

· 水文地质工程地质 ·

地质时间序列的频谱分析方法及应用^{*}

杨永国 韩宝平 (中国矿业大学资源与环境科学学院 徐州 221008)

摘要 介绍了地质时间序列随机模型及研究时间序列相关关系的频谱分析方法,给出了显著周期检验的 Fisher 准则。研究结果表明,频谱分析能够很好地确定地质时间序列的隐含周期及相关关系,从而揭示地质时间序列的内部特征。

关键词 时间序列分析 频谱分析 互功密度谱 迟延时间

中国图书资料分类法分类号 P628.1

作者简介 杨永国 男 35岁 副教授 硕士 数学地质

1 引言

地质时间序列由于受各种因素的影响,其变化受各种影响因素的控制。要完全揭示或量化其内在变化的机理,目前尚有困难。确定性数学方法,如解析法、模拟法、有限元法等的使用,在许多方面已取得一定的成果。但由于确定性数学法建立在众多的假设前提下,使其应用效果受到了很大影响。

大量的事实和实践证明,人们不能精确地描述地质现象的发生和内在机理。因此目前较多的是从时间序列本身的研究出发,采用随机的数学方法,对其变化规律作出定量的分析,从而对其内在机理作出一定程度的估计,以指导实际工作。

2 地质时间序列的随机模型

设 $\{y_t, t=1, 2, \dots, N\}$ 为地质观测数据,它是按等时间隔观测记录得到的离散数据序列。一般地质时间序列可看作由四种要素组成:趋势因子 y_t^T , 季节因子 y_t^C , 周期因子 y_t^S , 不规则因子 y_t^I 。依具体的组合模式有两种模型:

$$\text{加法模型: } y_t = y_t^T + y_t^C + y_t^S + y_t^I; \quad (1)$$

$$\text{乘法模型: } y_t = y_t^T \cdot y_t^C \cdot y_t^S \cdot y_t^I. \quad (2)$$

由于乘法模型经过对数变换可化为加法模型,因此不失一般性,只须讨论加法模型即可。

在地质时间序列的分解要素中,趋势因子表征序列的长期发展趋势及潜在水平;季节因子是季节变动影响的短期周期性波动;周期因子反映了序列长期变化过程中的周期性变化;不规则因子是各种偶然因素造成的随机扰动。通过对地质时间序列的频谱分析,可以辨识它们的周期因子及它们之间的相互关系。

频谱分析方法把时间序列看作是无数具有随机振幅和相位的周期振荡的叠加。为了描述一个平稳时间序列的统计特征,在时域分析中是研究其协方差(或相关)函数:

$$R(k) = E\{y_{t+k} - \mu(y_t - \mu)\}. \quad (3)$$

式中 $\mu = E(y_t)$ 为均值函数。在频谱分析中是考察功率谱密度函数(即功率谱) $S(\omega)$, 功率谱与协方差函数通过富里叶变换建立 1-1 对应关系:

$$R(k) = \int_{-\pi}^{\pi} S(\omega) e^{i\omega k} d\omega, k=0, \pm 1, \pm 2, \dots; \quad (4)$$

$$S(\omega) = \frac{1}{2\pi} \int_{-\pi}^{\pi} R(k) e^{-i\omega k} dk = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} R(k) e^{-i\omega k}, \quad (5)$$
$$-\pi \leq \omega \leq \pi$$

协方差函数和功率谱分别是在时域分析和频域分析中对时间序列统计特性的不同描述方法,而功率谱对于分析时间序列的隐含周期和时间序列之间的相关关系是十分有力的工具。

3 地质时间序列频谱分析原理

3.1 地质时间序列的自谱分析

设时间序列的数学模型为:

^{*} 煤炭科学基金资助项目

$$y_t = \sum_{i=1}^k c_i \cos(2\pi f_i t + \varphi_i) + a_t. \quad (6)$$

式中 k, c_i, f_i —— 常数;

φ —— 在 $(-\pi, \pi)$ 内均匀分布的独立随机变量;

a_t —— 独立于 φ 的纯随机过程, 且

$$E(a_t) = 0, E(a_t^2) = \sigma_a^2.$$

(6) 式利用三角函数展开则有:

$$y_t = \sum_{i=1}^k (a_i \cos 2\pi f_i t + b_i \sin 2\pi f_i t) + a_t. \quad (7)$$

式中 $a_i = c_i \cos \varphi, b_i = -c_i \sin \varphi$, 为待估参数;

$\varphi = \arctg(-b_i/a_i)$ 。如果已有 N 个观测值 $\{y_t, t=1, 2, \dots, N\}$, 用最小二乘法将

$$Q = \sum_{t=1}^N \{y_t - \sum_{i=1}^k (a_i \cos 2\pi f_i t + b_i \sin 2\pi f_i t)\}^2 \quad (8)$$

极小化来估计相应的 a_i 和 b_i , 结果为:

$$a_i = \frac{2}{N} \sum_{t=1}^N y_t \cos 2\pi f_i t, b_i = \frac{2}{N} \sum_{t=1}^N y_t \sin 2\pi f_i t. \quad (9)$$

定义 N 个观测序列的周期图为:

$$I_N(f) = \frac{2}{N} \left| \sum_{t=1}^N y_t e^{-2\pi i f t} \right|^2, (-\frac{1}{2} \leq f \leq \frac{1}{2}). \quad (10)$$

等价地有:

$$I_N(f) = [A(f)]^2 + [B(f)]^2 \quad (-1/2 \leq f \leq 1/2); \quad (11)$$

$$A(f) = \frac{2}{N} \sum_{t=1}^N y_t \cos 2\pi f t; \quad (12)$$

$$B(f) = \frac{2}{N} \sum_{t=1}^N y_t \sin 2\pi f t; \quad (13)$$

实际计算中 $I_N(f)$ 的频率取值是离散的值: $0, 1/N, 2/N, \dots$ 。故记 $I_p = I_N(f_p), f_p = p/N, p = 0, 1, \dots, [N/2], I_p = [A(f_p)]^2 + [B(f_p)]^2$ 。当 f_p 的值等于 y_t 中周期分量的频率 f_i 时, 周期图有大的峰值。

3.2 地质时间序列的互功率谱分析

设 $\{x_t, t=1, 2, \dots, N\}$ 和 $\{y_t, t=1, 2, \dots, N\}$ 为两组不同的地质时间序列, x_t 为输入, y_t 为输出, 其中 y_t 为 x_t 的响应, 两者构成一个系统。如果系统的所有基本性质不随时间变化且响应是线性的, 则称为常数系数线性系统。常数系数线性系统不引起任何频率变

换, 只改变输出的振幅和相位。本文利用互功率谱来考察两个时间序列的频率响应特征。互功率谱密度函数是互相关函数的富里叶变换, 一般互功率谱密度函数是一复数, 记为:

$$S_{xy}(f) = R_{xy}(f) + I_{xy}(f)i. \quad (14)$$

式中 $R_{xy}(f), I_{xy}(f)$ —— 互功率谱的实部和虚部。

y_t 在频率 f 处相对 x_t 滞后的相位角为:

$$\varphi_{xy}(f) = -\arctg(I_{xy}(f)/R_{xy}(f)). \quad (15)$$

当谱峰较显著时, 可以认为这是两个时间序列共有的且呈响应的周期谱峰。

3.3 隐含周期的检验

虽然一般总可以从周期图上找到一些峰值, 但还不能立即判定哪些对应着 x_t 的真正周期分量, 因为甚至在 $c_i = 0$ 的情况下, 也会因随机采样引起周期纵坐标的波动而出现虚假的峰值。为此, 需用统计检验方法来确定所出现的峰值是否比 $c_i = 0$ 时情况要显著得多。

先检验其中的最大峰值。作原假设 H_0 : 不存在隐含周期。作统计量

$$g = \max(I_p) / \sum_{p=1}^{[N/2]} I_p, \quad (16)$$

则当 H_0 成立时,

$$P\{g > z\} = n(1-z)^{n-1} - \frac{n(n-1)}{2}(1-2z)^{n-1} + \dots + (-1)^{m-1} \frac{n!}{m!(n-m)!} (1-mz)^{n-1}. \quad (17)$$

式中 $n = [N/2]$; m 为小于 $1/z$ 的最大整数。给定显著性水平 α 由 $P\{g > z_\alpha\} = \alpha$ 定出 z_α 。若 $g > z_\alpha$ 则拒绝 H_0 , 即 $\{y_t\}$ 含有周期分量。

当检验出周期图上最大的一个周期分量后, 我们考虑第二大的一个, 但这时必须将式(16)的分母中去掉刚才最大的那一个, 也就是说, 若 I_{p_2} 是第二大的, 则考虑:

$$g' = I_{p_2} / [\sum_{p=1}^{[N/2]} I_p - I_{p_1}],$$

用同样的方法进行检验, 此时式(17)中的 n 应改为 $n-1$, 以上程序可以一直进行下去, 直到无法判定有隐含周期为止。

3.4 谱的平滑化

离散时间序列的功率谱大都呈锯齿状, 使谱的峰点不好确定。对谱进行平滑化可以去掉锯齿, 使谱变得光滑。平滑化方法很多, 一般采用汉恩(Hannan)公式:

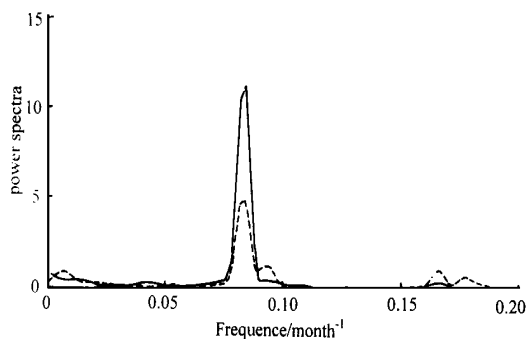


图1 红岩煤矿降雨量和涌水量功率谱图

实线——降雨量谱, 虚线——涌水量谱

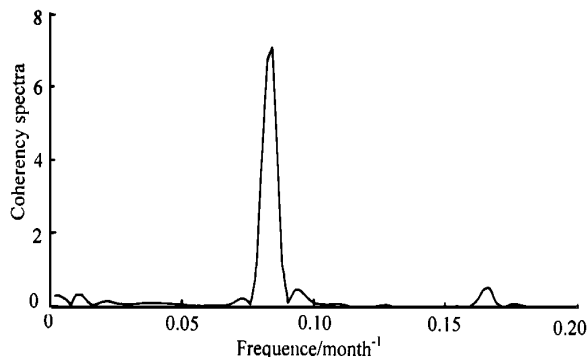


图2 红岩煤矿降雨量和涌水量互功率谱图

实线——降雨量谱; 虚线——涌水量谱

$$S_k = 0.25S_{k-1} + 0.5S_k + 0.25S_{k+1} \quad (18)$$

如有必要还可把得到的加权平均序列 S_k 再按式(18)进行下去, 直到满意为止。

4 应用实例

4.1 自谱分析结果

对红岩煤矿 1966~1995 年月最大涌水量和月降雨量观测序列分别进行谱频分析, 计算结果如图 1。由图 1 知, 降雨量谱在频率 $f = 0.083\ 915$, 即 $T = 1/f = 12$ 个月处有一显著峰值, 经检验为显著周期, 涌水量谱在频率 $f_1 = 0.083\ 915$, $f_2 = 0.165\ 88$, 即 $T_1 = 1/f_1 = 12$ 个月, $T_2 = 1/f_2 = 6$ 个月两处有显著峰值, 经检验均为显著周期。故红岩煤矿降雨动态变化以一年为周期, 涌水量动态变化以一年及 6 个月为周期。

4.2 互谱分析结果

把降雨量作为系统的输入, 涌水量作为系统的输出, 对它们进行互谱分析, 结果如图 2。由图 2 知, 互功率谱在功率 $f = 0.083\ 915$, 即 $T = 1/f = 12$ 个月处有一显著峰值, 经检验为显著周期。由于 $T = 12$ 与前面降雨量与涌水量具有一个周期相同, 故其为降雨量与涌水量共有的且呈线性响应的周期谱峰。在 $T = 12$ month 处, 涌水量相对降雨量滞后 $P = 0.208$ month = 6 d。

5 结语

a. 地质时间序列是由本身蕴含着的各种因素

对其综合作用的结果, 因此建立地质时间序列的随机模型, 对于地质过程的定量及揭示地质过程的内部特征是非常必要的。

b. 通过红岩煤矿降雨量和涌水量的频谱分析, 得到它们的动态变化共有一年的周期, 反映了红岩煤矿涌水量受降雨的严密控制。涌水量有一个 6 个月的次一级周期, 说明涌水量的变化受降雨控制外, 还受其它因素, 如附近地表河流水位变化周期, 人工开采周期等的影响。

c. 红岩煤矿涌水量与降雨量的互谱分析结果表明, 降雨量与涌水量呈线性响应, 即它们的补给关系是线性的。而且存在滞后 6 d 的关系, 即降雨量出现峰值后, 涌水量只须 6 d 就出现峰值。说明本矿井降雨对涌水量的补给关系是管道式的快速补给, 因此应特别注意丰水季节矿井水的防治。

参考文献

- 1 杨位钦, 顾岚. 时间序列分析与动态数据建模. 北京理工大学出版社, 1988; 177~196
- 2 杨永国. 灰色时序组合模型及其在矿井涌水量预测中的应用. 水文地质工程地质, 1996; (6): 36~38
- 3 陈兆国. 时间序列及其谱分析. 科学出版社, 1988; 253~354
- 4 余志伟, 丁仲礼, 刘东升. 2.5 Ma 以来地球轨道参数变化对黄土粒度的线性驱动. 第四纪研究, 1992; (2): 118~127
- 5 Hannan E. J. Multiple time series analysis. New York and London, John Wiley and Sons, 1970; 16~58

(收稿日期 1997-10-17)

混合随机序列模型在地下水位预报中的应用

虎维岳 (煤炭科学研究总院西安分院 710054)

宋金良 (广州大学建工系 510091)

刘其声 马培智 (煤炭科学研究总院西安分院 710054)

摘要 通过对北方某矿区水文地质条件的分析,特别是对底板承压水多年水位变化规律及其控制因素的分析,建立了预测地下水位变化趋势的混合随机序列模型,并对所建模型进行了考核,分析了残差序列的随机性,通过对模型的预报应用,取得了良好的效果。

关键词 随机模式 水位 预报

中国图书资料分类法分类号 P641.4

作者简介 虎维岳 男 35岁 高级工程师 博士研究生 水文地质工程地质

1 引言

在地下水位中长期预报理论和方法研究方面,人们已做了大量的工作,主要成果可分为以数值法为代表的分布参数模型和以系统理论为基础的集中参数预报模型。但不同模型方法有其各自不同的应用条件,本文所研究的北方某矿区有着其特殊的水文地质条件和30年以上的反映含水层系统输入、输

出情况的随机序列数据。通过对其水文地质条件的分析,可简单地将该含水系统概化为图1形式。其中大气降水是含水层的主要补给水源,且其呈现出明显的多年周期性随机变化。含水系统的四周基本为隔水边界,只有东部边界有少量的较稳定的侧向补给,矿井排水构成含水层的主要排泄边界。由此可见,反映含水层状态输出随机序列主要受大气降水控制。图2为大气降水与地下水位年变幅关系图。

由于煤矿开采深度的不断加大,煤层底板承受

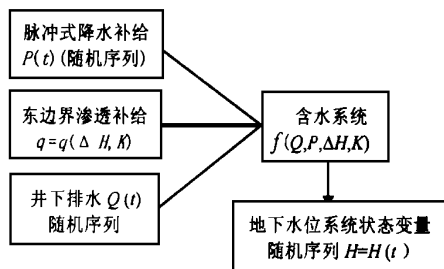


图1 含水系统输入-输出关系示意图

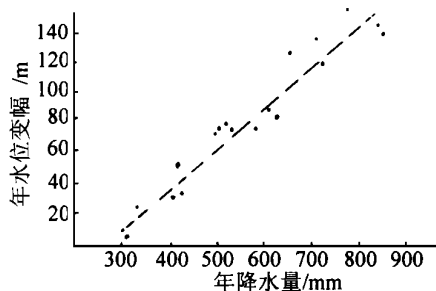


图2 年大气降水与水位变幅关系图

THE SPECTRAL ANALYSIS METHOD OF THE GEOLOGICAL TIME SERIES AND ITS APPLICATION

Yang Yongguo Han Baoping (China University of Mining and Technology)

Abstract The random model of the geological time series and the spectral analysis method used to study its correlation are introduced, and the Fisher criteria for testing the significance periods of the time series are given. The application results shown that the unknown periods of the geological time series and their correlation could be well determined by the spectral analysis, thus the inner characteristics of the geological time series could be find out.

Keywords time series analysis; spectral analysis; cross-power density analysis; delay time