

水质检测中数据处理与误差分析

张少武 周志杨 杨洪兴 吴超越

浙江九安检测科技有限公司 浙江 杭州 310053

摘要: 本文围绕水质检测中的数据处理与误差分析展开研究。阐述了水质检测数据采集与预处理的关键步骤,包括传感器监测、日志记录、数据清洗与标准化等。深入探讨了误差的来源,涵盖系统误差、随机误差和过失误差,并详细分析了各类误差产生的原因。针对误差提出了相应的控制方法,如规范操作、仪器校准、空白实验等。同时,介绍了多种数据分析方法在水质检测中的应用,如描述性统计、相关性分析、时间序列分析等,并探讨了这些方法在提升水质检测准确性和科学性方面的作用。

关键词: 水质检测; 数据处理; 误差分析; 数据分析方法

引言

水是生命之源,水质安全直接关系到人类的健康和生态环境的稳定。随着工业化和城市化的快速发展,水污染问题日益严重,对水质进行准确检测和科学分析显得尤为重要。水质检测数据是评估水质状况、制定水质管理策略的重要依据,然而,在水质检测过程中,数据处理和误差分析是两个关键环节。数据处理的质量直接影响检测结果的准确性和可靠性,而误差分析则有助于识别和减少误差,提高检测精度。因此,深入研究水质检测中的数据处理与误差分析具有重要的现实意义。

1 水质检测数据采集与预处理

1.1 数据采集方式

水质检测数据采集作为整个水质监测工作的基础环节,其准确性和全面性对于后续的水质评估和管理决策至关重要。目前,水质检测数据采集主要依赖于多种在线监测设备和手动记录操作信息相结合的方式。在线监测设备中,传感器监测是最为常用的方式之一。pH计作为水质检测的基础设备,能够实时、精准地测量水体的酸碱度,其测量范围广泛,可满足不同水质环境下的检测需求。溶解氧(DO)传感器则专注于监测水体中溶解氧的含量,这对于评估水体的自净能力和生物活性具有重要意义。浊度仪通过测量水体中悬浮颗粒物的散射光强度,来反映水体的浊度,为判断水体的清澈程度提供了直观的数据支持。这些传感器具备连续、实时监测的能力,能够及时捕捉水质参数的细微变化,为水质变化的预警和快速响应提供了可能。除了传感器监测,手动或自动记录设备运行时间、药剂投加量、污泥产量等操作信息也是数据采集的重要组成部分。设备运行时间记录有助于了解设备的使用频率和负荷情况,为设备的维护和管理提供依据。药剂投加量记录能够反映水处理工

艺中化学药剂的使用情况,对于优化药剂投加方案、降低处理成本具有重要作用。污泥产量记录则有助于评估污水处理过程中污染物的去除效果和污泥处理系统的运行状况。这些操作信息与水质参数数据相互关联,共同构成了水质检测的完整数据体系,对于深入分析水质变化的原因和评估处理工艺的效果具有不可替代的意义^[1]。

1.2 数据预处理步骤

采集到的原始数据往往存在各种问题,如重复、错误或无效的数据点,这些问题会严重影响后续数据分析的准确性和可靠性。因此,数据预处理是水质检测数据处理过程中不可或缺的环节。数据清洗是数据预处理的首要步骤,其目的是去除原始数据中的噪声和异常值,提高数据的质量。去除重复值是数据清洗的基本操作之一,通过识别和删除重复的数据记录,可以避免数据冗余对分析结果的影响。填充缺失值则是针对数据中存在的缺失数据进行处理,常见的方法包括均值填充、中位数填充和基于模型的预测填充等。均值填充是将缺失值用该列数据的平均值进行替代,适用于数据分布较为均匀的情况;中位数填充则是用中位数替代缺失值,对于存在极端值的数据更为稳健;基于模型的预测填充则是利用回归模型、时间序列模型等对缺失值进行预测填充,能够更好地保留数据的内在规律。异常值检测和处理也是数据清洗的重要内容。异常值可能是由于传感器故障、测量误差或数据录入错误等原因导致的,其存在会对数据分析结果产生较大的干扰。常用的异常值检测方法包括标准偏差法、控制图法等。标准偏差法通过计算数据的标准偏差,将偏离均值一定倍数标准偏差的数据视为异常值;控制图法则是根据数据的分布特征绘制控制图,将超出控制界限的数据视为异常值。对于检测到的异常值,可以采用删除、替换或修正等方法进行处

理,具体方法的选择应根据异常值的产生原因和对分析结果的影响程度来确定。数据标准化是将不同来源和单位的数据转换为统一格式和标准的过程。由于水质检测数据可能来自不同的监测设备、不同的监测时间或不同的监测地点,其数据格式和单位可能存在差异。数据标准化可以消除这些差异,使数据具有可比性和可分析性。常见的数据标准化方法包括最小-最大标准化、Z-score标准化等。最小-最大标准化是将数据线性变换到[0,1]区间内,适用于数据分布较为集中的情况;Z-score标准化则是将数据转换为均值为0、标准差为1的标准正态分布,能够更好地保留数据的分布特征。

1.3 水质检测误差来源分析

在水质检测过程中,误差是不可避免的。误差的存在会影响检测结果的准确性和可靠性,因此深入了解误差的来源对于采取有效的误差控制措施具有重要意义。系统误差是由某种固定原因造成的,具有单向性、重复性和可测性。方法误差是系统误差的主要来源之一。在水质检测的滴定分析中,指示剂的选择和使用方法不当会导致滴定终点与化学计量点不完全符合,从而造成测定结果偏高或偏低。此外,副反应的发生也会干扰主反应的进行,影响测定结果的准确性。例如,在测定水中的氯离子含量时,如果存在其他能与银离子发生反应的离子,就会产生副反应,导致测定结果不准确。仪器误差也是常见的系统误差因素。天平、砝码或容量器皿的刻度不准,在使用过程中会使测定结果产生误差。例如,天平的零点漂移、砝码的质量不准确或容量器皿的容积不准确等,都会导致测量结果的偏差。试剂误差则由于试剂不纯或蒸馏水中含有微量杂质引起。不纯的试剂可能会引入干扰物质,影响测定反应的进行;蒸馏水中的微量杂质可能会参与反应,导致测定结果偏高或偏低。操作误差是由操作人员的主观因素导致的。分析人员所掌握的分析操作与正确实验条件不符,如加热温度控制不准确、反应时间掌握不当等,都会影响测定结果的准确性。视觉原因也是操作误差的常见因素之一,在辨别滴定终点颜色和读取刻度时,由于光线、视力等因素的影响,可能会出现偏差。随机误差是由一些偶然和意外的原因产生的,在同一条件下多次测定所出现的随机误差,其大小、正负不定,是非单向性的^[2]。温度、压力等外界条件的突然变化会对测量结果产生影响。例如,温度的变化可能会导致溶液的体积膨胀或收缩,从而影响溶液的浓度和测量结果;压力的变化可能会影响气体的溶解度和扩散速度,进而影响气体成分的测定结果。仪器性能的微小变化,如传感器的灵敏度下降、电

子元件的噪声等,也会引起随机误差。操作技术上的微小差别,如取样的位置、取样的量等,同样会对测量结果产生随机影响。分析仪器最小读数的不确定性也是随机误差的来源之一,例如,数字式仪器的最小显示位数有限,会导致测量结果存在一定的舍入误差。由于随机误差的不可预测性,无法用校正的方法减小或避免,但可以通过增加测量次数,采用统计方法(如计算平均值、标准差等)来降低其对测量结果的影响。

2 水质检测误差控制方法

2.1 规范操作

规范操作是水质检测误差控制的基础,贯穿于水质检验工作的始终。水质检验工作人员在实际工作中,需从多个方面严格把控操作细节,以确保检测数据的准确性和可靠性。仪器设备的清洁与干燥是关键环节。每次检测前,工作人员应仔细检查所选取的仪器,确保其表面无灰尘、污渍等杂质,内部无残留化学物质。残存物质可能会与待测水样发生化学反应,改变水样成分或性质,导致检测数据出现误差。例如,玻璃器皿内壁残留上次实验试剂,在本次检测中,残留试剂可能与新加入试剂或水样反应,影响检测结果准确性。因此,工作人员应使用合适清洁剂和清洗方法对仪器进行彻底清洗,并采用干燥箱、干燥剂等方式将仪器干燥,避免水分残留对检测造成干扰。试剂的取用和实验操作必须严格按照操作规程进行。试剂的错用、过量使用或使用顺序不当等操作失误,都可能导致检测结果出现偏差。在化学分析中,不同试剂具有不同化学性质和作用,若将指示剂错用为标准溶液,或者将标准溶液浓度计算错误,都会使滴定终点判断不准确,影响最终检测结果。所以,工作人员在取用试剂前,应仔细核对试剂的名称、浓度、有效期等信息,按照规定的用量和顺序进行取用。在实验操作过程中,要严格遵循操作步骤,注意操作顺序、反应条件(如温度、时间、搅拌速度等)的控制,避免因操作不当导致数据误差。读数的准确性也是规范操作的重要内容。以读取滴定管数据为例,读数人员的视线应与滴定管内液体的凹面处于同一水平线,这是因为液体表面存在表面张力,会使液体表面形成凹面,若视线不水平,会导致读数偏高或偏低。同时,读取数字时应保留到小数点之后两位,以提高读数的精度。为了进一步减小读数误差,应多次读取数据后选取中位数定为最终检验结果。通过多次读取数据,可以减小偶然误差的影响,而选取中位数则可以避免个别异常数据对最终结果的影响,从而使读取的数据更加有效、准确。

2.2 仪器校准

仪器校准是保证水质检测数据准确性的关键措施。检测仪器在长期使用过程中,由于环境因素(如温度、湿度、灰尘等)、机械磨损、电子元件老化等原因,其性能会逐渐发生变化,导致测量结果出现偏差。因此,定期对检测仪器进行校准和维护至关重要。制定科学的校准计划是仪器校准工作的前提。根据仪器的使用频率和说明书要求,对于使用频率较高的仪器,应适当缩短校准周期;对于使用频率较低的仪器,可适当延长校准周期,但也不能超过说明书规定的最长校准间隔时间。在校准过程中,应严格按照校准方法和标准进行操作,确保校准结果的准确性。以分光光度计为例,它是水质检测中常用的仪器之一,其波长准确性对吸光度测量结果有重要影响。若分光光度计的波长不准确,会导致测量得到的吸光度值与真实值存在偏差,从而影响水质检测结果的准确性。因此,需要定期使用标准波长滤光片对分光光度计的波长进行校准,确保其波长偏差在允许范围内。pH计也是水质检测中常用的仪器,其电极的老化会使测量的pH值偏离真实值。电极在使用过程中,会受到溶液中离子、微生物等因素的影响,导致电极的性能逐渐下降^[3]。因此,需要定期对pH计的电极进行校准和维护,如使用标准缓冲溶液对电极进行校准,清洗电极表面的污垢等,以保证电极的精度和稳定性。

2.3 空白实验

空白实验是一种简单而有效的误差控制方法,其基本原理是通过消除试剂和仪器本身带来的干扰,提高检测结果的准确性。空白实验的核心公式是:实测结果 = 样品结果 - 空白值。在进行空白实验时,需要采用与样品检测完全相同的检测条件进行试验。这意味着要使用相同的试剂、相同的仪器设备、相同的操作步骤和相同的反应条件。例如,在进行水样的化学需氧量(COD)检测时,空白实验应使用与样品检测相同的重铬酸钾溶液、硫酸银-硫酸溶液等试剂,使用相同的消解装置和滴定仪器,按照相同的消解时间和滴定步骤进行操作。将空白实验的试验结果定为空白值,这个空白值包含了试

剂中的杂质、仪器设备的本底噪声以及实验操作过程中可能引入的误差等因素。然后,将样品检测结果减去空白值,就可以得到更为精准的检测数据结果。空白实验可以有效地消除试剂和仪器本身带来的干扰。例如,试剂中可能含有微量的待测物质或其他干扰物质,这些物质会在检测过程中与样品中的物质发生反应,导致检测结果偏高。通过空白实验,可以测定出试剂中的这些干扰物质的含量,并在计算最终结果时予以扣除。同样,仪器设备的本底噪声也会检测结果产生影响,空白实验可以消除这种影响,使检测结果更加准确可靠。此外,空白实验还可以帮助工作人员发现实验过程中可能存在的问题。如果空白值异常偏高或偏低,可能意味着试剂质量存在问题、仪器设备存在故障或实验操作过程中存在误差。通过对空白值的分析,工作人员可以及时采取相应的措施,如更换试剂、维修仪器设备或改进实验操作方法,以提高水质检测的质量。

结束语

水质检测中的数据处理与误差分析是保障水质检测准确性和可靠性的重要环节。通过科学的数据采集与预处理方法,可以获得高质量的数据;深入了解误差的来源,并采取有效的误差控制方法,可以减少误差对检测结果的影响;合理运用数据分析方法,可以挖掘数据背后的信息,为水质评估、处理工艺优化和管理决策提供有力支持。未来,随着水质检测技术的不断发展和数据分析方法的不断创新,水质检测工作将更加科学、准确和高效,为保护水资源和生态环境做出更大的贡献。

参考文献

- [1]胡志勇.废水水质检测化验误差分析与数据处理[J].环境与发展,2018,30(6):174+176.
- [2]孔小禹.废水水质检测化验误差分析与数据处理[J].资源节约与环保,2018(2):50-51.
- [3]张美琴.废水水质检测误差分析及数据处理研究[J].资源节约与环保,2018(08):137-138+111.