降雨年际气候变化的非线性动力统计模拟预测*

严华生　曹　杰　谢应齐　季崇萍　郭世昌

(云南大学地球科学系, 昆明, 650091)

摘　要

文中应用非线性动力统计方法来模拟降水年际气候变化。经过对中国 42 个测点近百年降雨变化的模拟预测检验, 取得了较好的效果, 并得出了一些有意义的结论。

关键词: 非线性, 动力统计, 模拟预测。

1　引　言

以往常用的气候模拟预测方法有动力模拟方法和统计模拟方法, 但由于这两种方法的模拟观测结果与实际观测结果之间均存在一定误差, 因而值得进一步研究改进模拟预测方法。

随着非线性科学的发展和人们逐渐认识到气候模拟的复杂性, 于是提出了非线性动力统计相结合的模拟预测方法[1]。在气候模拟预测研究中, 可把多年来积累的实际观测资料看成气候动力系统的一系列特解, 可根据这些特解来反演所研究的气候动力系统, 并用于模拟预测实验。文中就是根据这一设想, 把年雨量气候变化设为非线性时变动力系统, 虽不知道描述这一系统的具体数学物理模型, 但却掌握了它的一系列特解, 即多年积累的降雨观测资料。于是以变系数非线性动力微分方程作为描述年雨量气候变化系统的数学模型, 利用观测资料求得微分方程的时变参数序列, 然后建立时变参数自身振动和外界强迫作用影响的非线性回归方程, 来模拟和预测动态系统时变参数的变化。通过以上动力统计模式来模拟和预测年雨量的变化。

作为例子, 对中国大陆共 42 个站 95 a(1901～1995 年), 每个站的年雨量气候变化进行了模拟预测研究。通过大范围实验表明, 此方法的历史模拟预测准确率已达到或略高于目前公认的水平, 值得在模拟预测中进一步研究。

2　数学模型

在考虑年雨量变化时, 先对降雨资料进行距平百分率处理后, 再分析降雨距平百分率变化情况。

设雨量的一般非线性时变梯度动力系统可写为如下微分方程[2]:

---

*　初稿时间: 1997 年 7 月 9 日; 修改稿时间: 1998 年 5 月 12 日。

资助课题: 国家自然科学基金(49865001)。

$$\frac{\mathrm{d}R(t)}{\mathrm{d}t} = -f(R(t), C(t)) = 0 \tag{1}$$

式(1)说明雨量变化受自身变化规律 $R(t)$ 和系统内外控制参数因素 $C(t)$ 的综合影响。对非线性函数 f 展开为 $R(t)$ 的 Taylor 级数形式, 设其为物理上的双稳态结构, 于是取展开式的截断项最高次数为 3 次, 根据突变论中的微分同胚和拓扑等价性质, 则可把式(1)写为如下月尖突变方程形式:

$$\frac{\mathrm{d}R(t)}{\mathrm{d}t} = -(4R^3(t) + 2R(t)u(t) + v(t)) = 0 \tag{2}$$

在式(2)中, $u(t)$ 和 $v(t)$ 为待定的随时间变化的时变微分方程控制参数, 它表示 t 时刻系统自身变化及外界的影响作用。把式(2)写成差分形式, 取 $\Delta t = 1$, 即得:

$$R(t+1) - R(t) = -(4R^3(t) + 2u(t)R(t) + v(t)) \tag{3}$$

应用多年降雨观测资料, 把它列为如下数据观测资料矩阵:

$$\boldsymbol{D} = \begin{bmatrix} R(2) - R(1) \\ R(3) - R(2) \\ \\ R(n) - R(n-1) \end{bmatrix}; \quad \boldsymbol{G} = \begin{bmatrix} R^3(1) & R(1) & 1 \\ R^3(2) & R(2) & 1 \\ & & \\ R^3(n-1) & R(n-1) & 1 \end{bmatrix};$$

$$\boldsymbol{C} = \begin{bmatrix} -4 & -2u(1) & -v(1) \\ -4 & -2u(2) & -v(2) \\ & & \\ -4 & -2u(n-1) & -v(n-1) \end{bmatrix}$$

于是, 对应于式(3)的观测数据矩阵方程可写为

$$\boldsymbol{D}(t) = \boldsymbol{G}(t) \quad \boldsymbol{C}^{\mathrm{T}}(t) \tag{4}$$

应用观测资料数据矩阵方程式(4), 就可以来反演时变动力系统的控制参数 $\boldsymbol{C}(t)$ 。

首先, 根据描述运动变化的一般性原理

$$A(t) = \bar{A} + A'(t) \tag{5}$$

在式(5)中, $A(t)$ 为任意随时间变化的系统, $\bar{A}$ 为该系统的平均状态, $A'(t)$ 为该系统随时间变化状态。

于是, 可把时变动力系统式(4)的时变参数反演问题归结为: 求在约束条件 $\boldsymbol{D}(t) = \boldsymbol{G}(t) \quad \boldsymbol{C}^{\mathrm{T}}(t)$ 下, 使残差函数 $E = \| \hat{\boldsymbol{C}}(t) - \bar{\boldsymbol{C}} \|$ 达到最小, 应用微积分中求解条件极值的拉氏乘子法即可导得:

$$\hat{\boldsymbol{C}}(t) = \bar{\boldsymbol{C}} + \frac{1}{\boldsymbol{G}(t) \quad \boldsymbol{G}^{\mathrm{T}}(t)} \boldsymbol{G}(t) \quad \boldsymbol{D}(t) - \boldsymbol{G}(t)\bar{\boldsymbol{C}}^{\mathrm{T}} \tag{6}$$

式中 $\bar{\boldsymbol{C}}(t) = [\boldsymbol{G}\boldsymbol{G}^{T}]^{-1}\boldsymbol{G}^{T}\boldsymbol{D}$。即假设参数不随时间变化的定常模式为 $\boldsymbol{D}(t) = \boldsymbol{G}(t)\bar{\boldsymbol{C}}^{T}$, 应用最小二乘法, 即可求得 $\bar{\boldsymbol{C}}$ 的表达式。

由此可见, $\bar{\boldsymbol{C}}$ 为系统处于定常状态的常系数, 可理解为动态系统的平均状态, 把它作为初值代入式(6), 就可求得动态系统时变参数序列。

根据式(6)反演出 $u(t)$ 和 $v(t)$ 后, 由于认为 $u(t)$ 和 $v(t)$ 是系统自身变化与外界强迫影响共同作用而造成的, 通过控制参数的时间演变体现出来, 于是可建立系统自身变化因素、大气环流、海温等外界变化对 $u(t)$ 和 $v(t)$ 产生影响的非线性回归方程[3]

$$\begin{cases} \hat{u}(t) = f_1(u(t-k), \cdots, v(t-k), x_1(t-k), \cdots, x_m(t-k)) \\ \hat{v}(t) = f_2(u(t-k), \cdots, v(t-k), x_1(t-k), \cdots, x_m(t-k)) \end{cases} \tag{7}$$

从而完成系统建模和参数反演。

然后将式(3)改写为

$$R(t+1) = R(t) - (4R^3(t) + 2R(t)u(t) + v(t)) \tag{8}$$

即可用于气候模拟和预测。首先, 应用回归方程式(7)求得 $u(t)$ 和 $v(t)$ 值, 再代入式(8)求得年雨量模拟预测值。

3 具体应用

3.1 资料来源

选取中国大陆 42 个站点 95 a(1901～1995)的年降雨量为研究对象。所选取得 42 个站点的地理位置如表 1 所示:

表 1 42 站点地理位置 单位: °

站点	经度	纬度	站点	经度	纬度	站点	经度	纬度
博克图	121	55	海拉尔	119	49	齐齐哈尔	123	47
哈尔滨	126	45	牡丹江	129	36	长春	125	43
延吉	129	42	沈阳	123	41	朝阳	120	41
大连	121	38	承德	117	40	张家口	115	40
北京	116	39	烟台	121	37	青岛	120	41
济南	117	36	蚌埠	117	32	南京	118	32
上海	121	31	杭州	120	30	九江	116	29
汉口	114	30	岳阳	113	29	温州	120	28
福州	119	25	长沙	112	46	厦门	118	24
汕头	116	23	曲江	113	24	河源	114	50
广州	113	23	南宁	108	22	北海	109	21
百色	106	23	贵阳	106	26	重庆	106	29
成都	104	30	宜宾	104	28	昆明	102	41
蒙自	103	23	西安	108	34	兰州	103	36

由表可见, 这 42 个站地理分布范围在(20～55°N, 100～130°E), 包括了中国大部分地区。在如此广大的范围进行方程的反演、建模、模拟、检验有助于验证本方法的可行性。

3.2 具体步骤

3.2.1 降雨量系统的反演

由公式(6)可获得控制参数 $u(t)$, $v(t)$ 的时间演变序列。

3.2.2 利用统计方法建立因子与 $u(t)$, $v(t)$ 的回归方程

先综合考虑外界强迫的影响。这里主要选取了 1901～1995 年的北半球西风带大气环流 W, C, E 型年频数和南方涛动(SOI)年指数, 考虑这些因子之间存在的非线性影响和南北环流因子间不同配置相互作用的综合影响, 对环流因子进行多项式展开, 共生成 11 个影响因子[2], 如表 2 所示:

表 2　外界强迫因子说明

函数类型	因子号	因子意义
线性项	x_1	南方涛动指数(SOI)
	x_2	大气环流 W 型年频数
	x_3	大气环流 C 型年频数
	x_4	大气环流 E 型年频数
非线性二次项	x_5	x_1x_1
	x_6	x_2x_2
	x_7	x_3x_3
	x_8	x_4x_6
大气环流因子相互作用项	x_9	x_1x_2
	x_{10}	x_1x_3
	x_{11}	x_1x_4

再考虑系统内部自身变化的影响, 取二次项展开, 如表 3 所示:

表 3　系统外部因子说明

x_{12}	x_{13}	x_{14}	x_{15}	x_{16}
$u(t)$	$v(t)$	$u(t)\quad v(t)$	$u^2(t)$	$v^2(t)$

于是系统内外部共取 16 个影响因子。再考虑前两时刻对后时刻的积累影响, 即 $(t-1)$, $(t-2)$ 时刻对 t 时刻的影响, 最终生成 32 个预报因子, 用这 32 个因子分别与 $u(t)$ 和 $v(t)$ 进行逐步回归, 计算结果列于下表 4。表中 U, V 列分别表示对 $u(t)$ 和 $v(t)$ 建立的逐步回归方程的检验结果。

表 4 逐步回归检验

	博克图		海拉尔		齐齐哈尔		哈尔滨		牡丹江		长春		延吉	
	U	*V*	*U*	*V*	*U*	*V*	*U*	*V*	*U*	*V*	*U*	*V*	*U*	*V*
R 复相关系数	0.77	0.76	0.84	0.65	0.88	0.77	0.72	0.77	0.68	0.71	0.71	0.61	0.87	0.77
方程 *F* 检验	14.53	9.64	14.10	6.18	21.55	8.26	46.16	8.40	4.83	11.53	9.06	6.69	36.54	8.14

	沈阳		朝阳		大连		承德		张家口		北京		烟台	
	U	*V*	*U*	*V*	*U*	*V*	*U*	*V*	*U*	*V*	*U*	*V*	*U*	*V*
R 复相关系数	0.77	0.63	0.87	0.69	0.67	0.76	0.72	0.67	0.73	0.70	0.88	0.70	0.65	0.70
方程 *F* 检验	8.41	6.42	20.30	5.79	22.83	7.86	6.10	4.67	11.05	5.51	32.20	7.21	6.26	5.32

	青岛		济南		蚌埠		南京		上海		杭州		九江	
	U	*V*	*U*	*V*	*U*	*V*	*U*	*V*	*U*	*V*	*U*	*V*	*U*	*V*
R 复相关系数	0.87	0.76	0.74	0.73	0.74	0.53	0.79	0.72	0.83	0.79	0.75	0.69	0.76	0.65
方程 *F* 检验	20.47	11.88	9.23	6.23	12.02	3.87	11.25	6.56	15.04	11.66	7.13	9.69	9.66	8.27

	汉口		岳阳		温州		福州		长沙		厦门		汕头	
	U	*V*	*U*	*V*	*U*	*V*	*U*	*V*	*U*	*V*	*U*	*V*	*U*	*V*
R 复相关系数	0.58	0.74	0.75	0.64	0.67	0.78	0.71	0.81	0.68	0.71	0.67	0.72	0.75	0.69
方程 *F* 检验	4.33	7.42	7.33	3.40	6.16	8.91	10.17	12.12	5.14	8.64	17.10	6.07	7.42	5.30

	曲江		河源		广州		南宁		北海		百色		贵阳	
	U	*V*	*U*	*V*	*U*	*V*	*U*	*V*	*U*	*V*	*U*	*V*	*U*	*V*
R 复相关系数	0.70	0.78	0.84	0.74	0.74	0.70	0.66	0.72	0.84	0.69	0.70	0.71	0.65	0.71
方程 *F* 检验	6.00	8.87	32.00	6.95	6.85	6.06	4.32	6.10	69.04	8.81	5.43	6.90	8.40	7.63

	重庆		成都		宜宾		昆明		蒙自		西安		兰州	
	U	*V*	*U*	*V*	*U*	*V*	*U*	*V*	*U*	*V*	*U*	*V*	*U*	*V*
R 复相关系数	0.67	0.75	0.63	0.48	0.74	0.75	0.64	0.71	0.82	0.81	0.70	0.68	0.84	0.75
方程 *F* 检验	5.15	11.29	10.90	3.39	6.95	7.19	4.10	6.99	11.68	10.52	8.10	5.91	13.97	11.20

由表 4 中数据可得出如下几点结论:

(1) 方程均通过 $\alpha = 0.01$ 的显著性检验。

(2) 就控制参数 $u(t)$ 的回归方程的复相关系数来看: 在全国 42 个站点中只有 1 个站小于 0.60, 占总站数的 2%; 有 13 个站在 0.60～0.70 之间, 占总站数的 31%; 有 17 个站在 0.70～0.80 之间, 占总站数的 40%; 有 11 个站大于 0.80, 占总站数的 26%。

(3) 就控制参数 $v(t)$ 的回归方程的复相关系数来看: 在全国 42 个站点中只有 2 个站小于 0.60, 占总站数的 5%; 有 13 个站在 0.60～0.70 之间, 占总站数的 31%; 有 25 个站在 0.70～0.80 之间, 占总数的 59%; 有 2 个站大于 0.80, 占总站数的 5%。上述结果说明此方法在全国大范围模拟效果较好。

3.2.3 历史模拟检验

先用回归方程预报出 u, v, 再用 u, v 做年雨量 R 的预测。为检验本方法进行历史模拟的准确性, 进行如下两种检验:

(1) 分别计算年降雨量的标准化值 $\frac{R(t) - \bar{R}}{\bar{R}}, \frac{\hat{R} - \bar{R}}{\bar{R}}$, 若两者同号, 则认为历史模拟与实际观测相符, 记为对。若两者异号, 但差的绝对值小于 0. 10, 也认为历史模拟与实际观测相符, 亦记为对, 否则为错。对的年份累计为 n_1, 检验总年份记为 n, 计算 n_1/n 的值。

(2) 利用相关公式:

$$r(R(t), \hat{R}(t)) = \frac{\sum (R(t) - \bar{R}) \quad \sum (\hat{R}(t) - \bar{R})}{\sum (R(t) - \bar{R})^2 - (\hat{R}(t) - \bar{R})^2} \tag{16}$$

计算 $r(R(t), \hat{R}(t))$ 。两种检验的结果列于表 5。

表 5　历史模拟检验

站点	相关系数	n_1/n	站点	相关系数	n_1/n
博克图	0. 652	0. 659	汉口	0. 602	0. 693
海拉尔	0. 605	0. 636	岳阳	0. 516	0. 682
齐齐哈尔	0. 669	0. 568	温州	0. 619	0. 705
哈尔滨	0. 643	0. 702	福州	0. 670	0. 682
牡丹江	0. 618	0. 727	长沙	0. 638	0. 648
长春	0. 613	0. 648	厦门	0. 587	0. 511
延吉	0. 570	0. 591	汕头	0. 620	0. 705
沈阳	0. 550	0. 648	曲江	0. 724	0. 727
朝阳	0. 508	0. 636	河源	0. 511	0. 670
大连	0. 597	0. 739	广州	0. 521	0. 659
承德	0. 650	0. 648	南宁	0. 606	0. 739
张家口	0. 647	0. 727	北海	0. 408	0. 636
北京	0. 445	0. 636	百色	0. 628	0. 614
烟台	0. 688	0. 773	贵阳	0. 562	0. 682
青岛	0. 549	0. 670	重庆	0. 515	0. 682
济南	0. 609	0. 739	成都	0. 430	0. 648
蚌埠	0. 319	0. 568	宜宾	0. 675	0. 670
南京	0. 572	0. 705	昆明	0. 572	0. 636
上海	0. 661	0. 705	蒙自	0. 705	0. 693
杭州	0. 582	0. 670	西安	0. 607	0. 693
九江	0. 577	0. 670	兰州	0. 667	0. 534

由表 5 中数据可见, 在年降雨量的历史模拟和观测资料的相关中:

(1) 有 23 个站的相关值有 0. 6～0. 7 之间, 占总站数的 55%; 15 个站的相关值在 0. 5～0. 6 之间, 占总站数的 36%; 只有 4 个站的相关值在 0. 5 以下, 占总站数的 9%。

(2) n_1/n 值除了有 5 个站小于 0. 6 以外, 其它绝大部分位于 0. 6～0. 7 之间, 占总数

的 89%。这一结果说明用本方法进行短期气候预测, 已达到或略高于目前这一领域的公认水平。

3. 2. 4 预测

应用以上方法对 1996 年的年降雨量做了预测, 结果列于表 6。

表 6 1996 年的年降雨量距平预测

站点	距平	站点	距平	站点	距平	站点	距平	站点	距平	站点	距平
博克图	+	海拉尔	–	齐齐哈尔	–	哈尔滨	–	牡丹江	–	长春	+
延吉	+	沈阳	+	朝阳	+	大连	+	承德	–	张家口	–
北京	+	烟台	–	青岛	–	济南	–	蚌埠	+	南京	+
上海	–	杭州	–	九江	+	汉口	+	岳阳	+	温州	–
福州	–	长沙	+	厦门	–	汕头	–	曲江	–	河源	–
广州	–	南宁	–	北海	–	百色	–	贵阳	–	重庆	–
成都	–	宜宾	–	昆明	+	蒙自	–	西安	–	兰州	+

为了更进一步讨论, 还绘制了关于方程控制参数 $u(t)$, $v(t)$ 的反演值与回归值的相关场、年降雨量和实际观测值相关场(图略)。由图中可看出, 相关场分布很有规律, 高低值中心明显。

4 小结和展望

(1) 在气候模拟预测中, 有气候动力模拟预测和气候统计模拟预测两种方法。文中应用动力统计相结合的方法来预测年降雨量的年际气候变化, 结果显示模拟预测效果达到或略高于目前业务水平, 说明在气候模拟预测研究中, 发展动力统计预报方法是有重要意义的。

(2) 以往的气候动力模式, 多为常系数微分方程模式; 以往的气候统计预测模式, 多为线性统计预测模式。文中考虑到气候预测系统的非线性复杂性, 引入非线性时变模式, 结果表明, 非线性时变模式能更好地模拟和解释气候变化的复杂性。

(3) 目前在气候业务预报中, 存在两种预报思路, 一种是考虑预报因子对预报对象影响作用的预报思路, 另一种是考虑预报对象自身变化规律的预报思路, 每种思路都有一定的依据和效果, 但都存在预报不准的现象, 说明都有一定的局限性, 有必要进一步改进和发展。文中提出考虑系统自身变化规律和外界预报因子相互作用形成的综合影响的预报模式, 从而改进了预报效果。

(4) 文中尝试应用多年历史预测资料, 反演建立描述降雨气候变化的非线性微分时变动力系统, 对它的气候学意义, 动力学分析和突变性质探讨还有待进一步研究。

参考文献

1 丑纪范. 长期数值天气预报. 北京: 气象出版社, 1986, 329pp
2 严华生, 陈兴芳, 曹杰, 周传喜. 中国近百年雨量与大气环流因子关系. 热带气象学报, 1998, 14(3): 251～257
3 严华生, 谢应齐, 曹杰. 非线性统计预报及其应用. 昆明: 云南科技出版社, 1998, 129pp