

五台山站历史气候资料的均一性分析

高晓容¹ 李庆祥² 董文杰³

(1 兰州大学大气科学学院, 兰州 730000; 2 国家气象信息中心, 北京 100081; 3 国家气候中心, 北京 100081)

摘要 以山西五台山地面观测站为例, 对建站(1957 年 1 月)至 2005 年的年、季节平均气温、最高和最低气温、相对湿度和降水观测资料序列作均一性分析, 结合历史沿革变化(包括迁站、仪器变化和观测业务改变等), 得出这些因素对不同气候要素均一性的影响情况。从与该站不同要素序列相关性较好、水平距离最近的台站中选取若干站点作为待检台站的参考台站建立了各自的参考序列作为均一性检查的对比序列, 采用国际上应用较为广泛的标准正态检验(SNHT)方法分别对各种要素进行了统计学检验。对统计方法得到的不连续点(不均一性产生年份)进行了总结分析, 发现台站迁移对于温度、相对湿度观测序列的均一性带来的影响非常明显, 其中对最高气温序列影响最为显著, 达到 5.5 °C。深入分析还表明, 迁址对各要素的影响程度各有区别; 对相对湿度的年均和春季序列也产生了较为明显的非均一性, 而对降水量的影响则并不显著。综合分析表明, 常用观测要素的均一性检验和订正是局地气候变化检测的基础环节, 但不同要素的均一性受到各种因素影响的表现程度各不相同, 需要加以区别和深入研究。

关键词 台站迁址 SNHT 方法 均一性检验

引言

气候资料是进行气候变化研究及气候预测的基础, 只有在对气候资料进行均一性分析和订正的基础上, 才能揭示真实的气候变化特征与规律。长期以来, 台站迁移、仪器变更、观测与资料处理过程中采用的方法变化等因素都会导致单站气候资料产生非均一现象, 而单站气候资料的不均一直接导致了局域气候资料的不均一性。近年来我国不少研究人员开始注重气候资料的均一性研究^[1~4], 分别对我国的探空资料、气温、降水以及风速资料等进行了一些均一性研究, 并对气温数据作了均一化订正^[5], 取得了一些重要进展(李庆祥, 张洪政, 刘小宁等的中国均一化历史气温数据集(1951~2004), 1.0 版, 国家气象信息中心气象资料室印)。同时崔宜少等^[6]也研究了威海站在历史上经过多次迁站、仪器变化及计算方法的改变后, 哪些因素是对气象资料产生非均一性的主要因子。吴利红等^[7]对浙江省 36 个代表站的年平均气温序列进行均一性检验, 分析表明 36 个站中有 39% 的测站表现出非均一性, 而台

站迁移、环境变化是产生非均一性的主要原因。虽然研究人员已经做了不少工作, 但是目前在我国, 对于单站气候资料均一性的研究工作开展得还很不充分, 对于单站地面、高空各气象要素历史观测序列系统性的均一性分析和订正还存在着明显的差距和不足。因此, 深入分析影响气候资料均一性的各种因子, 研究它们对不同气候数据的影响, 以及各要素气候资料的均一性检验、订正方法, 是当前气候学和气候变化研究面临的紧迫任务。

1 资料和元数据信息

本文采用的数据为国家气象信息中心气象资料室收集整理的山西省五台山气象观测站(53588)的地面逐月平均气温、最高和最低气温、相对湿度、风速、降水序列, 时间尺度是 1957 年 1 月至 2005 年 12 月, 共 49 年。该数据来源于台站上报的气象报表, 国家气象信息中心气象资料室对此进行了初步质量控制, 剔除了个别明显极端值。

五台山站建于 1955 年 10 月, 当时建于五台山中台顶, 纬度和经度分别是 39°02'N、113°32'E, 标注

海拔高度为 2720.4 m(在 1/50000 地图上查得), 1960 年 1 月 1 日国家测绘局用三角测量法测得海拔 2895.8 m, 1998 年 1 月 1 日迁至五台山木鱼山, 纬度和经度变化为 $38^{\circ}57'N$ 、 $113^{\circ}31'E$, 海拔高度变为 2208.3 m(测量)。这次迁址, 台站海拔高度下降了近 700 m, 水平距离达到 20 km。在近 50 年中, 除了站址迁移外, 其他如仪器、环境(均为山地)等均没有太大的变化。故本文主要考察这种台站迁移对各种气候要素序列均一性的影响。

2 对各气象要素的均一性检验

气象要素的均一性检验方法有很多种, 大都是基于最大似然方法。在诸多方法中, 标准正态检验(SNHT)方法是一个应用范围非常广泛的方法, 它对于幅度较小的不连续点的检测较为敏感, 故本文采取该方法进行均一性检验^[8,9]。

2.1 参考序列的建立及质量状况

气候序列均一性检验的基础和关键是建立待检序列的参考序列和检验序列, 通过待检序列和参考序列的对比来判断序列的均一性。参考序列首先必须是均一的, 才可以代表待检序列真实的气候序列变化, 但这种理想状况是不可能达到的。因此如何建立一个尽可能均一的参考序列是本文检验工作的基础。本文将五台山站的年、季各气象要素作为待检序列 Y_i ($i=1, 2, \dots, 49$), 从距离五台山站水平距离最近的 10 个相邻台站(表 1)中挑选 5 个与待检站的各要素序列相关性最好的台站, 作为待检台站的参考台站($j=1, 2, 3, 4, 5$), 表 2 列出了不同要素待检站与各参考台站观测序列的相关系数 ρ_j , 利用下文中式(1)与式(3)建立各要素年、季的参考序列 f_i 。

表 1 五台山站(53588)及其相邻台站位置

区站号	站名	经度(E)	纬度(N)	海拔/m
53588	五台山	$113^{\circ}31'$	$38^{\circ}57'$	2208.3
53673	原平	$112^{\circ}43'$	$38^{\circ}44'$	828.2
53782	阳泉	$113^{\circ}33'$	$37^{\circ}51'$	741.9
53487	大同	$113^{\circ}20'$	$40^{\circ}06'$	1067.2
53698	石家庄	$114^{\circ}25'$	$38^{\circ}02'$	80.5
53593	蔚县	$114^{\circ}34'$	$39^{\circ}50'$	909.5
53478	右玉	$112^{\circ}27'$	$40^{\circ}00'$	1345.8
53772	太原	$112^{\circ}33'$	$37^{\circ}47'$	778.3
53663	五寨	$111^{\circ}49'$	$38^{\circ}55'$	1401.0
53787	榆社	$112^{\circ}59'$	$37^{\circ}04'$	1041.4
54602	保定	$115^{\circ}31'$	$38^{\circ}51'$	17.2

$$f_i = \left[\sum_{j=1}^5 \rho_j^2 X_{ji} \right] / \sum_{j=1}^5 \rho_j^2 \quad (1)$$

其中 X_{ji} 为第 j 个参考台站的要素序列, 对于温度、湿度和风速的均一性检验, 我们采用差值法构造检验序列 P_i :

$$P_i = Y_i - f_i \quad (2)$$

而对于降水量这种累积量, 则一般用式(3)建立参考序列, 检验序列则用比值法。构造的比值序列 Q_i 表示为式(4):

$$f_i = \left[\sum_{j=1}^5 \rho_j^2 X_{ji} / \bar{X}_j \right] / \sum_{j=1}^5 \rho_j^2 \quad (3)$$

$$Q_i = \frac{Y_i / \bar{Y}}{f_i} \quad (4)$$

其中 \bar{X}_j 为第 j 个参考台站要素序列的平均值, \bar{Y} 为待检序列的平均值。表 2 也给出了构造参考序列 f_i 与待检序列 Y_i ($i=1, 2, \dots, 49$) 的相关系数 r' 。

本文首先分析区域内不同气象要素的空间代表性, 对 53 区($105^{\circ} \sim 115^{\circ}E$, $35^{\circ} \sim 45^{\circ}N$) 内每个台站之间的年平均温度、相对湿度、风速和降水量各要素的相关系数与台站距离的关系分别进行了分析, 发现气温要素(包括平均、最高和最低气温)、相对湿度以及降水量, 基本呈现出类似的变化特点: 相关系数随着水平距离的增大而减小, 并且距离较近的台站间的相关系数较大, 尤其是气温要素, 近距离台站间的相关性非常强, 降水和相对湿度略低; 对于风速则明显不同, 一方面随着水平距离增加, 其相关性没有明显减弱, 另一方面近距离台站间序列的相关性并不很强。这说明: 在上述区域内, 气温序列的代表性最强, 降水和相对湿度序列次之, 而风速序列的代表性较差。因此本文中五台山站的风速要素不适宜用该方法构造参考序列作为均一性检验的依据。下面直观地比较一下五台山站的平均气温、相对湿度和降水量的待检序列、参考序列均一状况。图 1 给出的是各要素的逐年平均变化曲线。

图 1a 平均气温待检序列出现了非常明显的跳跃, 而参考序列比较均一, 由表 2 可见待检台站与参考台站的相关系数都在 0.750 以上, 相关性较好, 由参考台站构造的参考序列与待检序列的相关系数达到了 0.841。图 1b 为相对湿度, 由表 2 可见对于年平均相对湿度, 待检站与参考站的相关系数介于 0.500~0.700, 参考序列与待检序列的相关系数为 0.677。图 1c 反映出降水量的变化幅度比较大, 由

表2 五台山站(53588)与参考台站各要素观测序列(1957~2005年)的相关系数

	年		春		夏		秋		冬	
	站号	相关系数	站号	相关系数	站号	相关系数	站号	相关系数	站号	相关系数
平均温度	53673	0.846	53673	0.837	53593	0.755	53593	0.791	53772	0.810
	53593	0.844	53593	0.828	53487	0.727	53673	0.750	53673	0.779
	53772	0.840	53772	0.822	53478	0.682	53772	0.729	53593	0.764
	53487	0.764	53487	0.793	53673	0.681	53698	0.671	53663	0.741
	53698	0.758	53478	0.782	53698	0.643	53487	0.647	53487	0.683
r'		0.841		0.844		0.738		0.759		0.790
最高温度	53663	0.725	53663	0.711	53593	0.682	53663	0.621	53663	0.686
	53593	0.693	53673	0.709	53487	0.607	53593	0.617	53478	0.624
	53772	0.675	53772	0.702	53478	0.581	53478	0.592	53787	0.610
	53487	0.657	53593	0.689	53663	0.577	53487	0.562	53487	0.587
	53787	0.652	53487	0.685	53673	0.527	53772	0.562	53772	0.587
r'		0.714		0.731		0.632		0.614		0.653
最低温度	53593	0.886	53772	0.881	53593	0.847	53593	0.830	53772	0.852
	53772	0.865	53673	0.879	53673	0.825	54602	0.777	53593	0.846
	53673	0.864	53593	0.872	53487	0.824	53698	0.769	53673	0.842
	53487	0.824	53487	0.838	53698	0.767	53673	0.753	53698	0.783
	53698	0.815	53698	0.770	54602	0.745	53772	0.702	54602	0.763
r'		0.889		0.900		0.855		0.826		0.864
相对湿度	53673	0.670	53772	0.831	53487	0.827	53593	0.663		
	53593	0.667	53673	0.820	53673	0.809	53673	0.598		
	53772	0.594	53787	0.796	53593	0.779	53663	0.584	<0.300	
	53698	0.521	53593	0.767	53663	0.761	53487	0.572		
	54602	0.518	53663	0.765	53478	0.758	53772	0.572		
r'		0.677		0.847		0.870		0.658		
风速	53478	0.846	53478	0.851	53478	0.613	53478	0.828	53478	0.786
	53673	0.768	53673	0.747	53673	0.536	53673	0.706	53673	0.739
	53782	0.635	53782	0.737	53782	0.346	53782	0.696	53782	0.659
	53698	0.562	53698	0.660			53698	0.641	53698	0.638
	53663	0.544	53593	0.639			53487	0.557	53487	0.633
r'		0.871		0.900				0.865		0.854
降水量	53673	0.697	53698	0.597	53673	0.754	53663	0.647		
	53663	0.646	53772	0.579	53663	0.710	53673	0.621		
	53593	0.618	53673	0.554	53772	0.691	53487	0.617	<0.500	
	53772	0.614	53782	0.519	53487	0.610	53593	0.608		
	53478	0.610	53663	0.511	53478	0.607	53478	0.586		
r'		0.751		0.625		0.824		0.703		

注： r' 为各要素参考序列与待检序列的相关系数。

表2可见对于年降水量,待检站与参考台站的相关系数介于0.600~0.700,参考序列与待检序列的相关系数为0.751。由此看来,除风速以外的其它要素根据邻近台站建立的参考序列应该能够代表五台山地区真实的气候变化序列。但值得注意的是,由于相对湿度和降水要素的空间代表性差于气温要素,因此其检验结果的可靠性也随之略差,需要进一步分析、确认。

2.2 检验序列(差值/比值序列)

图2给出几种要素的检验序列(对于气温、相对湿

度为观测原始序列与所建立的参考序列的差值,对于降水则为相应的比值序列)。从图2a可以直观地看到,年平均气温差值序列从20世纪90年代后期起明显高于以前,明显地看到有一个跳跃(不连续点),春季平均气温差值序列同期也存在同样的现象。年、季最高和最低气温差值序列的变化趋势与平均气温差值序列十分相似(图略),都在同期有一个显著的变化,差值明显变小。从图2b、c可以看到其它要素不同季节也出现了不连续。为了找出可能存在的间断点,采用SNHT方法对比值/差值序列进行检验。

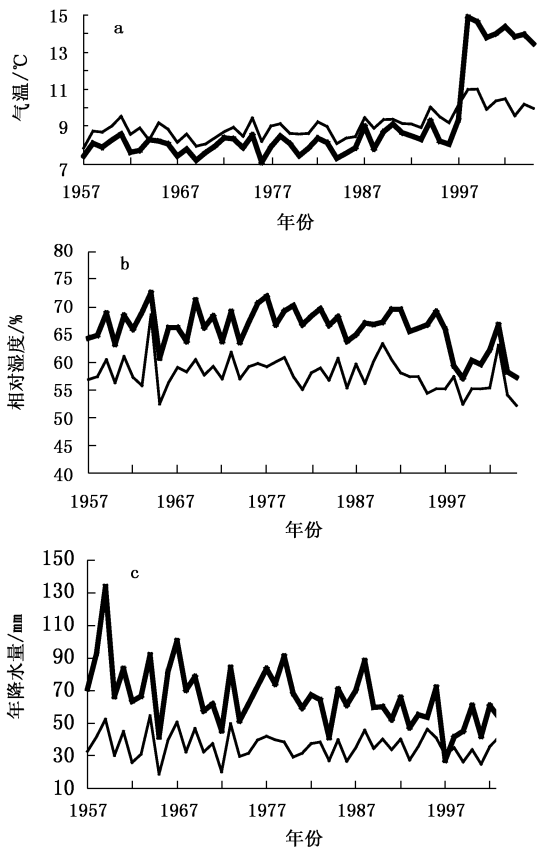


图 1 五台山站(53588)各要素逐年平均的待检序列、参考序列:(a)平均温度;(b)相对湿度;(c)降水量(粗线为待检序列,细线为参考序列)

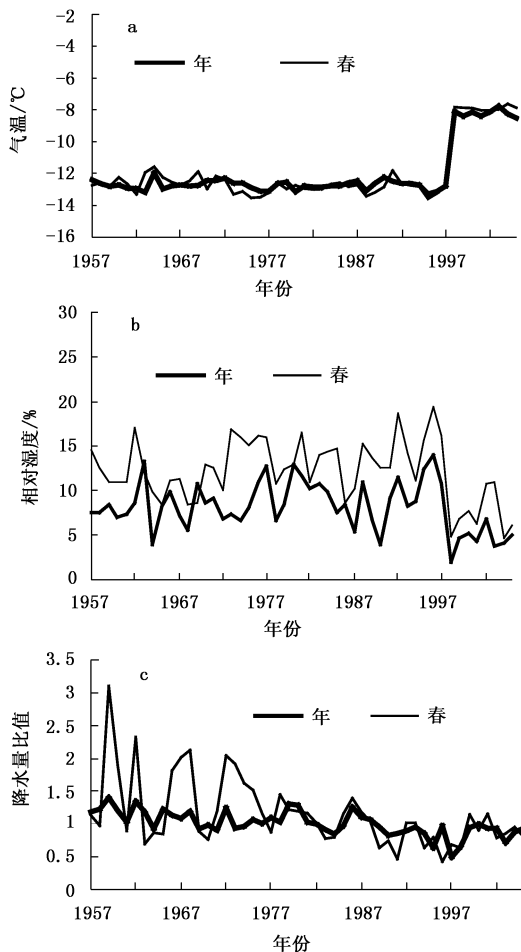


图 2 五台山站(53588)各要素年、季原序列与参考序列之差值(a)、(b)及比值序列(c)

2.3 统计假设检验

对于序列 $\{X_i\} (i=1, 2, \dots, n)$ 假设:

如果 $\{X_i\}$ 序列没有断点存在, 则统计假设为 $H_0: \forall t, X_t \in N(0, 1)$

如果 $\{X_i\}$ 有一个间断点 a , 则统计假设为:

$$\begin{cases} X_i \in N(\mu_1, 1) & i \in \{1, \dots, a\} \\ X_i \in N(\mu_2, 1) & i \in \{a+1, \dots, n\} \end{cases} \quad (5)$$

式中 μ_1, μ_2 分别为假设间断点 a 前后两个序列的平均值 ($\mu_1 \neq \mu_2$), n 为样本总量。根据最大似然比率构造统计量 T^s 作为显著性判据:

$$T^s = a \bar{z}_1^2 + (n - a) \bar{z}_2^2 \quad (6)$$

\bar{z}_1, \bar{z}_2 分别表示 a 前后的平均值, 如果 T^s_{max} 大于选定的显著性信度水平, 假设被拒绝, 即存在非均一断点。对于多个断点的检验、判断, 本文采用半级分段法, 如检验到突变点 T_1 , 再判断是否有其他突变点时, 将该序列分为两子段 $[1, a)$ 和 $[a+1, n]$, 如果再

得到两个突变点 T_2 和 T_3 , 需要检验 $[T_2, T_3]$ 之间突变点是否存在。如果不存在, 下一步要检验 $[1, T_2], [T_2+1, T_3], [T_3+1, N]$; 反之, 就要按照 $[T_2, T_3]$ 间检验到的突变点重新划分检验区间, 直到子段不能再被划分为止。

2.4 T^s 值序列及检验结果

图 3 给出的是平均气温、相对湿度和降水量的年平均 T^s 值序列。

从图 3 可以看到平均气温、相对湿度在 1997 年值都达到了峰值, 远远超过 95% 的显著性水平。为了防止此前可能还存在另一个不连续点被屏蔽掉, 本文对 1957~1996 年序列重新进行了检验, 结果显示: 最低气温的 T^s 值在 1990 年为 8.3, 超过了 95% 的显著性水平, 连续几年持续走高, 1993 年达到最大值 13.8; 同期平均气温 T^s 序列的峰值出现在 1994 年, 为 6.3, 没有达到 90% 的显著性水平; 最高

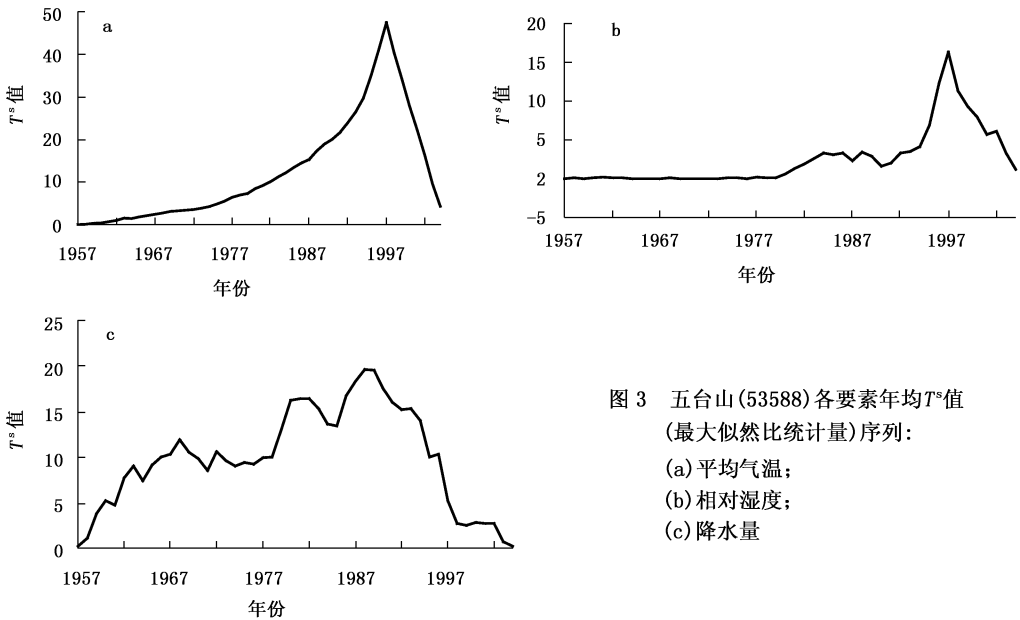


图3 五台山(53588)各要素年均 T^* 值(最大似然比统计量)序列:
(a)平均气温;
(b)相对湿度;
(c)降水量

气温的峰值出现在1988年,为6.5,也没有达到90%的显著性水平;相对湿度也在1994年出现峰值6.9,几乎达到90%的显著性水平(7.00)。这是由于最低温度一般出现在夜间,在非均一性检验中最为敏感。

表3列出了统计量 T^* 序列检验的五台山站观测序列中的不连续点。再对 T^* 统计量检验的不连续点的前后时间序列求平均,用间断点之后的平均值减去间断点之前的平均值(对于降水量,用间断点之后的平均值减去间断点之前的平均值,再除以间断点之前的平均值),就得到间断点的变化幅度,表3也列出了各要素观测序列不连续点的变化幅度。

(1)从气温的 T^* 检验结果来看,1997年这个不连续点在平均气温的年均和四季中均有体现,在最高、最低气温序列中,也得到了这个间断点,由此可见在该站气温观测序列中,1997年这个不连续点是较为一致的。再从1997年不连续点的变化幅度来看,最高气温的年均和四季变化幅度比同期最低气温要高 $1.0\text{ }^{\circ}\text{C}$ 以上,说明了1998年迁站给最高气温带来的不均一性最显著,平均气温次之,最低气温再次之;此次迁站对气温(不论是平均、最低还是最高)春季序列造成的不均一最显著,达 $5\text{ }^{\circ}\text{C}$ 以上,相比较而言对夏季序列带来的不均一最弱。

(2)从相对湿度的 T^* 检验结果来看,1997年这个不连续点在年均和春季也有反映,夏、秋季没有检

测到,这应该与夏、秋季相对湿度比较大有关。参考台站的夏季平均相对湿度都在50%以上,五台山站夏季平均相对湿度最低也接近70%;参考台站的秋季平均相对湿度在45%以上,五台山站在50%以上,1998年的台站变化没有给夏秋季节带来不均一。冬季五台山站与参考台站的相关性较差(相关系数都小于0.500),没有建立参考序列,因而无法检验台站变化是否给冬季的相对湿度带来不均一。

(3)对于降水量,统计量 T^* 序列没有能够检测出1997年这个点。其原因可能有:1998年1月1日的迁站对降水量带来的变化不明显;五台山站周围台站密度不够大;降水量的空间连续性比较差,空间代表性差于气温、湿度等其它要素等。

对其它的不连续点是否都是真正的非均一性断点进行进一步的判断和确认。气温要素和相对湿度的春、夏季序列检测到1997年以外的不连续点,而这些点在相应要素的年序列中并没有体现出来,换句话说,年平均气温和湿度序列在这些年份并没有表现出明显的不连续性,因此,这些不连续点的出现应该是由季节序列的波动导致的。年降水量序列检测到1988年为不连续点,经查阅台站历史沿革,1988年并没有出现仪器更换等导致不连续产生的事件,再与其它要素进行对比,发现其它要素并没有检测到1988年这个点,那么这个点应该是由统计方法的干扰带来的。

表 3 五台山站(53588)各要素观测序列不连续点及其变化幅度

气象要素	时间	不连续点(年)	变化幅度	不连续点(年)	变化幅度	不连续点(年)	变化幅度
平均气温/℃	年	1997	4.5				
	春季	1972	-0.5	1997	5.1		
	夏季	1997	4.2				
	秋季	1997	4.8				
	冬季	1996	4.5				
最高气温/℃	年	1997	5.5				
	春季	1997	6.0				
	夏季	1965	0.4	1997	5.1		
	秋季	1997	5.6				
	冬季	1996	1996	5.8			
最低气温/℃	年	1997	3.4				
	春季	1972	-0.6	1994	-1.4	1997	5.1
	夏季	1991	-0.5	1997	3.5		
	秋季	1997	3.6				
	冬季	1997	4.2				
相对湿度/%	年	1997	-4				
	春季	1972	3	1997	-7		
	夏季						
	秋季						
	冬季	(因相关较差,没有建参考序列)					
降水量(比值)	年	1988	-23%				
	春季	1992	-35%				
	夏季						
	秋季	1961	-24%				
	冬季	(因相关较差,没有建参考序列)					

3 结论

(1) 由于台站迁移的影响,该站气候资料序列存在着明显的非均一性。在研究区域气候变化特征(尤其是局地气候变化)时,必须认真考虑序列本身的均一性问题,以减少研究结果的不确定性。但是,由于不同要素的空间代表性不同,因此在针对不同要素的均一性研究时,必须从要素场的空间结构分析入手,区别对待。

(2) 1998 年台站迁移造成了五台山站气温序列明显的不连续,经对比分析,发现对最高气温序列,不管是年均还是四季带来的不均一性最显著,平均气温次之,最低气温最弱;对气温要素(不论是平均、最低还是最高气温)春季序列造成的不均一性最显著,达 5℃ 以上,相比较而言对夏季序列带来的影响最小。

(3) 1998 年台站变化对相对湿度的年序列和春季序列均产生了较明显的不连续,而对湿度比较大的夏、秋季节则没有产生明显的不连续;此次台站迁移对降水量带来的变化不明显,由于五台山站周

围台站密度不够大;降水量的空间连续性又比较差,降水量检验序列没有出现明显的不均一现象。但值得注意的是,由于降水和湿度要素的空间代表性差于气温要素,因此,上述检验结果还有待于利用更大密度的序列加以进一步验证。

参考文献

- [1] 李庆祥,刘小宁,张洪政,等. 定点观测气候序列的均一性研究[J]. 气象科技,2003,31(1):3-10.
- [2] 翟盘茂. 中国历史探空资料中的一些过失误差及偏差问题[J]. 气象学报,1997,55(5):563-572.
- [3] 李庆祥,Menne M J. Williams Jr C N,等. 利用多模式对中国气温序列中不连续点的检测[J]. 气候与环境研究,2005,10(4):736-742.
- [4] 刘小宁. 我国 40 年年平均风速的均一性检验[J]. 应用气象学报,2000,11(2):27-34.
- [5] Li Qingxiang, Liu Xiaoning, Zhang Hongzheng, et al. Detecting and adjusting on temporal inhomogeneity in Chinese mean surface air temperature dataset [J]. Adv. Atmos. Sci., 2004, 21(4): 260-268.
- [6] 崔宜少,李建华,丛美环,等. 威海气象站年平均气温等气象要素资料的非均一性检验[J]. 应用气象学报,2006,17(4):511-513.

- [7] 吴利红,康丽莉,陈海燕,等. 地面气象站环境变化对气温序列均一性影响[J]. 气象科技, 2007,35(1):154-158.
- [8] Alexanderson H. A homogeneity test applied to precipitation data [J]. J. Climatol. 1986,6:661-675.
- [9] 黄嘉佑. 气象统计分析与预报方法[M]. 北京:气象出版社, 2003.

Homogeneity Analysis of Historical Meteorological Data at Wutaishan

Gao Xiaorong¹ Li Qingxiang² Dong Wenjie³

(1 College of Atmospheric Sciences, Lanzhou University, Lanzhou 730000; 2 National Meteorological Information Center, Beijing 100081; 3 National Climate Center, Beijing 100081)

Abstract: According to the metadata of in situ observation at Wutaishan (53588) of Shanxi Province, including information about station relocation, instrumentation change, and observing and data processing procedure changes, etc., the homogeneities of annual and seasonal mean, maximum and minimum temperatures, relative humidity, and precipitation time series are analyzed from January 1957 to 2005. The reference series for different meteorological variables are constructed by combining the corresponding observed data at five nearby stations with the largest correlation in candidate stations, and then based on the Standard Normal Homogeneity Test (SNHT) technique, each meteorological variable is statistically tested, respectively. The discontinuous points obtained from SNHT are analyzed and compared. Results indicate that the influence of station relocation on 1 January 1998 on the homogeneities of different meteorological variables mentioned above is very obvious, and that on the maximum temperature time series is the most significant, being 5.5 °C. Further analysis shows that this station relocation has different effects on the other meteorological variables time series; for example, the influence on the homogeneities of annual and spring relative humidity time series are also obvious, but not on precipitation. The tests and adjustment of the in situ climatic data series used frequently are important to local climate change researches. The effects of various “non-climate” factors on different meteorological variables vary, so more careful distinction and further research are needed.

Key words: station relocation, Standard Normal Homogeneity Test (SNHT), homogeneity test.