

敦煌戈壁气象塔站资料的质量控制

王超, 韦志刚, 李振朝

(中国科学院寒区旱区环境与工程研究所, 甘肃 兰州 730000)

摘要:讨论了气象观测资料的质量问题,并分析和强调了塔站资料质量控制的重要性,结合地面观测资料的质量控制方法,提出了逻辑极值检查、僵值检查、时间一致性检查、相似一致性检查、决策算法、质量控制码的标注和人工干预检查等组成的一套针对塔站资料的质控方法。利用提出的质量控制流程,对敦煌戈壁塔站 2001 年 4 月至 2009 年 4 月共 9 a 的资料进行质量控制。结果表明:该流程能很好的找出缺测、错误、可疑的数据并给与标注,特别是在处理过程中,结合了综合判别法,使得该检验结果更为可靠。结果表明:塔站观测资料的质量是比较好的,正常数据占总数据的 91.2%,非正常数据只是少数情况,其中大部分为缺测数据,占其 78.7%。错误数据占总数据的 1.93%,其中大部分为僵值数据,占其 87.02%,其次分别为一致性和逻辑极值错误。

关键词:敦煌;气象塔;质量控制;质量监视评估

中图分类号:P413.2

文献标识码:A

引言

质量控制主要根据气象学、天气学、气候学原理,以气象要素的时间、空间变化规律和各要素间相互联系的规律为线索,分析气象资料是否合理。质量检查的主要目的是检测数据的有效性,并做到对错误资料及时纠正^[1]。气象观测资料是气象业务和气象研究的基础资源^[2],各种数据序列的质量直接影响着气象业务和气象研究的应用效果、准确性和可靠性。气象资料在提供给用户前,经历了仪器观测、记录、传输,甚至编码、解码等重要环节,各个环节都可能引起数据错误^[3]。一般来讲未经质量检查的数据是不能提供使用的,因为不准确的数据往往会让使用者的研究得不到准确的结果,有时甚至是完全相反的结论。例如在陆面过程参数化中,用梯度风、温度资料计算输送系数或粗糙度时,往往微小的误差通过一系列计算而放大到我们不能接受。因此,质量控制(QC)非常重要,它是观测站点资料质量保证的基础,应该贯穿于从站点建立到数据产品生成以及应用的整个过程^[4]。近年来,随着国际上对数值预报的重视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也越来越高。因此,在国内外,质量控制(QC)和质量评估(QA)技术发展很快,由此而发展了许多气象观测资料质量控制的方法^[3-12]。

大气边界层是指大气层最底下的一个薄层^[13],它是地球—大气间物质和能量交换的桥梁,在大气边界层特别是近地层的研究中,目前也许还没有比气象观测塔更能提供丰富、翔实资料的观测手段,仅就资料的连续性一项指标,就很难有其它观测方式相比拟。因此,研究塔层资料的质量控制问题,对实现边界层实测资料的质量控制具有相当的实用价值^[14]。伴随着新技术的出现以及计算机的更新,仪器设备发展的更加灵敏,精度更高,但是利用塔层进行测量所带来的一些问题至今依旧存在。气象观测塔出现错误资料的几率远远高于常规地面观测资料^[15]。例如:数据集中有不合常理的气象要素值、连续几个时次甚至几天的值都一样、或是要素值之间出现矛盾现象等。这些问题都可以使资料的准确性降低到令人难以置信的程度,如果不修正或剔除这些异常值,在随后的理论分析中必然会出现许多

收稿日期:2010-03-05;改回日期:2010-04-08

基金项目:国家重点基础研究发展计划资助(2009CB421402);国家自然科学基金项目(40730952)和中国科学院知识创新工程重要方向项目 KZCX2-YW-220 共同资助

作者简介:王超(1986-),男,四川绵阳人,在读博士,主要从事干旱气候研究。E-mail:wangchao@lzb.ac.cn

问题。因此,在对原始资料进行理论分析前,有必要修正和剔除这些异常值,这就是所谓的质量控制,尽管质量控制工作是非常基础性的,然而却是十分复杂,对于边界层资料,这种复杂性尤为突出。对此本文结合了地面资料的常用的质量控制方法,提出对于气象塔站观察资料的一套较为完整的资料质量控制系统,对国家 973 项目设在敦煌戈壁的气象塔站资料进行了质量控制的分析研究。

1 试验数据介绍

1.1 试验介绍

从 2000 年 5 月开始,科技部首批启动的《国家重点基础研究发展规划项目》之一“我国重大气候和天气灾害形成机理和预测理论的研究”项目在对我国气候形成和演变关键作用的区域规划了陆上和洋面 2 大野外试验,其中在跨甘肃和青海 2 省的西北广大干旱和高寒地区开展了西北干旱区陆气相互作用野外观测试验“NWC - ALIEX”^[16]。其中设在敦煌戈壁的气象观测塔站在相关项目和部门的支持下一直延续至今,并成为 2009 年启动的《国家重点基础研究发展规划项目》“全球变暖背景下东亚能量和水分循环变异及其对我国极端气候的影响”的一个野外观测站,继续进行相关的观测试验。

1.2 观察场地地理、气候背景

敦煌位居甘肃河西走廊西端,甘、青、新 3 省交界的三角地带,境内东有三危山,南有鸣沙山,西临库姆塔格沙漠,北接戈壁与天山余脉相连,地势南北高中间低,自东南向西北略微倾斜,历来就有“敦煌盆地”和“戈壁绿洲”之称。

敦煌双墩子戈壁观测站(图 1)在塔克拉玛干沙漠的东端,位于 40°10'N 和 94°31'E,海拔 1 150 m,距敦煌绿洲边缘的最近距离大约为 7 km;距离偏南边的鸣沙山大约 20 ~ 30 km,微气象中心站的观测场地为平坦的沙石戈壁。温度日较差较大,白天有记录的最高地表温度接近 60 °C,有记录的最高气温高达 43.6 °C,最低气温曾达 -28.5 °C;天空少云,太阳辐照强,日照充分,年无霜期 178 d;该地区多风,为沙尘暴多发区之一^[17]。地表年平均大气压为 873 hPa,该观测站处在极端干旱的气候区,年平均降水量约为 39 mm^[18],盛行风向为东风,该地区干旱少雨,地表干燥,蒸发力极强,属于大陆干旱性干旱气候^[19]。

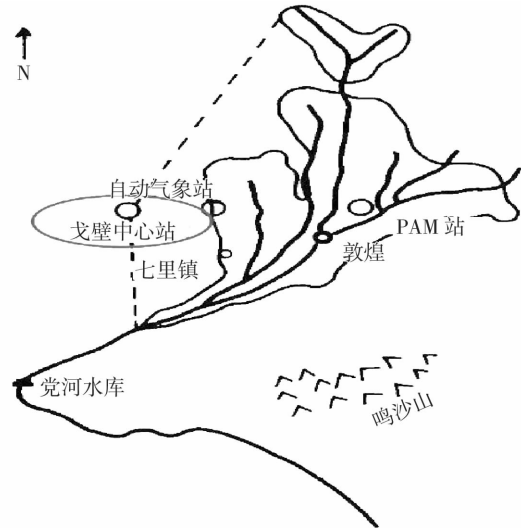


图 1 敦煌塔站的地理位置^[20]

Fig. 1 The location of Dunhuang meteorology tower

1.3 数据介绍

双墩子戈壁微气象观测点观测项目主要包括铁塔的风、温、湿和地温的梯度观测,铁塔的风、温、湿共有 4 层,分别在 18、8、4 和 2 m 高(4 m 和 2 m 在 2000 年 6 月 5 日后分别移到 2 m 和 1 m 高处),并在 10 m 高处有风向观测;地温有 7 层,分别为 0、5、10、20、40、80、180 cm。资料的采样频率在 2001 年 8 月 19 日从 1 h 一次加密到 0.5 h 一次。表 1 给出在塔站上使用的仪器的型号、精度、范围和生产厂家。

表 1 各传感器的型号和概况

Tab. 1 Specifications and performances of the sensors

仪器型号	精度	测量范围	生产厂家
风向传感器 EL15-2	3°	0~360°	德国蒂斯天津生产
风速传感器 1022S-1	0.1 m/s	0.3~60 m/s	美国 BELFORT 公司
空气温度 PT-100-385	0.05 °C	-40~80 °C	芬兰 VAISALA 公司
空气湿度 HMP45D	2%	0~100%	芬兰 VAISALA 公司
地表温度 PT-100-385	0.05 °C	-40~80 °C	芬兰 VAISALA 公司
土壤温度 PT-100-385	0.05 °C	-40~80 °C	芬兰 VAISALA 公司
数据采集器 DT-600			澳大利亚 DE 公司

2 质量控制方法

质量控制是一个系统工程,它包括数据的采集以及数据的校对、审核、质量检查和标注分级等环节,最终成为合格的具有“三性”(代表性、准确性、比较性)的气象资料。其方法包括:范围检查、逻辑

极值检查、时间一致性检查、相似一致性检查等,这些方法被普遍应用到地面气象资料的质量控制中。

以前数据的质量控制通过人工审查、经验分析等方法剔除可疑数据。在工作中,预报人员凭经验就可以判断出严重的错误数据,但随着计算机的普及,信息量的增大,观测数据的质量控制就必须由计算机来完成。因此,研究和采用各种数据质量控制的方法,是非常重要的实际工作。好的资料无需特别出色,重要的是其质量应该是已知的和可证实的。

2.1 误差来源

气象数据误差按产生原因,可分为4种类型^[21],即系统误差、粗大误差、随机误差和微气象误差。(1)系统误差:对同一被测量进行多次测量,误差大小或符号保持恒定,或按一定规律变化,这类误差称为系统误差。(2)粗大误差:明显偏离真值的误差为粗大误差。在自动观测系统中,粗大误差主要由自动站观测仪器异常及在数据编码、处理、传输、存储及解码等业务过程中产生,例如传感器由于电压的不正常变化而产生错误记录,这类误差的观测值一般很明显,如果不剔除,将对后面的分析产生很大的影响,这种误差在质量控制过程中容易被发现。(3)随机误差:在相同条件下对同一被测量进行多次测量,由于受到大量微小随机因素影响,测量误差大小和符号没有一定规律,且无法估计,即为随机误差。随机误差是无法避免的,其总体服从正态分布。(4)微气象误差:是由小尺度天气系统扰动引起,由于观测系统时空分辨率原因,这些天气系统一般不会被完全观测到。但当观测到这种天气系统时,其观测数据和周围台站同时间相比,被认为是异常值。但是在实际工作中,以上分类并不严格,不同误差的分界也并不明显。

系统误差由一段时间的对比观测容易确定,可根据确定结果对资料进行修正。粗大误差应给予删除,然后可根据情况,前后观测时次或空间范围插值运算加以补充。随机误差是仪器和测量方法造成的,是不能被避免的。微气象误差在极端的天气条件下是可以出现的,而在一般的质量控制中,很容易被诊断成虚假值,所以在气象数据质量控制时,查出的异常资料,应进一步做综合性质量控制与分析,以判断这些资料是否代表了真正的极端天气气候事件。

2.2 处理流程

在本文的自动质量控制流程(图2)中包括:逻辑

极值检查、僵值检查、时间一致性检查、相似一致性检查、决策算法、质量控制码的标注和人工干预检查。前6个过程是有计算机自动完成,最后一步需要了解仪器性能、具有观测经验的人员来完成。

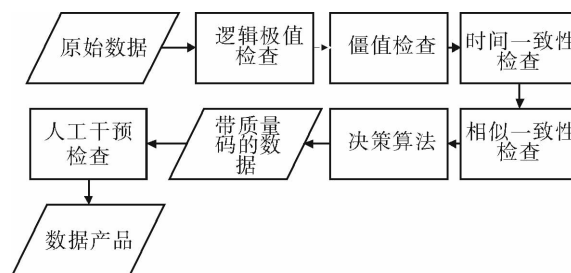


图2 质量控制流程图

Fig. 2 The flow chart of quality control routine

2.2.1 逻辑极值检查

这项检查中包括逻辑检查和极值检查。首先,对数据采集时间的合法性进行检查。数据采集是按时间和格式进行记录的,并找出缺测记录,以便于后期数据的处理和判断仪器是否正常工作^[22]。其次,我们给观测值一定范围(表略),也就是观测极值。极值的给定是观测仪器的范围和根据当地历史上曾出现过的最大值和最小值,这跟当地的纬度、海拔高度、气候特点有很大关系。

国家气象信息中心使用的质量控制方案^[23],在给定台站极值参数时,根据各个站各要素的平均值加减4倍的标准差作为各站的要素极值。但在本文,由于戈壁地区气象要素变化幅度很大,如果按上述方法会出现很多“可疑值”。所以,在本文的极值给定时,我们根据仪器的观测范围和已观测的历史数据,按季节给出对应的极值,以便精确的检查,保证数据的质量。其中风向、风速和湿度的极值范围在各个季节都是固定的,而温度要素的极值按季节(3、4、5月为春季,6、7、8月为夏季,9、10、11月为秋季,12、1、2月为冬季)给出。

2.2.2 僵值检查

僵值检查,也称为最小变率检查,由于观测仪器的失灵可能导致各要素观测值长时间不变,从而数据记录不真实,产生僵值记录。僵值检查采用相邻时次资料做差法,它要求给定时间范围内观测要素所有2相邻时次资料差的最大值大于给定判据。在本文规定在48次记录(24h)中,要素的变化幅度都在仪器的精度范围内,则认为是僵值。也就是某要素值与前一时刻的值之差连续48次小于特定的值

D 。值得注意的是,在底层地温变化中,经常出现连续 24 h 的变化在精度范围内,所以我们进一步缩小范围,如果不这样做,会误将正常观测值作为僵值处理。表 3 是对敦煌塔站资料设定的僵值检查的阈值。

表 3 僵值检查的阈值
Tab.3 The limit of static value check

要素	阈值	要素	阈值
风向	2	10 cm 地温	0.05
风速	0.5	20 cm 地温	0.001
气温	0.1	40 cm 地温	0.001
湿度	0.5	80 cm 地温	0.001
地表温度	0.1	180 cm 地温	0.001
5 cm 地温	0.05		

2.2.3 时间一致性检查

与僵值检查相反,时间一致性检查是从观测要素时间变化率的另一个角度检查观测的合理性。气象要素随时间的变化具有一定的规律,利用连续变化原理来检查观测信息或观测要素的时间变化率,识别出不理想的突然变化。当要素资料超出一定时间内的变化范围,则该资料视为可疑。它适用于高的时间分辨率,因为相邻样本的相关性随着时间分辨率增加而加强。检查判据与样本的时间分辨率有关。本文采用时变检查^[7]来实现时间一致性检查方法,时变检查主要是根据要素在某一时段内可能变化范围判断该要素值的质量,其中风向资料由于变化太大,故此不对其做此检查。该方法是考察某一要素观测值 X_i ,与其前一时刻和后一时刻观察值分别作差的绝对值的积的大小,也就是考察 $|X_i - X_{i+1}| |X_i - X_{i-1}|$ 的大小,并规定阈值 D (表略),如果 $|X_i - X_{i-1}| |X_i - X_{i+1}| < D$ 则标记为 0 (正常),如果该值大于 D ,则标记为 1 (怀疑),如果该值大于 $2D$,则标记为 2 (错误),当前后有缺测时标记为 8 (无法判断)。这种方法能有效地找出粗大误差记录。

2.2.4 相似一致性检查

气象塔的观测中,同一要素往往有几层不同高度的观测值,于是可以在相邻不同层次要素间进行比较,当这 2 个观测值之差大于给定的判据,则认为数据可疑。相似性检查方法可应用到温度、湿度和风速等要素。在本文的应用中,把 4 层要素分为 2

组,即 1 m、2 m 为一组;8 m、18 m 高度为一组,分别求其差的绝对值 A ,然后规定阈值 D ,在 $< D$ 范围内标记为 0 (可信),在 D 与 $2D$ 之间的标记为 1 (怀疑), $> 2D$ 的标记为 2 (错误);如果一组中有一个数据没有通过前面的检查,则标记为 8 (无法判断)。虽然资料集中也有 7 层土壤温度,但是表层温度变化很大,而低层地温变化较小,所以没有对土壤温度进行此检查(表略)。

2.2.5 决策算法

又称质量控制码综合判断,它是根据前面 4 个步骤,逻辑极值检查、僵值检查、时间一致性检查、相似性检查的质量控制结果,采用“气象资料质量控制综合判别法^[11]”进行一次质量控制码的综合性质量判断。如果该资料没有通过逻辑极值检查或僵值检查,标记为 2 (错误),如果没有通过时间一致性检查或相似性检查,质量码由表 4 给出。

表 4 决策算法的质量码

Tab.4 The quality code of synthetic method

时间一致性	相似一致性	综合质量码	时间一致性	相似一致性	综合质量码
0	1		8	1	
1	0	0	1	8	
0	0		1	1	
0	2		2	1	2
2	0		1	2	
0	8	1	8	2	
8	0		2	8	
8	8	8	9	9	9

注:0—正确,1—怀疑,2—错误,8—无法判断,9—缺测

2.2.6 质量控制码的标注

质量控制标识主要是为资料的使用者提供有关资料的质量信息。不仅提供资料是否可用,或可信度有多高的信息,同时也需提供资料在此之前的质量控制中所采用的方法、技术及资料质量控制过程,以便正确使用观测资料。检查结果以质量控制码的方式表示,质量码直接写入每个数据记录之后,一个要素值对应一个质量控制码。根据质量控制码生成新的数据集,标记出那些质量有问题的资料,同时也提供被标记的那些资料。本文的标注意义为:0——正确、1——怀疑、2——错误、8——无法判断、9——缺测。

2.2.7 人工干预检查

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。有时因中小尺度现象或局地气候,某一观测要素产生较大突变,自动质量控制过程有可能把正确的观测资料误判为可疑或错误的;对似是而非的资料,自动观测质量控制也无法做出决策,这些都需要借助人工干预才能做出最后的判断。同时人工干预检查过程中,需要了解仪器性能、具有经验的人员来完成。例如,在本文的敦煌实验中,有时由于敦煌地区空气湿度过于干燥,而湿度计的脉冲电流造成记录为负,在某些情况下湿度可能是负数,对于该问题,可认定湿度为 0,而不是仪器出错。

综上所述,逻辑极值、僵值和人工干预检查对每一观测项目都有检验,时间一致性检查中无风向,相似一致性检查中无风向和地温。

3 对敦煌数据的检查与结果分析

利用上面介绍的质量控制方法,对敦煌气象塔 9 a 的塔站资料数据集进行了全面的质量检查。

图 3a 给出总数据、正常数据和非正常数据的数量情况,从图 3a 来看,塔站观测资料的质量还是比较好的,非正常数据的出现只是少数情况。正常数据占总数据的 91.2%,而非正常数据占总数据的 8.8%。这里的非正常数据指的是缺测和错值数据,实际上缺测的数量占非正常值的大部分,以下将详细分析各类数据的分布情况。

非正常数据占总数据的 8.8%,在非正常数据分布中(图 3b),缺测数据占 78.7%,特别是在建站初期出现很多,由于仪器或软件的问题,会出现几天甚至 1 个月都没数据记录的情况,所以导致缺测的数据个数较多。错值占非正常数据的 21.3%,图 5 会对其分布进行详细讨论。而程序辨别出的无法判断数据和怀疑数据都经人工检查,以确定其质量码。

错误数据占总数据的 1.93%,在错误数据分布中(图 3c),大部分是因为出现了僵值数据而导致的错值,这种情况占错误数据中的 87.02%,风速和湿度资料出现僵值的次数较大。一致性错误占 6.18%,该错误的大部分出现在 2003 年 3 月和 4 月,18 m 湿度数据一直很小,没有通过相似一致性检查。超出逻辑极值的数据占 6.8%,多数为湿度资料,因为在戈壁地区的湿度很少达到饱和,偶尔会出现传感器有凝结水的时候,导致记录出现饱和的状

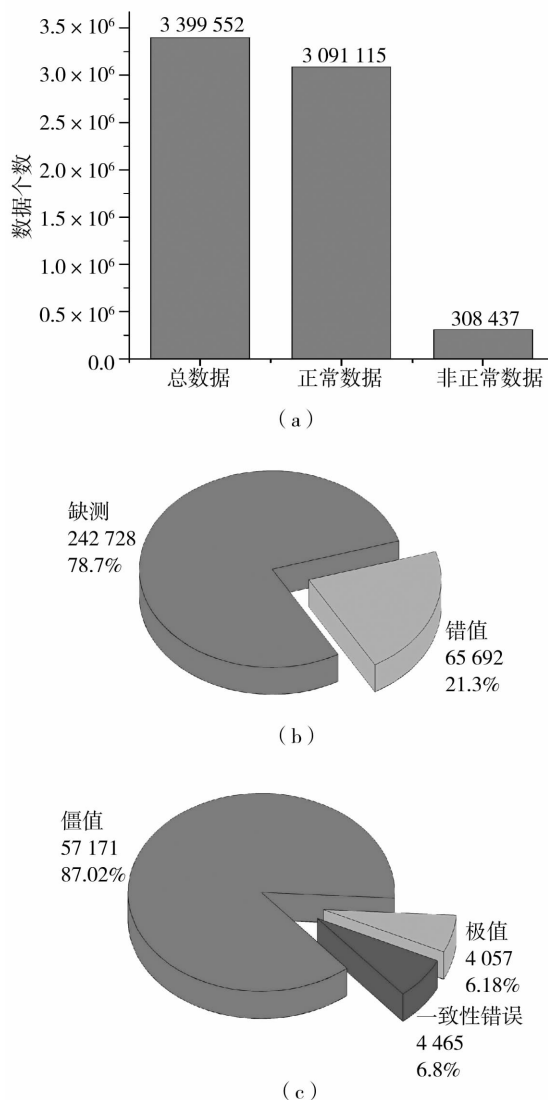


图 3 总的分布 (a)、非正常数据分布 (b)、和错误数据分布 (c)

Fig. 3 General data distribution (a), Non-normal data distribution (b), and Error data distribution (c)

况,所以规定 >95% 的湿度资料为错值。

以下对错值的 3 种情况分别给出典型的例子。

(1) 未通过逻辑极值检查的例子

图 4 给出不符合逻辑极值的数据,地表温度记录在 2002 年 8 月 22 日 8:05、13:35 和 18:05 毫无根据的降为 0 °C。如果按一般的地表温度的极值判定,地表温度为 0 °C 也是很正常的,但这是 8 月份夏季数据,显然 0 °C 是不可能出现的。我们根据按季节的极值判定可以轻易的辨别出来,这也说明逻辑极值检查的优越性。在野外观测中,由于环境恶劣,仪器受外界影响很大,通常仪器温度过高或过低导

致仪器的暂时性失灵,而产生错误记录,在后面的实验数据集整理中,应给与标注。

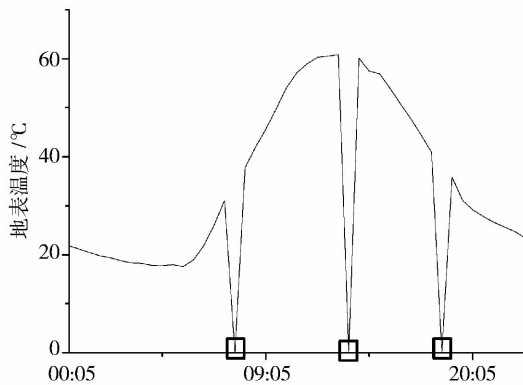


图4 典型的极值数据
Fig. 4 Typical extreme errors

(2) 未通过僵值检查的例子

图5给出僵值数据,1 m的风速记录从2001年6月13日18:05开始到6月22日10:35一直保持为同一个值,如此长时间的观察值不变,可认定为其僵值。由于观测仪器的失灵可能导致各要素观测值长时间不变,从而产生错误的记录。在野外观测中,多为无人值守,所以当仪器出现问题时,往往不能及时的发现和维护,而是在随后的资料分析中才发现。传感器或数采器出现故障都可导致僵值数据的产生,我们可以根据数据进行简单的判定:如果多种要素记录都为僵值,则很可能是数采器出问题;如果只有一种要素出现僵值(如上图),只有1m风速出现僵值,则很可能是风速传感器出现故障。

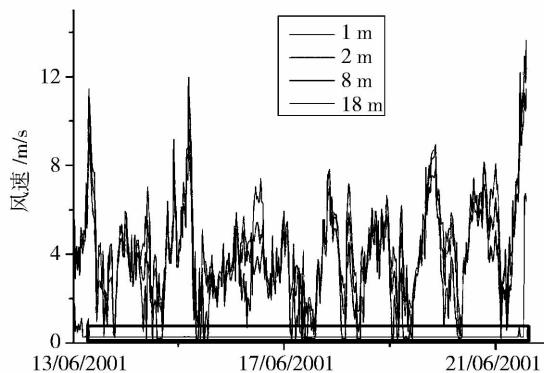


图5 典型的僵值数据
Fig. 5 Typical static errors

(3) 未通过一致性检查的例子

图6为2003年12月1~19日的4层湿度资料

料,很明显地看出18 m的湿度资料明显偏小,而且与其它3层的湿度变化趋势也大不相同,所以根据一致性原则认定该段时间的18 m湿度资料为错值。因为铁塔观测的4层高度相差不大,所以4层的要素变化幅度和趋势应该大致相同,当某一层出现了与其它3层差异很大或者不同的变化规律时,可认定该层的记录有误。

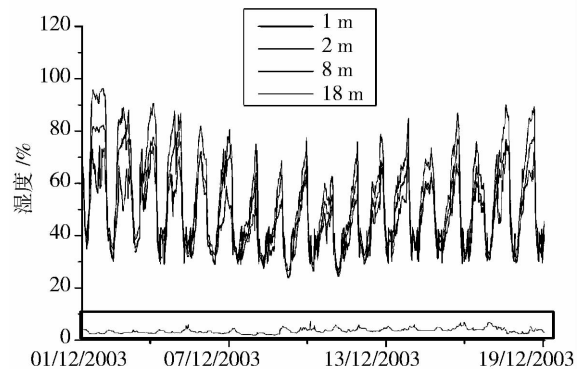


图6 典型不符合一致性数据
Fig. 6 Typical consistency errors

4 结论与讨论

(1) 讨论了气象观测资料的质量问题,并分析和强调了塔站资料质量控制的重要性,结合地面观测资料的质量控制方法,提出了逻辑极值检查、僵值检查、时间一致性检查、相似一致性检查、决策算法、质量控制码的标注和人工干预检查等组成的一套针对塔站资料的质控方法。

(2) 利用提出的质量控制流程,对敦煌戈壁塔站的资料进行质量控制。该流程能很好地找出缺失、错误、可疑的数据并给与标注。

(3) 对敦煌戈壁近10 a塔站资料的质量控制,塔站观测资料的质量还是比较好的,非正常数据只是少数情况。正常数据占总数据的91.2%,非正常数据占总数据的在8.8%,其中大部分为缺测数据,占其78.7%;无法判别数据和怀疑数据都通过人工检查,归为正常数据或错误数据。错误数据占总数据的1.93%,其中大部分为僵值数据,占其87.02%,其次分别为一致性错误占6.8%,超出逻辑极值的数据占6.18%。

(4) 本文阐述的塔站资料质量控制的基本思路和做法,在使用的时候,应着重掌握其精神,在不同的塔站使用时,应根据当地的气候背景、观测要素和

采样频率等相对应的修改参数,以适合当地的塔站使用。

(5)通过对塔站数据的质量检查,不仅可以保证观测资料的正确性,还可以检测出塔站的运行状态,以便及时对其维护,这对无人值守的气象塔站尤为重要。

致谢:感谢卫国安、侯旭宏等在敦煌试验中付出的艰苦劳动。

参考文献:

- [1] WMO. Manual on the global data - processing and forecasting system[R]. WMO - No. 485, 1992.
- [2] 俞卫平,周诗健,王伟. 地面气象观测规范[M]. 北京:气象出版社,2003.22.
- [3] WMO. Guidelines on quality control procedures for data from automatic weather stations. [R]. WMO - No. 305, 2004.
- [4] Foken T, Wichura B. Tools for quality assessment of surface - based flux measurements[J]. Agricultural and Forest Meteorology, 1996, 78(1-2):83-105.
- [5] Gandin L. Complex quality control of meteorological observations [J]. Monthly Weather Review, 1988,116(5):1137-1156.
- [6] Rissanen P, Jacobsson C, Madsen H. Nordic methods for quality control of climate data[R]. KLIMA Report, 2000,10.
- [7] Vejen F, Jacobsson C, Fredriksson U, et al. Quality Control of Meteorological Observations, Automatic Methods Used in the Nordic Countries[R]. KLIMA Report, 2002,8.
- [8] 方炳兴. 常规气象资料质量的综合控制[J]. 气象,1994,20(2):33-36.
- [9] 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.
- [10] 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.
- [11] 王伯民. 基本气象资料质量控制综合判别法的研究[J]. 应用气象学报,2004,15(B12):50-59.
- [12] 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.
- [13] 胡非. 湍流,间歇性与大气边界层[M]. 北京:科学出版社,1995.14.
- [14] 陈明,洪钟祥. 大气边界层观测资料的质量控制方案研究及其应用[J]. 大气科学,1993,17(1):97-105.
- [15] 陶士伟,徐枝芳. 加密自动站资料质量保障体系分析[J]. 气象,2007,33(2):34-41.
- [16] 张强,黄荣辉,王胜,等. 西北干旱区陆—气相互作用试验(NWC-ALIEX)及其研究进展[J]. 地球科学进展,2005,20(4):427-441.
- [17] 韦志刚,黄荣辉,陈文. 敦煌戈壁不同近地层大气稳定度下的感热交换系数[J]. 高原气象,2006,25(5):834-839.
- [18] 王胜,张强,卫国安. 敦煌干旱区一次降水过程陆面特征模拟[J]. 干旱气象,2004,22(4):46-50.
- [19] 李红寿,汪万福,张国彬,等. 影响莫高窟小气候的环境因子对比分析[J]. 干旱气象,2008,26(2):25-29.
- [20] 张强,卫国安. 荒漠戈壁大气总体曳力系数和输送系数观测研究[J]. 高原气象,2004,23(3):305-312.
- [21] 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.
- [22] 陈柏堃,郜庆林. 加密自动站市级中心站实时数据质量控制分析[J]. 安徽农业科学,2008,36(29):12994-12996.
- [23] 任芝花,许松,孙化南,等. 全球地面天气报历史资料质量检查与分析[J]. 应用气象学报,2006,17(4):412-420.

A Quality Control Routine For Dunhuang Gobi Meteorology Tower Data

WANG Chao, WEI Zhigang, LI Zhenchao

(Cold and Arid Regions Environmental and Engineering Research Institute,
Chinese Academy of Sciences, Lanzhou 73000, China)

Abstract: This paper discussed the problem of quality control of meteorological observation data and emphasized its importance. Integrated with the ground quality control, a series of quality control methods were advanced to deal with observation data, such as logical extremum check, static check, time consistency check, similarity check, synthetic method, quality code and manual interference check. Especially in the treatment process, using the logical extremum check and synthetic method can make the test results more reliable. And the quality control flow method was applied on data from Dunhuang meteorology tower from April 2001 to April 2009. The results demonstrated that the method could differentiate the missing, error, dubious data and mark it. The tower data was reliable with normal data accounting for 91.2% of the total data. Most of the abnormal data were missing observations accounting for 76.8%; error data accounted for 1.93% of the total data. Most of error data are static data, accounting for 87.02%, the consistency error and logical extreme errors coming next.

Key words: Dunhuang; meteorology tower; quality control; quality assurance procedures