

年降水量序列非均一性

检验方法探讨¹⁾

刘小宁 孙安健

(国家气候中心,北京 100081)

提 要

使用我国 400 余站 1951—1990 年年降水量序列,利用比值法检验其非均一性问题。结果表明,该方法对判断我国年降水量是否均一是较为有效的。我国台站迁移及雨量器的改变是引起年降水量序列非均一的主要原因。

关键词: 年降水量序列 非均一性 检验

气候变化研究的基础是建立能反映气候真实变化的均一性的气候资料序列。但是由于人为因素的影响往往会造成气候资料序列的非均一,使整个序列所显示的气候变化规律,尤其是气候突变,变得不真实,甚至虚假。因此检验气候资料序列的非均一性,建立均一的气候序列在气候研究中有重要意义。国内外许多学者已作过大量研究^[1-4],但比较多的是涉及温度要素。而降水序列,由于年际变化大,地区差异明显,有关它的非均一性的研究尚不多见。

本文利用 Alexandersson 提出的比值法^[5],使用我国 400 余站 1951—1990 年年降水量资料,探讨该方法对判断年降水量序列是否均一的适用性。并依据相应的台站历史沿革库资料,对产生非均一的原因进行分析。

1 检验方法

由于降水过程主要是天气系统活动所致,因而邻近地区的降水量序列之间有一定的可比性。所以通常使用被检验站降水序列与邻近地区参照站的降水序列的比值序列来检验。检验按如下两个步骤进行。

1.1 序列资料的预处理

对于被检验站降水序列 x_i 作如下处理:
 $f(x_i) = x_i / \bar{x}$ \bar{x} 为被检验站降水量平均值。

对于参照站函数 $g(y_j)$,用 k 个参照站的加权平均求取。本文采用相关系数平方作为加权因子。即

$$g(y_i) = \frac{\sum_{j=1}^k v_j (y_{ij} / \bar{y}_j)}{\sum_{j=1}^k v_j}$$

其中 $v_j = r_j^2$ r_j 为被检验站与第 j 个参照站之间的相关系数。 \bar{y}_j 为第 j 个参照站的算术平均值。 k 为参照站的站数。 y_{ij} 为第 j 个站第 i 年的年降水量。

就年降水量序列,求出其比值序列 q_i :

$$q_i = f(x_i) / g(y_i)$$

对 q_i 序列进行标准化处理,形成 z_i 序列:

$$z_i = (q_i - \bar{q}) / S_q$$

其中 \bar{q} 为比值序列 q_i 的算术平均值, S_q 为该序列的标准差。

$$z_i = 0 \quad S(z_i) = 1 \quad (1)$$

即 z_i 序列为平均值为 0, 标准差为 1 的正态分布。

1.2 序列假设检验

(1) 如果 $\{z_i\}$ 序列无间断点,统计检验为:

$$\text{零假设 } H_0: z \in N(0, 1) \forall i$$

1) 本文系国家科技攻关项目《气候变化规律及数值模拟的研究》资助。

(2)如果 $\{z_i\}$ 序列有一间断点,且出现在序列 v 处;统计检验为:

$$H_1: \text{对某些 } 1 \leq v < n \text{ 和 } \mu_1 \neq \mu_2 \text{ 有}$$

$$\begin{cases} z \in N(\mu_1, 1) & i \leq v \\ z \in N(\mu_2, 1) & i > v \end{cases} \quad (2)$$

这里 $z \in N(0, 1)$ 为正态分布, μ_1, μ_2 分别为

$$\max_{\mu_1, \mu_2, v} \frac{(2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2} [\sum_{i=1}^v (z_i - \mu_1)^2 + \sum_{i=v+1}^n (z_i - \mu_2)^2]}}{(2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2} \sum_{i=1}^n z_i^2}} > C \quad (4)$$

将式(3)代入式(4)

$$\max_{1 \leq v < n} [v \bar{z}_1^2 + (n-v) \bar{z}_2^2] > 2 \ln C = C' \quad (5)$$

$$\text{令 } T_v = [v \bar{z}_1^2 + (n-v) \bar{z}_2^2]$$

可构造检验统计量

$$T_0 = \max_{1 \leq v < n} \{T_v\}$$

$$= \max_{1 \leq v < n} [v \bar{z}_1^2 + (n-v) \bar{z}_2^2] \quad (6)$$

间断点 v 前后两个序列的平均值, v 为假设的间断点。 n 为样本数。

$$\mu_1 = \bar{z}_1 \quad \mu_2 = \bar{z}_2$$

$$\bar{z}_1 = \frac{1}{v} \sum_{i=1}^v z_i \quad \bar{z}_2 = \frac{1}{n-v} \sum_{i=v+1}^n z_i \quad (3)$$

根据式(6)可计算出检验统计量 T 序列, T_0 为该序列的最大值。

当 T_0 大于某临界水平值 T_{95} ,则该序列为该水平上的非均一。 T_0 临界值与序列 n 长度有关。表1给出了 n 从25到41的临界值(T_{90}, T_{95})。

表1 不同 n 值下的临界值(T_{90}, T_{95})

	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41
T_{90}	6.55	6.58	6.60	6.63	6.66	6.69	6.71	6.74	6.77	6.79	6.82	6.85	6.88	6.91	6.93	6.96	6.99
T_{95}	7.75	7.78	7.81	7.85	7.88	7.91	7.94	7.97	8.01	8.04	8.07	8.10	8.13	8.17	8.20	8.23	8.26

2 检验前的数据处理

2.1 建立初始序列

使用上述比值法进行降水序列非均一性检验,要求原序列必须服从正态分布。但是在气候较为干旱的地区,年降水量并不遵从正态分布,而是服从Pearson IV型分布。鉴此,根据文献[6]的方法,通过某种变换,将原始的年降水量变换为服从正态分布的 u 变量。 u 变量序列作为比值检验的初始序列。

Alexandersson 提出原序列作平方根变换后的序列近似地服从正态分布。为了比较,检验中也采用平方根变换的序列作为年降水量的初始序列。

2.2 选择参照站

使用比值法检验,选择好的参照站是检验的一个重要前提。显然,被检验站周围可供选用的站点越多、关系越密切越好。

参照站的选择。第一步,对全部参照站进行 u 变换或平方根变换,并对参照站序列进行正态分布检验。第二步,求出被检验站与参照站的相关系数,并做相关系数的可靠性检验($\alpha=0.05$),在达到此显著性水平的站中选

出5个相关系数最高的站,作为被检验站的参照站。西部地区,由于站点较稀,参照站数可以少于5个,但至少应有2个。第三步,对所选出的参照站进行均一性检验,若所选参照站为非均一,则删除该站,另行增补。为了实现这一点,只有利用计算机巡回检验的程序功能。

3 检验分析

3.1 中断原因

对400余站年降水量序列按上述方法和步骤,进行检验,共检验出15个站的序列有间断点。根据间断点的年份,查找台站沿革库,分析中断原因,影响年降水量序列中断的主要原因为台站迁移及雨量器的改变(表2)。从表2可见,因雨量器变化引起的中断年份大都在1960年前后。这显然和我国1949年以来雨量器变化有关。在1950—1960年,雨量器安装高度多数为2.0m高,且有防风罩。自1960年以来,雨量器高度一般由2.0m降至0.7m,并无防风罩。于是,在我国风力较大的三北地区(东北、华北、西北)的一些台站就在1960年前后出现了年降水量序列的非

均一。

表 2. 检验结果及间断原因

省	站名	区站号	间断年	*	原因*
黑龙江	克山	50658	1953	①	1954年雨量器安装变化
黑龙江	海伦	50756	1960	①	1961年1月检查雨量器修正并调整
新疆	乌鲁木齐	51463	1959	①	1960年9月雨量器由2.0m改为0.7m
新疆	若羌	51777	1966	③	
新疆	安德河	51848	1966	③	
青海	德令哈	52737	1973	②	1974年1月迁站高差100m距原址30km
甘肃	乌鞘岭	52787	1957	③	
青海	都兰	52836	1959	①	1960年10月雨量器高度由2.0m改为0.7m
	北京	54511	1960	③	
河北	乐亭	54539	1960	①	1960年10月雨量器高度由2.0m改为0.7m
四川	理塘	56257	1959	③	
西藏	林芝	56312	1968	④	无1969年资料,无法查原因
四川	金佛山	57622	1956	②	1956年1月迁站,高度由2041.8m改为1798.9m
浙江	定海	58477	1970	②	1969年9月迁站由山脚到山顶
浙江	括苍山	58653	1963	③	

* ①雨量器变化引起;②迁站变化引起;③高山站或西部站或原因不明;④缺资料无法查原因

表 3 被检验站与参照站情况

被检验站		参照站			
区站号	站名	区站号	站名	海拔高度	距被检验站距离
51777	若羌 (888.3m)	51765	铁千里克	847.1m	209km
		51656	库尔勒	932.7m	350km
		51711	阿合奇	1986.0m	900km
		51716	巴楚	1117.4m	810km
52787	乌鞘岭 (8045.1m)	52866	西宁	2262.2m	120km
		52889	兰州	1518.3m	140km
		52681	民勤	1368.5m	170km
		52495	巴音毛道	1329.0m	370km
		53149	满都拉	1226.1m	830km

表 2 中原因不明的 6 个站,将其中两个被检验站与相应的参照站情况列入表 3。从表 3 看出,西部地区,由于台站较稀,地形复杂,被检验站很难找到水平距离与海拔高度均很相近的若干参照站(即使降到 2 个)。由这些参照站判断被检验站的间断年本身就存在无法避免的缺陷,故而所诊断的间断年份往往是虚假的。因此统计检验的结果必须得到该站历史沿革信息的印证。

3.2 个例分析

图 1 给出了乌鲁木齐站的比值序列 q_i 与检验统计量 T 值。由图可见,1959 年处的 $T_{max} = 9.5$ 远大于 T_{05} 值(该序列样本数为 39,由表 1 查得 $T_{05} = 8.20$)。可以判断序列在 1959 年有明显间断。从历史沿革库查,1960 年该站迁过站,但环境变化不大(查 1960 年

11 月该站从 43°49'N, 87°33'E, 海拔 850.5m 西郊迁至 43°51'N, 87°31'E, 海拔 653.5m 西郊机场),更主要的是该站 1960 年 9 月雨量器安装高度由 2.0m 降至 0.7m,且由有防风罩改为无防风罩。据此可认为雨量器的改变是造成该站年降水量序列在 1959 年出现间断的主要原因。

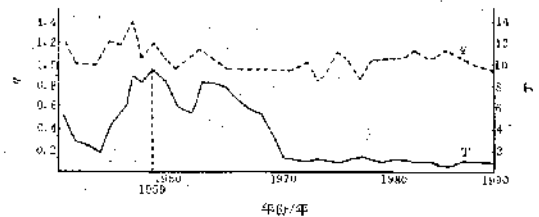


图 1 乌鲁木齐站年降水量序列的 q_i, T 图

图 2 为新疆安德河站的比值 q_i 和检验统计量 T 值。从图中可看出,1966 年 T 值有一个明显峰值。 $T_{max} = 8.5$, 大于 $n = 21$ 时 T_{05} 值 7.63。序列在 1966 年出现间断。从参照站看,海拔高度与安德河站相近,只是水平距离相差较远。南疆地区降水量稀少,水平梯度很小,水平梯度差异所致的影响程度远小于东部地区。故所选参照站还是可利用的。但从所掌握的台站历史沿革库看,还未找到引起 1966 年间断的明显原因。

3.3 几个探讨的问题

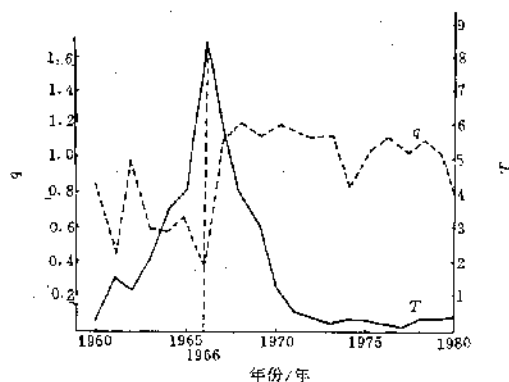


图2 安德河站年降水量序列的 q, T 图

3.3.1 建立初始序列方法的比较

根据比值法检验的要求,序列必须服从正态分布。用原序列及用平方根变换分别做为初始序列,用以比较检验效果。计算结果表明,用原序列共多判6个站出现间断年²⁾。经查台站历史沿革,这6个站未发现引起间断的原因。这些站基本位于干旱、半干旱气候区。经检验除腾冲外,原始年降水量序列不服从正态分布。原始序列必须经过变换使其服从正态分布才能使用比值法检验。

为比较用 u 变换和平方根变换建立初始序列方法的优劣,选取了巢湖(58326, 1966年5月迁站)、金川(56168 站址无变化)、富蕴(51087, 1961年6月迁站)、北京(54511, 1965, 1981年迁站)4个站。这4个站均有若干更邻近的参照站。检验结果见表4。

表4 u 变换与平方根变换序列检验比较

	u 变换	平方根变换
58326	未间断	1965年间断
56168	未间断	未间断
51087	未间断	未间断
54511	未间断	未间断

从表4可看出,平方根变换做为初始序列,其效果稍佳于 u 变换。

3.3.2 参照站的选择

被检验站与参照站的相关系数与其间距离有着密切关系。一般距离愈远相关系数愈小,如原平站(53673)与参照站相关系数为0.81时,相距160km;当 r 为0.67时,距离达560km;蔚县(53593)站 r 为0.75时,距离为200km; r 为0.52时,距离为600km。这种差异直接影响到检验效果,甚至真伪。表4中北京站所用参照站为市区或近郊的台站(如朝阳、门头沟等),检验未见间断。但在表2中,所用的参照站较远(最近为天津),检验结果北京在1960年为间断。统计检验只是显著,必须经台站历史沿革资料查证才能最终确定。查北京沿革资料,并未在1960年前后找到解释原因。据此可以确定表4的结果是真的,表2的结果是伪的。参照站的选择对检验结果是十分关键的。

参考文献

- Potter, K. W. Illustration of a new test for detecting a shift in mean in precipitation series. Mon. Wea. Rev., 109, 2040.
- A Review David R. Easterling Thomas C. Peterson. Techniques of detecting and adjusting for artificial discontinuities in climatological time series. Global Climate Laboratory, National Climate Data Center, Asheville, Nc. 28801.
- 宋超辉, 张谦. 气候序列非均一性检查和资料订正方法的初步探讨. 国家气象中心技术报告 9003号.
- 刘小宁, 宋超辉, 李集明. 气象资料序列非均一性检查的试验及初步分析. 国家气象中心技术报告.
- Alexandersson, H. A Homogeneity test applied to precipitation data. Journal of climatology, 1986, 6: 661-675.
- 曲建和. 华北平原旱涝气候特征的研究. 山东气象(增刊), 1989, 5.

An Inhomogeneity Test Study of Annual Precipitation Series

Liu Xiaoning Sun Anjian

(National Climate Center, Beijing 100081)

Abstract

The inhomogeneity has been tested by using ratio method, based on annual precipitation series from 1951 to 1990 for four hundred stations in China. The results suggest that this method is effective to judge the inhomogeneity of annual precipitation series changes of stations and gauges are major reasons which caused the inhomogeneity of annual precipitation series.

Key Words: annual precipitation series inhomogeneity test

2) 分别为:阿巴嘎旗(53192)1957年;阿勒泰(51076)1957年;理塘(56257)1965年;腾冲(56739)1973年;那曲(55299)1956年;朱日和(53276)1955年。