

# 干、湿月游程的Markov链模拟

丁裕国 牛涛\*\*

## 提 要

本文以文献[1]中提出的游程和转折点的概念,利用Markov链研究我国各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。此外,文中还对干、湿月转移概率矩阵的稳定性及其极限分布作了计算和讨论,其结果可直接提供长期预报的气候背景。

## 一、引 言

长期预报或气候预报的一个至关重要的问题就是,一种气候类型或状态(或天气状态)在何时发生转折,从一种状态转至另一种状态有多长的持续时间。例如干、湿时期的持续与转折,某种环流模态的持续与转折等等,都是气候预报或长期预报的关键。所谓游程,是指在离散(时间)序列中,同类性质的元素持续出现所占据的一个阶段,例如,某气候要素持续出现一段时间的正距平可称为正游程,其持续时间长度就称为正游程长度;反之,负距平持续时间长度就称为负游程长度。显然,在不同性质的游程之间必有转折点(在时间轴上即为转折时间)。游程理论最早由Mood(1940)创立<sup>[2]</sup>,其后Wald等人, Feller(1957)<sup>[3]</sup>, Yevjevich(1972)<sup>[4]</sup>, Hunter(1983)<sup>[5]</sup>都曾做过许多研究。杨鉴初早在1951年就曾从实践预报经验总结出有关气象要素历史曲线演变的规律性,指出在长期气候记录中存在着循环过程,在循环过程中必有转折点<sup>[6]</sup>。么枕生早就注意到长期气候预报的统计学理论,对于游程转折点规律做过许多研究。在文献[7]中,曾将马尔科夫链理论结合到游程计算中。又从不同的角度推导出干、湿游程转折点间的等待时间 $T_i$ 所具有的概率分布、游程转折时间的平均数和方差、转折周期(干湿循环)长度的理论公式<sup>[8-11]</sup>。对于统计气候预报来说,研究气候序列的游程和转折点统计规律并寻求其统计预报模式具有理论和应用价值。

从统计学观点来看;若将记录的数值时间序列按某一临界值划分为性质不同的游

1989年1月24日收到,5月9日收到修改稿

\* 本文是国家自然科学基金资助课题“中国统计气候研究”项目的论文之一

\*\* 我院1988届本科毕业生,现为国家气象局气科院硕士生

程二分序列(如正、负距平序列),则可使序列内部的非线性变化所造成的影响减小<sup>[1]</sup>。对这种简化时间序列(即游程序列)不但可以建立线性自回归(AR(p))模式,也可建立各种非线性自回归模式。同时,研究游程长度的概率分布,游程转折周期的概率分布以及研究不同环流背景或前期天气气候条件下的游程随时间的转折规律,无疑也是非常有意义的。

本文将Markov链理论应用于游程转折点研究中,对我国不同气候区的主要代表性测站计算干、湿月游程概率及其分布,利用经验拟合选配各站游程概率分布模式,进一步讨论了各地转移概率矩阵的极限分布和稳定性规律。

## 二、资 料 和 方 法

为了全面了解我国各大气候区域逐月干、湿演变的统计规律,选取哈尔滨、北京、宜昌、南通、上海、杭州、成都、兰州、银川、南宁等十个代表测站(1951—1980年)近30年逐月降水资料,其中上海站还另外选取1881—1950年逐月降水资料作为研究长短样本的对比资料。所以选取有限的代表站,一是为了研究本方法的可行性,二是为了减少计算量。因此,本文不作有关统计参数在地理空间分布的探讨,而仅针对各站本身或各站之间统计规律的对比分析。

首先,对降水记录作如下的预处理:凡月降水量大于该站该月累年平均值,定为湿月,记为“1”;凡月降水量小于、等于该站该月累年平均值,定为干月,记为“0”。这样就得到以“0~1”标记的干、湿月游程序列。如图1绘出十个代表站干、湿月气候概率在不同地区的分布。由图可见,湿润气候区湿月气候概率大于干旱气候区,而干旱气候区不但干月气候概率较大,且干湿月气候概率差值也较大。

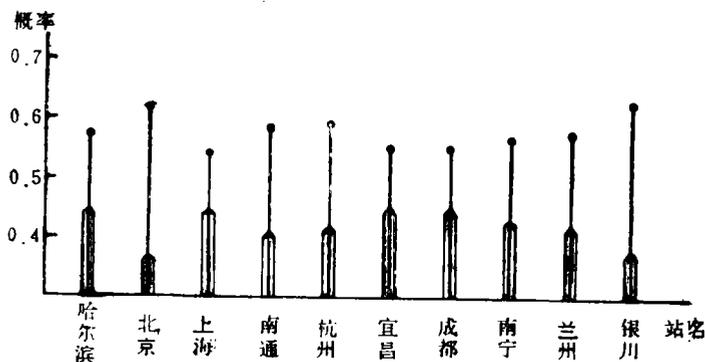


图1 各站干、湿月气候概率  
(图中单实线表示干月,双实线表示湿月)

从Markov链的观点来看,游程长度就是游程转折点之间的等待时间,即所谓首次通过时间,因此,游程长度的分布就是首次通过时间的分布。就两状态一阶Markov链而言,长度为k的干、湿月游程出现概率就是从干、湿月出发经过k步转移,首次到达湿、干月的概率,因而有

$$\left. \begin{aligned} P\{D_r=k\} &= (1 - P_{d|d}) P_{d|d}^{k-1} \\ P\{W_r=k\} &= (1 - P_{w|w}) P_{w|w}^{k-1} \end{aligned} \right\} \quad (1)$$

其中 $D_r$ 和 $W_r$ 分别表示干月和湿月游程长度,  $P_{d|d}$ 表示前1月为干月, 后1月仍为干月的条件概率,  $P_{w|w}$ 表示前1月为湿月, 后1月仍为湿月的条件概率。根据[10]的研究, 利用高阶Markov链, 也可推得长度为 $k$ 的干、湿月游程概率分布, 它们分别为\*

$$\left. \begin{aligned} P\{D_r=k\} &= (1 - P_{d|dd}) P_{d|wd} P_{d|dd}^{k-2} \\ P\{W_r=k\} &= (1 - P_{w|ww}) P_{w|dw} P_{w|ww}^{k-2} \end{aligned} \right\} \quad k \geq 2 \quad (2)$$

$$\left. \begin{aligned} P\{D_r=k\} &= (1 - P_{d|ddd}) P_{d|wd} P_{d|wdd} P_{d|ddd}^{k-3} \\ P\{W_r=k\} &= (1 - P_{w|www}) P_{w|dw} P_{w|dww} P_{w|www}^{k-3} \end{aligned} \right\} \quad k \geq 3 \quad (3)$$

.....

$$\left. \begin{aligned} P\{D_r=k\} &= (1 - P_{d|d\dots d}) P_{d|wd} P_{d|wdd} \dots P_{d|d\dots d}^{k-s} \\ P\{W_r=k\} &= (1 - P_{w|w\dots w}) P_{w|dw} P_{w|dww} \dots P_{w|w\dots w}^{k-s} \end{aligned} \right\} \quad k \geq s \quad (4)$$

式(2)、(3)、(4)分别代表在2、3、...、 $s$ 阶Markov链意义下, 干、湿月游程的理论分布模式。不难看出, 它们都是(1)式的推广, 而(1)式正是文献[9]和[3]从不同观点推导的一致结果。在上述公式中概率 $P_{d|dd}$ 表示前2月为干月, 后1月仍为干月的条件概率, 而 $P_{d|wd}$ 表示初始月为湿月, 第2月为干月, 第3月又为干月的概率, 类似地, 对于湿月也有相应条件概率 $P_{w|ww}$ 、 $P_{w|dw}$ ...等等, 其余类推之。

任何气候要素的全部统计特征都可通过概率分布来描述, 而其主要统计特征如平均值和方差, 则常常是人们最关心的, 它们在一定的程度上代表了一地气候的主要特征。游程的平均长度、平均返回时间、天气气候循环的平均长度正是从不同侧面来描述干、湿月转折的平均状态, 而这三者在概念的本质上一致, 因此, 可借助文献[1]推导的相应公式计算各地干、湿月游程的这类统计参数。其公式为

$$\left. \begin{aligned} L_d &= \frac{1}{1 - P_{d|d}} \\ L_w &= \frac{1}{1 - P_{w|w}} \end{aligned} \right\} \quad (5)$$

$$\left. \begin{aligned} \sigma_d^2 &= \frac{P_{d|d}}{(1 - P_{d|d})^2} \\ \sigma_w^2 &= \frac{P_{w|w}}{(1 - P_{w|w})^2} \end{aligned} \right\} \quad (6)$$

\* 文献[10]中未给出推导过程, 但公式有误, 作者在本文除引用其思路外, 重新推导出(2)(3)(4)式

$$\left. \begin{aligned} T_d &= 1 + \frac{P_w | d}{P_d | w} \\ T_w &= 1 + \frac{P_d | w}{P_w | d} \end{aligned} \right\} \quad (7)$$

$$\left. \begin{aligned} \mu_c = E[c] &= \frac{1}{1 - P_d | d} + \frac{1}{1 - P_w | w} \\ \sigma_c^2 = \text{Var}[c] &= \frac{P_d | d}{(1 - P_d | d)^2} + \frac{P_w | w}{(1 - P_w | w)^2} \end{aligned} \right\} \quad (8)$$

式(5)、(6)分别为干、湿月游程平均长度和方差的计算公式，式(7)、(8)分别为干、湿月平均返回时间的计算公式和天气气候循环平均长度和方差计算公式。其中 $L_d$ 、 $L_w$ 分别为干月和湿月游程平均长度， $\sigma_d^2$ 、 $\sigma_w^2$ 分别为干月游程和湿月游程长度的方差， $T_d$ 和 $T_w$ 分别为干月和湿月的平均返回时间， $\mu_c$ 和 $\sigma_c^2$ 则为天气气候循环平均长度和其方差。

### 三、计算结果

#### (一)干、湿月游程概率分布

对哈尔滨、北京、宜昌、上海、杭州、成都、兰州、银川、南宁等十个站分别应用(1)—(4)式，计算在1—4阶Markov链假设下的干、湿月游程概率(理论值)与干、湿月游程频数(实测值)。表1中给出几个站(1—4阶链假设下)游程长度的理论频数与实测频数分布。由表可见，各地游程长度的理论分布与实测分布基本吻合，其特点是大多

表1 a 干月游程长度频数分布表

单位：月

项 目	游程长度 (月)	游程长度															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
北京	一阶	31.9	19.8	12.3	7.6	4.7	2.9	1.8	1.1	0.7	0.4	0.3	0.2	0.1	0.0	0.0	0.0
	二阶	31.9	17.8	11.4	7.3	4.7	3.0	1.9	1.2	0.8	0.5	0.3	0.2	0.1	0.0	0.0	0.0
	三阶	31.9	16.8	12.3	7.4	4.8	3.1	1.8	1.1	0.5	0.4	0.2	0.3	0.1	0.0	0.1	0.0
	四阶	31.9	17.8	12.1	8.5	5.2	3.1	1.9	1.2	0.7	0.4	0.3	0.2	0.1	0.0	0.0	0.0
	实测	34.0	17.0	11.0	8.0	4.0	4.0	4.0	4.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
哈尔滨	一阶	35.9	20.9	12.2	7.1	4.1	2.4	1.4	0.8	0.5	0.3	0.2	0.1	0.1	0.0	0.0	0.0
	二阶	35.9	22.5	12.8	7.2	4.1	2.3	1.3	0.7	0.4	0.2	0.1	0.1	0.0	0.0	0.0	0.0
	三阶	35.9	21.0	11.9	7.3	3.6	2.2	1.5	0.8	0.5	0.4	0.1	0.1	0.0	0.0	0.0	0.0
	四阶	35.9	22.5	12.4	5.8	3.6	2.2	1.3	0.5	0.4	0.3	0.2	0.1	0.0	0.0	0.0	0.0
	实测	34.0	23.0	14.0	6.0	2.0	1.0	0.5	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0



续表 1b

项 目	游 程 长 度 (月)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
		银川	一阶	42.4	18.1	7.7	3.3	1.4	0.6	0.3	0.2	0.1	0.0	0.0	0.0	0.0	0.0
	二阶	42.4	15.9	7.4	3.4	1.6	0.7	0.3	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	三阶	42.4	15.9	7.4	3.3	1.7	0.7	0.3	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	四阶	42.4	15.9	7.4	3.3	1.7	0.7	0.3	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	实测	44.0	16.0	7.0	3.0	3.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
兰州	一阶	36.5	18.5	9.4	4.7	2.4	1.2	0.6	0.3	0.2	0.1	0.0	0.0	0.0	0.0	0.0	0.0
	二阶	36.5	22.4	10.2	4.6	2.1	1.1	0.0	0.5	0.2	0.1	0.0	0.0	0.0	0.0	0.0	0.0
	三阶	36.5	22.4	10.2	2.6	2.2	3.0	0.9	0.5	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	四阶	36.5	22.4	10.3	6.2	2.3	1.0	0.5	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	实测	32.0	23.0	9.0	7.0	1.0	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
宜昌	一阶	57.5	23.1	9.3	3.7	1.5	0.6	0.2	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	二阶	57.5	28.2	9.5	3.2	1.1	0.4	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	三阶	57.5	28.2	10.1	3.4	1.3	0.4	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	四阶	57.5	28.2	13.0	1.8	0.7	0.3	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	实测	53.0	26.0	14.0	2.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

例，我们仅取一阶链所拟合的频数点绘几个测站频率分布曲线(如图 2)。根据各站频率分布形状，还可配合如下的经验指数曲线，作为离散型概率分布的一种近似：

$$P_i(r) = k_i e^{\alpha_i r} \quad i = 1, \dots, 10$$

$$r = 1, 2, \dots$$

其中  $k_i, \alpha_i$  分别为第  $i$  站的两个待定参数，表 2 中列出各代表站参数值(根据最小二乘法拟合计算结果)。

表 2

站		哈尔滨	北京	上海	宜昌	南宁	成都	南通	兰州	银川
干	$k_i$	0.7159	0.6115	0.7455	0.9702	0.7647	1.4929	0.6608	0.5566	0.4902
月	$\alpha_i$	-0.5396	-0.4771	-0.5570	-0.6782	-0.5589	-0.9135	-0.5012	-0.4429	-0.3989
湿	$k_i$	1.2839	1.6998	0.9648	1.4929	0.6945	0.7817	1.1884	0.9738	1.3395
月	$\alpha_i$	-0.8259	-0.9932	-0.6755	-0.9135	-0.5043	-0.5574	-0.7833	-0.6799	-0.8499

(二)干、湿月游程平均长度及其方差

表 3 中已经列出各站干、湿月游程平均长度及其均方差, 这些结果由公式 (5) 和 (6) 算得。由表可见, 银川、兰州、杭州无论干或湿月游程长度都较长, 而宜昌、南宁

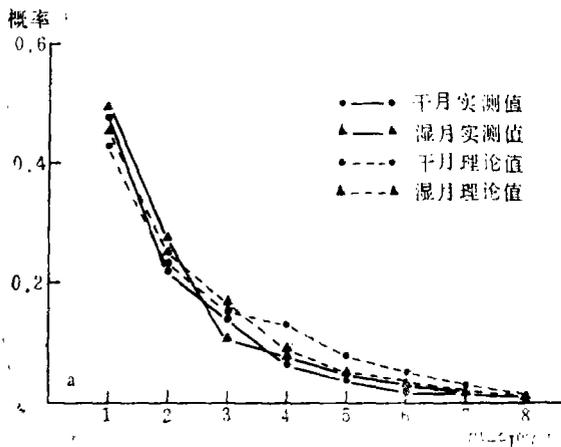


图 2 a 上海干、湿月游程概率分布曲线

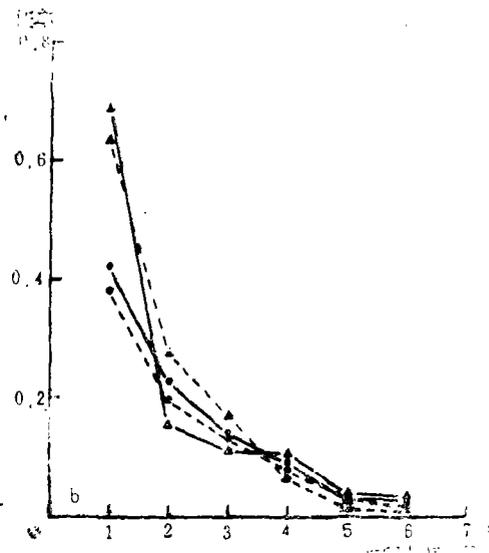


图 2 b 北京干、湿月游程概率分布曲线 (同例图 a)

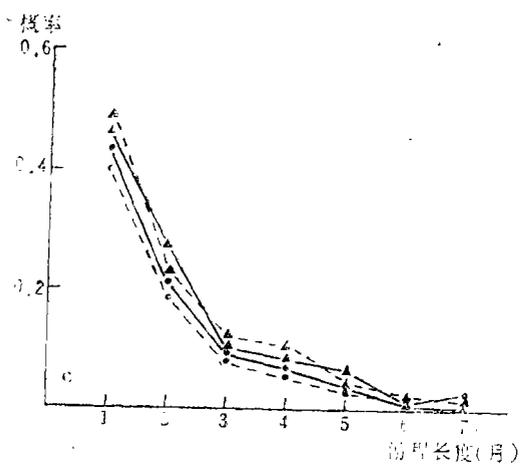


图 2 c 兰州干、湿月游程概率分布曲线(同例图 a)

则相对较短。

(三)天气气候循环平均长度与平均返回时间

在表 3 中, 还列出了由公式 (7) 和 (8) 计算得到的各站平均返回时间  $T_d$  或  $T_w$  和天气气候循环平均长度  $\mu_d$  (见表 3 下部)。这些结果与干、湿月游程平均长度  $L_d$  或  $L_w$  是互相匹配的。例如, 银川、兰州、杭州三站干、湿月游程相对较长, 其天气气候循环的平均长度也相应较长; 反之, 如宜昌、南宁干、湿月游程相对都较短, 其天气气候循环周期也较短。这与表 3 倒二行所列计算结果 ( $\mu_d$ ) 恰好相符。另一方面, 从表 4 给出的各站

天气气候循环频数分配也可找到同样的佐证。在表 4 中, 列出各站天气气候循环的理论频数和实测频数分配。由表可见, 宜昌、南宁两地天气气候循环长度集中在 3 个月以下, 换言之, 两地月际干湿交替大约在 3 个月以内循环的机会为最多。例如, 宜昌 4 个月以下的天气气候循环实测频率约占 72.7%, 南宁则占 62.5%; 而银川、兰州、杭州分别只有 36.7%、41.5%、45.7%。这种比例关系说明宜昌、南宁两地干湿游程平均较短, 而银川、兰州两地天气气候循环的干湿游程平均来说比上述两地稍长。再以表 3 中干月

表 3 干、湿月游程统计特征理论计算值 (单位: 月)

项 目 \ 站		北京	哈尔滨	银川	兰州	宜昌	成都	南通	上海	杭州	南宁
		干 月	$L_d$	2.64	2.40	3.04	2.80	2.03	2.28	2.54	2.34
	$\sigma_d$	2.08	1.83	2.49	2.24	1.45	1.71	1.98	1.77	2.18	1.72
	$T_d$	1.60	1.75	1.55	1.72	1.82	1.83	1.73	1.87	1.70	1.72
湿 月	$L_w$	1.59	1.78	1.75	2.03	1.67	1.90	1.84	2.04	1.92	1.72
	$\sigma_w$	0.97	1.18	1.14	1.44	1.06	1.30	1.25	1.45	1.33	1.15
	$T_w$	2.66	2.35	2.74	2.38	2.22	2.20	2.38	2.15	2.43	2.33
	$\mu_c$	4.22	4.18	4.79	4.82	3.70	4.18	4.38	4.38	4.66	4.01
	$\sigma_c$	2.29	2.17	2.74	2.67	1.79	2.14	2.34	2.29	2.56	2.05

表 4 天气气候循环长度的频数分布

站名与项目		循环长度 (月)													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
北 京	理论值	0.0	20.1	19.9	15.1	10.4	6.8	4.4	2.8	1.7	1.1	0.7	0.4	0.0	0.0
	实测值	0.0	22.4	15.4	15.9	10.4	7.5	3.5	5.0	0.5	3.0	0.5	0.0	0.0	0.0
银 川	理论值	0.0	13.9	15.3	12.8	9.7	7.0	4.9	3.4	2.3	1.6	1.0	0.7	0.5	0.3
	实测值	0.0	13.9	12.4	10.4	13.4	8.9	6.5	4.0	3.5	0.0	0.5	0.0	0.5	0.0
哈 尔 滨	理论值	0.0	19.9	20.4	15.1	10.8	7.0	4.4	2.7	1.6	1.0	0.6	0.3	0.0	0.1
	实测值	0.0	17.4	19.4	13.9	7.5	5.5	4.0	5.0	0.0	1.0	0.5	1.0	0.0	0.0
兰 州	理论值	0.0	13.1	15.0	13.0	10.0	7.3	5.1	3.5	2.4	1.6	1.0	0.7	0.4	0.3
	实测值	0.0	13.5	13.5	14.5	6.5	13.0	4.0	4.0	1.5	1.0	0.5	0.5	0.5	0.0
宜 昌	理论值	0.0	28.0	25.7	17.6	10.8	6.2	3.4	1.9	1.0	0.5	0.3	0.1	0.1	0.0
	实测值	0.0	22.4	26.4	23.9	10.9	8.5	2.5	1.0	0.0	0.5	0.0	0.0	0.0	0.0
成 都	理论值	0.0	19.7	20.3	15.8	10.9	7.1	4.5	2.7	1.6	1.0	0.6	0.3	0.2	0.1
	实测值	0.0	24.9	15.4	15.4	10.4	4.5	8.5	1.5	2.5	0.0	1.0	0.0	0.5	0.5
南 通	理论值	0.0	17.3	18.4	14.8	10.6	7.2	4.7	3.0	1.9	1.2	0.7	0.4	0.3	0.1
	实测值	0.0	17.4	14.9	19.4	11.4	7.0	4.5	2.5	2.0	0.5	0.5	0.0	0.0	0.0
上 海	理论值	0.0	17.0	18.4	14.9	10.8	7.3	4.8	3.0	1.9	1.2	0.7	0.4	0.1	0.0
	实测值	0.0	17.9	17.4	18.4	7.5	5.0	5.5	4.0	1.0	1.0	2.5	0.0	0.0	0.0
杭 州	理论值	0.0	14.4	16.1	13.5	10.2	7.2	5.0	3.3	2.2	1.4	0.9	0.6	0.6	0.2
	实测值	0.0	12.4	12.4	20.9	10.9	7.0	3.0	4.5	1.0	1.5	0.0	1.5	0.0	0.0
南 宁	理论值	0.0	25.5	22.2	16.5	10.9	6.9	4.2	2.5	1.4	0.8	0.5	0.3	0.1	0.0
	实测值	0.0	25.5	20.5	16.5	9.0	10.5	5.5	3.5	0.5	0.5	0.0	0.0	0.0	0.0

游程频数来看,也与前面分析的结论一致,例如宜昌长度为1个月的干湿月游程频数比其它站都高,而其次为南宁,相反,频数最少的三个站为兰州、杭州、银川,这一事实表明,宜昌、南宁干湿月转移频繁,而兰州、杭州、银川的干湿转移相对较少。从气候学观点来看,反映出宜昌、南宁出现持续性旱涝的机会不多,而其余三个站则持续出现旱涝的可能性相对较大。干、湿月游程的各种统计特征(含其概率分布),从平均意义上反映了各地旱涝气候特点。尽管这些统计参数本身因受抽样的影响而具有抽样振动,但其气候学意义却仍十分清楚。

此外,干、湿月游程长度的方差以及天气气候循环长度的方差,则从另一侧面反映一地月际干湿交替的稳定程度。由表3可见,北京干月游程平均长度为2.64(月),而其均方差则达到了2.08(月),它表明北京持续干月的变动幅度也是很大的。相反,北京湿月游程平均长度为1.59(月),而其均方差仅为0.97(月),它表明该地持续湿月不但时间不长,而且变动幅度也不大。由此就可大致描述北京月际降水正负距平的演变特点,即偏旱的月份持续性强,且持续月数变化性也大,而偏涝的月份持续性弱,且持续月数的变化性小。类似的情况还可从其它地区看出。

#### (四)干、湿月转移概率矩阵的稳定性及其极限分布

由于涉及游程分布模式及其参数的所有计算都必须首先计算转移矩阵

$$P = \begin{pmatrix} P_{d|d} & P_{w|d} \\ P_{d|w} & P_{w|w} \end{pmatrix} \quad (9)$$

其中 $P_{d|d}$ 代表由干月转干月的概率, $P_{w|d}$ 代表由干月转湿月的概率, $P_{d|w}$ 为由湿月转干月的概率, $P_{w|w}$ 为由湿月转湿月的概率。若各站转移矩阵随样本增大到达稳定状态,则计算结果较为可靠,否则,因转移概率抽样波动太大会造成计算结果出现较大误差。为了论证转移矩阵的稳定性,本文以上海资料为例,计算了样本为25年(25×12月)、30年(30×12月)和100年(100×12月)的转移矩阵如下:

$$P_{25\text{年}} = \begin{pmatrix} 0.6848 & 0.3152 \\ 0.5703 & 0.4297 \end{pmatrix}$$

$$P_{30\text{年}} = \begin{pmatrix} 0.5729 & 0.4271 \\ 0.4910 & 0.5090 \end{pmatrix}$$

$$P_{100\text{年}} = \begin{pmatrix} 0.5708 & 0.4292 \\ 0.5455 & 0.4545 \end{pmatrix}$$

由上可见,当样本达到30年以上时,转移矩阵已基本稳定,这一结果与文献[12]所计算的逐日降水转移概率矩阵的稳定样本基本上是一致的。在那里,作者曾证明10年(某月)逐日降水转移矩阵已趋于稳定,因为10年即指10×30日,亦即样本容量达到300以上。本文针对逐月降水转移矩阵,实质上30×12月,也恰好达到样本容量为300以上。由此可见,这并非是一种偶然的巧合。作者在最近所做的其他研究中也发现有类似的规律。

本文所计算的各站样本均已达到上述要求，因此，我们可以粗略地认为抽样引起的随机误差较小。

遍历性是Markov链的重要性质。对于本文所涉及的干湿月游程模拟来说，遍历性表明无论初始状态如何，在相当长时间以后，过程可转移到任何状态。一般说来，天气气候状态的随机转移过程可用所谓正则Markov链来描述，正则链就是一种各态遍历链。对于模拟干、湿月转移规律来说，转移矩阵的极限分布是指，当转移步数 $n \rightarrow \infty$ 时，干、湿月状态的转移概率是否收敛于一组极限概率 $P^*$ ，而 $P^*$ 与初始状态无关。文献[7]已经证明当满足一定条件时，这样的极限概率总是存在的，即有

$$P^* \parallel P \cdot P \quad (10)$$

其中  $P = \begin{pmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{pmatrix}$  为两状态Markov链转移概率矩阵。上式给出一组矩阵方程组

$$\left. \begin{aligned} P_0^* &= P_{00}P_0^* + P_{10}P_1^* \\ P_1^* &= P_{01}P_0^* + P_{11}P_1^* \end{aligned} \right\} \text{(当 } P_0^* + P_1^* = 1)$$
 (11)

从理论上说，稳定状态的极限概率 $P_0^*$ 、 $P_1^*$ 将等于气候概率。计算表明，这两者是十分接近的(见表5)。本文对所选测站求得的极限分布，一般都经过几步转移后，干、湿月状态的概率就已收敛于某一极限 $P^*$ ，而与初始状态无关。根据柯尔莫哥洛夫方程，我们可以求得各步转移矩阵，例如

$$\begin{aligned} \text{上海: } & \begin{pmatrix} 0.5729 & 0.4271 \\ 0.4910 & 0.5090 \end{pmatrix} \xrightarrow{(2)} \begin{pmatrix} 0.5379 & 0.4621 \\ 0.5317 & 0.4683 \end{pmatrix} \xrightarrow{(3)} \begin{pmatrix} 0.5348 & 0.4652 \\ 0.5348 & 0.4652 \end{pmatrix} \\ \text{北京: } & \begin{pmatrix} 0.6205 & 0.3795 \\ 0.6295 & 0.3705 \end{pmatrix} \xrightarrow{(2)} \begin{pmatrix} 0.6239 & 0.3761 \\ 0.6238 & 0.3762 \end{pmatrix} \\ \text{兰州: } & \begin{pmatrix} 0.6425 & 0.3575 \\ 0.4934 & 0.5066 \end{pmatrix} \xrightarrow{(2)} \begin{pmatrix} 0.5899 & 0.4101 \\ 0.5700 & 0.4300 \end{pmatrix} \xrightarrow{(3)} \begin{pmatrix} 0.5799 & 0.4201 \\ 0.5799 & 0.4201 \end{pmatrix} \\ \text{宜昌: } & \begin{pmatrix} 0.5076 & 0.4924 \\ 0.5988 & 0.4012 \end{pmatrix} \xrightarrow{(2)} \begin{pmatrix} 0.5525 & 0.4475 \\ 0.5443 & 0.4557 \end{pmatrix} \xrightarrow{(3)} \begin{pmatrix} 0.5488 & 0.4512 \\ 0.5488 & 0.4512 \end{pmatrix} \end{aligned}$$

类似地，求得哈尔滨、银川、成都、南宁的极限分布，其转移步数除哈尔滨外，均需3步转移达到极限概率。表5中列出各站极限概率及其转移步数。由表可见，一般说来，华北、东北地区的临界转移步数约为2步，其它地区约为3步。这就意味着，对华北、东北而言，干、湿月预报应考虑到前1个月的干、湿状态，而其它地区干、湿月预报，则应考虑到前两个月的干、湿演变。

#### (五)上海近百年干、湿月游程演变特点

对上海近百年(1881—1980)月降水量资料，按上述方法详细分析计算，得到如下有意义的结果：

表5 各站极限概率与转移步数临界值

站		北京	哈尔滨	银川	兰州	宜昌	成都	南通	上海	杭州	南宁
干月	气候概率	0.62	0.57	0.63	0.58	0.55	0.55	0.58	0.53	0.59	0.56
	极限概率	0.62	0.57	0.63	0.58	0.55	0.55	0.58	0.56	0.59	0.56
湿月	气候概率	0.38	0.43	0.37	0.42	0.45	0.45	0.42	0.47	0.41	0.44
	极限概率	0.38	0.43	0.37	0.42	0.45	0.45	0.42	0.44	0.41	0.44
临界转移步数		2	2	3	3	3	3	3	3	3	3

(1)百年平均而言,干、湿月游程长度几乎不超过10个月,且短游程出现机会多。若将年景状况加以划分,以年降水量累年均值为基准,正负一个标准差上下为界划分旱、涝和正常年三种年景状况,计算干、湿月游程及其统计量,有如下特点:1)在旱年中,干月持续4个月以上的概率为0.39;2)而在涝年中,湿月持续3个月以上的概率为0.41;3)在旱年中,天气气候循环平均长度为4.961(方差为2.941)而涝及正常年中分别为4.459(方差为2.359)、4.043(方差为2.060)。这些特点表明,一旦出现旱、涝异常年,约有40%的可能出现持续旱、涝月,且旱月持续往往比涝月更长(平均约4个月以上);而所有旱涝异常年的天气气候循环长度都比正常年长,平均几乎将近5个月时间。这正是长江中下游地区常见的一个天气气候事实,它与我们的实践经验非常吻合。

(2)从旱、涝年的转移矩阵来看,干、湿月转移的极限分布其临界步数约为3步,而正常年仅为2步,表明在异常年干、湿月转移较慢,所以预报应充分考虑前两个月的影响。这一点与上面分析的特点也是一致的。

#### 四、小 结

1.本文用 Markov 链模拟各地干、湿月游程统计特征,所得结果可作为气候背景参考。

2.分析和计算结果表明,在一阶链假设下,已知转移概率矩阵,就可从理论上计算出一地干、湿月游程的统计特征:平均干、湿月游程长度及方差、天气气候循环平均长度、平均返回时间、达到极限分布(与初始状态无关)的临界步数等等,这些指标,从统计气候的观点,全面描述了一地月际降水演变的特征,对于中长期天气预报有参考意义。

3.所选站点干、湿月游程的具体特点,已在文中评述。

## 参 考 文 献

- [ 1 ] 么枕生, 应用转折点与游程的气候分析与预报, 地理研究, 5卷, 3期, 1—10, 1988.
- [ 2 ] Mood, A.M., The distribution theory of runs, *Ann. of Math. Stat.*, 11, 367, 1940.
- [ 3 ] Feller, W., *An introduction to Probability theory and its applications*, I—II, Wiley and Sons, New York, 1957—1971.
- [ 4 ] Yevjevich, V., *Stochastic processes in hydrology*, Fort Collins, 1972.
- [ 5 ] Hunter, J.J., *Mathematical techniques of applied probability*, I, Academic Press, 1983.
- [ 6 ] 杨鉴初, 运用气象要素的历史变化作一年以上的长期天气预报, 天气月刊, 13, 1951.
- [ 7 ] 么枕生, 气候统计学基础, 科学出版社, 1984.
- [ 8 ] Yao, C.S., II International meeting on Stati. Climatology, 1983.
- [ 9 ] 么枕生, 湿润期、干旱期与干湿气候循环的概率, 南京大学学报(气象学), 2, 1—18, 1962; 高校自然科学学报, 1964.
- [ 10 ] Berger, A. and C. Goossens, Persistence of wet and dry spells at Uccle (Belgium), *J. Climate*, 3, 21—34, 1983.
- [ 11 ] Lowry, W.P. and D. Guthrie, Markov chains of order greater than one, *Mon. Wea. Rev.*, 96, 1968.
- [ 12 ] 丁裕国、张耀存, 降水气候特征的随机模拟试验, 南京气象学院学报, 1989, 2, 146—155.

## A MARKOV CHAIN SIMULATION FOR DRY AND WET MONTH RUNS

Ding Yuguo    Niu Tao

Much climatological-statistical information of practical significance has been obtained through the study of the statistical features of dry/wet month runs at a number of representative stations by using Markov chain and from the concepts of runs and turning points presented in Literature 1. In addition, the stability of the dry/wet month transition probability matrix and its limit distribution are discussed. The results in this study provide a climatic background for making long-term forecasts.