

doi:10.3799/dqkx.2016.118

基于变系数回归模型的三峡库区滑坡位移预测

喻孟良^{1,2}, 梅红波^{3*}, 李冀骅⁴, 赵慧², 吴润泽⁵

1. 中国地质大学信息工程学院, 湖北武汉 430074

2. 中国地质环境监测院, 北京 100081

3. 中国地质大学资源学院, 湖北武汉 430074

4. 中国地质大学计算机学院, 湖北武汉 430074

5. 三峡库区地质灾害防治工作指挥部, 湖北宜昌 443000

摘要:降雨—库水联合作用影响着三峡库区滑坡, 而降雨、库水分别对滑坡演化的贡献及作用规律迄今尚不明确。以库区树坪滑坡和八字门滑坡为例, 通过分析降雨和库水位资料, 采用变系数回归模型, 对滑坡位移进行预测。实验结果表明: 经过改进的变系数回归模型方法不仅比传统的线性回归模型、自回归积分滑动平均模型、支持向量机模型方法具有更高的预测精度, 而且能定量地给出各影响因素对滑坡位移的贡献。

关键词: 三峡库区; 降雨; 库水; 变系数回归模型; 滑坡位移预测; 数字模型。

中图分类号: P694

文章编号: 1000—2383(2016)09—1593—10

收稿日期: 2016—01—22

Landslide Displacement Prediction Based on Varying Coefficient Regression Model in Three Gorges Reservoir Area

Yu Mengliang^{1,2}, Mei Hongbo^{3*}, Li Jihua⁴, Zhao Hui², Wu Runze⁵

1. Faculty of Information Engineering, China University of Geosciences, Wuhan 430074, China

2. China Institute of Geo-Environment Monitoring, Beijing 100081, China

3. Faculty of Earth Resources, China University of Geosciences, Wuhan 430074, China

4. School of Computer Science, China University of Geosciences, Wuhan 430074, China

5. Command of Geological Hazard Control in Three Gorges Reservoir Area, Yichang 443000, China

Abstract: Landslides in Three Gorges reservoir area are affected by both the rainfall and the reservoir water, but their respective contribution and the laws of the landslide evolution remain unclear. In this paper, we analyzed the rainfall and water level data, and applied a varying coefficient regression model for landslide displacement prediction taking Shuping landslide and Bazimen landslide in the reservoir area as the study area. The results of experiments show that the improved varying coefficient regression model is not only more accurate than traditional linear regression model, auto-regressive Integrated moving average model and support vector machine model, but also it can give quantitative contribution of each factor on the landslide displacement.

Key words: Three Gorges reservoir area; rainfall; reservoir water; varying coefficient regression model; landslide displacement prediction; numerical models.

三峡工程是人类历史上最大规模的水利工程。三峡工程在发挥巨大综合效益的同时, 在地质灾害防治等方面还存在一些亟需解决的问题(国务院办

公厅 2011 年《三峡后续工作规划》)。三峡大坝蓄水后, 地质灾害活动加剧, 已严重影响到三峡库区的社会和经济生活。

基金项目: 三峡库区地质灾害预警指挥系统数据库及管理系统建设项目(No.SXJC-3ZH1B1)。

作者简介: 喻孟良(1978—), 男, 博士, 高级工程师, 主要从事水工环地质信息化工作, E-mail: yuml@mail.cigem.gov.cn

*** 通讯作者:** 梅红波, E-mail: hbmei@cug.edu.cn

引用格式: 喻孟良, 梅红波, 李冀骅, 等, 2016. 基于变系数回归模型的三峡库区滑坡位移预测. 地球科学, 41(9): 1593—1602.

降雨—库水联合作用影响着三峡库区滑坡,而降雨、库水分别对滑坡演化的贡献及作用规律迄今尚不明确.在滑坡演化过程中,多大的雨量及雨强在多久后给滑坡产生多大的影响? 库水位及其变化速度对滑坡发展演化的贡献如何?

这两个问题与滑坡的监测预警紧密相关,直接关系到人民的生命财产安全;且库水位问题还涉及三峡水电站的经济效益,水位快降有利于发电,但降速过快又可能引发沿岸的滑坡灾害,目前电站采用的水位日降幅指标是经验值,该指标已成为三峡水库扩大运行调度自由度的瓶颈问题.

近年来,围绕库区滑坡位移预测,国内外学者已提出了数十个滑坡预测预报理论模型和方法.与本文相近的研究有:Du *et al.* (2009,2012)通过移动平均法将位移分解为趋势项及周期项,前项利用多项式位移函数进行拟合预测,后项利用 BP 神经网络进行多变量位移预测,在三峡库区八字门和白水河滑坡位移预测中取得了较好的效果;易庆林等(2013)收集了某水库滑坡 24 期的 GPS 地表位移监测数据及其诱发因素,即水库水位、降雨等资料,采用 BP 神经网络模型对该水库滑坡变形进行建模,最后将 6 期水库水位、降雨等资料输入模型进行滑坡变形预测,结果表明预测结果与实测数据符合性好,总体上能较好地反映变形趋势;Lian *et al.* (2013,2014)提出了改进的基于聚类经验模态分解的极端学习机算法(modified ensemble empirical mode decomposition based extreme learning machine,M-EEMD-ELM),该算法应用于三峡库区白水河滑坡位移的预测中,结果优于传统的神经网络算法和未改进的 EEMD-ELM 的算法;此外,Lian *et al.* (2014)的另一个改进是运用灰色关联分析方法筛选出更具影响力的因素作为集成极端学习机的输入,也取得了良好的效果;彭令等(2013)构建出核主成分分析和粒子群优化支持向量机协同模型,对滑坡相对位移进行预测,预测结果与实际监测值具有良好的一致性;黄发明等(2015)提出了一种基于相空间重构和小波分析—粒子群向量机的滑坡地下水位预测方法,该模型引入小波变换法对地下水位序列进行时频分解,再利用 WA-PSVM 模型对地下水位进行预测.

上述非线性的分析方法很好地符合了滑坡行为的复杂的、开放的和耗散的非线性动力学演化特征,是目前通用的、拟合精度较高的几种预测方法,但这些方法都属于“黑箱”方法,不具有机理可解释性.

本文以三峡库区树坪滑坡与八字门滑坡两个典型滑坡为例,拟采用变系数模型进行滑坡变形影响因素的分析和位移预测,该方法假定滑坡变形过程中,影响滑坡位移的各个因素的权重是随时间动态变化的,并据此建立滑坡位移预测模型,找出各因素权重的变化趋势并进行位移预测.该方法以数理统计学为理论基础,算法透明,便于解释机理.

1 变系数回归理论及方法

1.1 变系数回归模型

变系数回归模型是近年发展起来的广受人们重视的一类模型,该模型通过假定线性回归模型中的回归参数是其他自变量的未知函数以增加模型的灵活性和适应性,同时由于回归系数的函数是某个自变量的一元函数,有效避免了在拟合问题中出现的维数灾难问题(梅长林与王宁,2012).

传统的变系数回归模型可以表示如下:

设 Y 为因变量, $X_1, X_2, \dots, X_p$ 和 U 为自变量(Hastie and Tibshirani,1993),若 Y 与 $X_1, X_2, \dots, X_p$ 及 U 满足:

$$Y = \beta_1(U) X_1 + \beta_2(U) X_2 + \dots + \beta_p(U) X_p + \epsilon = \sum_{j=1}^p \beta_j(U) X_j + \epsilon, \tag{1}$$

式中: $\beta_j(U)$ ($j=1,2,\dots,p$) 为自变量 X_j 的系数,同时也是关于自变量 U 的未知函数, ϵ 为误差项,其数学期望和方差分别满足 $E(\epsilon | U, X_1, \dots, X_p) = 0$, $Var(\epsilon | U, X_1, \dots, X_p) = \sigma^2(U)$.在实际运用中通常将 U 设为时间,即可理解为随着时间的变化,自变量的系数也随着变化,可了解 Y 关于每个 X_j 的变化强度随时间变化的动态规律.

1.2 变系数模型的局部线性估计

对于变系数回归模型,常用的拟合方式有 3 种,包括样条估计、局部多项式估计和核估计方法.变系数模型本质上是局部线性模型,因此更适合用局部多项式方法来拟合(卢骏等,2015).

设 $(Y_i; X_{i1}, X_{i2}, \dots, X_{ip}, U_i)$ ($i=1,2,\dots,n$) 为 n 组独立的观察数据,则变系数模型的样本形式为:

$$Y_i = \sum_{j=1}^p \beta_j(U_i) X_{ij} + \epsilon_i, i=1,2,\dots,n, \tag{2}$$

式中: $\beta_j(u)$ ($j=1,2,\dots,p$) 为关于自变量 u 连续的一阶导数, u 的取值范围为 μ ,对于任一给定的 $u_0 \in \mu$,在 u_0 的邻域内由 Taylor 公式展开有:

$$\beta_j(u) \approx \beta_j(u_0) + \beta_j'(u_0)(u - u_0),$$

$j=1,2,\cdots,p$. (3)

设 $K(t)$ 为给定的核函数,通常取为对称单峰的概率密度函数且满足 $\lim_{t \rightarrow \infty} K(t)=0$,若 U 代表着时间,则其光滑思想为与 u_0 时间越接近的观测点,它对 u_0 处的参数所起的作用越大(卢辰龙等,2014).令:

$$K_h(t)=\frac{1}{h}K\left(\frac{t}{h}\right), \tag{4}$$

式中: $h>0$ 被称为光滑参数或带宽,常用的核函数有:

(1)Gauss 核: $K(t)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}t^2\right)$,即标准正态分布的概率密度函数.

(2)对称 Beta 函数族: $K(t)=\frac{1}{\text{Beta}\left(\frac{1}{2},\gamma+1\right)}$

$(1-t^2)_+^\gamma,\gamma=0,1,2,\cdots$,其中 $(1-t^2)_+^\gamma$ 表示函数 $(1-t^2)^\gamma$ 的正部.

则变系数模型的局部线性拟合(Fan and Huang,2005)即选择 $\beta_j(u_0)$ 和 $\beta_j'(u_0)$ ($j=1,2,\cdots,p$),使得:

$$\sum_{i=1}^n\left(Y_i-\sum_{j=1}^p(\beta_j(u_0)+\beta_j'(u_0)(U_i-u_0))X_{ij}\right)^2\cdot K_h(U_i-u_0)=\min, \tag{5}$$

令

$$X(u_0)=\begin{pmatrix} X_{11} & \cdots & X_{1p} & X_{11}(U_1-u_0) & \cdots & X_{1p}(U_1-u_0) \\ X_{21} & \cdots & X_{2p} & X_{21}(U_2-u_0) & \cdots & X_{2p}(U_2-u_0) \\ \vdots & & \vdots & \vdots & & \vdots \\ X_{n1} & \cdots & X_{np} & X_{n1}(U_n-u_0) & \cdots & X_{np}(U_n-u_0) \end{pmatrix};$$

$$W(u_0)=\text{Diag}(K_h(U_1-u_0),K_h(U_2-u_0),\cdots,K_h(U_n-u_0));$$

$$Y=(Y_1,Y_2,\cdots,Y_n)^T.$$

则在公式(5)中加权最小二乘问题的解中,系数向量 $\beta(u)=(\beta_1(u),\beta_2(u),\cdots,\beta_p(u))^T$ 在 u_0 处位置的局部线性估计可表示成:

$$\hat{\beta}(u_0)=\left(\hat{\beta}_1(u_0),\hat{\beta}_2(u_0),\cdots,\hat{\beta}_p(u_0)\right)^T=\left(I_p,O_p\right)\left(W^T(u_0)W(u_0)X(u_0)\right)^{-1}X^T(u_0)W(u_0)Y, \tag{6}$$

式中: I_p 和 O_p 分别表示 p 阶单位矩阵和 p 阶零矩阵.因此根据公式(6)可得到因变量 Y 在各个点 U_i ($i=1,2,\cdots,n$)的估计值为:

$$\hat{Y}_i=\hat{\beta}_1(U_i)X_{i1}+\hat{\beta}_2(U_i)X_{i2}+\cdots+\hat{\beta}_p(U_i)X_{ip}=\left(X_i^T,0_{1\times p}\right)\left(X^T(U_i)W(U_i)X(U_i)\right)^{-1}X^T(U_i)W(U_i)Y. \tag{7}$$

1.3 交叉确认

在公式(4)的核函数中,不同的光滑参数 h 会对核函数产生影响,进而会影响到回归函数的拟合.如若光滑参数选得过大,会使拟合的曲线过于光滑以至于出现欠拟合现象,从而不能正确体现曲线的真实趋势;而当光滑参数选得过小时,会使估计的曲线震荡加剧从而导致过拟合现象,同样不能真实地反映回归函数的趋势.

交叉确认(cross validation)方法可以用来产生相对较优的光滑参数.对于给定的光滑参数 h ,在 n 个设计点 $X_1,X_2,\cdots,X_n$ 处,因变量的拟合值所构成的向量可表示为:

$$\hat{Y}(h)=\left(\hat{Y}_1(h),\hat{Y}_2(h),\cdots,\hat{Y}_n(h)\right)^T=L(h)Y$$
,对于给定的 h ,去掉第 i 组观测值 (Y_i,X_i) ,用其余 $n-1$ 组数据在给定的光滑参数 h 下利用局部加权拟合求出回归函数 $m(x)$ 在 X_i 点的拟合值记为:

$$\hat{Y}_{(-i)}(h)=\hat{m}_{(-i)}(X_i),$$

依次对 $i=1,2,\cdots,n$ 执行上述估计过程得

$$\hat{Y}_{(-1)}(h),\hat{Y}_{(-2)}(h),\cdots,\hat{Y}_{(-n)}(h)$$
,令

$$CV(h)=\frac{1}{n}\sum_{i=1}^n\left(Y_i-\hat{Y}_{(-i)}(h)\right)^2.$$

交叉确认法即选取 h_0 使得:

$$CV(h_0)=\min_{h>0}CV(h). \tag{8}$$

2 树坪滑坡与八字门滑坡

2.1 滑坡地质概况

树坪滑坡位于秭归县沙镇溪镇树坪村三峡水库右岸,距三峡大坝上游 47 km 处,滑坡体积达到 2 360 万 m^3 .树坪滑坡属于老滑坡,其主要成分为崩坡积碎块土石.滑坡自北向南长约 800 m,东西宽约 700~900 m,滑坡前缘直接延伸至长江.滑坡前缘较为平缓,坡度为 $20^\circ\sim25^\circ$,滑坡中部与滑坡后部较为陡峭,坡度为 $25^\circ\sim30^\circ$.自 2003 年不到 1 个月蓄水 135 m,树坪滑坡就发生大面积变形,裂缝不断加长加宽,至今仍在变动.滑坡平面形态和工程地质剖面如图 1 所示.从图 1a 可以看出,树坪滑坡共设有 6 个 GPS 监测点,其编号分别为 ZG85、ZG86、ZG87、ZG88、ZG89、ZG90,本文选取变形量较大、位于滑坡体前缘的 ZG85 监测点来进行滑坡位移预测.

八字门滑坡位于秭归县归州镇香溪村附近,距

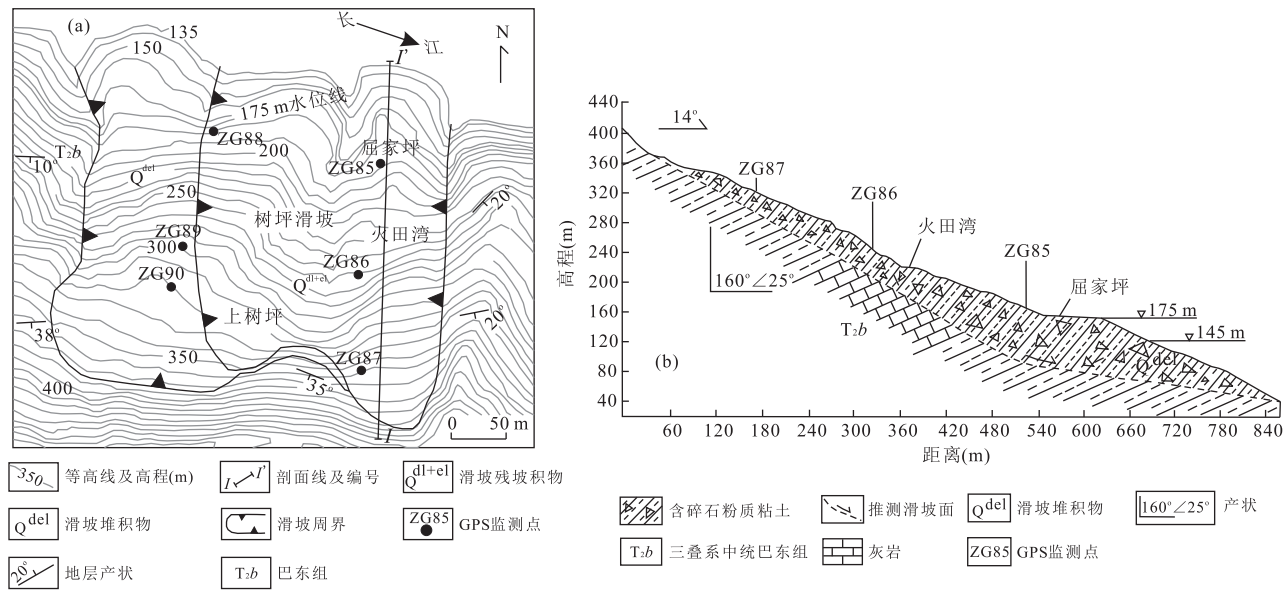


图 1 树坪滑坡 GPS 监测点布置平面图(a)和工程地质剖面(b)
Fig.1 Layout of GPS monitoring points (a) and geological profile (b) of Shuping landslide
来源于三峡库区地质灾害防治工作指挥部,经过适当修改整饰

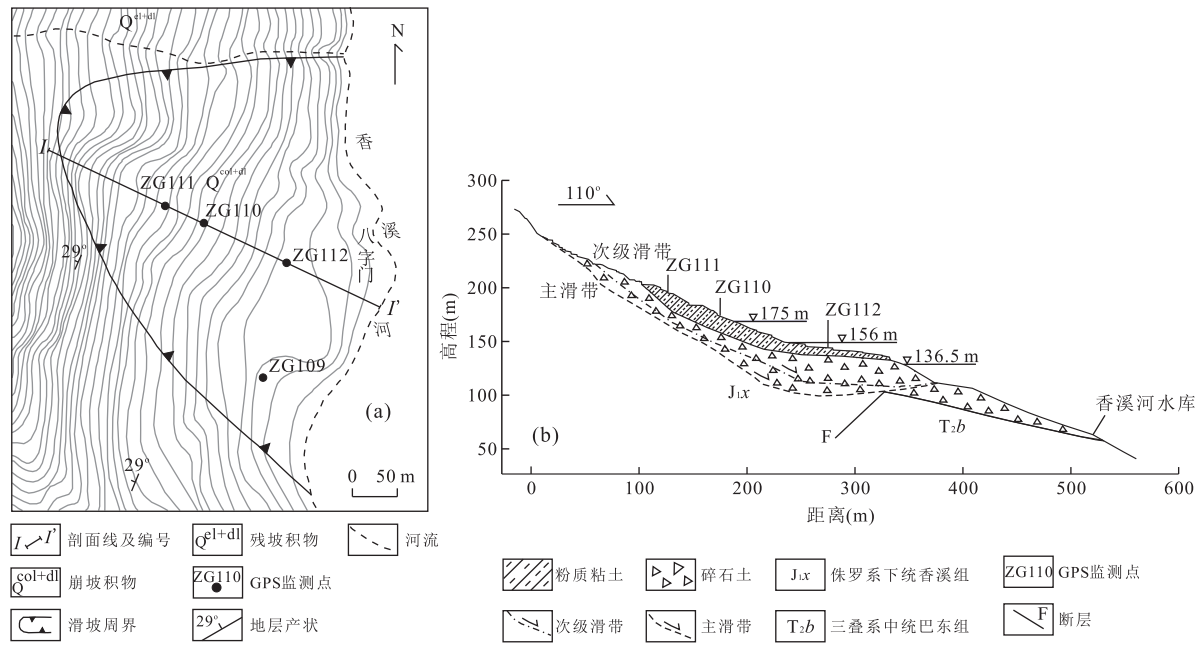


图 2 八字门滑坡 GPS 监测点布置平面(a)和工程地质剖面(b)
Fig.2 Layout of GPS monitoring points (a) and geological profile (b) of Bazimen landslide
来源于三峡库区地质灾害防治工作指挥部,经过适当修改整饰

离三峡大坝约 38 km.滑坡体源于早期侏罗纪砂页岩、砂岩、燧石、砾石等组成的斜坡中,坡度大概为 10°~30°范围内,自从 2003 年 6 月三峡水库蓄水到 135 m 后,滑坡体下缘被水淹没,水上部分滑坡长 380 m,宽 100~500 m,厚度约为 10~35 m,滑坡平面形态和工程地质剖面如图 2 所示.从图 2a 中可以

看出,在八字门滑坡上共设有 4 个 GPS 监测点,其编号分别为 ZG109、ZG110、ZG111、ZG112,其中后 3 个监测点组成一条自西北向东南垂直于江岸的监测剖面,然而随着三峡水库水位高度的不断增加, ZG109 与 ZG112 监测点分别在 2006 年 10 月和 2008 年 10 月,库水位高度达到 153 m 左右时被淹

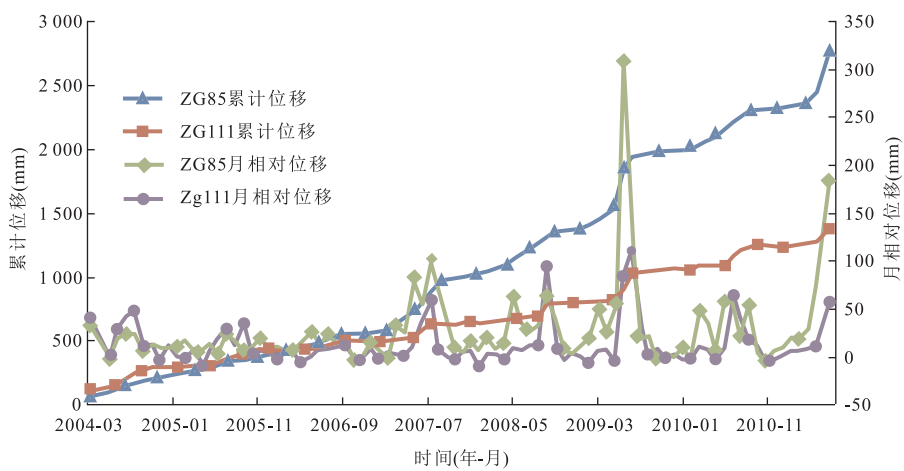


图 3 监测点累计位移与相对位移

Fig.3 Cumulative displacement and relative displacement of monitoring points

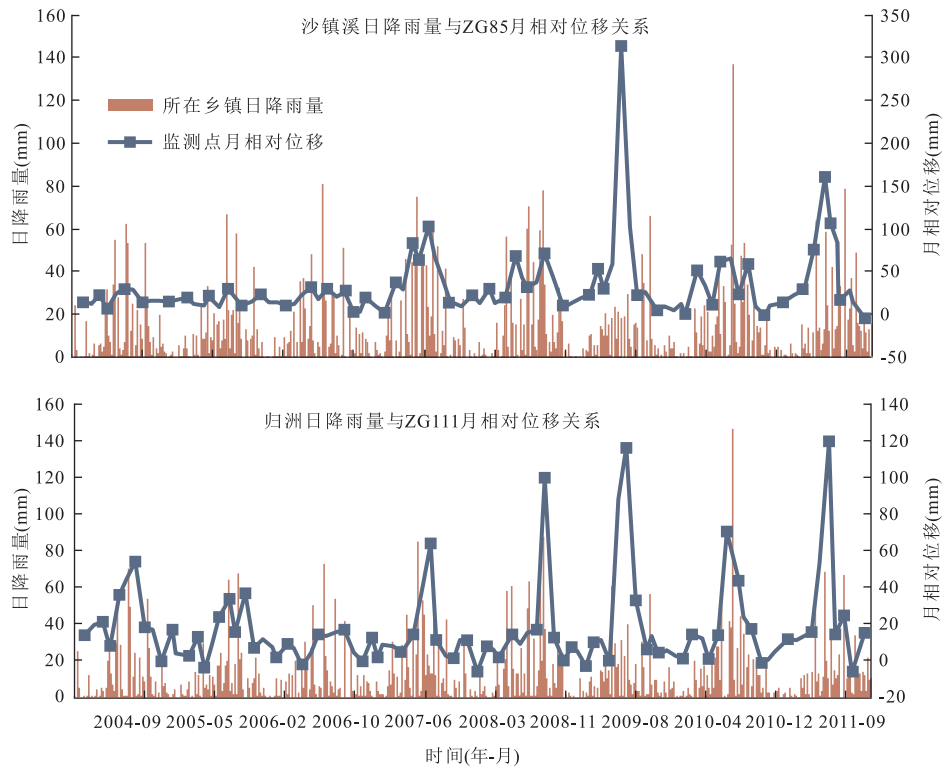


图 4 降雨与监测点相对位移的关系

Fig.4 The relationship between rainfall and relative displacement of monitoring points

没,只剩下 ZG110 和 ZG111 两个监测点的设备能正常运行,本文选取滑坡体后缘的 ZG111 监测点来进行滑坡位移预测.

2.2 监测点位移

监测点位移数据的记录方式为基于时间间隔与基于增量相结合的方式,即每隔一个时间间隔记录一次,如果在该时间间隔内的滑坡位移达到某一上限时,也记录一次.采样时间间隔为 1 个月.ZG85 及 ZG111 的累计位移和相对位移如图 3 所示.

2.3 降雨

降雨是诱发滑坡变形的主要因素之一,本文选取两个滑坡监测点所在位置(沙镇溪镇和归州镇)同期的降雨数据作为计算依据.降雨数据是按天采集,位移数据是按月采集,两个监测点日降雨量与月相对位移的关系如图 4 所示.通过图 4 可以看出,当某时间段内的降雨较多时,其位移也相对较大.考虑到降雨对滑坡的影响可能会存在一定的延迟效应,具体的延迟天数根据公式(8)的交叉验证来选出

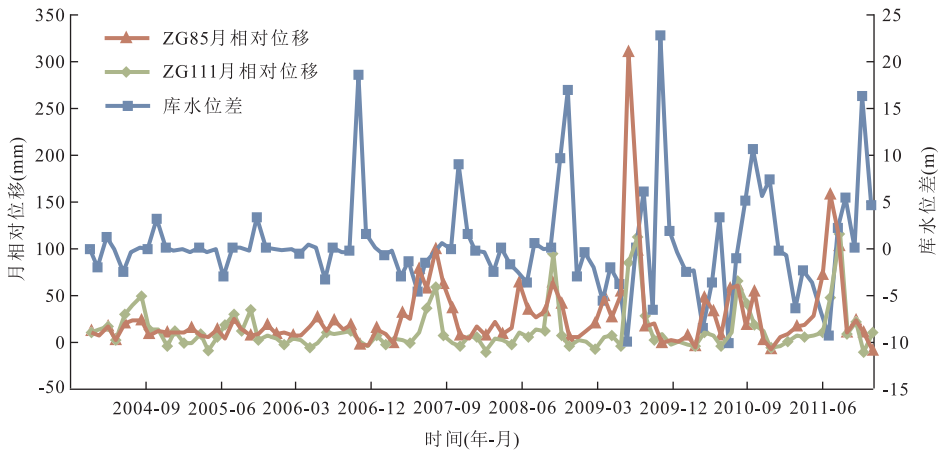


图 5 库水位差与监测点相对位移的关系

Fig. 5 The relationship between water level difference and relative displacement of monitoring points

相对较优的数据。

2.4 库水位

库水位的上升与下降常常诱发边坡失稳破坏而发生滑坡地质灾害(卢书强等, 2014; 罗红明等, 2008). 本文所选滑坡数据来自三峡库区秭归县, 并且位于长江支流的沿岸, 所以三峡库区水位的变化同样会对滑坡产生一定的影响. 三峡水库自从开建到现在主要可以分为 3 个阶段: (1) 2003 年 6 月份开始初次蓄水, 坝前库水位达到 135.0 m; (2) 2006 年 5 月大坝全线建成开始第 2 次蓄水, 并且水位上升至 156.0 m; (3) 2008 年 9 月开始实验性蓄水, 库水位达到 172.8 m. 由于库水位的记录间隔为天, 所以需要转换为月数据来与位移数据匹配, 库水位的变化用库水位差(本月末库水位—上月末库水位)来表示. 图 5 显示了库水位差与监测点 ZG85 和 ZG111 相对位移变化的关系, 可以看出两者间的关联性, 而且库水位变化对于位移的影响同样应该是有延迟效应的, 具体的延迟时间同样应该由公式(8)交叉确认得出.

3 滑坡位移预测

根据影响滑坡的内外因素, 本文选取累计降雨、库水位、库水位差、上月位移作为变系数回归模型的自变量, 累计位移作为因变量. 同时, 由于外在因素的变化会有延迟效应, 具体到选取前多少天累计降雨, 前多少天库水位, 前多少天的库水位差, 会通过公式(8)交叉确认来找到较优的排列组合. 经过交叉确认计算得出: 对于树坪滑坡, 沙镇溪前 58 d 累计降雨、前 24 d 的库水位、过去 52 d 的水位差的组合

为最优组合, 在拟合时方差最小; 对于八字门滑坡, 归洲前 59 d 累计降雨、前 61 d 库水位、前 20 d 水位差的组合为最优组合. 本文选取观测样本为 ZG85 观测点和 ZG111 观测点, 其中 ZG85 的样本是从 2004 年 3 月到 2011 年 6 月共 88 组数据, 用其中前 80 组数据来做拟合, 后 8 组数据用来做预测. ZG111 的样本是从 2004 年 3 月到 2009 年 1 月共 59 组数据, 用其中前 46 组数据来做拟合, 后 13 组数据用来做预测.

(1) 变系数回归模型: 根据变系数回归模型, 对 ZG85 监测点和 ZG111 监测点进行预测的结果如表 1 所示. 真实值与拟合/预测值的对比如图 6 所示.

该模型中, 在拟合 ZG85 时最大误差为 13.07%, 平均误差为 1.97%, 在预测 8 组数据时, 最大误差为 8.2%, 平均误差为 2.7%. 在拟合 ZG111 时最大误差为 7.77%, 平均误差为 2.58%, 预测 13 组数据, 最大误差为 3.36%, 平均误差为 1.25%.

根据变系数回归模型, 各自变量因素随时间的变化如图 7 所示.

通过各自变量因素的系数变化情况可以看出来, 对于 ZG85 监测点来说, 库水位对位移的影响权重是持续增加的, 库水位差对位移的影响权重是持续下降的, 而累计降雨对位移的影响权重缓慢增加的. 对于 ZG111 八字门滑坡, 水位差对位移的影响权重是持续增加的, 而水位对位移的影响权重是持续下降的, 而累计降雨对位移的影响权重基本保持不变.

(2) 改进的变系数回归模型: 在变系数模型中, 各个变量的系数是根据公式(6)得到的, 由于其本质上是属于局部加权, 在距离估计点比较近的点所占

表 1 变系数回归模型和改进的变系数回归模型预测结果

Table 1 Prediction results from varying coefficient regression and improved varying coefficient regression

日期	变系数回归			改进的变系数回归		
	原始值(mm)	预测值(mm)	误差(%)	原始值(mm)	预测值(mm)	误差(%)
ZG85						
2010-11-12	2 330.8	2 316.532	−0.612 16	2 330.8	2 325.660 762	−0.220 490
2010-12-11	2 339.3	2 321.932	−0.742 43	2 339.3	2 344.587 696	0.226 038
2011-01-14	2 349.9	2 316.524	−1.420 30	2 349.9	2 355.631 293	0.243 895
2011-02-21	2 369.6	2 330.334	−1.657 08	2 369.6	2 389.211 497	0.827 629
2011-03-13	2 390.5	2 337.750	−2.206 66	2 390.5	2 419.628 201	1.218 498
2011-04-09	2 420.1	2 344.962	−3.104 77	2 420.1	2 451.231 368	1.286 367
2011-05-12	2 493.8	2 389.472	−4.183 51	2 493.8	2 519.672 382	1.037 468
2011-06-12	2 653.8	2 435.680	−8.219 18	2 653.8	2 596.337 785	−2.165 280
ZG111						
2008-01-11	662.4	672.041 2	1.455 49	662.4	672.410 5	1.511 24
2008-02-16	668.0	666.478 6	−0.227 76	668.0	666.906 2	−0.163 74
2008-03-11	672.2	660.134 2	−1.794 98	672.2	660.290 9	−1.771 66
2008-04-10	671.8	660.483 5	−1.684 50	671.8	660.183 1	−1.729 22
2008-05-12	683.8	681.754 7	−0.299 11	683.8	680.924 2	−0.420 56
2008-06-18	691.6	691.406 7	−0.027 96	691.6	690.482 8	−0.161 54
2008-07-13	706.0	709.959 6	0.560 86	706.0	709.399 2	0.481 47
2008-08-15	719.9	743.671 2	3.302 01	719.9	744.446 6	3.409 72
2008-09-17	816.9	789.395 1	−3.366 99	816.9	792.501 8	−2.986 68
2008-10-20	826.2	816.195 9	−1.210 86	826.2	822.810 1	−0.410 30
2008-11-21	823.7	827.996 7	0.521 63	823.7	838.704 7	1.821 62
2008-12-21	828.4	826.307 4	−0.252 61	828.4	838.750 1	1.249 41
2009-01-08	831.2	821.230 5	−1.199 41	831.8	831.200 0	−0.081 76

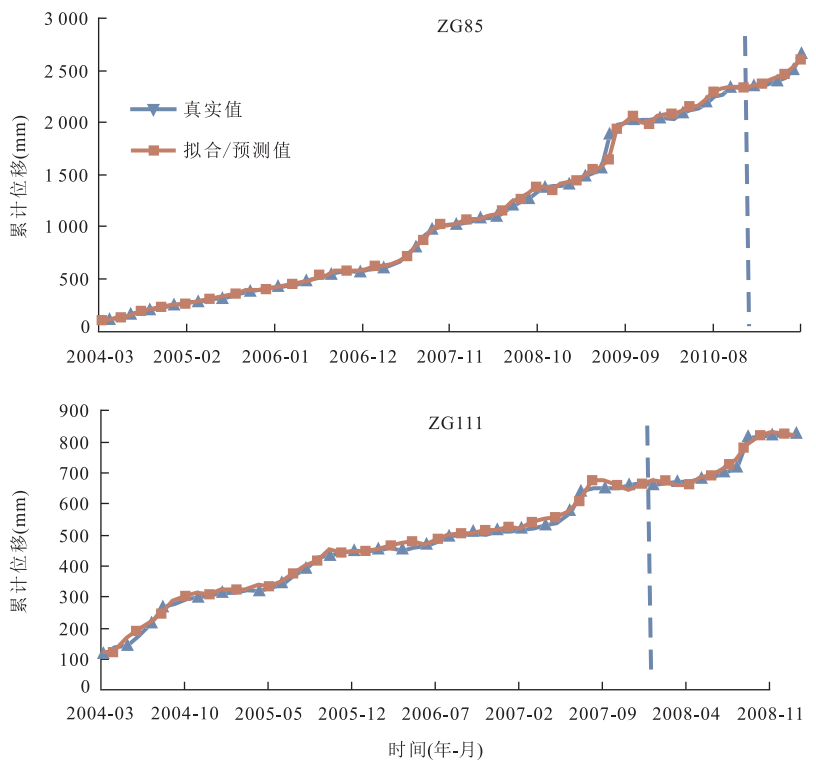


图 6 监测点累计位移真实值与拟合/预测值对比

Fig.6 The comparison of true values and fitting/predicted values of monitoring points

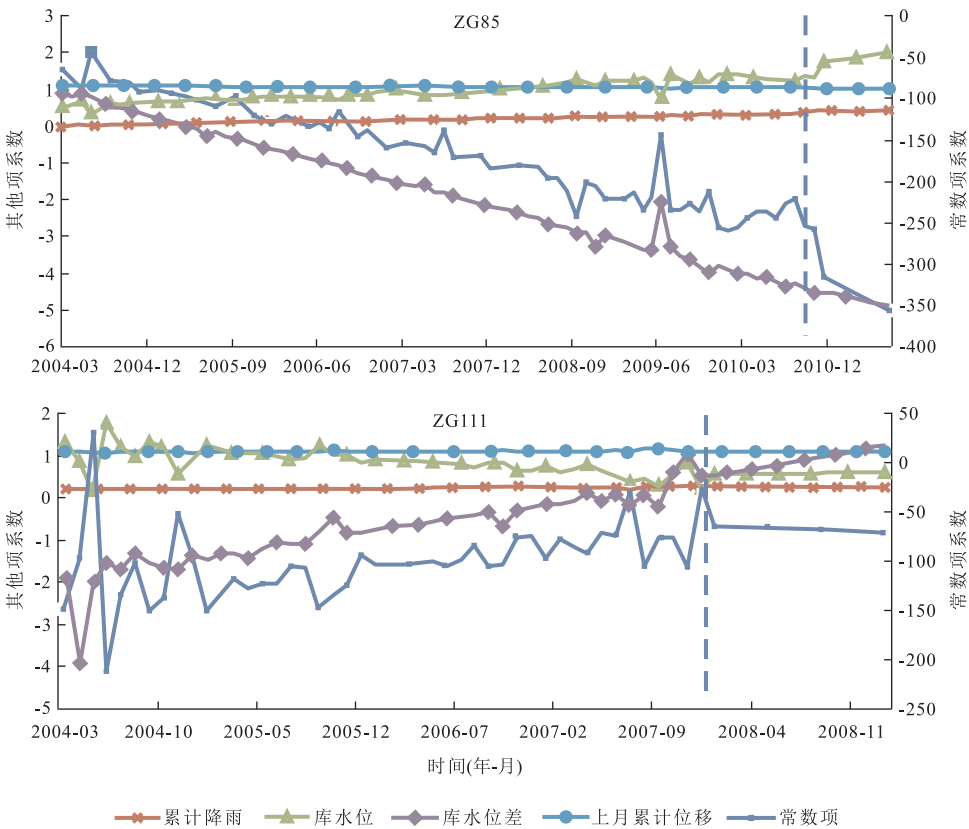


图 7 各自变量因素的系数变化

Fig.7 Coefficient changes of each independent variable factor

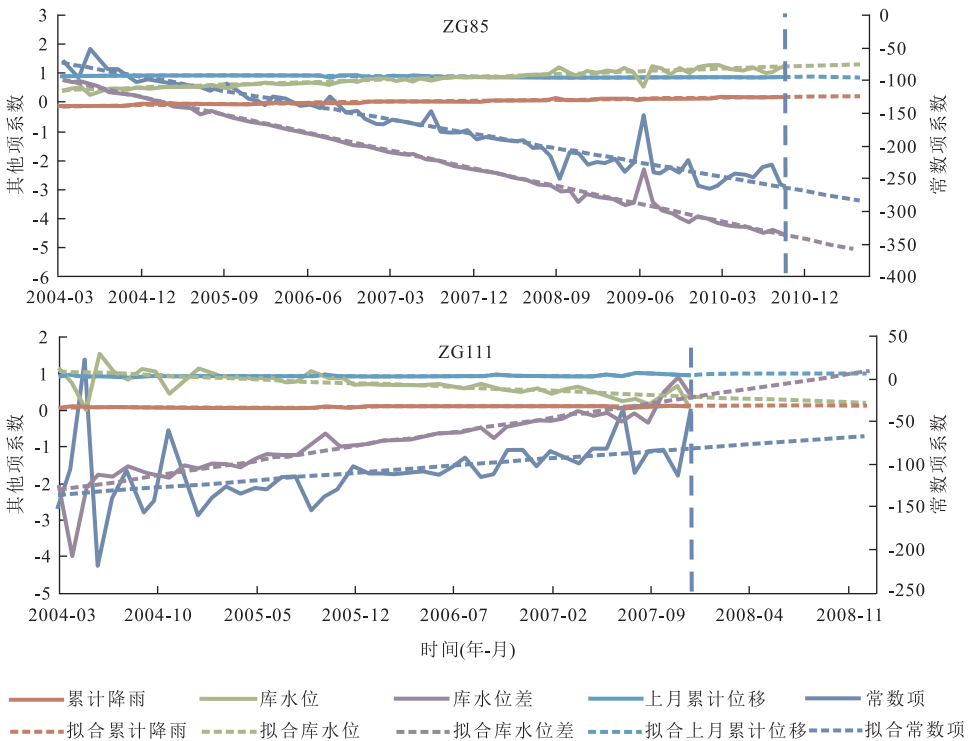


图 8 各自变量因素系数变化及预测值

Fig.8 Predicted value and coefficient changes of each independent variable factor

表 2 监测点 5 种模型的预测对比

Table 2 Prediction and comparison of five models

模型	ZG85				ZG111			
	最大误差	最小误差	平均绝对误差	均方根误差	最大误差	最小误差	平均绝对误差	均方根误差
ARIMA	4.09	0.88	2.29	61.74	7.51	0.210	4.60	36.43
SVR	6.80	0.38	1.70	68.04	6.80	0.380	1.70	68.04
线性回归	3.24	0.25	1.25	37.21	7.90	0.260	1.36	19.07
变系数回归	8.21	0.61	2.76	93.57	3.36	0.027	1.22	12.19
改进的变系数回归	1.28	0.22	0.63	19.23	3.40	0.080	1.24	12.22

权重比较大,在距离估计点比较远的点所占权重比较小.然而在图 7 中可以发现,在预测时各个自变量的斜率几乎是不变的,因此,对于这两个案例而言,可以将拟合时的系数通过最小二乘拟合成一次函数.用此方法得到的各自变量因素系数变化及预测值如图 8 所示.通过此方法得到的预测结果如表 1 所示.

通过表 1 可以看出,监测点 ZG85 预测的精度有所提高,最大误差从 8.2%降到了 2.1%,平均误差也同样有所下降;对于监测点 ZG111,在预测的点中,运用传统的变系数回归对预测点的误差平方和为 1 932.71,而通过改进的变系数回归对预测点的误差平方和为 1 941.28,两者相差不大.然而观察 2008 年 9 月,在该月内滑坡位移增加了接近 100 mm,此处是一个突变点,而突变点的预测对于预测滑坡、预警预报、帮助专业人员决策有比较重要的意义.通过传统的变系数回归对于此突变点的预测误差为 3.36%,改进的变系数回归对突变点预测误差为 2.98%,相比较而言更好一点.并且无论是传统的变系数模型还是改进的变系数模型,都比 Du *et al.* (2009, 2013) 采用 BP 神经网络模型得到的 4.4% 的最大误差要好.

本文在对树坪 ZG85 监测点与八字门 ZG111 监测点使用变系数回归模型进行预测的同时,也使用了自回归积分滑动平均模型(auto-regressive integrated moving average model,简称 ARIMA)模型、支持向量机(support vector machine,简称 SVM)模型、传统的线性回归模型来进行对比实验,表 2 显示了 ZG85 与 ZG111 使用这些模型的对比结果.

其中 ARIMA 模型与 SVR 模型,都是在不断调整各种参数(例如 ARIMA 的模型阶数;SVR 的核函数、惩罚参数等)后选出最优结果.可以看出,在这几种模型中,改进的变系数回归模型预测效果最好,其平均绝对误差与均方根误差均为最小.

4 结论

(1)变系数回归通过假定回归模型的系数为另一个自变量 U 的未知函数,从而既能够在一定程度上满足非参数回归模型的灵活性和适应性,又避免了拟合过程中出现维数灾难的问题.滑坡变形过程的非线性特征明显,适于变系数回归分析和预测.

(2)考虑到外在因素对滑坡位移的延迟效应,通过对累积降雨、库水位、库水位差等影响因素进行交叉确认分析,得出这些影响因素取值情况较优的排列组合.

(3)通过对传统的变系数回归模型进行改进,将拟合的系数通过最小二乘拟合,预测精度明显优于 ARIMA、SVR 等模型,也优于文献中的神经网络方法;同时,更为可贵的是,该方法可以识别出各影响因素的时变规律,从而具有一定的机理可解释性.

References

Du, J., Yin, K. L., et al., 2009. Study of Displacement Prediction Model of Landslide Based on Response Analysis of Inducing Factors. *Chinese Journal of Rock Mechanics and Engineering*, 28 (9): 1783 — 1789. doi: 10.3321/j.issn:1000-6915.2009.09.007

Du, J., Yin, K. L., Lacasse, S., 2012. Displacement Prediction in Colluvial Landslides, Three Gorges Reservoir, China. *Landslides*, 10 (2): 203 — 218. doi: 10.1007/s10346-012-0326-8

Fan, J. Q., Huang, T., 2005. Profile Likelihood Inferences on Semiparametric Varying-Coefficient Partially Linear Models. *Bernoulli*, 11 (6): 1031 — 1057. doi: 10.3150/bj/1137421639

Hastie, T., Tibshirani, R., 1993. Varying-Coefficient Models. *Journal of the Royal Statistical Society Series*, 55 (4): 757 — 796.

Huang, F. M., Yin, K. L., Zhang, G. R., et al., 2015. Landslide Groundwater Level Time Series Prediction Based on Phase Space Reconstruction and Wavelet Analysis—

- Support Vector Machine Optimized by Pso Algorithm. *Earth Science*, 40 (7): 1254—1265 (in Chinese with English abstract).
- Lian, C., Zeng, Z. G., Yao, W., et al., 2012. Displacement Prediction Model of Landslide Based on a Modified Ensemble Empirical Mode Decomposition and Extreme Learning Machine. *Natural Hazards*, 66 (2): 759—771. doi: 10.1007/s11069—012—0517—6
- Lian, C., Zeng, Z. G., Yao, W., et al., 2013. Ensemble of Extreme Learning Machine for Landslide Displacement Prediction Based on Time Series Analysis. *Neural Computing and Applications*, 24 (1): 99—107. doi: 10.1007/s00521—013—1446—3
- Lian, C., Zeng, Z. G., Yao, W., et al., 2014. Extreme Learning Machine for the Displacement Prediction of Landslide under Rainfall and Reservoir Level. *Stochastic Environmental Research and Risk Assessment*, 28 (8): 1957—1972. doi: 10.1007/s00477—014—0875—6
- Lu, C. L., Kuang, C. L., Dai, W. J., et al., 2014. Extracting Seasonal Signals from Continuous GPS Time Series Based on Varying-Coefficient Regression Models. *Journal of Geodesy and Geodynamics*, 34 (5): 94—100 (in Chinese with English abstract).
- Lu, J., Dai, W. J., Zhang, Z. T., 2015. Modeling Dam Deformation Using Varying Coefficient Regression. *Geomatics and Information Science of Wuhan University*, 40 (1): 139—142 (in Chinese with English abstract).
- Lu, S. Q., Yi, Q. L., Yi, W., et al., 2014. Study on Dynamic Deformation Mechanism of Landslide in Drawdown of Reservoir Water Level—Take Baishuihe Landslide in Three Gorges Reservoir Area for Example. *Journal of Engineering Geology*, 22 (5): 869—875 (in Chinese with English abstract).
- Luo, H. M., Tang, H. M., Zhang, G. C., et al., 2008. The Influence of Water Level Fluctuation on the Bank Landslide Stability. *Earth Science*, 33 (5): 687—692 (in Chinese with English abstract).
- Mei, C. L., Wang, N., 2012. Modern Regression Analysis and Method. The Science Publishing Company, Beijing (in Chinese).
- Peng, L., Niu, R. Q., Yang, Y. N., et al., 2013. Landslide Displacement Prediction Based on Kernel Principal Component Analysis and Particle Swarm Support Vector Machine. *Journal of Wuhan University (Information Science Edition)*, 38 (2): 148—152 (in Chinese).
- Yi, Q. L., Zeng, H. E., Huang, H. F., 2013. Reservoir Landslide Deformation Forecast Using BP Neural Network. *Hydrogeology & Engineering Geology*, 40 (1): 124—128 (in Chinese with English abstract).

附中文参考文献

- 黄发明, 殷坤龙, 张桂荣, 等, 2015. 基于相空间重构和小波分析—粒子群向量机的滑坡地下水位预测. *地球科学*, 40 (7): 1254—1265.
- 卢辰龙, 匡翠林, 戴吾蛟, 等, 2014. 采用变系数回归模型提取 GPS 坐标序列季节性信号. *大地测量与地球动力学*, 34 (5): 94—100.
- 卢骏, 戴吾蛟, 章浙涛, 2015. 大坝变形变系数回归建模. *武汉大学学报(信息科学版)*, 40 (1): 139—142.
- 卢书强, 易庆林, 易武, 等, 2014. 库水下降作用下滑坡动态变形机理分析——以三峡库区白水河滑坡为例. *工程地质学报*, 22 (5): 869—875.
- 罗红明, 唐辉明, 章广成, 等, 2008. 库水位涨落对库岸滑坡稳定性的影响. *地球科学*, 33 (5): 687—692.
- 梅长林, 王宁, 2012. 近代回归分析方法. 北京: 科学出版社.
- 彭令, 牛瑞卿, 杨艳南, 等, 2013. 基于核主成分分析和粒子群优化支持向量机的滑坡位移预测. *武汉大学学报(信息科学版)*, 38 (2): 148—152.
- 易庆林, 曾怀恩, 黄海峰, 2013. 利用 BP 神经网络进行水库滑坡变形预测. *水文地质工程地质*, 40 (1): 124—128.