

# 自动气象站运行效能统计

李 雁 梁海河 孟昭林 裴 翀 石 城

(中国气象局气象探测中心,北京 100101)

## 摘 要

利用中国气象局气象探测中心综合气象观测系统运行监控平台中对自动气象站的监控情况,将监控平台上 2100 余套国家级自动气象站的运行效能分华北、东北、华东、中南、西南、西北 6 个区域从数据到报率、可用性和可靠性 3 方面进行统计,并从数据报文格式错误和数据要素错误两方面对影响自动站效能的因素进行分析。结果表明:2007 年 9—10 月全国国家级自动站的数据到报率、可用性和可靠性保持在 80% 以上,且三者的变化趋势基本一致;东北区域自动站的运行效能最高,而西南区域的运行效能最低;数据到报率是数据可用性和可靠性的前提条件,在数据到报率一定的情况下,数据报文格式错误较数据要素质量错误更多地影响数据可用性和可靠性;第 1 行格式错误较第 2 行格式错误对自动站效能的影响程度更大,相对于其他观测要素,地温要素是影响自动站效能高低的主要因子;在地温要素中,不同层次地温对自动站效能影响也存在一定差异,320 cm 地温影响程度最大,而 5 cm 地温影响程度最小,基本呈现出越往地下深处地温要素对效能影响程度越大的趋势。

**关键词:** 综合气象观测系统运行监控平台;自动气象站;运行效能

## 引 言

地基、空基和天基一体化的综合观测系统是现代观测体系建设的目标<sup>[1]</sup>,自动气象站是地基观测的重要组成部分。中国气象局自国家“十五”计划开始,陆续在全国各台站装备多要素自动气象站,在某些环境恶劣或高海拔地区替代人工观测。然而,受各种因素制约,自动站观测的数据报文格式和观测要素数据出错率较高,影响了自动站本身效能更大程度发挥,给预报工作带来很大不便。

本文以中国气象局气象探测中心综合气象观测系统运行监控平台中监控的全国 2100 余套国家级自动气象站为基础<sup>[2]</sup>,构建了数据到报率、数据可用性和可靠性 3 类运行指标,统计了华北、东北、华东、中南、西南、西北 6 个区域的自动站效能情况,从数据报文格式错误(第 1 行格式错误、第 2 行格式错误和时间错误)情况和数据要素错误(温、压、湿、风等)两方面分析了影响自动站效能的主要因素,并提出了可行性建议。

## 1 综合气象观测系统运行监控与保障平台简介

综合气象观测系统运行监控平台(ASOM)是在

中国气象局现行的两级管理(中国气象局职能司、省气象局职能处)和三级保障体系(中国气象局保障部门、省气象局保障部门、地市局台站等)框架上,按照气象设备技术保障职责和业务流程,满足气象探测业务运行监控和技术保障的需要,为实现技术保障工作向管理信息化、流程规范化、业务体系化方向发展而建立的一个技术支撑平台<sup>[3-4]</sup>。

该系统定位为全国气象探测网业务运行监控及技术保障的实时业务系统,服务于探测网技术保障人员和业务管理人员,是现代气象综合观测系统的重要组成部分。目前,该系统监控的设备包括全国各省多普勒天气雷达 110 余部、探空雷达 120 部、自动气象站 2100 多套<sup>[3]</sup>,具有设备运行监控、探测数据监控、维护维修管理、探测业务运行评估、站网信息管理、装备保障管理等功能。

## 2 运行效能评估指标的建立

综合气象观测系统运行监控平台中监控的自动站包括气候观象台——原国家基准气候站、气象观测一级站——原国家基本气象站、气象观测二级站——原国家一般气象站 3 类。

自动站运行状态通过台站上传的数据报文中的观测数据质量反映。在监控平台上将自动站的运行状态分为 4 类:①数据正常(绿色) 正常接收到自动

站探测数据报文,经相关判定方法检查发现,探测数据无任何数据质量问题发生,数据完全正常,数据质量控制等级为 0。② 数据异常(橙色) 正常接收到自动站探测数据报文,经相关判定方法检查发现,观测要素数据可疑,数据基本可用,数据质量控制等级为 1 或 2。③ 数据错误(红色) 正常接收到自动站探测数据报文,但经相关数据质量判定方法检查发现,观测要素数据错误或数据格式错误,数据不可用,数据质量控制等级为 3。④ 传输异常(灰色) 由于自动站设备故障或网络传输故障,无任何数据传输。

自动气象站观测数据的准确情况将直接影响天

$$\text{到报率} = \frac{\text{应工作时次} - \text{未到报时次}}{\text{应工作时次}} \times 100\% \quad (1)$$

$$\text{可用性} = \frac{\text{应工作时次} - \text{数据错误时次} - \text{未到报时次} - \text{格式错误时次}}{\text{应工作时次}} \times 100\% \quad (2)$$

$$\text{可靠性} = \frac{\text{应工作时次} - \text{数据错误时次} - \text{数据异常时次} - \text{未到报时次} - \text{格式错误时次}}{\text{应工作时次}} \times 100\% \quad (3)$$

其中,应工作时次为自动站自投入使用到考核结束所经历的时间;未到报时次为在应工作时次内缺少报文的时次,和上述自动站运行状态中的传输异常相对应;数据错误时次为在应工作时次内对上传数据报文中观测数据经数据质量判定方法发现存在数据质量错误的时次,和上述自动站运行状态中的数据错误相对应;数据异常时次为在应工作时次内对上传数据报文中观测数据经数据质量判定方法发现数据质量存在可疑或异常的时次,和上述自动站的运行状态中的数据异常相对应;格式错误时次为在应工作时次内由于自动站的原始报文不符合自动站报文编码规则,无法对原始报文进行解析入库的时次。

### 3 自动站运行效能统计

#### 3.1 自动站运行效能统计

首先,将监控平台上的自动站按我国行政区划划分到华北、东北、华东、中南、西南和西北 6 个区域。其次,对全国 2100 余套自动站自 2007 年 9 月 1 日—2007 年 10 月 31 的运行情况分区域从数据到报率、数据可用性和可靠性 3 个方面进行统计,如图 1 所示。

从图 1 可以看出,西南地区自动站的数据可用性和可靠性最低,远远低于 81.96% 和 81.51% 的平均水平;东北地区的效能最高。

气预报的准确性<sup>[5]</sup>。为了客观、定量评价自动气象站运行情况,使用数据到报率、数据可用性和可靠性 3 类指标。

数据到报是指自动站观测数据报文在规定时间内由台站每天 24 个时次上传到中国气象局;数据可用是指到报数据中尽管有可疑和异常情况发生,但不影响对设备的总体评价;数据可靠是指在设备申报故障或申报维护而关机时段之外,数据及时到报,且无异常、错误和传输异常情况发生。

根据上述定义,3 类指标表示如下:

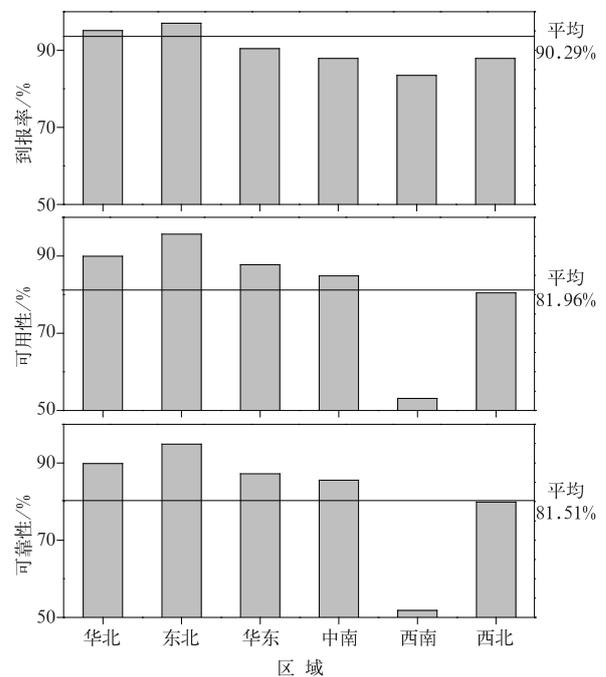


图 1 2007 年 9—10 月自动站运行效能统计  
Fig. 1 The statistic of the AWS's efficiency in September and October of 2007

#### 3.2 自动站运行效能

影响自动站运行效能高低的因素主要有自动站故障和网络传输情况、数据报文的格式错误和观测要素自身数据错误情况。自动站故障和网络传输直接影响数据到报。

在此,统计了2007年9月1日—2007年10月31自动站数据报文格式错误和数据质量错误,结果如图2所示。

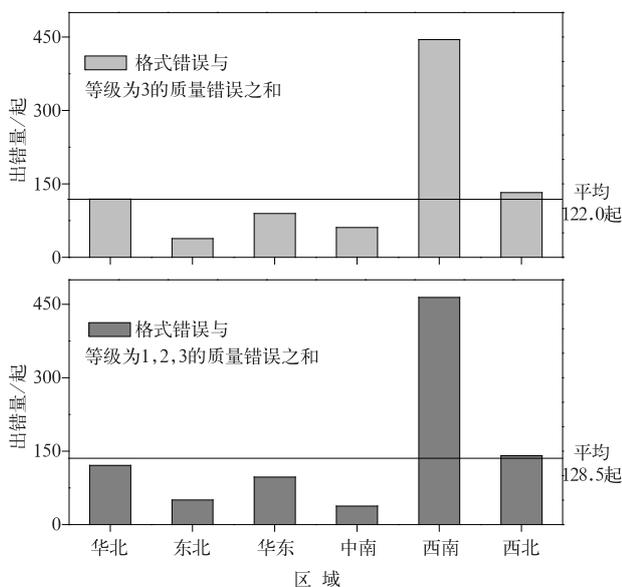


图2 2007年9—10月数据报文格式错误、数据质量错误和异常情况

Fig. 2 The datagram format error, data quality error and data abnormal in September and October of 2007

从图2可以看出,不同区域自动站出错量之间存在较大差异,西南地区错误情况最明显,其错误总量远远高于128.5起的平均水平;东北和中南地区出错量最少,远远低于平均水平。

自动站数据报文或数据要素出错情况与自动站运行效能之间存在很大关系。从图1和图2的对比中可以发现,西南地区数据出错情况最频繁,而且该区域数据到报率较其他区域相比也最低,导致该区域自动站的运行效能停留在一个较低的水平;相反,东北和中南区域数据出错量最小,不但如此,这两区域数据到报率也保持较高的水平,所以,这两个区域自动站效能也较高。

自动站的观测要素数据错误和报文异常来自于以下几个方面:设备故障、软件故障、外界干扰、设备维护不当或欠维护、台站参数设置不当等。在这些因素中,设备故障引起的观测数据错误和报文异常情况最严重,而引起设备故障的主要因素包括传感器老化、极端天气影响(风杯被冻结)、雷击等。西南地区一方面自动站数量较多,另一方面该区域地形、天气现象复杂多变,雷电现象多发;此外,由于自动站本身防雷方面存在缺陷,雷击导致传感器、采集器

损坏也较其他区域频繁。

### 3.2.1 数据报文格式错误

自动站采集的数据每小时正点向中国气象局以报文形式上传,在中国气象局对上传的原始报文进行解析,解析后的报文由4条记录组成:

第1行(本站基本参数)包括区站号、经度、纬度、观测场海拔高度、气压传感器海拔高度、观测方式共6组项目34个字,每组用一个半角字符分开。

第2行(器测项目)包括观测时间、风速、风向、气温、地温以及气压和能见度等共52个要素值262个字节,每组用一个半角字符分开。

第3条记录(小时中分钟降水量)120个字节,每分钟2个字节,即1~2位为第1分钟的记录,3~4位为第2分钟的记录……,如此类推,119~120位为第60分钟的记录;每分钟内无降水时存入“00”,微量存入“.”,降水量 $\geq 10$  mm时,一律存入99,缺测存入“//”。

第4条记录共23个要素值,每组用1个半角空格分隔,无该条记录内容时则省略,“=”加在分钟降水量之后。文件最后一条记录为NNNN,表示文件结束<sup>[6]</sup>。

根据上述编码规则,将这段时间内数据报文错误类型列举如下:①第1行和第2行观测要素字符长度超出或少于相应字段报文编码长度。例如:56144 314800 0983500 ~~31840~~31850 4。阴影部分长度应该为6个字节,实际文件中为7个字节。②观测要素缺失,用半角空格表示,不符合报文编码规范。③第1行中观测时间、站点经纬度及海拔高度信息缺失。④报文中出现乱码,如出现字母、“\*”、“-”等。⑤DOS状态下混杂UNIX符号,只回车不换行。⑥文件名中日期错误或不完整。例如:Z\_O\_AWS\_ST\_C\_BETJ\_02331204220000.TXT。处理日期显示为233年,严重错误。⑦无数据,缺测无标注。⑧分钟数据雨量为“0”,累积雨量为“////”。⑨分钟雨量数据为无法识别“-”,而累积雨量为“0”。⑩分钟数据雨量之和不等于累积雨量值。

为研究数据报文格式错误与数据可用性和可靠性之间的关系,在数据到报率不变情况下,做数据报文格式错误和数据要素质量错误与数据可用性和可靠性之间的偏相关分析,结果如表1所示。

从表1中可以发现,数据报文格式错误是影响数据可用性和可靠性的主要因子。数据报文格式错误可分为第1行格式错误、第2行格式错误和时间

错误等几类,为了研究数据运行效能究竟受哪类格式错误控制,还做出同期几类格式错误的分布图,如图 3 所示。

表 1 2007 年 9—10 月数据可用性和可靠性与数据质量情况和报文格式错误之间偏相关分析  
Table 1 The partial correlations analysis between AWS's availability, reliability and data quality abnormal and datagram format error in September and October of 2007

控制变量	变量	数据质量情况(质量控制等级为 1,2 和 3)	数据报文格式错误
到报率	可用性	-0.274	-0.961*
	可靠性	-0.49	-0.957*

注: \* 相关系数达到 0.01 的显著性水平。

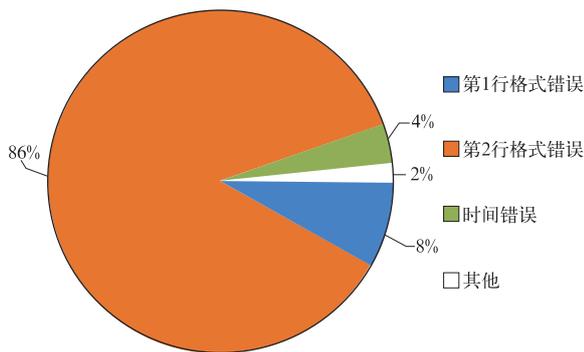


图 3 2007 年 9—10 月数据报文格式错误中各类格式错误出错量对比图

Fig. 3 The percentage graph of each datagram format errors in September and October of 2007

从各类格式错误所占比例来看,不同类型的格式错误出错的几率之间存在较大差别,第 1 行格式错误出错率占有报文格式错误的 86%,而第 2 行格式错误和时间错误的出错情况所占比例较低,说明第 1 行格式错误是自动站可用性和可靠性的主要限制因子。

第 1 行格式错误基本是由于数据处理软件造成的,而引起第 2 行格式错误的因素较多,设备损坏、软件处理故障及线路连接等都有可能引起此类错误发生。因此,按时对软件升级、对设备定期维护都有可能避免此类错误发生。

### 3.2.2 数据质量错误

自动站探测的气象要素有:2 min 平均风向、2 min 平均风速、10 min 平均风向、10 min 平均风速、最大风向、最大风速、瞬时风向、瞬时风速、极大风向、极大风速、小时降水量、气温、最高气温、最低

气温、相对湿度、最小相对湿度、水汽压、露点、本站气压、最高气压、最低气压、草温、最高草温、最低草温、地温、最高地温、最低地温、5 cm 地温、10 cm 地温、15 cm 地温、20 cm 地温、40 cm 地温、80 cm 地温、160 cm 地温、320 cm 地温、海平面气压等<sup>[6]</sup>。目前自动气象站数据质量控制的方法有要素连续性检查、气候极值检查、气候学界限值检查、要素一致性检查<sup>[7-23]</sup>。在综合气象观测系统运行监控平台中,通过要素连续性检查、气候极值检查、气候学界限值检查、要素一致性检查和综合检查方法对上述要素进行质量控制。

在此,对两个月数据质量的错误情况进行统计。从统计结果来看,各要素出错量之间存在很大差别,地温要素占各要素出错比例最大,其次为气温要素,而降水、风速、水汽压和湿度要素错误量很低(图 4)。

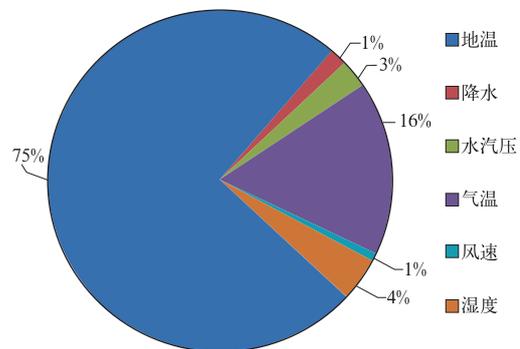


图 4 2007 年 9—10 月数据质量错误中各类要素出错量对比图

Fig. 4 The percentage graph of each data quality errors in September and October of 2007

不但如此,各地温要素之间出错几率也存在较大的差别,其中 320 cm 地温出错量所占比例较大,其次为 160 cm 地温,5 cm 地温出错量所占比例最小,各层次地温要素之间基本呈现随深度加深出错几率上升的趋势(图 5)。

通过对 2007 年 9—10 月自动站各观测要素出错情况分析发现,自动站观测数据错误主要呈现如下特点:① 观测数据往往表现为同一时段各类要素同时显示为异常,且变化趋势一致。此种情况主要是因数据采集器故障引起,可能是采集器遭雷击或数据采集器与传感器之间的连接线松动,而遭雷击的现象比较普遍。② 观测要素出错时还表现为单个要素超极值或保持恒定,此现象往往是因传感

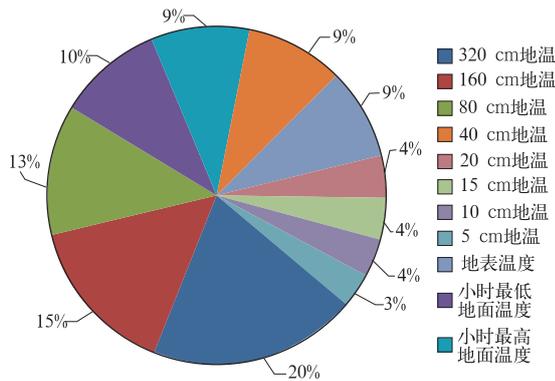


图5 2007年9—10月数据质量错误  
中地温各要素之间出错量对比图

Fig.5 The percentage graph of each ground  
temperature errors in September  
and October of 2007

器本身造成。③观测要素出错还表现为多个或单个要素在短时间内以较大的幅度波动变化。

#### 4 小 论

通过上述分析,发现影响自动站运行效能的主要因素如下:

1) 数据到报情况是影响数据可用性和可靠性的主要因子,它是评价数据可用性和可靠性的前提。

2) 在数据到报率一定的情况下,数据报文格式错误较数据要素质量错误更多地影响数据的可用性和可靠性。

3) 在数据格式错误中,第1行格式错误和第2行格式错误对自动站效能的影响程度也不尽相同,相比较而言,第1行格式错误对效能影响程度更大。

4) 从数据质量情况来看,各观测要素之间对自动站效能影响程度也存在一定差别,地温要素错误是主要影响因子,其次为气温;在地温要素中,不同层次的地温对自动站效能的影响也存在差别,320 cm地温的影响程度最大,而5 cm地温的影响程度最小,基本呈现出越往地面深处,地温要素对效能的影响程度越大的趋势。

#### 参 考 文 献

[1] 冯筠,高峰,黄新宇. 构建天地一体化的全球对地观测系统——三次国际地球观测峰会与 GEOSS. 地球科学进展, 2005, 20(12): 127-133.  
[2] 中国气象局大气探测技术中心全网运行监控与保障室. 气象

探测全网运行 2007 年度监控分析报告. 北京: 中国气象局, 2007: 1-3.

- [3] 中国气象局大气探测技术中心. 气象探测全网运行监控系统技术报告. 北京: 中国气象局, 2007:1-14.  
[4] 梁海河,张春晖,孟昭林. 实时气象观测监控和分析系统//2008 年度气象探测技术学术年会文集. 北京: 中国气象局气象探测中心, 2008.  
[5] 中国气象局. 地面气象观测规范. 北京: 气象出版社, 2003: 133-139.  
[6] 中国气象局. 地面气象观测数据文件和记录簿表格式. 北京: 气象出版社, 2005: 9-46.  
[7] 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. 气象, 2006, 32(3): 107-112.  
[8] 刘小宁,任芝花. 地面气象资料质量控制方法研究概述. 气象科技, 2005, 33(3): 199-203.  
[9] 王伯民. 基本气象资料质量控制综合判别法的研究. 应用气象学报, 2004, 15(增刊): 50-59.  
[10] 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. 应用气象学报, 2006, 17(1): 37-43.  
[11] 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. 气象, 2007, 33(1): 19-24.  
[12] 方炳兴. 常规气象资料质量的综合控制. 气象, 1994, 20(2): 33-36.  
[13] 杨贤为. 气候应用专用数据库气象资料质量检验. 气象, 1998, 24(12): 33-36.  
[14] 陶士伟,张跃堂,陈卫红,等. 全球观测资料质量监视评估. 气象, 2006, 32(6): 53-58.  
[15] 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. 气象学报, 2005, 63(4): 526-533.  
[16] 李铁,邹立尧,国世友. 东北地区低温气象资料数据集及其质量控制. 应用气象学报, 2004, 15(增刊): 164-167.  
[17] 王伯民. 基本气象资料质量控制综合判别法的研究. 应用气象学报, 2004, 15(增刊): 50-59.  
[18] 熊安元. 北欧气象观测资料的质量控制. 气象科技, 2003, 31(5): 314-320.  
[19] Secretariat of WMO. Guide To Meteorological Instruments and Methods of Observation (Sixth Edition). 1996.  
[20] Lanzante J R. Resistant, robust and nonparametric techniques for the analysis of climate data: Theory and examples, including applications to historical radiosonde station data. *Int J Climatol*, 1996, 16:1197-1226.  
[21] Gandin L S. Complex quality control of meteorological observation. *Mon Wea Rev*, 1988, 116(5): 1137-1156.  
[22] Vejen F, Jacobsson C, Fredriksson U, et al. Quality control of meteorological observations automatic methods used in the Nordic countries. *Climate Report*, 2002, No. 8/2002, KLI-MA.  
[23] Eischeid J K, Baker C B, Karl T R. The quality control of longterm climatological data using objective data analysis. *J Appl Meteor*, 1995, 34: 2787-2795.

## The Statistic of Automatic Weather Station's Efficiency

Li Yan Liang Haihe Meng Zhaolin Pei Chong Shi Cheng

(CMA Meteorological Observation Center, Beijing 100081)

### Abstract

Over 2100 state automatic weather stations (AWS) can be monitored by the Atmospheric Observing System Operations and Monitoring (ASOM) platform of CMA Meteorological Observation Center at present. They are divided into six groups by areas, which are north, northeast, east, south central, southwest, northwest China. The efficiencies of the AWS are statically analyzed in the respect of data arrival rate, equipment availability and reliability. Influencing factors on the data format errors of datagram and the qualities of data element errors are also analyzed. The results indicate that the data arrival rate, equipment availability and the reliability of all the 2100 state AWS maintain over 80%, with a generally consistent trend. The AWS operation efficiencies of the northeast region are the highest and those of the southwest are the lowest. The data arrival rate is the precondition of the equipment availability and reliability. With the data arrival rate fixed, the format error of datagram has more influence on the equipment availability and reliability comparing to the data elements errors. The 1st line data format errors have more influence on AWS operation efficiencies than the 2nd line data format errors. The ground temperature is the main elements affecting AWS operation efficiencies. Among the ground temperature factors, the different levels of ground temperature have different impacts on AWS efficiencies; 320 cm ground temperature has the most significant influence, while the 5 cm ground temperature's influence is relatively lower. It shows a trend that deeper layers of the geothermal elements have greater influences on AWS efficiencies.

**Key words:** ASOM; AWS; operation efficiency