

# 土性随机过程线性模型与相关函数

周小文 包承纲

(土工研究所)

**摘 要** 用随机序列原理讨论了土性随机过程的线性模型及相关函数, 探讨了几种常用相关模型的理论基础, 并认为单指数型、指数余弦型是二种比较合适的土性相关模型

**关键词** 土 随机过程 模型 相关

## 0 前 言

在实际计算土性相关函数及相关距离时, 需要将随机过程离散化, 得到离散随机过程样本, 一般取样为等间距, 因此可视为随机序列 (或称时间序列) 例如, 某土性沿深度方向随机过程样本曲线, 等间距取样后得到一系列观测值:  $X(0), X(\Delta h), X(2\Delta h), X(3\Delta h), \dots, X(n\Delta h)$ 。为方便起见, 记为  $X_0, X_1, X_2, X_3, \dots, X_n$ 。此即为对过程离散采样所得随机序列样本, 如过程为平稳的, 间距为  $K\Delta h$  的两点间的自相关函数 (相关系数) 为

$$\hat{c}_k = \frac{C_k}{C_0} = E \left[ \frac{X_t - \bar{X}}{e} \frac{X_{t+k} - \bar{X}}{e} \right] = \frac{C_k}{e^2}$$

式中,  $C_k = E[(X_t - \bar{X})(X_{t+k} - \bar{X})]$ ,  $k = 0, 1, 2, \dots$ , 为协方差函数;  $\hat{c}_k$  用样本序列估值为

$$\hat{c}_k = \frac{\sum_{i=0}^{N-k} [(X_i - \bar{X})(X_{i+k} - \bar{X})]}{\sum_{i=0}^n [(X_i - \bar{X})^2]}$$

除了白噪声之外, 所有随机过程相邻两取值之间都存在相关性, 线性相关函数  $\hat{c}_k$  即是描绘相关的型式及相关的程度, 例如  $e^{-\gamma k}$  和  $e^{-\gamma k} \cos U_k$  是实际中二种最常遇到的关联形式<sup>[1]</sup>, 为什么不同随机过程或序列会有不同的相关形式? 如果不从频域角度而是从时域角度去解释, 是由于过程或序列具有不同的线性模型和模型参数, 所谓线性模型是指过程或序列点取值间的线性依赖关系。

均值为零, 具有有理谱密度的平稳过程或序列必属于以下三类模型之一 (工程中常见的平稳过程或序列大多数都具有有理谱密度)<sup>[2]</sup>。

## 1 三类线性模型

### 1.1 AR(p)自回归模型

任何一个时刻  $t$  的数值  $X_t$  可表示为过去  $p$  个时刻上数值  $X_{t-1}, X_{t-2}, \dots, X_{t-p}$  的线性组合加上  $t$  时刻的白噪声 (或称  $t$  时刻的扰动), 表示为

$$X_t = H_1 X_{t-1} + H_2 X_{t-2} + \dots + H_p X_{t-p} + a_t \quad (2)$$

常数系数  $H_1, H_2, \dots, H_p$  为模型参数,  $p$  为模型阶数,  $\{a_t, t = 0, 1, 2, \dots\}$ ,  $a_t$  是与  $X_t$  独立的均值为零的白噪声, 它表示  $t$  时刻的附加波动或外加扰动

### 1.2 MA(q)滑动平均模型

$X_t$  可表示成白噪声  $\{a_t\}$  在  $t$  和  $t$  以前  $q$  个时刻上数值 (扰动)  $a_t, a_{t-1}, \dots, a_{t-q}$  的加权和, 即表示为

$$X_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (3)$$

### 1.3 ARMA(p, q)自回归滑动平均模型或混合模型

模型表示为

$$\begin{aligned} X_t &= H_1 X_{t-1} + H_2 X_{t-2} + \dots + H_p X_{t-p} \\ &= a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \end{aligned} \quad (4)$$

最常使用的几种模型为 AR(1), AR(2), MA(1), MA(2), ARMA(1, 1), ARMA(2, 1), ARMA(1, 2)等。

三类模型中参数  $H_i, \theta_i$  及白噪声  $a_t$  的方差可由自相关函数  $\hat{c}_k$  或协方差函数  $C_k$  求得, 如

$$\text{AR}(1): H_1 = \hat{c}_1, \sigma_a^2 = C_0(1 - \hat{c}_1^2) \quad (C_0 = e^2)$$

$$\text{AR}(2): H_1 = \hat{c}_1(1 - \hat{c}_2) / (1 - \hat{c}_1^2), H_2 = (\hat{c}_2 - \hat{c}_1^2) / (1 - \hat{c}_1^2), \sigma_a^2 = C_0(1 - H_1 \hat{c}_1 - H_2 \hat{c}_2)$$

$$\text{MA}(1): \theta_1 = -2\hat{c}_1 / (1 + \sqrt{1 - 4\hat{c}_1^2}), \sigma_a^2 = C_0(1 + \sqrt{1 - 4\hat{c}_1^2}) / 2$$

$$\text{ARMA}(1, 1): H_1 = \hat{c}_2 / \hat{c}_1, \theta_1 = (b \pm \sqrt{b^2 - 4}) / 2, b = (1 - 2\hat{c}_2 + H_1^2)$$

选  $\theta_1$  表示式中正负号, 使  $|\theta_1| \leq 1$

以上模型实际描述的是过程的静特性,以 AR(1)为例,只考虑了  $X_{t-1}$  对  $X_t$  的影响,如果考虑  $X_{t-1}$  的数值与  $X_{t-2}$  有关,而  $X_{t-2}$  的数值又与  $X_{t-3}$  有关,这样依次递归代换之后,将得到过程的动态表达式。三类线性模型的动态表达式为

$$X_t = \sum_{j=0}^{\infty} (G_j a_{t-j})$$

式中  $G_j$  为格林函数,  $G$  反映了各点上附加波动或扰动的影响程度或称为过程对抗动的记忆能力。  $G$  值越大,说明扰动噪声  $a_{t-j}$  对  $X_t$  的影响越明显(可见  $G$  与相关距离有内在的联系)

格林函数可采用比较系数法或因式分解法,由模型参数计算得出,图 1 图 2 为二组 AR(1), ARMA(2, 1)模型的格林函数曲线

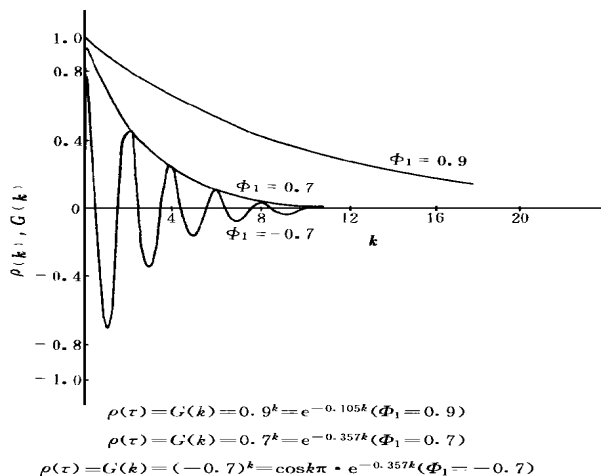


图 1 AR(1)模型格林函数及自相关函数

## 2 协方差与相关函数

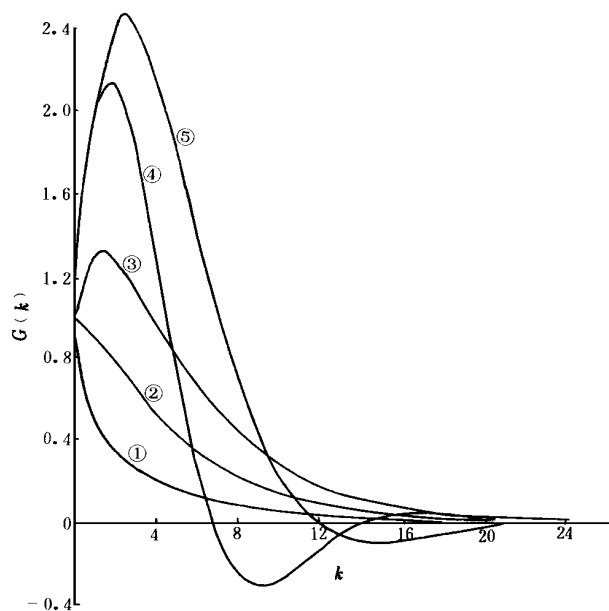
### 协方差函数

$$\begin{aligned}
 C_k &= E[(X_t - \bar{X})(X_{t+k} - \bar{X})] \\
 &= E\left[\left(\sum_{j=0}^{\infty} G_j a_{t-j}\right)\left(\sum_{j=0}^{\infty} G_j a_{t+k-j}\right)\right] \\
 &= \left(\sum_{j=0}^{\infty} G_j G_{j+k}\right) \sigma_a^2 \quad (6)
 \end{aligned}$$

经推导得出下列模型

#### 2.1 AR(1)模型

$$\begin{aligned}
 \because C_0 &= \sigma^2 = \sigma_a^2 / (1 - H_1^2); \\
 C_k &= C_0 H_1^k = \sigma_a^2 / (1 - H_1^2) H_1^k \\
 \therefore d_k &= H_1^k = G
 \end{aligned}$$



①:  $H_1 = 0.85, H_2 = -0.04, \theta_1 = 0.4$  ②  $H_1 = 1.3, H_2 = -0.4, \theta_1 = 0.4$  ③  $H_1 = 1.3, H_2 = -0.4, \theta_1 = 0$  ④  $H_1 = 1.4, H_2 = -0.59, \theta_1 = -0.55$  ⑤  $H_1 = 1.5, H_2 = -0.6, \theta_1 = -0.5$

图 2 ARMA(2, 1)模型格林函数

可见, AR(1)模型的自相关函数  $d_k = d_k^*$ , 也就是格林函数, 如图 1 所示。此外, 由于 AR(1)模型满足  $d_k = d_n d_{(m+n-k)}$ , 若  $\{X_t\}$  为高斯平稳过程(或序列), 则  $\{X_t\}$  为马尔科夫过程(链), 也就是说, 具有 AR(1)线性模型的高斯平稳过程为马氏过程。

$d_k = d_k^*$  可改写为:

$$\text{若 } d_k > 0, d_k = e^{k \ln \rho_1} = e^{-k |\ln \rho_1|} \quad (\text{单指数型式}) \quad (7)$$

$$\text{若 } d_k < 0, d_k = (-1)^k e^{k \ln(-\rho_1)} = e^{-k |\ln(-\rho_1)|} \cos k\pi \quad (8)$$

对于连续随机过程, 只能取 (7) 式, 对于离散随机过程或随机序列, 可取 (7) 或 (8) 式。

由于其它线性模型都不满足  $d_k = d_n d_{(m+n-k)}$ , 因此可以反过来说, 高斯平稳的连续马尔科夫过程线性模型为 AR(1)型, 而其相关函数  $d_k$  可以表示为单指数衰减形式:

$$d(f) = e^{-\tau f} \quad (9)$$

$f$  为两点间距离。

#### 2.2 ARMA(2, 1)模型

$$G = g_1 \lambda_1^k + g_2 \lambda_2^k \quad (10)$$

式中:  $g_1 = (\lambda_1 - \theta_1) / (\lambda_1 - \lambda_2)$ ;  $g_2 = (\lambda_2 - \theta_1) / (\lambda_2 - \lambda_1)$

$$\text{而 } \lambda_1, \lambda_2 = (H_1 \pm \sqrt{H_1^2 - 4H_2}) / 2 \quad (11)$$

协方差函数

$$C_k = d_1 \lambda_1^k + d_2 \lambda_2^k \quad (12)$$

式中:  $d_1 = \frac{e^2}{g_1^2} (g_1^2 / (1 - \lambda_1^2) + g_1 g_2 / (1 - \lambda_1 \lambda_2))$ ;  
 $d_2 = \frac{e^2}{g_2^2} (g_2^2 / (1 - \lambda_2^2) + g_1 g_2 / (1 - \lambda_1 \lambda_2))$ ;

自相关函数

$$\hat{c}_k = C_k / C_0 =$$

$$(d_1 \lambda_1^k + d_2 \lambda_2^k) / (d_1 + d_2) = f_1 \lambda_1^k + f_2 \lambda_2^k \quad (13)$$

式中:  $f_1 = d_1 / (d_1 + d_2)$ ;  $f_2 = d_2 / (d_1 + d_2)$

根据式(11), 由不同的模型参数  $H_1, H_2$ , 可以得到  $\lambda_1, \lambda_2$  的几种取值情况。当  $(H_1^2 + 4H_2) \geq 0$  时,  $\lambda_1, \lambda_2$  为实数, 否则为复数。取实数时  $\lambda_2$  又可能为正数或负数。因此  $\hat{c}_k$  可改写为如下几种形式:

$$\hat{c}_k = f_1 e^{-k|\lambda_1|} + f_2 e^{-k|\lambda_2|} \quad (14)$$

$$\hat{c}_k = f_1 e^{-k|\lambda_1|} + f_2 e^{-k|\lambda_2|} \cos k\theta \quad (15)$$

$$\hat{c}_k = e^{-kT} (\cos U_k + b \sin U_k)$$

$\lambda_1, \lambda_2$  为复数, 且  $|\lambda_1| \leq 1, |\lambda_2| \leq 1, T, U, b$  为由  $H_1, H_2, \theta_1$  决定的正实数 (16)

由此可知, 线性模型为 ARMA(2, 1)型的平稳随机过程或序列, 其自相关函数(相关系数)可有以下几种型式: (1)二个单指数衰减函数的加权和; (2)一个单指数衰减函数和一个指数余弦函数的加权和; (3)指数余弦指数正弦型。

由(14), (15)式可见, 对第(1), (2)两种相关型式, 自相关函数实际上是两独立马尔科夫过程相关函数的加权和, 因此具有这二种相关型式的随机过程描绘的是两独立马尔科夫过程的组合。

图3为几组 ARMA(2, 1)模型的自相关函数曲线。

### 2.3 AR(2), ARMA(1, 1), MA(1)模型

AR(1), AR(2), ARMA(1, 1), MA(1)均可认为是 ARMA(2, 1)的特例。如 ARMA(2, 1)中:  $H_2 = \theta_1 = 0$  时, 退化为 AR(1)型;  $\theta_1 = 0$  时, 退化为 AR(2)型;  $H_2 = 0$  时, 退化为 ARMA(1, 1)型;  $H_1 = H_2 = 0$  时, 退化为 MA(1)型;

### 2.4 ARMA(1, 2)模型

经推导得

$$G_0 = 1, G_1 = H_1 - (\nu_1 + \nu_2),$$

$$G_j = H_1^j - (\nu_1 \nu_2) H_1^{j-1} + \nu_1 \nu_2 H_1^{j-2} \quad (j \geq 2)$$

式中  $V_1, V_2 = (\theta_1^2 + 4\theta_2) / 2$

$$\hat{c}_k = A_1 H_1^{k-1} + A_2 H_1^k + A_3 H_1^{k-1} + A_4 H_1^{k-2} + A_5 H_1^{k-3} + A_6 H_1^{k-4} \quad (17)$$

其中  $A_1, A_2, \dots, A_6$  是由模型参数决定的系数:

$$A_1 = (\frac{e^2}{e^2}) [H_1 - (\nu_1 + \nu_2)]; A_2 = (\frac{e^2}{e^2}) \{1 - [H_1 - (\nu_1 + \nu_2)](\nu_1 + \nu_2) + H_1^2 / (1 - H_1^2)\}; A_3 = (\frac{e^2}{e^2}) \{[H_1 - (\nu_1 + \nu_2)]\nu_1 \nu_2 - (\nu_1 + \nu_2) - 2H_1^2(\nu_1 + \nu_2) / (1 - H_1^2)\}; A_4 = (\frac{e^2}{e^2}) \{\nu_1 \nu_2 + H_1^2 [2\nu_1 \nu_2 - (\nu_1 + \nu_2)^2] / (1 - H_1^2)\}; A_5 = (\frac{e^2}{e^2}) \{-2H_1^2 \nu_1 \nu_2 (\nu_1 + \nu_2) / (1 - H_1^2)\}; A_6 = (\frac{e^2}{e^2}) \{H_1^2 \nu_1^2 \nu_2^2 / (1 - H_1^2)\}$$

由(17)式可以看出, ARMA(1, 2)模型与 ARMA(2, 1)模型相似, 其自相关函数亦具有几种型式, 且亦为单指数、指数余弦、指数余弦指数正弦三种相关型式的组合。

## 3 土性随机过程相关函数

土性随机过程比较复杂, 其线性模型如要精确可能需要有较高的阶数, 如 AR(4), ARMA(4, 2), MA(4)等, 但用 AR(1), AR(2), ARMA(2, 1), ARMA(1, 2), MA(1), MA(2), ARMA(1, 1)几种低阶模型可以近似地模拟(这样的近似模型应可以用于孔间地质特征推断或孔外特征估计)。

同样, 土性自相关亦有各种形式, 但用所拟合线性模型的相关函数亦能比较准确地近似, 也就是如上说的, 可以用单指数、指数余弦、指数余弦指数正弦三种相关模型的加权和近似。

但实用起来这样作仍很复杂, 还需简化。土性自相关函数应选用一二种简便且有一定准确度并有比较广泛代表性的函数形式(所谓代表性是指不同的场地、不同的土层、不同的土性指标)。

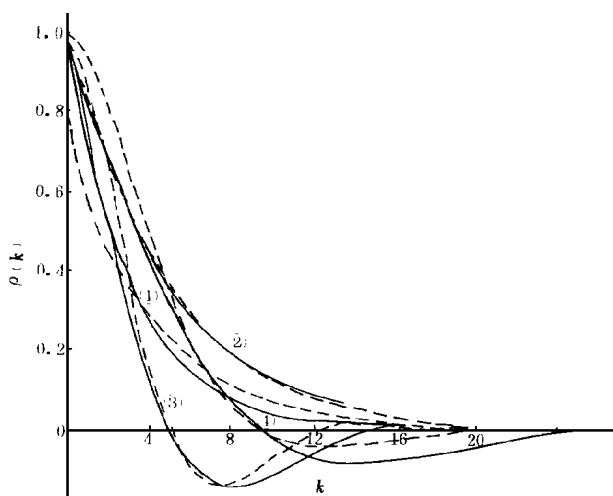
Vanmarcke 建议的几种土性自相关函数型式为: 单指数、双指数、三角形、二次自动回归、指数余弦、指数余弦指数正弦<sup>[3]</sup>。从以上不很全面的理论推导可以看出, Vanmarcke 的建议是有其理论根据的, 同时可见这几种是近似的。

目前比较常用的是单指数型, 对于高斯平稳随机过程, 由于只有马尔科夫过程的自相关函数才是单指数型, 因此这种型式的相关函数严格说只有一定的代表性。

指数余弦型表达式比单指数型略为复杂, 它能反映当两点间距增大到一定程度呈现负相关的情况, 因此有更好的适用性和代表性。单指数型可视为指数余弦型的特例。当  $\cos U_f$  中的  $U_f \rightarrow 0$  时, 指数余弦型即退化为单指数型。

在图3中, 用单指数型、指数余弦型近似模拟

ARM A(2,1)模型的相关函数,可见近似程度是比较好的。



①:  $H_1 = 0.85, H_2 = -0.04, \theta_1 = 0.4$

“——”表示  $d(k) = e^{-0.314k}$

“---”表示  $d(k) = 0.687e^{0.223k} + 0.313e^{-2.223k}$

②:  $H_1 = 1.3, H_2 = -0.4, \theta_1 = 0.4$

“——”表示  $d(k) = e^{-0.20k}$

“---”表示  $d(k) = 1.164e^{-0.223k} - 0.184e^{-0.593k}$

③:  $H_1 = 1.4, H_2 = -0.59, \theta_1 = 0.55$

“——”表示  $d(k) = \cos 0.314k \cdot e^{-0.22k}$

“---”表示  $d(k) = e^{-0.254k} (\cos 0.424k + 0.642 \sin 0.42k)$

④:  $H_1 = 1.5, H_2 = -0.6, \theta_1 = -0.5$

“——”表示  $d(k) = e^{-0.15k} \cos 0.165k$

“---”表示  $d(k) = e^{-0.255k} (\cos 0.253k + 1.025 \sin 0.253k)$

(实线为模拟曲线)

图 3 几组 ARM A(2,1)模型的自相关函数曲线

土性的变异比较复杂,在不同地区、不同的土类、甚或不同的土性指标,其自相关可能有各种形式,且不规则,再加上钻探工具、测试手段的极不完善,根据土性测值计算得到的自相关曲线,可能不单独符合任何一种典型模型,而是几种模型的组合。针对这种情况,有必要进行去粗取精、去伪存真的平滑处理,并尽量以最简单的模型拟合实际的相关函数曲线。比较而言,单指数型和余弦指数型是二种便于采用、近似程度较好的模型。

## 4 算 例

图 4 为一静力触探曲线,采样间距  $\Delta h = 0.30\text{m}$ ,样本点数  $N = 59$ ,计算得其自相关函数如图 5 中实线所示。

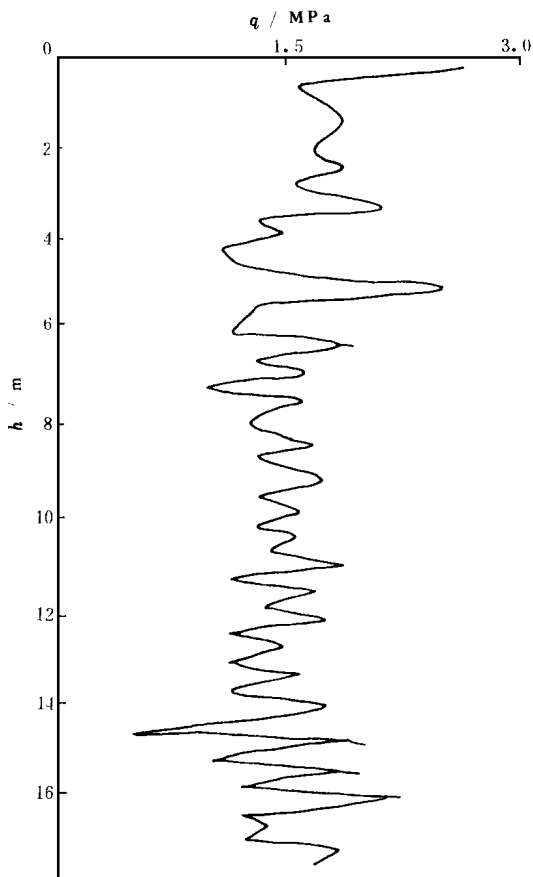


图 4 静力触探曲线

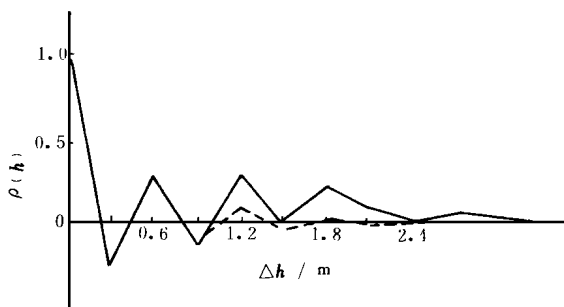


图 5 自相关函数

经判别该土性过程为 AR(2)模型,  $H_1 = -0.172, H_2 = 0.253$

线性模型

$$q_h = 15 - 0.172q_{h-0.3} + 0.253q_{h-0.6} + a_t$$

该线型模型的相关函数

$$\begin{aligned} d &= 0.359 - 0.421k + 0.641(-0.593)^k \\ &= 0.359e^{-0.865k} + 0.641\cos\pi \cdot e^{-0.523k} \end{aligned}$$

将其绘于图 5 如虚线表示,可以认为与实际相关曲线是足够近似的。

从图 5 可见,相比之下用指数余弦模拟实际相

关曲线精度要稍高一些，但显然有一定误差

参 考 文 献

1 Vanmarcke E H Random Field Analysis and Synthe-

sis. M. I. T. Press, Cambridge, Mass, 1984  
2 汪荣鑫. 随机过程. 西安交通大学出版社, 1987  
3 Vanmarcke E H. Probabilistic Modeling of Soil Profiles.  
Journal of Geotechnical Engineering Division, ASCE,  
Vol. 103, No. GT11, 1977

Linear Model Studying Soil Random Process  
And Relation Functions

Zhou Xiaowen Bao Chenggang  
( Yangtze River Scientific Research Institute, Wuhan 430010)

**Abstract** Using random sequence theory, a linear model studying the properties of soil random process and relation functions are discussed and the theoretic bases of several typical relation models are investigated. It is found that the single index model and index cosine model are regarded as two better appropriate models for studying soil property relation

**Keywords** soil; random process; model; relation

(上接第 5页)

Studies on Hydrodynamical Character of Horizontal Type  
Gate of Three Gorges Project's Ship Lift

Zhou Chi, Liu Dunhuang, Yang Chun  
( Yangtze River Scientific Research Institute, Wuhan 430010)

**Abstract** Using the method combined physical model and mathematic model , under the different conditions of submerged water depth and support position of oil cylinder, a primary research on effect of the operating duty of horizontal type gate upon the hydrodynamic load, opening or closing force and reaction of supports, etc. , was conducted. The results showed that the variation of water level between gates is basically restrained by using the synchronous operating duty with variable velocities of two gates , and so, the opening or closing force and reaction of supports may be reduced. The research result provides reliable reference data for optimizing design.

**Keywords** ship lift; horizontal type gate; operating duty; hydrodynamical character; physical model; mathematic model