第29卷第1期
2007年3月

地球科学与环境学报
Journal of Earth Sciences and Environment

No. 1 Vol. 29
Mar. 2007

基于混沌时间序列的地下水位多步预测模型

陈南祥[1]，张海丰[1]，李松海[2]

（1. 华北水利水电学院 岩土工程系，河南 郑州 450008；2. 黄河勘测规划设计有限公司，河南 郑州 450011）

摘要：利用相空间重构技术，并借助G-P算法、C-C方法和Wolf方法从宁陵地区地下水位一维时间序列中提取Lyapunov指数，结果表明此时间序列具有混沌特征。计算了宁陵地区地下水位时间序列的关联维数、时间延迟和最大Lyapunov指数，将局域加权一阶多步预测模型应用于地下水位预测。预测表明，此模型可有效应用于地下水位时间序列的多步预测。

关键词：时间序列；相空间重构；混沌；地下水位；Lyapunov指数；关联维数

中图分类号：P641.1；TV124　文献标志码：A　文章编号：1672-6561(2007)01-0066-04

Multi-Steps Prediction Model of Underground Water Table Based on Chaotic Time Series

CHEN Nan-xiang[1], ZHANG Hai-feng[1], LI Song-hai[2]

(1. *Department of Geotechnical Engineering, North China Institute of Water Conservancy and Hydroelectric Power, Zhengzhou* 450008, *Henan, China*; 2. *Yellow River Engineering Consulting Company, Zhengzhou* 450011, *Henan, China*)

Abstract: Applying phase space reconstruction method, G-P arithmetic, C-C arithmetic and Wolf method, this paper distills Lyapunov exponents from one dimension time series of underground water table in Ningling county. The result indicates that this time series possesses the character of chaos. The correlation dimension of time series, time delay and the largest Lyapunov exponent of underground water table in Ningling county are calculated. The add-weighted one-rank multi-steps prediction model is developed for the prediction of underground water table. The prediction indicates that this model can be easily used in multi-steps prediction of underground water table time series.

Key words: time series; phase space reconstruction; chaos; underground water table; Lyapunov exponent; correlation dimension

0 引言

从时间序列研究混沌，始于Packard[1]提出的重构相空间理论。对于决定系统长期演化任一变量的时间演化，均包含了系统所有变量长期演化的信息。近年来，混沌理论、分形理论以及非线性预测方法的发展为研究地下水动态提供了新的思路。通过决定系统长期演化的任一单变量时间序列来研究系统的混沌行为，使基于混沌时间序列的预测成为可能。

地下水文系统可以通过数值模拟求解渗流模型来确定地下水动态变化规律。由于这种方法要求具备较多的水文地质参数和实测数据，加之影响地下水动态的因素存在着很大的时空变异，因而对于一些资料缺乏、时空变异较大的研究区域来说，该方法很难获得满意的预测结果。随机性模型成本小、周期短，适用于长序列的地下水动态模拟预报，因此得到了不断发展。但由于地下水动态时间

收稿日期：2006-04-01
基金项目：河南省杰出人才创新基金项目(0421000500)
作者简介：陈南祥(1958-)，男，江苏张家港人，教授，从事水文地质与水文水资源研究。E-mail: chennanxiang@ncwu.edu.cn

序列常常是非平稳的或非线性的, 目前常用的随机模型大多是基于平稳序列或线性序列建模, 效果不甚理想。地下水文系统实质上是一个受水文地质条件控制, 并受降雨、气候和人类工程活动等多种因素影响而发展演化的非线性耗散动力系统[2]。

由于混沌系统内在的有序性和规律性, 依据地下水位的历史数据, 利用重构相空间方法, 在一定时间内对地下水位的预测不但是可能的, 而且可能比基于一般统计方法的预测更好。在一定的时间尺度内, 这些要素的变化呈现出一定的规律性, 可以对地下水动态进行分析、预测。

1　相空间重构

对于时间序列 $x_1, x_2, \cdots, x_n$, N 是序列的长度。根据 G-P(Grassberger-Procaccia) 算法[3] 计算出时间序列的关联维 D, 再由 Takens 定理选取嵌入维数 $m=2D+1$, 然后用 G-C 方法对时间序列进行分析得到时间延迟 τ[4] (即观测数据时间间隔的 m 倍), 从而得到重构相空间

$$Y_i = (x_i, x_{i+\tau}, \cdots, x_{i+(m-1)\tau}) \quad i = 1, 2, \cdots, M \tag{1}$$

式中: M 是重构相空间中相点的个数, $M = N - (m-1)\tau$。

1.1　关联维数的计算

在实际应用中, 事先不知道吸引子的维数, 所以也无法知道吸引子的嵌入维数。1983 年, Grassberger-Procaccia 提出了一种计算嵌入维数的方法, 即 G-P 算法: 定义关联积分函数 $C_m(r)$

$$C_m(r) = \lim_{N\to\infty} \frac{1}{N} \sum_{i,j=1}^{N} H(r - \| Y_i - Y_j \|) \tag{2}$$

式中: Y_i 为相空间中的相点; r 为给定的距离; H 为 heaviside 函数, 即

$$H(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leqslant 0 \end{cases} \tag{3}$$

在给定 r 后, $C(r)$ 表示两点间距小于 r 的概率, 在适当的范围内, 当 r 增加时, $C_m(r)$ 将以 r^{D_m} 的比率增加, 即

$$C_m(r) \propto r^{D_m} \tag{4}$$

两端同时取对数后

$$\ln C_m(r) \propto D_m \ln r \tag{5}$$

所以选取若干个不同的 r, 分别计算相应的 $C_m(r)$, 将这些不同的 r 和 $C_m(r)$ 代入式(5), 即可拟合出关联维数 D_m。

1.2　时间延迟 τ 的选取

时间延迟 τ 的选取采用 G-C 方法, 它是计算时间延迟的一种总有效方法[5], 则定义嵌入时间序列的关联积分为

$$C(m, N, r, t) = \frac{2}{M(M-1)} \cdot \sum_{1\leqslant i\leqslant j\leqslant M} H(r - \| Y_i - Y_j \|) \quad r > 0 \tag{6}$$

式中: m 是嵌入维数; $M = N - (m-1)\tau$; Y_i 为相空间中的相点; r 为给定的距离; H 为 heaviside 函数。

对于一般的时间序列, 将其分成 t 个不相交的子序列, 然后定义每一个子序列的 $S(m, N, r, t)$ 为

$$S(m, N, r, t) = \frac{1}{t}\sum_{s=1}^{t}\left[C\left(m, \frac{N}{t}, r, t\right) - C_s^m\left(1, \frac{N}{t}, r, t\right)\right] \tag{7}$$

令 $N\to\infty$, 且选择对应值最大和最小两个半径 r, 定义差量为

$$\Delta S(m, t) = \max[S(m, r_j, t)] - \min[S(m, r_j, t)] \quad m = 2, 3, \cdots \tag{8}$$

根据 BDS[6] 的统计结论计算

$$\Delta\overline{S}(t) = \frac{1}{4}\sum_{m=2}^{5}\Delta S(m, t) \tag{9}$$

依据上述计算结果作图, 其中 $\Delta\overline{S}(t)$ 的第一个极小值时间 t 为该时间序列的延迟。

1.3　Lyapunov 指数计算

利用 Wolf[4] 方法从单变量的时间序列提取 Lyapunov 指数的方法仍然是基于时间序列的重构相空间。$Y(t_0)$ 为混沌时间序列的初始点, 设其最邻近点 $Y_0(t_0)$ 的距离为 L_0, 追踪这两点的时间演化, 直到 t_1 时刻, 其间距超过某规定值 $r > 0$, $L_0' = |Y(t_1) - Y_0(t_1)| > r$, 并在 $Y(t_1)$ 邻近另找一点 $Y_1(t_1)$, 使得 $L_0' = |Y(t_1) - Y_0(t_1)| < r$, 并且与其夹角尽可能的小, 继续上述过程, 直至 $Y(t)$ 到达时间序列的终点 N, 这时追踪演化过程总的迭代次数为 M, 则最大的 Lyapunov 指数为

$$\lambda_1 = \frac{1}{t_M - t_0}\sum_{i=0}^{M}\ln\frac{L_i'}{L_i} \tag{10}$$

2　加权一阶局域法一步多步预测模型

2.1　加权一阶局域法一步预测模型

设中心点 Y_M 的邻近点为 Y_{Mi}, $i=1, 2, \cdots, q$, 并且到 Y_M 的距离为 d_i, 设 $d_{\min}$ 是中的最小值, 定义点 Y_{Mi} 的权值为

$$P_i = \frac{\exp[-(d_i - d_{\min})]}{\sum_{i=1}^{q}\exp[-(d_i - d_{\min})]} \quad (11)$$

则一阶局域线性拟合为

$$Y_{Mi+1} = ae + bY_{Mi} \quad i = 1, 2, \cdots, q \quad (12)$$

式中：a、b 为拟合所需的实系数；e 为 q 维向量，$e=(1, 1, \cdots, 1)^{\mathrm{T}}$；$Y_{Mi+1}$ 是 Y_{Mi} 演化一步后的相点。

当嵌入维数 $m=1$ 时，应用加权最小二乘法有

$$\sum_{i=1}^{q} P_i (x_{Mi+1} - a - bx_{Mi})^2 = \min \quad (13)$$

将式(13)看成 a、b 的二元函数，两边求偏导即可得到方程组系数 a、b，将 a、b 代入一步预测公式 $Y_{M+1}=ae+bY_M$，即可得到演化一步后的相点预测值

$$Y_{M+1} = (x_{M+1}, x_{M+1+\tau}, \cdots, x_{M+1+(m-1)\tau}) \quad (14)$$

这里，Y_{M+1} 中前$(m-1)$个元素为原序列中已知值，其第 m 个元素 $x_{M+1+(m-1)\tau}$ 即为原序列的一步预测值 $\hat{x}_{N+1}$。此即加权一阶局域法一步预测模型[5]。

2.2 加权一阶局域法多步预测模型

加权一阶局域法一步预测模型的实质是在重构的相空间中找到与参考点最相似的$(m+1)$个相点，并根据这$(m+1)$个相点演化一步的规律进行一步预测[7]。

对于混沌时间序列，相空间中一对最近邻点随时间演化遵循的是一种指数规律 $e^{-\lambda}$，参数 λ 是最大 Lyapunov 指数，体现了系统初始闭轨道的指数发散速率。因此，当嵌入维数 $m>1$，且需进行 $k(k>1)$ 步预测时，可以类似加权一阶局域法一步预测模型，根据这$(m+1)$个相点演化 k 步的规律进行 k 步预测。具体推导步骤为：

设中心点 Y_M 的参考向量集$\{Y_{Mi}\}$，$i=1, 2, \cdots, q$，其演化 k 步后的相点集为$\{Y_{Mi+k}\}$，一阶局域线性拟合为

$$Y_{Mi+k} = a_k e + b_k Y_{Mi} \quad i = 1, 2, \cdots, q \quad (15)$$

根据加权最小二乘法有

$$\sum_{i=1}^{q} P_i \left[\sum_{j=1}^{m} (x^j_{Mi+k} - a_k - b_k x^j_{Mi})^2\right] = \min \quad (16)$$

将式(16)看成是关于未知数 a_k、b_k 的二元函数，两边求偏导得

$$\begin{cases} \sum_{i=1}^{q} P_i \sum_{j=1}^{m} (x^j_{Mi+k} - a_k - b_k x^j_{Mi}) = 0 \\ \sum_{i=1}^{q} P_i \sum_{j=1}^{m} (x^j_{Mi+k} - a_k - b_k x^j_{Mi}) x^j_{Mi} = 0 \end{cases} \quad (17)$$

根据式(17)求得 a_k、b_k，代入 k 步预测公式 $Y_{M+k}=a_k e+b_k Y_M$，即可得到演化 k 步后的相点预测值

$$Y_{M+k} = (x_{M+k}, x_{M+k+\tau}, \cdots, x_{M+k+(m-1)\tau}) \quad (18)$$

这里，Y_{M+k} 中的第 m 个元素 $x_{M+k+(m-1)\tau}$，即为原序列的 k 步预测值 $\hat{x}_{N+k}$。

3 实例计算

以河南省商丘市宁陵县的地下水位时间序列为例，建立地下水位数学模型。宁陵县在黄河下游冲积平原区，位于东经 115°15′～115°30′，北纬 34°14′～34°37′。浅层潜水含水层以中细砂为主，富水性中等。地下水主要接受大气降水补给和地表水入渗补给，消耗主要是人工开采。地下水位动态比较复杂，具有周期性变化规律，每年最高水位在雨季后期的 9 月份出现，最低水位在 12 月份到翌年 1 月份之间，水位年变幅 0.20～3.00 m。

根据灌区 7 号地下水长观孔 1982～2002 年 5 日观测 1 次的地下水位序列共 21 年资料构成单变量时间序列。图 1 为地下水位对应的给定 $m=3$, 4, 5, …, 时，$\ln C(r)$ 和 $\ln r$ 之间的变化关系。图 2 为地下水位对应的分维数 D 与嵌入维数 m 之间的变化关系。

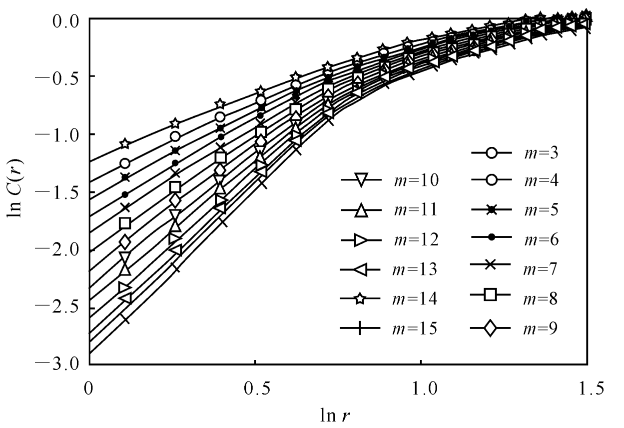

图 1 7 号长观孔地下水位 $\ln C(r)$-$\ln r$ 曲线

Fig. 1 ln C(r)-ln r Curve of Underground Water Table of the 7th Long Term Observation Pole

由图 2 可以得出，地下水位时间序列在 $m=7$ 以后关联维数不再随嵌入维数增大而增大，这时所求的嵌入维数即是所求的饱和嵌入维数。因此，宁陵 7 号长观孔的地下水位时间序列可以嵌入到 7 维状态空间向量序列。根据 C-C 方法求得时间延迟 $\tau=9$；Wolf 方法求得的最大 Lyapunov 指数 $\lambda_1=0.0975$，所以 7 号井地下水位时间序列为一混沌时间序列。这里用加权一阶局域法多步预测模型对

2002 年 1 ~3 月的地下水位进行 18 步预测(表 1)。

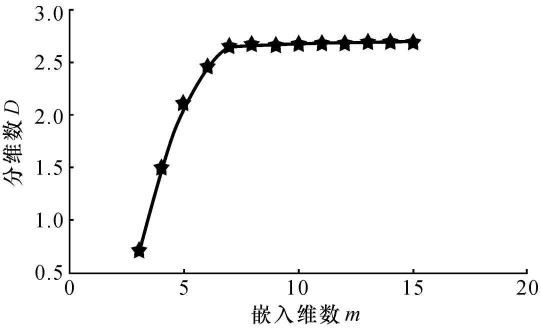

图 2　7 号长观孔地下水位 D- m 曲线

Fig. 2　D- m Curve of Underground Water Table of the 7th Long- Term Observation Pole

表 1　2002 年 7 号长观孔地下水位的实际观测值及预测值

Tab. 1　Practise Observation Value and Prediction Value of Underground Water Table of the 7th Long Term Observation Pole in 2002

序号	实测值 /m	预测值 /m	绝对误差 /m
1	44. 47	44. 51	- 0. 04
2	44. 48	44. 54	- 0. 06
3	44. 62	44. 58	0. 04
4	44. 64	44. 60	0. 04
5	44. 66	44. 62	0. 04
6	44. 72	44. 56	0. 06
7	44. 74	44. 55	0. 19
8	44. 70	44. 55	0. 15
9	44. 82	44. 62	0. 20
10	44. 47	44. 67	- 0. 20
11	44. 48	44. 46	- 0. 18
12	44. 62	44. 70	- 0. 08
13	44. 64	44. 71	- 0. 07
14	44. 66	44. 71	- 0. 05
15	44. 72	44. 73	- 0. 01
16	44. 74	44. 70	0. 04
17	44. 70	44. 70	0. 00
18	44. 82	44. 71	0. 11

从表 1 可知, 误差能够满足预测要求, 说明混沌时间序列作短期多步预测是可行的, 并且预测精度较高。从误差数据的大小分布来看, 前 5 步的预测精度都非常高, 但随着预测时间和步长的增多, 预测误差越来越大, 这个结论与混沌时间序列的本质特征是相符的。因为, 一方面混沌时间序列的确定性表明许多混沌现象短期是可以预测的; 另一方面, 混沌现象所固有的对初值敏感性又意味着预测能力受到根本性限制。混沌现象长期是不能预测的。

4　结语

加权一阶局域法多步预测模型是基于一步预测模型的近一步推广, 预测结果表明, 多步预测模型预测精度较高。但同时可以看出混沌系统的耗散性和对初值的极端敏感性, 在短期内精度高, 随着预测时间的增长, 预测精度逐渐下降。同时正的最大 Lyapunov 指数存在表明: 地下水位观测序列中除了通常认为的趋势性、周期性及随机扰动性成分外, 还存在较为明显的混沌成分。混沌系统的不可长期预测性决定了地下水位的不可长期预测。不同地下水位序列对应不同的最大 Lyapunov 指数, 所以不同的序列存在不同的最大预测步长。尽管可以进行多步预测, 但预测长度仍是非常有限的; 基于混沌成分的存在, 地下水位的预测方法研究不能局限于原有的趋势性、周期性外延方法的研究上, 应该考虑混沌理论与其他非线性动力系统的相关理论进行结合, 这应该是一个积极可行的研究方向。

参考文献:

[1]　Packard N H. Geometry from a Time Series[J] . Phy sica Rev Lett, 1980, 45(9): 712 716.

[2]　宋　宇, 陈家军, 孙　雄. 地下水水位时间序列中的混沌特征[J] . 水文地质工程地质, 2004, 31(1): 14 18.

[3]　Grassberger P, Procaccia I. Measuring the Strangeness of Strange Attractors [J] . Physics D, 1983, 9: 189 208.

[4]　Wolf A, Swinney H L, Vastano J A. Determining Lyapunov Exponents fom Time Series[J] . Physics, 1985, 16(D): 285 317.

[5]　陶诏灵, 陈国华. 基于 C C 方法的 Lyapunov 指数计算[J] . 南京气象学院学报, 2002, 25(4): 555 559.

[6]　吕金虎, 陆君安, 陈士华. 混沌时间序列分析及其应用[M] . 武汉: 武汉大学出版社, 2002.

[7]　丁　涛, 周惠成. 混沌时间序列局域预测方法[J] . 系统工程与电子技术, 2004, 26(3): 338 340.