,2023,45(5):118-122.