

基于多状态 Markov 链模式的极端降水模拟试验^{*1}

丁裕国 张金铃 江志红
DING Yuguo ZHANG Jinling JIANG Zhihong

南京信息工程大学,江苏省气象灾害重点实验室,210044

Jiangsu Key Laboratory of Meteorological Disaster, Nanjing University of Information Science & Technology, Nanjing 210044, China

2007-09-04 收稿,2007-10-11 改回.

Ding Yuguo, Zhang Jinling, Jiang Zhihong. 2009. The simulation experiments of extreme rainfall based on a multi-status Markov chain model over east region of China. *Acta Meteorologica Sinica*, 67(1):20-27

Abstract A multi-status Markov chain model is proposed to simulate daily rainfalls records and the extreme rainfalls in summer are fitted from the simulated daily rainfall records by using a Generalization Pareto Distribution (GPD) model. The research results show that this statistical simulation method display most good dependability and the statistic climatic features from the simulated daily rainfall records have reached more high precision for most stations, specially the pluvial region over east area of China. The analysis comparatively shows that the multi-status Markov chain model is excellent with the two status Markov chain model for its simulation ability, specially, the simulated climatic features of extreme rainfalls in east region of China have reached most high precision. The experiment results for the selected six stations demonstrate the excellent simulation effects of above statistical model, for example, the following features; standard deviation of monthly rainfall, maximum daily rainfall, mean number of rain days during a month spell, standard deviation of daily rainfall, mean value of daily rainfall amount etc. are consistent with the real observation values. Both the multi-status Markov chain model and the two status Markov chain model.

Key words Extreme precipitation, Simulation test, Markov chain, Generalization Pareto Distribution model

摘要 文中建立了基于多状态一阶 Markov 链的逐日降水量随机模式并结合广义帕雷托分布(GPD)产生夏季逐日极端降水量的模拟资料,结果所显示的各种气候特征表明,绝大多数站点(尤其是中国东部多雨地区)都达到较高的精度。分析表明,该模式对中国东部极端降水特征的模拟能力在某些方面优于两状态一阶 Markov 链模式。对东部 6 个代表站模拟试验结果表明,月降水均方差、日降水极大值、月平均降水日数、日降水均方差、日平均降水量等指标与实况比较,均证明该模式对逐日降水量的模拟效果较好,基本模拟出降水量的各种特征。对中国东部 78 个代表站采用的两种模式模拟结果对比发现,除日平均降水量以外,月平均降水日数、日降水平均极大值都与实际观测结果较为一致,总体上优于两状态模式,说明用该模式在全国范围内模拟逐日降水特征尤其是极端降水特征有较高的可行性。例如,由其中 6 个代表站模拟资料所拟合的极端降水 GPD 模式具有较高的拟合优度。无论从门限值或重现期值来看都可发现模拟与实测结果有较好的相似性,且两者门限值的误差越小,重现期极值的差距也越小。证明 Markov 链模式对极端降水的模拟有广泛的适用性。

关键词 极端降水量,模拟试验,Markov 链,广义 Pareto 分布模式

中图法分类号 P435 P456.7

1 引言

根据 IPCC 第 4 次科学评估报告,全球持续增

暖且有加剧的趋势,已成为大多数科学家的共识。鉴于目前以 AOGCM (Atmosphere-Ocean General Circulation Models)为基础的耦合气候系统模

* 资助课题:国家自然科学基金项目(40675043)。

作者简介:丁裕国,长期从事气候和极端气候及其变化的诊断与预测研究。E-mail:dygnim@jlonline.com

式其空间分辨率的局限性,模式输出信息并不足以有效地描述全球各区域或局地的气候变化特征,如何考察全球变暖背景下各区域(或局地)气候的变化趋势已成为当务之急。目前正在兴起的降尺度技术方法为解决这一难题提供了有力帮助,所谓降尺度方法是基于下述观点:区域气候变化不但受制于全球大尺度(如大陆尺度,甚至行星尺度)气候强迫,而且还受制于本地区(或局地)气候影响下的区域强迫(例如:地形、海岸线、湖泊和陆面特征),因此,通过降尺度方法可将大尺度、低分辨率的 AOGCM 输出信息转化为区域尺度的地面气候变化信息(如气温,降水),从而弥补 AOGCM 对区域气候变化情景预测的局限性。目前动力降尺度技术虽然有了进一步发展,但也有不可避免的缺点,统计降尺度技术具有低成本,简便易行,且效果并不亚于动力学方法的优点而为大多数学者所能接受(Wilby, et al, 1997; Wilks, 1999; Wilks, et al, 1999)。在全球变暖的背景下,区域气候或局地气候的响应变化更加复杂,目前有关这一领域的研究已成为气候科学界关注的焦点之一。当前人们公认的统计降尺度技术主要包括函数转换法、天气分型法、天气发生器 3 大类方法。后者是一系列可以构建气候要素随机过程的统计模型的总称,它们可以被看作复杂的随机数发生器。其指导思想是通过直接拟合气候要素的观测值,得到统计模型的拟合参数,然后用统计模型(如某种随机过程)模拟生成随机的逐日气候要素时间序列。尽管随机天气发生器方法已经取得了较好的效果,其中很多模拟方法都借助于 Markov 过程,但这些模型中关于逐日气象数据的仿真模拟仍较为粗糙,往往与真实天气气候过程还有一定的差距(Gregory, et al, 1993; Palutikof, et al, 2002)。

20 世纪中期以来,国内外学者早就应用 Markov 链方法在拟合气温、降水等气候要素方面取得了丰硕成果(么枕生等, 1990; Gabriel, et al, 1962; Katz, 1974; Gates, et al, 1976)。例如么枕生(1966)为了改进 Markov 链无后效性的缺点,以原有的各个状态的历史演变来重新定义状态,取得较好效果。周家斌等用 Markov 链预报浙江北部汛期逐日晴雨变化(王宗皓等, 1974)。丁裕国等(1989, 1990)从逐日降水过程干湿日演变的一阶 Markov 链及日降水量的 Γ 分布模式出发,在理论上得到任意给定时期 N 日降水总量及最大日降水量的理论分布函数,同

时得到 N 日总降水量和最大日降水量的数学期望与方差,初步证实所得到模式的普适性(张耀存等, 1990a, 1990b)。此外,丁裕国等(1989, 1990)还利用 Markov 链研究中国各大气候区若干代表测站干、湿月游程的统计特征,得到许多有实际应用意义的气候统计信息。利用多状态 Markov 链建立一种能够生成单站逐日降水量模拟记录的随机模式,结果表明,所模拟的气候统计参数同实测结果十分吻合。廖要明等(2004)建立了一个适用于中国广大地区的随机天气发生器,以两状态一阶 Markov 链模拟为基础,根据各地不同月份计算的 4 个降水模拟参数对中国各地的逐日降水进行模拟并利用年的实测数据对 30 a 模拟结果在统计意义上进行了检验,模拟结果较好。

尽管基于两状态一阶 Markov 链和两参数 Γ 分布的天气发生器取得较理想的结果(廖要明等, 2004; 丁裕国, 1994; 徐钟济, 1985)。然而,由于两状态一阶 Markov 链模式过于简化了天气演变过程,并不能很好地分辨逐日降水量的细节特征,因而在此基础上要建立极端降水的统计特征就有一定的困难。本文正是针对这一问题的关键所在,提出一种改进的多状态一阶 Markov 链作为产生随机天气发生器的主要工具基础,借助于这种统计模式建立较为复杂的随机天气发生器,以便模拟出更加符合当地气候条件的逐日气候要素时间序列,从而更加真实地模拟出极端降水量的时间序列(么枕生, 1984; Ding, et al, 2008)。

2 模式的建立

任何一个测站的逐日降水过程时间序列,其内部演变规律本质上都包含了两方面因素:其一,干湿日序列的时间分布即承替规律;其二,每一湿日的降水量一般都有大小不同的量级,其总体概率分布可找到一种对应的分布模式。因此,仅用简单的两状态 Markov 链模拟逐日降水量时间序列必然有其局限性。为了更加符合实际,本文从上述两个方面入手,首先将湿日按降水量大小分成若干等级,并规定湿日的 n 个状态为 $s_1, s_2 \dots s_n$ 。显然,连同干日 s_0 就可组成降水过程较为完整的 $n+1$ 个状态的状态空间($s_0, s_1, s_2 \dots s_n$)。这种状态划分有利于客观模拟真实的逐日降水过程。

设状态转移符合一阶 Markov 链,并将转移概

率矩阵写为

$$\mathbf{P} = \begin{pmatrix} p_{00} & p_{01} & \cdots & p_{0n} \\ p_{10} & p_{11} & \cdots & p_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ p_{n0} & p_{n1} & \cdots & p_{nm} \end{pmatrix} \quad (1)$$

式中 $p_{ij} = \mathbf{P}(x_t = s_j | x_{t-1} = s_i)$ 表示在第 $t-1$ 时刻(日期), 天气处于状态 s_i , 经一步转移在第 t 时刻, 天气处于状态 s_j 的转移概率。进一步假定 p_{ij} 与转移发生的时间坐标的位置无关, 而仅与前后两个时刻之差值有关, 即具有平稳转移的均匀 Markov 链。由于湿状态 $s_1, s_2 \cdots s_n$ 各自对应于不同的降水量级, 因而每个 $s_i (i=1, 2 \cdots n)$ 必对应于一种概率分布。丁裕国(1994)曾从理论上详细研究降水量符合 Γ 分布的普适性, 并用实测逐日降水资料作了拟合分布的验证。由于 Γ 型逐日降水量的正偏性, 出现小降水量或较少降水量的机会较大, 而出现极端日降水量的机会很少, 特别是由于降水极端值远离坐标原点, 使得逐日降水量的这种总体上的 Γ 分布呈现出一种拖尾状或长尾状的分布曲线。因此, 降水状态的划分应当以递增区间由小到大排列为宜, 即对于小量的降水区间应该取得小一些, 随着降水量的增大, 日降水量等级区间可逐步扩大, 这种不等距划分区间的结果势必会形成一种几何级数的区间分布, 其优点在于可使落入各区间的频数不致于差异太悬殊, 从而有利于构建均匀 Markov 链。为此, 本文采用近似的几何级数作为划分湿状态量级的界限标准, 尽管各月降水量平均状况有差异, 但基本上都符合几何级数的等级区间。为了计算方便, 对所划等级区间也可适当作些调整。此外, 为了考虑总体概率分布的正偏态长尾状特点, 假设状态 s_0 为两点分布, $s_1, s_2 \cdots s_{n-1}$ 为均匀分布, 而 s_n 则定义为位移指数分布, 这在一定程度上都与实测资料相符。之所以考虑上述假设的根据在于, 日降水量符合长尾状的 Γ 分布, 即使在夏季也基本上呈 L 型分布, 其特点是正偏长尾状。因此, 除 s_0 和 s_n 以外, 其余各状态所对应的概率分布密度直方图显示出较为均匀的频数分布特点, 故假定为均匀分布型, 而 s_n 则取指数函数逼近其概率密度。

北京逐日降水量实测频率分布大体上为正偏长尾状曲线(图 1), 其直方图大致反映出上述规律(图 1—图 3)。处于 $s_1, s_2 \cdots s_{n-1}$ 状态和最后的第 n 状态 s_n 基本上符合均匀分布和指数分布型。其具

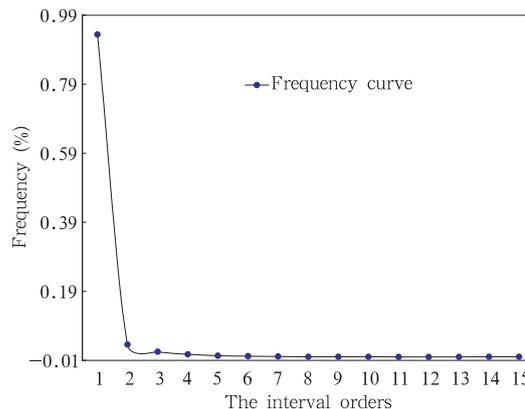


图 1 北京逐日降水量频率分布曲线

Fig. 1 The frequency distribution curve for all daily precipitation in Beijing

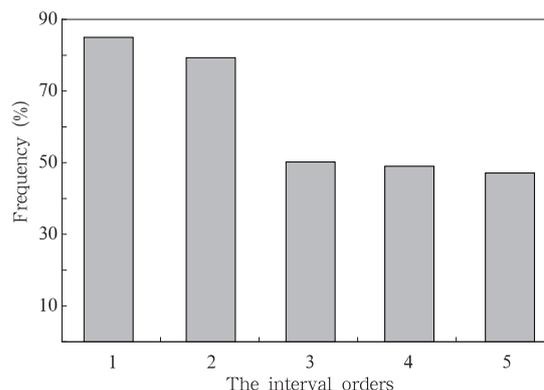


图 2 北京第 4 状态降水量直方图

Fig. 2 The histogram of state s_4 for daily precipitation in Beijing

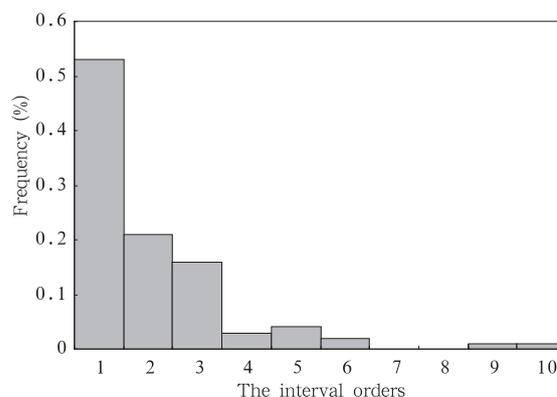


图 3 北京第 n 状态降水量直方图

Fig. 3 The histogram of state s_n for daily precipitation in Beijing

体理论模型为:

s_0 状态的概率分布为

$$p_0 = \begin{cases} P(X = 0.0) = q_{0.0} & \text{无降水} \\ P(X = 0.1) = q_{0.1} & \text{微量降水} \end{cases} \quad (2)$$

s_1 至 s_{n-1} 状态近似服从均匀分布密度

$$f_i(x) = \frac{1}{m_i - m_{i-1}} \quad m_{i-1} < x \leq m_i$$

$$i = 1, 2, \dots, n-1 \quad (3)$$

式中, m_i 为 $x \in s_i$ 的上限。

而 s_n 状态则近似服从下列指数分布

$$F_n(x) = \lambda e^{-\lambda(x-c)} \quad x \in s_n \quad (4)$$

其中 $F_n(x)$ 又称为位移指数分布函数, 式中 λ 为分布参数, 而 c 为状态 s_{n-1} 的上限。

根据给定的历史资料首先可估计 $s_0, s_1 \dots s_n$ 的转移概率矩阵, 求得式(1)。然后利用这些历史资料分别估计各状态 s_i 所对应的概率分布。显然, 实际上只要知道各状态区间界限, 均匀分布即可确定, 而指数分布则主要估计参数 λ 。在上述基础上, 利用离散随机变量的模拟方法(徐钟济, 1985)即可产生模拟的逐日降水记录。

根据均匀 Markov 链的平稳转移循环公式即 Chapman-Kolmogoroff 方程(么枕生, 1966), 当给定初始状态及其概率向量 $\mathbf{p}(0)$ 时可推求任何 k 步转移概率矩阵, 因此, 结合式(1), 我们有

$$\mathbf{P}^k = \mathbf{P}\mathbf{P}^{k-1} = \begin{pmatrix} \mathbf{p}_{00}^{(k)} & \mathbf{p}_{01}^{(k)} & \dots & \mathbf{p}_{0n}^{(k)} \\ \mathbf{p}_{10}^{(k)} & \mathbf{p}_{11}^{(k)} & \dots & \mathbf{p}_{1n}^{(k)} \\ \dots & \dots & \dots & \dots \\ \mathbf{p}_{n0}^{(k)} & \mathbf{p}_{n1}^{(k)} & \dots & \mathbf{p}_{nn}^{(k)} \end{pmatrix} \quad (5)$$

$$k = 1, 2, \dots$$

将行向量记为 $\mathbf{p}_j^{(k)} = (\mathbf{p}_{j0}^{(k)}, \mathbf{p}_{j1}^{(k)} \dots \mathbf{p}_{jn}^{(k)})$, 若已知初始向量 $\mathbf{p}(0) = (\mathbf{p}_0(0), \mathbf{p}_1(0) \dots \mathbf{p}_n(0))$, 则逐日降水过程在 k 步以后转移到各个状态 $s_0, s_1 \dots s_n$ 的概率向量必为

$$\mathbf{p}(k) = \mathbf{p}(0)\mathbf{p}(0, k) = \mathbf{p}(0)\mathbf{p}^k \quad k = 1, 2, \dots \quad (6)$$

其中 $\mathbf{p}(0, k)$ 表示从时刻为 0 到时刻为 k 的转移概率矩阵。或者写为

$$\mathbf{p}_j^{(k)} = \mathbf{p}(0)\mathbf{p}^k \quad (7)$$

其中 $\mathbf{p}_j^{(k)}$ 就是式(5)中的行向量, 而 $\mathbf{p}(0)$ 则是初始向量, 并有 $\mathbf{p}_j(0) = 1$, 而其余为 0, j 可为 $0, 1 \dots n$ 中任一序号。例如初始状态为 s_1 , 则有 $\mathbf{p}(0) = (0, 1, 0 \dots 0)$ 。

在模拟计算中, 笔者根据当地气候季节的年内变化将一年的逐日演变划分成若干阶段, 以便消除

季节对模拟结果的影响。例如, 采用按月计算的方案, 即对各月分别用上述模型产生模拟记录。由式(1)~(7), 原则上可采用两种算法确定逐日状态。

(1) 对历史观测记录, 在干湿多状态划分的基础上计算一阶转移矩阵 \mathbf{P} , 根据给定的初始日状态 $s_i (i=0, 1, 2 \dots n)$, 本文中的初始状态是统计了该日多年的状态频率分布之后, 确认初始状态为概率最大的状态, 利用式(7)并据概率行向量中的最大转移概率确定下一日(最可能)状态 $s_j (j=0, 1, 2 \dots n)$, 又由该日状态 s_j 为新的初始日状态, 重复利用一阶转移矩阵 \mathbf{P} 和式(7)推求再下一日的状态, 如此重复运算, 逐次求得全部各日所处状态。

(2) 对历史观测记录计算一阶转移矩阵 \mathbf{P} , 根据给定的初始概率向量, 利用式(6)或(7), 推求任意 k 步的概率向量, 从而一次确定全部各日所处(最可能)状态, 其中 $k=1, 2 \dots K$ 。

在推得逐日状态的基础上, 利用两点分布、均匀分布和位移指数分布产生各状态相应的随机数, 从而得到逐日模拟记录。

为了使模拟记录更加符合实测记录, 对初始状态的选择应考虑其代表性和客观性。统计样本资料中初始状态的频数分布, 例如某站初始日状态频率分布为单峰铃形分布, 则模拟程序中对初始日的选取即可按这种分布产生初始日状态。

3 模拟试验实例

3.1 资料

(1) 采用齐齐哈尔、北京、南京、郑州、武汉、广州等 6 个代表站 7 月份(1961—1990 年)的逐日降水量资料进行模拟试验, 首先, 确定干湿状态数以及区间的划分, 其次, 估计各月逐日转移概率矩阵 \mathbf{P} 及位移指数分布参数 λ 。状态划分结果表明, 当 n 从 6 增至 15, 并以等差级数和几何级数不同分级区间进行随机模拟效果比较后, 发现取状态数 8 或 9, 用几何级数作分级区间界限最为理想, 其分级区间界限可用下列经验公式

$$m_h = \frac{(x_{m1} + x_{m2})}{2^{n-1}} 2^{(h-2)} \quad (8)$$

式中, m_h 为分级区间上界, $h=2, 3 \dots n-1$, 而 $h=1$ 为日降水量不大于 0.1 mm 的特定状态, 即 s_0 状态, x_{m1} 和 x_{m2} 分别为日降水量的样本极大值和次极大值, n 为分界点数。换言之, s_1, s_2 直到 s_{n-1} 的上界

均可用上述经验公式来估计。表 1 列出了各代表站 7 月逐日降水量状态划分标准。为了获得稳定的降

水模拟记录,我们对齐齐哈尔等 6 个代表站分别重复产生 5 次模拟记录,然后求其平均。

表 1 7 月各代表站逐日降水量状态划分标准(单位:mm)

Table 1 The criterions of each state for daily rainfall of the stations during July

状态	s_0	s_1	s_2	s_3	s_4	s_5	s_6	s_7
齐齐哈尔	≤ 0.1	≤ 1.17	≤ 2.35	≤ 4.69	≤ 9.38	≤ 18.78	≤ 37.55	> 37.55
北京	≤ 0.1	≤ 1.66	≤ 3.32	≤ 6.64	≤ 13.29	≤ 26.58	≤ 53.15	> 53.15
郑州	≤ 0.1	≤ 2.50	≤ 5.01	≤ 10.02	≤ 20.04	≤ 40.07	≤ 80.15	> 80.15
南京	≤ 0.1	≤ 2.72	≤ 5.45	≤ 10.89	≤ 21.79	≤ 43.58	≤ 87.15	> 87.15
武汉	≤ 0.1	≤ 2.76	≤ 5.53	≤ 11.06	≤ 22.11	≤ 44.22	≤ 88.45	> 88.45
广州	≤ 0.1	≤ 1.54	≤ 3.07	≤ 6.15	≤ 12.29	≤ 24.59	≤ 49.17	> 49.17

3.2 与两状态 Markov 链模式的对比试验

若将降水天气随机发生器以两状态 Markov 链为基础。其模拟可分为两步:①确定干湿日出现的概率。根据输入的气候数据,该模型就可以确定某天的干湿(有雨或者无雨)情况,如果无降水则数值为 0。如果有降水则转到第 2 步。②模拟降水量。用来描述逐日降水的概率分布模式主要有偏态分布和负指数分布。基于两状态 Markov 链模型的降水发生器其主要缺点就是对雨日的模拟处理过于简单化,尤其在多雨地区,因降水有强弱不同,而降水日

数又往往较多,可能会出现较大的误差。本文用两状态 Markov 链产生的模拟结果与多状态 Markov 链产生的模拟结果作了比较,发现尽管两种模型的模拟记录与实测记录在各个降水量等级都有一致性,不同模型对不同区域的降水虽都有一定的模拟能力,但是对逐日强降水的模拟却有较大差异,除个别测站(如齐齐哈尔、北京等少雨地区)外,多状态 Markov 链模式的模拟效果比两状态 Markov 链的模拟效果好。这就表明多状态模式对极端降水有更好的模拟效果。

表 2 主要降水气候指标的模拟与实测结果对比(括号内为相对误差)

Table 2 Comperation between modeling and observation for main climatic indices of precipitation

站名	平均日降水量			日降水极大值			月平均降水日数		
	实测	多状态	两状态	实测	多状态	两状态	实测	多状态	两状态
齐齐哈尔	4.51	8.13(0.80)	7.42(0.64)	77.8	58.3(0.25)	96.0(0.23)	10.3	9.3(0.10)	11.1(0.08)
北京	6.55	10.1(0.54)	6.77(0.03)	84.9	71.4(0.16)	93.4(0.10)	12.6	13.9(0.10)	10.1(0.20)
郑州	4.86	11.6(1.38)	5.37(0.10)	100	83.0(0.17)	140.(0.40)	9.7	11.4(0.18)	7.8(0.20)
南京	5.69	8.68(0.52)	6.94(0.22)	114.	124.0(0.08)	102.(0.10)	9.9	12.3(0.24)	7.6(0.23)
武汉	9.6	11.5(0.19)	6.35(0.34)	222.	198.(0.11)	101(0.54)	10.6	13.7(0.29)	7.1(0.33)
广州	9.02	12.2(0.35)	7.42(0.18)	75.1	72.8(0.03)	96.0(0.28)	17.6	16.6(0.05)	10.4(0.43)

从表 2 可见,平均日降水量的模拟效果差异并不明显,但是,日极大降水量及其相对误差和月平均雨日数相对误差大都以多状态模式效果较好,平均相对误差都比两状态模式小。究其原因,可能是由于多状态模式对降水量的详细描述,使得所模拟的雨日状况的细节有所改善。为了更加说明这一特点,进一步采用中国东部地区 78 个测站的逐日降水资料借助于上述两种模式模拟逐日降水量,基于上述两种模式进一步针对极端降水量作 GPD 拟合,结果表明,所模拟的夏季(7 月)逐日极端降水量序列,一般都具有较高的精度。对所模拟的资料统计

其各种气候特征量表明,绝大多数测站都达到较高的精度,其中多状态模式对中国东部极端降水特征的模拟能力基本上优于两状态模式,平均相对误差一般都小于两状态模式。实况与模式模拟的结果在一定精度上非常相似,多状态和两状态的模拟方法都基本能模拟出降水量的各种特征。其中日降水均方差的模拟效果最好。但是多状态 Markov 链模式模拟的日降水极大值、月平均降水日数明显优于两状态 Markov 链的效果,特别是对中国多雨地区的模拟。说明多状态的 Markov 链更适用于对极端降水的模拟。图 4 中绘出了 30 个测站,多状态

Markov 链模式模拟的平均极大日雨量相对误差比较图。表 3 列出了模拟降水极大值相对误差小于

15% 的大部分站点名称,它们约占总站数(78 站)的 40%左右。

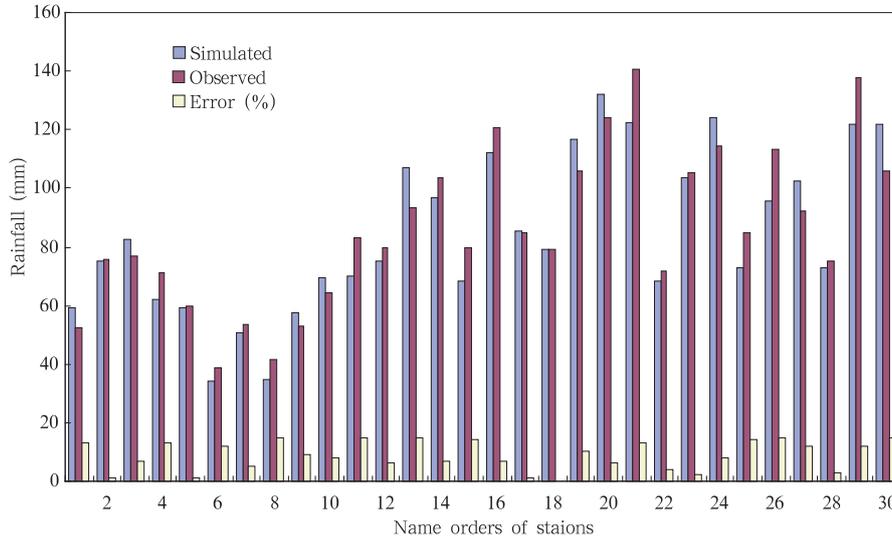


图 4 多状态 Markov 链模拟与实测的平均极大日雨量误差比较

Fig. 4 The mean error of extreme daily rainfalls between simulated from the multi-status Markov chain model and observed data

表 3 相对误差 < 15% 的站名序号表

Table 3 The names and its ordering for each station (relative error)

站名	序号	站名	序号	站名	序号
嫩江 (Nenjiang)	1	牡丹江 (Modanjiang)	11	莒县 (Juxian)	21
克山 (Keshan)	2	通辽 (Tonliao)	12	信阳 (Xinyang)	22
齐齐哈尔 (Qiqihaer)	3	四平 (Siping)	13	恩施 (Enshi)	23
佳木斯 (Jiamusi)	4	长春 (Changchun)	14	衡阳 (Hengyang)	24
白城 (Baicheng)	5	鞍山 (Anshan)	15	淮阴 (Huangyin)	25
靖远 (Jinyuan)	6	沈阳 (Shenyang)	16	南京 (Nanjing)	26
华家岭 (Huajialing)	7	承德 (Chengde)	17	浦城 (Pucheng)	27
银川 (Yinchuan)	8	兴城 (Xingcheng)	18	柳州 (Liuzhou)	28
榆林 (Yulin)	9	北京 (Beijing)	19	百色 (Baise)	29
西峰镇 (Xifengzhen)	10	潍坊 (Wefang)	20	广州 (Guangzhou)	30

4 多状态模拟降水数据的极值分布模式及其概率特征

对于 Markov 链模式模拟逐日降水的总效果, 虽已作了充分的验证, 结果表明, 利用该模式可以较好地模拟出日降水量, 但我们更加关注的是, 该模式对于极端降水量的模拟效果究竟如何? 在文献 (Ding, et al, 2008) 的研究中已经发现, 广义 Pareto 分布 (简记为 GPD) 对极端降水有较好的拟合效果, 相比之下, 优于 GEV 和 Gumbel 分布。假如应用由多状态 Markov 链模式生成的模拟日降水量序列抽

取极值, 是否完全符合 GPD? 为此, 我们验证 Markov 链对极端降水的模拟效果。

仍以齐齐哈尔、北京、南京、郑州、武汉、广州等 6 个代表站 7 月份 (1961—1990 年) 30 a 的逐日降水量为例, 并用模拟资料和实测资料所拟合的 GPD 特征作对比, 如前所采用的柯尔莫哥洛夫检验、相关系数、均方差等 3 种检测指标, 对模拟和实测结果作对比分析 (如表 4), 由表 4 可见, 多状态 Markov 链所模拟的日降水量完全符合 GPD 模式, 并具有较高的拟合优度。经过多状态 Markov 模拟结果所得的 GPD 均能通过柯尔莫哥洛夫检验, 这就充分说明了

模拟和实际观测的 GPD 特征量十分相近。这一试验的进一步意义和价值在于,一旦我们拥有全球数值模拟的输出信息,经过降尺度技术处理就能取得某一区域或局地平均气候信息,在此基础上借助于

平均气候与极端气候的关系或者直接利用 Markov 链模拟逐日气候资料序列,就可推求 GPD 的特征量及其极端值概率和分位数,从而预估未来气候极值的一系列特征。

表 4 多状态 Markov 链模拟的日降水量 GPD 拟合效果检验
Table 4 The efficiency test for fitting GPD of the simulated extreme daily rainfalls from the multi-status Markov chain model

效果检验		齐齐哈尔	北京	郑州	南京	武汉	广州
柯氏检验	模拟值	0.11	0.12	0.12	0.14	0.09	0.12
	实测值	0.14	0.07	0.12	0.08	0.08	0.08
相关系数	模拟值	0.97	0.92	0.97	0.98	0.98	0.96
	实测值	0.94	0.99	0.98	0.99	0.99	0.99
均方差	模拟值	0.01	0.03	0.02	0.01	0.01	0.02
	实测值	0.02	0.01	0.01	0.01	0.01	0.01

表 5 Markov 链模拟的日降水量与观测数据的 GPD 拟合效果比较
Table 5 The comparison of fitting GPD effects between simulated from the Markov chain model and observation data

站名		齐齐哈尔	北京	郑州	南京	武汉	广州	平均百分误差
门限值	模拟值	38.0(19%)	52.0(16%)	44.0(22%)	54.0(17%)	48.0(17%)	48.0(14%)	17%
	实测值	32.0	45.0	36.0	45.0	41.0	42.0	
50 a 重现期极值	模拟值	63.9(5%)	99.7(3%)	91.2(23%)	178.3(11%)	146.1(1%)	134.8(7%)	8.3%
	实测值	60.9	96.9	118.2	160.5	144.6	126.4	
100 a 重现期极值	模拟值	69.9(10%)	107.1(1%)	117.8(11%)	196.3(8%)	171.6(5%)	149.6(6%)	6.8%
	实测值	63.3	106.3	133.0	181.1	163.1	141.0	

* 括号内数字为模拟结果的相对误差。

此外,选取模拟和实况数据的门限值及 50 和 100 a 重现期值来分析两者(模拟和实况)拟合 GPD 的情况(表 5),结果发现,模拟值和实测值的 50 和 100 a 重现期值非常相似,但模拟结果普遍高于实测值的结果,从相对百分误差可以看出其差别很小。而门限值的对比结果表明,模拟资料的门限值高于实际资料的门限值,并存在一定的误差。这也表明,模拟和实测门限值的误差越小,重现期极值的差距也越小。说明多状态 Markov 链模式模拟的极端降水与实况非常相似,即该方法能较好地模拟中国的极端降水。

5 结 论

(1) 利用多状态一阶 Markov 链产生逐日降水序列的模式,对中国东部 6 个代表站进行模拟试验,其结果表明,月降水均方差、日降水极大值、月平均降水日数,日降水均方差、月平均降水日数、日平均降水量等指标与实况比较,证明该模拟方法对未来

的逐日降水量的模拟效果较好,能基本模拟出降水量的各种特征。

(2) 对中国东部 78 个代表站采用的两种模式模拟结果对比发现,除日平均降水量以外,月平均降水日数、日降水平均极大值都与实际观测结果较为一致,且总体上优于两状态 Markov 模式的模拟结果,说明用该模式在中国范围内模拟逐日降水特征尤其是极端降水特征有较高的可行性。

(3) 多状态 Markov 链模式,对所选取的 6 个代表站逐日降水量模拟资料所拟合的极端降水 GPD 模式具有较高的拟合优度。无论从门限值或是 50 和 100 a 重现期值来看都可发现模拟结果与实测结果有较好的相似性,且模拟和实测结果的门限值误差越小,重现期极值的差距也越小。证明用 Markov 链模式对极端降水的模拟有广泛的适用性。

References

Ding Yuguo, Zhang Yaocun. 1989. A stochastic simulation test for

- climatological features of precipitation. *J Nanjing Inst Meteor (in Chinese)*, 12(2):146-154
- Ding Yuguo, Niu Tao. 1990. A Markov chain simulation for dry and wet month runs. *J Nanjing Inst Meteor (in Chinese)*, 13(3): 286-297
- Ding Yuguo. 1994. Research of universality for Γ distribution model of precipitation. *Scientia Atmospherica Sinica (in Chinese)*, 18 (5):552-560
- Ding Yuguo, Cheng Bingyan, Jiang Zhihong. 2008. A newly-discovered GPD-GEV relationship together with comparing their modelings of extreme precipitation in summer. *Adv Atmos Sci*, 25(3):507-516
- Gabriel K R, Neurpann J. 1962. A Markov chain model for daily rainfall occurrence at Telaviv. *Quart J Roy Met Soc*, 188:90-95
- Gates, P Tong H. 1976. On Markov chain modeling to some weather data. *J Appl Meteor*, 15: 1145-1151
- Gregory J M, Wigley T M L, Jones P D. 1993. Application of Markov models to area-average daily precipitation series and interannual variability in seasonal totals. *Climate Dynamics*, 8: 299-310
- Katz R W. 1974. Computing probabilities associated with Markov chain model for precipitation. *J Appl Meteor*, 13:953-954
- Liao Yaoming, Zhang Qiang, Chen Deliang. 2004. The precipitation simulation based on using a stochastic weather generator cover China region. *J Geography (in Chinese)*, 59 (5):698
- Palutikof J P, Goodess C M, Watkins S J, et al. 2002. Generating rainfall and temperature scenarios at multiple sites: examples from the mediterranean. *J Climate*, 15: 3529-3548
- Wang Zonghao, Li Maicun. 1974. *The Methods of Probability Statistics in Weather Forecasting*. Beijing: Siceince Press, 116-128
- Wilby R L, Wigley T M L. 1997. Downscaling general circulation model output: a review of methods and limitations. *Progress Physi Geography*, 21:530-548
- Wilks D S. 1999. Multisite downscaling of daily precipitation with a stochastic weather generator. *Climate Res*, 11: 125-136
- Wilks D S, Wilby R L. 1999. The weather generation game: A review of stochastic weather models. *Progress Physi Geography*, 23: 329-357
- Xu Zhongji. 1985. *Monte-Carlo Method (in Chinese)*. Shanghai: Shanghai Scientific & Technical Publishers, 324pp
- Yao C S. 1963. Some applications of the fundamental laws in the theory of probability to the cliamatology. *Acta Meteor Sinica (in Chinese)*, 33(2), 245-255
- Yao C S. 1966. The probabilities of stochastic change between wet and dry day. *Acta Meteor Sinica (in Chinese)*, 36(2):249-256
- Yao Zhensheng, Ding Yuguo. 1990. *Climatic Statistics(in Chinese)*. Beijing: China Meteorology Press, 698-721
- Zhang Yaocun. 1990a. A stochastic distribution model for N daily rainfall. *J Nanjing Inst Meteor (in Chinese)*, 13(1):23-30
- Zhang Yaocun, Ding Yuguo. 1990b. Statistical characteristics of daily rainfall series at five representative stations in Eastern China. *J Nanjing Inst Meteor (in Chinese)*, 13(2):194-203

附中文参考文献

- 丁裕国, 张耀存. 1989. 降水气候特征的随机模拟试验. *南京气象学院学报*, 12(2):146-154
- 丁裕国, 牛涛. 1990. 干、湿月游程的 Markov 链模拟. *南京气象学院学报*, 13(3):286-297
- 丁裕国. 1994. 降水量分布模式的普适性研究. *大气科学*, 18(5):552-560
- 廖要明, 张强, 陈德亮. 2004. 中国天气发生器的降水模拟. *地理学报*, 59(5):698
- 王宗皓, 李麦村. 1974. 天气预报中的概率统计方法. 北京: 科学出版社, 116-128
- 徐钟济. 1985. 蒙特卡罗方法. 上海: 上海科学技术出版社, 324pp
- 么枕生. 1963. 概率论基本定理在气候统计中的应用. *气象学报*, 33 (2):245-255
- 么枕生. 1966. 湿日与干日随机变化的概率. *气象学报*, 36(2):249-256
- 么枕生, 丁裕国. 1990. 气候统计. 北京: 气象出版社, 698-721
- 张耀存. 1990a. N 日降水量的随机分布模式. *南京气象学院学报*, 13 (1):23-30
- 张耀存, 丁裕国. 1990b. 我国东部地区几个代表测站逐日降水序列统计分布特征. *南京气象学院学报*, 13(2):194-203