

北京自动气象站实时数据质量控制应用

窦以文^{1,2} 屈玉贵¹ 陶士伟³ 胡保昆²

(1. 中国科学技术大学, 合肥 230026; 2. 北京市气象局; 3. 国家气象中心)

提 要: 自动气象站(AWS)观测资料是中尺度、短时效天气预报不可缺少的重要资料来源之一。但由于AWS观测资料的特殊性,其质量问题不同程度地阻碍了它的使用。因此,AWS观测资料的质量控制越显重要。北京市气象局设计的自动气象站观测数据质量控制的方案主要是物理极值检查、历史极值检查、内部一致性检查、时间一致性检查、空间一致性检查等方法并把它应用于北京地区2007年奥运气象服务演练,取得了令人鼓舞的效果。结果表明:该质量控制方案的建立,能够实时反映北京市气象局自动站观测数据的可靠性,同时也反映出数据的缺测和及时性也存在一些问题。此系统的运行对2008年奥运气象服务,特别是对INFO2008和城市运行实时数据服务的可靠性将会发挥重要作用。

关键词: 自动气象站 质量控制 气象信息处理

The Application of Quality Control Procedures for Real-time Data from Automatic Weather Stations

Dou Yiwen^{1,2} Qu Yugui¹ Tao Shiwei³ Hu Baokun²

(1. University of Science and Technology of China, Hefei 230026;

2. Beijing Meteorological Bureau; 3. National Meteorological Center)

Abstract: Quality Control for data from automatic weather stations is very important with the development of society and technology. Beijing Meteorological Bureau has created real-time Quality Control system for automatic weather stations (AWS) which has applied on the 2007 Olympic Games Drilling. The scheme of Quality Control for data mainly includes methods, such as plausible value check, time consistency check, internal consistency check, history value check, spatial check, etc. The application of the data quality control system to 2007 Olympic Games Drilling in Beijing area shows that Beijing AWS data is reliable, but there are some problems on missing and timely data. This system will become a good example of real-time processing for AWS Quality Control in Beijing 2008 Olympic Games weather services.

Key Words: automatic weather stations quality control meteorology information processing

引 言

随着现代化技术和世界经济的发展,中尺度、短时效、特殊行业天气预报的需求日益高涨。气象预报及资料同化技术也日新月异。要满足和适应日益高涨的气象服务需求,描述中尺度、短时效天气现象的观测资料是必不可少的。北京市气象局为了适应2008年奥运会和中尺度预测的气象观测的需要,到2005年底已建立了城区8个四要素自动气象站,18个六要素自动气象站,9个多要素自动气象站;郊区有44个四要素自动气象站,12个六要素自动气象站,15个多要素站。共106个自动气象观测站和2个道面监测站。见图1(上为城区图,下为郊区图)

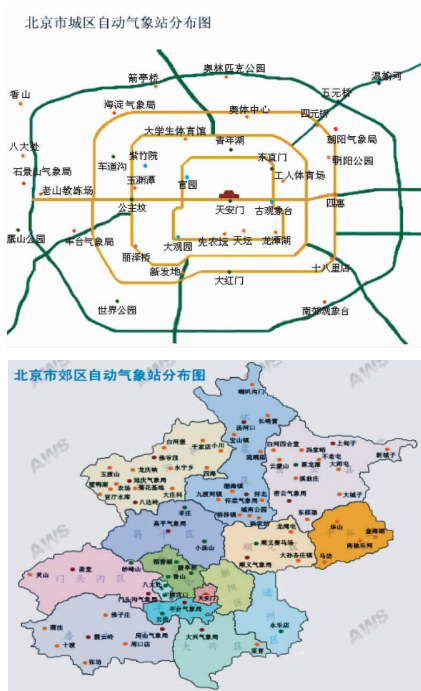


图1 北京市自动气象站分布

由人工观测到自动气象站观测是地面观测手段的重大变动,有必要对自动观测资料

进行实时审核,便于资料使用。自动与人工观测数据的差异是多种原因造成的。这些原因包括仪器的原理不同,观测时间和空间不同,采样方式与算法不同,观测时次不同等等^[1-2,5]。目前,国内对自动气象站数据质量控制主要是着眼于历史数据,而对实时数据质量控制的应用较少。

建立一套实时 AWS 观测资料数据质量控制系统,为2007年奥运气象服务演练和2008年奥运气象服务,特别是 INFO2008 数据和城市运行数据服务的可靠性具有非常重要的意义。本文介绍了 AWS 观测资料数据质量控制系统的结构,并分析该系统应用效果,得出一些结论。

1 AWS 观测资料数据质量控制系统

现在我们已把各种气象数据进行实时上网显示、奥运气象服务和业务使用。为了保证气象观测的代表性、客观性、及时性、完整性,在资料使用前,必须对自动观测资料进行实时审核。

1.1 网络设计

北京市气象局数据传输主要是使用 GPRS/CDMA 网络、同步数字系列 SDH 专线、主动拨号三种传输方式。基于 GPRS/CDMA 网络是新一代无线通讯方式,覆盖面广,可靠性高,投入、维护运行成本低,扩展性强。特别适合如自动站数据传输频繁、一次数据传输少的远程系统^[3]。SDH 网络在传输正常业务时,同时传输自动站数据,这种传输方式快速、稳定、可靠,但是费用较高,传输地点固定。拨号传输作为现代网络传输方式虽然已落后,费用高,但是作为备份线路是很好的选择,拨号网络一向比较稳定。

1.2 自动质量控制方法

目前常用的数据质量控制方法有很多,主要有极值的检查、时间序列的检查、空间序列的检查、利用神经网络等人工智能技术进行质量控制检查等^[4-5]。用于历史数据质量控制的方法中有些方法在实时自动站质量控制同样适用。如内部一致性,历史极值的判定等。但是实时信息的检验和历史质量控制有一些不同,如对于历史质量控制可以知道该时刻前后时刻和空间的数据,而实时数据实时性强,只知道过去时刻和空间的数据。自动站数据质量控制是一个非常复杂的过程,在自动站设备采集的数据传输到通信服务器后的质量控制一般经过这样几个过程:

(1) 原始数据的检查:包括年、月、日的检查、数据字符的检查、数据格式的检查、区站号的检查。

(2) 极值的检查:包括气候极限值的检查和历史极值的检查。

表 1 气候极限值

气象要素	最小值	最大值
空气温度	-80℃	60℃
露点温度	-80℃	35℃
地面温度	-80℃	80℃
土壤温度	-50℃	50℃
相对湿度	0	100%
本站气压	500hPa	1100hPa
风向	0	360°
2 分钟平均风速	0	75m·s ⁻¹
5 分钟间隔降水总量	0	40mm

对于气候极限值的检查,如果要素不在表 1 范围中,可以标定为错误值。

历史极值选自 1971—2006 年地面各要素观测资料的最大值和最小值,超过历史极值范围的资料为可疑资料。历史极值的检查包含的要素是本站气压、空气温度、相对湿度、风速、5 分钟累计雨量。

初始极值库的确定是在原有极值的基础

上加减 3 倍标准差。通过实时信息的追踪,更新极值库。标准差 σ 的计算方法如下:

$$\sigma = \left[\frac{1}{n-1} \sum_{i=1}^n (x_i - x)^2 \right]^{1/2}$$

其中 n 为年数, x_i 为月观测平均值, x 为多年同月观测平均值。

(3) 时间序列的检查:气象要素的变化在时间上具有连续性。

最大变化值检查(见表 2):各气象要素当前时刻与过去 5 分钟的值变化,如果超过表 3 中怀疑极限值且小于错误极限值,作为怀疑值。如大于等于表 3 中错误极限值,作为错误值。

表 2 临近时间(5 分钟时间间隔)序列最大变化值

气象要素	怀疑极限值	错误极限值
空气温度	3℃	4℃
露点温度	3℃	4℃
地面温度	5℃	10℃
5cm 土壤温度	0.5℃	1℃
10cm 土壤温度	0.5℃	1℃
20cm 土壤温度	0.5℃	1℃
80cm 土壤温度	0.3℃	0.5℃
100cm 土壤温度	0.1℃	0.2℃
相对湿度	10%	15%
本站气压	0.5hPa	2hPa
平均 2 分钟风速	10m·s ⁻¹	20m·s ⁻¹

表 3 临近时间序列最小变化值^[5]

气象要素	怀疑极限值
在过去一个小时的空气温度	0.1℃
在过去一个小时的露点温度	0.1℃
在过去一个小时的地面温度	0.1℃
土壤温度	没有最小值
在过去一个小时的相对湿度	1%
在过去一个小时的大气压	0.1hPa
在过去一个小时的风向	10 度
在过去一个小时的风速	0.5m·s ⁻¹

也可以使用下列算法检查:

使用的算法是:

$$|V_i - V_{i-1}| + |V_i - V_{i+1}| \leq 4\sigma V$$

其中, V_i 是当前值, V_{i-1} 是上一时刻值, V_{i+1} 是下一时刻值, σV 为至少 5 分钟估计的标准

的偏差。

最小允许变化值检查。各气象要素当前时刻与过去一个小时的值变化,如果超过表 3 中怀疑极限值,作为怀疑值。

(4) 空间序列的检查:根据气象要素具有的空间相关性的检查。选择 5 个临近站作参考,求出被检要素相关系数,此方法也可用于时间序列和不同要素相关检查。

$$Y_{i,j} = a_j + b_j X_{i,j}$$

其中 $X_{i,j}$ 为第 j 个参考站第 i 时刻要素实测值, $Y_{i,j}$ 为被检站第 i 时刻要素估计值。通过真实数据与估计值进行比较,判定是否在一定范围内。

(5) 一致性检查:是指内部一致性。

内部一致性是指根据同一测站同时测得的参数之间应保持一致性原则。如风速为 0,则风向为静风,如风向为静风,则风速为 0;最大风速小于等于极大风速。露点温度小于等于空气温度等^[6]。分钟累计降水量小于等于小时累计降水量。

(6) 利用曲线趋势人工检查奇异值。

图 2 中曲线从上到下依次为凤凰岭、昌平、北安河、居庸关、八达岭的 2007 年 8 月 12 日至 8 月 13 日气压曲线图,如图 2:居庸关当日气压到达极小值 963.3hPa,从系统单点历史极值检查,这是一个怀疑值,但是从曲线的连续性来看,图中的值与前后观测值在时间上是连续的,空间上也没有突变,因此这是一个可靠值。

(7) 综合决策算法:综合分析各质量控制方法的结果。把(2)至(5)算法都加上一个权重或优先级别,综合各种结果,最后给出数据能否使用的结论。

对于经过质量控制后的数据如何使用,方法有很多种。有的是采用概率百分比,有的定义数据的等级等。应用系统中主要使用的是标志来定义数据的质量(见表 4)。对于这些经过质量控制的数据,不同的需求对应

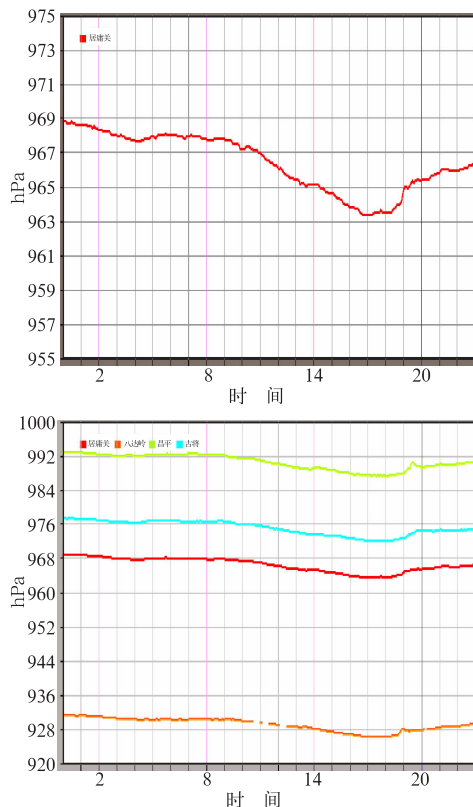


图 2 利用对比曲线人工检查奇异值

于不同的数据应用。如 GPS 气象数据对于质量控制达到错误标志的数据用周围 10km 的数据填充。奥运气象网站的数据如果标志为错误,则填充缺测。对于怀疑数据,不进行处理。一般情况下,不更改原始数据,每条数据都同时生成 netCDF 和文本两种格式的质量控制标志,任何使用自动站数据的用户,主要是通过标志来识别数据的质量^[7]。

表 4 质量控制标志

标志	说明
0	要素数据没有审核
1	要素数据可靠
2	要素数据怀疑
3	要素数据不可用
4	要素数据缺测
5	没有此观测要素
6	人工修改要素数据
7—9	保留使用

2 AWS 观测资料数据质量控制系统应用效果分析

图3中为2007年8月1日至26日奥运气象演练的自动气象站数据。

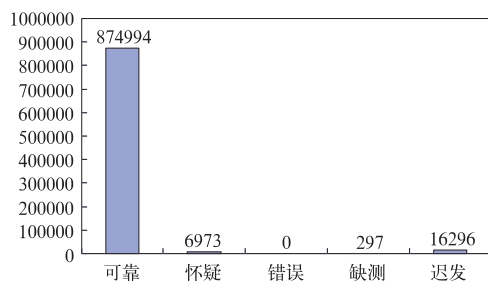


图3 自动站质量控制结果

图3横坐标为质量控制结果的5种情况(可靠、怀疑、错误、缺测、迟发),纵坐标为15个代表站 \times 8个要素 \times 26天 \times 24小时 \times 60分钟/5分钟。

要素为气压、温度、相对湿度、瞬时风向、瞬时风速、极大风速的风向、极大风速、分钟的降水量;定义观测5分钟后的数据为迟发。26天是指2007年8月1日至26日奥运气象演练期间。从图3中可以看到北京市气象局自动气象站数据可靠性非常高,但是数据缺测和及时性是主要问题。

3 结论

(1) 由于奥运气象服务数据量非常庞大,不可能完全像过去一样进行人工审核,而且人工审核从时效和工作量上远远不能满足

奥运气象服务需求。而目前很少有这方面的应用,本系统通过气候极值检查、历史极值检查、要素相关性检查、时间一致性检查、空间一致性检查等方法,基本能够满足2008年奥运自动站质量控制气象服务需要。

(2) 北京市气象局自动气象站数据可靠性非常高,但是数据缺测和及时性是主要问题。这是2008年自动气象站信息服务需要想办法改进的地方。

(3) 历史极值的检查是一个不断完善的过程,需要根据实况数据更新历史极值表,这只有实时系统才能做到的。

(4) 对于自动检查的怀疑数据,可以在时间连续性和空间一致性进行人工检查。

参考文献

- [1] 胡玉峰. 自动与人工观测数据的差异[J]. 应用气象学报, 2004, 15(6): 719-726.
- [2] 连志鸾. 自动站与人工站观测记录的差异分析[J]. 气象, 2005, 31(3): 48-52.
- [3] 刘聪, 顾建, 吴国平, 等. 基于GPRS的远程气象观测数据实时采集传输系统及其应用[J]. 应用气象学报, 2004, 15(6): 712-717.
- [4] 杨贤为. 气候应用专用数据库气象资料的质量检验[J]. 气象, 1998, 24(12): 33-36.
- [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(03): 199-200.
- [6] Igor Zahumensky. Guidelines on Quality Control Procedures for Data from Automatic Weather Stations. Expert team on surface technology and measurement techniques first Session Geneva, Switzerland, 13-16 October 2004.
- [7] Mark A. Shafer, Christopher A. Fiebrich, AND Derek S. Arent, etl. Quality Assurance Procedures in the Oklahoma Mseonetwork[J]. J. Atmo. Oceanic Technol., April 2000: 474-494.