

**JJF** XXXX—20XX

**中华人民共和国国家计量技术规范**

取水监测计量数据质量控制技术规范

Technical Specification for Quality Control of Water Intake Monitoring and Metering Data

发 布

××××-××-××实施

××××-××-××发布

**国家市场监督管理总局**

发 布

发 布

|  |  |
| --- | --- |
| 取水监测计量数据质量控制技术规范  Technical Specification for Quality Control of Water Intake Monitoring and Metering Data | JJF XXXX—20XX **代替JJF1033-2001** |

归口单位：全国能源资源计量技术委员会水资源计量分技术委员会

主要起草单位：

参加起草单位：

本规范委托全国能源资源计量技术委员会水资源计量分技术委员会负责解释

本规范主要起草人：

参加起草人：

目  录

[引言 I](#_Toc4792)

[1 范围 1](#_Toc5112)

[2 规范性引用文件 1](#_Toc241)

[3 术语和定义 1](#_Toc2378)

[4 质量控制责任 2](#_Toc29979)

[4.1 监测计量设备 2](#_Toc29825)

[4.2 安装与调试 2](#_Toc25567)

[4.3 数据报送 2](#_Toc4)

[4.4 日常管理 2](#_Toc5127)

[4.5 监督检查 2](#_Toc26910)

[5 数据质量标准 2](#_Toc5927)

[5.1 取水监测计量分类 2](#_Toc8975)

[5.2 取水监测计量精度 2](#_Toc9083)

[5.3 监测计量数据存储 3](#_Toc17209)

[5.4 监测计量数据质量检查 3](#_Toc6313)

[5.5 监测计量数据准确性要求 3](#_Toc29516)

[6 监测计量 3](#_Toc10169)

[6.1 在线监测计量 3](#_Toc32228)

[6.2 其他监测计量 3](#_Toc8131)

[7 数据质量检查 4](#_Toc4339)

[7.1 异常数据类型 4](#_Toc11559)

[7.2 数据标记方法 4](#_Toc27829)

[7.3 入库检查 4](#_Toc1853)

[8 质量控制落实 5](#_Toc11004)

[8.1 异常发现 5](#_Toc10473)

[8.2 异常跟踪处理 5](#_Toc27153)

[8.3 异常数据处理策略 5](#_Toc21116)

[附录A 应用数据质量检查 6](#_Toc16263)

[A.1 质量检查内容 6](#_Toc18357)

[A.2 检测结果处理要求 7](#_Toc15018)

[A.3 质量检查时间安排 7](#_Toc4856)

[附录B 质量检查方法 10](#_Toc15226)

[附录C数据质量检测报表 14](#_Toc7907)

## 

引  言

取水监测计量是水资源开发利用管理的重要环节，在取水监测计量数据采集、传输、存储等过程，对数据质量存在多种干扰因素，直接影响水资源日常管理工作。

《取水许可管理办法》(2008年发布，2015年修正，2017年修正)规定，取水单位或者个人（以下简称取用水户）为监测计量数据质量责任主体，各级水资源管理部门负责对数据质量进行监管。为了有效提高监测计量数据质量，服务于水资源日常管理及国家水资源调配宏观决策，制订本规范。

本规范给出了监测计量现场数据质量控制方式及标准，取用水户对异常情况应及时发现并修复处理；取水监测计量数据传输至水资源管理部门指定的存储节点后，相关管理部门应对数据进行跟踪管理。

本规范给出了常见异常数据类型定义及检查方法，水资源相关管理部门可依照本规范，对原始数据库、应用数据库等开展数据质量控制，发现异常数据时，应及时敦促取用水户进行现场检查，并确认异常原因、排除故障。

在水资源开发利用取水统计工作中，可应用本规范给定方法对所采用的数据开展质量审定。

本规范依据JJF 1071—2010给出的规则起草，首次发布。

取水监测计量数据质量控制技术规范

1. 范围

本标准对取水监测计量过程中的数据采集、传输、存储等环节数据质量控制给定了规范方法及数据质量报表。

本标准适用于水资源管理和开发利用过程中自动采集传输的数据质量控制管理。

1. 规范性引用文件

下列文件中的内容通过文中的规范引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 30943-2014 水资源术语

GB/T 33113-2016 水资源管理信息对象代码编制规定

GB/T 28714-2012 取水计量技术导则

SL/Z 349-2015 水资源监控管理系统建设技术导则

SL 365-2015 水资源水量监测技术导则

SL/Z 376-2007 水利信息化常用术语

SL/T 426-2021 水量计量设备基本技术条件

1. 术语和定义

取水监测 water intake monitoring

利用取水工程或设施从地表、地下或其他水源取用水，并使用设备对水位、水深、流速、流量等进行测量转换为水量的过程。



取水计量 water intake metering and Monitoring

按照有关法律法规和技术规范实现取水水量单位统一、量值准确可靠的活动。



在线监测 automatic monitoring

采用自动在线采集设备获取数据，通过网络实时或定时将数据发送至接收处理端的数据采集方式。



离线监测 off-line monitoring

安装仪器设备用于对取水过程的取水量进行测量，所测量结果仅在现场指示，不进行实时远程传输的测量方式。



人工监测 manual monitoring

由观测人员携带计量设备或利用现场已安装的设备，通过人工施测，并参与实现监测数据入库存储的监测方式，主要包括驻测和巡测两类。



取用水户 water user

利用取水工程或设施从地表、地下或其他水源等取水单位或者个人。

1. 质量控制责任
   1. 监测计量设备

取水设施应按照SL/T 426规范选型与配置监测计量设备，所配置的计量设备应符合国家规定的质量要求。取用水户应对设备质量进行检查，确定合格。供货单位应向购货单位提供质量承诺。

* 1. 安装与调试

取用水户负责组织开展取用水监测计量设备安装与调试，各操作环节应符合对应设备的安装规范技术要求，调试期运行稳定后方可进行验收。

* 1. 数据报送

取用水户应将监测计量数据报送至水资源行政管理部门（或其指定的部门）给定的地址，报送频次、数据格式与单位等应符合其要求。

* 1. 日常管理

取用水户为监测计量数据质量责任主体，对取水监测计量设备应加强日常维护管理，出现故障及时修复。

* 1. 监督检查

各级水资源行政管理部门负责对数据质量进行监管。为了有效提高监测计量数据质量，服务于水资源日常管理及国家水资源调配宏观决策，水资源行政管理应配置相关软硬件技术装备，及时发现异常数据，并督促取用水户检查维修。

1. 数据质量标准
   1. 取水监测计量分类

取水监测计量可分为折算计量、直接计量，以及在线监测、离线监测、人工监测等方式，通过在线监测、离线监测、人工监测等方式是实现直接计量的主要途径。

* 1. 取水监测计量精度

取水监测设备的监测要素测量精度按照SL426中规定技术要求执行，不在SL426范围的监测设备应不低于同类设备精度要求。

取水计量精度应满足GB/T28714规定。

明渠流取水计量流量测验成果抽检及管道流量计检定应符合GB/T28714规定。

* 1. 监测计量数据存储

取水监测计量数据宜分为原始数据库、应用数据库存储，原始数据库存储在线监测或离线监测采集的原始记录；应用数据库存储经质量检查、分析、统计、计算、修正异常后的监测记录，直接服务于取水管理业务软件系统。

* 1. 监测计量数据质量检查

在利用监测计量数据进行计费、用水统计、信息发布等应用前，应对监测数据进行质量检查，确保各监测数据符合准确性要求。

* 1. 监测计量数据准确性要求

考虑仪器设备自身精度条件下，与工程取水能力、实际用水需求量、前后用水过程、同期对比等处理分析后，不得出现明显偏差。

1. 监测计量
   1. 在线监测计量

6.1.1 取水监测仪器设备配置应符合SL426技术条件。

6.1.2 取水监测仪器设备安装应满足GB/T28714规定。

6.1.3 取水计量设施检定校准周期和方式应符合*GB/T28714监督管理规定。*

6.1.4 应定期开展对取水监测仪器设备的系统性全面维护，数据远程设备的易损易耗件、避雷、电磁防护等应按照相关规范定期开展维护保养，维护保养周期一般不超过18个月。

6.1.5 监测数据应及时上传，在断电、信号差、无信号等状况发生时，应及时采取措施排除故障。

* 1. 其他监测计量

6.2.1 其他监测计量包括离线监测、人工监测、折算计量等方式。

6.2.2 采用以电折水间接折算（以用电量折算为水量）的取水计量，或其他折算方式，在对数据可靠性加强管理时，可采用以下一种或多种方式实施。

a）建立计量电表、取用水户、用水范围等对应关系；

b）提高数据采集频次，常用频次为每月、半月或旬；

c）用水定额核算控制；

d）甄别其他用电，提高实际取水用电量准确性；

e）建立用电量与取水量关系模型；

f）建立取水自动监测站网，确定折算系数。

6.2.3 在数据采集方面，若采用人工现场读数的，应采用以下一种或多种方式加强数据质量控制管理。

a）双方现场核对仪表读数，并用手持终端录入或手写记录；

b）水政人员单方记录时，应提醒取用水户进行核对；

c）取用水户现场读数并报出，水政人员应不定期进行现场核对；

d）以取用水规模、用水类型等类似区域，建立自动监测与人工现场读数上报比对关系。

1. 数据质量检查
   1. 异常数据类型

7.1.1 数据缺失，是指取水正常发生，而产生的数据连续性或间断性缺失。

7.1.2 按照与前后数据、同期数据等进行比较，达到一定变化幅度可认为发生突变。与前期一段时间监测的数据比较发生了较大幅度变化，且前一段时间监测数据均较正常；该变幅超过平均变幅（标准差）数倍以上，可认为发生突变。

数据突变类型可分为突变大、突变小、突变零等，具体分别如下：

（1）突变大。与通常监测数据相比该监测数据变大了很多，增大量达到平均变幅数倍以上（如3倍）。

（2）突变小。与通常监测数据相比该监测数据变小了很多，减少量达到平均变幅数倍以上（如3倍）。

（3）突变零。该监测数据为零值，而前期监测数据不为零。

7.1.3 持续不变数据，一定时段内多个连续监测数据保持不变，各数据值相同。

7.1.4 波动变化数据，一定时段内数据连续波动变化、规律明显，波峰、波谷数据各保持不变。

7.1.5 各水资源管理单位可根据辖区内监测实际，增加异常类型，以便给出合适的异常确定方法。

* 1. 数据标记方法

7.2.1 各监测数据接收中心，宜分别建立原始监测数据库、应用数据库，以及异常分析记录数据库。不具备条件分库建设的应对数据处理属性做好标记，分为原始数据、修正数据等。

7.2.2 异常分析记录须包括异常数据、跟踪处理过程、处理结果、主要参与人员等内容。

7.2.3 按照异常类型，对异常数据分别给出标记。异常标记代码参照使用如下所示。

表1 异常标记代码表

| 代码 | 异常说明 |
| --- | --- |
| 1 | 数据缺失，在非季节性或企业生产周期影响下，产生的数据连续性或间断性缺失监测值 |
| 2 | 突变大，相对前一段时间内监测数据突然变大很多，达到平均变幅数倍以上，可根据行业特点、用户特点等确定具体倍数 |
| 3 | 突变小，相对前一段时间内监测数据突然变小很多，达到平均变幅数倍以上，可根据行业特点、用户特点等确定具体倍数 |
| 4 | 突变零，该监测数据为零值，而前期监测数据不为零 |
| 5 | 多个连续监测数据保持不变，各数据值相同，该数据为其中之一 |
| 6 | 连续波动上下起伏跳跃变化、规律明显，该数据为其中之一 |
| 9 | 其他异常现象 |

7.2.4 在异常分析记录数据库设计时，应保证按照异常数据时间、监测点代码等特征属性，在原始数据库、应用数据库及异常分析记录数据库等之间可进行关联查询。

* 1. 入库检查

对数据质量控制宜以原始数据库为基础，对各行业、各类取用水户等可能产生异常数据类型进行分析，在该基础上建立异常数据检测软件系统。

1. 质量控制落实
   1. 异常发现

8.1.1 取用水户是异常数据发现的第一责任主体。

在测站日常巡检中若发现导致数据异常的隐患，应及时处理；导致数据异常的隐患主要有避雷设施遭破坏、外界新增高压磁场环境、信号线接触不良等。

8.1.2 水资源行政管理部门有责任监督检查取用水数据的准确可靠性。

由相关数据管理、业务应用部门发现的数据异常，应及时通知取用水户，将异常信息告知监测现场维护部门。

8.1.3 在异常天气环境、强对流天气现象发生后，对易受威胁的站点应及时开展现场检查，确定是否有导致数据异常的隐患存在。

8.1.4 取水户管理及生产经营需要，暂停取水导致监测数据出现异常，该类异常应进行备案，并按照数据存储规范在数据记录中备注异常原因。

* 1. 异常跟踪处理

8.2.1 取用水户应建立数据质量控制制度。

8.2.2 取用水户落实质量控制时，应按照监测要素及传感器类型对异常数据进行分类，并对不同异常类型的处理过程进行总结，积累异常处理经验。

8.2.3 运行维护责任单位应在接到异常情况通知后，及时至监测现场开展问题排查处理，记录排查处理过程。

8.2.4 当发生异常时，应对远程传输单元（RTU）进行详细检查并做相应处理，检查内容包括避雷设施、电源、数据连接线、接口模块等，并确定附近周围是否有高电流、高磁场设备。

8.2.5 对于明渠流监测，应分别根据不同测量方式、测量设施按照GB/T28714安装规范要求进行检查，确定是否各传感器、计量设施是否运行正常。

8.2.6 对于管道流量计应检查测量管段是否淤积，传感器是否出现松懈、脱落等现象。

* 1. 异常数据处理策略

8.3.1 对各类异常数据宜采取人工干预处理。

8.3.2 经调查判别后，若软件系统自动检测出的异常数据属于实际取水行为所致，则该数据应存入应用数据库，视为实际取水行为的监测结果。

8.3.3 经调查判别后，因仪器设备受自然环境、非人为有意干扰造成的突变，可按近期的均值修正后存入应用数据库，并对仪器设备开展维护检测。

8.3.4 经调查，是人为有意干扰破坏造成突变，经相关责任人确认后，须按相关管理办法进行处罚，对该异常数据按近期均值修正后存入应用数据库。

8.3.5 对各异常数据存入应用数据库，须在对应的数据记录中增加备注信息，说明异常原因及处理方式。

附录A

应用数据质量检查

A.1 质量检查内容

A.1.1 水资源行政管理部门应加强对应用数据质量检查。包括过程数据质量检查、汇总数据质量检查、数据集质量检查三种。

A.1.2 过程数据质量检查，主要是当前入库数据的异常检查和一段时间监测序列数据的异常检查。

（1）当前入库数据的异常检查，宜采取判断其突变类型为主的的快速检查，并与原始数据库的异常检查相互对照，若与原始数据库异常检查结论不一致时，应及时启动告警反馈程序；

（2）一段时间监测序列数据的异常检查，以最新入库的一段时间数据序列为检查数据集，分析判断序列数据的整体特征，判断序列数据中的点异常和系列异常情况，加以标注和告警反馈。

A.1.3 汇总数据质量检查，包括单监测点分时段汇总数据的异常检查、多监测点分时段汇总数据的异常检查、分行业分时段汇总数据的异常检查。

（1）单监测点分时段汇总数据的异常检查，是对单一监测点上报数据，按照时、日、月三种间隔进行汇总，分别对形成的三类汇总数据序列，按照日、月、年进行整体特征分析，判断异常情况，加以标注和告警反馈；

（2）多监测点分时段汇总数据的异常检查，是对两个以上相关监测点（相关监测点宜人工设定）上报数据，按照时、日、月三种间隔进行汇总，将多监测点两两组合，分别将汇总数据，按照同一时段数据组织为二元组，对二元组序列，进行相关性分析判断。判断异常点，加以标注和告警反馈。

分组示例：对、、三个相关监测点的日过程数据汇总分组

第一步：形成分小时的汇总数据

监测点的按小时汇总的序列为：

监测点的按小时汇总的序列为：

监测点的按小时汇总的序列为：

第二步：形成二元组序列

*、*监测点组合的二元组序列：<>,<>,….<>

*、*监测点组合的二元组序列：<>,<>,….<>

*、*监测点组合的二元组序列：<>,<>,….<>

（3）分行业分时段汇总数据的异常检查，是对一定区域内各用水户，先按照一定时段进行取水量汇总，再按照行业进行分类，对比分析是否存在用水量异常情况。

A.1.4 数据集质量检查，是对按区域分类的数据集，进行及时率、完整率、异常率的检查。

（1）监测数据集应采取县、市、省三级分类；

（2）及时性，是对各监测点上报数据，是否依照设定的应上报时间及时上报情况的统计，即未及时上报数和应上报数的比值；

示例：计算县级上报数据及时率

第一步：统计县辖监测点数据上报及时率

若监测点设定的在时间段内应上报条数据，实际按时上报数为条

则监测点x的及时上报率为

第二步：统计县级上报数据集及时率

某县级区域辖区有n个监测点，x1，x2，，，xn

及时率 =监测点及时上报数合计/各监测点应上报数合计

（3）完整率，是各监测点上报数据，是否按照规约要求完整上报的情况，即，各要素完整上报数据条数和上报数据总条数的比值。

（4）异常率，是对各监测点上报数据的异常数据条数和上报数据总条数的比值。

A.2 检测结果处理要求

A.2.1 应用数据库数据检查异常，应与原始库联动、异常分析记录数据库联动，建立关联查询方式。

A.2.2 应用数据库不宜在数据记录上增加异常处理相关信息。

A.2.3 应用数据库异常数据，根据异常处理情况，其内容应及时更正。其原始数据、异常告警、异常分析、处理方案、处理结果、应用数据更正情况等分别在原始数据库和异常分析记录数据库如实记录，永久存储。

A.2.4 过程数据异常，宜软件系统自动告警，报送单位需每日检查告警记录，并参照6.2规定检查异常原因，及时处理，并将处理结果登记到异常分析记录数据库。

A.2.5 各类异常告警，均需要人工确认并处理。

A.3 质量检查时间安排

A.3.1 应用数据库数据质量检查，分为即时检查和定期检查。即时检查是对单一监测点监测过程当前入库数据的质量检查；定期检查，是对一段时间过程数据以及时段统计数据的规定时间内的质量检查，进一步分为单一监测点时段统计数据检查、多监测点时段统计数据对比检查、分行业时段统计数据对比检查。

A.3.2 按照质量检查类型，各类检查的时间安排，见表2。

表A.1 质量检查时间安排

| 检查类型 | 序号 | 检查内容 | 检查方法 | 检查时间 |
| --- | --- | --- | --- | --- |
| 过程数据检查 | 1 | 当前入库数据的异常检查。检查是否超出设定上下限阈值，检查是否与均值偏离幅度过大 | 上下限阈值判断  标准差法 | 即时检查 |
| 2 | 一段时间监测序列数据的异常检查。检查1个严格的异常数据 | 格拉布斯检验法 | 定期检查  日 |
| 3 | 一段时间监测序列数据的异常检查。检查序列数据中多个疑似异常数据 | 狄克逊检验法 | 定期检查  日 |
| 4 | 一段时间监测序列数据的异常检查。检查较长时间无变化或变化幅度微小片段、检查较长时间波动规律性较强，峰值变化幅度微小片段 | 皮尔逊相关系数法 | 定期检查  日 |
| 5 | 对各监测点过程数据异常处理情况检查 | 处理痕迹检查  统计图表 | 定期检查  月 |
| 汇总数据检查 | 6 | 单监测点分时段汇总数据的异常检查。检查一天（当天0:00至当天24:00）分小时汇总数据异常情况，从汇总数据点异常，进一步探究过程异常 | 格拉布斯检验法  狄克逊检验法 | 定期检查  日 |
| 7 | 单监测点分时段汇总数据的异常检查。对一月数据，分日、分时汇总，并按每日的同时间分组，检查汇总数据异常情况 | 格拉布斯检验法  狄克逊检验法  统计图表 | 定期检查  月 |
| 8 | 单监测点分时段汇总数据的异常检查。对一年数据，分月、日、时汇总，检查汇总数据异常情况 | 格拉布斯检验法  狄克逊检验法  散点、折线统计图 | 定期检查  年 |
| 9 | 多监测点分时段汇总数据的异常检查。对两个以上相关监测点数据，按照时、日、月间隔进行汇总，进行相关性分析判断。 | 皮尔逊相关系数法  柱状对比统计图 | 定期检查  月、季 |
| 10 | 分行业分时段汇总数据的异常检查。对一定区域内各用水户，先按照日、月、季进行取水量汇总，再按照行业进行分类，对比分析是否存在用水量异常情况。 | 皮尔逊相关系数法  柱状对比统计图 | 定期检查  月、季、年 |
| 11 | 对汇总数据异常处理情况检查 | 处理痕迹检查  统计图表 | 定期检查  月、年 |
| 数据集检查 | 12 | 及时性检查和更新 | 按7.1.4 规定进行汇总统计  统计报表 | 日、月、年 |
| 13 | 完整性检查和更新 |
| 14 | 异常率检查和更新 |

附录B

质量检查方法

B.1 在应用数据库进行数据质量检查可根据异常特点采用相应的较为成熟方法，本标准仅对标准差法、格拉布斯、狄克逊、皮尔逊等常用方法介绍，鼓励对其他方法开展研究探索。

B.2 标准差法

（1）本方法适用于单一监测点当前入库数据的快速异常数据检查。

（2）检查步骤如下：

第一步：提取除待检查数据之外，且最近已入库并判定为正常的连续监测数据30条，若不足30条，全部提取，形成数据序列

第二步：计算数据序列的平均值 和标准差

第三步：判断新入库数据与平均值的偏差，若 ，则标注为异常数据

（3）算法流程

根据异常数据类型，采用堆栈方式建立数据结构，每个计算周期与监测周期一致（监测频率过密，应抽稀），其算法基本流程如下。

（1）初始化。建立原始数据堆栈Y[n]，根据监测数据的频率设置堆栈长度，为了掌握数据变化过程，一般原始数据堆栈一般应存间隔在10分钟以上的30个数据，若频率很大，则将数据抽稀；建立W[m]波动数据检查堆栈。n=0；m=0；

（2）自当前时间点顺序读取原始数据；

（3）计算Y[n]平均值及标准差；

（4）监听是否有新监测数据到达，读取新到监测数据F；

（5）判断F是否为零，若是则按异常标记该监测数据，并转至（4）；若不是零，则继续执行；

（6）计算|F-|，并与比较，若是的3倍以上，则按异常标记该监测数据；否则，该数据入栈Y[n]，W[m]=F，m++；

（7）若m>10，则判断W[m]是否为持续不变或波动变化，若是，则按异常标记该段数据，并清空Y[n]、W[m]，等待n个计算周期，并读入Y[n]；

（8）转至（4）。

B.3 格拉布斯（Grubbs）检验法

（1）本方法适用于单一监测点系列入库数据的异常值检查，可对系列数据分多组，每组可检查1个异常数据。

（2）待检查系列数据应超过30个。

（3）将待检查数据，按入库时间先后依次分组，每组100个数据。若最后一组未达到31个，该组数据作为下一次检查的待查数据。

（4）对每组数据，检查步骤如下：

第一步：将待检查数据组的个数据从小到大排列

第二步：按照下列公式，计算统计量和计算

其中，和 分别是序列数据的平均值和标准差，计算公式见7.2.1

第三步：按数据组数据的数量 值查表2，获得临界值

第四步： 当 且 则 为异常值；当 且 则 为异常值；

第五步：若第四步发现异常数据，进行标注。本组检查结束

第六步：若还有待检查的数据组，返回第一步，否则本次检查终止

表B.1 格拉布斯检验临界值（显著水平0.01）

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***n*** | ***G*0.01(*n*)** | ***n*** | ***G*0.01(*n*)** | ***n*** | ***G*0.01(*n*)** | ***n*** | ***G*0.01(*n*)** | ***n*** | ***G*0.01(*n*)** |
| 31 | 3.119 | 45 | 3.292 | 59 | 3.405 | 73 | 3.487 | 87 | 3.551 |
| 32 | 3.135 | 46 | 3.302 | 60 | 3.411 | 74 | 3.492 | 88 | 3.555 |
| 33 | 3.150 | 47 | 3.310 | 61 | 3.418 | 75 | 3.496 | 89 | 3.559 |
| 34 | 3.164 | 48 | 3.319 | 62 | 3.424 | 76 | 3.502 | 90 | 3.563 |
| 35 | 3.178 | 49 | 3.329 | 63 | 3.430 | 77 | 3.507 | 91 | 3.567 |
| 36 | 3.191 | 50 | 3.336 | 64 | 3.437 | 78 | 3.511 | 92 | 3.570 |
| 37 | 3.204 | 51 | 3.345 | 65 | 3.442 | 79 | 3.516 | 93 | 3.575 |
| 38 | 3.216 | 52 | 3.353 | 66 | 3.449 | 80 | 3.521 | 94 | 3.579 |
| 39 | 3.228 | 53 | 3.361 | 67 | 3.454 | 81 | 3.525 | 95 | 3.582 |
| 40 | 3.240 | 54 | 3.368 | 68 | 3.460 | 82 | 3.529 | 96 | 3.586 |
| 41 | 3.251 | 55 | 3.376 | 69 | 3.466 | 83 | 3.534 | 97 | 3.589 |
| 42 | 3.261 | 56 | 3.383 | 70 | 3.471 | 84 | 3.539 | 98 | 3.593 |
| 43 | 3.271 | 57 | 3.391 | 71 | 3.476 | 85 | 3.543 | 99 | 3.597 |
| 44 | 3.282 | 58 | 3.397 | 72 | 3.482 | 86 | 3.547 | 100 | 3.600 |

B.4 狄克逊（Dixon）检验法

（1）本方法适用于单一监测点系列入库数据的异常值检查，可对系列数据分多组，每组可检查10个以内的异常数据。

（2）待检查系列数据应超过31个。

（3）将待检查数据，按入库时间先后依次分组，每组100个数据。若最后一组未达到31个，该组数据作为下一次检查的待查数据。

（4）对每组数据，检查步骤如下：

第一步：将待检查数据组的个数据从小到大排列

第二步：按照下列公式，计算和计算

第三步：按数据组数据的数量 值查表3，获得临界值

第四步： 当 且 则 为异常值；当 且 则 为异常值；若未能找出异常值，则本组检查终止

第五步：将检查出的异常值进行标注。若异常数据达到10个，将本组数据标记为系列异常，并终止检查；否则，从检查数据组中剔除该异常数据，返回第一步重复进行检查

表B.2 狄克逊检验临界值（显著水平0.05）

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***n*** | ***D0*.05(*n*)** | ***n*** | ***D0*.05(*n*)** | ***n*** | ***D0*.05(*n*)** | ***n*** | ***D0*.05(*n*)** | ***n*** | ***D0*.05(*n*)** |
| 31 | 0.371 | 45 | 0.323 | 59 | 0.295 | 73 | 0.277 | 87 | 0.263 |
| 32 | 0.367 | 46 | 0.321 | 60 | 0.294 | 74 | 0.276 | 88 | 0.262 |
| 33 | 0.362 | 47 | 0.318 | 61 | 0.292 | 75 | 0.275 | 89 | 0.262 |
| 34 | 0.358 | 48 | 0.316 | 62 | 0.291 | 76 | 0.274 | 90 | 0.261 |
| 35 | 0.354 | 49 | 0.314 | 63 | 0.289 | 77 | 0.273 | 91 | 0.260 |
| 36 | 0.350 | 50 | 0.312 | 64 | 0.288 | 78 | 0.272 | 92 | 0.259 |
| 37 | 0.347 | 51 | 0.310 | 65 | 0.287 | 79 | 0.271 | 93 | 0.259 |
| 38 | 0.343 | 52 | 0.308 | 66 | 0.285 | 80 | 0.270 | 94 | 0.258 |
| 39 | 0.340 | 53 | 0.306 | 67 | 0.284 | 81 | 0.269 | 95 | 0.257 |
| 40 | 0.337 | 54 | 0.304 | 68 | 0.283 | 82 | 0.268 | 96 | 0.256 |
| 41 | 0.334 | 55 | 0.302 | 69 | 0.282 | 83 | 0.267 | 97 | 0.255 |
| 42 | 0.331 | 56 | 0.300 | 70 | 0.280 | 84 | 0.266 | 98 | 0.255 |
| 43 | 0.328 | 57 | 0.298 | 71 | 0.279 | 85 | 0.265 | 99 | 0.254 |
| 44 | 0.326 | 58 | 0.297 | 72 | 0.278 | 86 | 0.264 | 100 | 0.254 |

B.5 皮尔逊（Pearson）相关系数法

（1）本方法适用于多组监测数据序列、分时段汇总数据序列的两两相关性对比分析，可判断序列间的相似程度，或判别出与其他序列明显不同的数据序列。

（2）要求比较的各序列数据个数相同且超过30个，数据排列时序相同。

（3）每对数据序列，按照以下步骤进行相关性分析：

第一步：从序列组中提取两组数据序列，和

第二步：按照下列公式，计算相关系数，记录组合计算的相关系数值

其中，和分别是两组序列数据的平均值

第三步：按照表4确定两组数据的相关程度，r>0为正相关，r<0为负相关，r=0为无相关性

表B.3 相关系数分段描述

|  |  |
| --- | --- |
| **相关系数|r|** | **相关性描述** |
| 0.00-0.19 | 极低相关 |
| 0.20-0.39 | 低度相关 |
| 0.40-0.69 | 中度相关 |
| 0.70-0.89 | 高度相关 |
| 0.90-1.00 | 极高相关 |

第四步：所有序列组均以两两组合完成相关性计算，则终止，否则返回第一步继续进行

（4）在数据序列同期比较分析中，若r<0.20 或r>0.89 宜于异常告警；在分析数据序列是否存在长期不变或幅度变化微小，或数据序列规律性波动且极值不变或变幅微小时， 若r>0.89 宜于异常告警。在同行业用水量分析中，若r<0.20 宜于异常告警。

附录C

数据质量检测报表

表C.1 原始数据库数据质量检测报表

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 检测开始时间 | | | | | （1） | 检测截止时间 | | | | | （2） |
| 总监测点（个） | | | | | （3） | 总数据量（个） | | | | | （4） |
| 数据缺失异常 | 监测点代码 | | | | | 数据个数 | | | | | |
| （5） | | | | | （6） | | | | | |
|  | | | | |  | | | | | |
|  | | | | |  | | | | | |
| 小计 | *（7）监测点个数* | | | | 小计 | *（8）总数据个数* | | | | |
| 数据突变异常 | 监测点代码 | | | | | 数据个数 | | | | | |
|  | | | | |  | | | | | |
|  | | | | |  | | | | | |
|  | | | | |  | | | | | |
| 小计 |  | | | | 小计 | |  | | | |
| 连续不变异常 | 监测点代码 | | | | | 数据个数 | | | | | |
|  | | | | |  | | | | | |
|  | | | | |  | | | | | |
|  | | | | |  | | | | | |
| 小计 |  | | | | 小计 | | |  | | |
| 起伏跳变异常 | 监测点代码 | | | | | 数据个数 | | | | | |
|  | | | | |  | | | | | |
|  | | | | |  | | | | | |
|  | | | | |  | | | | | |
| 小计 | |  | | | 小计 | | |  | | |
| 其他异常异常 | 监测点代码 | | | | | 数据个数 | | | | | |
|  | | | | |  | | | | | |
|  | | | | |  | | | | | |
|  | | | | |  | | | | | |
| 小计 | | |  | | 小计 | | |  | | |
| 监测点个数合计 | | | | | （9） | 数据个数合计 | | | | （10） | |

注：（1）填写检测开始时间，包括年月日时分；（2）填写检测截止时间，包括年月日时分；（3）该检测时间段内参与检测的监测点个数；（4）检测时段内参与检测的监测数据总个数；（5）出现异常的监测点代码；（6）对应监测点的异常数据个数；（7）该类异常监测点总个数；（8）该类异常数据总个数；（9）以上各类异常监测点总个数；（10）以上各类异常数据个数。

表C.2 异常数据交办单

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 监测点代码 |  | 测站代码 | | |  | | | 测站名称 | |  |
| 检测时段 | （1） | 检测人 | | | （2） | | | 联系电话 | |  |
| 序号 | 异常值 | | 发生时间 | | | 备注 | | | | |
|  | （3） | | （4） | | | （5） | | | | |
|  |  | |  | | |  | | | | |
|  |  | |  | | |  | | | | |
|  |  | |  | | |  | | | | |
| 报文入原始库检查情况 | | | | | | | | | | |
| 检查人 |  | 联系电话 | |  | | | 检查时间 | |  | |
| 报文解释入库是否正常 |  | | | | 处理结果 | | | （6） | | |
| 监测点现场检查情况 | | | | | | | | | | |
| 检查人 |  | 联系电话 | |  | | | 检查时间 | |  | |
| RTU是否正常 |  | | | | 处理结果 | | |  | | |
| 传感器是否正常 |  | | | | 处理结果 | | |  | | |
| 其他相关设备是否正常 |  | | | | 处理结果 | | |  | | |

注：（1）检测开始时间至检测截止时间；（2）填写检测人姓名；（3）填写该监测点出现的异常值，从原始数据库读取，若是数据缺失异常，则置空；（4）该异常值对应记录写入原始数据库的时间，包括年月日时分秒；（5）填写异常类型、出现个数、上次出现时间等信息；（6）若不正常，填写如何处理、是否解决或需要其他支持；若正常，则置空。

表C.3 应用数据库数据质量检测报表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 检测人 |  | | 联系电话 | |  |
| 过程数据质量检查情况 | | | | | |
| 监测点个数 |  | | 检测时间段 | |  |
| 上报数据总量 |  | | 超阈值数据数量 | |  |
| 异常数据数量 |  | | 疑似异常数据数量 | |  |
| 连续无变化数据数量 |  | | 显著周期变化数据数量 | |  |
| 汇总（统计）数据质量检查情况 | | | | | |
| 时段单位 | （1） | | 检测时段 | |  |
| 单日分时汇总异常点数量 |  | | 月内同日汇总数据对比异常 | |  |
| 年内数据异常时段 |  | | 年内数据异常时段内数据汇总值 | |  |
| 数据集质量检查 | | | | | |
| 检测时段 |  | | 检测区域 | |  |
| 及时率 | | 完整率 | | 异常率 | |
| （2） | | （3） | | （4） | |

注：（1）按时、日、月统计，对应时段单位分别为时、日、月；（2）、（3）、（4）涵义见7.1.4

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_