

# 自动气象站的动态试验及其测量准确度

郭锡钦 曾书儿 王金钊

(中国气象科学研究院, 北京 100081)

## 提 要

计算了自动气象站动态对比试验结果的相关系数和差值误差区间, 从而论证了对比试验的可信度. 提出了计算自动气象站测量结果的标准差、合格率、准确度的方法. 并用隐含周期分析与自相关模型得出了自动气象站的某些测量结果优于人工观测结果的结论.

**关键词:** 自动站; 试验; 准确度.

## 1 引 言

自动气象站在投入业务使用之前, 必须在自然条件下进行动态对比试验. 这种试验应该与人工观测基本上在同一时间, 同样条件下进行. 如果两种观测结果有较高的相关性, 其差值又有较小的误差区间, 就可以证明对比试验的结果有较高的可信度.

由于这种试验是在既不知道被测对象的真值, 又是在被测对象不断变化的条件下进行的, 因而使得试验结果的分析产生很大困难. 本文采用多种数理统计方法, 得出了估算自动气象站测量结果的精密性、合格率、准确度的方法<sup>[1]</sup>. 尤其是隐含周期分析与自相关模型的应用, 使得对自动气象站测量结果的分析有了新的进展.

需要说明的是, 本文论述的是自动站的观测方法, 不是全部自动站的试验结果, 所举的观测数据, 只是用来说明其观测方法. 较为详细的试验结果请见作者的其它论述<sup>[2]</sup>.

## 2 对比试验的可信程度

### 2.1 自动站测量值与人工观测值的相关特性

设  $n$  为观测次数,  $r$  为相关系数; 并按下式进行  $t$  检验<sup>[3]</sup>:

$$t = r \sqrt{n-2} / \sqrt{1-r^2} \quad (1)$$

利用对比试验的数据,按月和旬所做的  $t$  检验结果见表 1、表 2。旬相关系数及其检验因数数据太多,只抽样列入表内。

表 1 两种测量值月相关系数及其检验

杭州站	1991年7月	8月	9月	10月	11月	12月	1992年1月	2月	
气压	$n$	180	248	228	242	229	233	237	232
	$r$	1.000	1.000	0.958	1.000	0.976	0.970	1.000	0.985
	$t$	$\infty$	$\infty$	50.23	$\infty$	67.52	60.65	$\infty$	86.55
气温	$n$	193	247	229	244	227	232	215	232
	$r$	0.997	0.997	0.998	0.997	0.997	0.998	0.998	0.998
	$t$	178.94	201.50	237.92	200.22	193.07	239.49	230.47	239.49
相对湿度	$n$	194	246	290	245	237	231	216	232
	$r$	0.987	0.990	0.989	0.992	0.993	0.982	0.982	0.988
	$t$	84.84	109.67	113.48	122.25	128.68	78.34	76.05	97.00
风速	$n$	196	247	290	244	239	246	247	232
	$r$	0.926	0.927	0.954	0.956	0.946	0.956	0.996	0.920
	$t$	34.12	38.64	53.99	50.76	44.95	50.97	174.48	35.59
新民站	1991年8月	9月	10月	11月	1992年1月				
气压	$n$	205	687	740	707	748			
	$r$	0.998	0.995	0.982	1.000	0.976			
	$t$	224.99	260.68	140.63	$\infty$	120.80			
气温	$n$	205	691	740	707	728			
	$r$	0.999	0.998	0.999	0.999	0.999			
	$t$	318.36	414.50	607.00	594.24	602.17			
相对湿度	$n$	205	690	740	707	728			
	$r$	0.995	0.994	0.995	0.978	0.948			
	$t$	141.91	236.58	270.65	123.97	80.25			
风速	$n$	205	690	740	707	728			
	$r$	0.902	0.869	0.896	0.934	0.922			
	$t$	29.76	45.73	54.82	69.42	64.16			

注:①相关系数出现 1.000,是四舍五入的原因。②新民站 1991 年 12 月因自动站整机故障而无数据。

表 2 两种测量值旬相关系数及其检验

杭州站	1991年8月			1991年10月			1991年12月			
	上	中	下	上	中	下	上	中	下	
气压	$n$	80	80	88	79	75	88	/	66	88
	$r$	0.993	1.000	0.981	0.936	0.993	1.000	/	1.000	1.000
	$t$	74.73	$\infty$	46.89	23.33	71.83	$\infty$	/	$\infty$	$\infty$
气温	$n$	80	80	87	78	78	88	79	66	87
	$r$	0.996	0.997	0.997	0.994	0.995	0.998	0.997	0.997	0.999
	$t$	98.45	113.76	118.76	79.22	86.86	146.44	113.03	103.05	206.00
相对湿度	$n$	80	79	87	79	78	88	79	65	87
	$r$	0.988	0.990	0.992	0.989	0.987	0.994	0.987	0.987	0.763
	$t$	56.50	61.57	72.47	58.68	53.54	84.27	53.89	48.74	10.92
风速	$n$	80	80	87	78	78	88	80	65	87
	$r$	0.933	0.893	0.956	0.935	0.923	0.951	0.951	0.955	0.961
	$t$	22.90	17.53	30.04	22.99	20.91	28.49	27.16	25.56	32.04

续表 2

新民站	1991年9月			1991年11月			1992年1月		
	上	中	下	上	中	下	上	中	下
气压	$n$ 216	239	232	238	234	235	240	226	262
	$r$ 0.983	0.989	0.877	1.000	1.000	0.994	1.000	0.981	0.989
	$t$ 78.32	102.95	27.68	$\infty$	$\infty$	138.71	$\infty$	75.68	107.82
气温	$n$ 215	236	240	238	234	235	240	226	262
	$r$ 0.997	0.997	0.998	0.999	0.997	0.999	0.999	0.999	0.998
	$t$ 188.00	197.04	243.57	343.60	196.20	341.40	345.05	334.76	254.58
相对湿度	$n$ 215	236	239	238	234	235	240	226	262
	$r$ 0.993	0.996	0.995	0.982	0.980	0.977	0.893	0.930	0.957
	$t$ 122.72	170.52	153.38	79.86	75.00	69.95	30.61	37.87	53.21
风速	$n$ 215	236	239	238	234	235	240	226	262
	$r$ 0.865	0.869	0.878	0.905	0.915	0.936	0.893	0.912	0.912
	$t$ 25.16	26.87	28.25	32.68	34.54	40.59	30.61	33.29	35.85

由上表可见,旬、月的两种测量结果都有较好的线性相关特性.

## 2.2 差值月平均值的误差区间估计

自动站的各要素在动态试验中都无法参考标准与之对比,因而无法得知各要素的真值,也就不能直接得出这些测量结果的准确度.但是,可以近似地认为常规观测值与自动站测量值是在同一地点、同一时间获得的.它们之间的一一对应的差值是相对稳定的.如果把常规仪器近似地认为是标准的话,两种观测值的差值就是一种重复测量的结果.在样本数较大时,这些差值近似正态分布.

差值平均值是差值  $X(t)$  的数学期望的点估计,在概率为 0.95 的情况下,我们去计算覆盖该数学期望的误差区间大小.误差区间愈小,算术平均值就愈接近真值,说明对比记录的质量愈高.在置信度为 95% 时,我们利用试验数据计算出了差值月平均值的误差区间<sup>[4]</sup>(表 3).

表 3 两测量值差值月平均值的误差区间

杭州站	1991年7月	8月	9月	10月	11月	12月	1992年1月	2月
气压(hPa)	$\pm 0.01$	$\pm 0.01$	$\pm 0.11$	$\pm 0.07$	$\pm 0.11$	$\pm 0.04$	$\pm 0.01$	$\pm 0.01$
气温( $^{\circ}\text{C}$ )	$\pm 0.04$	$\pm 0.03$	$\pm 0.03$	$\pm 0.04$	$\pm 0.04$	$\pm 0.03$	$\pm 0.03$	$\pm 0.04$
相对湿度(%)	$\pm 0.29$	$\pm 0.24$	$\pm 0.27$	$\pm 0.23$	$\pm 0.27$	$\pm 0.29$	$\pm 0.41$	$\pm 0.39$
风速(m/s)	$\pm 0.18$	$\pm 0.19$	$\pm 0.06$	$\pm 0.06$	$\pm 0.07$	$\pm 0.06$	$\pm 0.06$	$\pm 0.05$
新民站	1991年8月	9月	10月	11月	1992年2月			
气压(hPa)	$\pm 0.05$	$\pm 0.03$	$\pm 0.02$	$\pm 0.02$	$\pm 0.02$			
气温( $^{\circ}\text{C}$ )	$\pm 0.03$	$\pm 0.02$	$\pm 0.02$	$\pm 0.02$	$\pm 0.02$			
相对湿度(%)	$\pm 0.22$	$\pm 0.16$	$\pm 0.15$	$\pm 0.30$	$\pm 0.50$			
风速(m/s)	$\pm 0.09$	$\pm 0.05$	$\pm 0.04$	$\pm 0.04$	$\pm 0.03$			

由表 3 可见,误差区间很小,对比测量结果是可信的.表明我们进行的对比试验有较高的可信度,由此推出的结果有可靠的基础.

## 3 自动气象站测量值的标准差计算

### 3.1 常规仪器测量值的标准差估计

根据我们掌握的资料,常规仪器的检定标准差如下:气压 $\pm 0.1\text{hPa}$ ;气温 $\pm 0.1^{\circ}\text{C}$ ;相

对湿度 $\pm 2\%$ ;风速不大于 $\pm 0.3\text{m/s}$ 。

检定标准差实际上是系统误差的标准差. 因此,在常规仪器符合使用要求的情况下,它们的观测误差的标准差不会小于下列数值<sup>[5]</sup>:气压 $\pm 0.2\text{hPa}$ ;气温 $\pm 0.2^\circ\text{C}$ ;相对湿度 $\pm 4\%$ ;风速 $\pm 0.6\text{m/s}$ 。

### 3.2 自动站测量值的标准差估计

设 $\sigma_1$ 为常规仪器测量值的标准差, $\sigma_2$ 为自动站测量值的标准差,其它如观测位置、防护方法等差异引起的误差忽略不计,则对比差值的方差为:

$$D_x = \sigma_1^2 + \sigma_2^2 \quad (2)$$

而

$$D_x = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

其中 $n$ 为观测次数, $X_i$ 为第 $i$ 次观测差值, $\bar{X}$ 为差值平均值。

$$\text{则} \quad \sigma_2 = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 - \sigma_1^2} \quad (3)$$

表4给出用自动站测量资料计算的标准差。

表4 自动站测量值的月标准差( $\sigma_2$ )

杭州站	1991年7月	8月	9月	10月	11月	12月	1992年1月	2月
气压(hPa)	$\pm 0.09$	$\pm 0.12$	$\pm 0.86$	$\pm 0.54$	$\pm 0.83$	$\pm 0.32$	$\pm 0.06$	$\pm 0.08$
气温( $^\circ\text{C}$ )	$\pm 0.27$	$\pm 0.22$	$\pm 0.25$	$\pm 0.28$	$\pm 0.29$	$\pm 0.23$	$\pm 0.26$	$\pm 0.32$
相对湿度(%)	$\pm 1.80$	$\pm 1.86$	$\pm 1.92$	$\pm 1.74$	$\pm 2.02$	$\pm 2.01$	$\pm 2.95$	$\pm 2.88$
风速(m/s)	$\pm 0.30$	$\pm 0.29$	$\pm 0.30$	$\pm 0.29$	$\pm 0.30$	$\pm 0.28$	$\pm 0.28$	$\pm 0.30$
新民站	1991年8月	9月	10月	11月	1992年1月			
气压(hPa)	$\pm 0.36$	$\pm 0.34$	$\pm 0.30$	$\pm 0.34$	$\pm 0.22$			
气温( $^\circ\text{C}$ )	$\pm 0.22$	$\pm 0.28$	$\pm 0.30$	$\pm 0.21$	$\pm 0.23$			
相对湿度(%)	$\pm 1.60$	$\pm 1.98$	$\pm 1.76$	$\pm 3.92$	$\pm 5.76$			
风速(m/s)	$\pm 0.57$	$\pm 0.45$	$\pm 0.45$	$\pm 0.46$	$\pm 0.36$			

显然, $\sigma_1$ 的估计在一般情况下是保守的,其它的外界因素引起的误差又忽略不计,因此计算出的自动站的标准差实际上是偏大的. 即在一般情况下,自动站测量值的标准差会比表4中的数值小一些。

## 4 自动站雨量、风向记录分析

在自动站中,雨量测量的问题较多. 这是因为年降水量、月降水量都是累计值. 而在自动站中采用的翻斗雨量计,要完全避免干扰信号的影响还有技术上的困难. 只要有一次测量错误,就使月、年降水量产生错误. 此外,正当下雨时,人工观测过程中就有雨量损失,致使产生较大的对比误差. 以上两种误差,在降水量小时,其百分误差很大. 根据5台资料收集平台(DCP)一年的对比结果,年降水量的相对误差在20%—25%之间。

因此我们认为,根据目前自动站的实际情况,雨量的对比观测应以一次降水过程为起止点. 对于风向记录,台站现用的电接风向风速计测量原理与自动站中风向测量原理很不相同,两者的分辨率相差很大. 加之在自然条件下,风向在时间和空间分布上差别也很大,

因此,只需求出两者相符的百分率.

$$\text{相符百分率}(\%) = (\text{相符次数} / \text{总次数}) \times 100\% \quad (4)$$

电接风向风速计启动风速为 2m/s,因此在统计中不计 2m/s 以下的记录. 它的风向以 22.5°为 1 个方位. 而自动站中的风向的测量准确度为 ±10°. 当自动站所测风向对应一定的角度范围时,如 NNE 为 1.25°—43.75°,NE 为 23.75°—66.25°,以此类推,即可认为两者相符.

实验表明,风向对比相符率在不同地区相差甚大. 在高山站,由于地形复杂,相符率甚低. 如安装在泰山站的资料收集平台,12 个月的风向相符率平均只有 36%. 在新民、杭州,有线遥测站 1991 年共 12 个月的风向平均相符率为 67%. 而资料收集平台在内蒙朱日和一年的试验中,风向的平均相符率高达 96%. 通常认为,风向的相符率在 60%—70%之间就是满意的结果.

## 5 隐含周期分析与自相关模型的应用

### 5.1 自动站与人工观测资料对比

我们把自动站和人工观测的对比资料看作是两个平行的时间系列. 其中的缺测资料用常规方法进行插补.

现在以七角井站气温为例予以说明. 设在  $t$  次观测时,气温的真值为  $\eta_t$ ,人工观测值为  $y_t$ ,自动站观测值为  $z_t$ ,人工观测的误差为  $e_t$ ,自动站的观测误差为  $\tilde{e}_t$ . 则

$$\begin{aligned} y_t &= \eta_t + e_t \\ z_t &= \eta_t + \tilde{e}_t \end{aligned} \quad (t = 1, 2, \dots, n) \quad (5)$$

其中气温的真值部分由周期性变化量  $f(t)$  和趋势项(非周期性变化量) $\epsilon_t$  组成. 则

$$\eta_t = f(t) + \epsilon_t \quad (t = 1, 2, \dots, n) \quad (6)$$

由式(5)有

$$y_t - z_t = e_t - \tilde{e}_t$$

若记

$$\zeta_t = y_t - z_t$$

那么

$$E\zeta_t = Ee_t - E\tilde{e}_t$$

则方差为

$$V_{ar}\zeta_t = V_{ar}e_t + V_{ar}\tilde{e}_t = \sigma_1^2 + \sigma_2^2 \quad (7)$$

显然,  $\sigma_1^2$  是人工观测值的方差,  $\sigma_2^2$  是自动站测量值的方差. 这是常规方法的计算结果.

用隐含周期和自相关模型求出  $\sigma_1^2 - \sigma_2^2$ , 然后与式(7)联解,就可分别求出  $\sigma_1$  和  $\sigma_2$ . 由式(6),周期部分可根据傅立叶级数理论,则有

$$f(t) = \sum_{j=1}^k (\alpha_j \cos \lambda_j t + \beta_j \sin \lambda_j t) \quad (t = 1, 2, \dots, n) \quad (8)$$

其中,  $k$  是周期项的个数,  $\lambda_j$  为频率,  $\alpha_j, \beta_j$  是与  $\lambda_j$  相应振幅有关的量.

实际计算结果是:  $k = 4, \lambda_1 = 0.0, \lambda_2 = 0.0043035, \lambda_3 = \pi/2, \lambda_4 = \pi$ . 所以,

$$f(t) = \alpha_1 + \sum_{j=2}^3 (\alpha_j \cos \lambda_j t + \beta_j \sin \lambda_j t) + \alpha_4 \cos \lambda_4 t \quad (9)$$

由于三角函数以  $2\pi$  为周期,所以  $T_j$  为  $\lambda_j$  相应的周期,则有

$$\lambda_j(t + T_j) = \lambda_j t + 2\pi$$

所以,  $T_j = 2\pi/\lambda_j$ , 即  $T_2 = 1461.4$ ,  $T_3 = 4$ ,  $T_4 = 2$ .

由于每天有 4 次测量,若以日为单位,气温周期变化为 365.3, 1 和 0.5,这与日常经验相符.

根据经验,气温的趋势项(非周期部分)是自相关的. 故用一阶自回归模型来描述它.

$$\varepsilon_t = \rho \varepsilon_{t-1} + u_t \quad (t = 1, 2, \dots, n) \quad (10)$$

其中  $\rho$  为自相关系数,  $u_t$  是一个正态分布的随机变量. 经计算,  $\hat{\rho} = 0.813$ . 最后的计算结果:  $\hat{\sigma}_1 = 0.512$ ,  $\hat{\sigma}_2 = 0.377$ .

## 5.2 与 CIMO 数据的比较

该方案的各要素计算结果与 CIMO 指南中的数据比较<sup>[7]</sup>如表 5(均以七角井站的资料为例).

表 5 人工站、自动站诸要素测量标准差与 CIMO 相应数据的比较

要素	CIMO 数据	人工观测	自动站测量
气温( $^{\circ}\text{C}$ )	0.25—0.50	0.51	0.38
相对湿度(%) *	5.7—8.3	8.9	4.4
气压(hPa)	0.18	0.25	0.27
风速(m/s)	0.50—0.65	0.97	0.89

注: CIMO 在该项上列出的是湿球温度 0.35—0.60 $^{\circ}\text{C}$ . 在气温为 -10—+40 $^{\circ}\text{C}$  范围内,上述数值相当于相对湿度 5.7%—8.3%

CIMO 的数据表明,由于大气中高频干扰的存在,即使用目前最好的滤波方式,也无法完全避免高频干扰混迭到观测数据中去. 因此,不管仪器的测量准确度如何的好,都无法将测量误差减下去. 计算表明,对比资料也反映出了这种特征. 但是,与人工观测相比较,由于自动站的测量值采用了多次取样后的平均值. 因此,它的测量结果多少抑制了高频干扰的混迭,这是它比人工观测结果优越之处. 正如 CIMO 指南中所说,在大气中,气压波动的幅度很小,求平均值的意义不大,所以两种观测结果十分相近.

## 6 自动站的动态测量准确度

### 6.1 用二倍标准差作为准确度指标

在 95% 的置信度条件下,准确度可用  $(X - \bar{X}) \pm 1.96\sigma$  表示. 其中  $X$  表示测量值的真值,  $\bar{X}$  表示  $n$  次观测的平均值,  $\sigma$  为观测值的标准差. 显然,  $(X - \bar{X})$  为系统误差,它是可以修正的. 根据 CIMO 指南,为了简易起见,把  $2\sigma$  作为准确度指标.

用表 4 中的数据平均后,其测量准确度与 CIMO 的要求相比较的结果如表 6.

表6 自动站测量准确度与 CIMO 要求比较表

要素	标准差	二倍标准差	CIMO 要求	是否符合 CIMO 要求
气压(hPa)	±0.43	±0.86	±1.00	是
气温(°C)	±0.26	±0.52	±1.00	是
相对湿度(%)	±2.73	±5.46	±8(平均)	是
风速(m/s)	±0.37	±0.74	±2.00	是

## 6.2 求对比差值的合格率

由于常规仪器不是标准仪器,其置信度应按 95%考虑.而 CIMO 的要求系指二倍标准差,即 95%的测量值应落在此要求范围内.因此,对比测量的数据的合格率为  $95\% \times 95\% = 90\%$ 时,即认为符合要求.现以泰山 DCP1988 年对比数据为例,气压、气温、相对湿度的合格率分别为 98.3%、93.6%、90.0%,符合 CIMO 要求,只有风速的合格率为 66.1%,不符合 CIMO 要求.

气温在 0°C 以下时,人工观测相对湿度的准确度很低,全年合格率只有 83.1%.若只计算气温 0°C 以上时,则合格率为 90%.风速合格率低的原因是泰山气象站的风场在时间和空间上变化较大.

## 7 结 论

自动站的动态对比试验(除泰山站风的数据外)是在可靠性很高的情况下进行的,完全可以用来论证自动站的观测效果.通过对自动站的精密度、合格率、准确度的计算,可以认为我国的自动站的测量结果能基本满足 CIMO 对自动站的要求,为自动站投入业务使用提供了可靠的依据.目前存在的问题是,我国的自动站故障较多,上述的结论是在自动站正常运行时得出的.另外,雨量测量中的防干扰问题较为突出.高山地区风的测量也有观测方法的问题需要解决.关于自动站测量记录与人工观测记录的连续性问题,作者将另文论述.用隐含周期分析和自相关模型只对七角井站一年的资料进行了计算,这是因为人工观测记录的存盘,各站使用的机型不同而无法兼容,致使工作量太大.但初步计算结果表明,自动站的测量由于多次取样平均,在扼制大气中的高频干扰方面比人工观测更优越.

## 参 考 文 献

- 1 郭锡钦,曾书儿,王金钊.资料收集平台(DCP)现场对比试验数据分析、整理方法初探.气象科技,1990,(4):79—82.
- 2 郭锡钦,曾书儿,王金钊.引进的资料收集平台(DCP)的试验评估.气象,1990,16(1):22—27.
- 3 屠其璞,王俊德,丁裕国等.气象应用概率统计学.北京:气象出版社,1984.117—118,377—384.
- 4 WMO-No. 8 Accuracy of measurements. Geneva, 1983. 1.12—1.15.
- 5 WMO-No. 8 Groups of instruments. Geneva, 1983. 1.15—1.16.
- 6 陈善敏.南极自动气象站.气象科技,1985,(3):85—87.
- 7 WMO-No. 8 Techniques for Sampling Surface Variables. Geneva, 1983. 25.1—25.8.

## FIELD TESTS AND MEASUREMENT ACCURACY OF AUTOMATIC WEATHER STATIONS (AWS)

Guo Xiqin Zheng Shu'er Wang Jinzhao

*(Chinese Academy of Meteorological Sciences, Beijing 100081)*

### Abstract

Based on the computed correlation coefficients and the error of the differences between the tested values from AWS and the conventional observations, the reliability of the comparative tests were demonstrated. A method of calculating the standard deviation, qualified rate and accuracy of the AWS measurement results were suggested. The hidden periodic analysis and autocorrelation model were used to compute the AWS measurement data. The results show clearly that some measurement parameters from AWS are better than the man-made observations.

**Key words:** Automatic Weather Station(AWS); Test; Accuracy.