

南海区域站气象资料质量控制方法改进

江益 王立俊 羊清雯 王小光 陈海莉
陈珍莉 严经敏 施晨晓

(海南省气象信息中心, 海口 570203)

摘要 为确保南海海岛气象站资料的可靠性、可用性, 本文通过对南海区域站实时气象资料原有质量控制方法得出的疑误信息进行分析, 利用增加双站对比检查和对空间一致性检查方法进行改进, 对南海区域站实时气象资料进行质量控制, 以降低南海区域站正确数据被检测为疑误数据的概率。应用新质量控制方法对 2015 年 9 月至 2016 年 7 月南海区域站数据进行质量控制, 剔除的误检信息占原疑误信息的 71.7%, 疑误检出后的准确率有较大提高。

关键词 区域站; 南海; 实时气象资料; 质量控制

中图分类号: P413 **DOI:** 10.19517/j.1671-6345.20170362 **文献标识码:** A

引言

南海是我国领土的重要组成部分, 对我国有着重大的战略意义, 因此南海气象资料十分重要, 为确保海岛气象站资料的可靠性、可用性, 需探索适合南海区域自动站的实时气象资料质量控制方法。国内外有许多学者对实时气象资料质量控制方法进行了研究^[1-8], 全国各气象部门现使用实时历史气象资料一体化质量控制系统, 依托气象资料业务系统(MDOS)对实时历史气象观测资料进行质量控制。但南海区域站情况特殊, 站点距离远, 使用 MDOS 进行质量控制时, 仍常出现大量最终被确认为正确的疑误信息, 给质量控制人员造成较大的困扰, 本文针对南海区域站的特殊性对该质量控制方法进行改进。

1 资料

南海区域站距离远, 交通不便利, 无法及时对区域站进行维护维修, 因此在建有区域站的岛礁上都装有 2 套自动站进行气象观测, 其中对于交通更为不便的南沙群岛, 每套自动站都有 2 种通信方式(GPRS、北斗 DPC)对数据进行传输, 并为方便管理, 每条通信线路各赋予了 1 个站号, 最大限度保证

海岛区域站数据的完整性。

本文所用资料为南海西沙群岛和南沙群岛共 12 个岛礁(图 1)上的 38 个区域站数据, 数据来源于

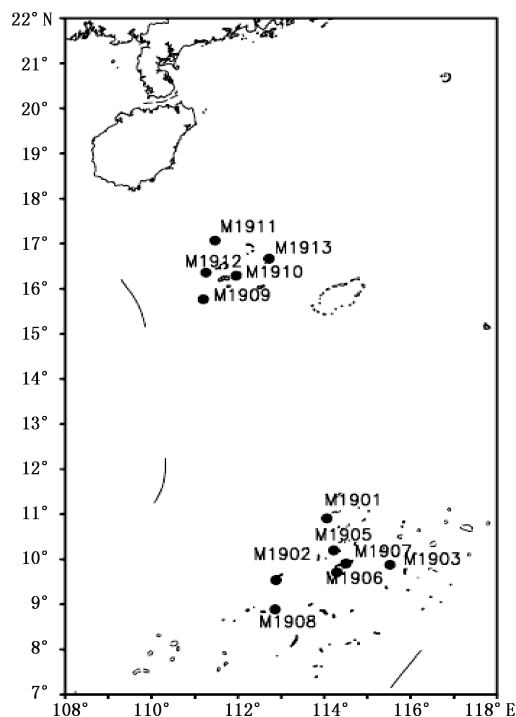


图 1 南海区域站分布

<http://www.qxkj.net.cn> 气象科技

海南省气象局青年基金项目(HNQXQN201405)资助

作者简介: 江益, 女, 1988 年生, 硕士, 工程师, 从事气象资料审核工作, Email: moer1003@126.com

收稿日期: 2017 年 6 月 20 日; 定稿日期: 2018 年 2 月 2 日

气象资料业务系统数据库,区域站站号分布如表 1 所示,1.1 和 1.2 列的区域站为第 1 套自动站的 2 条不同的通信线路,其中 1.1 使用北斗 DPC 通信,1.2 使用 GPRS 通信,2.1 和 2.2 列的区域站为第 2 套自动站的 2 条不同的通信线路,其中 2.1 使用北斗 DPC 通信,2.2 使用 GPRS 通信。

表 1 南海区域站站号

站名	1.1	2.1	1.2	2.2
渚碧礁	M1901	M1914	M1925	M1932
永署礁	M1902	M1904	M1926	M1933
美济礁	M1903	M1915	M1927	M1934
南薰礁	M1905	M1921	M1928	M1935
赤瓜礁	M1906	M1922	M1929	M1936
东门礁	M1907	M1923	M1930	M1937
华阳礁	M1908	M1924	M1931	M1938
中建岛	M1909	M1916		
深航岛	M1910	M1917		
北礁	M1911	M1918		
金银岛	M1912	M1919		
东岛	M1913	M1920		

注:1.1 和 2.1 列使用北斗 DPC 通信,1.2 和 2.2 列使用 GPRS 通信方式。

2 原质量控制情况

现行的实时资料质量控制系统为 MDOS,该系统主要的质量控制方法是对气象观测资料进行气候界限值检查、极值检查、要素内部一致性检查、时间

一致性检查和空间一致性检查,检查出的疑误信息再经由人工审核确认后进行处理。

对 2015 年 6 月至 2016 年 7 月 MDOS 中出现的南海海岛区域站的疑误信息进行统计,共有疑误信息 33347 条,其中最终确认为错误的信息有 7745 条,被确认为错误的信息基本上为仪器故障所造成的数据错误,而大部分的疑误信息最终都被确认为正确数据。这些疑误信息多集中出现,为质量控制人员造成了较大的困扰。现行的质量控制系统中,由于南海区域站站点之间距离较远,且站点较为稀少,排除在空间一致性检查之外,即现有的质量控制系统并未对南海区域站进行空间一致性检查。因此对于南海区域站气象资料的质量控制方法仍需进一步研究与改进,降低正确数据被误检为疑误数据的概率。

对被确认为正确的疑误信息的情况进行分析,从图 2a 中可以看出,按照疑误信息的要素进行分类,气压相关的疑误信息有 22990 条,占总数量的 89.8%;其次为气温相关的疑误信息,共 2518 条,占 9.8%;风、降水量、相对湿度的疑误信息很少,总共只占 0.4%。从图 2b 中可以看出,2015 年 6—10 月被确认为正确的疑误信息共有 25098 条,占总数的 98.1%,而其余月份只占很少比率。6—10 月为南海低压和台风的多发月份,受南海热带低压和台风的影响,南海区域站的气压值与正常情况相比有明显偏差,造成大量疑误信息的产生。

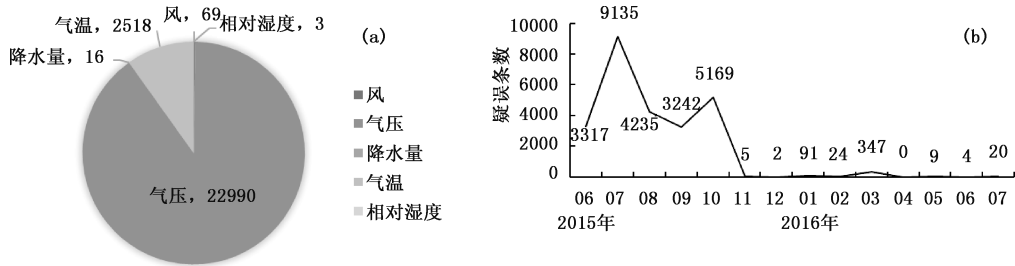


图 2 南海区域站 2015 年 6 月至 2016 年 7 月确认正确的疑误信息数量按要素(a)和按月份(b)分布

3 质量控制方法改进

现行的 MDOS 系统分 3 层质量控制码(简称质控码),其中为方便数据处理,综合质控码(0~100)对各类检查都设置了相应的质控码,0 为数据正确,1 为数据可疑,2 为数据警告,3 为数据错误,11 为未通过持续性检查,12 为未通过内部一致性检查等,

最后再综合确定数据最终的质控码^[7]。

3.1 双站对比检查

由于南海岛礁的双套自动站位置在一起,同一岛礁上自动站的对应数据理应相差不大,在对南海区域站小时数据进行质量控制时,对同一时次同个岛礁上的两套自动站相同气象要素的数据进行对比,以此判别数据是否正确。该方法最重要的是要确认两套自

动站之间各要素所允许的误差值,若同一要素的差值在误差值范围内,则视为数据正常。参考自动气象站测量性能指标要求^[9]和自动站与备份站之间各要素的差值范围,确认同一岛礁上双站之间所允许的差值(表 2),对两套站的数据(降水量除外)进行对比检查,对于降水量,若两套站之中一套站有降水、另一套站无降水,则两套站该时次降水量质控码均为 2;若两套站均有降水,则根据式(1)决定其质控码。

$$f = \begin{cases} 0 & |E_1 - E_2| \leq |\Delta E| \\ 1 & |E_1 - E_2| > |\Delta E| \end{cases} \quad (1)$$

式中, f 为经过双站对比检查后的质量控制码, E_1 、 E_2 分别为第 1 套自动站和第 2 套自动站的要素值, ΔE 为最大允许误差。

表 2 同一岛礁上双套站之间允许差值

测量要素(E)	允许差值(ΔE)
气压/hPa	± 0.6
气温/ $^{\circ}\text{C}$	± 0.6
相对湿度/%	± 10
风向/ $^{\circ}$	± 10
风速/($\text{m}\cdot\text{s}^{-1}$)	± 2
降水量/mm	± 0.8

3.2 空间一致性检查

在南海有低压或高压系统维持,或有台风的影响时,原有的质量控制系统会出现大量最终被确认为正确的气压疑误信息,或气温大幅度变化时,出现大量气温的疑误信息,因此针对该类大范围气压或气温变化事件需要进行再次质量控制。在通过原有质量控制系统检查过后,出现的疑误数据的质量控制码为 1 或 11。以 10 月为例,10 月 5169 条疑误信息中,有 5032 条出现在 10 月 2—5 日,疑误要素主要为气压和气温,疑误原因主要为未通过持续性检查,质控码为 11,有 83 条气压疑误信息出现在 12 日,疑误原因主要为快速质控(界限值、范围值等),质控码为 1。对此类事件发生时南海区域站的气压进行分析,10 月 2—5 日台风“彩虹”影响海南,从图 3 中可以看出,西沙各站气压的变化趋势基本一致,南沙各站气压的变化趋势也基本一致。

因此,当大范围气压发生变化或维持不变时,或气温大幅变化时,南海上同一区域的自动站会在同一时间被认定为疑误数据,针对此类情况,需要对气

压、气温进行空间一致性检查。采用空间插值的方法^[10],根据最近邻域法和反向距离法的基本原理,选取疑误站点周围相邻的几个站点,根据所选取台站与疑误台站的水平距离的远近来计算其要素值的插值权重,如式(2),并通过插值计算给出疑误站点要素的估算值,如式(3);将估算值与实际要素值相比较,通过比较结果来确定其质控码。

$$W(r) = \begin{cases} \exp\left[-\alpha\left(\frac{r}{R_p}\right)^2\right] - e^{-\alpha} & (r \leq R_p) \\ 0 & (r > R_p) \end{cases} \quad (2)$$

$$x_p = \frac{\sum_{i=1}^N W_i [x_i + \beta(z_p - z_i)]}{\sum_{i=1}^N W_i} \quad (3)$$

式(2)中, $W(r)$ 为邻近站点的要素值对疑误站点 p 的插值权重函数, r 为参考站点与疑误站点之间的水平距离, R_p 为截断距离。 α 为高斯形态系数,是与气象要素距离相关性衰减率有关的 1 个参数。式(3)中, W_i 即 $W(r)$, 为邻近站点 i 的要素值对疑误站点 p 的插值权重, N 为疑误站点 p 周围的邻近参考站点总数, x_p 为疑误站点最后估算数据, x_i 为参考站点 i 的要素值, β 为要素随海拔高度变化的垂直变化率,对于气压为 -0.1 hPa/m , 对于气温为 $-0.006 \text{ }^{\circ}\text{C/m}$, z_p 、 z_i 分别为疑误站点和其邻近站点的海拔高度。

由于南海区域站之间距离相距较远,分布较为稀疏,截断距离 R_p 根据海岛站密度选取 150 km , N 的范围为 $2 \sim 4$ 。根据刘宇等^[10]的研究,高斯形态系数 α 和参考站点总数 N 的变化对于插值平均绝对误差大小有较为明显的影响,在质量控制过程中,不需要将插值误差降到最小,只需误差稳定在一个范围内即可,根据插值平均绝对误差随 α 和 N 的变化规律,当 N 为 $2 \sim 4$ 时,对于气温 α 选取 3.0 ,对于气压 α 选取 5.0 。在空间一致性检查过程中,被其他质控方法检测出被检要素数据错误,即质控码为 3 的站点不做为邻近参考站,且邻近参考站个数小于 2 时,不进行空间一致性检查。将计算得出的 x_p 与疑误站点的实际观测值相比较来确认疑误数据的最终质控码。

$$\Delta = |x_p - x_0| \quad (4)$$

$$f = \begin{cases} 0 & \Delta < 1 \\ 1 & 1 \leq \Delta < 2 \\ 2 & \Delta \geq 2 \end{cases} \quad (5)$$

式中, x_0 为疑误站点的观测值, f 为质量控制码, 为避免漏检的情况发生, 根据对南海历史气象资料的

分析, 将 Δ 的阈值设置为较小的值, Δ 的单位对于气压为 hPa, 对于气温为 $^{\circ}\text{C}$ 。

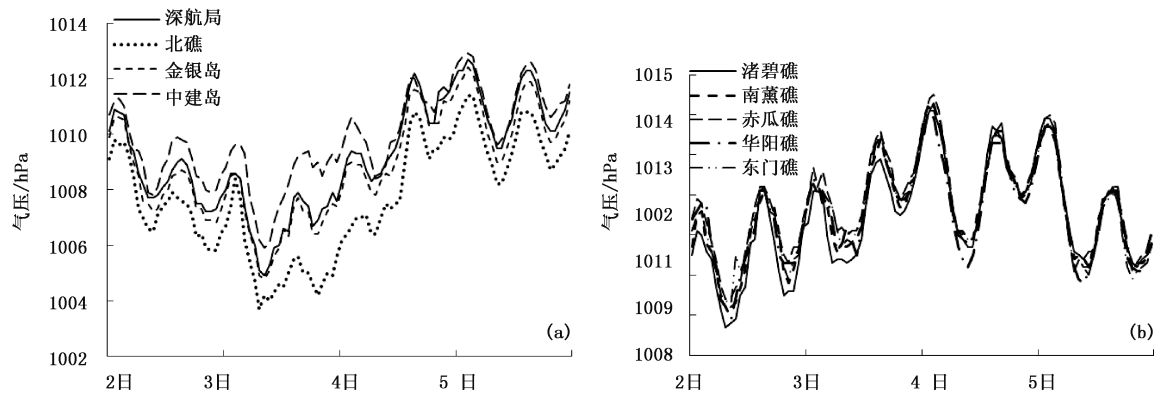


图 3 南海 2015 年 10 月 2—5 日西沙(a)和南沙(b)各岛礁气压

4 应用结果及实例分析

用添加上述质量控制检查的质量控制程序, 对南海区域站的气象资料进行质量控制, 将原质量控制方法与新质量控制方法质控后的结果进行对比分析。以 2015 年 9 月至 2016 年 7 月为例, 原质量控制方法检测出 10585 条疑误信息, 其中最终被确认为正确数据的有 8917 条; 新质量控制方法检测出 2995 条疑误信息, 其中最终被人工确认为正确的数据有 1327 条, 最终被确认为错误的的数据为 1668 条, 与原质量控制方法相同。由此可见, 新质量控制方法对于正确数据被检测为异常的情况有较大的改善, 剔除的 2015 年 9 月至 2016 年 7 月的误检信息共 7590 条, 占原疑误信息的 71.7%, 其中双站对比检查剔除 5598 条误检信息(气压疑误 3832 条, 气温疑误 1746 条, 其他疑误 20 条), 空间一致性检查剔除 1992 条(气压疑误 1556 条, 气温疑误 436 条), 正确数据被检测为疑误数据的误检率大为降低。

挑取 2015 年 10 月 5 日 08:00 的观测资料进行详细分析, 原质量控制系统对该时次的气象数据进

行质量控制, 得出疑误数据 120 条, 主要疑误要素为气压、气温, 质量控制码均为 11, 最终均被确认为正确数据, 出现疑误信息的站号如表 3 所示。

表 3 南海区域站 2015 年 10 月 5 日 08:00 出现疑误站点

站名	站号	站号	站号	站号
渚碧礁	M1901	M1914	M1925	M1932
南薰礁	M1905	M1935		
赤瓜礁	M1906	M1922	M1929	M1936
东门礁	M1923	M1937		
华阳礁	M1908	M1931		
中建岛	M1916			
深航岛	M1910	M1917		
北礁	M1911			
金银岛	M1912	M1919		

表 4 为该时次能够进行双站对比检查的站的疑误数据, 对比表 2 中数据可知, 这些站的双站误差均在双站对比检查允许误差范围内, 因此这些数据均为正确数据, 经过双站对比检查后, 将表中数据质控码置为 0。

表 4 南海区域站 2015 年 10 月 5 日 08:00 双套站气压和温度对比

站名	站号	气压 hPa	气温 $^{\circ}\text{C}$	站号	气压 hPa	气温 $^{\circ}\text{C}$	站号	气压 hPa	气温 $^{\circ}\text{C}$	站号	气压 hPa	气温 $^{\circ}\text{C}$
渚碧礁	M1901	1013.6	29.7	M1914	1013.4	29.7	M1925	1013.6	29.7	M1932	1013.4	29.7
南薰礁	M1905	1013.4	28.9	M1935	1013.5	28.8						
赤瓜礁	M1906	1013.5	30.4	M1922	1013.6	30.1	M1929	1013.5	30.4	M1936	1013.6	30.1
深航岛	M1910	1012.3	28.0	M1917	1012.4	27.9						

对其余疑误数据进行空间一致性检查,除华阳礁 M1908 站由于邻近参考站数小于 2,不进行空间一致性检查,其余站点均可进行检查。表 5 为疑误台站观测值和空间一致性检查的估算值之间的对比,从表中可以看出,该时次疑误要素的观测值与估算值之间的差值均小于 1,即数据均可直接判定为正确,质量控制码置为 0。综上所述,经过所有质量控制程序过后,2015 年 10 月 5 日 08:00 最终剩余疑误信息 12 条,比之原质控出的疑误信息条数减少了 90%。

表 5 南海区域站 2015 年 10 月 5 日 08:00 疑误站点
观测值与空间一致性检查估算值对比

站名	站号	气压/hPa		气温/℃	
		观测值	估算值	观测值	估算值
东门礁	M1923	1013.3	1013.5	29.9	29.7
中建岛	M1916	1012.6	1012.3	28.5	28.0
北礁	M1911	1010.7	1011.1	27.8	27.9
金银岛	M1912	1011.8	1011.9	27.9	28.1

5 结论

本文根据南海区域站建有双套站和站间距离大的特性,对原有的质量控制方法进行了改进。

(1)增加了双站对比检查,若同一岛礁上的两套自动站之间各要素的差值在误差允许范围内范围内,则视为数据正常。

(2)改进了空间一致性检查的方法,使南海区域站能够进行空间一致性检查,降低了特殊情况下气压或气温要素疑误的误检率。

(3)通过新的质量控制方法对南海区域站 2015 年 9 月至 2016 年 7 月的数据进行质量控制,其结果与原质量控制结果相对比,改进后的质量控制方法能准确地检查出错误数据,且对于正确数据被检测为异常的情况有较大的改善,剔除的误检信息占原疑误信息的 71.7%,正确数据被检测为疑误数据的误检率大为降低。

参考文献

- [1] Wolfson N, Erez J, Alpers Z. Automatic real-time quality control of surface synoptic observation [J]. Journal of applied meteorology, 2010, 17 (4): 449-457.
- [2] 韩海涛,李仲龙. 地面实时气象数据质量控制方法研究进展 [J]. 干旱气象, 2012, 30(2): 261-265.
- [3] 熊安元. 北欧气象观测资料的质量控制 [J]. 气象科技, 2003, 31 (5): 315-320.
- [4] 李良富,王汉杰,刘金玉,等. 基于黑板模型的地面气象数据质量控制 [J]. 气象科技, 2006, 34(2): 199-204.
- [5] 刘小宁,任芝花. 地面气象资料质量控制方法研究概述 [J]. 气象科技, 2005, 33(3): 199-203.
- [6] 任芝花,熊安元. 地面自动站观测资料三级质量控制业务系统的研制 [J]. 气象, 2007, 33(1): 19-24.
- [7] 王海军,杨志彪,杨代才,等. 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.
- [8] 何志军,封秀燕,何利德,等. 气象观测资料的四方位空间一致性检验 [J]. 气象, 2010, 36(5): 118-122.
- [9] 黄思源,刘钧. 新型自动气象站观测业务技术 [M]. 北京:气象出版社, 2014: 509-510.
- [10] 刘宇,陈洋勤,张稳,等. 一种地面气温的空间插值方法及其误差分析 [J]. 大气科学, 2006, 30(1): 146-152.

(下转第 496 页)

Improvement of Quality Control Method for Real Time Meteorological Data from Regional Automatic Stations on South China Sea

Jiang Yi Wang Lijun Yang Qingwen Wang Xiaoguang Chen Haili
Chen Zhenli Yan Jingmin Shi Chenxiao

(Meteorological Information Center of Hainan Province, Haikou 570203)

Abstract: The South China Sea is an important part of China's territory, and it has a great strategic significance for China. In order to ensure the reliability and availability of the data from the island weather stations, this paper analyzes the error messages that are generated by the original quality control system, and in order to reduce the number of false checks for the regional automatic stations on the South China Sea, a method is developed, which can check the data difference between two stations on the same island, and the spatial consistency check method is improved. By using the new method to check the data of the regional automatic stations on the South China Sea from September 2015 to July 2016, the removed false check messages accounts for 71.7% of the original false check messages. The accuracy of error detection is greatly improved.

Keywords: regional automatic station; South China Sea; real time meteorological data; quality control