

重建历史降水量场的统计模拟方法

丁裕国 冯燕华*

(南京气象学院)

摘要 从相邻地区旱涝分布特征的相互关联性出发,提出一种借助于 Monte-Carlo 随机模拟思想重建历史降水量场的统计模拟方法。据此,首先拟合各级旱涝所对应的降水量条件分布模式,由多次数值试验产生相应于各级旱涝的模拟降水量记录。这些模拟记录在一定程度上重现了历史上大范围降水量场的特征,从而使历史旱涝等级场转化为降水量场。试验表明,重建的降水量场对于气候变化的研究是有益的。

关键词 统计模拟,降水量,旱涝等级

历史时期气候变化是现代气候变化的背景,历史记载所获得的长年代旱涝(或冷暖)资料信息,可弥补器测记录年代太短的缺陷。70年代中期,我国气象工作者曾以地方史志为基础,整编出版了各省旱涝历史资料和近五百年旱涝等级分布图集^[1],为研究我国近五百年气候变化规律提供了依据。许多学者在此资料基础上对我国五百年旱涝时空变化作了大量的研究^[2-6],这些研究表明,该资料集具有基本的可靠性。但是,随着气候变化研究的深入,仅借助于旱涝等级资料来研究气候变化似显不足。因为,各地旱涝的实际状况与降水量气候特征有密切关系,同一等级旱(或涝)的降水量在各地并不相同,因此,旱(涝)等级序列往往显示不出各地降水量的气候差异。事实上,我国各地降水量不但其平均值随地理空间分布的差异甚大,且其他多种统计特征(如降水变率、极值、概率等)也随地区而变化。所以,从理论研究或气候预报的角度来看,迫切希望能有一套更为完整具体地反映各地降水量(或距平)变化规律的气象记录。本文基于上述目的,提出恢复历史降水量场记

* 魏春秀(吉林省气象局)、王 霞(贵州省气象局)参加了本文的部分计算工作
收稿日期:1991-09-09

录的一种统计模拟方法。

1 基本资料和模拟方法

统计实验表明,在气候状况基本相近的区域内,旱涝等级记录与降水量记录(它的标准化量)具有时空域上的相关性。例如,我们选取华北区(含 11 个站)、华东区(含 11 个站)和西北区(含 14 个站),近 30 年(1951—1980 年)旱涝等级记录和同期夏季(6 至 9 月)降水量作为试验样本资料,计算各区域内降水量标准化距平(即标准化量)与旱涝等级的相关系数,结果发现,各区域内旱涝等级资料与降水量标准化距平均有较高的正相关,其中华北区为 0.83,华东区为 0.89,西北区为 0.76。由于降水量已标准化,在同一区域上所选的站点全部有 30 年记录,这等于在时空域上取了一个相当大的样本(即华北区 $N=11 \times 30$;华东区 $N=11 \times 30$;西北区 $N=14 \times 30$)。这种时空域上的高相关实际上表明,就统计总体而言,各等级旱涝分别对应有降水量的一个条件概率分布,其条件方差较小。绘制各区域不同旱涝等级所对应的标准化降水距平的频数直方图就可获得相应等级下的降水

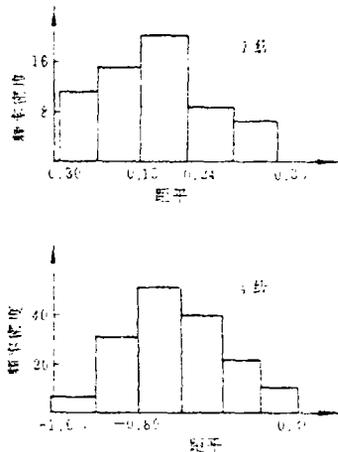


图 1 华北区降水量标准化距平频数

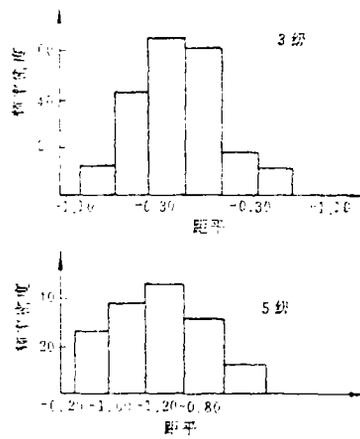


图 2 华东区降水量标准化距平频数

量样本条件分布(见图 1、图 2、图 3)。

根据统计学理论,由已知样本频率(数)分布可适度配合它所服从的某总体概率分布。例如,分析图 1、2、3 发现,华北区 1、3、4 级,华东区 2、3、4、5 级和西北区 2、3、4、5 级均大致符合正态分布,而 3 个区域的其余等级则可配合均匀分布(图略)。χ² 拟合优度检验表明,这些分布模式的选择基本适当。一般说来,旱涝等级记录应是历史上逐年降水量的一种假括。利用已知样本的条件直方图所拟合的分布模式,原则上代表了历年各级旱涝所对应的降水统计特征。因此,根据这些条件分布模式产生相应的模拟记录(随机数),经多次

重复试验就可近似地重现出历史降水量及其区域分布(降水量距平场)序列。这种以条件分布为基础的随机模拟方法(即 Monte-Carlo 试验)是文献[7]提出的随机模拟方法的一个推广。

利用文献[8—10]首先可产生(0, 1)均匀分布随机数,再经适当变换就可得到符合各等级旱涝的降水量条件分布模式的模拟值。例如,采用乘同余法产生(0, 1)均匀分布随机数 r_1, r_2, \dots, r_n 后,由反函数法可求任意区间 (a, b) 上的均匀分布随机数,即由变换公式

$$x_i = a + (b - a)r_i \quad i = 1, 2, \dots, n \quad (1)$$

得到任意区间 (a, b) 上的均匀分布随机数 x_i 。又由中心极限定理,可构造统计量

$$z_i = \frac{\sum_{i=1}^n r_i - \frac{n}{2}}{(n/2)^{1/2}} \quad i = 1, 2, \dots, n \quad (2)$$

当 n 足够大时, z_i 渐近地服从 $N(0, 1)$ 分布。因此,由 z_i 很容易得到非标准化正态变量^[9]。

本文在计算中注意到随机数与模式的一致性,取 $n=12$ (见文献[10]),求得均匀随机数,同时,在采用乘同余法时,注意到递推公式中

$$x_i \equiv \lambda x_{i-1} \pmod{M} \quad i = 1, 2, \dots, n$$

其中参数 x_0 (初值)、 λ (乘子)、 M (模)的选取要适当,才能更优地符合概型的要求。经验表明,为了使随机数具有较好的统计性能,参数的选取应试验多次。本文最终选取了 3 组已通过频率检验和独立性检验的参数(见表 1),从而求得 3 列(0, 1)均匀分布和 $N(0, 1)$ 正态分布随机数。将 3 列不同参数下所求得的均匀分布随机数序列加以平均即得到相当稳定的均匀分布随机数序列。同理,对于相应的正态分布也可求得其平均的随机数序列。根据各个试验区域不同等级旱涝所提供的分布模式参数(见表 2 和表 3),便可逐站恢复并延长历年(6—9 月)降水量模拟记录。

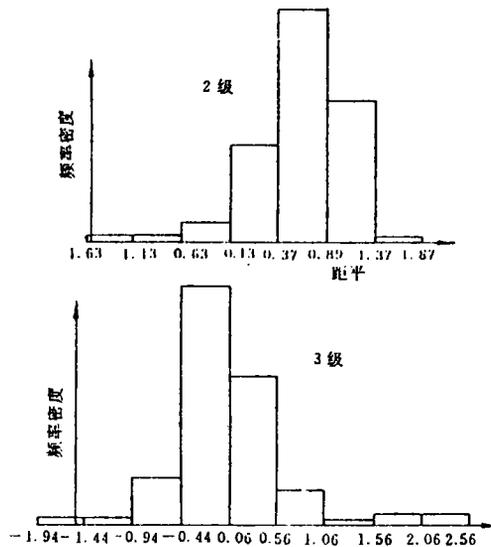


图 3 西北区降水量标准化距平频数

表 1 产生(0,1)均匀分布的参数

序号	M	λ	x_0	L(周期)
1	2 ¹¹	5 ¹¹	1	10 ⁹
2	2 ¹⁶	5 ¹²	1	2×12 ¹⁰
3	2 ¹²	5 ¹⁷	1	10 ¹²

表 2 各区域不同级(旱涝)的降水正态分布参数*

级 别	华 北			华 东				西 北			
	1	3	4	2	3	4	5	2	3	4	5
均值	1.53	-0.08	-0.56	0.71	0.10	-0.53	-1.27	0.63	0.05	-0.53	-0.9
方差	0.72	0.45	0.47	0.42	0.38	0.40	0.36	0.48	0.59	0.59	1.0

*表中均以降水标准化距平记录统计,表3同

表 3 各区域不同级(旱涝)的降水均匀分布参数

级 别	华 北		华 东		西 北	
	2	5	1	1		
a	0.20	1.9	0.7		0.6	
b	1.1	-1.0	2.4		2.1	

2 模拟效果分析

统计试验获得的模拟降水量距平场与实测降水量距平场具有相当好的拟合效果。就各站总拟合结果来看,历年(6—9月)实测降水量与统计模拟降水量一般都有较高的相关性,从表4看,相关系数平均维持在0.84左右。其中80%以上的测站相关系数都在0.80以上,有的测站达到0.90以上。为了全面考察模拟效果,进一步从时间变化和空间分布两个方面来检查其拟合误差。

表 4 部分测站 30 年实测降水与模拟降水的相关系数 r 值

站 名	太 原	北 京	天 津	保 定	德 州	安 阳	洛 阳	郑 州	济 南	临 沂	南 阳	信 阳	徐 州	南 京	上 海	阜 阳	合 肥	安 庆	杭 州	宜 昌	武 汉	平 均
r	0.81	0.82	0.84	0.89	0.85	0.85	0.82	0.91	0.91	0.79	0.88	0.84	0.78	0.91	0.81	0.89	0.78	0.88	0.84	0.88	0.84	0.84

就时间变化来看,各测站实测降水与模拟降水的年际变化趋势十分一致(参见图4)。此外,计算各站30年中历年实测与模拟降水量的相对误差平均值(百分比)发现,约90%的站的平均相对误差在10%以内,其中约67%的测站相对误差仅在5%以内,且误差的分布近似于正态(表略)。这种误差的分布状况,除个别测站以外,基本满足恢复历史降水

量场所要求的精度。

就空间分布来看,我们可绘制和比较逐年各区域内模拟场与实测场(本文以距平场代表)及早涝等级场。如我们以1955年(西北区)和1978年(东部区)为代表绘制了模拟和实测曲线图(略)。可见,该两个年份都能较好的拟合。限于篇幅,本文不再一一列举。

3 资料延长情况下的模拟记录

为了考虑仅有等级记录而无降水记录的重建,对于资料延长情况作了验证和对比。例

如,1949年(偏涝年)、1929年(偏旱年)的西北区和1943年(大旱年),1931年(大涝年)的东部区这类典型年的验证结果表明,除个别测站旱涝等级记录与模拟记录偏差较大(如1931年阜阳)外,其地理空间的等值线分布趋势和范围与旱涝等级的地理分布的高低值中心和分布状况非常吻合(如图5a、b,余略)。由此可见,统计模拟方法恢复历史上无降水记录的各个年份的降水量具有一定的功能。换言之,用现有的各个区域上全部测站短年代资料构成样本条件分布,借助于随机模拟就可推断出较长年代的逐年降水量空间分布。

同样,对于单个测站来说,恢复相应年份的旱涝等级记录为降水量记录,只要将已有样本条件分布所产生的模拟值按历年旱涝等级逐年确定即可得到。例如,假定上海仅有近30年降水记录(即假定1950年前仅有旱涝等级记录),利用华东区已获得的参数和分布模式对照上海历史上逐年旱涝等级,分别抽取模拟值(随机数)就可延长到较长的年代。同理,利用华北区已得到的模式和参数对照北京历史上逐年旱涝等级分别抽取模拟值(随机数)就可使北京降水量记录延长。之所以选择上海、北京两站作试验,就是为了实际考察上述方法对于只有旱涝等级、文字记载的那些测站是否可行。图6中模拟降水量是假定上海这段年份无降水记录而仅有等级记录,用上述随机模拟方法推断得到的,因此,它等价于对单站降水记录的恢复。由图6可见,无降水记录年代的降水量恢复值与实测值相当吻合。同样,对北京作的验证(1894—1950年)效果亦佳。表5给出上海和北京近百年恢复降

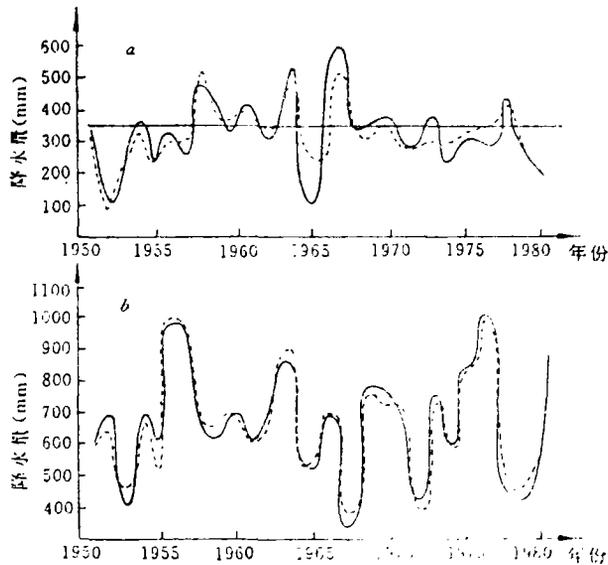


图4 2个典型测站的降水量模拟值的年际变化

实线为实测值,虚线为模拟值

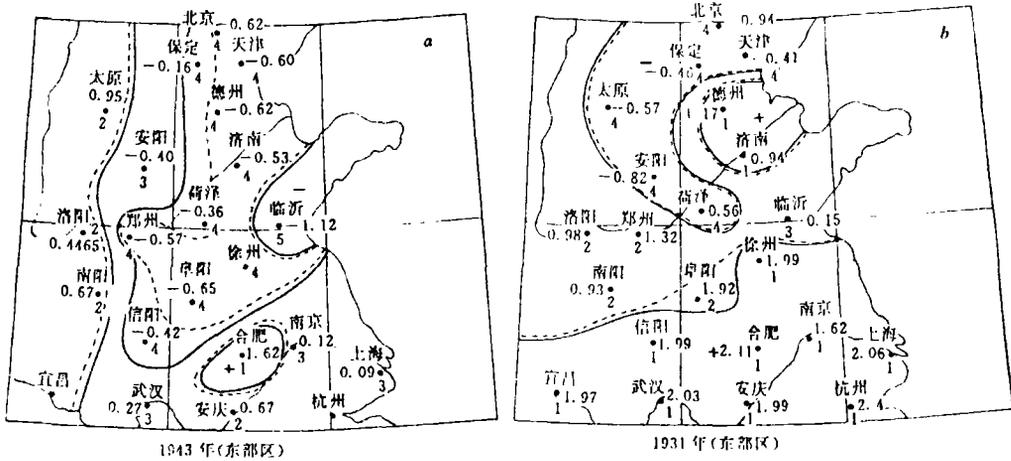


图 5 降水等级场与模拟场的空间分布 — 等级场 ... 模拟场

水量的平均相对误差。由表可见,实际降水记录与模拟记录每 10 年的平均相对误差一般都不太大。最大相对误差不超过 12%。经计算,上海和北京两站模拟值与实测值(以近百

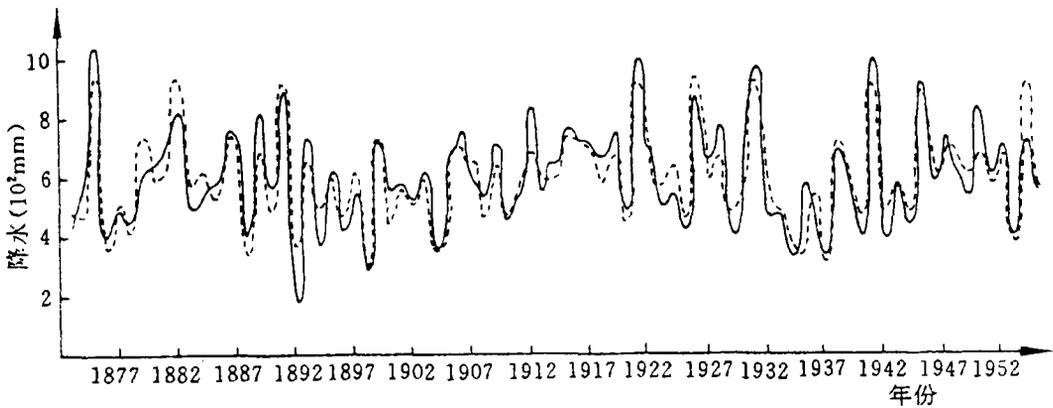


图 6 上海(1873—1950 年)模拟降水量(虚线)与实测降水量(实线)年为据)的相关系数分别达到 0.91 和 0.90。

表 5a 上海模拟降水量记录每 10 年相对误差(%)

起止年份	1873—1882	1883—1892	1893—1902	1903—1912	1913—1922	1923—1932	1933—1942	1943—1952	1953—1962	1963—1972
误差	2.0	2.0	1.1	6.0	6.3	4.6	1.3	0.7	6.8	10.9

表 5b 北京模拟降水量记录每 10 年相对误差(%)

起止 年份	1899— 1908	1909— 1918	1919— 1928	1929— 1938	1939— 1948	1949— 1958	1959— 1968	1969— 1980
误差	4.8	11.2	6.5	5.1	5.7	8.1	6.5	7.5

4 结论与讨论

(1) 本文试验表明,借助于 Monte-Carlo 统计模拟将等级旱涝记录恢复为降水量记录具有良好的效果和一定的可行性。

(2) 从理论上说,各级旱涝正是某一时段内(如 6—9 月)降水量分档的一种结果,而每一级旱涝所对应的降水量,就时空域来说,必然对应有一种概率分布模型。假定关于旱涝状况历史记载的概念基本不变(含排除历史记载的不真实性或社会原因),则在某一较长历史时期可认为各级旱涝对应有一个条件概率分布。从总体上说,同一等级的自然降水量的年际时间分配已经由历史记载确定了,所以,根据它所服从的条件分布模式产生该级(旱涝)的模拟降水量,其气候统计特征(时间平均、方差、极值、概率等)必然符合它原有的分布模式。众所周知,由实测样本记录所估计的条件概率分布模式在一定的信度下可作为总体概率分布的推断。因此,这就构成了统计模拟恢复历史降水场的理论基础。

参 考 文 献

- 1 国家气象局气象科学研究所. 中国近百年旱涝分布图集, 北京: 地图出版社, 1981
- 2 徐瑞珍等. 全国气候变化学术讨论会文集, 北京: 科学出版社, 1981: 52—63
- 3 柳又春, 史慧敏. 南京气象学院学报, 1983, 6(1), 83—93
- 4 柳又春, 李 骅. 气候学研究——“天、地、生”相互影响问题, 北京: 气象出版社, 1989: 80—89
- 5 盛承禹. 气候学研究——“天、地、生”相互影响问题, 北京: 气象出版社, 1989: 75—79
- 6 屠其璞. 南京气象学院学报, 1990, 13(3), 266—273
- 7 丁裕国, 张耀存. 南京气象学院学报, 1989, 12(2), 146—155
- 8 Allen DM, Hean CT et al. *Stochastic Simulation of Daily Rainfall Research Report*, 1975, (72), 1—50
- 9 黄克中, 毛善培. 随机方法与模糊数学应用, 上海: 同济大学出版社, 1987: 208—211
- 10 张建中, 周琴芳等. 常用时间序列分析软件包, 北京: 气象出版社, 1986: 90—100

A STATISTICAL SIMULATION METHOD FOR RECONSTRUCTING PRECIPITATION FIELDS IN HISTORICAL TIME

Ding Yuguo Feng Yanhua

(Nanjing Institute of Meteorology)

Abstract A statistical simulation method for reconstructing precipitation fields in the historical time is suggested from Monte—Carlo random simulation idea and the interrelation of distribution of drought and flood order in neighboring regions. With this method, the conditional distribution model of precipitation amount corresponding to the different orders of drought and flood is fitted and the simulated precipitation records corresponding to the different orders of drought and flood are obtained through many numerical experiments for the last 500 years in China. These simulated records reproduce, to a certain extent, the characteristics of the large—scale precipitation fields in the history, thus turning them into precipitation fields. Results show that the reconstructed precipitation fields are useful to the study of climatic variations during very long historical time.

Key words statistical simulation, precipitation amount, drought and flood order