

# 中国历史探空资料中的一些过失误差及偏差问题\*

翟 盘 茂

(国家气候中心, 北京, 100081)

## 摘 要

利用综合静力学质量控制技术, 分析中国历史探空资料集中的一些过失误差情况; 着重研究了 1960 年代探空仪器变化给温度和湿度时间序列带来的不均一性问题; 进一步讨论了中国近期探空温度和湿度资料中垂直层次上存在的一些偏差问题。

关键词: 探空资料, 误差, 不均一。

## 1 引 言

气候变化诊断与模拟研究依赖于可靠的长期历史气候资料。较高质量的全球地表面气象资料已经比较清楚地揭示了近百年来气候变化特点<sup>[1]</sup>。Angell 利用 1958–1987 年 63 个站的探空资料分析了近 30a 全球对流层与平流层大气的温度变化的特点<sup>[2]</sup>。随着资料的逐步积累与气候研究的不断深入, 中国一些学者也开始关注研究中国高空气候变化规律<sup>[3]</sup>。中国探空观测大多始于二十世纪五十年代初, 其中以南京和北京两站为最早, 测风观测则更早。在过去四十多年的观测中, 与其它国家和地区一样, 中国探空观测也经历了观测仪器、观测时间和观测规范的变化, 随着气象现代化事业的发展, 也经历了资料记录方式和传输方式的变化。这些变化可能会直接影响观测资料质量, 严重时带入资料中的“噪音”会超过气候变化信号从而影响气候分析结果。鉴于观测资料质量对气候变化观测研究的重要影响, 探空资料的误差问题已经引起了国际上不少学者的重视<sup>[4–6]</sup>。

观测资料中通常包含有随机误差、过失误差及系统误差。随机误差是观测资料中普遍存在的固有特性, 在统计上其分布常被假设为正态分布, 其均值为零。因此, 在气候平均处理中, 随机误差的影响程度会基本消除。系统误差与随机误差的主要差别在于其均值不为零, 因此, 在气候平均处理时资料就会出现与实际状况明显的偏差, 在时间序列或空间分布上系统误差不一致时, 就会产生资料中的不均一, 影响气候变化的研究结果。过失误差是没有任何气象学意义的资料错误, 它通常由资料操作过程中的失误或故障引起, 这类错误虽然发生频率不高, 但也会影响实际气候分析结果。

本文将针对中国建站到 1990 年的探空信息化资料中的温度、高度和湿度要素, 在资料质量检查和时空不均一性检测的基础上研究中国探空资料中的一些过失误差以及系统偏差情况, 以引起中国气象工作者在气候研究中的重视。

\* 初稿时间: 1996 年 5 月 23 日; 修改稿时间: 1996 年 8 月 8 日。

## 2 关于过失误差的分析

研究结果表明,探空资料中包含有 5%–20% 的错误<sup>[7]</sup>。本文使用的质量控制方法为综合静力学质量控制技术(CHQC),该技术由 Collins 和 Gandin 发展而成<sup>[7]</sup>,又经过 CARDS(综合高空数据集)工作的改进<sup>[6]</sup>。CHQC 利用高度、温度以及湿度之间的静力学关系可以成功地检测 30 多种类型的资料错误并能对其中部分错误进行订正。其主要思路如下:

在等压面  $p_i$  和  $p_{i+1}$  之间对大气静力学方程积分可以得到

$$Z_{i+1} - Z_i = - \frac{R}{g} \frac{p_{i+1}}{p_i} - T_v d(\ln p) \quad (1)$$

式中  $Z$  为位势高度,  $R$  为气体常数,  $g$  为重力加速度,  $T_v$  为虚温。如果  $T_0 = 273.15\text{K}$ , 设

$$A_{i,i+1} = (RT_0/g) \ln(p_i/p_{i+1})$$

$$B_{i,i+1} = (R/2g) \ln(p_i/p_{i+1})$$

则式(1)可表达为

$$Z_{i+1} - Z_i = A_{i,i+1} + B_{i,i+1}[(T_i + T_{i+1}) + 2t_{i,i+1}] \quad (2)$$

其中  $T_i$  为第  $i$  层的温度,  $t_{i,i+1}$  为从式(1)到式(2)转换过程中忽略湿度等处理引入的误差。因此,根据位势高度计算所得静力学余差为:

$$s_{i,i+1} = Z_{i+1} - Z_i - A_{i,i+1} - B_{i,i+1}(T_i + T_{i+1}) \quad (3)$$

根据温度计算的静力学余差为

$$X_{i,i+1} = (Z_{i+1} - Z_i)D_{i,i+1} - (T_i + T_{i+1}) - 2T_0 \quad (4)$$

其中  $D_{i,i+1} = 1/B_{i,i+1}$ 。如果温度的过失误差为  $T$ , 位势高度的过失误差为  $Z$ , 则余差为

$$s_{i,i+1} = Z_{i+1} - Z_i - A_{i,i+1} - B_{i,i+1}(T_i + T_{i+1}) \quad (5)$$

或

$$X_{i,i+1} = (Z_{i+1} - Z_i)D_{i,i+1} - (T_i + T_{i+1} - 2t_{i,i+1}) \quad (6)$$

根据各层高度和温度过失误差的出现情况,对式(5)和式(6)进行分析即可得出各类错误的存在条件与订正方法<sup>[7-8]</sup>。

例如在第 2 层高度和第 3 层温度出错时,  $Z_2$  和  $T_3$  不等于 0。式(6)可写成:

$$\begin{cases} Z_2 D_{12} = X_{12} - 2t_{12} \\ - Z_2 D_{23} - T_3 = X_{23} - 2t_{23} \\ - T_3 = X_{34} - 2t_{34} \end{cases} \quad (7)$$

假设  $t_{12}$ ,  $t_{23}$ , 和  $t_{34}$  相互独立,把式(7)各方程结合消去  $Z_2$  和  $T_3$ , 两边取平方和平均处理得到:

$$\begin{aligned}
 & - [(B_{12}/B_{23})X_{12} - X_{23} + X_{34}]^2 \\
 & = 4[(B_{12}/B_{23})^2(\overline{t_{12}})^2 + (\overline{t_{23}})^2 + (\overline{t_{34}})^2] \quad (8)
 \end{aligned}$$

观测结果表明,

$$[(\overline{t_{i,i+1}})^2]^{1/2} = \overline{t_{i,i+1}} \quad (9)$$

且几乎与气压层独立, 因此式(8)可写成

$$|(B_{12}/B_{23})X_{12} - X_{23} + X_{34}| = 2[2 + (B_{12}/B_{23})^2]^{1/2} \overline{t} \quad (10)$$

设  $\overline{t_{all}} = 3.5$  (适合于所有错误类型), 用  $\overline{t_{all}}$  代替  $\overline{t}$  时, 可以得到这类错误的存在条件为,

$$|(B_{12}/B_{23})X_{12} - X_{23} + X_{34}| < 2\overline{t_{all}}[2 + (B_{12}/B_{23})^2]^{1/2} \quad (11)$$

当式(7)中余差  $X_{12}$ ,  $X_{23}$  和  $X_{34}$  平方和达到最小时, 可以得到订正值  $Z_2$  和  $T_3$

$$\begin{cases} Z_2 = \frac{D_{23}X_{23} - 2D_{12}X_{12} - D_{23}X_{34}}{2D_{12}^2 + D_{23}^2} \\ T_3 = \frac{D_{12}^2X_{23} + D_{12}D_{23}X_{12} + (D_{12}^2 + D_{23}^2)X_{34}}{2D_{12}^2 + D_{23}^2} \end{cases} \quad (12)$$

CHQC 通过决策算法(DMA)实现对探空资料的质量控制, 其基本步骤是:

- (1) 对每条探空廓线从最底层开始按顺序取出四层进行检查;
- (2) 计算相邻层次之间三个静力学余差;
- (3) 确定错误类型, 并给出订正值(在有把握订正情况下对错误资料直接订正, 否则给出建议订正值);

(4) 对一条探空廓曲线订正完毕以后, 再次进行上述步骤, 以确保质量控制结果。

CHQC 对中国不同信息化方式的探空资料质量检查的结果见表 1, 表 1 中可见在中国电码资料中错误率最高, 达 8.1%; 而卡片及键盘输入资料错误率较低。中国卡片及键盘输入的资料都是按照气象台站上报的高空观测报表资料信息化而成, 其错误明显偏少的原因主要是在报表制作过程中人工审核及信息化过程中认真检查的结果。而电码资料为通过通讯线路取得的实时资料, 虽然其资料搜集速度快, 匆忙之中也带入了较多的错误。这些错误与资料收集、传输及处理过程是有关的。据分析, 这几种资料中的主要错误类型为等压面层上的温度或高度错误, 并有 1/3- 1/2 的错误可以进行可靠的订正。

表 1 CHQC 对中国探空资料的质量检查结果

资料记录方式	卡片资料	键盘输入资料	电码资料
期限	1957- 1963	建站- 1957	1980- 1990
	1971- 1979	1964- 1970	
错误占探空报的%	1.6	1.5	8.1
错误中可订正的%	36.8	29.0	53.1

对比中国电码资料与其它国家的电码探空资料中的错误出现情况(见表 2), 可以发现, 资料错误率在美国国家气象中心(NMC)的 1971- 1972 年资料集与澳大利亚 1990-

1993 资料集中最高; 中国 1980– 1990 年资料出错率与美国 NMC 的 1988 年资料以及前苏联的 1989– 1990 资料相当; 美国国家气候资料中心(NCDC) 资料中出错率最低, 这可能是在此之前进行过某些其它的质量控制步骤的结果。总的说来, 中国历史探空资料中的错误率在国际上处于中等程度。

表 2 利用 CHQC 对不同国家的探空资料的过失错误的检查结果

资料来源	中国	美国 NMC	美国 NMC	美国 NCDC	澳大利亚	前苏联
	1980– 1990	1971– 1972	1988	所有资料	1990– 1993	1989– 1990
错误占探空报的%	9.1	18.0	10.1	1.1	16.8	8.8

### 3 中国探空资料中的系统偏差

#### 3.1 时间序列中的不均一性

Gaffen 和 Elloit<sup>[4]</sup> 与 Paker 和 Cox<sup>[5]</sup> 指出, 在长期的观测过程中, 观测仪器、观测时间、辐射订正方法以及资料演算方法的变化都可能给探空资料序列引入偏差, 使资料时间序列产生不均一性。不均一性研究的难点之一在于如何选取参照资料。许多学者建议利用相邻测站资料进行比较以显露序列中的不均一性, 这种方法对于台站迁移造成的不均一性显示比较有效。但对于同一国家或地区来说, 仪器及观测规范的变化常常是同时或几乎同时发生的, 因而不容易通过邻站序列比较来检测其不均一性。本文一方面试图利用香港与广州的探空资料序列比较, 来显示一些不均一性, 另一方面着重采取中国测站 08 时和 20 时的探空温度和湿度资料序列的比较来揭示中国探空资料中的不均一性。为了比较客观地确定序列中的偏差值出现位置及不均一性程度, 采用先进的 Easterling–Peterson 技术<sup>[9]</sup> 配合分析。该方法原为地面资料不均一性检测与订正而设计, 其最大优点是可以自动检测序列中的不均一性, 但不适用于在短期之内频繁发生的不均一性检测。Zhai 和 Eskridge 对此方法进行了改进并成功地引入了 CARDS 的时间序列不均一检测过程中<sup>[10]</sup>。其主要技术步骤是:

- (1) 首先计算被研究序列  $X_i$  与参照序列  $Y_i$  的差值, 得到差值序列  $Z_i$ ;
- (2) 对整个差值序列进行线性回归, 并计算其剩余方差和  $RSS_1$ ;
- (3) 对整个差值序列在某检查点处进行二相线性回归, 并计算其剩余方差和  $RSS_2$ ;
- (4) 利用统计量

$$U = [(RSS_1 - RSS_2) / 3] / [RSS_2 / (n - 4)] \quad (13)$$

对二相回归的显著性进行检验, 同时利用 t 检验对检查点前后的均值差异显著性进行检验, 确定可能的不连续点;

(5) 在可能的不连续点处把序列断开, 重新进行步骤(2)–(4), ……, 直到序列不足 10 点无法再分为止;

(6) 在找出所有可能不连续点后, 用无参数的 MRPP(Multi-Response Permutation Procedures) 技术<sup>[11]</sup> 进一步确定不连续点;

(7) 以序列的最新资料为标准, 把断点前后的均值差作为偏差, 对断点前序列进行订正;

- (8) 对经过调整的序列, 再次进行不均一性检测与订正, 直到偏差序列完全均一为止;  
 (9) 利用订正后的均一的偏差序列与参照序列对被研究的不均一序列进行订正。

考虑 1950 年代期间, 中国探空观测仪器和规范变化较大, 1957 年 4 月中国探空观测时间从北京时间 11 时和 23 时变为 08 时和 20 时, 1959 年前后又经历了辐射订正的变化<sup>[12]</sup>, 以下只分析 1959 年以来的序列。对照图 1 中 08 时和 20 时两条年平均温度序列, 非常清楚地显示出在 1965 年前后 100hPa 高度上 08 时和 20 时的两条序列间出现明显的不一致性, 如果以 20 点温度序列为标准, 利用 Easterling-Peterson 不均一性检测方法得出在广州和北京 08 点温度序列在 1965 年前后年平均偏差值分别为 1.1 和 1.3。根据这两站台站历史记载, 发现广州站在 1965 年 8 月、北京站在 1964 年 6 月经历了 P3-049(国际上称为 RZ-049) 型仪器向 GZZ-2 型探空仪器的更换, 但中国从 1966 年 1 月 1 日起统一开始对新仪器进行辐射订正, 辐射订正是以日夜差值为基础的, 其开始时间与订正量级同温度序列中的偏差相当一致。因此广州和北京 08 点(白天)温度序列在 1965 前后的偏差与仪器更换后的辐射订正有关。

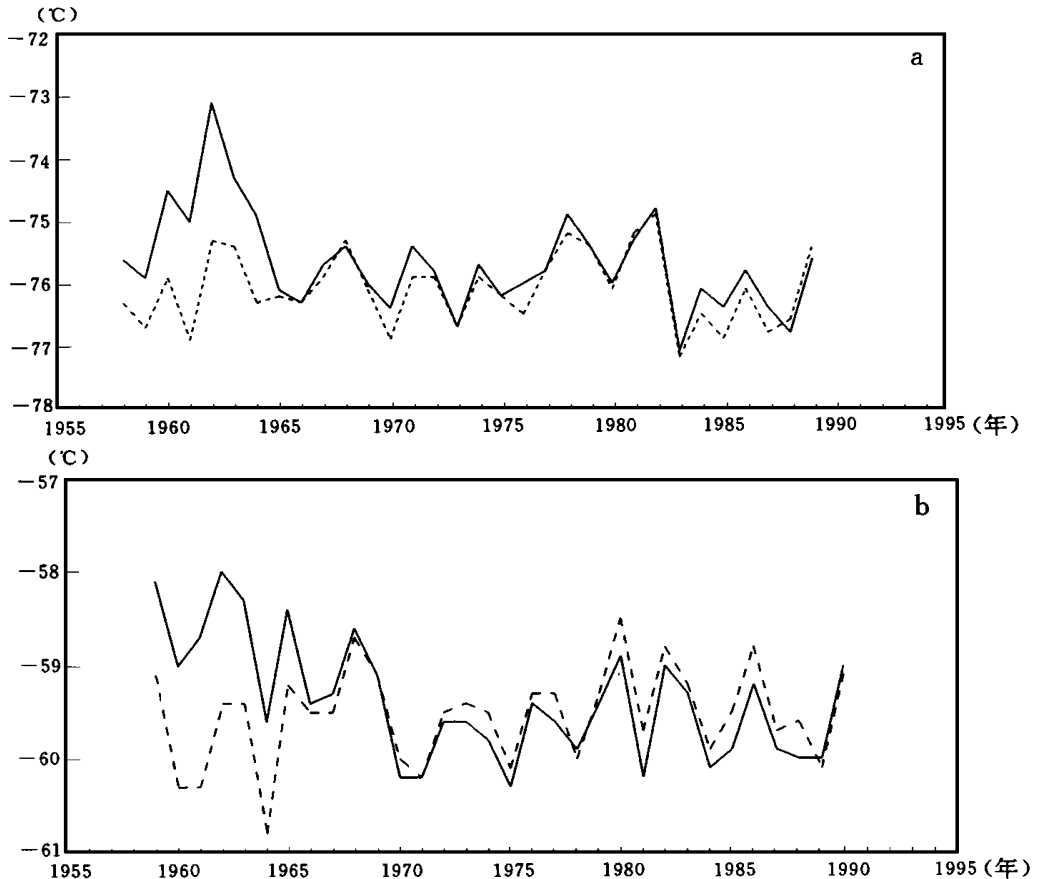


图 1 100hPa 08 时(实线)与 20 时(虚线)年平均温度序列比较

(a. 广州; b. 北京)

不均一性检测结果还表明,探空仪器的变化也影响了中国湿度序列的均一性。图 2 反映出在广州 1965 年前后以及武汉 1966 年前后 08 时和 20 点序列的温度露点差序列之间的差值发生明显的反相变化,这些变化也都与 P3-049 型仪器向 GZZ-2 型探空仪器的转换有关。对中国 1959 年以后的不同测站的探空资料的不均一性研究结果表明,1960 年代探空仪器变化造成的探空温度序列和湿度序列的不均一性现象是普遍存在的,其中温度的不均一性现象在对流层低层也同样有所显示。这种不均一性现象在气候变化研究中如不加以考虑或消除,将会带来错误的气候变化分析结果。

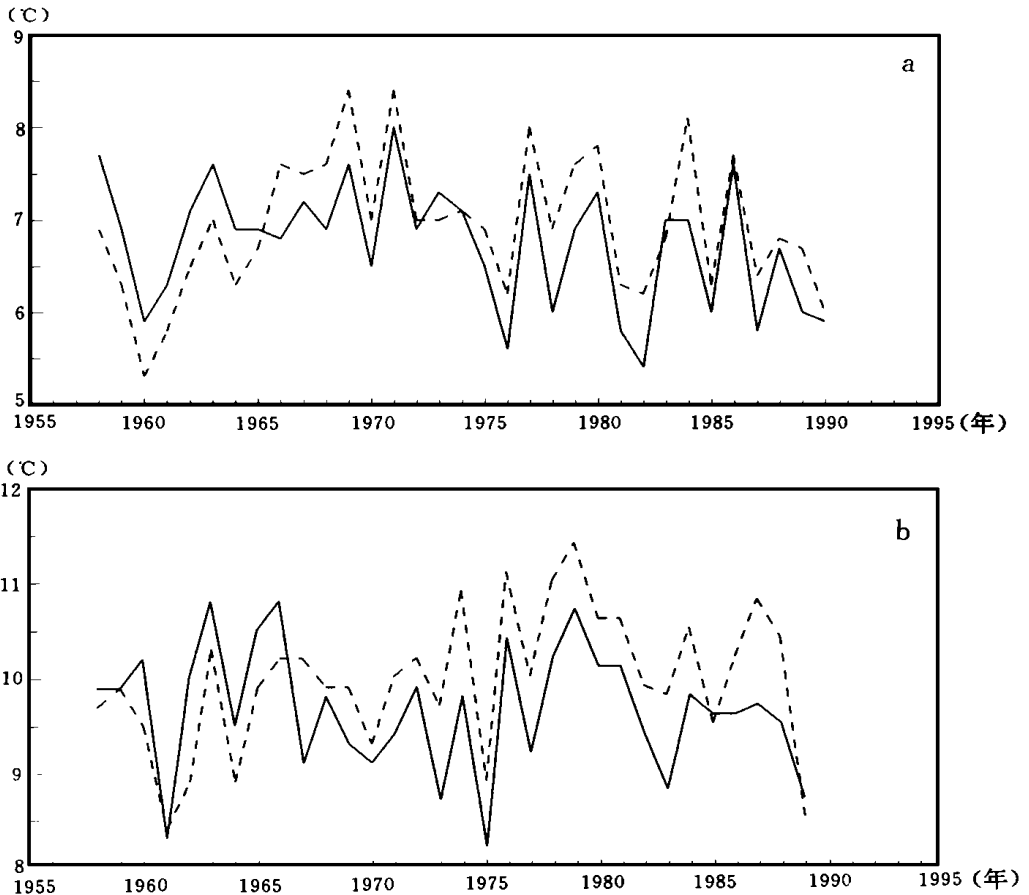


图 2 08 时(实线)与 20 时(虚线)年平均温度露点差序列比较

(a. 广州 700hPa; b. 武汉 500hPa)

详细的台站历史对于序列的不均一性研究具有十分重要的意义。图 3 中表示出 1968 年(A 处)香港探空仪从 Kew Mark IIB 型变成 Vaisala RS13 型; 1974 年(B 处)又变成 Vaisala RS18 型; 1981 年(C 处)又变成 Vaisala RS21 型; 1983 年(D 处)又变成 Vaisala RS80 型。对于广州来说, 1966 年以后没有经历重大的仪器变化, 因此序列相对均一。在两个地区的温度时间序列对照中不难发现, 与仪器变化对应的温度序列也产生了一定的偏差。

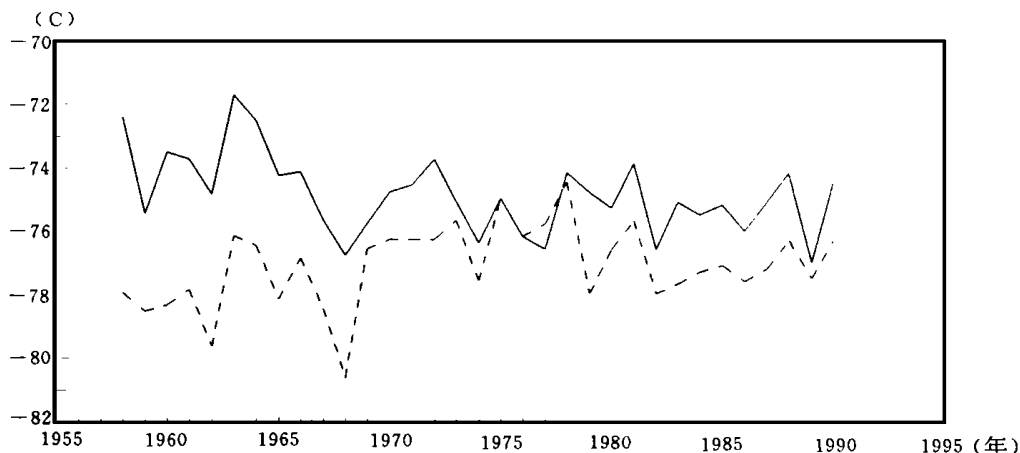


图3 广州(实线)和香港(虚线)7月份08时100hPa平均温度序列

### 3.2 探空资料空间上的一致性

为了显示近期探空资料在空间上的一致性,通过广州与香港两个地理上非常接近但使用仪器和规范完全不同的地区的资料对比来讨论温度与湿度观测之间的差异。以下采用的资料为1984–1990年,在这一时期内广州和香港两站均没有经历仪器变化。考虑到香港探空湿度资料表示为相对湿度而广州使用的是用相对湿度转换成温度露点差,本文按中国探空观测规定的相对湿度与露点和温度的关系<sup>[12]</sup>反算了相对湿度然后把香港与广州的资料进行比较。为了分析湿度观测在垂直层次上的一致性,还通过中国部分测站在晴天和云天条件下的探空湿度差异随高度的变化来讨论湿度观测中的一些潜在问题,其具体表示方法为:

$$PW = (q_{cd} - q_{cr}) / q_{cr} \quad (14)$$

这里  $PW$  为平均状况下与晴天条件下的大气水分差值比率,  $q_{cd}$  为平均状况下的大气比湿,  $q_{cr}$  为晴天条件下的大气比湿,比湿的计算方法为:

$$q = 0.622E / (P - 0.278E) \quad (15)$$

其中  $E$  为大气水汽压,  $P$  为气压。 $E$  可以根据相对湿度  $U$  和饱和水汽压  $E_s$  计算,

$$E = U \times E_s / 100 \quad (16)$$

中国信息化探空资料中湿度资料以温度露点差  $D$  表示,在已知摄氏温度  $t$  和温度露点差  $D$  时,依据相对湿度和温度露点差的转换公式,

$$\begin{cases} D = t - a/b \\ a = 237.3 \left( \frac{7.5t}{237.3 + t} + \log U - 2 \right) \\ b = 7.5 \left( \frac{7.5t}{237.3 + t} + \log U - 2 \right) \end{cases} \quad (17)$$

可以得到相对湿度  $U$ , 大气饱和水汽压  $E_s$  可以利用下列方法<sup>[13]</sup> 计算

$$E_s = 1.00071e^{0.000045p} E_w \quad (18)$$

$$E_w = 6.1094e^{\frac{17.625t}{243.04+t}} \quad (19)$$

当气温低于  $-40$  时,

$$E_s = 0.99882e^{0.000008p} E_i \quad (20)$$

$$E_i = 6.1121e^{\frac{22.587t}{273.86+t}} \quad (21)$$

比较广州和香港 7 月份 850hPa 到 700hPa 相对湿度资料(表 3), 可以发现, 在 500hPa 及其以下层次广州和香港观测的相对湿度结果相当接近, 但到了 500hPa 以上, 广州站的相对湿度观测值明显偏高, 两站之间的偏差值具有随高度增加而增加的趋势。

表 3 广州—香港 7 月份平均(1984—1990 年)相对湿度(%)比较

层次(hPa)	850	700	500	300	200
广州 08 时	81	67	55	45	44
香港 08 时	76	63	51	33	25
广州 20 时	77	66	56	48	47
香港 20 时	74	63	53	39	29

进一步研究中国一些测站有云和晴天情况下的大气水分差异(表 4), 可以看出在有云情况下大气水分明显偏多, 尤其在 700—300hPa 之间其偏高比率达到最高值。这显然与云层中大气水分明显偏多相关联。另一方面, 在云顶以上高度大气水分含量仍明显偏多, 反映了中国以肠衣为元件的湿度探测仪器在进入高水分含量气层后无法准确测量较干燥的气层大气湿度这一事实。这表明中国对流层高层湿度资料的精度不是十分可靠的。

表 4 夏季(1972—1990 年 6—8 月平均)有云条件与晴天条件大气比湿比较(%)

层次(hPa)	地面	850	700	500	400	300	200
兰州	37	26	37	56	<u>61</u>	60	26
济南	43	49	55	61	64	<u>68</u>	42
昆明	36	-	38	64	<u>64</u>	53	55
长沙	18	33	48	47	47	<u>49</u>	36
南京	31	46	57	61	<u>61</u>	60	42

为了比较广州探空温度与香港所用 Vaisala RS80 型探空仪观测温度的差异, 表 5 给出利用 1984—1990 年 7 月份 08 时和 20 时大气各层平均温度差值情况。表 5 中显示出在 850hPa 与 700hPa 上温度差值比较一致; 但由于广州对 300hPa 以下探空温度没有进行辐射订正, 在 500hPa 广州 08 点观测结果偏高, 使得广州 08 点和 20 点的温度差值比香港偏高 0.4; 300hPa 高度上两个测站 08 点和 20 点的温度差异又相一致; 之后随着探空高度的增加, 两测站的探空温度之间的差异又进一步增大。这显然与采用的探空仪器及相应的辐射订正方法等有关。国外一些学者正在对不同类型的探空仪的温度测量的系统偏差研制



物理学模型进行订正<sup>[14]</sup>来解决这一问题。

表 5 广州和香港 7 月份(1984-1990 年平均)探空温度比较

层次(hPa)	850	700	500	300	200	100
广州 08-20 点	-1.0	-0.4	0.0	-0.5	-0.2	0.8
香港 08-20 点	-0.9	-0.4	-0.4	-0.6	-0.7	0.2
08 点广州-香港	0.1	0.2	0.6	0.7	1.6	1.7
20 点广州-香港	0.2	0.2	0.2	0.6	1.1	1.0

#### 4 总结与讨论

经过综合静力学质量控制(CHQC)技术检查与不均一性检测技术对中国历史探空资料分析,得到了以下几点结论:

(1) 中国历史探空资料中错误率在世界上属于中等程度,在中国资料中以气象报表信息化的资料错误最少。

(2) 在中国 1960 年代探空序列中存在着明显的不均一性,这种不均一性是探空仪的变换与辐射订正相联系的。

(3) 中国近期探空资料尤其是探空湿度资料,在对流层高层存在着过分偏高的情况,这可能与探空仪肠衣湿度元件经过高湿层次后的滞后影响有关。

致谢: 孙安健同志给本文提出了宝贵意见、王树廷同志帮助查阅有关观测规范,作者在此深表感谢。

#### 参考文献

- [1] Paker D E, Jones P D, Folland C K and Bevan A. Interdecadal changes of surface temperature in the nineteenth century. *J Geophys Res*, 1994, 99: 14373-14739.
- [2] Angell J K. Variations and trends in tropospheric and stratospheric global temperatures, 1958-87. *J Climate*, 1988, 1: 1296-1313.
- [3] 王绍武等. 北半球及中国上空自由大气温度的变化. 气候变化及其数值模拟研究论文(第三集). 北京: 气象出版社, 1996. 12-16.
- [4] Gaffen D J. Temporal inhomogeneities in radiosonde temperature records. *J Geophys Res*, 1994, 99: 3667-3676.
- [5] Parker D E and Cox D I. Towards a consistent global climatological radiosonde data-base. *International Journal of Climatology*, 1995, 15: 473-496.
- [6] Eskridge R E, et al. A comprehensive arological reference data set (CARDS). *Bull Amer Meteor Soc*, 1995, 76: 1759-1775.
- [7] Collins W G and Gandin L S. Comprehensive hydrostatic quality control at the National Meteorological Center. *Mon Wea Rew*, 1990, 118: 2752-2767.
- [8] 翟盘茂, 周琴芳. 高空资料质量控制的 CHQC 方法. *气象科技*, 1993, 3: 92-97.
- [9] Easterling D R and Peterson T C. A new method for detecting and adjusting for undocumented discontinuities in climatological time series. *Int J Climatol*, 1994, 15: 369-377.
- [10] Zhai P-M and Eskridge R E. Analyses of inhomogeneities in radiosonde temperature and humidity time series. *J Climate*, 1996, 9(6): 884-894.
- [11] Mielke P W. The application of multicariate permutation methods based on distance functions in the earth sciences. *Earth-Sci Rev*, 1991, 31: 51-71.

- [12] Gaffen D J. Historical Change in Radiosonde Instruments and Practices, Final report, WMO Instruments and Observing Methods Reports No. 50, WMO/TD No. 541, World Meteorological Organization, Geneva, 1993. 123 pp.
- [13] Alduchov O A and Eskridge R E. Improved Magnus' form approximation of saturation vapor pressure. *J Appl Meteor*, 1996, 35: 601– 609.
- [14] Luers J K and Eskridge R E. Temperature corrections for VIZ and Vaisala radiosondes. *J Appl Meteor*, 1995, 34: 1241– 1253.

## SOME GROSS ERRORS AND BIASES IN CHINA'S HISTORICAL RADIOSONDE DATA

Zhai Panmao

(*National Climate Center, Beijing, 100081*)

### Abstract

By employing comprehensive hydrostatic quality control technique, this paper firstly analyzed gross errors and biases in China's historical radiosonde data sets. The significant inhomogeneities related to instrument change during 1960's were also evaluated. Moreover, the temperature and humidity biases at the vertical pressure levels in the recent period were discussed.

**Key words:** Radiosonde data, errors, inhomogeneities.