

王佳强,赵煜飞,任芝花,等,2018.中国自动土壤水分观测资料质量控制方法设计与效果检验[J].气象,44(2):244-257.

# 中国自动土壤水分观测资料质量控制 方法设计与效果检验<sup>\*</sup>

王佳强 赵煜飞 任芝花 高 静

国家气象信息中心,北京 100081

**提 要:** 土壤湿度资料对气候变化、农业干旱监测、农业气象预报与服务等研究具有重要意义,为剔除土壤湿度观测资料中的异常数据,本文提出了一套适用于全国自动土壤水分观测资料的质量控制方法。首先,以2014年全国自动土壤水分观测资料为基础,根据资料中异常数据的特征将异常数据分为四类。其次,从界限值检查、内部一致性检查、时间一致性检查等方面提出:异常极值检查、异常增大检查、异常减小检查、异常恒定检查四类方法。最后,利用2014—2015年全国观测资料以及中国气象局陆面数据同化系统(CLDAS)土壤体积含水量数据集产品(V2.0)对各检查方法应用效果进行检验,结果表明:(1)四类检查方法均可判断出自动土壤水分观测资料中的疑误数据;(2)四类检查方法的判定结果在时间连续性上及空间分布上具有一定的一致性;(3)该质量控制方法可减小观测数据与CLDAS数据之间的均方根误差(RMSE)。目前,该方法已应用于我国气象资料处理业务系统。

**关键词:** 中国区域,自动土壤水分观测站,逐小时资料,土壤体积含水量,质量控制,CLDAS

**中图分类号:** P413

**文献标志码:** A

**DOI:** 10.7519/j.issn.1000-0526.2018.02.004

## Design and Verification of Quality Control Methods for Automatic Soil Moisture Observation Data in China

WANG Jiaqiang ZHAO Yufei REN Zhihua GAO Jing

National Meteorological Information Centre, Beijing 100081

**Abstract:** Soil moisture data play a key role in the study of climate change and agricultural drought monitoring, agricultural weather forecast and service. In order to effectively eliminate the abnormal data in observations, this paper puts forward a set of quality control (QC) methods which could be applied to the data of automatic soil moisture observation station (ASMOS) in China. First, based on the data of ASMOS 2014 in China, the abnormal data are divided into four categories according to their characters. Secondly, under the consideration of three aspects: threshold value check, internal consistency check, time consistency check, the QC methods are designed, which include abnormal extreme check, abnormal increase check, abnormal decrease check and abnormal constant check. Finally, the QC methods are verified by using the data of ASMOS and soil volumetric water content products of CMA Land Data Assimilation System (CLDAS-V2.0) in China in 2014—2015. The results show that: (1) the four kinds QC methods can effectively identify the four types of abnormal data. (2) The results from the four kinds QC methods are in good agreement in temporal continuity and spatial distribution. (3) The QC methods can effectively reduce the root mean square error (RMSE) between observation and the CLDAS data. At present, the methods have been applied to the Meteorological Data Processing Service System.

<sup>\*</sup> 公益性行业(气象)科研专项(GYHY201106038)和国家自然科学基金项目(91637313)共同资助

2017年1月20日收稿; 2017年11月3日收修定稿

第一作者:王佳强,主要从事气象资料质量控制及评估工作. Email: wangjiaq@cma.cn

**Key words:** China region, automatic soil moisture observation station (ASMOS), hourly data, soil volumetric moisture content, quality control, CMA Land Data Assimilation System (CLDAS)

## 引言

土壤湿度是研究陆气相互作用的重要影响因素之一,它通过影响地表的感热和潜热输送,进而对天气和气候产生较大影响(Entekhabi et al,1996;马柱国等,2001;郭维栋等,2007),同时,掌握土壤湿度的变化规律,也是开展农业干旱监测、农业气象预报与服务等课题的重要依据(王万秋,1991;刘宪锋等,2015;孙小龙等,2015)。自动土壤水分观测站可以方便、快速、连续地在同一地点进行不同层次土壤水分的观测,具有实时性及连续性强的特点,目前我国已建成2000多个站的观测站网并实现观测数据的实时上传(吴东丽等,2014)。该站网的业务运行,可提供一套我国时空分辨率最高的土壤水分地面观测数据集。然而,目前观测业务系统中还没有建立相应的质量控制流程,因此,为提高该资料的实用性,需要建立一套质量控制方法对土壤湿度质量进行检查。

关于气象资料的质量控制方法研究,目前较成熟的方法主要包含:界限值检查、内部一致性检查、时间一致性检查以及空间一致性检查等(WMO,2008;任芝花等,2010;中国气象局,2010a)。其中,针对中国土壤水分观测资料的质量控制,国内外学者做了一系列研究。Dorigo et al(2013)根据土壤含水量的时间序列特征,以及土壤含水量与地温、降水的一致性关系等,设计了一套质量控制方法并将该方案应用于全球土壤水分数据网,然而,该数据网仅包含中国40个人工观测站2000年之前的观测数据。元保军(2014)通过对河南省历年人工观测土壤水分数据的统计分析,确定了针对河南省各站各层土壤的阈值检查方法;王晓东等(2015)对安徽省土壤水分人工观测站与自动观测站两者间绝对误差及相关性进行了分析,建立了安徽省土壤水分自动观测数据质量控制方法;张喜英等(2015)通过分析常德地区9个自动土壤水分观测站观测数据,建立了适合当地条件的各层土壤水分阈值和质量控制方案。遗憾的是该系列研究主要依据某一地区的成果设计方案,并未分析各质量控制方案在全国范围内的适用性。对于全国自动土壤水分观测资料的质量

控制,张志富(2013)利用全国人工观测和自动观测土壤水分资料的统计特点,分别从界限值检查、内部一致性检查、时间一致性检查三个方面建立了一套质量控制方法;吴东丽等(2016)利用范围检查、时变检查和持续性检查方法,对2012—2014年表层土壤(0~10 cm)数据进行了检查和分析。然而,在内部一致性检查方面,各方案均未引入与地温、降水的关系检查,检查方法相对单一。

针对以往研究中主要依据中国某一区域,并且尚未引入地温、降水数据来增加内部一致性检查等不足,本文提出了一套可适用于全国自动土壤水分观测资料的质量控制方法。首先,以2014年全国自动土壤水分观测站数据为基础,根据资料中的异常数据特征,将异常数据分为四类。其次,从界限值检查、内部一致性检查、时间一致性检查等方面提出:异常极值检查、异常增大检查、异常减小检查、异常恒定检查共四类检查方法,并在内部一致性检查中增加了土壤含水量与逐小时地温、降水的一致性检查,以及土壤含水量逐小时变率的内部一致性检查。最后,利用2014—2015年全国观测资料以及中国气象局陆面数据同化系统(CMA Land Data Assimilation System,CLDAS)土壤体积含水量数据集产品(V2.0)对本文方法的应用效果进行检验。

## 1 资料与术语

### 1.1 土壤水分资料、质量控制对象及质量控制码设定

自动土壤水分观测站(automatic soil moisture observation station, ASMOS)逐小时资料包含四个要素:土壤体积含水量、土壤相对湿度、土壤重量含水率、土壤有效水分贮存量,均为小时正点前10分钟的平均值。土壤测量深度共有8层:0~10、10~20、20~30、30~40、40~50、50~60、70~80、90~100 cm,个别台站由于土壤类型等因素,其50 cm以下可不观测。自动土壤水分观测站采用频域反射法(FDR)测量土壤体积含水量,其他三个要素可结合土壤水文参数计算得到(中国气象局,2010b),因此,本文以土壤体积含水量(soil volumetric mois-

ture content,  $Q$ )为质量控制对象,其他三个要素直接采用土壤体积含水量的质量控制结果。

本文选取2014年全国土壤水分数据作为试验数据,设计质量控制方法,并选取2015年全国数据以及CLDAS生成的2014—2015年逐小时、空间分辨率为 $0.0625^\circ \times 0.0625^\circ$ 的土壤体积含水量产品(V2.0)对方法的应用效果进行检验。在对数据进行质量控制(quality control, QC)的过程中,随着质量控制进程的推进,需要不断地对被检数据设置或修改QC码。QC码的规定如下:数据正确, QC码=0;数据可疑, QC码=1;数据错误, QC码=2;数据缺测, QC码=8;数据未做质量控制, QC码=9(中国气象局, 2010a)。

## 1.2 辅助资料与术语

在设计质量控制方法时,为增加土壤体积分水量的内部一致性检查,本文还使用了地温与降水数据,该数据主要采用国家基准站、基本站及一般站的逐小时观测数据。2014—2015年自动土壤水分观测站共约1500站,本文选择10 km范围内距离自动土壤水分观测站最近的地面观测站逐小时数据作为辅助数据,经计算共匹配成功1175站,分布覆盖全国(图1)。

同时,为后文分析方便,本文约定:以 $L=1, 2, \dots, 8$ 分别指代各层土壤;以 $Q(t)$ 表示某台站 $t$ 时次的土壤体积含水量(单位: $0.01 \text{ g} \cdot \text{cm}^{-3}$ ),其逐小时变率 $\Delta Q(t)$ 可由式(1)计算得到:

$$\Delta Q(t) = Q(t) - Q(t-1) \quad (1)$$

当 $\Delta Q(t) > 0$ 称为逐小时正变率, $\Delta Q(t) < 0$ 称为逐小时负变率。若需明确指定某层土壤,则分别用 $Q(L, t)$ 、 $\Delta Q(L, t)$ 表示( $L=1, 2, \dots, 8$ );以 $RH(t)$

表示某台站 $t$ 时次的土壤相对湿度(单位:%),同样 $RH(L, t)$ 特指某层土壤;以 $RR(t)$ 表示某台站 $t$ 时次的小时降水量(单位:mm);文中时间均采用世界时。

## 2 质量控制方法设计

为建立适用于全国自动土壤水分观测资料的质量控制方法,本文以2014年全国自动土壤水分观测数据为基础,通过分析资料中的异常数据特征,将异常数据分为四类:异常极值、异常增大、异常减小和异常恒定。在此基础上,根据四类异常数据的特征分别设计相应的质量控制方法。

### 2.1 异常数据特征

图2a为甘肃环县站(53281)正常的土壤体积含水量逐小时序列图,由图可以看到:当有降水事件时,土壤体积含水量会有明显的增加;降水事件结束后,土壤体积含水量开始逐渐缓慢的减小(如图AB段所示)。图2b为甘肃崆峒站(53915)土壤体积含水量逐小时序列图,与图2a对比发现:该小时序列存在明显的跳跃断点,主要表现为突然的增大或减小;同时,在跳跃断点之间还存在观测值长时间保持不变的恒定现象,如图2b中CD、EF、GH各段所示。通过人工核查可以确定图2b中的突然增大、突然减小以及恒定数据均为疑误数据,本文分别称为异常增大、异常减小和异常恒定。另外,观测数据超出其合理范围之外的极大/极小值,也是一种常见的异常数据类型,本文简称异常极值(图略)。

在实际业务中,该四类异常数据比较常见且特征明显,对此,本文分别设计了相应的检查方法:异常极值检查(abnormal extreme value check, AEC)、异常增大检查(abnormal increase value check, AIC)、异常减小检查(abnormal decrease value check, ADC)和异常恒定检查(abnormal consistent value check, ACC)。其中AEC方法属于界限值检查范畴,AIC和ADC方法属于内部一致性检查范畴,ACC方法则属于时间一致性检查范畴。

### 2.2 异常极值检查

#### 2.2.1 与地温关系检查

与地温关系检查(relationship check between  $Q$

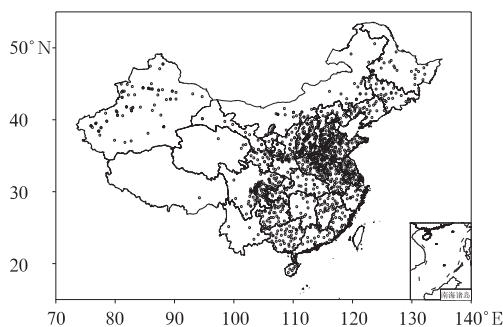


图1 自动土壤水分观测站空间分布图(1175站)

Fig. 1 Distribution of ASMOS in China  
(1175 stations)

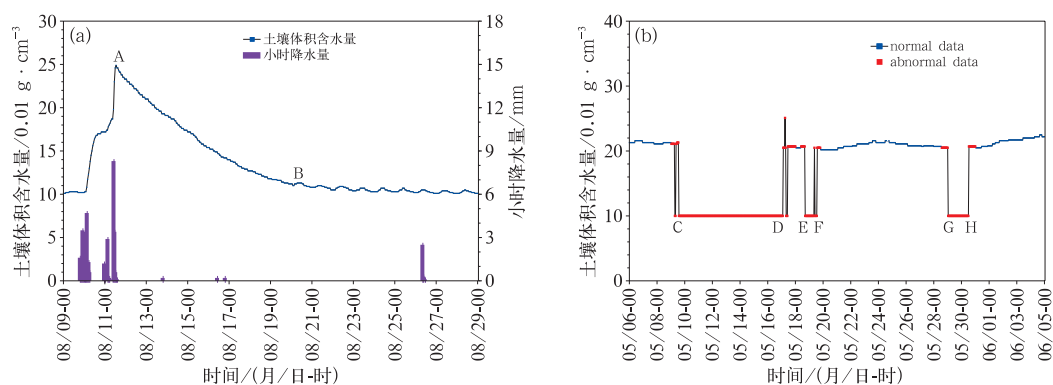


图2 土壤体积含水量逐小时序列

(a) 甘肃环县站(53281)2015年第一层土壤正常个例,

(b) 甘肃崆峒站(53915)2015年第三层土壤异常个例

Fig. 2 Hourly sequence of  $Q$ 

(a) the 2015 1st layer soil normal case of Huanxian Station, Gansu (code 53821),

(b) the 2015 3th layer soil abnormal case of Kongtong Station, Gansu (code 53915)

and soil temperature of AEC, AECst), 是利用地温数据对土壤体积含水量异常低值进行判别的方法, 目前已在国际上得到广泛应用(Dorigo et al, 2013; Xia et al, 2015)。本文直接采用该方法, 即: 当观测层的土壤地温 $\leq 0^{\circ}\text{C}$ 时, 则判定该层土壤体积含水量观测值为错误, QC 码=2。图3为内蒙古鄂温克站(区站号 50525)2014年第一层土壤体积含水量与地温的逐小时序列散点图, 由图可以看到: 当土壤温度 $< 0^{\circ}\text{C}$ 时, 土壤体积含水量观测值存在明显的偏小现象。

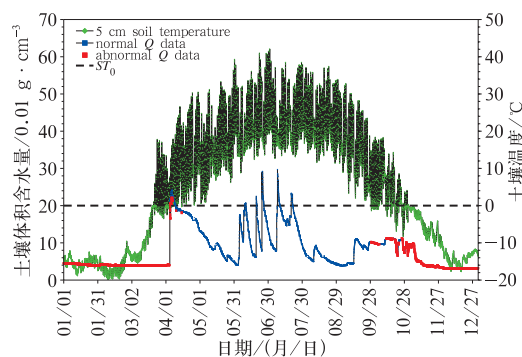


图3 2014年内蒙古鄂温克站(50525)第一层数据与地温关系检查异常个例

(绿色: 地温; 蓝色: 土壤体积含水量; 红色: 异常土壤体积含水量;  $ST_0$ : 土壤温度  $0^{\circ}\text{C}$  线)

Fig. 3 The 1st layer soil abnormal case of AEC method on ASMOS data: Ewenke Station, Neimenggu (code 50525) in 2014

(Green:  $ST$ ; blue:  $Q$ ; red: abnormal  $Q$ ,  $ST_0$ :  $0^{\circ}\text{C}$  line of  $ST$ )

## 2.2.2 界限值检查

界限值检查(threshold value check of AEC, AECtrh)主要是判断土壤体积含水量  $Q$  是否在其合理范围内的方法, 由此可剔除观测数据中明显的异常极值。本文采用王良字和何延波(2015)提出的利用土壤饱和含水量、土壤最大吸湿量分别估算  $Q$  的上/下限值的方法, 计算公式见式(2)和式(3)(耿增超和戴伟, 2011)。若  $Q$  不在上/下限值范围内, 则数据判为错误(QC 码=2)。

$$\text{土壤饱和含水量} = \left(1 - \frac{\rho}{\rho_s}\right) \rho_w \times 100 \quad (2)$$

$$\text{土壤最大吸湿量} = \frac{\omega_k}{1.5} \times \rho \quad (3)$$

式中,  $\rho_s$  为土粒密度, 一般在  $2.5 \sim 2.8 \text{ g} \cdot \text{cm}^{-3}$  范围内,  $\rho$  为土壤容重(单位:  $\text{g} \cdot \text{cm}^{-3}$ ),  $\rho_w$  为水的密度(本文采用  $1.0 \text{ g} \cdot \text{cm}^{-3}$ ),  $\omega_k$  为凋萎湿度(单位: %, 采用重量含水率表示), 由式(2)和式(3)计算得到的土壤饱和含水量以及土壤最大吸湿量, 单位均为:  $0.01 \text{ g} \cdot \text{cm}^{-3}$ 。另外, 由于土壤体积含水量上限值一般不会大于 60(单位:  $0.01 \text{ g} \cdot \text{cm}^{-3}$ )(Dorigo et al, 2013; 吴东丽等, 2014; Xia et al, 2015), 因此, 当土壤饱和含水量估算值大于 60 时本文即取 60。

图4为贵州德江站(57637)第四层土壤 2014 年及 2015 年界限值检查结果, 通过对比图 4a 与 4b 可以明显看到: 2014 年大部分数据均存在偏大现象, 至 2015 年数据则处于合理范围内, 数据偏大的现象



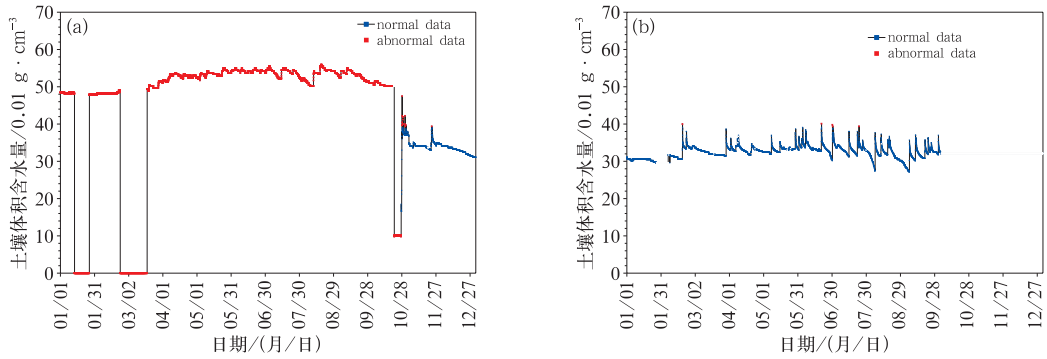


图 4 2014 年(a)和 2015 年(b)贵州德江站(57637)第四层数据界限值检查个例

Fig. 4 The 4th layer soil abnormal cases of AECtrh method on ASMOS data;  
Dejiang Station, Guizhou (code 57637) in 2014 (a) and 2015 (b)

消失。需注意,若土壤水文参数本身存在误差,将直接影响上/下界限值的估算结果,进而可能引起本项检查的误判。因此,若某台站观测数据经该项检查后绝大部分数据被判定为错误,应首先确认是否与土壤水文参数的不准确有关。

2.3 异常增大检查

异常增大检查主要是判断逐小时正变率  $\Delta Q$  是否在其合理范围内的方法。假设  $\Delta Q$  的上限值为  $Q_a$  ( $Q_a > 0$ ),那么当  $\Delta Q(t) > Q_a$  时即为异常增大。由于降水是  $Q$  的主要来源,因此,本文使用  $\Delta Q$  与降水的关系对  $Q_a$  进行估算,称为  $\Delta Q$  与降水关系检查 (check the relationship between  $\Delta Q$  and  $RR$  of AIC, AICrr)。此外,针对个别台站难以匹配到相应的逐小时降水数据,本节进一步设计了  $\Delta Q$  与上层  $\Delta Q$  的关系检查,称为  $\Delta Q$  与上层变率关系检查 (check the relationship between QC layer and upper layer on element  $\Delta Q$  of AIC, AICqq)。

2.3.1  $\Delta Q$  与降水关系检查

对于某层土壤,当土壤体积含水量采用单位“ $0.01 \text{ g} \cdot \text{cm}^{-3}$ ”、小时降水量采用单位“mm”时,两者在数值上相等。记  $t$  为质量控制时次,由  $t$  时次向前推若干时次(记为  $m_1$ ),若  $t-m_1$  至  $t$  时段内总的降水量为  $RR_x$ ,并假设  $RR_x$  通过某种途径直接进入质量控制层土壤(记为  $L$  层)且不会继续下渗,那么  $RR_x$  即可作为  $t$  时次该层土壤  $\Delta Q(L, t)$  的上限值  $Q_a$ ,其数值可由式(4)计算得到。

$$Q_a = RR_x = \sum_{T=t-m_1}^{T=t} RR(T) \quad (4)$$

对于不同的土壤层,式(4)中  $m_1$  的取值至关重

要:若  $m_1$  取值较大,则  $Q_a$  值较大,此时错误数据被漏判的风险则较大;相反,  $m_1$  取值较小,则  $Q_a$  值较小,此时正确数据被误判的风险则较大。本文使用 2014 年全国观测数据对参数  $m_1$  进行统计并确定其取值,详见附件 A. 1。将该方法应用于全国 2014 年观测数据,发现当  $\Delta Q > 2.0$  时,本项检查的漏判率及误判率均相对较低,因而选择  $\Delta Q > 2.0$  作为检查的起判条件。以第  $L$  层土壤为例,当  $\Delta Q(L, t) > 2.0$  且  $\Delta Q(L, t) > Q_a$  时,则  $t$  时次及  $t-1$  时次的土壤体积含水量判为错误, QC 码 = 2。需注意,参数  $m_1$  及起判条件均针对全国台站而设计,在将本方法应用于个别台站时,可根据该台站的实际观测数据确定合适的参数值。此外,本项检查可适用于所有层土壤,而 Dorigo et al(2013)提出的  $\Delta Q$  与降水关系检查仅适用于表层土壤。

由于表层土壤(0~10 cm)容易发生龟裂,并且易受气温、降水等环境因素的影响,因此,在起判条件之外( $2.0 \geq \Delta Q > 0$ )的观测数据中出现异常数据的概率较高,为进一步剔除该部分异常数据,本节增加了表层土壤的可疑数据判定。以  $t$  时次数据为例,根据  $RR(t)$  的不同取值给出  $Q_a$  的取值(表 1),当  $2.0 \geq \Delta Q > Q_a$  时,则  $t$  时次及  $t-1$  时次的土壤体积含水量判为可疑, QC 码 = 1。表 1 中  $Q_a$  参数的统计方法见附件 A. 2。

表 1 表层土壤体积含水量与降水关系检查的参数阈值  
Table 1 Threshold value for the relationship  
check between surface  $Q$  and precipitation

编号	$RR(t)$ 取值	$Q_a$ 取值
1	0	0.5
2	$(0, 2.0]$	$RR(t-1) + RR(t)$

图 5a 和 5b 为本项检查发现的疑似数据示例。由图 5a 可以看到:无论土壤处于干旱或湿润状态疑似数据均普遍存在,其中图 5b 为图 5a 中黑框标记序列的放大图,可进一步看到在无降水的情况下土壤体积含水量存在明显的有规律的波动,且与地温(绿线)的波动趋势具有很高的 consistency。Davis and Chudobiak(1975)曾指出:当砂土或黏土的温度从 1℃ 升高到 40℃ 时,土壤介电常数会增加约 10%。经查该站各层土壤均为黏土类型,说明该类异常数据与地温有关。

图 5c 和 5d 为本项检查的错误数据示例,由图可以看到:对于明显的跳跃数据本项检查均可识别。另外,图 5d 所示错误数据中有个别数据出现在小降水或无降水情况下,该现象可能与土壤龟裂有关(陈海波,2013),但也不能排除与降水的非均匀性有关。由于逐小时灌溉数据难以获取,该方法尚未考虑由灌溉引起的土壤水分突然增大,此时本方法会产生

误判。但是,该方法主要是针对跳跃数据进行判定,所以,灌溉事件仅会引起当前时次及前一时次数据的误判,不会影响灌溉事件之后的各时次数据。

### 2.3.2 $\Delta Q$ 与上层变率关系检查

若以第  $L-1$  层土壤为例,记  $t-m_2$  ( $m_2>0$ ) 至  $t$  时刻范围内进入以及移出该层土壤的水分总量为  $RR_y$ ,假设  $RR_y$  全部进入该层土壤的下邻层( $L$  层)且继续下渗,那么  $t$  时次第  $L$  层土壤  $\Delta Q(L, t)$  的上限值  $Q_a$  即为  $RR_y$ 。根据土壤体积含水量定义, $t$  时次进入/移出某层土壤的水分与  $\Delta Q(t)$  的数值相当,因此  $RR_y$  数值可由式(5)计算得到。

$$Q_a = \alpha \times RR_y = \alpha \times \sum_{T=t-m_2}^{T=t} |\Delta Q(L-1, T)| \quad (5)$$

式中, $\alpha$  为权重系数,可调整  $Q_a$  的大小; $m_2$  为时间窗参数,与 2.3.1 节  $m_1$  参数类似。为计算方便本文取  $m_2 = m_1$ , 参数  $\alpha$  的取值详见附件 A.3。同样,本

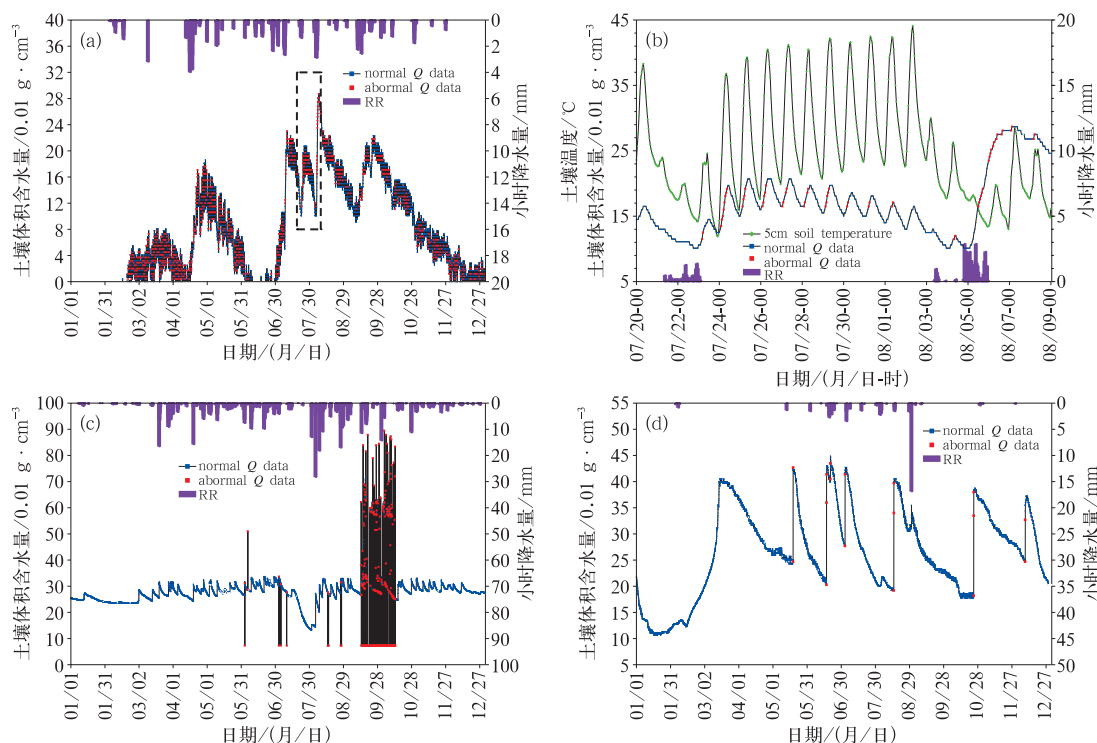


图 5 与降水关系检查的异常个例:(a)甘肃景泰站(52797)2014 年第一层土壤,(b)甘肃景泰站(52797)2014 年 7 月 20 日至 8 月 9 日第一层土壤,(c)重庆龙洲湾站(A8802)2014 年第二层土壤,(d)新疆伽师站(51707)2014 年第二层土壤

Fig. 5 The 1st layer soil abnormal cases of AICrr method on ASMOS data: Jingtai Station, Gansu (code 52797) in 2014 (a) and from 20 July to 9 August 2014 (b), the 2nd layer soil cases of Longzhouwan Station, Chongqing (code A8802) (c) and Jiashi Station, Xinjiang (code 51707) (d) in 2014

项检查也以  $\Delta Q > 2.0$  作为起判条件,即:当  $\Delta Q(L, t) > 2.0$  且  $\Delta Q(L, t) > Q_a$  时,则  $t$  时次及  $t-1$  时次的土壤体积含水量判为错误, QC 码 = 2。需注意,若质量控制层的上邻层数据也存在异常,本项检查可能会引起数据的漏判或误判。

图 6 为本项检查的异常数据示例,由图可见:对于明显的异常跳跃数据,本项检查均可识别。另外,图 6a 虚线框中所示异常跳跃数据并未作出准确判断,经查,其上邻层土壤在相应时次也存在明显的跳跃数据(图略),此情况下本项检查产生漏判。但是,该漏判数据在前文的与降水关系检查中均已被识别和判定,如图 5c 所示。

### 2.4 异常减小检查

异常减小检查是判断逐小时负变率  $\Delta Q$  是否在其合理范围内的方法。假设  $\Delta Q$  的下限值为  $Q_b$  ( $Q_b < 0$ ),当  $\Delta Q < Q_b$  则为异常减小。土壤的干旱是

一个缓慢的过程,土壤湿度越小其逐小时负变率的数值也越小(如图 2a AB 段所示),因此,  $Q_b$  值可根据相邻时次土壤湿度状况进行设定。土壤湿度可采用  $RH$  定量描述,本节根据  $RH(t-1)$ 、 $RH(t)$  的不同取值范围,分别给出  $\Delta Q(t)$  的异常范围及对应 QC 码取值,详见表 2。表中各参数适用于各层土壤,其设定方法详见附件 A. 4。当  $RH(t-1)$ 、 $RH(t)$  及  $\Delta Q(t)$  同时满足表 2 任意一组条件时,  $t$  时次及  $t-1$  时次  $Q$  的质量控制码即取相应的 QC 值。

图 7 为本项检查的异常数据示例,由图 7a 可以看到土壤体积含水量在全年时段内以恒定步长 0.6 进行频繁的上下跳跃;图 7b 所示错误数据则表现为无规律的上下跳跃。图 7c 和 7d 为一种典型的异常数据类型,其中图 7d 为图 7c 中虚线框标记序列的放大图。由图 7d 可以看到在连续的小时序列中存在很小且不规则的向下跃变,该类异常数据在全年时段内均存在(图 7c),Dorigo et al(2013)称此类向

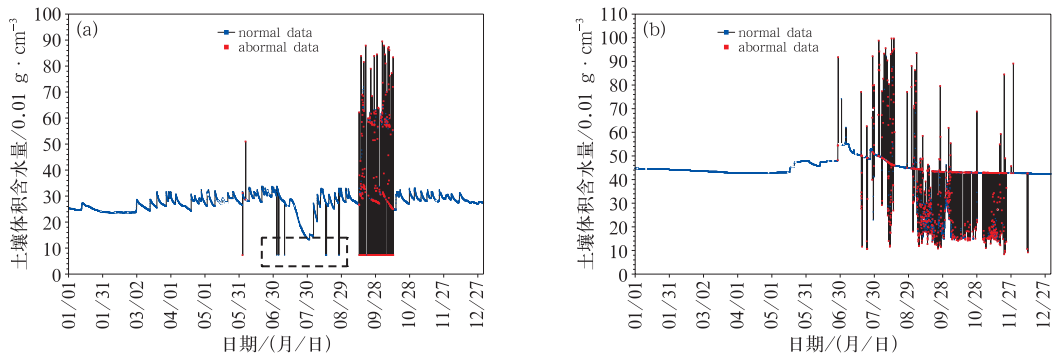


图 6 逐小时正变率与上层变率关系检查的异常个例:

- (a)重庆龙洲湾站(A8802)站 2014 年第二层土壤,
- (b)吉林和龙西南站(E0121)站 2014 年第十层土壤

Fig. 6 The 2nd layer soil abnormal case of AICqq method on ASMOS data: Longzhouwan Station, Chongqing (code A8802) in 2014 (a), the 10th layer soil abnormal case of Helong Southwest Station, Jilin (code E0121) in 2014 (b)

表 2 异常减小检查的参数阈值及质量控制结果

Table 2 Threshold value and QC results of ADC method

编号	$RH(t-1)$ 取值范围 /%	$RH(t)$ 取值 范围/%	$\Delta Q(t)$ 取值 范围/ $0.01 \text{ g} \cdot \text{cm}^{-3}$	QC 码 取值
1	$\geq 95$	$< 90$	$< -2.0$	2
2	(50, 95)	—	$[-1.0, -0.5)$	1
			$< -1.0$	2
3	$\leq 50$	—	$< -0.5$	2

上或向下的跳跃数据为“钉子数据”。另外,图 7c 有一段时间内并未检查出异常数据,经查,该段时间内的土壤相对湿度全部缺测,因此不满足本项检查的条件。

### 2.5 异常恒定检查

异常恒定检查主要是判断连续多个时次  $Q$  值保持不变的数据是否为异常的方法。通过分析全国

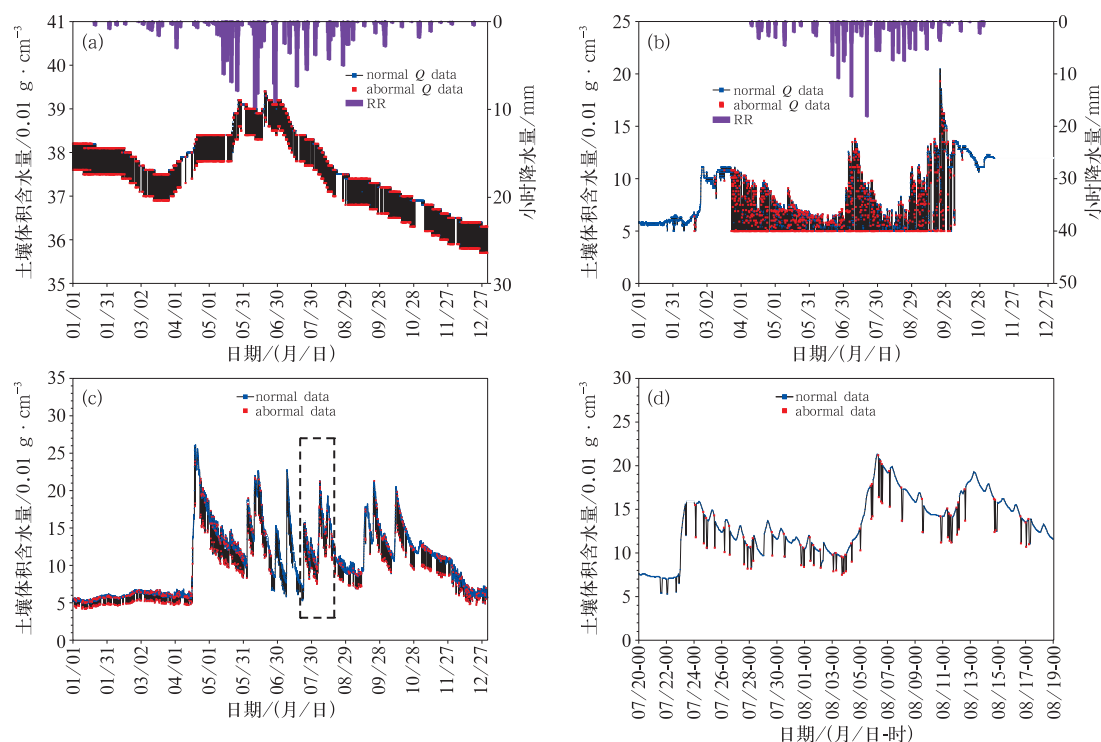


图7 逐小时负变率与土壤相对湿度关系检查结果逐小时序列:

(a) 吉林和龙站(54286)2014年第十层土壤逐小时,(b)山西河曲固定站(53564)

2014年第二层土壤,(c)青海民和站(52876)2014年第一层土壤,

(d)青海民和站(52876)2014年7月20日至8月19日第一层土壤

Fig. 7 The 10th layer soil abnormal case of ADC method on ASMOS data:

Helong Station, Jilin (code 54286) in 2014 (a), the 2nd layer soil case of Hequ Fixed Station,

Shanxi (code 53564) in 2014 (b), the 1st layer soil case of Minhe Station,

Qinghai (code 52876) in 2014 (c) and from 20 July to 19 August 2014 (d)

2014年实际观测数据,发现存在两类明显异常恒定数据:(1)当 $Q$ 异常跳跃(增大或减小)时,在其开始前或结束后的一段时间内常常出现恒定不变的僵值,如图2b所示;(2)当目标层与相邻层土壤的 $RH$ 在垂直方向存在较大差值,且目标层 $Q$ 值在连续多个时次出现恒定不变时即为异常恒定数据。针对上述两类异常恒定数据,本节分别设计了第一类恒定值检查ACC1和第二类恒定值检查ACC2方法。

以某台站第 $L$ 层土壤为例,假设 $t$ 为质量控制时次,当 $\Delta Q(L, t) \neq 0.0$ 时则需进行本项检查。从质量控制时次的前一时次( $t-1$ )开始向前推,若连续多个时次 $Q$ 保持恒定不变,令 $N$ 为连续时次总数,那么 $t-N$ 至 $t-1$ 时段范围内的观测数据即为本项检查的数据对象。

### 2.5.1 第一类恒定值检查

本项检查以界限值检查(AECtrh)、异常增大检查(AIC)、异常减小检查(ADC)的结果为基础,分为

以下两步进行检查:

当 $Q(L, t-N)$ 或 $Q(L, t-1)$ 对应QC码=2时,则 $t-N$ 至 $t-1$ 时段范围内的 $Q$ 均判为错误, QC码=2;

当 $Q(L, t-N)$ 或 $Q(L, t-1)$ 对应QC码=1时,则 $t-N$ 至 $t-1$ 时段范围内的 $Q$ 均判为可疑, QC码=1。

图2b即为本项检查的效果示例,图中CD、EF、GH各段恒定值均可被有效判定。

### 2.5.2 第二类恒定值检查

为剔除该类异常恒定数据,需分别设定连续不变时次数 $N$ 以及相邻层土壤相对湿度最大绝对差值 $\Delta RH_{\max}$ (计算方法见附录A.5)的上限值。本节根据 $RH(L, t-1)$ 的不同取值范围,分别给出参数 $N$ 及 $\Delta RH_{\max}$ 的上限值,详见表3。表中各参数适用于各层土壤,其统计方法详见附录A.6。当 $RH(L, t-1)$ 、 $\Delta RH_{\max}$ 及 $N$ 同时满足表3任意一组条件时,



$t-N$  至  $t-1$  时段范围内质量控制层( $L$  层)土壤体  
积含水量均判为可疑, QC 码=1。

表 3 恒定值异常检查方法 2 的参数取值

Table 3 Parameters for ACC2 method

编号	$RH(L, t-1)$ 取值范围/%	$\Delta RH_{\max}$ 取 值范围/%	$N$ 取值 范围/h	QC 码 取值
1	(0, 50]	>10	$\geq 60$	1
2	(50, 100]	>20	$\geq 60$	1
3	(100, $+\infty$ )	>10	$\geq 60$	1

图 8 为广东番禺站(59481)2014 年第二、三、四  
层小时序列值,其中第三层为质量控制层,第二和四  
层为参考层。图中标记为红色的数据为本项检查判  
定的可疑恒定数据,由图可以看到:土壤含水量相对  
较少的第三层与土壤含水量相对较高的第二层和第  
四层相邻,当第三层土壤的水分含量连续 60 h 及以  
上保持不变时,本方法即判定为可疑数据。

3 质量控制方法效果检验

本节采用 2015 年全国观测数据以及 CLDAS  
数据检验质量控制方法的效果,同时,给出 2014 年  
全国观测数据的相应 QC 结果作为补充对比。由于  
AECst 方法比较成熟且在国际上已得到广泛应用,  
本节仅对除 AECst 之外的其他方法进行检验。此  
外,为排除我国南北台站土壤冻结、解冻现象的不  
一致带来的干扰,本节仅针对 4 月下旬至 10 月上  
旬(后文简称检验期)的数据进行分析。

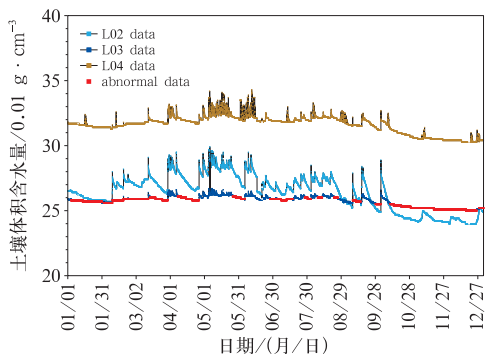


图 8 2014 年广东番禺站(59481)第二、三、四层  
小时序列值 ACC2 检查的异常个例  
(第三层为质量控制层)

Fig. 8 The 2nd, 3rd, 4th layer hourly abnormal  
cases of ACC2 method on ASMOS data:  
Panyu Station, Guangdong (code 59481) in 2014  
(The 3rd layer is QC layer)

3.1 观测资料空间分布检验

当观测仪器在 2014—2015 年时段内未做调整  
时,其数据质量应存在一定的连续性,表现为台站疑  
误率在全国区域内的空间分布具有一致性。另一方  
面,“钉子数据”常表现为异常增大与异常减小的成  
对出现,偶尔还伴随异常恒定数据的出现,使得各质  
量控制方法对应的疑误率空间分布也存在一定的一  
致性。因此,本小节采用上述两种空间分布一致性,  
来检验质量控制方法的可信度。疑误率计算公式见  
式(6),当疑误数据仅包含错误数据或可疑数据时,  
疑误率即指错误率或可疑率。

疑误率 = 疑误数据量 / 非缺测数据总量  $\times 100\%$   
(6)

第一种空间分布一致性检验。首先,根据单站  
疑误率对 1175 个台站进行由小到大排序,并记  
P80、P90、P95 分别为排位第 80%、90%、95% 台  
站的疑误率。其次,对比分析单站疑误率分别落在  
(0, 0, P80]、(P80, P90]、(P90, P95]、(P95, 100%]  
四个范围内的台站在全国的空间分布,结果如图 9  
所示。对比图 9a 和 9b,可发现 2014 与 2015 年的  
逐站疑误率分布具有较好的一致性,疑误率大值区  
域均在贵州、四川、安徽、河南等地。

第二种空间分布一致性检验。由于第一类恒定  
值检查(ACC1)与界限值检查(AECtrh)、异常增大  
检查(AIC)、异常减小检查(ADC)有关,因此将  
AECtrh、AIC、ADC 三种方法的质量控制结果进行  
合并,通过对比 ACC1 方法应用前后疑误率的增加  
值来检验 ACC1 方法。本小节统计了 2015 年各质  
量控制方法对应的单站疑误率空间分布,如图 10 所  
示(2014 年结果类似,图略),对比图 10a~10f,可  
发现各质量控制方法对应的疑误率空间分布一致,  
其中与“钉子数据”联系最紧密的 AICrr、AICqq、ADC  
对应的疑误率分布最为一致(图 10b~10d)。

3.2 与 CLDAS 对比检验

CLDAS 土壤体积含水量数据集产品(V2.0)共  
包含 0~5、0~10、10~40、40~80 和 80~200 cm 共  
5 层土壤体积含水量。其中,仅有 0~10 cm 与观测  
资料的土壤深度一致,因此,本小节仅选取 0~  
10 cm 层的数据用来检验分析。

首先,以 CLDAS 数据作为基准值,利用双线性  
插值方法将 CLDAS 数据插值到观测站点,进而得

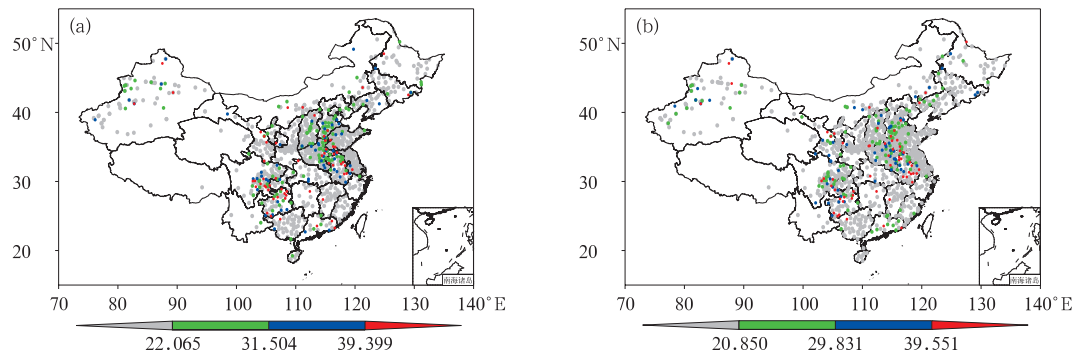


图9 2014年(a)和2015年(b)质量控制结果逐站疑误率(单位:%)空间分布

Fig. 9 Distribution of doubt/error rates of the QC result over China in 2014 (a) and 2015 (b) (unit: %)

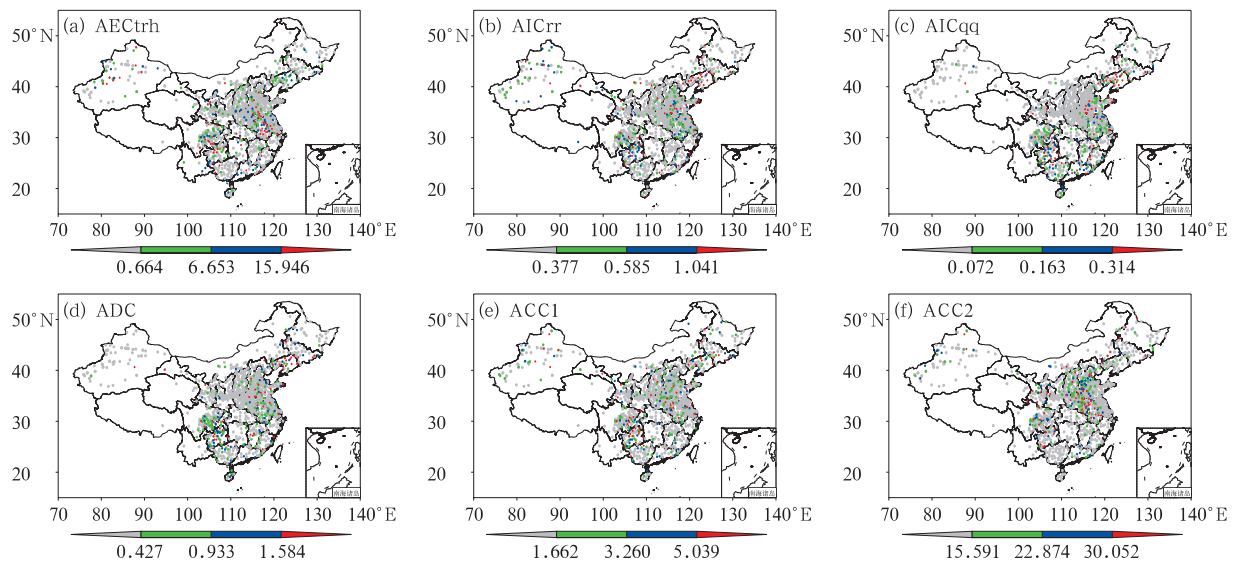


图10 各检查方法逐站疑误率(单位:%)全国区域分布

Fig. 10 Distribution of doubt/error rates by each method over China (unit: %)

到各台站对应的基准值序列。其次,分别计算 QC 前后各检验台站观测序列与基准值序列的差值序列。最后,选用平均偏差(MBE)和均方根误差(RMSE)作为误差指标(刘梦娟和刘舜,2016;赵煜飞和朱亚妮,2017),分别计算两组差值序列的 MBE 和 RMSE,以此对质量控制效果进行检验。因个别台站观测数据的缺测率较高,在计算 MBE、RMSE 过程中剔除了部分台站(2014 年剔除 33 个、2015 年剔除 103 个)。

2014 和 2015 年 QC 前后的 MBE、RMSE 以及相应的差值如图 11 所示,结果表明:(1)单站疑误率越大, QC 前后 MBE、RMSE 的差异越明显,见图 11a 和 11c;(2)QC 前后 MBE 的差值有正有负, RMSE 的差值则以负值为主,见图 11b 和 11d。一

般来说,疑误数据的出现具有一定随机性,因此,剔除疑误数据仅会减小 RMSE。本次检查的结果符合这一特征,说明本文提出的质量控制方法能有效判定疑误数据。

## 4 结论与讨论

为剔除土壤湿度资料中的异常数据,本文提出了一套适用于全国自动土壤水分观测资料的质量控制方法。首先,根据 2014 年全国观测资料中的异常数据特征将异常数据分为四类;其次,针对每类分别设计了质量控制方法:异常极值检查、异常增大检查、异常减小检查、异常恒定检查。其中,异常极值检查直接采用 Dorigo et al(2013)以及王良宇和



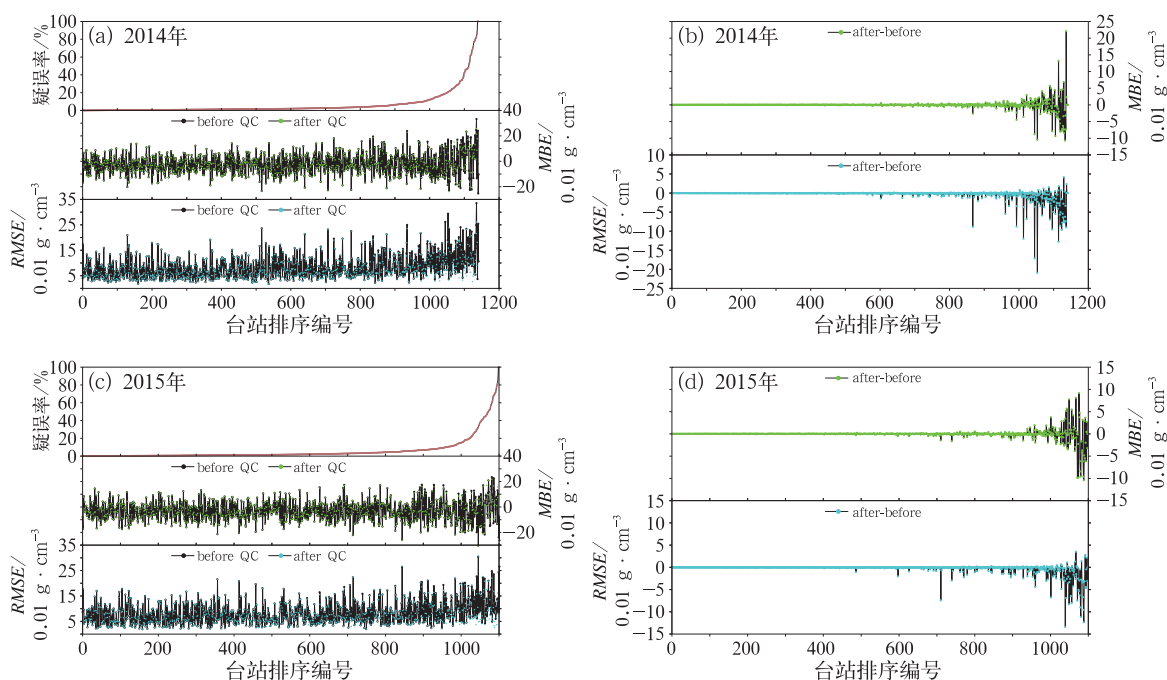


图 11 2014 年(a, b)和 2015 年(c, d)QC 前(a, c)和 QC 后(b, d)的 MBE、RMSE 以及相应的差值随疑误率的散点图

Fig. 11 The scatter plots of MBE, RMSE and corresponding difference of doubt/error rates before QC (a, c) and after QC (b, d) in 2014 (a, b) and 2015 (c, d)

何延波(2015)提出的相关方案;在异常增大检查中,设计了  $\Delta Q$  与降水关系检查、 $\Delta Q$  与上层变率关系检查;在异常减小检查中,设计了  $\Delta Q$  与土壤相对湿度关系检查;在异常恒定检查中,分别设计了第一类恒定值检查与第二类恒定值检查。

最后,利用 2014 和 2015 年全国土壤水分观测资料以及 CLDAS 数据集产品(V2.0)对质量控制方法的应用效果进行检验,结果表明:

(1) 四类检查方法均可判断出自动土壤水分观测资料中相应的疑误数据;

(2) 2014 年与 2015 年质量控制结果疑误率空间分布一致,各质量控制方法对应的疑误率空间分布也一致,说明质量控制方法可信;

(3) 利用本文质量控制方法对观测数据进行 QC,可明显减小观测数据与 CLDAS 数据的 RMSE,说明质量控制方法有效。

目前,该方案已应用于我国气象资料处理业务系统。此外,本文设计的质量控制方法仍存在不足。

(1) 在与降水的关系检查中,当有灌溉事件发生时,可能产生误判。

目前,气象观测业务上还没有灌溉的实时监测数据,因此,在设计异常增大检查时,本文尚未考虑

由灌溉引起的土壤水分突然增大。由于异常增大检查主要是针对跳跃数据进行判定,所以,灌溉事件仅会引起当前时次及前一时次数据的误判,不会影响灌溉事件之后的各时次数据。未来,若个别台站有灌溉数据,可将其视为相应时次的降水量,利用本方法进行判定。

(2) 存在异常数据的漏判及误判。

本文提出的四类异常数据类型并不能包含目前观测资料中所有的异常数据,因此,存在异常数据的漏判问题。进一步分析资料中异常数据的类型,可设计相应的新检查方案进行判断。此外,本文各方法采用的参数值均针对全国台站设定,对于个别台站可能会产生漏判或误判。因而,将本文方法应用于个别台站时,可根据该台站的实际数据确定合适的参数值。

**致谢:**余予博士在画图方面提供的帮助,李艺苑博士对本文提出的修改意见,师春香研究员、姜志伟博士在 CLDAS 资料处理方面提供的帮助,在此一并表示感谢!

## 参考文献

陈海波,2013. DZN2 型自动土壤水分观测仪常见问题分析[J]. 气象与环境科学,36(3):54-57.

- 耿增超,戴伟,2011. 土壤学[M]. 北京:科学出版社:79-102.
- 郭维栋,马柱国,王会军,2007. 土壤湿度——一个跨季度降水预测中的重要因子及其应用探讨[J]. 气候与环境研究,12(1):20-28.
- 刘梦娟,刘舜,2016. 上海组网风廓线雷达数据质量评估[J]. 气象,42(8):962-970.
- 刘宪锋,朱秀芳,潘耀忠,等,2015. 农业干旱监测研究进展与展望[J]. 地理学报,70(11):1835-1848.
- 马柱国,符淙斌,谢力,等,2001. 土壤湿度和气候变化关系研究中的某些问题[J]. 地球科学进展,16(4):563-568.
- 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.
- 孙小龙,宋海清,李平,等,2015. 基于 CLDAS 资料的内蒙古干旱监测分析[J]. 气象,41(10):1245-1252.
- 王良宇,何延波,2015. 自动土壤水分观测数据异常值阈值研究[J]. 气象,41(8):1017-1022.
- 王万秋,1991. 土壤温湿异常对短期气候影响的数值模拟试验[J]. 大气科学,15(5):115-123.
- 王晓东,杨太明,吴必文,等,2015. 安徽省自动观测土壤水分质量控制方法[J]. 气象科技,43(3):399-404.
- 吴东丽,曹婷婷,薛红喜,2016. 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,4(1):1-10.
- 吴东丽,梁海河,曹婷婷,等,2014. 中国自动土壤水分观测网运行监控系统建设[J]. 气象科技,42(2):278-282.
- 元保军,2014. 土壤水分自动观测数据质量控制与评估系统阈值研究方法[J]. 电子设计工程,22(2):16-18.
- 张喜英,罗敏,李德华,2015. 常德地区土壤水分自动站数据质量控制
- 的阈值确定[J]. 安徽农业科学,43(28):109-111.
- 张志富,2013. 自动站土壤水分资料质量控制方案的研制[J]. 干旱区地理,36(1):101-108.
- 赵煜飞,朱亚妮,2017. 中国地面均一化相对湿度月值格点数据集的建立[J]. 气象,43(3):333-340.
- 中国气象局,2010a. QX/T 118-2010 地面气象观测资料质量控制[S]. 北京:气象出版社.
- 中国气象局,2010b. 自动土壤水分观测规范(试行)[R]. 北京:中国气象局.
- Davis J L, Chudobiak W J, 1975. In situ meter for measuring relative permittivity of soil[R]. Ottawa: Geological Survey of Canada: 75-79.
- Dorigo W A, Xaver A, Vreugdenhil M, et al, 2013. Global automated quality control of in situ soil moisture data from the international soil moisture network[J]. Vadose Zone J, 12(3). DOI: 10.2136/vzj2012.0097.
- Entekhabi D, Rodriguez-Iturbe I, Castelli F, 1996. Mutual interaction of soil moisture state and atmospheric processes[J]. J Hydrol, 184(1-2):3-17.
- WMO, 2008. Guide to meteorological instruments and methods of observation[R]. WMO-No. 8, Geneva, Switzerland: WMO.
- Xia Youlong, Ford T W, Wu Yihua, et al, 2015. Automated quality control of in situ soil moisture from the North American soil moisture database using NLDAS-2 products[J]. J Appl Meteor Climatol, 54(6):1267-1282.

## 附录 A:质量控制方法各参数设定说明

### A.1 与降水关系检查之参数 $m_1$ 设定

参数  $m_1$  主要与各层土壤对降水事件的响应时间有关:假设  $t_0$  时次为连续降水的开始时刻,  $t$  时次( $t > t_0$ )为  $L$  层土壤  $Q(L, t) > 2.0$  首次出现的时刻,  $t_1$  时次为  $t_0$  至  $t$  时段内的最大小时降水量首次出现的时刻,那么  $L$  层土壤对降水的相应时间介于  $t - t_1$  至  $t - t_0$  之间( $t_0 < t_1 < t$ )。对于各层土壤参数  $m_1$  的选取,本文以 2014 年全国观测数据为基础进行统计,初步确定参数  $m_1$  的取值范围;其次,在其取值范围内设定不同的取值,并将其应用于 2014 年全国数据中,针对各参数对应的检查结果,通过人工抽查的方式进行核查,最终选择漏判率及误判率均相对较低的取值,详见附表 A1。

表 A1 参数  $m_1$  取值说明

Table A1 The value of parameter  $m_1$

土壤层 编号	土壤层深度/ cm	$m_1$ 取值 范围	本文 $m_1$ 取值
1	0~10	[1, 6]	1
2	10~20	[2, 7]	2
3	20~30	[2, 8]	3
4	30~40	[2, 8]	4
5	40~50	[3, 8]	5
6	50~60	[3, 8]	6
7	70~80	[3, 9]	8
8	90~100	[4, 10]	10

### A.2 与降水关系检查之可疑数据判定的参数设定

#### (1) $RR(t) = 0$ 时 $Q_a$ 设定

对于表层土壤,本文以 2014 年全国范围内  $\Delta Q(t) > 0$  且  $RR(t)$  不为缺测的数据为样本(共约 60 万个样本),并设定  $\Delta Q(t)$  步长为 0.5,  $RR(t)$  步长为 1,对  $\Delta Q(t)$  不同取值条件下  $RR(t)$  取不同值的频率进行统计,结果如图 A1 所示。由图可以看到  $\Delta Q(t) > 0.5$  时,对应降水事件( $RR > 0$ )发生的频率则较高。同样,通过人工抽查方式对  $\Delta Q(t) > 0.5$  且  $RR(t) = 0$  的样本进行了核查,发现主要以异常数据为主,因此设定  $Q_a = 0.5$ 。

#### (2) $0 < RR(t) \leq 2.0$ 时 $Q_a$ 设定

当降水量处于  $(0, 2.0]$  范围内时,此时本文将  $m_1$  取值设定为 2,即:  $Q_a = RR(t-1) + RR(t)$ 。同时,对参数  $Q_a$  的应用效果进行了人工核查,结果表明该方案可有效识别异常数据。

### A.3 逐小时正变率与上层变率关系检查之参数 $\alpha$ 设定

关于参数  $\alpha$  的调试。本文从 1.0~3.0 每隔 0.2 设定一个  $\alpha$  值,并将该方案应用于 2014 年全国数据,经统计发现当

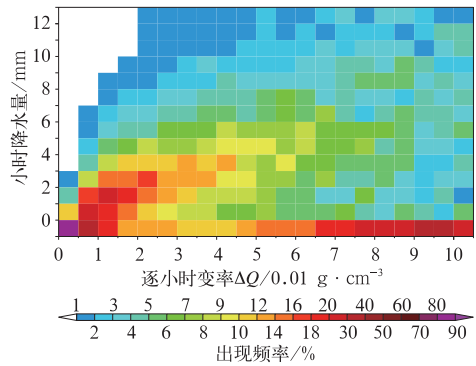


图 A1 不同  $\Delta Q$  取值条件下  $RR$  不同取值的出现频率( $RR < 0$  表示无降水情况)  
Fig. A1 Frequency of different values of  $RR$  under different  $\Delta Q$  values ( $RR < 0$  represents no precipitation)

$\alpha > 2.0$  之后疑误数据量的相对变化比较平稳,如图 A2 所示。其次,对  $\alpha$  取 2.0 附近数值时的检查结果进行对比和人工核查,最终选择漏判率及误判率均相对较小的  $\alpha$  数值,本文取值为 2。

### A.4 异常减小检查的参数设定

针对每层土壤,本文以 2014 年全国范围内  $\Delta Q(t) < 0$  且  $RR(t-1)$  不为缺测的数据为样本,并设定  $\Delta Q(t)$  步长为 0.25,  $RH(t-1)$  步长为 10,对  $RH(t-1)$  不同取值条件下  $\Delta Q(t)$  取不同值的频率进行统计。以第一层土壤的统计结果为例,如附图 A3 所示(约 103 万个样本),其他各层结果类似(图略),由图可以看到土壤相对湿度在  $< 50\%$ 、 $50\% \sim 90\%$ 、 $> 90\%$  三种情况下,逐小时负变率  $\Delta Q(t)$  具有不同的分布规律。在此基础上,本文采用  $RH(t)$ 、 $RH(t-1)$  的数值大小将土壤干湿状况分为三种,并以 0.25 为步长对逐小时负变率  $\Delta Q(t)$  下限值进行逐步调试。利用人工核查方法,通过对检查结果的对比分析,最后选择漏判率及误判率均相对较小的  $\Delta Q(t)$  的下限值。

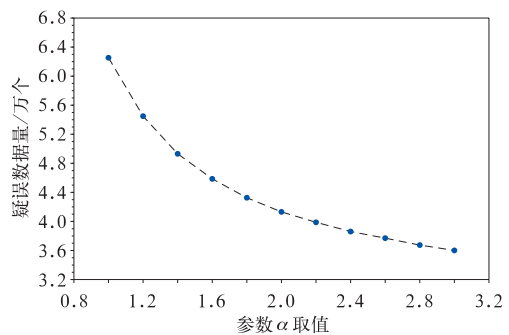


图 A2 参数  $\alpha$  与疑误数据量关系  
Fig. A2 Relation between parameter  $\alpha$  and the amount of doubt/error data

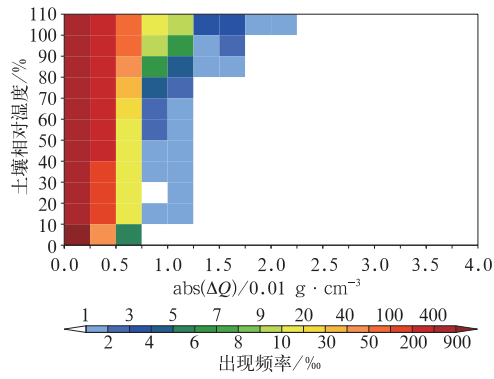


图 A3 不同  $RH$  取值条件下逐小时  
负变率  $\Delta Q$  不同取值的出现频率

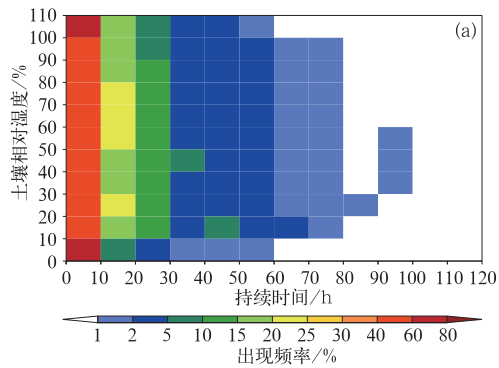
Fig. A3 Frequency of different values of  
 $\Delta Q$  under different  $RH$  values

#### A.5 第二类恒定值检查之参数 $\Delta RH_{\max}$ 计算方法

依据第二类恒定值检查方法的相关要求,记  $t-N$  至  $t-1$  时次范围内  $L-1$  层及  $L+1$  层土壤相对湿度数据序列为  $RH_i^{\text{up}}, RH_i^{\text{dn}} (i=1, 2, \dots, N)$ 。对于序列  $RH_i^{\text{up}}$ ,首先对其进行由小到大排序;其次,令  $m=[N/2]$ ,其中符号  $[\ ]$  表示向下取整;那么,令  $A_{\text{up}}$  为序列  $RH_j^{\text{up}} (j=1, 2, \dots, m)$  的算术平均值,令  $B_{\text{up}}$  为序列  $RH_j^{\text{up}} (j=m+1, m+2, \dots, N)$  的算术平均值,计算公式如下:

$$A_{\text{up}} = \sum_{i=1}^m RH_i^{\text{up}} / m \quad (\text{附 } 1)$$

$$B_{\text{up}} = \sum_{i=m+1}^N RH_i^{\text{up}} / (N - m) \quad (\text{附 } 2)$$



同样,对于序列  $RH_i^{\text{dn}}$  可以得到  $A_{\text{dn}}$  和  $B_{\text{dn}}$ 。 $\Delta RH_{\max}$  即为  $RH(L, t-1)$  与数组  $[A_{\text{up}}, B_{\text{up}}, A_{\text{dn}}, B_{\text{dn}}]$  的差值绝对值的最大值,计算公式如下:

$$\Delta RH_{\max} = \max(|RH(L, t-1) - [A_{\text{up}}, B_{\text{up}}, A_{\text{dn}}, B_{\text{dn}}]|) \quad (\text{附 } 3)$$

#### A.6 第二类恒定值检查的参数设定

关于土壤体积含水量连续保持不变的小时数(记为  $N$ )与土壤相对湿度的分布规律统计,本文选取 2014 年全国范围内  $N>1$  且  $RH$  不为缺测的数据为样本(约 22 万个样本),对  $RH$  不同取值条件下  $N$  取不同值的频率进行统计( $RH$  步长为 10,  $N$  步长为 10)。对于深层土壤,土壤湿度一般比较稳定,其连续不变的时次数一般较长,以第六层土壤的统计结果为例,如图 A4a 所示。同样,关于相邻层土壤相对湿度的分布规律统计,本文以某层土壤  $RH(L)$  及其与上层土壤  $RH(L-1)$  的绝对差值  $\text{abs}[RH(L) - RH(L-1)]$  为对象,记为  $\Delta RH'$ 。选取 2014 年全国范围内  $\Delta RH'$ 、 $RH(L)$  不为缺测的数据为样本[若连续多个时次  $RH(L)$  保持不变,则只记为一个样本,共选取约 28 万个样本],对  $RH(L)$  不同取值条件下  $\Delta RH'$  取不同值的频率进行统计[ $RH(L)$  步长为 10,  $\Delta RH'$  步长为 2.5]。也以第六层土壤的统计结果为例,如图 A4b 所示。

在此基础上,本文根据  $RH(L)$  大小将土壤干湿状况分为三种: $<50\%$ 、 $50\% \sim 100\%$ 、 $>100\%$ ,并在每种情况下对土壤相对湿度差值及连续时次数进行逐步调试,通过人工抽查方法对质控结果进行对比分析,最后选择漏判率及误判率均相对较小的参数取值,进而建立本项检查。

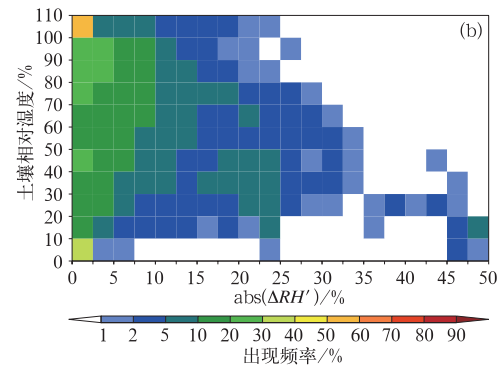


图 A4 土壤湿度特征的统计分析

(a)土壤相对湿度持续不变时次数的频率统计,(b)土壤相对湿度上下层差值的频率统计

Fig. A4 Statistical analysis of soil moisture characteristics frequency of the number  
of hours in which  $RH$  remains constant (a), frequency of  $RH$  difference  
between upper and lower layers (b)