暴雨灾害
Torrential Rain and Disasters
中国科技核心期刊
中国科技论文统计源期刊

一种自动气象站小时降水数据非实时质量控制方法及应用

闫荞荞 王必强 刘园园 严婧 李波 杨宏谦 刘莹

A non-real-time quality control method for hourly precipitation data from automatic weather stations and its application

YAN Qiaoqiao，WANG Biqiang，LIU Yuanyuan，YAN Jing，LI Bo，YANG Hongqian，LIU Ying

在线阅读 View online: https://doi.org/10.12406/byzh.2023-278

您可能感兴趣的其他文章 Articles you may be interested in

三种插值方法对区域自动气象站日气温缺测数据插补的适用性研究

Applicability of three interpolation methods in estimating daily temperature with missing data from regional automatic weather station

暴雨灾害. 2023, 42(1): 97-104 https://doi.org/10.12406/byzh.2022-183

基于改进卷积技术的雷达回波图像质量控制方法研究

Study on radar echo image quality control based on improved convolution technology

暴雨灾害. 2022, 41(5): 598-606 https://doi.org/10.12406/byzh.2022-011

基于小时分辨率的降水过程辨识方法研究

Research on precipitation process identification method based on hourly resolution

暴雨灾害. 2021, 40(1): 12-18 https://doi.org/10.3969/j.issn.1004-9045.2021.01.002

一种基于迁移学习和长短期记忆神经网络的降水预报方法

A precipitation forecast method based on transfer learning and Long Short Term Memory

暴雨灾害. 2024, 43(1): 45-53 https://doi.org/10.12406/byzh.2023-118

一种改进的数值模式降水预报时间降尺度方法

An improved time downscaling method for numerical model precipitation forecast

暴雨灾害. 2024, 43(5): 542-550 https://doi.org/10.12406/byzh.2023-210

毫米波云雷达观测和反演云降水微物理及动力参数方法研究进展

Reviews on retrieval methods for microphysical and dynamic parameters with cloud radar

暴雨灾害. 2021, 40(3): 231-242 https://doi.org/10.3969/j.issn.1004-9045.2021.03.002

关注微信公众号，获得更多资讯信息

闫荞荞,王必强,刘园园,等.2025.一种自动气象站小时降水数据非实时质量控制方法及应用[J].暴雨灾害,44(1):102−111. YAN Qiaoqiao, WANG Biqiang, LIU Yuanyuan, et al. 2025. A non-real-time quality control method for hourly precipitation data from automatic weather stations and its application [J]. Torrential Rain and Disasters,44(1):102−111 (in Chinese). doi:10.12406/byzh.2023−278

一种自动气象站小时降水数据非实时质量控制方法及应用

闫荞荞[1], 王必强[2], 刘园园[1], 严婧[1, 4], 李波[3], 杨宏谦[1], 刘莹[1, 4]

(1. 湖北省气象信息与技术保障中心, 武汉 430074; 2. 湖北省气象服务中心, 武汉 430070;
3. 中国长江电力股份有限公司, 宜昌 443002; 4. 三峡国家气候观象台, 宜昌 443199)

摘　要: 为分类检测隐藏较深的降水质量异常数据, 定义了非实时降水质量异常事件, 基于 Madsen-Allerupt 方法和百分位数法设计了一种自动气象站小时降水数据非实时质量控制方法, 基于该方法对湖北省 2 481 个区域自动气象站 2005—2022 年经实时质量控制的小时降水数据做进一步的非实时质量控制, 制作了小时降水质控数据集, 并初步做了应用评估。结果表明: (1) 设计的方法能够有效实现隐藏较深降水质量异常数据的分类检测, 降水量持续偏小占比为 87.9%, 降水时数异常占比为 8.7%, 降水量持续偏大占比为 3.4%。(2) 质量异常数据的总量逐年趋于减少, 空间分布无明显差异, 时间上 4—10 月质量异常事件总量低于其他月份, 冬季占比最高为 36.7%, 夏季最低为 13.1%。(3) 湖北省降水数据的质量建站初期整体欠佳, 从 2008 年开始稳步提升, 2008—2022 年年数据可用率平均为 95.6%, 年内 4—10 月的数据平均可用率为 97.8%, 高于其它月份的 92.6%。

关键词: 降水; 质量控制; Madsen-Allerupt 方法; 百分位法; 降水数据集

中图法分类号: P413.2　　**文献标志码**: A　　**DOI:** 10.12406/byzh.2023−278

A non-real-time quality control method for hourly precipitation data from automatic weather stations and its application

YAN Qiaoqiao[1], WANG Biqiang[2], LIU Yuanyuan[1], YAN Jing[1, 4], LI Bo[3], YANG Hongqian[1], LIU Ying[1, 4]

(1. *Hubei Meteorological Information and Technological Support Center*, *Wuhan* 430074; 2. *Meteorological Service Center of Hubei Province*, *Wuhan* 430070; 3. *China Yangtze Power Co., Ltd.*, *Yichang* 443002; 4. *National Climate Observatory of Three Gorges*, *Yichang* 443199)

Abstract: In order to detect the deeply hidden precipitation quality anomaly data, the non-real-time precipitation quality anomaly events are defined, and a non-real-time quality control method for hourly precipitation data from automatic weather stations is proposed by combining Madsen-Allerupt and percentile method. The hourly precipitation data from 2 841 regional automatic weather stations in Hubei Province during 2005 to 2022 are further quality-controlled by the proposed method and the hourly precipitation quality-controlled dataset is produced. The results are as follows. (1) The designed method can effectively classify and detect the deeply hidden quality anomaly data. The continuous low precipitation events account for 87.9%, precipitation hours anomaly events account for 8.7% and continuous high precipitation events account for 3.4%. (2) The total amount of quality anomaly data tends to decrease year by year, and there is no obvious differences in the spatial distribution. The total amount of quality anomaly events from April to October is lower than that in other months with the highest proportion of 36.7% in winter and the lowest of 13.1% in summer. (3) The quality of precipitation data in Hubei Province is generally poor at the beginning of the construction of the station. Since 2008, it has been steadily improved. The average annual data availability ratio from 2008 to 2022 is 95.6%, and the average data availability ratio from April to October of the year is 97.8%, which is higher than 92.6% in other months.

Key words: precipitation; quality control; Madsen-Allerupt method; percentile method; precipitation dataset

收稿日期: 2023−12−20; **定稿日期**: 2024−09−08

资助项目: 长江流域开放基金项目 (CJLY2022Y08); 湖北省气象局科技发展基金项目 (2022Z05-01); 武汉城市圈气象联合科技创新基金项目 (WHCSQZ202308); 中国长江电力股份有限公司科研项目 (Z242302005)

第一作者: 闫荞荞, 主要从事气象资料加工处理及质量控制方法研究。E-mail: 18971220528@163.com

通信作者: 王必强, 主要从事应用气象和气象服务研究。E-mail: 164323517@qq.com

© Editorial Office of *Torrential Rain and Disasters*. OA under CC BY−NC−ND 4.0

引　言

气候变化加剧了极端天气和水文事件的发生, 降水是区域干旱和洪水事件最直接的驱动因素, 降水数据在气象、水文和农业等领域均有广泛应用。降水数据可通过自动气象站、雷达、卫星等多种方式获取, 但仅气象站属于直接观测 (WMO, 2008)。自动气象站降水数据作为降水观测的“真值”, 其精度要优于雷达、卫星等遥感观测数据, 对灾害预警 (周伟东等, 2022)、预报检验 (武英娇等, 2019)、区域气候分析 (伍红雨等, 2019) 及防灾减灾服务 (郭帷等, 2022) 等至关重要。

受仪器故障、数据传输、软件错误及观测环境等因素影响, 自动气象站观测降水数据的质量问题客观存在。质量控制是保障观测数据应用质量的重要手段, 既适用于实时态, 也可非实时运作, 例如定期质量控制 (任之花等, 2015)。国内外学者围绕自动站降水数据的质量控制已做了大量工作, 根据使用资料种类可将现有质量控制方法划分为常规质量控制和多源数据质量控制两类。前者主要使用自动站观测资料标注降水质量异常数据, 包括界限值检查、气候极值检查、内部一致性检查、时间一致性检查和空间一致性检查等方法 (任芝花等, 2010; 陈正旭等, 2012; 吴书成等, 2016; Blenkinsop et al., 2017)。后者引入不同观测系统的资料, 通过分析雷达估算降水与自动站降水的差异, 建立关联指标对自动站降水的质量进行综合研判, 提高问题数据的自动识别能力 (丛芳和刘黎平, 2011; 王红艳等, 2015; Qi et al., 2016; 张乐坚等, 2016; 张志强等, 2017)。

现有工作的开展有效提高了降水数据的应用质量, 但由于降水时空分布不均匀, 有时降水质量异常数据的时空分布与真实的降水极值非常相似, 方法设计不当容易错误剔除真实的降水极值, 影响对极端降水的监测 (Schneider et al., 2014; Sha et al., 2021)。因此, 质量控制仍然是自动站观测降水数据加工处理的重难点。近年来, 气象业务的高质量发展对自动站降水数据的质量提出了更高的要求, 譬如网格实况等新型业务对虚假零值降水和小值降水的敏感性越来越高。但现有质量控制方法主要侧重于异常大值降水的检测, 对异常零值降水或小值降水的检测研究较少。Yan 等 (2024) 利用卫星、雷达多源数据设计了仪器空翻及雾、露、霜等导致的 0.1~0.3 mm 虚假小值降水的实时标注, 但未讨论错误零值降水的检测。考虑到小时尺度降水时空异质性更强且降水生消演变过程中小值降水的空间差异更为突出 (Morrison et al., 2019), 实时质量控制对异常零值降水或小值降水的检测存在较大的不确定性, 且该类问题具有一定的隐蔽性, 需要通过分析较长时段的数据才能准确识别, 较难满足实时质控对算法时效的要求。非实时质量控制对方法运行时效要求不高, 更适用于该类问题的检测。

鉴于此, 考虑不同时间尺度降水变化的差异, 本文通过对累计降水量和累计降水时数的统计分析, 综合 Madsen-Allerupt 方法和百分位数法设计了一种自动站小时降水数据非实时质量控制方法, 结合降水质量异常事件定义对隐藏较深的降水质量问题进行分类检测。利用设计的方法对湖北省 2005—2022 年区域自动气象站 (以下简称区域站) 已经实时质量控制的小时降水数据做进一步的非实时质量控制, 研制湖北省区域站小时降水质控数据集, 为区域站数据整编和数据应用提供支撑。

1　资料和方法

1.1　资料说明

本文所用到的资料均来源于湖北省气象资料档案馆, 具体包括: (1) 2005—2022 年湖北省区域站观测的逐小时降水数据; (2) 站点所处位置的经度、纬度及海拔高度等元数据信息。从 2005 年区域站建站开始, 湖北省有降水观测任务的区域站数量逐年上升, 2013 年基本稳定, 截止 2022 年有降水观测区域站共计2 481 站 (图 1)。所使用资料已依托气象资料业务系统 (刘莹等, 2017) 进行了实时质量控制, 利用质量控制码对数据质量进行了标注, 可用数据的质量控制码为 0 或 4, 错误数据为 2, 缺测数据为 8。

为方便数据处理, 将对某站点进行质量控制时使用参考资料的空间范围和时间范围统称为质量控制参考域, 其空间范围取以待评估站点为中心、给定半径长的圆形区域。综合数据加工处理技术规定和业务应用需求, 兼顾站点密度, 通过大量试验分析, 质量控制参考域空间范围的半径取为 50.0 km, 以 20 时 (北京时, 下同) 为日界作为质量控制参考域的时间范围。

1.2　有关定义及方法

1.2.1 非实时质量异常事件定义

为使质量控制的结果能更好地反映异常数据的表现形式和产生原因, 为数据处理和质量问题改进等提供参考, Yan 等 (2024) 通过对大量降水质量异常数据成因、分布特征及应用敏感性的分析定义了降水质量异常事件, 将其划分为实时事件和非实时事件两类: 实时事件对应明显错误的数据, 对短临预

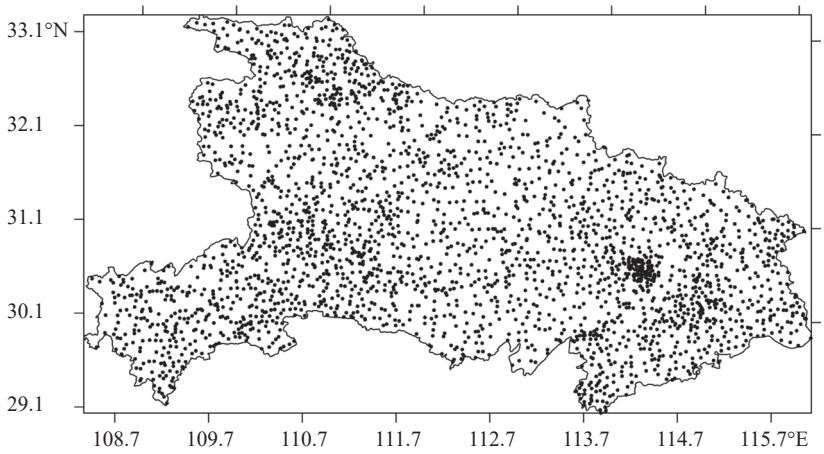


图 1 湖北省降水观测区域自动气象站空间分布图

Fig. 1 The spatial distribution of regional automatic meteorological stations with precipitation observation in Hubei Province

报预警等高时效应用影响较大，通过设计实时质量控制方法利用当前或邻近时次数据进行检测；非实时事件对应隐藏较深的错误数据，通过分析特定时段内数据的变化特征进行检测，该类数据的剔除可为仪器维护、数据整编等对质量要求高于时效的应用提供支撑。

Yan 等 (2024) 给出了实时降水质量异常事件的详细定义，并设计了质量控制方法进行检测，但未对非实时降水质量异常事件进行深入探讨。本文在 Yan 等 (2024) 工作的基础上，通过归类分析将非实时降水质量异常事件细分为降水量持续偏大、降水量持续偏小及降水时数异常 3 种 (表 1)。

表 1 非实时降水质量异常事件的定义及产生原因

Table 1 Definition and causes of non-real-time precipitation quality anomaly events

序号	事件类型	定义	产生原因
1	降水量持续偏大	站点观测降水误差持续高于可接受误差范围的上限值，表现为特定时段内站点累计降水量明显偏大。	仪器故障，观测环境不良，昆虫爬行或浇灌等。
2	降水量持续偏小	站点观测降水误差持续低于可接受误差范围的下限值，表现为特定时段内站点累计降水量明显偏小。	漏斗堵塞、干簧管短路、观测环境不良、降雪或冻结等原因。
3	降水时数异常	站点观测有降水时段与实际降水时段不一致，即降水停止后观测仍为有降水或降水发生过程中观测记录为无降水。	漏斗堵塞、观测环境不良、降雪或冻结等原因。

1.2.2 质量控制方法

为使质量控制方法描述更加清晰，首先定义质量控制因子作为方法设计的基本参量；再基于质量控制因子综合 Madsen-Allerupt 方法 (Madsen, 1992; 1993) 和百分位法 (孔锋等, 2019) 设计质量控制参数；最后通过对质量控制参数的分析计算质量控制指标，并结合质量控制流程设计实现不同质量异常事件的分类检测。

(1) 质量控制因子。考虑到降水时空分布不均匀，结合非实时降水质量异常事件对应数据的分布特征，从降水时间尺度效应出发，遴选累计降水量和累计降水时数作为质量控制因子。对于给定的待质控站点，质量控制因子定义如下

$$R=\sum_{t=1}^{T} r_t \tag{1}$$

$$H=\sum_{t=1}^{T} h_t \tag{2}$$

式中R为累计降水量，H为累计降水时数，T为质量控制参考域内的总小时数，t为时序，r_t为第t小时的降水量，h_t表示第t小时是否有降水，若降水量大于 0，则h_t取值为 1，否则取值为 0，数据缺测当无降水处理。

(2) 质量控制参数。Madsen-Allerupt 方法是经典的降水空间一致性检查方法 (Vejen et al., 2002; 何志军等, 2010; 丛芳和刘黎平, 2011)，百分位法常用于极端降水事件的分析。将一组数据按升序排列并计算相应的累计百分位，某一百分位对应数据的值就称为该百分位的百分位数，即处于 Q% 位置的值就称为第 Q 百分位数，第 95 百分位数常作为极端降水事件的阈值。Madsen-Allerupt 方法利用序列的第 25、50、75 百分位数建立异常检测参数，当异常检测参数

超出给定阈值时就判定为数据异常。考虑因天气气候变化导致的极端降水与降水质量异常数据的相似性, 综合分析质量控制参考域内参考站点累计降水量和累计降水时数序列的百分位数, 设计了非实时降水数据质量控制参数。对于给定的待质控站点, 质量控制参数定义如下

$$F_r = \frac{R-(Q_{r_75}-Q_{r_25})}{Q_{r_75}-Q_{r_25}} \tag{3}$$

$$L_r = \alpha \times Q_{r_5} \tag{4}$$

$$U_r = \beta \times Q_{r_95} \tag{5}$$

$$F_h = \frac{H-(Q_{h_75}-Q_{h_25})}{Q_{h_75}-Q_{h_25}} \tag{6}$$

$$U_h = \beta \times Q_{h_95} \tag{7}$$

式中F_r、L_r、U_r为降水量质控参数, Q_{r_5}、Q_{r_25}、Q_{r_50}、Q_{r_75}和Q_{r_95}为参考站点累计降水量序列的第 5、25、50、75、95 百分位数; F_h、U_h为降水时数质控参数, Q_{h_25}、Q_{h_50}、Q_{h_75}和Q_{h_95}为参考站点累计降水时数序列的第 25、50、75、95 百分位数。其中, F_r、F_h为 Madsen-Allerupt 方法推荐使用的异常检测参数, 为无量纲值, 绝对值越大, 累计降水量、累计降水时数为异常值的概率越大; L_r、U_r、U_h为限定参数, L_r和U_r单位与降水量一致, 分别对异常降水量的下限值和上限值进行约束, U_h的单位与降水时数一致, 对异常降水时数的上限值进行约束; α、β为约束参数, 试验分析取值为 1.5。

(3) 质量控制指标。通过对降水数据质量控制参数统计特征的分析, 结合非实时降水质量异常事件定义设计了降水质量控制指标 (表 2), 实现非实时降水质量异常事件的分类检测。其中, I_r和I_h分别为降水量质控指标和降水时数质控指标。α_r、α_h为质控调节参数, Madsen-Allerupt 方法推荐取值为 3.0。

表 2 降水质量控制指标及计算规则

Table 2 Calculation rules of precipitation quality control index

序号	质控指标	计算规则	含义
1	$I_r=1$	$F_r \geqslant \alpha_r$且$R \geqslant U_r$	降水量持续偏大
2	$I_r=-1$	$F_r < -\alpha_r$且$R < L_r$	降水量持续偏小
3	$I_h=1$	$F_h \geqslant \alpha_h$且$H \geqslant U_h$或 $F_h < -\alpha_h$且$R > 0.0$	降水时数异常

(4) 质量控制流程。当某站降水数据持续异常时, 其累计降水量、累计降水时数的时空分布通常都会有所体现, 这也是数据质量控制方法设计的主要依据, 但也大概率会使质量控制结果不唯一, 譬如当某站因仪器堵塞导致降水量持续偏小时对应的降水时数也可能出现异常。为解决该问题, 综合考虑不同降水质量异常事件的耦合关系及应用敏感性, 设计了小时降水数据非实时质量控制流程 (图 2)。对于给定的待质控站点, 首先检索质量控制参考域内的数据, 接着依次计算质量控制因子、质量控制参数和质量评估指标, 最后按顺序关系对降水量持续偏大、降水量持续偏小和降水时数异常事件进行标注。

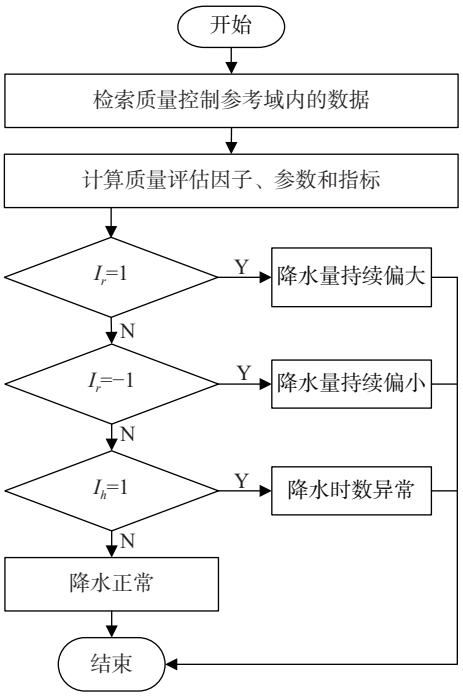


图 2 小时降水数据非实时质量控制流程图 (Y 和 N 为判断语句, 分别表示 yes 和 no)

Fig. 2 Flow chart of the non-real-time quality control for hourly precipitation data (Y and N are judgment statements, representing yes and no respectively)

1.2.3 分析方法

本文以“站月”作为数据量的统计单位, 设计“异常站月比”作为指标对方法检出的错误数据的数量和分布特征进行分析。具体来讲, “1 站月”代表 1 个站点 1 个月的数据, “1 站月数据异常”代表 1 个站点 1 个月的数据未通过非实时质量控制被判定为异常数据, 也对应发生了 1 次降水质量异常事件。“异常站月比”的计算公式如下

$$R_e = \frac{M_e}{M} \times 100\% \tag{8}$$

式中R_e为异常站月比, M_e为异常站月数, M为总站月数。

根据业务规定, 利用数据可用率作为指标对数据的应用质量进行评估, 为降低缺测数据对质量控制准确性的影响, 仅对数据可用率超过 40.0% 的站

点开展非实时质量控制，且计算质量控制参数时仅选取质量控制参考域内数据可用率高于 90.0% 的站点作为参考站。数据可用率计算公式如下

$$R_a = \frac{N_a}{N} \times 100\% \tag{9}$$

式中R_a为数据可用率，N_a为可用数据数，N为应到数据数。

2 应用结果与分析

利用上述质量控制方法对湖北省 2 481 个区域站 2005—2022 年经过实时质量控制的小时降水数据做了进一步的非实时质量控制，将非实时质量控制检出的异常数据的质量控制码更新为 2，制作了湖北省区域站 2005—2022 年的小时降水质控数据集。首先，利用“异常站月比”对非实时质量控制检出的异常数据进行分析，了解异常数据的时空分布特征。其次，为更好地应用质量控制结果，对非实时质量异常事件的分布特征进行分析，并结合典型个例对每类事件的数据表现形式及产生原因做具体阐述，为直观理解和改进降水质量问题提供参考。最后，利用数据可用率对小时降水质控数据集的质量进行评估，提出数据应用建议。

2.1 质量异常数据分析

经统计，区域站数据总量为 398 129 站月，其中参与非实时质量控制的数据总量为 371 074 站月，27 055 站月数据因数据可用率低于 40% 未进行非实时质量控制。结果显示，共 19 305 站月数据未通过非实时质量控制，占非实时质控总数据量的比例为 5.2%。

为了解非实时质量控制检出异常数据的站点分布情况，计算了各站的异常站月比R_e，统计结果显示，非实时质控未检出异常的站点共 326 站，占比为 13.1%；$0 < R_e \leqslant 5.0\%$ 的站点共 1 525 站，占比为 61.5%；$5.0\% < R_e \leqslant 10.0\%$ 的站点共 461 站，占比为 18.6%；$R_e > 10.0\%$ 的站点共 169 站，占比为 6.8%。从中可知，占比 93.2% 的站点异常站月比低于 10.0%，数据质量整体较好。各站点异常站月比空间分布图 (图 3) 显示，不同区间异常站月比的站点全省各地均有出现，异常数据的空间分布差异不明显。

为了解非实时质量控制检出异常数据的时间变化，统计了逐年的异常站月比，结果显示年平均异常站月比为 10.3%。从异常站月比的逐年变化情况 (图 4) 可知，降水数据质量状况整体在逐年向好。其中，2005—2007 年异常站月比较低，经分析这 3 a 参与质量控制的数据总量较之后年份偏少很多，结果缺乏代表性。从 2008 年开始逐步稳定，2008—2013 年异常站月比较高，2014 年开始明显下降，2020—2022 年异常站月比最低，平均为 4.4%。

2.2 质量异常事件分析

2.2.1 特征分析

为了解非实时质量异常事件的分布情况，统计了不同降水质量异常事件对应的异常站月数，结果显示不同质量异常事件的数量差异较大，降水量持续偏小占比最高为 87.9%，降水时数异常占比次之，为 8.7%，降水量持续偏大占比最低，为 3.4%。究其原因，主要是区域站观测环境较为复杂，易出现堵塞导致降水量持续偏小或产生滞后降水导致降水时数异常，而降水量持续偏大多由仪器故障及灌溉等特殊原因引起，出现概率相对较低。为减少事件数量差异对变化趋势分析的影响，对异常站月数作了对数函数转换，给出了降水质量异常事件的月变化 (图 5)，由此可见，4—10 月异常事件数量整体低于其

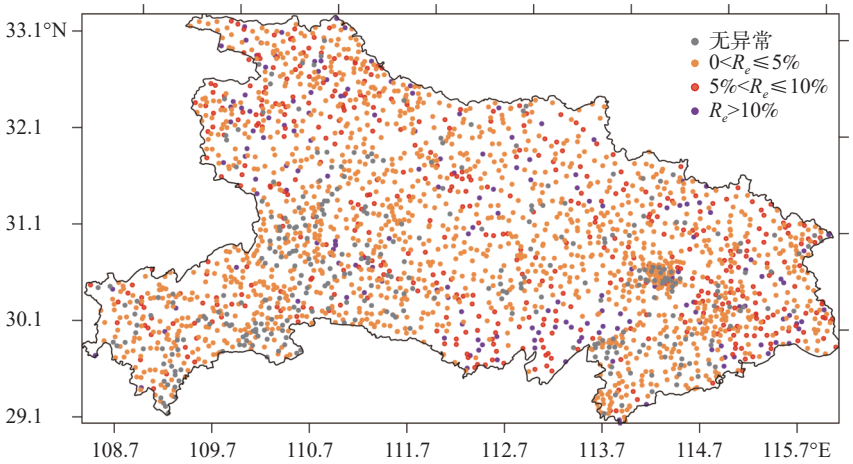


图 3 湖北省区域站各站点异常站月比空间分布图

Fig. 3 The spatial distribution of anomaly station-month ratio for regional stations in Hubei Province

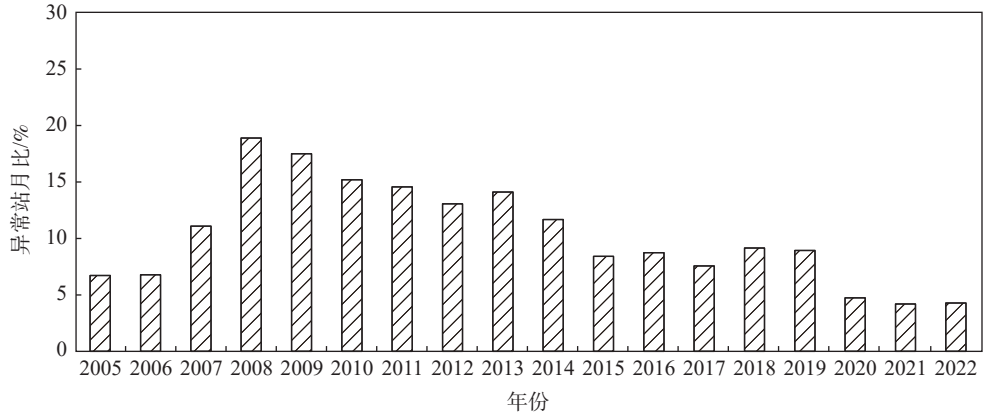


图 4 2005—2022 年湖北省区域站降水数据异常站月比年变化

Fig. 4 Annual variation of the anomaly station-month ratio of precipitation data of regional stations from 2005 to 2022 in Hubei Province

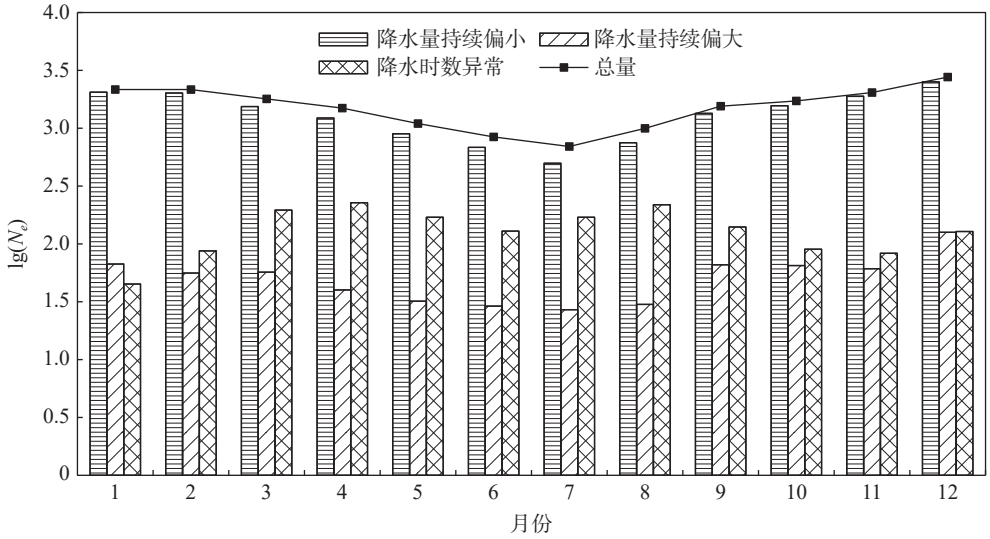


图 5 湖北省区域站降水质量异常事件月变化

Fig. 5 Monthly variation of precipitation quality anomaly events of regional stations in Hubei Province

他月份, 降水时数异常事件 3—9 月占比略高于其他月份。按季节进一步分析显示, 冬季异常事件占比最高为 36.7%, 夏季占比最低为 13.1%。

2.2.2 个例分析

下面挑选部分降水质量异常事件个例进行分析。

(1) 降水量持续偏大。芳畈站位于孝感大悟罗汉坡水库管理处, 海拔高度 96.8 m。2017 年 10 月 (9 月 30 日 21 时—10 月 31 日 20 时) 芳畈站的累计降水量为 450.9 mm, 累计降水时数为 152 h, 有效参考站共 102 站, 降水量质控指标为 1, 降水时数质控指标为 0, 按照质控流程标注为降水量持续偏大。图 6 为芳畈站与邻近参考站大悟、阳平的逐小时降水时序图, 其中大悟站海拔高度 120.0 m, 距芳畈站 17.5 km, 阳平站海拔高度 46.6 m, 距芳畈站 7.6 km。月内 4 次降水过程芳畈站的降水时间与邻近站一致但降水量均明显偏大, 结合历史数据处理记录及调查分析确定为观测环境不良所致。该站附近绿化树木长时间未修剪影响到雨量观测, 雨水经树木导引落入雨量筒使观测降水比实际降水偏大。

(2) 降水量持续偏小。根据湖北省气候影响评价报告, 2020 年 7 月 (6 月 30 日 21 时—7 月 31 日 20 时) 湖北省降水异常偏多, 全省平均降水量为 437.1 mm, 较常年同期偏多 1.1 倍, 排 1961 年以来同期第 1 位, 月内强降水过程频繁, 多站日降水、持续降水突破历史极值, 其中恩施中部、江汉平原东部、鄂东大部 500.0~849.2 mm, 其他地区 136.8~500.0 mm。当月评估结果显示共 14 站降水量持续偏小。以腾龙洞外站为例, 其位于恩施利川腾龙洞停车场山坡, 海拔高度 1 070.0 m, 月累计降水量为 1.3 mm, 累计降水时数为 4 h, 有效参考站共 88 站, 降水量评估指标为 −1, 降水时数评估指标为 1, 根据质控流程标注为降水量持续偏小。图 7 为该站与其邻近参考站累计降水量和累计降水时数的空间分布图, 可见累计降水量和累计降水时数均明显偏小, 经调查发现为雨量

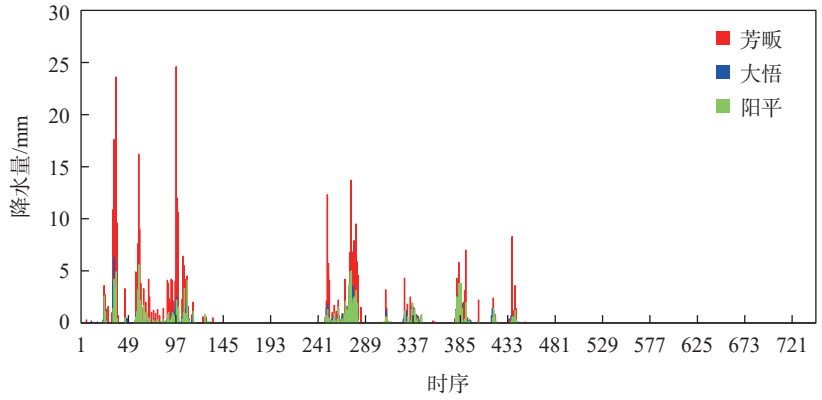


图 6 2017 年 9 月 30 日 21 时—10 月 31 日 20 时芳畈站与其邻近参考站大悟、阳平的逐小时降水量柱状图

Fig. 6 The histogram of hourly precipitation from Fangfan and its neighbors Dawu and Yangping station from 21:00 BT on 30 September to 20:00 BT on 31 October, 2017

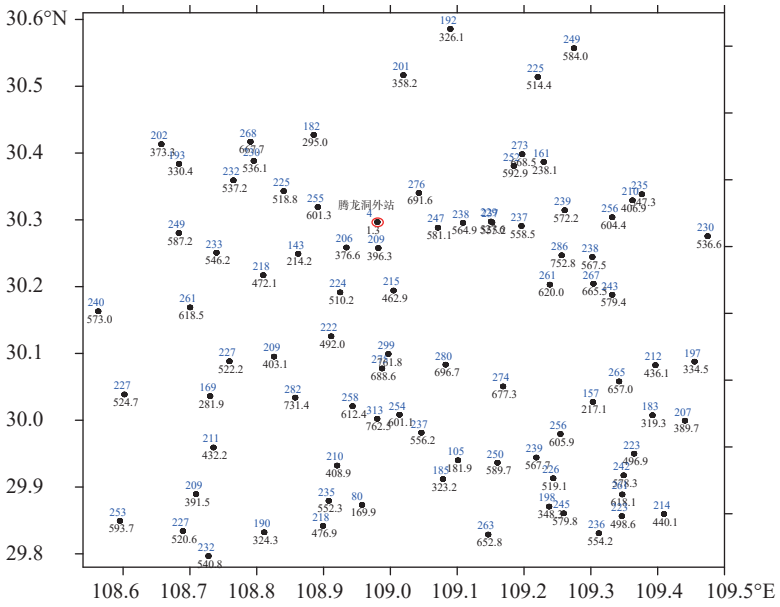


图 7 2020 年 7 月腾龙洞外站与其邻近参考站累计降水量 (黑色数字, 单位: mm) 和降水时数 (蓝色数字, 单位: h) 空间分布图

Fig. 7 The spatial distribution of accumulated precipitation (black number, unit: mm) and precipitation hours (blue number, unit: h) of Tenglongdongwai station and its neighbor stations in July 2020

筒严重堵塞所致。

(3) 降水时数异常。尺八站位于荆州监利尺八镇人民政府, 海拔高度 28.9 m。2014 年 7 月 (6 月 30 日 21 时—7 月 31 日 20 时) 尺八站的累计降水量为 106.5 mm, 累计降水时数为 453 h, 有效参考站共 38 站, 降水量质控指标为 0, 降水时数质控指标为 2, 根据质量控制流程标注为降水时数异常。图 8 为尺八站与邻近参考站监利、三洲的逐小时降水时序图, 其中监利站海拔高度 26.2 m, 距尺八站 33.1 km, 三洲站海拔 29.0 m, 距尺八站 8.8 km。月内 3 次降水过程尺八站的降水量均明显偏低且在过程结束后持续出现了0.1~0.3 mm 的小值降水, 经统计位于该区间的降水时数占其全月累计降水时数的比例为 83.0%。结合当月降水实况分析, 经调查发现与腾龙洞外站类似, 为雨量筒堵塞所致, 但该站堵塞程度较轻, 降水结束后雨水慢慢渗出导致出现虚假滞后降水。

2.3 数据集质量评估

为了解湖北省 2005—2022 年区域站小时降水质控数据集的质量, 统计了数据的年可用率和月平均可用率。从年可用率变化 (图 9) 可知, 2005—2007 年数据可用率明显偏低, 主要是建站初期数据的传输、存储及加工处理等环节尚不成熟, 数据完整性较差导致。2008 年开始随着各业务环节的不断优化完善, 数据质量出现了明显提升, 2008—2022 年数据年可用率平均为 95.6%。从 2008—2022 年月平均可用率 (图 10) 可知, 4—10 月数据可用率高于其他月份, 4—10 月平均为 97.8%, 其他月份平均为 92.6%。湖北省地处南北气候过渡带, 冬季易出现雪或雨夹雪天气, 湖北省区域站使用的翻斗雨量传感器无法准确测量固态降水, 降水过程中会出现虚假零值降水

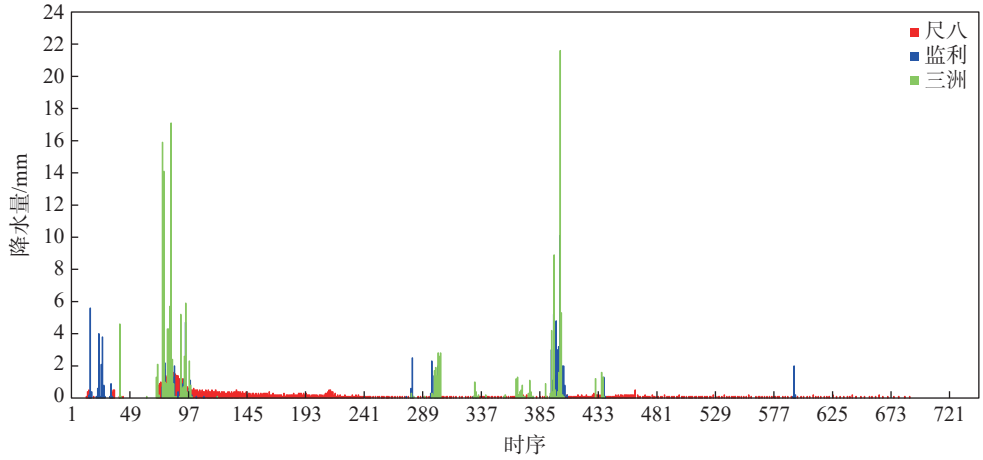


图 8 2014 年 6 月 30 日 21 时—7 月 31 日 20 时尺八站与其邻近参考站监利、三洲的逐小时降水量柱状图

Fig. 8 The histogram of hourly precipitation from Chiba and its neighbors Jianli and Sanzhou station from 21:00 BT on 30 June to 20:00 BT on 31 July, 2014

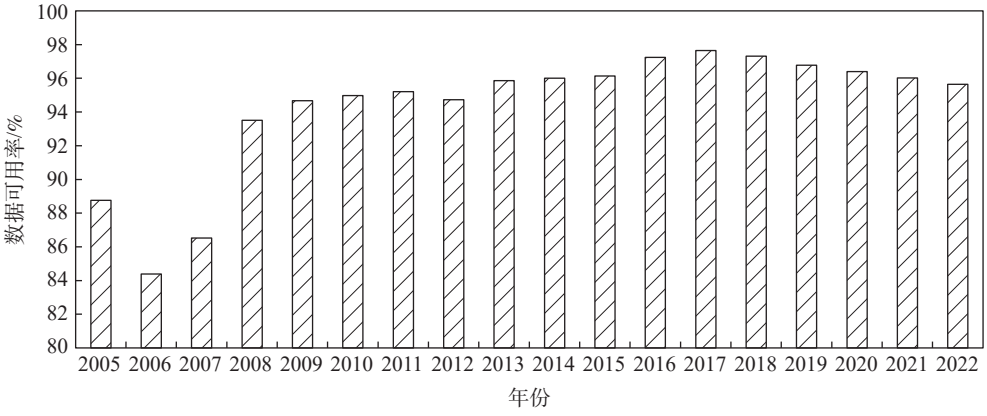


图 9 2005—2022 年湖北省区域站小时降水数据年可用率

Fig. 9 Annual availability of hourly precipitation data of regional stations from 2005 to 2022 in Hubei Province

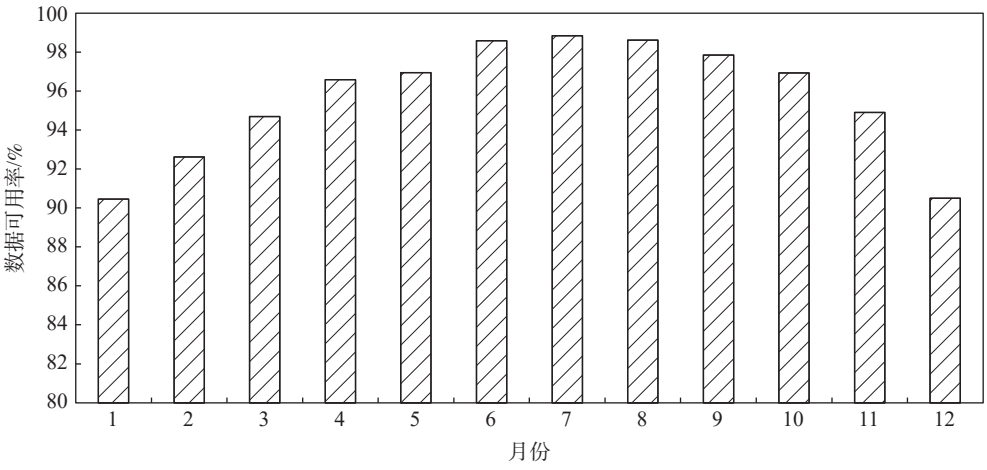


图 10 2008—2022 年湖北省区域站小时降水数据月平均可用率

Fig. 10 Monthly average availability of hourly precipitation data of regional stations from 2008 to 2022 in Hubei Province

或小值降水, 过程结束气温回升又会出现大量融雪性滞后降水, 使得冬季降水数据质量较差。非汛期区域站维护频次低也是冬季数据可用率偏低的原因之一。鉴于此, 建议慎用 4—10 月以外其他月份湖北省区域站的降水观测数据, 若必须使用应结合天气背景和应用场景综合研判。

3 结论与讨论

本文定义了非实时降水质量异常事件, 设计了一种自动气象站小时降水数据非实时质量控制方

法，实现了隐藏较深降水质量问题的分类检测。利用设计的方法对湖北省 2 481 个区域站 2005—2022 年的小时降水数据做了非实时质量控制，制作了湖北省区域站质控降水数据集，并对数据集的质量进行了评估，得到如下主要结论：

(1) 设计的方法能够有效检测出隐藏较深的降水数据质量问题。基于非实时降水质量异常事件的定义，对检出的异常数据进行了分类，探讨了引起各类事件发生的可能原因，提高了质量控制结果的参考价值，可为质量问题改进等提供指导。

(2) 各站点降水数据质量整体较好，异常数据空间分布差异不明显。非实时质控未检出异常的站点占比为 13.1%，站点异常站月比低于 5.0% 的站点占比为 61.5%，位于 5.0%~10.0% 占比为 18.6%，高于 10.0% 的占比为 6.8%，占比 93.2% 的站点异常站月比低于 10.0%。

(3) 不同类型降水质量异常事件的占比差异较大，降水量持续偏小占比为 87.9%，降水时数异常占比为 8.7%，降水量持续偏大占比为 3.4%。4—10 月异常事件总量低于其他月份，异常站月比冬季最高为 36.7%，夏季最低为 13.1%。

(4) 降水数据的应用质量从 2008 年开始稳步提升，2008—2022 年数据年可用率平均为 95.6%，年内 4—10 月平均可用率为 97.8%，高于其它月份的 92.6%。综合考虑湖北省的天气气候特征及冬季降水观测实际情况，建议慎用 4—10 月以外其他月份的数据。

本文设计的方法适用于非实时质量控制，聚焦隐藏较深降水质量异常数据的检测，为充分保障降水数据的质量，建议结合实时质量控制方法综合使用。需要说明的是，质量控制参考域相关参数的设定与站点密度密切相关，要结合具体应用场景确定参数取值。此外，方法对质量异常数据的识别主要以邻近参考站数据的分布特征为参考，在站点分布稀疏区的检测能力会受影响，可尝试将雷达估测降水产品等引入降水非实时质量控制方法的研制，提高方法的鲁棒性。

参考文献 (References)：

陈正旭,王俊骄,洪月英.2012.一种基于“莱因达”准则的区域自动站实时资料应用方法[J].江汉大学学报 (自然科学版),40(3):33−37. Chen Z X, Wang J J, Hong Y Y. 2012. An applied method based on the "PauTa" rule for real-time data of regional automatic station [J]. Journal of Jianghan University (Natural Science Edition),40(3):33−37 (in Chinese). doi:10.3969/j.issn.1673−0143.2012.03.009

丛芳,刘黎平.2011.新一代天气雷达与地面雨量资料的综合分析[J].气象,37(5):532−539. Cong F, Liu L P. 2011. A comprehensive analysis of data from the CINRAD and the ground rainfall station [J]. Meteorological Monthly,37(5):532−539 (in Chinese). doi:10.7519/j.issn.1000−0526.2011.05.003

郭帷,谢海涛,王旭,等.2022. 1981-2019 年北疆暴雨洪涝灾害的时空变化特征分析[J].暴雨灾害,41(1):101−108. Guo W, Xie H T, Wang X, et al. 2022. Spatio-temporal variation of rainstorm flooding disaster events in Northern Xinjiang from 1981 to 2019 [J]. Torrential Rain and Disasters,41(1):101−108 (in Chinese). doi:10.3969/j.issn.1004−9045.2022.01.012

何志军,封秀燕,何利德,等.2010.气象观测资料的四方位空间一致性检验[J].气象,36(5):118−122. He Z J, Feng X Y, He L D, et al. 2010. Examination of meteorological data by the horizontal space consistency analysis from four directions [J]. Meteorological Monthly, 36(5):118−122 (in Chinese). doi:10.7519/j.issn.1000−0526.2010.5.018

孔锋,乔枫雪,王瑞.2019.基于百分位数法的中国不同强度小时降雨频次空间分异格局及其变化模态研究 (1961-2013)[J].水利水电技术,50(6):37−48. Kong F, Qiao F X, Wang R. 2019. Percentile method-based study on spatial differentiation patterns of hourly rainfall frequency with different intensities and their changing modes in China (1961 to 2013) [J]. Water Resources and Hydropower Engineering,50(6):37−48 (in Chinese). doi:10.13928/j.cnki.wrahe.2019.06.005

刘莹,刘雯,闫荞荞,等.2017.气象资料业务系统 (MDOS) 使用手册[M].北京:气象出版社:1-137. Liu Y, Liu W, Yan Q Q, et al. 2017. Meteorological data operational system (MDOS) user manual [M]. Beijing: Meteorological Press: 1-137 (in Chinese)

任芝花,赵平,张强,等.2010.适用于全国自动站小时降水资料的质量控制方法[J].气象,36(7):123−132. Ren Z H, Zhao P, Zhang Q, et al. 2010. Quality control procedures for gourly precipitation data from automatic weather stations in China [J]. Meteorological Monthly, 36(7):123−132 (in Chinese). doi:10.7519/j.issn.1000−0526.2010.7.019

任芝花,张志富,孙超,等.2015.全国自动气象站实时观测资料三级质量控制系统研制[J].气象,41(10):1268−1277. Ren Z H, Zhang Z F, Sun C, et al. 2015. Development of three-step quality control system of real-time observation data from AWS in China [J]. Meteorological Monthly,41(10):1268−1277 (in Chinese). doi:10.7519/j.issn.1000−0526.2015.10.010

王红艳,王改利,刘黎平,等.2015.利用雷达资料对自动雨量计实时质量控制的方法研究[J].大气科学,39(1):59−67. Wang H Y, Wang G L, Liu L P, et al. 2015. Development of a real-time quality control method for automatic rain gauge data using radar quantitative precipitation estimation [J]. Chinese Journal of Atmospheric Sciences,39(1):59−67 (in Chinese). doi:10.3878/j.issn.1006−9895.1403.13295

吴书成,魏爽,王丽吉,等.2016.自动气象站降水实时快速质控算法设计[J].气象科技,44(6):882−888. Wu S C, Wei S, Wang L J, et al. 2016.

Algorithm design of fast and real-time quality control for precipitation data from automatic stations [J]. Meteorological Science and Technology,44(6):882−888 (in Chinese). doi:10.3969/j.issn.1671−6345.2016.06.005

伍红雨,翟志宏,张羽.2019. 1961-2018 年粤港澳大湾区气候变化分析[J].暴雨灾害,38(4):303−310. Wu H Y, Zhai Z H, Zhang Y. 2019. Analysis of climate change in Guangdong-Hong Kong-Macao Greater Bay Area from 1961 to 2018 [J]. Torrential Rain and Disasters,38(4):303−310 (in Chinese). doi:10.3969/j.issn.1004−9045.2019.04.002

武英娇,杨浩,钱仙桃,等.2019. ECMWF 集合预报在安徽大别山区降水预报中的检验[J].暴雨灾害,38(1):66−71 . Wu Y J, Yang H, Qian X T, et al. 2019. Verification of ECMWF ensemble prediction in precipitation forecast of Ta-pieh mountains area in Anhui province [J]. Torrential Rain and Disasters,38(1):66−71 (in Chinese). doi:10.3969/j.issn.1004−9045.2019.01.008

张乐坚,俞小鼎,李峰,等.2016.地面降水的多源数据辅助质量控制方法[J].气 象 ,42(3):363−371. Zhang L J, Yu X D, Li F, et al. 2016. Quality control method for multi-source data of surface rainfall [J]. Meteorological Monthly,42(3):363−371 (in Chinese). doi:10.7519/j.issn.1000−0526.2016.3.012

张志强,仲凌志,杨和平.2017.天气雷达在中国自动气象站实时质量控制系统中的应用[J].计算机应用,37(S2):298−300+326. Zhang Z Q, Zhong L Z, Yang H P. 2017. Application of weather radar in real-time quality control system of hourly guage precipitation in China [J]. Journal of Computer Applications,37(S2):298−300+326. (in Chinese)

周伟东,韩宁,戴建华,等.2022.上海地区暴雨预警评估及不同天气型暴雨预警差异分析[J].暴雨灾害,41(4):405−412. Zhou W D, Han N, Dai J H, et al. 2022. Investigation of rainstorm warning and its difference under different synoptic conditions in Shanghai [J]. Torrential Rain and Disasters,41(4):405−412 (in Chinese). doi:10.3969/j.issn.1004−9045.2022.04.005

Blenkinsop S, Lewis E, Chan S C, et al. 2017. Quality-control of an hourly rainfall dataset and climatology of extremes for the UK [J]. International Journal of Climatology,37(2):722−740. doi:10.1002/joc.4735

Madsen H. 1992. Semi-automatic quality control of daily precipitation measurements [C].Proceedings of 5th International Meeting on Statistical Climatology, Toronto, Canada, 22-26 June 1992

Madsen H. 1993. Algorithms for correction of error types in a semi-automatic data collection. Precipitation Measurements & Quality Control [C].Proceedings of Symposium on Precipitation and Evaporation, Bratislava, Slovakia, 20-24 September 1993

Morrison A, Villarini G, Zhang W, et al. 2019. Projected changes in extreme precipitation at sub-daily and daily time scales [J]. Global and Planetary Change: 103004. doi:10.1016/j.gloplacha.2019.103004

Qi Y, Martinaitis S, Zhang J, et al. 2016. A real-time automated quality control of hourly rain gauge data based on multiple sensors in MRMS system [J]. Journal of Hydrometeorology,17(6):1675−1691. doi:10.1175/JHM−D−15−0188.1

Schneider U, Becker A, Finger P, et al. 2014. GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle [J]. Theoretical and Applied Climatology,115:15−40. doi:10.1007/s00704−013−0860−x

Sha Y, Gagne D J, West G, et al. 2021. Deep-learning-based precipitation observation quality control [J]. Journal of Atmospheric and Oceanic Technology,38(5):1075−1091. doi:10.1175/JTECH−D−20−0081.1

Vejen F, Jacobsson C, Fredriksson U, et al. 2002. Quality control of meteorological observations automatic methods used in the Nordic countries [R]. Climate Report, No. 8/2002 KLIMA

WMO. 2008. Guide to instruments and methods of observation [R]. WMO-No. 8, WMO

Yan Q Q, Zhang B S, Jiang Y, et al. 2024. Quality control of hourly rain gauge data based on radar and satellite multi-source data [J]. Journal of Hydroinformatics,26(5):1042−1058. doi:10.2166/hydro.2024.272

(责任编辑　杨　浩)