

东江流域降水时间序列的混沌特征分析^{*}

石教智¹, 陈晓宏¹, 林汝颜²

(1 中山大学水资源与环境系, 广东 广州 510275)

(2 广东省水利电力勘测设计研究院, 广东 广州 510170)

摘 要: 以广东省东江流域月降雨序列为例, 在介绍相空间重构原理的基础上, 探讨了混沌分析的主要定量指标: 饱和关联维数 D_2 和最大 Lyapunov 指数 λ 。得到该时间序列的饱和关联维数 $D_2 = 3.93$ 最小嵌入维数 $m = 8$ 最佳嵌入滞时 $\tau = 3$ 个月, 最大 Lyapunov 指数 $\lambda = 0.253$ 并且采用主分量方法进一步验证了该序列具有混沌特性, 指出该序列的预测时限不应超过 4 个月, 对此结论则用 ARMA (p q) 模型作了验证, 为东江流域月降雨预测提供了较为科学的依据。

关键词: 东江流域; 相空间重构; 关联维数; 主分量分析; Lyapunov 指数

中图分类号: P333.6 **文献标识码:** A **文章编号:** 0529-6579 (2006) 04-0111+05

变化条件下的水循环特性一直是水文学研究的热点和难点之一。而降雨径流作为水循环的重要特征, 对其规律的探索就显得非常必要。尤其是降雨, 在我国大部分地区, 作为径流的主要来源, 直接影响水资源的开发和利用, 紧密关联与水有关的干旱、洪涝等自然现象。然而, 由于降水是一种复杂的水文现象, 其形成、发生、发展的全过程因受众多因素的影响而表现出并非随机却貌似随机的特征, 传统数学模型难以定量描述^[1]。混沌理论的出现为研究这种复杂水文现象提供了新的思路, 使得从复杂的非线性水文现象中提取确定性特征成为可能。所以近年来, 混沌理论在水文科学领域的应用研究开始呈现快速发展势头。

本文以广东省东江流域月降雨序列为例, 从相空间重构、饱和关联维数和最大 Lyapunov 指数等方面探讨降水时间序列演化和形成的混沌机制, 并应用主分量分析方法进一步验证其混沌特性, 从而为混沌理论进一步应用于水文时间序列的分析、预测等奠定基础。

1 混沌理论

混沌理论的早期研究可以追溯到 1963 年美国气象学家 Lorenz^[2]对两无限平面间的大气湍流的模拟。在求解过程中, Lorenz 发现当方程中的参数取适当值时, 解是非周期的且具有随机性, 表现出非常复杂的结构, 也就是著名的“蝴蝶”效应。即

由确定性方程可以得出随机性的结果, 这与几百年来统治人们思想的拉普拉斯确定论相违背 (确定性方程得出确定性结果)。此后混沌理论蓬勃发展, 其应用已渗透到力学、物理学、水动力学、化学、生物学及生理学等许多学科, 被认为是继相对论和量子力学之后发展起来的又一革命性的成就之一。

对于混沌目前尚无通用的严格定义^[3], 一般认为, 将不是由随机性外因引起的, 而是由确定性方程 (内因) 直接得到的具有随机性的运动状态称为混沌, 即混沌是由确定性产生的貌似随机性的非周期行为。

常用表征系统混沌特性的量有: 连续功率谱、饱和分维数、Lyapunov 指数和 Kolmogorov 熵, 相应的关于混沌时间序列判别的基本方法有: 功率谱法、主分量分析 (PAC) 方法、Poincare 截面法、Lyapunov 指数法、c-c 方法等, 都是从某一个方面判别水文序列是否为混沌序列的必要条件, 因而需要采用尽可能多的方法来鉴别。下面结合广东省东江流域降雨资料, 从相空间重构、饱和关联维数和最大 Lyapunov 指数等方面进行降水时间序列的混沌特性分析。

2 东江流域降水时间序列的混沌分析

2.1 资料来源

东江属珠江三大水系之一, 位于中国南方湿润

^{*} 收稿日期: 2005-08-05

基金项目: 国家自然科学基金资助项目 (50579078); 广东省自然科学基金资助项目 (04009805)

作者简介: 石教智 (1974 年生), 男, 工程师, 博士研究生; 通讯联系人: 陈晓宏; E-mail: eesxh@zsu.edu.cn

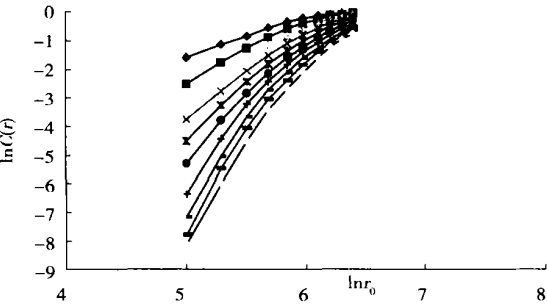


图 2 $\ln r_0 \sim \ln C(r)$ 关系图
Fig.2 Relationship between $\ln r_0$ and $\ln C(r)$

从图 2 看出，不同嵌入维 m 下， $\ln r_0 \sim \ln C(r)$ 曲线图中，存在直线相关的部分，即存在无标度区。因此东江流域月降雨系列分布具有分形特征，每一条曲线中的直线段部分的斜率，就是各自嵌入维 m 所对应的关联维数。绘制由此得到的关联维数 D_2 与不同嵌入维 m 之间的关系图，见图 3。

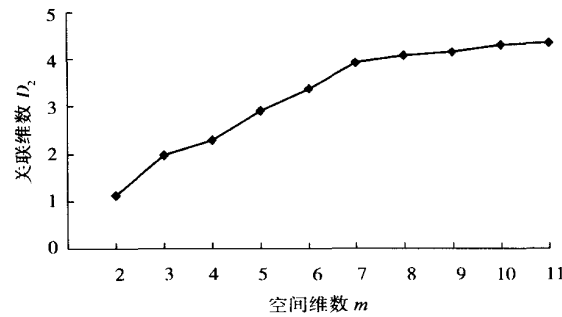


图 3 $m - D_2$ 关系图
Fig.3 Relationship between m and D_2

由图 3 分析，当嵌入维数 $m = 8$ 时，关联维数趋于稳定，达到一个饱和值，又称之为饱和关联维，即 $D_2 = 3.93$ 。与之对应的最小嵌入维数表征了动力系统的有效自由度维数。可见，东江流域月降雨序列具有混沌特性，当嵌入空间达到 8 维后，系统具有稳定的混沌吸引子维数 3.93。下面采用主分量分析法作进一步验证。

2.4 主分量分析（PAC 分布）

主分量分析（Principal Component Analysis）方法^[8]是一种能有效识别混沌和噪声的方法。具体计算步骤如下：若已知给定的时间序列为 $x_1, x_2, \dots, x_n$ ，选定滞时 τ 和嵌入维数 m ，用 2.2 所述方法进行相空间重构（符号意义同），则可得到序列轨线矩阵 $Y_{l \times m}$ ：

$$Y_{l \times m} = \frac{1}{\sqrt{l}} [Y_1, Y_2, \dots, Y_l]^T \quad (8)$$

计算协方差矩阵 A 为：

$$A_{m \times m} = \frac{1}{l} Y_{l \times m}^T Y_{l \times m} \quad (9)$$

然后计算协方差矩阵 A 的特征值 λ_i ($i = 1, 2, \dots, m$) 和相应的特征向量 U_i ($i = 1, 2, \dots, m$)。将特征值按大小排列： $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ 。则特征值 λ_i 及特征向量 U_i 称为主分量。求出所有特征值的和 γ ($\gamma = \sum_{i=1}^m \lambda_i$)。

以指标 i 为 X 轴， $\ln(\lambda_i / \gamma)$ 为 Y 轴得到主分量谱图。由于混沌信号和噪声的主分量分布之间存在着显著差异，即噪声的主分量谱图应是一条与 X 轴接近平行的直线，而混沌信号的主分量谱图应是一条过定点且斜率为负的直线（或含有类似直线部分）。

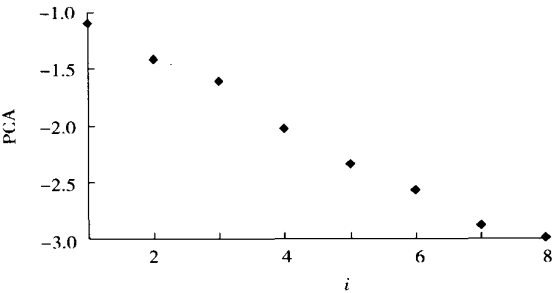


图 4 东江流域月降雨系列主分量谱图
Fig.4 Principal component spectrum of monthly rainfall serial in the Dongjiang River Valley

根据以上方法，绘制东江流域月降雨系列主分量谱图，见图 4。该图中存在斜率为负的直线部分，且与 Duffing 方程的主分量图相似，从而进一步确认该时间序列具有混沌特性，可以应用混沌理论建立月降雨预测模型。

2.5 最大 Lyapunov 指数

Lyapunov 指数是一个谱系 λ_i ($i = 1, 2, \dots, m - 1$)，若谱系中只有最大者为正值，则系统为一维混沌；若谱系中具有两个或两个以上正的 Lyapunov 指数，则系统为多维混沌或超混沌，否则系统为非混沌。因此，一般用最大 Lyapunov 指数来进行混沌分析。最大 Lyapunov 指数是判断混沌存在和混沌特性的一个重要依据，可以度量混沌系统对初值的敏感程度，即其值越大，混沌特性越强，对初值的敏感性也越强；反之敏感性越弱。而正由于混沌系统的初值敏感性，混沌系统的长期预测是不可能的。因此最大 Lyapunov 指数的倒数可作为系统的可预报尺度。对于时间序列 $x_1, x_2, \dots, x_n$ ，其求取方法如下^[11-12]：

先按 2.2 所述方法进行相空间重构，然后根据式（10）求得欧氏空间意义上 Y_i 的最近邻点 Y_j 及两点间距离 L_{ij} ，最后根据式（11）求最大 Lyapunov 指数 λ 。

$$L_i = \min_{j \neq i} [\|Y_j - Y_i\|], \quad i, j = 1, 2, \dots, l \quad (10)$$

$$\lambda = \frac{1}{l-1} \frac{1}{\tau} \sum_{i=1}^{l-1} \ln \frac{L_{i+1}}{L_i} \quad (11)$$

根据以上方法，求得东江流域月降雨系列的最大 Lyapunov 指数为 0.253，从而进一步说明了该水文时间序列存在混沌特性。 λ 表征了该水文系统的可预报长度大约为 4 个月。为了验证此结论，本

文对 1990 – 1999 年共 10a 的历年逐月雨量系列 $\{x_t \mid t=1, 2, \dots, 120\}$ 建立如下的 ARMA（2，1）模型：

$$x_t = 146.6529 - 3.5748x_{t-1} + 2.0767x_{t-1} + 2.0767x_{t-2} + a_t + 0.2467a_{t-1} \dots \dots \quad (12)$$

并用该模型对 2000 年的逐月资料进行验证。记 2000 年逐月实测资料为 $\{x_t \mid t=1, 2, \dots, 12\}$ ，预测系列为 $\{y_t \mid t=1, 2, \dots, 12\}$ ，先输入 1 月和 2 月的实测雨量值 x_1, x_2 ，利用式（12）预测 3 月份的雨量值 y_3 ，再利用 x_2, y_3 预测 4 月份的雨量值 y_4 ，依次类推，得表 1。

表 1 雨量值计算表

Tab. 1 Rain fall calculation

	1	2	3	4	5	6	7	8
实测值 x	69.12	51.71	67.39	490.66	118.84	260.58	239.83	296.14
预测值 y			62.00	441.60	99.83	218.89	196.66	230.96
$ y-x /\Delta$			0.08	0.10	0.16	0.16	0.18	0.22

由表 1 可见，随着预测期的增大，预测精度在降低，预测期为 4 个月的时候预测误差已接近 20%。该水文系统不能进行精度较高的长期预报。

3 结 语

混沌是一个内涵和外延十分丰富的概念，其本身处于快速发展阶段，尚有不少理论和实际计算问题亟代解决^[1]。目前应用混沌理论研究水文时间序列问题也存在争议。本文应用相空间重构原理，从饱和关联维数和最大 Lyapunov 指数等方面探讨广东省东江流域月降雨时间序列形成和演化的混沌机制，并应用主分量分析方法进一步验证其混沌特性。最后得到其嵌入滞时 $\tau=3$ ，饱和关联维数 $D_2=8$ ，最大 Lyapunov 指数 $\lambda=0.253$ 。根据最大 Lyapunov 指数的倒数可知该降水时间序列预测时限大约为 4 个月，不能进行精度较高的长期预报，并对此结论用 ARMA（ p, q ）模型作了验证。

参考文献:

[1] 王红瑞, 宋宇, 刘昌明, 等. 混沌理论及在水科学中的应用与存在的问题[J]. 水科学进展, 2004 15(3): 400 – 407

[2] LORENZ E. N. Deterministic Nonperiodic Flow [J]. Atmospheric Sci 1963(20): 29 – 34

[3] 唐巍, 李殿璞, 陈学允. 混沌理论及其应用研究[J]. 电力系统自动化, 2000(4): 67 – 70

[4] 丁涛, 周惠成, 黄健辉. 混沌水文时间序列区间预测研究[J]. 水利学报, 2004(12): 15 – 20

[5] UPMANU L. TA IYE B. Nonlinear dynamics of the Great Salt Lake: nonparametric short term forecasting [J]. Water Resource Research 1996 32(4): 975 – 985

[6] PACKARD N. H. CRUTCHFIELD J. B. FARMER J. D. et al. Geometry from a time series[J]. Phys Rev Lett 1980 45(9): 712 – 716

[7] ROSSLER O. E. An Equation for Continuous Chaos[J]. Phys Lett 1976(57): 13 – 15.

[8] 袁鹏, 李谓新, 王文圣, 等. 月降雨量时间序列中的混沌现象[J]. 四川大学学报: 自然科学版, 2002 34(1): 16 – 19

[9] 王海燕, 盛昭瀚. 混沌时间序列相空间重构参数的选取方法[J]. 东南大学学报, 2000 30(5): 113 – 117.

[10] 丁晶, 王文圣, 赵永龙. 长江日流量混沌变化特性研究[J]. 水科学进展, 2003 14(4): 407 – 416

[11] 窦春霞. 基于混沌—神经网络模型最优控制及应用[J]. 系统工程学报, 2004 19(3): 229 – 233

[12] LAI Y. C. LERNER D. Effective scaling regime for computing the correlating dimension for chaotic time series[J]. Physica D 1998(11): 10 – 18

Chaotic Characteristics of Precipitation Time Series in the Dongjiang River Valley

SHI Jiao-zhi¹, CHEN Xiao-hong¹, LIN Ru-yan²

(1. Department of Water Resources and Environmental Sciences

Sun Yat sen University, Guangzhou 510275 China

2. Guangdong Provincial Investigation Design and Research Institute of Water Conservancy and Electric Power, Guangzhou 510170 China)

Abstract Based on introducing the phase space reconstruction theory of chaotic time series, the main quantitative indexes of saturation correlation dimension D_2 and maximal Lyapunov exponent λ for chaotic analysis are discussed with the monthly rainfall time series of the Dongjiang River Valley in Guangdong province. The saturation correlation dimension, minimum embedding dimension, optimal built-in delay time and maximal Lyapunov exponent are calculated and given, that is $D_2=3.93$, $m=8$, $\tau=3$ and $\lambda=0.253$. On the basis of this primary component analysis method is applied to validate its chaotic character. It is suggested that the forecasting length for this rainfall time series should not exceed 4 months and this is verified with the model of ARMA. This time series chaotic analysis will provide a scientific gist for monthly rainfall forecasting in the Dongjiang River Valley.

Key words the Dongjiang River Valley; phase space reconstruction; saturation correlation dimension; primary component analysis; Lyapunov exponent