

# 气象要素时间序列的演化建模分析与短期气候预测

俞康庆<sup>1</sup>, 周月华<sup>1</sup>, 杨荆安<sup>1</sup>, 康卓<sup>2</sup>

(1. 中国气象局武汉暴雨研究所, 湖北 武汉 430074; 2. 武汉大学计算中心, 湖北 武汉 430072)

**摘要** 以武汉站(5~9月)汛期降水量观测数据序列为例, 将这类具有明显的不规则性(混沌特征)时间序列分解为宏观气候尺度周期的波动部分和迭加其上的微观气候尺度周期的波动部分, 分别采用演化建模方法和自然基小波方法模拟逼近。特别强调由演化建模方法得到的非线性常微分方程较之传统的线性建模具有更好的分析预测能力。

**关键词**: 时间序列, 演化建模, 短期气候预测

中图分类号: P456.1

文献标识码: A

## 引言

气象要素观测数据时间序列可以视作是一个精确完美的气候数值预报模式在完备定解条件下的解, 是各种气候因子(也包括其自身)相互间非线性作用的综合演化表现形式。随着气象观测方法的多元化、大型计算机的应用, 特别是气象学家对气候问题认识的拓展和深入, 短期气候数值预报模式取得了很大的进展。但是受到许多客观条件的限制, 对气候变化过程物理机制的未解决问题不少于已认识和解决的问题, 故气候数值预报模式尚不能像短期天气数值预报模式那样被普及, 因此, 对观测数据时间序列进行线性建模分析<sup>[1~2]</sup>在气象要素的短期气候预测业务中仍占相当的比重, 如使用AR模型、MA模型、ARMA模型、ARIMA模型。这些模型用于某些领域确实有很好的效果, 但从本质上讲它们是线性的, 存在一定的局限性。文献[1]也列举了一些拓宽的广义线性模型和非线性模型, 但是约定模型的非线性函数类型必须是已知的, 由人为主观估计。那么能否找到一种数学方法客观地模拟具有非线性特征的气象要素时间序列, 近十多年来, 应用数学领域中涌现出遗传算法<sup>[3]</sup>、及在遗传算法基础上进一步发展起来的适用于复杂系统分析的演化建模<sup>[4~5]</sup>先进算法。

演化建模方法是以遗传算法为基础, 模仿生物

演化过程(优化遗传、自然选择等规则)来求解复杂问题的一类计算方法。具有自适应、自组织、自学习以及内在的并行性等智能特征。在目前不能确定气候数值模式(非线性偏微分方程组)中气候因子非线性相互作用物理机制的显式数学表达式的情况下, 能够用非线性常微分方程数学模型去逼近气象要素历史观测时间序列, 并进而作短期气候预测不失为一种有效方法。本文以(5~9月)降水量(简称汛期降水量)作试验, 进行气象要素演化建模短期气候预测的方法研究。

## 1 思路和方法

### 1.1 演化建模思路

假设气象要素  $x$  的  $m$  个历史观测值(样本)的时间序列为

$$X = (x(t_0), x(t_1), x(t_2), \dots, x(t_{m-1}))^T$$
 式中  $t_i = t_0 + i \times \Delta t$  ( $i = 0, 1, 2, \dots, m-1$ ) 我们可以认为该序列是在有界可测集  $T$  (时间坐标上的  $t_0, \dots, t_0 + (m-1) \times \Delta t$ ) 上的可测函数  $x(t)$  的观测值, 如果我们在  $L^2(T)$  空间中能找到一个函数  $x^*(t)$ , 使  $x(t)$  与  $x^*(t)$  之间的距离  $\|x - x^*\|$  最小, 这是数学建模的基本原理。但是, 作为气象要素时间序列而言, 表现得很无序, 线性函数难以模拟逼近, 若能找到一个非线性函数, 不仅满足上述的距离  $\|x - x^*\|$  最小, 而且还能体现出序列随时间演变的规律, 于是可以将气象

要素动态数据建模问题转换为下述的数学问题。求解一个由初等函数组成的复合函数  $f$  表示的高阶非线性常微分方程 (HNODE),

$$\frac{d^n x^*}{dt^n} = f\left[t, x^*(t), \frac{dx^*}{dt}, \frac{d^2x^*}{dt^2}, \dots, \frac{d^{(n-1)}x^*}{dt^{(n-1)}}\right] \quad (1)$$

$f$  包括的初等函数集为  $O( +, -, \times, /, \sin, \cos, \exp, \ln )$  且满足初始条件

$$\begin{cases} x^*(t_0) = x(t_0) \\ \frac{d^j x^*(t_0)}{dt^j} = \frac{d^j x(t_0)}{dt^j} \quad (j = 1, 2, 3, \dots, n-1) \end{cases} \quad (2)$$

并使

$$\|X^* - X\| = \sqrt{\sum_{i=0}^{m-1} [x^*(t_i) - x(t_i)]^2} \quad (3)$$

达到最小值,则我们定义(1)式为气象要素  $x$  的历史观测时间序列  $X$  的非线性常微分方程数学模型。它表明,若  $n=1$ ,数学模型为一阶非线性常微分方程,即讨论气象要素  $x$  随时间  $t$  变化规律,必依赖于时间  $t$  和  $x$  自身;若  $n>1$ ,即讨论气象要素  $x$  随时间  $t$  的  $n$  阶(高阶)变化时,不仅依赖于时间  $t$  和  $x$  自身,而且依赖于  $x$  的一阶微商、二阶微商直至  $(n-1)$  阶微商。对(1)作  $m$  步数值积分,得到该气象要素  $x$  的时间序列的近似值

$X^* = (x^*(t_0), x^*(t_1), x^*(t_2), \dots, x^*(t_{m-1}))^T$   $t_i = t_0 + i \times \Delta t$  ( $i=0, 1, 2, \dots, m-1$ )。继续作  $l$  步数值积分,可得到该气象要素  $x$  在  $t_0 + m \times \Delta t$  到  $t_0 + (m+l-1) \times \Delta t$  各时刻的预测值  $\{x^*(t_m), x^*(t_{m+1}), x^*(t_{m+2}), \dots, x^*(t_{m+l-1})\}$ 。

与传统的时间序列分析方法<sup>[1]</sup>比较,气象要素的时间序列是复杂气候系统中状态变量非线性相互作用的综合演化表现形式,能够用非线性常微分方程数学模型去模拟逼近气象要素历史观测时间序列,并进而作短期气候预测,它可以比线性模型更有效地模拟气候系统的自然特征;并且模型的非线性常微分方程可以作为参考,有助我们开拓短期气候分析与研究的思路。而建立非线性常微分方程数学模型的关键是如何寻求方程(1)式等号右边的函数  $f$  表达式,下一节简要介绍演化建模方法。

### 1.2 演化建模方法

作为初值问题的高阶常微分方程(1)式可以转换成求解一阶常微分方程的方程组。为书写简便,省略(1)式中的 \* 符号,并作变量替换,

$$y_1 = x, y_2 = \frac{dx}{dt}, \dots, y_{n-1} = \frac{d^{(n-2)}x}{dt^{(n-2)}}, y_n = \frac{d^{(n-1)}x}{dt^{(n-1)}} \quad \text{则}$$

$$\begin{cases} y_2 = y'_1 \\ y_3 = y'_2 \\ \vdots \\ y_n = y'_{n-1} \\ y_{n+1} = y'_n = f(t, y_1, y_2, \dots, y_n) \end{cases} \quad (4)$$

且仍满足初始条件

$$\begin{cases} y_1(t_0) = x(t_0) \\ y_j(t_0) = \frac{dx^{(j-1)}(t_0)}{dt^{(j-1)}} \quad (j = 2, 3, \dots, n-1, n) \end{cases}$$

因此,决定数学模型的最关键部分是方程组(4)式中的最后一个方程的函数表达式  $y_{n+1} = f(t, y_1, y_2, \dots, y_n)$ 。由于它是由  $+, -, \times, /, \sin, \cos, \exp, \ln$  初等函数组成,在演化建模过程中,由遗传程序 GP (genetic programming) 依据变量  $x$  搜索上述基本函数,随机地形成若干数学模型结构个体;与此同时,参数优化处理程序不断地对模型结构个体中的参数进行估算并且优化,并以(3)式作为判别标准,淘汰差的模型结构个体,保留好的模型结构个体,最终求得一个误差最小的数学模型,这就是文献[5]作者提出的混合演化建模算法 (hybrid evolutionary modeling algorithm) 中心思想。

因此,混合演化建模算法具有自适应、自组织、自学习以及内在的并行性等智能特征。

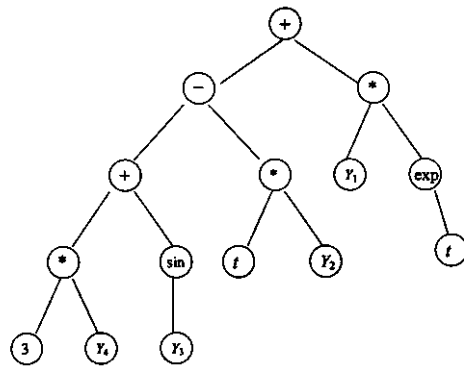


图1 高阶常微分方程的二叉树形式 (引自文献[5])

Fig. 1 The representation of a binary tree of  $y_5 = 3y_4 + \sin(y_3) - ty_2 + y_1 e^t$  (offered by Hongqing, et al 2000)

在遗传算法中,一个数学模型结构个体对应唯一的二叉树 (binary tree 或 parse trees) 形式,假定某个数学模型结构个体表达式是一个四阶常微分方程  $y_5 = 3y_4 + \sin(y_3) - ty_2 + y_1 e^t$  (符号含义如前所述),对应唯一的二叉树形式如图1所示,可见二叉树的

最大分叉层数  $d$  表示了微分方程的复杂程度,必须先设定;根据计算机条件的许可,还要事先设定需要多少模型结构个体,即初始化种群数  $b$  (breeds);在演化模拟过程中,遗传变异的代数  $g$  也需事先设定。综上所述,演化建模所需的参数为:样本数  $m$ ;非线性常微分方程的阶数  $o$ ;二叉树深度  $d$ ;初始化种群个数  $b$ ;遗传变异的代数  $g$  和预测步数  $l$ ,简洁表示为  $EM(m, o, d, b, g, l)$ 。

### 1.3 单气象要素时间序列 $X$ 的预测

本文中我们以汛期降水量代表气象要素  $x$ ,就短期气候预测而言,按照上一节内容中的定义:

$Y_1 = X$ , 气象要素观测数据的时间序列(时间间隔单位:年);

$Y_2 = (Y_1)'$ , 气象要素年际变化的时间序列;

$Y_3 = (Y_2)'$ , 气象要素年际变化趋势的时间序列;

在短期气候预测中,我们关心的是气象要素随时间的变化,即气象要素对时间的一阶微分方程  $\frac{dx}{dt}$

$=f(x, t)$ 。从方程中,我们不仅希望得到气象要素的预测值,还希望了解气象要素随时间变化的规律。就此意义而言,过高阶数的微分方程在气候预测问题中并没有实际意义,一阶微分方程足矣。本文采用文献[5]作者研制的演化建模软件,运行得到最好的数学模型表达式(一阶常微分方程)后,以  $Y$  的最后一行元素为初始值,对该表达式进行时间  $t$  积分  $l$  步,得到  $x^*(t_m), x^*(t_{m+1}), \dots, x^*(t_{m+l-1})$ ,作为未来时刻的预测值。在下一节中,以武汉站汛期降水量观测序列为例,介绍建模的具体做法。

## 2 个例计算

### 2.1 数据来源

选择处在梅雨带中的武汉观测站,由于历史原因,汛期降水量观测数据的长度以连续不间断的数据统计,有 42 a(1896~1937)和 57 a(1947~2003)两段序列,分别称为 Rwh42 和 Rwh57。直观看,2 段降水量时间序列曲线都表现了高度的不规则性(或混沌特征),其降水量年际变化规律差异较大。

### 2.2 数据序列的预处理——从数据序列 $X$ 中分解 $\bar{X}$ 和 $X'$

虽然汛期降水量观测序列具有混沌特征,但其中必然包含气候系统中各种强迫因子的演化信息,气候预测专家们长期研究分析<sup>[2,6]</sup>表明,从时间尺度而言,这些强迫因子使汛期降水量观测序列表现

出以百年而论的“世纪尺度”、以十年而论的“代尺度”、年尺度、月尺度等变化周期。虽然以精确的解析数学表达式表示几乎是不可能的,但为了能对这种具有混沌特征的时间序列作定量的描述,并根据历史观测数据进行预测,我们对数据作下述处理,将汛期降水量时间序列  $X$  分解为 2 部分,  $X = \bar{X} + X'$ 。其中  $\bar{X}$  为宏观气候尺度周期的波动部分(或称主值序列);  $X'$  为迭加在  $\bar{X}$  上的微观气候尺度周期的波动部分(或称扰动序列)。对宏观气候尺度周期的波动部分  $\bar{X}$  我们采用上一节中的演化建模方法模拟逼近;对微观气候尺度周期的波动部分  $X'$  我们采用自然基小波(或称自然分形)分解的方法进行模拟逼近(详见文献[7]),故使建立的数学模型体现气候因子多时间尺度影响的综合作用。

如何确定序列中的宏观气候尺度周期波动部分  $\bar{X}$ ,为了避免主观性和盲目性,首先采用文献[1]介绍的运用方差分析工具,从数据序列中识别提取整数周期的方法进行周期分析。

表 1 武汉站汛期降水量时间序列段包含的周期( $F$  检验的信度  $\alpha = 0.05$ )

Tab. 1 The periods included in the time series of precipitation in flood season (May - September) at Wuhan station (pass a standard "f test" with greater than 95% confidence)

观测数据序列 Rwhn( yy1 - yy2 )	周 期 ( 年 ) 分 析					
	T1	T2	T3	T4	T5	T6
Rwh42( 1896 - 1937 )	20	3				
Rwh57( 1947 - 2003 )	26	25	11	8	7	3

从表 1 的 2 段序列包含的周期可以看出:

(1) 过短的序列 Rwh42(1896~1937) 由于信息量有限,无法分析出多个周期。

(2) 序列 Rwh57(1947~2003) 包含的周期可以和气候分析专家们总结的规律联系,如太阳黑子活动的 11 a 周期<sup>[6]</sup>; ENSO 循环的显著时间尺度 3~5 a<sup>[5]</sup>; 使用功率谱分析发现的夏季西太平洋副热带高压南北位置存在 3~5 a 等的周期变化<sup>[2]</sup>; 文献[2]还从夏季降水类型的变化分析,认为“夏季多雨区偏北和偏南的准周期变化长度平均为 8.6 a, 实际在 7~10 a 内变动”。因此,在某种程度上,周期分析的结果具有一定的气候学意义。

在上述周期分析结果的基础上,我们考虑在  $\bar{X}$  中尽量保留较稳定的长年影响因子的作用,曾以平滑系数  $s$  为 9、10、11、12、13 分别对观测数据作滑动平均进

行试验,最终取在平滑序列之上叠加小波后拟合历史观测数据方差最小的  $s$  值。本个例  $s = 10$ ,对时间序列  $X = (x(t_0), x(t_1), x(t_2), \dots, x(t_{m-1}))^T$ ,  $t_i = t_0 + i \times \Delta t$  ( $i = 0, 1, 2, \dots, m-1$ ) 平滑,

$$\tilde{x}_0(t_j) = \begin{cases} \frac{1}{s} \sum_{j=0}^{s-1} x(t_j) & \text{当 } 0 \leq j \leq s-1 \\ \frac{1}{s} \sum_{j=j-s+1}^j x(t_j) & \text{当 } s \leq j \leq m-1 \end{cases}$$

(在演化建模中  $\tilde{x}_0(t_j)$  ( $j = 0, 1, \dots, s-1$ ) 不被使用)将周期长于 10 a 以上的气候因子的综合影响都归结到序列  $\tilde{X}$  中,由演化建模方法来模拟逼近; $x'(t_i) = x(t_i) - \tilde{x}_0(t_i)$  ( $i = 0, 1, 2, \dots, m-1$ ) 由文献 [7] 自然基小波方法模拟逼近。2 种方法得到的结果之和为最终的预测值。

### 2.3 计算结果分析

侧重对天文因子变化周期的考虑,将序列长度

定为 55 a,根据 1.2 节演化建模参数的定义  $EM(m, n, d, b, g, \rho, l)$  经过大量的计算分析,本文采用下列参数设置:

$m = 55$ , 观测数据序列长度;

$o = 1$ , 用一阶常微分方程模拟;

$d = 3$ , 二叉树深度为 3 层;

$b = 100$ , 初始化种群数取 100;

$g = 20$ , 演化运算中经过 20 代遗传即终止;

$l = 5$ , 预报未来 5 a 的汛期降水量(不宜过长);

由于每一次得到的常微分方程表达式是随机的,因此在试验中,对同一个数据序列运算 10 次,现以最简单的方式,10 次结果的算术平均数作为预测结论。表 2 列出了以 Rwh55(1949~2003) 序列进行演化建模得到的 10 个常微分方程及未来 3 a 的汛期降水量预测值,从运算结果可知:

表 2 Rwh55(1949~2003) 演化建模生成的 10 个随机常微分方程表达式及预报降水量

Tab. 2 Ten expressions of random NODE and their predicted figures in next 3 years based on time series of Rwh55(1949~2003)

	Expressions of NODE from ten runs of EM (depth 3, generation 20, breeds 100, order 1)	5~9 月降水量预测值(mm)		
		2004	2005	2006
1	$\frac{dx}{dt} = (t - 8.602427) \times \cos(x) \times x$	941	766	1 254
2	$\frac{dx}{dt} = \frac{x - t + 391.443512}{\sin(x)}$	943	764	1 342
3	$\frac{dx}{dt} = \frac{x}{\sin(x)} + \frac{t + 0.266459}{\sin(x)}$	1 020	879	1 484
4	$\frac{dx}{dt} = \cos(-113.667633 \times x) \times \frac{4278803.5}{x}$	981	806	1 391
5	$\frac{dx}{dt} = 6327.78479 \times \cos\left(\frac{841403.5625}{t}\right)$	917	738	1 288
6	$\frac{dx}{dt} = \frac{x}{\sin(x \times 17.44883)}$	961	783	1 383
7	$\frac{dx}{dt} = \frac{e^t - 0.453206 - x}{\sin(x)}$	959	793	1 396
8	$\frac{dx}{dt} = \frac{x + 12.219144}{\sin(x^3)}$	935	761	1 323
9	$\frac{dx}{dt} = (6.81838 \times x + x) \times \sin(x)$	988	780	1 346
10	$\frac{dx}{dt} = \ln( x ) - \frac{x + 92.175331}{\sin(x)}$	944	762	1 353
	(降水量预测值)平均	959	783	1 366

(1) 虽然常微分方程的表达式不同,但是其积分得到的数值基本上相接近(表略),叠加了自然基小波后,最终的降水量预测数值亦相对“集中”。

(2) 图 2 显示了表 2 中的数学模型 1 对 Rwh55

(1949~2003) 的 10 a 平滑序列  $\tilde{X}$  的拟合,在长年 ( $\geq 10$  a) 影响因子的作用下,汛期降水量主值序列部分随时间变化的规律表现为余弦形式的复杂周期函数,自 1960s 中期至 1980s 中期,汛期降水量宏观

气候尺度周期波动处于少雨期;自 1980s 中期至今

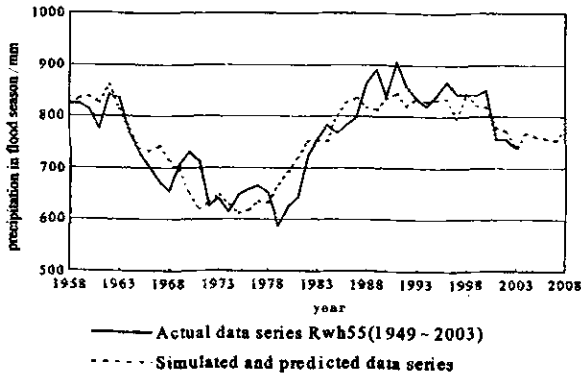


图2 (表2中)数学模型1对Rwh55(1949~2003)的10 a平滑序列 $\bar{X}$ 的拟合

Fig. 2 The comparison between actual figure with simulated one for  $\bar{X}$  of time series Rwh55 (1949~2003) by model (1) in Tab. 2

处于多雨期,演化建模得到的模拟曲线较好地反映了这一特点,并预测未来3 a(2004~2006)仍未有下降的趋势。图3为综合数学模型1和自然小波方法对Rwh55(1949~2003)拟合及未来3 a预测,显示了在较短周期影响因子的作用下,尽管在少雨宏观气候背景下,仍会出现汛期洪涝(1969);在多雨宏观气候背景下,仍会出现汛期偏旱(1985)。体现气候因子多时间尺度影响的综合作用,与实况较好地吻合。

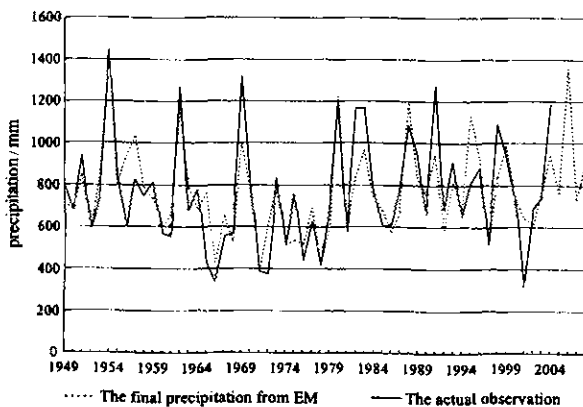


图3 (表2中)数学模型1与自然基小波对Rwh55(1949~2003)序列 $\bar{X} + X'$ 拟合及未来3 a预测,实线为实况值;虚线为拟合及预测值(单位: mm)

Fig. 3 Comparison between the predicted precipitation of  $\bar{X} + X'$  (dashed line) with the actual observation (solid line) of time series Rwh55 (1949~2003) by model (1) in Tab. 2

(3)汛期降水量宏观气候尺度周期波动随时间变化的规律表现为正弦、余弦形式的复杂周期函数,除数学模型表达式(5)之外,正弦、余弦函数中的自变量并非是我们预想的时间 $t$ ,而是汛期降水量 $x$ 自身。根据 $x$ 在不同的数值区间(汛期降水量 $x \neq 0$ ),确定了 $dx/dt$ 的符号,这就导致汛期降水量不会像以 $t$ 为自变量的正弦(余弦)函数那样,在某几年内持续地增大或持续地减小。而是呈不规则周期变化趋势,图2显示了这种不规则周期变化趋势,较好地吻合了实际观测数据序列特点。

(4)虽然这些数学模型仅仅显示数字序列变化规律,我们无法从中解释为什么汛期降水量随时间呈不规则周期变化的气候物理机制,但是它启示我们影响汛期降水量的宏观气候尺度变化的因子(包括我们已知和未知的)之间一定存在相互制约的关系。

为了考察演化建模方法的短期气候预测能力,仍取 $m = 55$ 序列段和同样的演化建模参数,我们对历史观测数据进行演化建模,由于资料的限制,表3中可以看到:在连续3个Rwh55序列预报结果中,虽然预测未来3 a的汛期降水量数值不同,但是它们的年际变化趋势是相同的。因此,使用演化建模方法可以了解降水量预测值在未来几年的变化趋势。

表3 连续3个Rwh55序列演化建模汛期降水量预测值及与实况比较(单位: mm)

Tab. 3 The comparison between the predicted precipitation of  $\bar{X} + X'$  with the observation values (mm)

演化建模的数据序列	2002	2003	2004	2005	2006
Rwh55(1947~2001)	预测值 631	846	1 058	811	1 407
	偏差百分率 -8%	+14%	-11%		
Rwh55(1948~2002)		预测值 794	945	733	1 311
		偏差百分率 +7%	-20%		
Rwh55(1949~2003)			预测值 939	783	1 366
			偏差百分率 -21%		
实况值	692	739	1 184		

## 2.4 与传统的数字时间序列线性建模原理的对比

上述的汛期降水量预测方法包括2步,以演化建模方法确定宏观气候尺度因子影响的主值序列部分 $\bar{X}$ ,采用自然基小波的方法模拟逼近微观气候尺度因子影响的扰动序列部分 $X'$ 。无论哪种方法都不苛求数字序列平稳、零均值。

文献 [1] 详尽叙述了线性建模方法 AR(p)、MA(q) 和 ARMA(p,q) 等的特性。它们是在线性最小方差意义下的平稳预报方法。其建模过程等价于对历史序列

$X = (x(t_0), x(t_1), x(t_2), \dots, x(t_{m-1}))^T$   $t_i = t_0 + i \times \Delta t$  ( $i = 0, 1, 2, \dots, m-1$ ) 进行相关函数及谱密度函数分析, 选择实系数  $\{c_0, c_1, c_2, \dots, c_{m-1}\}$ , 使预测值为观测值的线性组合, 并满足预测误差最小。以上述序列  $X$  做  $p = 5$  步预测, 它表示在时间序列  $X$  中任意一个数  $x(t_i)$  仅依赖于它前面的 5 个数的线性组合,

$$x(t_i) = c_1 x(t_{i-1}) + c_2 x(t_{i-2}) + c_3 x(t_{i-3}) + c_4 x(t_{i-4}) + c_5 x(t_{i-5})$$

5 个实系数  $c_1, c_2, c_3, c_4, c_5$  由最小均方 (least squares) 确定, 即

$$\sigma^2 = \frac{1}{(m-1-p)} \sum_{i=p}^{m-1} (x_i - c_1 x(t_{i-1}) + c_2 x(t_{i-2}) + c_3 x(t_{i-3}) + c_4 x(t_{i-4}) + c_5 x(t_{i-5}))^2$$

并求解方程组  $\frac{\partial \sigma^2}{\partial c_i} = 0$   $i = 1, 2, 3, 4, 5$  得到。

计算预测值  $X^*$  的递进公式为

$$\left. \begin{aligned} x^*(t_m) &= c_1 x(t_{m-1}) + c_2 x(t_{m-2}) + \dots + c_5 x(t_{m-5}), \\ x^*(t_{m+1}) &= c_1 x^*(t_m) + c_2 x(t_{m-1}) + \dots + c_5 x(t_{m-4}), \\ x^*(t_{m+2}) &= c_1 x^*(t_{m+1}) + c_2 x^*(t_m) + \dots + c_5 x(t_{m-3}), \\ &\vdots \end{aligned} \right\} (l < 5)$$

$$\begin{aligned} x^*(t_{m+l}) &= c_1 x^*(t_{m+l-1}) + c_2 x^*(t_{m+l-2}) + \dots + c_5 \\ &x^*(t_{m-l+5}), (l \geq 5) \end{aligned}$$

虽然这种方法简单, 也许在某些领域很有效, 但汛期降水量或者绝大多数其它气象要素的变化规律远非如此, 再者, 递进公式使  $X^*$  总是在  $X$  的平均值上下摆动, 很难表现出观测时间序列中出现的极端降水现象。

线性建模方法确定的模型参数只能确保方差最小, 从模型数学表达式中难以理解, 认识观测序列随时间的变化规律, 演化建模得到的非线性常微分方程是随机的, 多次计算得到的一批非线性常微分方程可以作为参考, 印证已知的规律或启发思考未知的规律。

### 3 结 论

(1) 针对汛期降水量这类具有混沌特征的数字

序列而言, 应以已有的气候学知识为依据, 合理地将它分解为宏观气候尺度的主值序列和微观气候尺度的扰动序列 2 部分。分别采用演化建模和自然基小波分解的方法模拟逼近, 可以得到较为令人满意的短期气候预测结果。

(2) 基于遗传算法的演化建模方法是模拟具有非线性特征复杂系统的有效方法, 它比线性建模更接近实际自然规律。从演化建模得到的非线性常微分方程可以作为参考, 开拓短期气候分析和预测的思路。

(3) 虽然单要素时间序列演化建模方法较之线性建模方法有其优越之处, 但毕竟它包含的历史信息量有限, 无法从模型中了解它与影响因子之间的统计关系, 这是它的局限性。但从计算技术上讲, 单要素演化建模方法可以推广到多要素演化建模。因为在气象问题中, 人们关注的是气象要素随时间变化的规律, 是对时间  $t$  的一阶微分方程。所以 (4) 式常微分方程组中的一个方程就可以表示一个变量随时间变化规律, 整个微分方程组就可以表示一个复杂系统中多个变量随时间变化规律, 而求解的遗传算法可以套用, 只是求解的定解条件变为多个变量在初始时刻的值。这正是本文作者正在研究的内容。

#### 参考文献:

- [1] 项静恬, 史久恩, 杜金观, 等. 动态和静态数据处理——时间序列和数理统计分析[M]. 北京: 气象出版社, 1991. 1125.
- [2] 陈兴芳, 赵振国. 中国汛期降水预测研究及应用[M]. 北京: 气象出版社, 2000. 241.
- [3] Koza J R. Genetic Programming II: Automatic Discovery of Reusable Programs[M]. Cambridge: MIT Press, 1994. 220.
- [4] Bäck T, Hammel U, Schwefel H P. "Evolutionary computation: comments on the history and current state." [J]. IEEE Trans Evol Comput, 1997, 1: 5-16.
- [5] Hongqing Cao, Lishan Kang, Yiping Chen. Evolutionary Modeling of Systems of Ordinary Differential Equations with Genetic Programming[A]. Genetic Programming and Evolvable Machines[M]. Kluwer Academic Publishers 2000. 309-337.
- [6] 林振山, 邓自旺. 子波气候诊断技术的研究[M]. 北京: 气象出版社, 1999. 174.
- [7] 康卓, 黄竟伟, 李艳, 等. 复杂系统数据挖掘的多尺度混合算法[J]. 软件学报, 2003, 14(7): 1229-1237.

(下转第 11 页)

