

地面自动站观测资料三级质量控制业务系统的研制

任芝花^{1,2} 熊安元²

(1. 北京大学物理学院, 北京 100871; 2. 国家气象信息中心)

提 要: 随着我国气象台站地面自动气象观测系统的逐步布设, 地面自动站观测资料质量控制系统的研制亦成为业务急需。2004年开始, 研制由台站, 到省级、国家级资料部门的地面自动站观测资料三级质量控制业务系统。各级质量控制系统融合了自动控制技术和交互式应用技术, 允许在必要时对疑误资料进行详细的人工分析、判断与修正。文章介绍了整个质量控制系统的业务流程设置、质量控制技术以及质量控制信息。

关键词: 地面自动站观测资料 质量控制系统 三级

Operational System Development on Three-step Quality Control of Observations from AWS

Ren Zhihua^{1,2} Xiong Anyuan²

(1. Physics of School, Peking University, Beijing 100871; 2. National Meteorological Information Center)

Abstract: With the increasing of Automatic Weather Stations (AWS) of China Meteorological Administration (CMA), it is very vital for meteorological operation to develop and set up a quality control system of observations from AWS. As a quality control flow from observational station to provincial meteorological data department, subsequently to national data department, the three-step quality control system in operation for observations from AWS has been studied since 2004. Each step has automatic control technique and interface application technique, allowing analyzing and amending some data. The three-step quality control system could improve the collection and disposition level of data observed at AWS and could insure the data quality. The operational flow setting in the whole quality control system, the quality control technique and the quality control information have been presented.

资助项目: 中国气象局新技术推广项目(CMATG2005M10)

收稿日期: 2006年6月22日; 修定稿日期: 2006年11月15日

Key Words: data measured at AWS quality control system three-step

引 言

数值预报技术的发展以及全球气候与大气环境的深入研究,对气象观测资料的质量要求越来越高。针对不同的气象资料,国内由此而研制了许多质量控制(QC)系统^[1-4]。随着地面观测自动化技术的发展,西方发达国家纷纷研制了地面自动观测资料质量控制系统。北欧5国(丹麦、芬兰、冰岛、挪威、瑞典)对自动站资料控制流程分4级:台站级资料质量控制、入库前实时资料质量控制、入库后非实时资料质量控制以及人工质量控制^[5]。美国由1100多个观测站组成的地面自动观测系统,在业务运行中,通过对资料进行台站级、州级、国家级三级质量控制,及时发现并反馈观测资料的质量问题,保证资料的准确性^[6]。截至2005年底,我国气象部门约有600多个国家级地面站、1200多个地面一般站已使用自动气象站进行地面气象观测。观测系统的变化使得观测仪器、观测方法、观测资料的密度及信息量、数据文件格式等均发生了巨大改变。随着中国气象台站地面自动气象观测系统的逐步布设,地面自动站观测资料质量控制系统的研制已成为业务急需。2004年,由国家气象信息中心牵头,联合安徽省气象局、湖北省气象局开始研制由气象台站,到省级、国家级资料部门的地面自动站观测资料三级质量控制业务系统(以下简称“三级质量控制系统”),制定各级数据质量控制方案,开发相应的业务适用软件。目前,该系统的研制工作已基本完成。各级质量控制系统融合了自动控制技术和交互式应用技术,提高了自动进程能力,允许在必要时对特殊资料进行详细的人工分析判断与修正。下文重点介绍三级质量控制系统业务流

程、有关质量控制技术以及质量控制信息。

1 三级质量控制系统业务流程

三级质量控制系统业务流程如图1所示。该流程的建立以现行业务观测体制中常规资料的采集传输流程为依据,包括台站级QC、省级QC和国家级QC。图中台站级QC是指台站在实时资料采集、质量控制的基础上,对本站形成的A、J、R月数据文件及年报文件Y文件,利用台站级质量控制软件进行质量检查和预审,认为合格后上报到省级资料处理部门;省级QC是指在每月规定时间内收集本省范围内由台站上传的上述数据文件,利用省级质量控制软件对其进行质量检查和人机交互审核,同时向台站反馈审核信息,经复查合格后,按要求上报至国家气象信息中心;国家级QC是指在每月规定时段内收集各省上传来的数据文件,利用国家级质量控制软件对其进行滚动格式检查和质量检查,利用人机交互界面进行数据检查处理,同时向各省反馈最终质量控制信息,确认合格后,归档存储到国家气象信息中心气象资料室。

(1)台站级质量控制的对象包括自动站实时采集数据、自动站采集数据文件(包括正点数据文件和分钟数据文件)和在前者基础上形成的地面A、J、R月数据文件及年报文件Y文件。台站级质量控制软件既可作为完整的独立软件,也可作为地面气象观测系统业务软件的组成部分。



图1 三级质量控制系统流程图

自动站实时采集数据的质量控制包括下列检查:

① 数据采集时间的合法性检验;② 要

素气候极值或允许范围值检查,当超出范围值时,相应数据按缺测处理;③ 字符的合法性检查,为非数据、缺测字符时,相应数据按缺测处理;④ 要素内部一致性检查,如:辐射、散射辐射、水平面直接辐射间的一致性检查;⑤ 极值出现时间与采集时间一致性、极值与正点值一致性的检查与处理;⑥ 利用各采集要素分钟曲线变化图,进行人工实时检测。

自动站实时采集数据在上述质量控制的基础上形成自动站采集数据文件。对自动站采集数据文件的质量控制包括:

① 数据文件的格式检查;② 要素气候界限值或允许范围值检查;③ 台站气候极值检查;④ 要素内部一致性检查;⑤ 要素时间一致性检查,包括要素分钟、小时、24 小时变化率检查;⑥ 利用要素日变化曲线图,进行人工分析与判断。

地面 A、J、R 月数据文件及年报文件 Y 文件的质量控制过程与自动站采集数据文件的质量控制过程类似,只是在人机交互界面,显示疑误信息、疑误数据及其与之有关的要素数据,便于审核人员分析判断。

(2) 省级和国家级质量控制的对象仅为地面 A、J、R 月数据文件及年报文件 Y 文件。所用的质量控制技术,除了台站级所涉及到的质量检查外,还有台站参数检查以及空间一致性检查。人机交互界面中除了显示疑误信息、疑误数据及其与之有关的要素数据外,还可在同一坐标系中显示多要素时间变化曲线图以及要素空间数值图。省级质量控制中,空间一致性检查仅涉及到本省内的所有台站,台站密度高,但地域范围小。国家级质量控制中空间一致性检查仅涉及全国所有的基准基本站,台站密度低,但范围大,可跨省、市检查。

由上述可见,台站级软件从数据采集开始到上报文件形成前,各个环节都有质量检测。地面 A、J、R 月数据文件及年报文件 Y 文件在各级质量控制中,各级所用的质量控制技术类似,但是在具体操作细节上有许多

不同,在质量控制参数(如相邻层间的地温差值控制参数)的设置上台站级和省级很细,细化到不同的站可能有不同的控制参数,而国家级则可能全国范围内用统一较宽的参数控制。因此,A、J、R、Y 文件在各级质量控制中,台站级检出的疑误信息最多,其次是省级,国家级检出的疑误信息最少,这符合把错误资料消灭在萌芽时期的精神。

2 质量控制技术

三级质量控制系统中,质量控制部分为整个系统的核心。国家级和省级的质量控制流程见图 2,台站级质量控制流程与图 2 类似,但无空间一致性检查。图 2 中,从格式检查到空间一致性检查的各个环节都为软件自动控制,人机交互检查为人工借助软件交互界面提供的多种信息对资料进行分析判断或对数据进行修正。

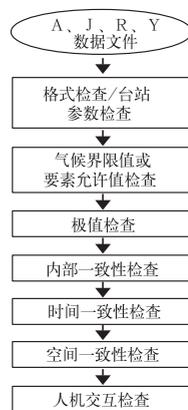


图 2 质量控制流程图

三级质量控制系统所涉及到的质量控制技术介绍如下。

(1) 格式检查

格式检查为对数据进行质量检查前的预处理检查。主要检查文件中的数据录入格式及数据范围是否与《地面气象观测数据文件格式和记录簿表格式》中规定的文件格式一致^[7]。另外进行台站参数中区站号、年、月与

文件名中的区站号、年、月的一致性检查,检查文件中的其它台站参数如经纬度、海拔高度等是否与台站参数表一致。

(2) 气候界限值或要素允许值检查

从气候的角度不可能出现的要素临界值称之为要素的气候界限值。资料形成时规定采用的表示方式和范围,为要素的允许值范围。超出要素的气候界限值或允许值范围的资料为错误资料。在三级质量控制系统中规定的部分常规要素允许范围如表 1 所示。允许值作为软件的外部数据,用户可根据具体情况修改。

表 1 各要素允许范围表

要素	允许范围
海平面气压	870~1100hPa
本站气压	400~1100hPa
气温	-75~80℃
地面温度	-90~90℃
露点温度	-70~40℃
水汽压	0~70hPa
相对湿度	0~100%
风速	0~75m·s ⁻¹
风向	十六方位或 0~360 度
云量	0~10 成
日降水量	0~1800mm
小时日照时数	0~1.0h
总辐射辐照度	0~2000W·m ⁻²
直接辐射辐照度	0~1408W·m ⁻²

(3) 台站气候极值检查

极值是指某个固定测站历史记录中某要素曾出现过的最大值(最小值),气象资料要素值是否超出极值的检查为极值检查。对超出测站要素极值的观测资料,必须做进一步检查,以判断资料正确与否。在三级质量控制系统中台站极值来自于建站至 2000 年的历史资料。业务上,在对三级质量控制系统的过程中,需根据实际情况对极值表进行更新。

(4) 内部一致性检查

对固定测站不同要素或项目之间是否符合某种物理关系的检查称之为内部一致性检查。

如:下面列出部分要素气象观测值必须符合的关系,若不符合时,则其中至少有一方为错误值。

(a) 日最低气压≤当日各时气压≤日最高气压;(b) 日最低气温≤当日各时气温≤日最高气温;(c) 日地面最低温度≤当日各时地面温度≤日地面最高温度;(d) 10 分钟平均风速≤日最大风速;(e) 总云量≥低云量;(f) 极大风速≥最大风速;(g) 极大风速≥17.0 m·s⁻¹时,应有大风现象;有大风现象时,极大风速≥17.0 m·s⁻¹;(h) 风向方位为“C”或角度为 0 度时,风速≤0.2 m·s⁻¹;(i) 散射辐射时总量≤总辐射时总量;(j) 反射辐射时总量≤总辐射时总量。

(5) 时间一致性检查

气象要素随时间的变化具有一定的规律。对气象资料变化是否符合这种规律的检查称之为时间一致性的检查。各要素资料不能超出一定时间内的变化范围,超出的资料为可疑资料。在三级质量控制系统中对器测数据进行了分钟变化率检查、小时变化率检查、日变化率检查以及连续 24 小时以上数据无变化检测。如:下列各要素必须符合以下时间变化率,否则,视为可疑。各种事件变化率参数作为外部数据,用户可根据具体情况修改。

(a) 气温 24 小时变温≤50℃;(b) 气压 24 小时变压≤50hPa;(c) 气温 1 小时变温≤8℃;(d) 气压 1 小时变压≤10hPa;(e) 相对湿度 1 小时变量≤50%;(f) 气温 1 分钟变温≤3℃;(g) 气压 1 小时变压≤3hPa;(h) 相对湿度 1 小时变量≤10%。

(6) 空间一致性检查

空间一致性检查是根据气象参数具有一定的空间分布特点而进行的检查。其有效性取决于观测站网的密度和被检参数与空间的相关程度。通常利用与被检查台站邻近的台站同一时间观测的气象要素值进行比较,或利用邻近测站观测值通过一定的插值方法计算出被检查台站的估计值,由观测值与估计值进行比较。下面介绍三级质量控制系统中所应用的空间质量控制方法,其有效性在所引用的文献中已作了较为详细的分析。

(a) Madsen-Allerupt 方法

Madsen-Allerupt 方法^[8]是对周围若干邻近站同一时刻某要素观测值进行由小到大排序,利用排序后的 $\frac{1}{4}$ 、 $\frac{2}{4}$ 、 $\frac{3}{4}$ 分位值,计算统计量 T_{it}

$$T_{it} = (x_{it} - q_{i,2/4}) / (q_{i,3/4} - q_{i,1/4}) \quad (1)$$

其中 x_{it} 为被检站 t 时刻某要素观测值, $q_{i,1/4}$ 、 $q_{i,2/4}$ 、 $q_{i,3/4}$ 为邻近站 t 时刻该要素观测值进行由小到大排序后的 $\frac{1}{4}$ 、 $\frac{2}{4}$ 、 $\frac{3}{4}$ 分位值。

当 $|T_{it}|$ 超过规定值时,认为被检站 t 时刻该要素观测值可疑。

(b) 空间回归检验法

在三级质量控制系统中,用空间回归检验法^[9]对日数据进行空间一致性检查。

首先,利用各站被检要素全月日数据,寻找与被检站相关系数较高的 5 个以上初步参考站。要求各参考站与被检站间的被检要素相关系数通过显著性检验。

其次,利用被检要素全月日数据逐对建立被检站与上述各参考站的一元线性回归方程:

$$\hat{x}_{i,j} = a_j + b_j y_{i,j} \quad (2)$$

其中, $y_{i,j}$ 为第 j 个初步参考站第 i 日要素实测值, $\hat{x}_{i,j}$ 为被检站第 i 日要素估计值。

最后,计算被检站全月要素观测值与各回归方程估计值间的均方根偏差 s_j^2 :

$$s_j^2 = \frac{1}{m-2} \sum_{i=1}^m (x_i - \hat{x}_{i,j})^2 \quad (3)$$

其中, x_i 为被检站第 i 日的实测值, m 为全月日数。

确定均方根偏差最小的 5 个相关站为被检站的最终参考站。

分别计算被检站被检要素第 i 日加权估计值 x_i' 及要素估计值的加权标准差 s' :

$$x_i' = \sqrt{\sum_{j=1}^n \hat{x}_{i,j}^2 s_j^{-2}} / \sqrt{\sum_{j=1}^n s_j^{-2}} \quad (4)$$

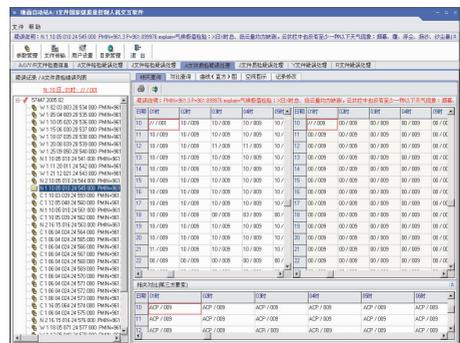
$$s' = \sqrt{n} / \sqrt{\sum_{j=1}^n s_j^2}$$

其中, j 为第 j 个最终参考站; n 为最终参考站的总数,在这里 $n=5$ 。

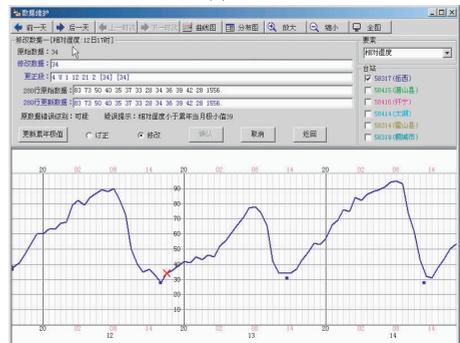
当 $|x_i - x_i'| > f_s'$ 时,表示被检站第 i 日的实测值 x_i 未通过空间一致性检查。 f_s' 为控制系数,取值范围为 3.0~5.0。

(7) 人机交互检查

完全自动化的质量控制不可能解决所有的数据质量问题,对于一些特殊情况或问题的判断还要辅助以人工检查。人工检查可以在任何一级的质量控制中进行,它是在上述自动控制后对可疑或错误的数据用人机交互的方式进行的质量控制。根据相关数据比较、图形分析、台站查询、邻站比较和人工分析判断,最终确定是否修改或订正数据并且确定要素值相应的质量控制码。三级质量控制系统提供了多种人机交互界面,有相关要素间的数据显示界面(图 3a)、单要素数据随时间变化曲线图或直方图界面(图 3b)、多种相关要素数据随时间变化的曲线图界面以及显示周围站同一时刻观测的数据的空间界面图等。人机交互界面为人工对疑误数据的分



(a)



(b)

图 3 三级质量控制系统中部分人机交互界面图

析判断提供了许多便利条件。各级质量控制系统中的人机交互界面均提供了对疑误数据和质量控制码的修改功能。

3 数据质量信息

在数据进行了质量控制后,对每个观测数据必须给出表示该数据质量状况的标识,保存数据质量的有关信息,以便数据处理者和用户参考。在月报文件中,均设置了保存观测数据质量信息的区间。数据质量信息包括质量控制码信息和数据更正信息。

质量控制码表示观测数据的质量状况,根据业务流程,将其分为三级:台站级、省级和国家级。质量控制码用三位整数表示,个位表示国家级,十位表示省级,百位表示台站级。各级质量控制码含义为:0:数据正确;1:数据可疑;2:数据错误;3:数据有订正值;4:数据已修改;8:数据缺测;9:数据未作质量控制。

在质量控制过程中,允许对有疑误或缺测的数据进行更正。在数据被更正的同时,系统自动记录了数据被更正的过程,包括更正级别(台站级、省级和国家级)、更正前的原始数据和更正后的数据等数据更正信息,以便资料处理时查询。数据更正包括数据订正和数据修改两种情况。数据订正是指原始观测数据疑误或缺测,通过一定的统计方法计算或估算的数据,该数据不替代观测数据中的原数据,只在数据更正信息中记录其订正状况。数据修改是指原始观测数据错误,经过查询确认正确的数据,该数据替代观测数据中的原数据,同时在数据更正信息中记录其修改状况。被更正后的数据质量控制码为 3 或 4。

4 结 语

由台站到省级、国家级资料部门的地面

自动站观测资料三级质量控制业务系统的研制工作已基本完成,目前处于试用与进一步完善阶段。该系统融合了自动控制技术和交互式应用技术,允许在必要时对特殊资料进行详细的人工分析判断与修正,自动记录质量控制信息。三级质量控制业务系统的研制适应了目前地面自动气象观测系统的业务需求,对提高我国自动观测站数据收集、处理水平,提高自动气象站观测资料的质量,进而提高自动站观测资料对天气预报、气候研究的服务能力,发挥自动站的整体效益,乃至对气象科学研究以及国计民生具有深远的意义。

致谢:参加该项目的还有国家气象信息中心的王新华、王颖、冯明农、刘小宁、孙化南、范邵华、鞠晓慧,安徽省气象局的吴必文、汪勇进、徐光清,湖北省气象局的张峻、杨志彪。在项目的研制过程中,局职能司的陈永清、刘海波、吴忠义等同志提供了许多建议与指导意见,在此谨表感谢。

参考文献

- [1] 方炳兴. 常规气象资料质量的综合控制[J]. 气象, 1994, 20(2): 33-36.
- [2] 杨贤为. 气候应用专用数据库气象资料质量检验[J]. 气象, 1998, 24(12): 33-36.
- [3] 陶士伟, 张跃堂, 陈卫红, 等. 全球观测资料质量监视评估[J]. 气象, 2006, 32(6): 53-58.
- [4] 任芝花, 刘小宁, 杨文霞. 极端异常气象资料的综合性质量控制与分析[J]. 气象学报, 2005, 63(4): 526-533.
- [5] 熊安元. 北欧气象观测资料的质量控制[J]. 气象科技, 2003, 31(5): 314-320.
- [6] 周林, 李湘. 关于美国自动地面观测系统(ASOS)的考察报告[J]. 气象科技合作动态, 2002, (4): 18.
- [7] 中国气象局. 地面气象观测数据文件和记录簿表格式[M]. 北京: 气象出版社, 2005: 18-65.
- [8] Lanzante, J. R. Resistant, robust and nonparametric techniques for the analysis of climate data: Theory and examples, including applications to historical radiosonde station data[J]. Int. J. Climatol., 1996, 16, 1197-1226.
- [9] 刘小宁, 鞠晓慧, 范邵华. 空间回归检验方法在气象资料质量检验中的应用[J]. 应用气象学报, 2006, 17(1): 37-43.