

李芬,周晓香,赖亮,等. 2019. 人工审核制作 A 文件与 MDOS 自动制作 A 文件对比分析[J]. 气象与减灾研究,42(1):75-80.
Li Fen, Zhou Xiaoxiang, Lai Liang, et al. 2019. Contrastive analysis between manual production of A File and automatic production with MDOS[J]. Meteorology and Disaster Reduction Research,42(1):75-80.

人工审核制作 A 文件与 MDOS 自动制作 A 文件对比分析

李 芬, 周晓香, 赖 亮, 杨立苑

江西省气象信息中心, 江西 南昌 330096

摘 要: 通过设计地面气象观测月数据文件对比软件,以江西省 91 个国家级自动站 2016 年全年经过省级人工质量控制后制作的地面气象观测数据 A 文件为参考标准,与气象资料业务系统(简称“MDOS 系统”)自动制作的同站同期数据 A 文件进行对比分析,并进一步对数据差异的时空分布、要素分布等进行了分析。结果表明:91 个国家级自动气象站全年数据对比差异个数总计为 16 490 个,占全年数据总量的 1.02%。MDOS A 文件中自动观测要素数据差异的主要原因有传感器系统性偏移、仪器安装错误、参数设置错误、疑误数据无法关联修订、异常记录未能有效判断并处理;人工观测要素数据差异的主要原因有数据格式错误、记录顺序错误、观测要素记录不匹配、漏记录等错误在 MDOS 平台未能有效检测出。最后提出了确保 MDOS 系统平台参数设置的正确性、完善 MDOS 系统平台数据处理能力等改进措施。

关键词: 月数据,质量控制,对比差异,分析

中图分类号: P413

文献标识码: A

文章编号: 1007-9033(2019)01-0075-06

doi: 10.12013/qxyjzj2019-012

Contrastive Analysis between Manual Production of A File and Automatic Production with MDOS

Li Fen, Zhou Xiaoxiang, Lai Liang, Yang Liyuan

Jiangxi Meteorological Information Center, Nanchang 330096, China

Abstract: Through comparing software design on ground meteorological observation data file, with 91 national automatic stations in Jiangxi province in 2016, the year after artificial quality control at the provincial level of the ground meteorological observation data in A file as the reference standard, and meteorological data service system (hereinafter referred to as MDOS system) automatically made to stand in the same period, comparing the data to A file the results showed that the number 91 national automatic station year data contrast differences amounted to 16490, or 1.02% total data for the whole year. The paper further analyzes the spatial and temporal distribution and factor distribution of data differences, points out the main reasons for the differences, and proposes improvement measures, aiming at providing more accurate observation data support for forecasting, service and scientific research.

Key words: monthly data; quality control; comparative difference; analysis

0 引 言

地面气象观测月数据文件是气象台站所积累的

气象情报资料原始档案,是国家的宝贵财富(中国气象局,2003)。随着观测业务自动化的发展,地面气象观测月数据文件经历了由台站制作并预审报送省

收稿日期: 2019-01-13; 修订日期: 2019-03-15.

基金项目: 2014 年江西省气象科技重点项目“地面气象观测数据分析服务系统研究”。

作者简介: 李 芬,高级工程师,主要从事气象数据处理与应用, E-mail:413563706@qq.com.

级人工审核,到省级利用气象资料业务系统(以下简称“MDOS 系统”)制作月数据 A、J 文件的变革。2017 年 7 月起,每月 5 日利用 MDOS 系统完成上月 A、J 文件制作并上报国家气象信息中心。由此,A、J 文件数据质量控制时间由以往的 2 个月缩短到 5 d,数据质量控制方式由人机交互质量控制向 MDOS 系统实时质量控制转变,A、J 文件审核制作的效率显著提高。

近年来,许多学者都对资料的质量控制进行了研究。王海军等(2007)对实时资料自动质量控制进行了分析研究。任芝花等(2010)、赵煜飞等(2011)分别对小时降水、小时相对湿度质量控制方法进行了分析研究。胡玉峰(2004)、王颖等(2007)分析了自动观测与人工观测差异,指出差异的主要原因有仪器原理差异、时空差异、样本差异及时次的差异。鞠晓慧等(2010)分析了自动与人工观测气压对比差异,指出高低温环境及海拔高度对差异有一定的影响。李亚丽等(2015)分析了自动与人工观测气温差异偏大的原因,指出传感器检定更换导致对比差值跳变,传感器零点漂移导致自动观测气温严重偏离人工观测值。茆佳佳等(2016)分析了自动与人工观测相对湿度对比差异,指出风速及气温对相对湿度的观测有一定的影响。任芝花等(2007)、于清平等(2008)分析了自动与人工观测的风速、降水对比差异,指出不同观测仪器及不同观测时间是造成差异的主要原因。然而针对 MDOS 系统、传统人工审核这两种不同质量控制方式形成的 A 文件数据差异对比分析比较少见。为提高 MDOS 系统自动生成制作 A 文件的数据质量,有必要对其与传统人工审核开展对比分析。本文对江西省 91 个国家自动气象站 2016 年全年经过省级人工质量控制后制作的地面气象观测数据 A 文件(以下简称“人工 A 文件”)为参考标准,与 MDOS 系统自动制作的同站同期数据 A 文件(以下简称“MDOS A 文件”)进行对比,对数据差异进行详细的分析,找出差异原因,在实际业务中针对性完善质量控制方法,提高 MDOS 自动生成 A 文件数据质量,为预报、服务、科研提供更加准确及时的观测数据。

1 对比评估的资料及方法

1.1 对比资料来源

文中使用江西省气象信息中心提供的江西省 91 个国家自动气象站 2016 年 1—12 月的人工 A 文件及 MDOS A 文件,样本文件数共计 2 184 个。

对比的要素包括气压、温度(气温、露点温度)、湿度(相对湿度、水汽压)、云量、能见度、降水量、蒸发量、积雪、风向风速、浅层地温(0—40 cm)、深层地温(80—320 cm)、日照、草面温度等(中国气象局,2005)。其中,有云量及蒸发量的观测台站为 26 个国家基本、基准站。

1.2 对比分析方法及程序设计

对人工 A 文件及 MDOS A 文件的源数据文件进行逐一对比,分别统计每份数据文件中台站参数、要素方式位和要素观测值的数据修改情况,并输出结果入库,通过数理统计方法,按站点、时间、观测要素等进行统计分析,输出可视化分析图表结果。图 1 为对比分析软件设计框架。

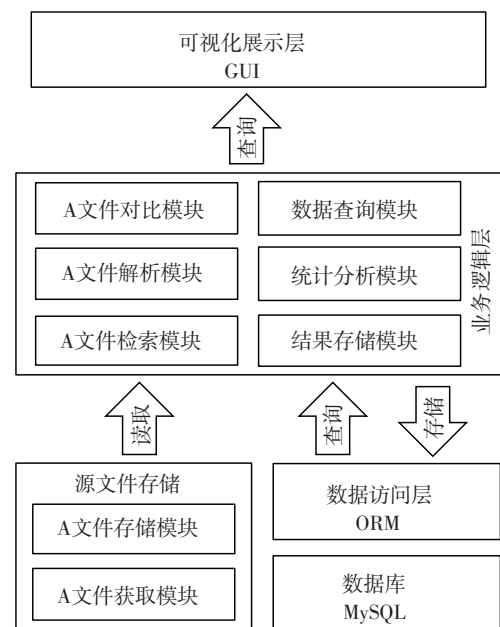


图 1 月报对比分析软件设计框架

Fig. 1 Design framework of the monthly comparative analysis software

由图 1 可知,软件由源文件存储、数据访问层、数据库、业务逻辑层、可视化展示层构成。采用 C/S 架构设计,使用 MySQL 数据库作为对比结果存储数据库,并通过 ORM 框架和上层业务逻辑层进行交互。首先获取人工 A 文件和 MDOS A 文件源文件并按照约定的目录结构及文件命名规则进行存储,启动对比算法,进行文件检索、解析、对比以及将对比结果入库。图 2 为对比算法流程。具体如下:根据报文的站号、日期等信息从源文件存储中检索出相应报文,按照《气象数据归档格式 地面》(中国气象局,2010)定义的地面气象数据归档格式标准来

完成报文解析,并将解析内容读入到内存中。通过报文对比模块将读入到存储中的两份报文内容数值进行逐个比较,将报文内容的不同处根据要素和日期进行计数,计数结果存储到 MySQL 数据库中。按照任意时间段、任意要素、任意站点进行统计,并以表格和柱状图的形式呈现对比结果。

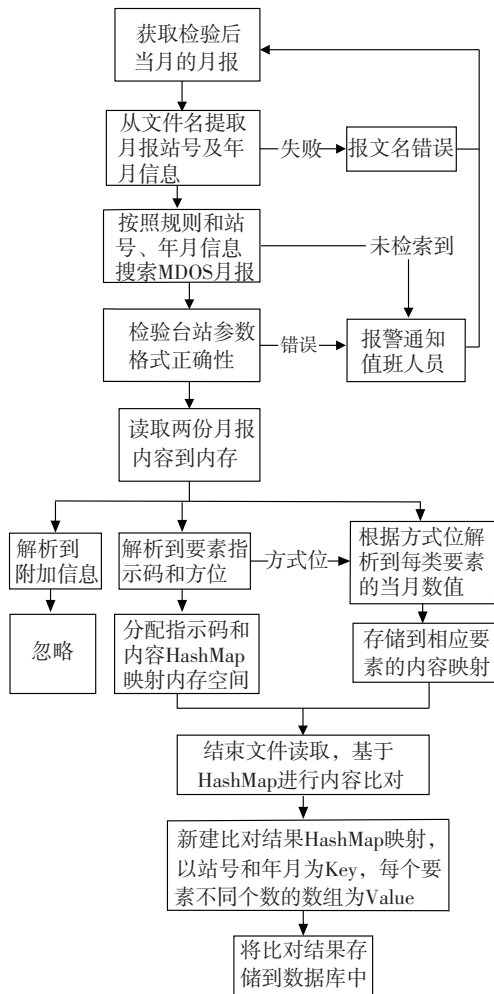


图 2 A 文件对比算法流程
Fig. 2 A file comparison algorithm flow chart

2 对比结果与分析

2.1 数据差异的空间分布

江西省 2016 年人工 A 文件与 MDOS A 文件数据差异总数为 16 490 个。其中,数据差异总数大于 200 个的有 14 个台站,总数为 12 721 个,占全省数据差异总数的 77%,其他 77 个台站数据差异个数均小于 200 个。

表 1 是江西省 2016 年人工 A 文件与 MDOS A

文件数据差异个数前 14 位的排位表。从表中可以看出,有 4 个台站数据差异明显偏多,分别是区站号为 57796 的台站数据差异总数为 4 215 个,区站号为 58517 的数据差异总数为 2 314 个,区站号为 58814 的数据差异总数为 1 732 个,区站号为 58622 的数据差异总数为 1 429 个,4 个台站数据差异总数为 9 690 个,占全省数据差异总数的 58.76%。

表 1 江西省 2016 年人工 A 文件与 MDOS A 文件数据差异个数排位表(前 14 位)

Table 1 Ranking list of data differences between manual A files and MDOS A files in Jiangxi province in 2016 (top 14)

差异排名	区站号	差异个数
1	57796	4 215
2	58517	2 314
3	58814	1 732
4	58622	1 429
5	57696	393
6	58600	372
7	59092	354
8	58612	348
9	58719	288
10	58701	283
11	59093	267
12	58607	252
13	57891	237
14	57894	237

2.2 数据差异的时间分布

图 3 是江西省 2016 年人工 A 文件与 MDOS A 文件数据差异个数随时间的变化分布。分析发现,各月数据差异分布不均匀。数据差异总数最多的是 1 月,为 3588 个,其次是 12 月,为 3216 个;数据差

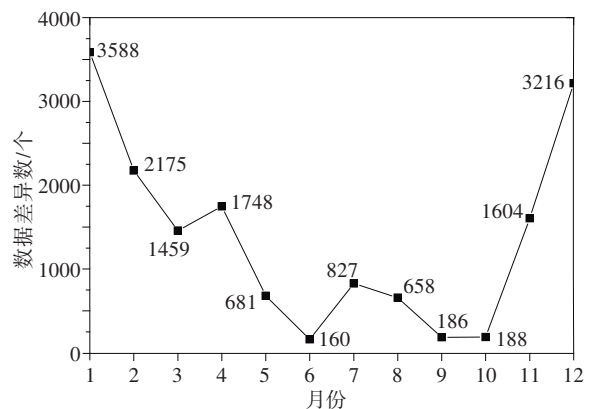


图 3 江西省 2016 年人工 A 文件与 MDOS A 文件数据差异个数随时间的变化分布

Fig. 3 Time distribution of data differences between manual A files and MDOS A files in Jiangxi province in 2016

异最少的是6月,为160个,其次是9月和10月,分别为186个和188个。冬季12月一次年2月数据差异较多,这主要是因为冬季雨雪天气较多,仪器受外界环境等影响出现问题的情况相对增加,且冬季观测要素记录处理较其他季节更为复杂,如,固态降水、风仪器冻结和蒸发结冰等异常记录的处理等;6、9、10月数据差异较少,主要是因为秋季天气总体稳定,温湿度等设备工作环境适宜,仪器发生故障的情况相对较少。

2.3 数据差异的各观测要素分布

图4是江西省2016年人工A文件与MDOS A文件数据差异个数要素分布。从图中可以看出,2016年月数据对比差异各要素分布不均,差异最多

的要素是风,差异总数为4469个,其次深层地温及蒸发量差异总数也较多,差异总数分别为3289个及2363个;差异总数最少的要素是电线积冰,为67个,云量及草面温度差异总数也较少,分别为105个及124个。

分析图4还可知,各观测要素数据差异个数分布不均。自动观测要素数据差异个数相对偏多,人工观测要素数据差异相对偏少。气压、温度、湿度、风、降水量、蒸发量、草面温度、浅层地温、深层地温自动观测要素数据差异共计14614个,占全省数据差异总数的88.62%,天气现象、云量、日照、积雪及电线积冰人工观测要素数据差异共计1876个,占全省数据差异总数的11.38%。

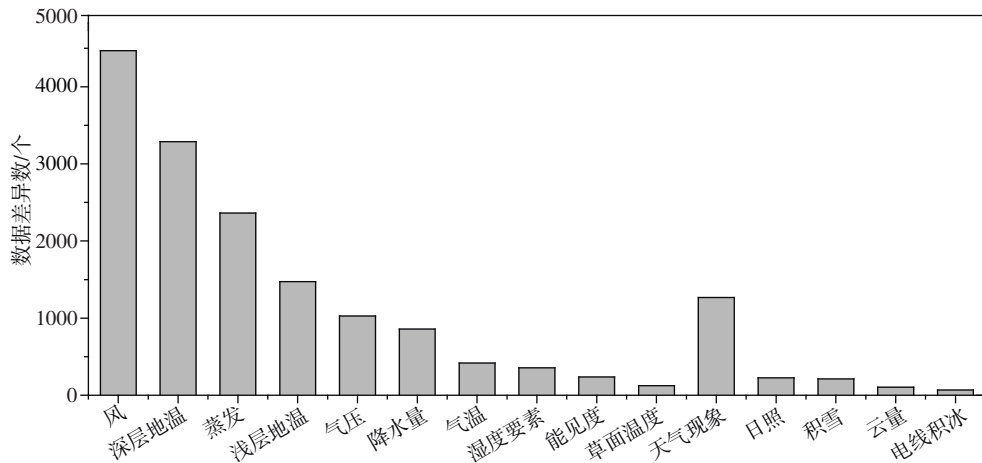


图4 江西省2016年人工A文件与MDOS A文件数据差异个数要素分布

Fig. 4 Key factors of data difference between manual A file and MDOS A file in Jiangxi province in 2016

3 数据差异原因分析

3.1 自动观测要素数据差异的主要原因

通过人工经验综合分析判断,查找MDOS A文件中自动观测要素数据差异的主要原因。

1) 传感器系统性偏移

仪器运行过程中,性能逐渐降低,产生传感器系统性漂移。传感器漂移有三种情况:正常状态漂移、失效状态漂移、失真状态漂移。传感器漂移是一个逐渐变化且持续时间较长的过程,传感器漂移期间异常记录短时间内不易被发现。此类异常记录比较“隐性”,在MDOS平台实时自动质量控制中未提示疑误信息,日常MDOS值班工作中不易被发现,需人工经验结合较长时间的多源资料综合分析判断,这类原因造成数据差异占全省数据差异总数的

11.62%。例如:区站号为58607的台站气压仪器较长时间漂移,故障期间与正常值偏差约2 hPa,造成数据差异总数443个,占气压数据差异总数的43.14%。区站号为58600及58622的台站连续多日小时蒸发量固定为“0”,此类异常造成该两站数据差异总数为1474个,占蒸发量数据差异总数的62.40%。

2) 仪器安装错误

气象台站在仪器设备的维修、维护及检定过程中,由于种种因素导致仪器安装错误,产生数据异常。这类原因导致的数据差异占全省数据差异总数的35.38%。通常情况下,某观测要素仪器安装错误,观测数据异常会比较明显,在MDOS平台实时自动质量控制中会提示疑误信息,疑误数据及时在MDOS平台日常值班中“时清”、“日清”及“月清”工

作中就能够得以解决。然而有些观测要素仪器安装错误后,短时间内不容易被发现,需较长时间人工经验综合分析判断。如:区站号为57796的台站2016年1月1日—3月22日,台站维护仪器时,风向传感器安装错误,风向角度偏差 180° ,导致数据差异4128个,占风数据差异总数的92.37%;区站号为58814的台站2016年1月1日—5月6日,40 cm地温与80 cm地温接线错误,导致数据差异1706个,占地温数据差异总数的35.82%。

3) 参数设置错误

出现MDOS平台中台站参数的设置错误时,MDOS平台自身不提示疑误信息,业务人员日常值班也不易发现此类问题。这类原因造成数据差异占全省数据差异总数的15.86%。例如:区站号为58517的台站2016年12月MDOS平台参数设置中,深层地温参数设置错误,导致80—320 cm地温数据差异2232个,占地温数据差异总数的46.86%;部分台站ISOS软件台站参数设置时,纬度的“秒”设置未按规定设置为“00”,导致气压数据差异384个,占气压数据差异总数的37.39%。

4) 疑误数据无法关联修订

MDOS平台提示小时极值疑误信息时,相关联的极值出现时间不匹配,但未提示有疑误信息,造成极值数据未得到正确修订;修订了小时值,相关的日合计值并不会同步修订,或者修订了日合计值,相应的小时值没有一并修订(中国气象局,2016)。气压、温度、湿度、草面温度、蒸发量、降水等要素均出现过上述情况,造成数据差异853个,占全省数据差异总数的5.17%。如:MDOS平台日常值班时,修订了本站气压、气温、相对湿度时,与其相关联的海平面气压、露点温度、水汽压要素值无法同步修订,造成海平面气压、露点温度、水汽压数据差异;台站修订了小时降水量后,对应日数据文件中相应的定时降水量(20—20时、20—08时、08—20时)同样不会同步修订,造成日定时降水量合计与日小时降水量累计值不一致。

5) 异常记录未能有效判断并处理

MDOS平台中的异常记录未能有效判断并及时处理,造成数据差异。如:小时蒸发量异常时,MDOS平台不提示疑误信息,造成数据差异总数为536个;MDOS平台软件生成A文件时,由于软件不完善,造成降水量上下跨连接值为缺测错误,滞后降水量、野值降水量、冬季固态降水等异常记录未有效处理,造成数据差异数650个;其他自动站故障、

雷击、仪器检定等原因造成的异常记录在MDOS平台未得到有效判断和处理,导致数据差异2486个。这类原因造成的数据差异占全省数据差异总数的22.27%。

3.2 人工观测要素数据差异主要原因

人工观测中数据格式错误、记录顺序错误、观测要素记录不匹配、漏记录等错误在MDOS平台未能有效检测出,是造成人工观测要素数据差异的主要原因。如:天气现象出现的时间先后顺序错误、分钟有降水量相应的无降水现象匹配、格式错误等造成天气现象要素数据差异总数1268个,占全省数据差异总数的7.69%;云量观测记录与天气现象、能见度、日照等记录不匹配造成云量要素数据差异总数105个,占全省数据差异总数的0.64%;漏录入日照时数、相邻时次日照记录录入错误等原因造成日照记录数据差异总数225个,占全省数据差异总数的1.36%。

4 小结

通过对江西省91个国家自动站2016年人工A文件与MDOS A文件对比分析,以人工A文件为参考标准,MDOS平台数据质量控制能力达到998.98%,存在1.02%的错误或可疑数据MDOS平台未能有效检测和质量控制,这类情况集中出现在少数台站,主要发生在冬季,且以自动观测要素为主。这类错误或可疑数据主要因传感器系统性漂移、仪器安装错误、参数错误、人工观测错误和记录错误等原因引起。各气象台站需要确保MDOS系统平台参数设置的正确性,完善MDOS系统平台数据处理能力,提高自动站设备业务可用性,同时结合多源数据分析及人工经验综合判断方法开发辅助质量控制程序,持续提高资料质量控制能力。

参考文献(References)

- 胡玉峰. 2004. 自动与人工观测数据的差异[J]. 应用气象学报, 15(6): 719-726. Hu Y F. 2004. Differences between data of automatic and manual observation[J]. J Appl Meteor Sci, 15(6): 719-726. (in Chinese)
- 鞠晓慧,任芝花,曹丽娟,等. 2010. 自动与人工观测气压的差异及原因分析[J]. 气象, 36(1): 104-110. Ju X H, Ren Z H, Cao L J, et al. 2010. Difference analysis between automatic and manual atmospheric pressure measurements[J]. Meteor Mon, 36(1): 104-110. (in Chinese)
- 李亚丽,任芝花,陈高峰,等. 2015. 自动与人工观测气温差异偏大的原因及影响分析—以143个国家基准站为例[J]. 气象, 41(8): 1007-1016. Li Y L, Ren Z H, Chen G F, et al. 2015. Causes

- and impact analysis of errors between temperatures obtained by automatic benchmark stations at 143 national automatic benchmark stations[J]. Meteor Mon,41(8):1007-1016.(in Chinese)
- 茆佳佳,莫月琴,张雪芬,等. 2016. 自动观测与人工观测相对湿度比对分析[J]. 应用气象学报,27(3):370-379. Mao J J, Mo Y Q, Zhang X F, et al. 2016. Comparison and analysis automatically-observation and manually-observed relative humidity[J]. J Appl Meteor Sci, 27(3):370-379.(in Chinese)
- 任芝花,冯明农,张洪政,等. 2007. 自动与人工观测降雨量的差异及相关性[J]. 应用气象学报,18(3):358-364. Ren Z H, Feng M N, Zhang H Z, et al. 2007. The difference and relativity and between rainfall by automatic recording manual observation[J]. J Appl Meteor Sci, 18(3):358-364.(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 hourly precipitation data from automatic weather stations in china[J]. Meteor Mon, 36(7):123-132.(in Chinese)
- 王海军,杨志彪,杨代才,等. 2007. 自动气象站实时资料自动质量控制方法及其应用[J]. 气象,33(10):102-106. Wang H J, Yang Z B, Yang D C, et al. 2007. The method and application of automatic quality control for real time data from automatic weather stations[J]. Meteor Mon, 33(10):102-106.(in Chinese)
- 王颖,刘小宁,鞠晓慧. 2007. 自动观测与人工观测差异的初步分析[J]. 应用气象学报,18(6):849-855. Wang Y, Lin X N, Ju X H. 2007. Differences between automatic and manual observation[J]. J Appl Meteor Sci, 18(6):849-855.(in Chinese)
- 于清平,黄文杰,李崇志,等. 2008. 南京自动气象站与人工观测风速差异分析[J]. 气象科学,28(5):577-580. Yu Q P, Huang W J, Li C Z, et al. 2008. The difference analysis of wind speed by AWS and manual observation in Nanjing[J]. Journal of the Meteorological Sciences,28(5):577-580.(in Chinese)
- 赵煜飞,任芝花,张强.2011.适用于全国自动站正点相对湿度资料的质量控制方法[J].气象科学,31(6):687-693. Zhao Yufei, Ren Zhihua, Zhang Qiang.2011.Quality control procedures for hourly relative humidity data from national automatic weather stations [J].Scientia Meteor Sinica,31(6):687-693.(in Chinese)
- 中国气象局. 2003. 地面气象观测规范[M]. 北京:气象出版社:104-111. China Meteorological Administration. 2003. Ground meteorological observation specifications[M]. Beijing:China Meteorology Press:104-111.(in Chinese)
- 中国气象局. 2005. 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:18-32. China Meteorological Administration. 2005. Format of surface meteorological observation data files and tables[M]. Beijing:China Meteorology Press:18-32.(in Chinese)
- 中国气象局. 2016. 地面气象观测业务技术规定实用手册[M]. 北京:气象出版社:25-26. China Meteorological Administration. 2016. Practical manual for technical regulations on surface meteorological observation[M]. Beijing:China Meteorology Press:25-26.(in Chinese)
- 中国气象局. 2010. QX/T 199-2010 气象数据归档格式 地面[S]. 北京:气象出版社. China Meteorological Administration. 2010. QX/T 199-2010 The format of ground meteorological data filed. [S]. Beijing:China Meteorology Press.