

实验室环境检测数据处理分析

段宝鹏¹ 吴扬^{2*}

1 泸州市叙永生态环境监测和应急服务中心 2 四川省天晟源环保股份有限公司

DOI:10.12238/eep.v6i5.1848

[摘要] 针对实验室环境检测数据的处理方法展开探讨,从数据分析误差角度入手,明确了误差种类以及表示方法,并从准确度、精密度,以及灵敏度三个方面,探讨了数据处理与误差之间的关系,并对协同数据的处理和分析进行研究,最后,提出了实验室数据修约处理要点和方法。根据研究结果可知,在实际进行实验室环境检测数据处理的过程中,影响数据处理质量的因素较多,为保障数据处理效果,需要明确数据的准确度、精密度的情况,并做好数据修约等处理,此外,还需要对数据展开一致性检验、离群值检验等,全面保障数据处理的有效性。

[关键词] 实验室; 环境监测; 数据处理

中图分类号: X83 **文献标识码:** A

Processing and Analysis of Laboratory Environmental Testing Data

Baopeng Duan¹ Yang Wu^{2*}

1 Xuyong Ecological Environment Monitoring and Emergency Service Center in Luzhou City

2 Sichuan Tianshengyuan Environmental Protection Co., Ltd

[Abstract] This article explores the processing methods of laboratory environmental testing data, starting from the perspective of data analysis errors, clarifies the types and representation methods of errors, and explores the relationship between data processing and errors from three aspects: accuracy, precision, and sensitivity. It also studies the processing and analysis of collaborative data. Finally, the key points and methods for laboratory data rounding processing are proposed. According to the research results, there are many factors that affect the quality of data processing in the actual process of laboratory environmental testing data processing. To ensure the effectiveness of data processing, it is necessary to clarify the accuracy and precision of the data, and carry out data rounding and other processing. In addition, consistency testing and outlier testing are also necessary to comprehensively ensure the effectiveness of data processing.

[Key words] laboratory; Environmental monitoring; data processing

前言

在当前可持续发展背景下,环境检测工作逐渐得到更多的重视,与此同时,实验室环境检测数据处理和分析工作的重要性逐渐得以凸显,用以反映当前环境质量情况,如空气、水中的化学污染物等,为环境治理以及保护提供可靠数据支持。由此可知,环境检测数据处理质量和效果,对于后续环境保护以及预测预报等相关工作有着重要影响。

1 实验室环境检测数据处理过程中的分析误差

环境检测的主要目的为准确测定各种环境当中的污染物情况,为后续环境预测和保护工作提供可靠支持,因此,对于检测数据处理结果的准确性具有较高要求。但在实际进行环境检测的过程中,环境污染本身具备一定流动性、变化性,而且由于检测技术手段、实验室分析方法、测量仪器、药品试剂等条件限

制,以及人为操作影响等,难以避免会出现检测结果与真实数据存在偏差的情况,即为分析误差^[1]。因此,需要展开大量数据的采集,然后通过统计处理后,供后续环节使用。

1.1 误差种类

结合误差来源以及性质,可将数据处理过程中的分析误差划分为以下两种类型:(1)系统误差,主要是在数据处理和分析过程中的确定因素引起的,属于相同条件下环境检测过程中会重复出现的误差,具有明确方向,与真实值相比总是偏高或者偏低。可通过实施空白试验、对照试验等,消除系统误差。(2)偶然误差,是由于不可测的随机误差造成的,如温度改变、电压波动等,通常难以控制,但是通过多次测定,能够发现偶然误差的规律,此类误差服从正态分布,具有单峰性和对称性。

1.2 表示方法

误差的表示方法主要包括绝对误差、相对误差,绝对偏差、相对偏差,极差,以及标准偏差和相对标准偏差。其中绝对误差和相对误差主要用于体现测定结果的准确性,即误差越小,测定结果的准确性越高。绝对误差和相对误差计算公式如下:

$$G = M - W$$

$$g = \frac{G}{W} \times 100\%$$

式中,G表示绝对误差;M表示测定值;W表示真值;g表示相对误差。

绝对偏差以及相对偏差主要是用于体现某次测量结果,与平均值之间的差异。因此,绝对偏差就是测定值与多次测量结果平均值之间的差异,相对偏差即绝对偏差在平均值当中的占比大小。

极差指的是同一个样品测定过程中,最大测定值与最小测定值之间的差,主要用于表示误差范围。

标准偏差是检测数据处理过程中的常用统计方式,主要用于体现测量结果的统计分布情况,简而言之就是各数据偏离平均数距离的平均数。计算公式如下:

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

式中,s表示标准偏差; x_i 表示第*i*次检测结果; \bar{x} 表示检测结果的均值;n表示检测的总次数,通常*n*>5。

相对标准偏差,即标准偏差与检测结果算术平均值的比,通常在实验室环境检测数据处理过程中,用于表示检测结果的精密度。

随着当前信息技术手段的发展和进步,在实际进行实验室环境监测数据处理的过程中,已经不再通过人工计算的方式处理上述数据,而是借助计算机功能完成计算,最基本的就是Excel,此外,还有专门用于环境检测数据处理和分析的软件系统,极大地提高了当前实验室检测数据处理和分析的准确性,降低了人为因素的影响,同时也能够有效提高数据处理效率。

2 实验室环境检测数据分析

以COD检测仪测量实验为例,实际测量过程中,温度、pH值等会对检测结果产生影响。在高温消解的过程中,消解温度通常为165℃,若温度过高会使得COD过度氧化,导致检测结果偏高;若温度过低,会使得COD反应不充分,导致结果偏低,产生数据误差。pH值方面,COD消解反应需在酸性条件下展开,通常要求pH值在2~4之间,若PH过低会使得COD反应不充分,结果偏低;若PH过高会使得COD过度氧化,结果偏高,因此实际检测过程中,应保持酸碱度恒定,以免影响数据的准确性。

因此,在实验室环境检测数据处理的过程中,为保障数据采集的质量,需要进一步对数据的准确度、精密度以及灵敏度等进行检验。相应数据处理方式如下。

(1)准确度。该指标主要是用于分析某次数据采集结果与假定或者真值之间的契合程度,同时准确度也能够反映系统误差与偶然误差的综合情况,以此检测采集到的数据的可靠性。准确度的评价方式通常包括以下两种:第一,用相同的方法分析标准物质;第二,通过加标回收法,测定准确度,计算公式如下:

$$p = \frac{m-n}{w}$$

式中,p表示回收率;m表示加标试样测定值;n表示试样测定值;w表示加标值。

(2)精密度。即环境检测数据的一致性,用于体现偶然误差的大小,通常采用极差、标准偏差等表示^[2]。在实验室环境检测当中,精密度通常包括以下两个含义:第一,同一个实验室内,分析人员、设备以及时间某一项因素不同的情况下,使用同一种分析方法进行多次独立检测,结果间的符合程度;第二,当处于不同实验室时,采用相同方法进行多次测定,结果间的符合度。

(3)灵敏度。主要用于体现某一方法或者仪器的灵敏程度。通常可用校准曲线的斜率表示灵敏度。

3 实验室环境检测协同数据处理分析

为保障实验室环境检测数据处理结果的可靠性以及有效性,在实际展开数据处理分析的过程中,需要对检测方法的精密度以及再现性进行检测,确保相应实验室检测方法能够得到有效运用。因此,需要展开协同试验,并对相应数据进行处理和分析,对此,国家制定了相应标准,用于实验室数据测量、计算和分析。

3.1 一致性检验

(1)k统计量检验。k是实验室检测过程中展开一致性检验的重要统计量,主要用于展开离散型统计分析。衡量测定结果离散程度的检验临界值计算公式如下:

$$k = \sqrt{\frac{p}{1+(p-1)/F}}$$

式中,k为衡量测定结果离散程度的检验临界值;p为实验室数量;F为F-F分布,可根据分布表获得,也可以通过Excel函数得到。

(2)h统计量检验。主要用于检验实验室数据的一致性,通过h统计量的计算,能够明确单个数据的利群情况。即h值越大,实验结果偏差越大,准确度也会相对更低。在实际展开数据处理和分析的过程中,还需要计算h的临界值,当h的大小超过其临界值时,表示实验室检测结果准确度较差。

3.2 离群值检验

离群值检验的常见方法包括科克伦检验、格拉布斯检验,以及狄克逊检验等。

其中科克伦检验法主要是针对标准差最大值的情况进行评判,在实际数据处理的过程中,需要将检测结果标准差,按照从小到大的顺序依次排列,然后按照科克伦检验法进行统计分析。

计算公式如下:

$$C_i = \frac{S_j^2}{\sum_{j=1}^p S_j^2}$$

式中, C_i 表示*i*水平下科克伦检验值; S_j 表示实验室内标

准偏差; S_3 表示 S_j 最大值。

结合上述计算结果, 进行离群值的检验分析: ①当统计结果小于5%的临界值时, 表示计算结果处于正常范围内; ②当统计结果超过5%的临界值, 并且在1%的临界值以内时, 表示该计算结果较为可疑; ③若统计结果大于1%的临界值, 则表示该值为离群值。

格拉布斯检验法主要用于判断数据当中的最大值或者最小值是否为离群值, 相较于科克伦检验法而言, 属于双侧检验法。实际计算的过程中, 仍然按照从小到大的顺序将实验室检测结果排列出来, 然后计算平均值、标准偏差以及格拉布斯检验法统计结果。在判定离群情况时, 相应判定标准和方法与科克伦检验法相同。案例项目当中各水平值的格拉布斯法的检验结果如表1所示。根据表2可知, 所有实验数据均不存在离群值和可疑值。

表1 各水平值的格拉布斯法的检验结果

统计量名称	不同水平统计量值			
	1	2	3	4
水平	1	2	3	4
实验室数	8	10	10	9
G_i	1.200	1.282	1.771	1.595
$G_{0.95,p}$	2.031	2.175	2.175	2.111
$G_{0.99,p}$	2.220	2.411	2.411	2.322

狄克逊检验是实验室环境检测数据处理过程中, 检验异常值的常见方法, 应用效果相对较好, 不仅具有单侧检验, 还可展开双侧检验。以双侧检验为例, 实际进行数据处理的过程中, 需要先进行数据排列, 然后计算相应统计量, 计算公式如下:

$$D_{\max} = \frac{x_{\max} - x_{\max-1}}{x_{\max} - x_2}$$

$$D_{\min} = \frac{x_2 - x_1}{x_{\max-1} - x_1}$$

$$D_i = \text{MAX}\{D_{\max}, D_{\min}\}$$

式中, D 表示统计量; x 表示实验室检测结果; x_{\max} 表示最大值。

在实际进行离群值的判定时, 仍以5%的临界值和1%的临界

值为参考标准。

4 实验室环境检测数据处理要点

在实验室环境检测数据处理过程中, 为实现对于数据的有效处理和清洗, 需要通过相应运算, 保留有效数字、处理可疑数据, 数据处理要点内容如下。

(1) 有效数字判定。有效数字来源于实验测量, 包括准确数字和最后一位的可疑数字, 主要用于反映数据的准确程度, 而有效数据的位数则表明了相应实验仪器的精密情况。在实际判定有效数字时, 需要根据其所处的位置。如使用分析天平称取试样时, 得到的质量为0.3460g, 该数据当中的有效数字的位数是四位, 最后一位“0”为可疑数字, 这不仅表明了质量的大小, 同时也体现了天命的精度。

(2) 数字修约处理。在实际进行数据处理的过程中, 确定有效数字的位数后, 需要舍去多余的数字, 即为修约。在实际进行修约的过程中, 若被修约数字小于等于4时, 则舍去; 等于5, 并且后面的数字不全为0时, 则进位; 等于5, 且后面的数字全为0时, 若进位后位数为偶数, 则进位, 若舍去后的尾数为偶数, 则舍去; 当被修约数字大于等于6时, 则进位^[3]。

5 结束语

综上所述, 实验室环境检测数据处理分析过程中, 由于各方面因素的影响, 难以避免会出现数据误差, 主要包括系统误差和偶然误差, 对于实验结果有着直接的影响。对此, 为保障实验结果的准确性和可靠性, 需要展开数据处理, 根据实际需求计算测量结果的绝对误差、相对误差, 绝对偏差、相对偏差, 极差, 以及标准偏差和相对标准偏差等, 实现对于实验室环境检测数据准确度、精密度以及灵敏度的分析和判断。此外, 还需要对检测结果进行协同数据处理, 展开一致性检验, 以及离群值的检验和分析, 在实际检验过程中, 应根据检测需求选择合理检验方法, 完成对于可疑数据的清理, 保障试验结果的准确性。值得注意的是, 在进行数据处理的过程中, 还应根据实际需求, 进行有效数字的判定以及数据的修约处理, 全面保障检测结果的可靠性以及数据的实用价值。

【参考文献】

[1]刘文琳.检测实验室检测数据质量控制关键技术探讨[J].云南化工,2021,48(12):86-88.

[2]冯淇.生态环境检测实验室现场采样质量管理技术数字化研究与应用[J].皮革制作与环保科技,2021,2(21):68-69.

[3]王伟荔.提高环境监测数据可靠性的策略探析[J].节能,2019,38(10):148-149.

作者简介:

段宝鹏(1982--),男,汉族,甘肃省镇原县人,大学本科,环境监测中级工程师,从事环境监测采样和分析、质量管理和环境管理等工作。

通讯作者:

吴扬(1991--),男,汉族,四川省仁寿县,硕士,环境工程工程师,从事土壤环境调查评估与修复等工作。