

文章编号: 1674-7054(2022)04-0367-09



海南省自动土壤水分站数据质量控制方法研究

施晨晓, 羊清雯, 陈珍莉, 刘霄燕, 王小洁

(海南省气象信息中心/海南省南海气象防灾减灾重点实验室, 海口 570203)

摘要: 为了提高土壤水分观测数据的准确性, 利用 2013—2018 年海南省自动土壤水分资料, 根据土壤水分自动站的观测原理及观测数据的特点并结合海南省气候特征, 统计分析得到不同土层不同土壤要素的阈值, 提出一种适用于小时土壤水分观测资料的可行质量控制方案。结果表明, 将质量控制方案应用到 2019 年全省自动站土壤水分观测小时资料中, 海南省土壤水分自动观测数据质控方案具有很好的适用性; 全年各月可用率均在 90% 以上; 除 9 月外, 平均可用率均在 97% 以上; 与现行的 RASM 审核软件相比, 降低了错检率和漏检率。

关键词: 海南省; 自动土壤水分站; 实时数据; S 文件; 质量控制; 质控评估

中图分类号: P 413 **文献标志码:** A

引用格式: 施晨晓, 羊清雯, 陈珍莉, 等. 海南省自动土壤水分站数据质量控制方法研究 [J]. 热带生物学报, 2022, 13(4): 367-375. DOI: [10.15886/j.cnki.rds wxb.2022.04.007](https://doi.org/10.15886/j.cnki.rds wxb.2022.04.007)

土壤水分是陆气相互作用中重要的一环, 它的变化不仅会影响地表水循环, 而且对地表蒸发也会产生一定的影响。它可以直接影响陆气间的水分与能量的交换, 同时与地表反照率存在正负反馈的关系^[1], 认识土壤水分对理解和预测地表温度、降水、旱涝以及未来气候变化的影响起到至关重要的作用^[2]。此外, 土壤水分是影响农业生产的重要因子之一, 是监测农业干旱的重要依据之一, 对农业旱灾监测预警具有重要意义^[3-4]。因此, 土壤水分观测数据的准确性显得尤为重要。国内外对于加强土壤水分数据的质控, 提高数据的准确性方面开展了很多研究与讨论。YILMAZ 等^[5]利用土耳其自 2007 年起, 共 149 个观测站点的每 10 分钟观测 1 次的土壤水分数据, 从时间序列上讨论数据的连续性和可靠性, 并结合气温数据辅助修正给出“可靠”与“不可靠”的质控标准。LIAO 等^[6]发现用于 NASMD(北美土壤水分数据集)的土壤水分数据质控程序在冻土等情况下, 不

能很好地判别数据的准确性, 因此, 基于 78 个土壤水分观测站与已进行质控的土壤水分历史数据对现有的质控程序改进。改进后的质控程序可以标记出虚假和错误的土壤水分观测数据, 特别是高海拔和高纬度的土壤水分观测站点数据。DORIGO 等^[7]利用全球土壤水分观测站数据分析发现, 土壤水分观测仪器不同以及受外界天气影响不同会造成各站点间的数据质量有很大的差异, 因此需建立一套合适的土壤水分数据质量控制方法对其进行统一的数据控制。2013 年起, 中国气象局为了适应和满足土壤水分数据的时效和准确需求, 取消土壤水分人工观测, 在全国范围内正式开展自动土壤水分观测^[8-13]。张志富^[14]利用全国人工观测和自动观测土壤水分历史资料, 统计分析了不同土层、不同土壤水分要素值的分布特点, 根据土壤水分自动观测原理, 提出适用于自动站土壤水分小时数据的质控方案。郭旭等^[15]利用人工与自动土壤水分观测资料, 分析统计不同

收稿日期: 2021-09-13

修回日期: 2022-05-07

基金项目: 海南省气象局青年基金项目(HNQXQN201801、HNQXQN201802)

第一作者: 施晨晓(1989-), 女, 硕士, 工程师。研究方向: 气象数据资料质量控制及应用研究。E-mail: hsmyhaha@163.com

通信作者: 羊清雯(1977-), 女, 本科, 工程师。研究方向: 气象数据资料质量控制及应用研究。E-mail: 402835097@qq.com

土层,不同土壤水分要素的气候极值和时变阈值,提出适用于四川地区自动土壤水分数据的质控方法。胡伟等^[16]利用数理统计学方法,统计分析不同土层,不同土壤水分要素的历年分布和极值情况,得到江西省不同土层,不同土壤水分要素的阈值范围,提出适用于江西省自动土壤水分数据的质控方法。

目前,海南省业务上还没有一套针对自动土壤水分观测数据的有效质量控制技术方案。主要问题在于土壤水分自动观测资料中的各种错误,难以实时自纠,且验证参考不易衡量,缺少必要的依据。因此,海南省土壤水分资料实际应用较少。基于这种现状,根据土壤水分自动站的观测原理及观测数据的特点并结合本省气候特征,提出一种适用于小时土壤水分观测资料的可行质量控制方案,试图寻找土壤水分资料的质控应用,以提高资料在服务农林、监测生态环境等方面的经济和社会效益。

1 土壤水分资料

自动观测土壤水分主要观测以下4种要素:土壤体积含水量、土壤重量含水率、土壤相对湿度、土壤有效贮水量。我省对土壤水分的观测土层分为5层,分别为0~10 cm、10~20 cm、20~30 cm、30~40 cm、40~50 cm。

海南省自2010年起开始安装建设自动观测土壤水分站,除海岛台站外,全省18个市县已完成建设安装,2012—2013年陆续投入业务运行,2013年起开始对自动观测土壤水分数据进行审核。全省18个站的具体站点分布情况见图1。

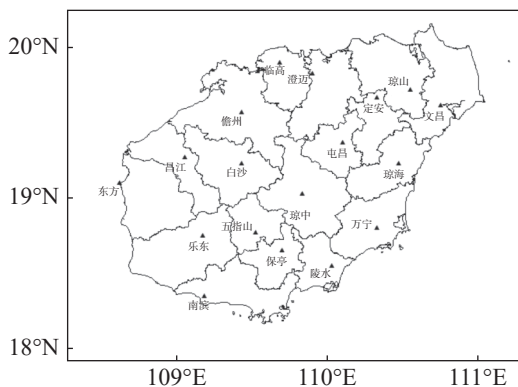


图1 海南岛自动土壤水分观测站点分布图

2 质量控制方法设计

为建立适用于海南省自动土壤水分观测资料

的质量控制方法,本研究以2013—2018年海南土壤水分数据为基础,首先,结合海南省土壤水分分布特点^[17]及本地气候条件,分析出质控要素的整体数据特点,并参考观测原理初步确定质控步骤;其次,对质控要素的历史资料分不同土层、不同土壤水分要素,同时结合不同天气特点进行统计分析,得到相应的阈值范围;再者,对初步设定的阈值范围进行质量评估,将阈值范围及质控步骤进行适当的调整,使设定的阈值范围及质控步骤更加合理;最后,拟形成一套完整且符合海南省的土壤水分质控方案。

通过分析海南省自动站土壤水分小时数据发现,自动观测土壤水分出现异常错误数据情况大致有以下几种:某一时次数据跳变;仪器不稳定或接触不良造成观测数据时常出现波动;土壤水文物理常数错误导致土壤水分数据异常;因各种客观原因造成数据大量缺测;由于土壤类型、土壤质地等不同,即使在同一地区,土壤水分也可能存在较大差异^[18]。根据以上数据特点,设定以下4种质量控制过程:数据缺测处理;土壤体积含水量界限值检查;各土壤水分要素气候极值检查;时间一致性检查。

2.1 土壤水分数据质量控制码的规定 在质控设计方案过程中,参考任之花等^[19]、张志富^[14]研制的全国自动站小时降水资料、全国自动站土壤水分数据的质控方法,对海南省自动观测土壤水分数据进行质控的进程中,需要不断对待检数据设置或是修改质量控制(QC)码。QC码0、1、2、8、9的含义分别为数据正确、数据可疑、数据错误、无观测数据/缺测数据、未作质量控制^[19]。

2.2 质控方法及质控码的设置 首先给每个即将进行质控的土壤水分数据分配一个初始质控码9(QC码=9),按照下列质控步骤进行检查。每一步的检查对象均是QC码为9的土壤水分数据,是否修改QC码则是根据检查结果来决定。

2.2.1 数据缺测检查 数据的缺测情况是反映观测数据完整性的重要指标之一,数据缺测量越少说明数据的完整性越高。在对2013—2018年S文件和实时土壤水分观测资料进行统计分析发现,海南省存在大量数据缺测的现象,是所有疑误数据类型中不可忽视的问题。因此,统计数据缺测情况是有必要的。主要方法是逐小时对所有土层的土壤体积含水量、土壤重量含水率、土壤相对

湿度、土壤有效贮水量进行检查,当遇到缺测符号“-”,直接判为某时次某土层某要素缺测, QC 码=8,之后的检查也无需再做。

表1为2013—2018年全省土壤水分数据缺测的情况。从表1可看出,年缺测总数在逐年增加,其中,2018年全省土壤水分观测数据年缺测数达26860个,为6a最多。1年中缺测数最多台站的缺测数最多月的缺测率为10%~32%。由此看出,土壤水分观测数据完整性较差,但造成完整性较差的原因是多方面的^[18]。

表1 2013—2018年海南省土壤水分观测数据缺测情况

年份	年缺测总数/个	缺测数最多台站	缺测数最多月	缺测数/个	缺测率/%
2013	10480	儋州站(59845)	1月	2400	23
2014	12219	儋州站(59845)	12月	1540	13
2015	12492	琼山站(59757)	1月	1280	10
2016	14396	儋州站(59845)	10月	1780	12
2017	18831	定安站(59851)	10月	6060	32
2018	26860	澄迈站(59843)	6月	3880	14

2.2.2 界限值检查 界限值检查主要的质控对象是自动土壤水分观测仪器直接观测的土壤体积含水量,即检查该土壤水分要素是否超出允许值范围,即检查是否存在土壤水分特征上、理论上不可能出现的异常值。

在对2013—2018年小时土壤体积含水量进行统计分析发现,海南省大部分台站各层土壤体积分含水量的上限值可达60%以上,下限值可至5%以下,结合DORIGO等^[7]、吴东丽等^[20]、XIA等^[21]的研究,土壤体积含水量一般不能大于60%,因此,将各土层的土壤体积分含水量的界限值范围设为0%~60%。当某一土层的土壤体积分含水量超过这一界限范围,判定该土层的土壤体积分含水量错误(QC码=2)。图2为琼山站2016年40~50cm土壤体积分含水量的界限值检查结果,可以看出,该站在10月18日04:00和06:00出现全年土壤体积分含水量的最大值69.1 g·cm⁻³(图2红线),数据存在超出界限值。经核实,该日受到1621号台风“莎莉嘉”的影响,出现较大的强降水导致仪器周边积水,需等雨水下渗后回落后,值的变化趋于平稳才算正常。经界限值检查,数据处于较合理的范围内,受降水影响、超出界限的异常值则剔除

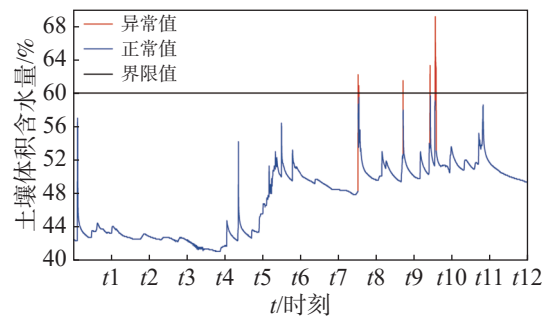


图2 2016年琼山站(59757)40~50cm土壤体积分含水量界限值检查个案

t1: 2016-01-31 7:00; t2: 2016-03-01 19:00; t3: 2016-04-01 7:00; t4: 2016-05-01 19:00; t5: 2016-06-01 7:00; t6: 2016-7-00-01 19:00; t7: 2016-08-01 7:00; t8: 2016-08-31 19:00; t9: 2016-10-01 7:00; t10: 2016-10-31 19:00; t11: 2016-12-01 7:00; t12: 2016-12-31 19:00

(图2蓝线)。

在该项检查中,有2点需要说明,一是当判定某土层的土壤体积分含水量为错误数据时,由于其他土壤水分要素值均是由土壤体积分含水量为基础计算得来,因此,其相应的其他土壤水分要素值均不做其他的后续检查。二是如果土壤水分物理常数本身错误,那么将直接影响上限和下限值的统计判断,从而影响该项检查的误判。因此,当某台站的土壤体积分含水量在经过此项检查后,大量数据被判定为错误数据,那么应首先确认土壤水文物理常数是否正确。

2.2.3 气候极值检查范围的确定 自动土壤水分数据中的气候极值,指的是在气候或理论的范围,某一固定的土壤水分观测站的某土壤水分要素值在历史上出现过的最大或最小值,这类似于地面质控中对气候学界限值的定义,即在气候学的角度上,不可能出现的临界值^[21-22]。由于考虑到实时数据资料的时效性,该检查采取的方法是针对海南省的气候、各土层土壤水分分布等特点,合理选择极值的上下界限范围,如果范围太大,则会导致异常偏大或异常偏小的值无法检查出来;如果范围太小,则会造成大量合理的数据被误检,无法达到质控程序能够合理和便捷检查数据的目的。

挑取和确定气候极值的数据集主要是2013—2018年的S文件和实时数据。由于万宁站和陵水站数据从2016年1月起才有完整的数据序列,因此这2个站的数据集时间是2016—2018年。挑

取和确定的具体方法步骤^[23-24]:

(1)根据上一项界限值检查,剔除不在界限范围内的土壤体积含水量,利用公式(1)对其他要素值进行计算^[25],并大致挑取各土层各要素气候极值范围(图3):

土壤质量含水率 w (%)可以用土壤体积含水量与土壤容重的比值来表示。

$$w = \frac{Q}{\rho} \quad (1)$$

其中, Q 表示土壤体积含水量(%), ρ 表示土壤容重($\text{g}\cdot\text{cm}^{-3}$)。

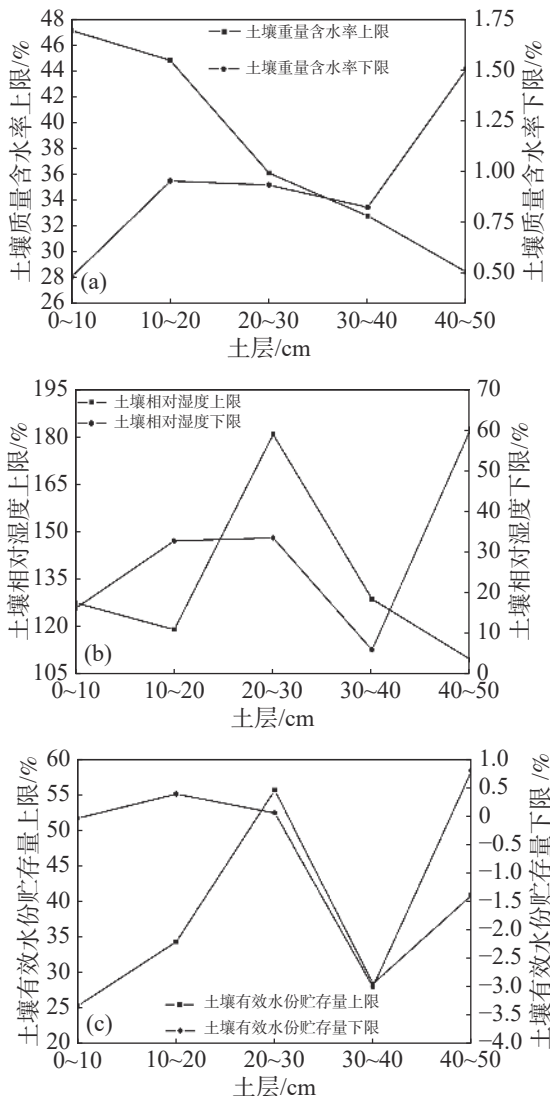


图3 土壤质量含水率(a)、土壤相对湿度(b)和土壤有效贮水量(c)不同土层的极值分布大致范围

土壤相对湿度 R (%)可以用质量含水率占田间持水量的百分比来表示,计算后取整数。

$$R = \frac{w}{f_c} \times 100\% \quad (2)$$

其中, f_c 表示田间持水量(% ,用质量含水率表示)。

土壤有效贮水量 u (mm)是指土壤中含有的大于凋萎湿度的水分贮存量。

$$u = \rho \times h \times (w - w_k) \times 10 \quad (3)$$

其中, h 表示土层间的厚度, w_k 表示凋萎湿度(% ,用质量含水率表示)。

(2)由于历史极值可能存在质量问题,直接利用现有的气候极值设定的范围,可能存在一定的误差、误检或漏检问题。因此为了获得较准确的气候极值的上下界限范围,应先求得各站各土层各土壤水分要素的平均值 \bar{x} 和标准差 σ ,当某要素值在 $(-3\sigma, 3\sigma)$ 内,保留;反之,则作为粗大值,剔除。剔除后剩下的值再次计算各站各土层各土壤水分要素的平均值 \bar{x} 和标准差 σ 。

(3)用再次计算得到的平均值 \bar{x} 减去 3σ ,重新挑取的气候极值的上下界值再结合人工审核,最终确定气候极值检查的阈值范围:

0% < 30 cm 以上(含 30 cm)各层小时平均土壤相对湿度 $\leq 200\%$;

0% < 30 cm 以下各层小时平均土壤相对湿度 $\leq 150\%$;

0% < 0 ~ 10 cm 层土壤质量含水率 $\leq 60\%$;

0% < 10 ~ 40 cm 层土壤质量含水率 $\leq 55\%$;

0% < 40 ~ 50 cm 层土壤质量含水率 $\leq 50\%$;

-5 mm < 0 ~ 30 cm 层土壤有效贮水量 ≤ 60 mm;

-10 mm < 30 ~ 50 cm 层土壤有效贮水量 ≤ 60 mm。

当不同土壤水分要素值超出其相应的阈值范围时,判定该土层某土壤水分要素错误(QC 码=2)。

2.2.4 时间一致性检查参数的建立 时间一致性检查包括时变检查和持续性检查两部分。

(1)时变检查

在无降水的情况下,土壤水分数据基本稳定在一定的范围内,小时间的变化较小。若有降水发生、传感器外壁破裂渗水等其他外界因素影响,土壤水分数据会出现波动或跳变。因此,时变检查就是检查某时次土壤水分数据的前一小时变化是否在一个合理的范围内。该项检查在理论上可以确保土壤水分数据的一致性和连贯性。

经过界限值检查和气候极值检查后,剔除合理范围外的数据,对剩下的数据进行时变检查。由于澄迈站和南滨站的数据存在长期异常的情况,为不影响时变检查阈值范围的挑取,故剔除两

站的数据, 仅对剩下的有降水发生且序列较为稳定的, 数据质量良好的 16 个台站, 进行全月数据的个例统计和分析。

以昌江站 2015 年 6 月各土层的土壤体积含水量与不同降水过程的对应变化情况为例(图 4)。16 日累积降水量为 13.1 mm, 造成 0~10 cm 的土壤体积含水量增加了 6.2%; 其次是 10~20 cm 增

加了 0.1%, 其余各层均未受到影响。由于在降水初期, 当小时雨量达到足够大时, 雨水会往下渗透, 至 10~20 cm 有小幅度的变化; 随着雨量的增加, 雨水不断向下渗透, 至降水过程结束后, 20~30 cm 才有小幅度的变化。20 日累积降水量为 9.0 mm, 造成 0~10 cm 的土壤体积含水量增加了 1.2%, 其余各层均未受到影响。22 日 11:00~23 日

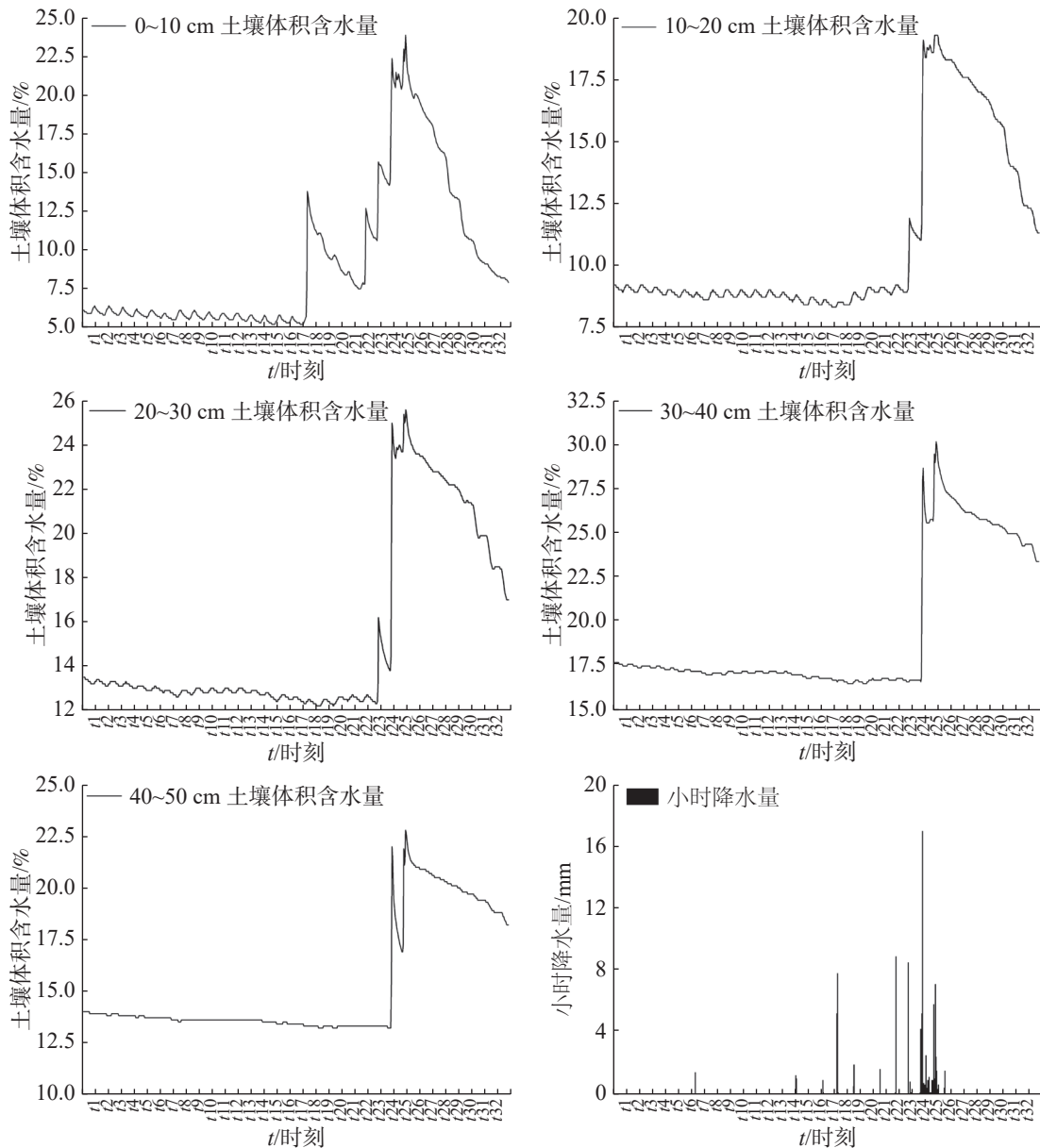


图 4 昌江站 2015 年 6 月小时降水量及各土层小时土壤体积含水量变化

t_1 : 2015-06-1 17:00 ; t_2 : 2015-06-2 15:00 ; t_3 : 2015-06-3 13:00 ; t_4 : 2015-06-4 11:00 ; t_5 : 2015-06-5 9:00 ; t_6 : 2015-06-6 7:00 ; t_7 : 2015-06-7 5:00 ; t_8 : 2015-06-8 3:00 ; t_9 : 2015-06-9 1:00 ; t_{10} : 2015-06-10 23:00 ; t_{11} : 2015-06-11 21:00 ; t_{12} : 2015-06-11 19:00 ; t_{13} : 2015-06-12 17:00 ; t_{14} : 2015-06-13 15:00 ; t_{15} : 2015-06-14 13:00 ; t_{16} : 2015-06-15 11:00 ; t_{17} : 2015-06-16 9:00 ; t_{18} : 2015-06-17 7:00 ; t_{19} : 2015-06-18 5:00 ; t_{20} : 2015-06-19 3:00 ; t_{21} : 2015-06-20 1:00 ; t_{22} : 2015-06-21 23:00 ; t_{23} : 2015-06-22 21:00 ; t_{24} : 2015-06-22 19:00 ; t_{25} : 2015-06-23 17:00 ; t_{26} : 2015-06-24 15:00 ; t_{27} : 2015-06-25 13:00 ; t_{28} : 2015-06-26 11:00 ; t_{29} : 2015-06-27 9:00 ; t_{30} : 2015-06-28 7:00 ; t_{31} : 2015-06-29 5:00 ; 2015-06-30 3:00。

18:00 累积降水量达到 59.9 mm, 各土层随时间均有明显的变化。其中, 22 日 15 时降水量达 17.0 mm, 该小时各层的土壤体积含水量分别增加了 2.3%、2.8%、3.7%、6.3%、6.4%。这与降水量、水分下渗速度有关。降水初期累积降水虽较少, 但仍存在下渗现象, 加之土壤本身涵养的水分, 对 20 cm 以下土层造成的影响相较于 20 cm 以上的要小, 但也导致了当出现较大的降水量时 20 cm 以下土层的土壤体积含水量会迅速增加。通过以上分析可知, 不同的累积降水量对各土层的土壤水分影响有较大的差别, 以 10~20 cm 为界, 当累积降水量在 10.0 mm 以下时, 20 cm 以下土层均不会受到影响; 由于土壤对降水有一定的滞后性^[17], 当累

积降水量在 10.0~15.0 mm 时, 30~40 cm 土层由于雨水下渗, 在降水发生后 1~2 h 内, 会有小幅度的变化, 40~50 cm 土层不会受到影响; 当累积降水量在 15.0 mm 以上(含 15.0 mm)时, 则各土层均会受到影响, 且当降水一直持续发生时, 所观测到的各土层的土壤体积含水量的值也会相应增大。

通过对当前现有的自动观测土壤水分数据进行个例统计分析后, 将当前时次与其前一时次进行差值计算和统计, 确定各土层当前时次与其前一时次差值的上下界限值范围, 即时变检查的阈值范围(表 2)。若超出时变检查的阈值范围, 则判定该时次与其前一时次的该土层该土壤水分要素值为可疑(QC 码=1)。

表 2 各层土壤水分要素当前时次与前一时次差值上下限

土层/cm	上限				下限			
	土壤质量 含水率/%	土壤相对 湿度/%	土壤体积 含水量/%	土壤有效 水分贮存量/mm	土壤质量 含水率/%	土壤相对 湿度/%	土壤体积 含水量/%	土壤有效 水分贮存量/mm
0~20	12	45	15	15	-8	-20	-10	-10
20~50	9	30	20	25	-10	-10	-12	-8

(2) 持续性检查

持续性检查与时变检查不同, 时变检查关注的是各小时间土壤水分数据的变化, 从而判定数据是否有跳变。持续性检查则关注的是土壤水分数据长时间是否有变化, 它是从数据随时间变化的另一个角度去判定数据是否合理, 即当数据的长时间不变时, 可能是由于仪器故障、观测环境变

化或是人工对比观测错误等原因造成, 那么可认为长时间不变的数据序列是可疑的。

通过分析全省 2013—2018 年 S 文件和实时数据, 发现各土层土壤水分要素值经常出现长时间维持一个值不变。这里说的“不变”可分为正常的“不变”(如图 5-a)和异常的“不变”(如图 5-b)。从图 5-a 可以看出, 临高站 2015 年 5 月 10~20 cm

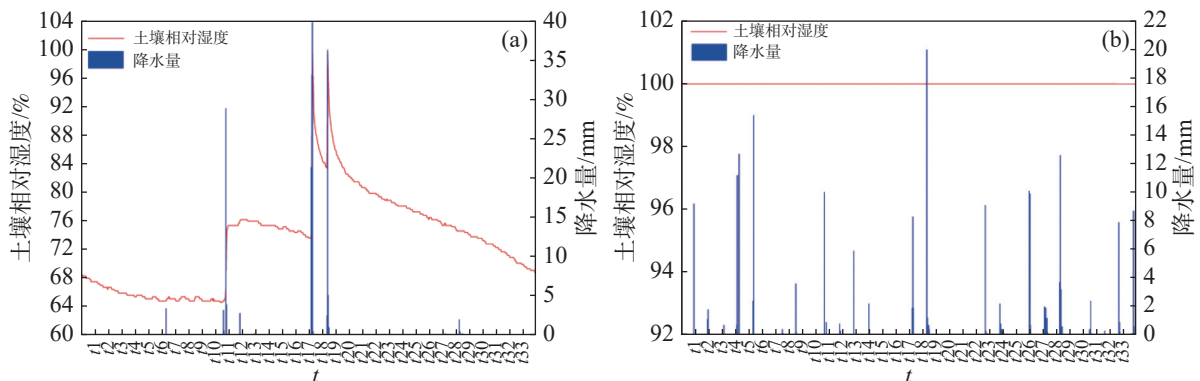


图 5 临高站(a)、澄迈站(b)2015 年 5 月 10~20 cm 相对湿度正常个例

t_1 : 2015-05-1T18:00; t_2 : 2015-05-2T16:00; t_3 : 2015-05-3T14:00; t_4 : 2015-05-4T12:00; t_5 : 2015-05-5T10:00; t_6 : 2015-05-6T8:00; t_7 : 2015-05-7T6:00; t_8 : 2015-05-8T4:00; t_9 : 2015-05-9T2:00; t_{10} : 2015-05-10T0:00; t_{11} : 2015-05-10T22:00; t_{12} : 2015-05-11T20:00; t_{13} : 2015-05-12T18:00; t_{14} : 2015-05-13T16:00; t_{15} : 2015-05-14T14:00; t_{16} : 2015-05-15T12:00; t_{17} : 2015-05-16T10:00; t_{18} : 2015-05-17T8:00; t_{19} : 2015-05-18T6:00; t_{20} : 2015-05-19T4:00; t_{21} : 2015-05-20T2:00; t_{22} : 2015-05-21T0:00; t_{23} : 2015-05-21T22:00; t_{24} : 2015-05-22T20:00; t_{25} : 2015-05-23T18:00; t_{26} : 2015-05-24T16:00; t_{27} : 2015-05-25T14:00; t_{28} : 2015-05-26T12:00; t_{29} : 2015-05-27T10:00; t_{30} : 2015-05-28T8:00; t_{31} : 2015-05-29T6:00; t_{32} : 2015-05-30T4:00; t_{33} : 2015-05-31T2:00。

出现相对湿度维持某一恒定值不变, 但恒定值维持的时间长度有长短之分。在降水或是长时间无降水致土壤中的水分蒸发等影响下, 相对湿度会有上升或是下降的过程, 数据变化正常。从图 5-b 可以看出, 澄迈站 2015 年 5 月 10~20 cm 无论是否有降水发生, 相对湿度一直维持 100% 不变, 数据异常变化。

在该项检查中对“长时间不变”的时间长度的设定很重要。若设置时间过短, 会出现大量的疑误信息, 无法提高质控的效率; 反之, 会漏检疑误

数据, 无法保证数据的准确性。假设 t 时次的某土层的土壤水分 Q 未缺测, 那么从该时次往前推, 若连续 n 个时次维持一个值不变, 那么 $t-n$ 至 $t-1$ 时段内的观测值均为该检查对象。若 $t-n$ 时次或 $t-1$ 时次的土壤水分 Q 判定为疑误 (QC 码=1) 或错误 (QC 码=2), 那么该时段内的观测值均判为疑误 (QC 码=1) 或错误 (QC 码=2)。该项检查也可以作为土壤水分观测仪器是否故障的重要依据。

图 5-b 中澄迈站 10~20 cm 的相对湿度的错误值均可被有效判定, 均判定为错误值。

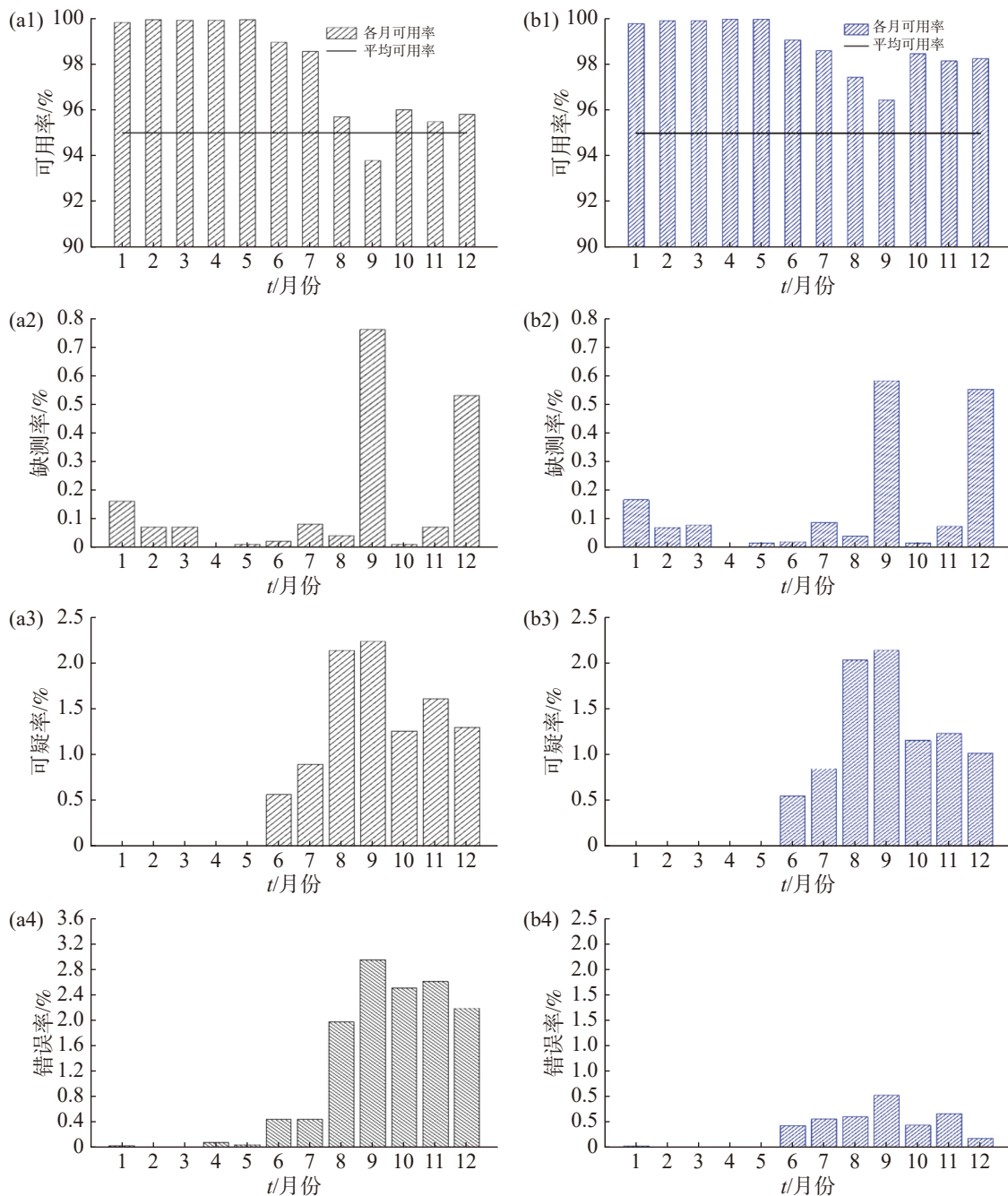


图 6 2019 年各月土壤水分数据质量状况

3 质量控制结果

为了检验质控方案对本省土壤水分自动站数据的应用效果,利用缺测率、可用率、可疑率和错误率对海南省 18 个土壤水分站 2019 年小时数据进行质量评估。各项质量控制检查方法的结果统计可以看出,2019 年数据缺测量为 22224 个,缺测率为 0.70%;数据可疑量为 8398 个,可疑率为 0.27%;数据错误量为 10495 个,错误率为 0.33%。与现用的 RASM(自动土壤水分观测报表系统)的审核结果比较(表 3),本质控方案不仅遍历了所有土壤水分要素及台站参数,并且针对不同的土壤水分要素值从历史气候、小时等不同时间维度进行质控;而 RASM 仅对一个土壤水分要素值进行质控,阈值范围设置较为宽泛,将一些正确的数据也判断为疑误数据,且错误数据并未指出。

表 3 本质控方案与 RASM 审核质控结果对比

质控方法	总数 数据/个	疑误 数据/个	错误 数据/个	错检 率/%	漏检 率/%
本质控方案	3 153 600	7 663	8 961	0.2	0.0
RASM	3 153 600	657 648	0	3.1	10.4

从图 6-a₁、a₂ 可知,2019 年全年土壤水分数据平均可用率为 97.80%,除 9 月外,其余各月数据可用率较高,均在平均可用率以上。结合各月缺测率、错误率和可疑率来看,6 月以后的缺测率、错误率和可疑率均较高,特别是 9 月的缺测率、错误率和可疑率分别为 0.76%、3.21%、2.24%。经核实发现,是由于琼海站在 6 月迁站,相对应的土壤水文物理参数并未进行及时修改,且安装土壤观测仪器时不规范,导致数据出现大量缺测、错误和可疑。为了进一步验证该结论,将琼海站的数据剔除后,重新对全年土壤水分数据进行质控和统计(图 6-b₁、b₂)发现,2019 年全年土壤水分数据平均可用率为 98.84%,较未删除琼海站数据前有所提高,全年各月均在平均可用率以上;各月的缺测率、错误率和可疑率均有所下降,特别是错误率,下降了 77.72%。由此说明,提出的质控方法能够提出疑误数据。

4 结论与讨论

本研究提出一套适合海南省自动土壤水分观

测资料的质量控制方法。首先,通过分析 S 文件和实时数据资料,得到 18 个土壤水分站的各土层各土壤水分要素的气候极值和时变检查的阈值范围;其次,结合土壤水分自动观测仪器原理及本省的气候特点,设置和确定了参数检查、数据缺测检查、界限值检查、气候极值范围检查、时变检查、持续性一致性检查 6 项检查;最后,利用 2019 年海南省土壤水分自动观测资料对质量控制方法的应用效果进行检验。结果表明:

(1)6 项检查均能够检查和判断出自动土壤水分观测资料中相应的疑误、错误和缺测数据;

(2)与现行的 RASM 审核软件相比,遍历所有土壤水分要素及台站参数,针对不同的土壤水分要素值从历史气候、小时等不同时间维度设置阈值范围,降低了漏检率和错检率。

除天气影响外,浇水、农业灌溉、地下水位的升高或降低、植被的覆盖率等客观因素都会对土壤水分造成一定的影响,因此,在后续的研究中还需针对不同地区、不同的植被覆盖类型、不同的土壤类型、不同的天气条件等方面,对海南省土壤水分数据进行更加细致的质量控制。

参考文献:

- [1] 陈海山,孙照渤. 陆气相互作用及陆面模式的研究进展[J]. 南京气象学院学报, 2002, 25(2): 277-288.
- [2] Robock A. Hydrology, floods and droughts: soil moisture[J]. Encyclopedia of Atmospheric Sciences (Second edition), Elsevier, 2015, 52(6): 232-239.
- [3] 鲁向东. 土壤湿度气象观测方法介绍[J]. 农业与技术, 2018, 38(7): 151-152.
- [4] 闫峰,覃志豪,李茂松,等. 农业旱灾监测中土壤水分遥感反演研究进展[J]. 自然灾害学报, 2006, 15(6): 114-121.
- [5] M TUGRUL YILMAZ, BURAK BULUT, MICHAEL H COSH, et al. Quality control of station-based soil moisture observations over Turkey[C]//Agu Fall Meeting. AGU Fall Meeting Abstracts, 2015.
- [6] WEILIN L, DAGANG W, GUILING W, et al. Quality control and evaluation of the observed daily data in the North American soil moisture database [J]. Journal of Meteorological Research, 2019, 33(3): 501-518.
- [7] DORIGO W A, XAVER A, VREUGDENHIL M, et al. Global automated quality control of *in situ* soil moisture data from the international soil moisture network [J]. Vadose Zone Journal, 2013, 12(3): 918-924.
- [8] 曹婷婷,邵楠,李巍,等. 全国土壤水分观测系统数据质量分析[J]. 现代农业科技, 2018, 24: 209-211.

- [9] 国家气象局气候监测应用管理司. 气象仪器和观测方法指南[M]. 北京: 气象出版社, 1992.
- [10] 宋连春, 李伟. 综合气象观测系统的发展[J]. 气象, 2008, 34(3): 3-9.
- [11] 吴东丽, 梁海河, 曹婷婷, 等. 中国自动土壤水分观测网运行监控系统建设[J]. 气象科技, 2014, 42(2): 278-282.
- [12] 曹婷婷, 邵楠, 李巍. 影响全国自动土壤水分站运行能力因素分析[J]. 气象水文海洋仪器, 2017, 34(4): 27-32.
- [13] 唐修熊, 董克非, 隋一勇, 等. 以介电常数为中心的自动土壤水分观测设备标定体系的探讨[J]. 气象水文海洋仪器, 2011, 12(4): 14-18.
- [14] 张志富. 自动站土壤水分资料质量控制方案的研制[J]. 干旱区地理, 2013, 36(1): 101-108.
- [15] 郭旭, 龙柯吉, 赵旋, 等. 四川地区自动土壤水分站数据质量控制方法研究[J]. 高原山地气象研究, 2017, 37(2): 90-95.
- [16] 胡伟, 熊凌云, 熊雄, 等. 自动土壤水分数据质量控制中的阈值确定[J]. 气象水文海洋仪器, 2012, 3: 14-16.
- [17] 施晨晓, 陈珍莉, 刘霄燕, 等. 海南岛土壤体积含水量变化规律及其与气候因子的关系[J]. 贵州农业科学, 2019, 47(9): 142-149.
- [18] 施晨晓, 陈珍莉, 刘霄燕, 等. 海南省自动土壤水分观测数据异常原因分析[J]. 气象科技进展, 2020, 10(4): 123-127.
- [19] 任芝花, 赵平, 张强, 等. 适用于全国自动站小时降水资料的质量控制方法[J]. 气象, 2010, 36(7): 123-132.
- [20] 吴东丽, 曹婷婷, 薛红喜. 自动土壤水分观测数据质量控制方法及其应用[J]. 土壤科学, 2016, 4(1): 1-10.
- [21] XIA YOU LONG, FORD T W, WU YIHUA, et al. Automated quality control of in situ soil moisture from the North American soil moisture database using NL-DAS-2 products [J]. Applied Meteorology and Climatology, 2015, 54(6): 1267-1282.
- [22] 王新华, 罗四维, 刘小宁, 等. 国家级地面自动站 A 文件质量控制方法及软件开发[J]. 气象, 2006, 32(3): 107-112.
- [23] 王海军, 杨志彪, 杨代才. 自动气象站实时资料自动质量控制方法及其应用[J]. 气象, 2007, 33(10): 102-109.
- [24] 王建荣, 江双伍, 董得保. 自动观测土壤体积分含水量实时资料自动质量控制方法[J]. 计算技术与自动化, 2013, 32(4): 134-137.
- [25] 任芝花, 许松, 孙化南, 等. 全球地面天气报历史资料质量检查与分析[J]. 应用气象学报, 2006, 17(4): 412-420.

Quality control of data from automatic soil moisture stations in Hainan Province

SHI Chenxiao, YANG Qingwen, CHEN Zhenli, LIU Xiaoyan, WANG Xiaojie

(Hainan Meteorological Information Center/Hainan Key Laboratory of South China Sea Meteorological Disaster Prevention and Mitigation, Haikou, Hainan 570203, China)

Abstract: Based on the data on automatic soil moisture in Hainan Province from 2013 to 2018, the threshold values of different soil elements at different soil layers were obtained by using statistical analysis, and a quality control scheme suitable for automatic soil moisture in Hainan Province was formulated. The quality control scheme was applied to the soil moisture data on an hourly basis from the automatic stations in 2019 and was found to have a good applicability in the quality control of automatic observation data on soil moisture in Hainan Province. The quality control scheme had an annual monthly availability rate of above 90%, and its average availability rate is above 97% except in September. This scheme reduced the false detection rate and missed detection rate, compared with the current audit software RASM.

Keywords: Hainan Province; automatic soil moisture; real-time data; S file; quality control; quality assessment

(责任编辑: 刘少军 责任编辑: 叶 静)