

doi:10.3799/dqkx.2015.105

基于相空间重构和小波分析一粒子群 向量机的滑坡地下水位预测

黄发明¹, 殷坤龙^{2*}, 张桂荣³, 周春梅⁴, 张俊²

1. 中国地质大学地质调查研究院, 湖北武汉 430074

2. 中国地质大学工程学院, 湖北武汉 430074

3. 南京水利科学研究所, 南京江苏 210029

4. 武汉工程大学资源与土木工程学院, 湖北武汉 430073

摘要: 预测滑坡地下水位的动态演变过程对滑坡稳定性分析具有重要意义, 三峡库区库岸滑坡地下水位时间序列受多种因素影响, 呈现出高度非线性非平稳的特征. 为对其进行预测, 提出一种基于相空间重构的小波分析一粒子群优化支持向量机(wavelet analysis-support vector machine, 简称 WA-PSVM)模型. 该模型引入小波变换法对地下水位序列进行时频分解, 将非平稳的地下水位序列转变为多个不同分辨率尺度下的较平稳的地下水位子序列; 然后重构各子序列的相空间, 再利用 PSVM(全称 support vector machine)模型对地下水位各子序列进行预测, 最后将各子序列预测值相加得到最终预测结果. 以三峡库区三舟溪滑坡前缘 STK-1 水文孔日平均地下水位序列为例, 首先分析滑坡前缘地下水位变化的影响因素, 再将 WA-PSVM 模型应用于地下水位预测, 并与单独 PSVM 模型和小波分析-BP 网络模型(wavelet analysis-back propagation, 简称 WA-BP)作对比. 结果表明: 滑坡前缘地下水位受降雨和库水位影响较大, 利用 WA-PSVM 模型对 STK-1 水文孔地下水位进行预测的均方根误差为 0.073 m、拟合优度为 0.966, WA-PSVM 模型预测精度高于单独 PSVM 模型和 WA-BP 模型. WA-PSVM 模型解决了地下水位序列非线性非平稳的问题, 在不考虑影响因素的情况下能获得满意的预测效果, 具有较高的建模效率和较强的实用性.

关键词: 库岸滑坡; 地下水位时间序列; 相空间重构; 小波分析; 粒子群算法; 支持向量机; 地下水; 地质灾害.

中图分类号: P694

文章编号: 1000-2383(2015)07-1254-12

收稿日期: 2014-11-18

Landslide Groundwater Level Time Series Prediction Based on Phase Space Reconstruction and Wavelet Analysis-Support Vector Machine Optimized by PSO Algorithm

Huang Faming¹, Yin Kunlong^{2*}, Zhang Guirong³, Zhou Chunmei⁴, Zhang Jun²

1. Institute of Geological Survey, China University of Geosciences, Wuhan 430074, China

2. Faculty of Engineering, China University of Geosciences, Wuhan 430074, China

3. Nanjing Hydraulic Research Institute, Nanjing 210029, China

4. School of Resource and Civil Engineering, Wuhan Institute of Technology, Wuhan 430073, China

Abstract: It is of great significance to predict the dynamic evolution process of landslide underground water level for landslide stability analysis. For the problem that the evolution process of groundwater level in reservoir landslide is a highly non-linear and non-stationary time series affected by many factors, to predict landslide groundwater level time series, a coupling model based on phase space reconstruction and wavelet analysis-support vector machine (WA-PSVM) optimized by particle swarm

基金项目: 中国地质调查局县域地质灾害风险管理研究项目(No.1212011220173); 国家自然科学基金项目(Nos.201271031415, 41240023, 41302230); 武汉市晨光计划项目(No.201271031415).

作者简介: 黄发明(1988-), 男, 博士, 主要研究方向为滑坡灾害预测预报. E-mail: huang1503518@sina.cn

* **通讯作者:** 殷坤龙, E-mail: yinklong@163.com

引用格式: 黄发明, 殷坤龙, 张桂荣, 等, 2015. 基于相空间重构和小波分析一粒子群向量机的滑坡地下水位预测. 地球科学——中国地质大学学报, 40(7): 1254-1265.

optimization is proposed. Firstly, the groundwater level time series was decomposed into several different frequency components to transform the non-stationary groundwater level time series into stationary time series. Secondly, the PSVM model was established for each component prediction based on the phase-space reconstruction. At last, the final prediction result was obtained by adding the predicted values of all frequency components. Taking daily average groundwater level time series of STK-1 hydrology hole on Sanzhouxi Landslide in the Three Gorges Reservoir Area for example, the influencing factors of landslide groundwater level fluctuation were analyzed and WA-PSVM model was used to predict the STK-1 groundwater level values. Meanwhile, the single PSVM model and wavelet analysis-back propagation neural network (WA-BP) model were also used for groundwater level prediction. The results show that reservoir water level fluctuation and rainfall are the main factors of groundwater level fluctuation in the reservoir landslide leading edge. We also find that the root-mean-square error (RMSE) of the proposed model for groundwater level time series prediction in STK-1 hydrology holes is 0.073 m, the goodness of fit is 0.966, respectively. The prediction accuracy of WA-PSVM model is higher than the single PSVM model and WA-BP model. What is more, WA-PSVM model solves the non-linear and non-stationary problem. WA-PSVM model also has a high operating efficiency and strong applicability without considering the impacts of reservoir water level fluctuation and seasonal rainfall.

Key words: reservoir landslide; groundwater level time series; phase-space reconstruction; wavelet analysis; particle swarm optimization; support vector machine; groundwater; geological hazard.

0 引言

我国滑坡灾害频发,加强对典型滑坡体的地质调查、监测和治理以及提高滑坡预警准确率,具有重大的经济和社会意义.自2003年三峡库区蓄水以来,大量库岸边坡前缘被淹没,库水位涨落和季节性强降雨引起边坡地下水位剧烈波动,导致边坡稳定性普遍下降(罗红明等,2008).地下水位变化引起边坡稳定性下降主要原因为:降雨入渗和库水位上升导致滑坡地下水位上升,造成坡体抗剪强度降低(王智磊等,2011);而地下水位的迅速下降,又会增加滑体中的动水压力,加速滑坡失稳变形的发展(刘新喜等,2005).因此对库岸滑坡地下水位动态变化过程进行预测具有重大意义.

三峡库区库岸滑坡地下水位的动态演变过程是一个受边坡水文地质条件控制,并受降雨、库水位等多种因素共同作用的非线性动力系统(彭令等,2012).本文在库岸滑坡前缘地下水位影响因素实例分析中,发现前缘地下水位主要受降雨和库水位的影响.目前还很难构建物理模型对这类复杂水文现象进行完全描述,需要借助数理统计、时间序列分析等方法来建模分析,以弥补物理模型的不足,但是传统线性模型难以模拟非线性的水文时间序列.近年来,随着非线性时间序列预测的发展,有效刻画复杂非线性系统的混沌理论为这类水文现象的研究提供了一种新思路.嵌入定理认为,在决定性系统长期演化过程中,其任一变量的发展历程均包含了所有变量长期演化的信息.即不需要考虑降雨、库水位等外部因素的影响,只需对地下水位混沌时间序

列选取合理的延迟时间和嵌入维,单变量的地下水位序列即可较好地重构原动力系统的相空间,并对地下水位未来发展趋势进行有效预测(Siek and Solomatine,2010).杨忠平等(2005)、Maiti and Tiwari (2014)建立单变量模型对地下水位进行预测,取得了较好的预测效果,但是他们未能详细阐述建模所需的相空间重构理论基础.针对这一问题,本文首先重建地下水位的相空间,再模拟其发展演化规律.

在实际的预测中,单变量混沌时间序列模型预测效果往往不够理想,原因在于滑坡地下水位等时间序列具有复杂的非线性非平稳特征,严重干扰了模型的预测效果.针对这一问题,本文引入小波变换法对地下水位序列进行分析.在时频两域都具有表征时间序列局部特征能力的小波变换法,被用于解决多种非线性非平稳时间序列的预测问题.李长冬等(2008)将小波变换法用于地基沉降预测,取得了较好的应用效果;刘辉等(2011)将小波分析用于非平稳风速信号的短期预测,提高了神经网络模型的预测精度;此外小波分析在径流量预测(Ren *et al.*, 2011)、电力需求预测(An *et al.*, 2011)和模式识别(任俊等,2005)、滑坡位移预测(Li *et al.*, 2009)等领域也获得了成功的应用.小波分析对非平稳时间序列具有多尺度分解能力,它可将非平稳时间序列分解成包含低频信号的趋势项子序列以及包含高频细节信号的周期项子序列和随机项子序列,各项子序列变得更加平稳,信噪比也得到了提高,因此能有效提高模型预测精度.本文将小波分析应用于非线性非平稳的滑坡地下水位序列预测中.

目前常用于非线性时间序列预测的数学模型主

要有神经网络模型(周翠英等,2005)、支持向量机(彭令等,2013)等,但传统神经网络模型如 BP(全称 back propagation)神经网络等,存在局部最优、泛化性能差的缺点.支持向量机 SVM(全称 support vector machine)模型在解决复杂非线性回归问题时效果良好,但也存在核函数和参数选取较困难等问题.粒子群向量机 PSVM(全称 particle swarm vector machine)模型采用全局搜索能力强大的粒子群算法优化选取 SVM 模型的参数,克服了传统神经网络和 SVM 模型存在的问题,具有全局最优、泛化性能较好等优点.PSVM 模型已在多个领域得到了广泛的应用(邱景平等,2012;谷文成等,2014),因此,在对地下水水位序列进行小波分析之后,本文将 PSVM 模型应用于滑坡地下水水位混沌序列预测.

综上所述,本文结合小波分析的时间序列分解功能和 PSVM 模型预测精度高的优点,提出基于相空间重构的 WA-PSVM(全称 wavelet analysis-support vector machine)预测模型.以三峡库区三舟溪库岸滑坡前缘 STK-1 水文孔日平均地下水水位时间序列为例,首先对其影响因素进行分析,揭示其表现出非线性非平稳特征的深层次原因,再探讨 WA-PSVM 模型对地下水水位值的预测性能,期待为滑坡地下水水位时间序列预测问题提供一种新思路.

1 基于相空间重构的 WA-PSVM 模型原理

WA-PSVM 模型建模较简单,思路也比较清晰,具体过程如下:第 1 步,将原始地下水水位序列归一化至 $[0, 1]$,再用小波分析法将其分解为低频子序列 aN 和高频子序列 $dt(t=1, 2, 3, \dots, N, t$ 为分解的尺度数);第 2 步,确定各子序列最佳嵌入维数值,对各子序列进行相空间重构;第 3 步,构建 PSVM 模型并分别对各个子序列进行预测;第 4 步,将各子序列的预测值相加得到地下水水位序列的最终预测结果.建模流程如图 1 所示.

1.1 地下水水位序列相空间重构

Tanks 定理认为,对于混沌系统,可以从单变量时间序列中重构出原始系统的相空间,因为时间序列本身蕴含了参与此动力系统的全部变量的有关信息.不失一般性,设某一混沌时间序列为 $X_i(i=1, 2, \dots, n)$,选取恰当的延迟时间 τ 和嵌入维数 m ,在此基础上重构其相空间为:

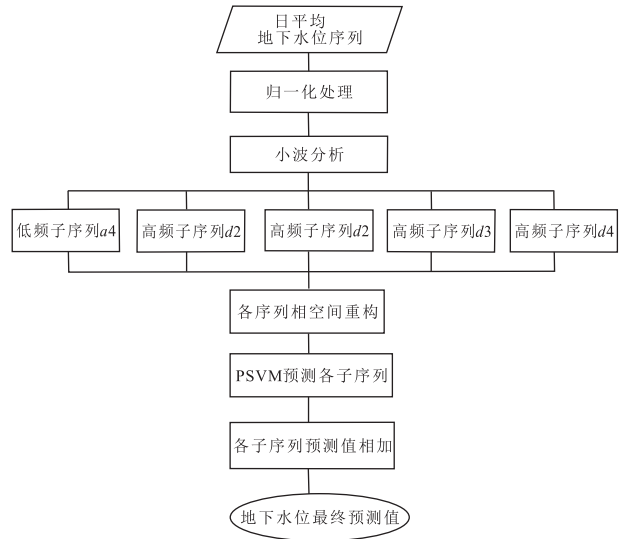


图 1 地下水水位预测 WA-PSVM 模型建模流程

Fig.1 Flowchart of WA-PSVM model for groundwater level prediction

$$y_t = \{x_t, x_{t+\tau}, x_{t+2\tau}, \dots, x_{t+(m-1)\tau}\}, \quad (1)$$

式中: $t=1, 2, \dots, n-(m-1)t$, y_t 是 m 维相空间中的相点.根据 Takens 嵌入定理,如果嵌入维数 $m > 2d+1$, d 是动力系统的维数,则重构的动力系统和原系统是拓扑等价的.因而存在一个反映一步预测规律的光滑映射函数 $F: R^m \rightarrow R$,使得:

$$y_{t+1} = F\{x_t, x_{t-\tau}, \dots, x_{t-(m-1)\tau}\}, \quad (2)$$

此时 $y_t \rightarrow y_{t+1}$ 的函数映射关系反映了原始未知动力系统的演化.理论上满足上式的 F 是唯一的,但是实际上可用数据总是有限的,因而不可能真正求得 F ,而只能根据有限的的数据构造映射 $G: R^m \rightarrow R$,使得映射 G 充分逼近 F .

相空间重构是否成功的关键在于 τ 和 m 的选取是否合理.非线性预测研究指出,重构相空间时, τ 即使未能取最佳延时,也只会影响重构吸引子的欧几里得几何形状进而影响关联维数的计算.对重构的吸引子无歧义地反映原始系统的动力学性质并无影响(Wolf et al., 1986).反而当 τ 取值较大时,非线性模型需要拟合的关系复杂,难度高,所以相空间重构的关键参数是嵌入维. τ 应视情况取较小的值,本文各个模型对所有子序列预测时均取 $\tau=1$ (黄发明和田玉刚,2014).

计算嵌入维值主要有 G-P 法(张淑清等,2010)、Cao 法、假邻近点法(韩敏,2007)等.本文采用应用范围广泛且实现较简单的假邻近点法计算原始地下水水位序列的嵌入维.其基本原理为:当 m 较

小时吸引子在相空间中不能完全打开,此时吸引子就会在嵌入空间产生一些投影点,投影点与相空间中的其他点形成最近邻点,但是在原系统中这 2 个点并非真正的最近邻点,因此称为虚假邻点.而当嵌入维数从 m 增加至 $m + 1$ 时,再判断轨线的邻点中哪些是真实的邻点和哪些是虚假的邻点,如果没有虚假邻点时,可以认为在此嵌入维数下完全打开了几何结构,这个嵌入维值即为最佳嵌入维数.

1.2 小波分析

小波分析是通过窗口大小随分解水平而改变的小波函数对信号进行时域局部化分析的,即把某一基本小波函数 $\psi(t)$ 平移 τ 后,再在不同尺度 a 下与待分析的信号 $x(t)$ 做内积:

$$f_x(a, \tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \psi\left(\frac{t-\tau}{a}\right) dt, a > 0. \quad (3)$$

等效的时域表达式为:

$$f_x(a, \tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(\omega) \psi(a\omega) e^{j\omega\tau} dt, a > 0. \quad (4)$$

唐明等(2007)采用 Mallat 算法将系统边际电价时间序列分解为不同尺度序列并进行预测,获得较好预测效果;且发现实际小波分解层数不宜太大或太小,以 4 层为宜.本文也采用 Mallat 算法对滑坡地下水位序列进行多尺度分解,对于小波分解的级数,也将滑坡地下水位序列分解为 4 层.

1.3 粒子群优化向量机模型

SVM 模型预测性能与其误差惩罚因子 c 、不敏感损失系数 ϵ 和核函数参数 φ 密切相关,本文利用 PSO(全称 particle swarm optimization)算法在某一范围内搜寻 SVM 参数的最佳组合,PSVM 模型计算流程如图 2 所示.

1.3.1 粒子群算法基本原理

PSO 算法是一种模拟群体智能的参数选取法(Kennedy and Eberhart, 1995).假设 D 维搜索空间中有 n 个粒子组成的种群 $X = (X_1, X_2, \dots, X_n)$, 其中第 i 个粒子表示为一个 D 维的向量 $\mathbf{X}_i = (X_{i1}, X_{i2}, \dots, X_{iD})^T$, 代表第 i 个粒子在 D 维搜索空间中的位置,即问题的一个潜在解.根据目标函数即可计算出每个粒子位置 X_i 对应的适应度值.第 i 个粒子的速度为 $V_i = (V_{i1}, V_{i2}, \dots, V_{iD})^T$, 其个体极值为 $P_i = (P_{i1}, P_{i2}, \dots, P_{iD})^T$, 种群的全局极值为 $P_g = (P_{g1}, P_{g2}, \dots, P_{gD})^T$.在每一次迭代过程中粒子通过个体极值和全局极值更新自身的速度和位置,更新公式如下:

$$V_{kd}^{k+1} = \omega V_{kd}^k + c_1 r_1 (P_{id}^k - X_{id}^k) + c_2 r_2 (P_{gd}^k - X_{gd}^k), \quad (5)$$

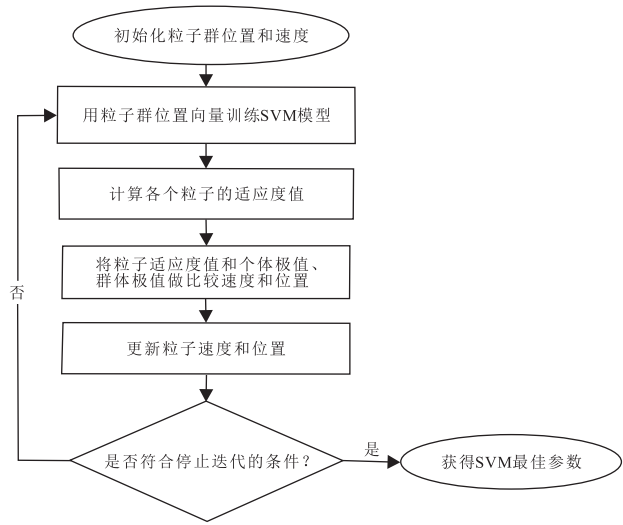


图 2 粒子群优化支持向量机流程

Fig.2 Flowchart of PSVM model

$$X_{kd}^{k+1} = X_{id}^k + V_{id}^{k+1}, \quad (6)$$

式中: ω 为惯性权重; $d = 1, 2, \dots, D$; $i = 1, 2, \dots, n$; k 为当前迭代次数; V_{id} 为粒子的速度; c_1, c_2 为非负的常数; r_1, r_2 为分布于 $[0, 1]$ 之间的随机数, 一般情况下, 粒子位置和速度都会被限制在一定的区间内.

1.3.2 支持向量机基本原理

支持向量机(SVM)模型(崔万照等, 2004)回归预测性能较好, 对于线性拟合情况, 函数首先考虑用线性回归函数 $f(x) = \omega \cdot x + b$ 拟合 $(x_i, y_i), i = 1, 2, \dots, n, x_i \in R^n$ 为输入量, $y_i \in R$ 为输出量, 即需要确定 ω 和 b . 标准 SVM 采用 ϵ -不灵敏度函数, 即假设所有训练数据在精度 ϵ 下用线性函数拟合:

$$\begin{cases} y_i - f(x_i) \leq \epsilon + \xi_i, \\ f(x_i) - y_i \leq \epsilon + \xi_i^*, i = 1, 2, \dots, n, \\ \xi_i, \xi_i^* \geq 0, \end{cases} \quad (7)$$

式中: ξ_i, ξ_i^* 是松弛因子, 当划分有误差时, ξ, ξ_i^* 都大于 0, 误差不存在取 0. 这时, 该问题转化为求优化目标函数最小化问题:

$$R(\omega, \xi, \xi^*) = \frac{1}{2} \omega \cdot \omega + C \sum_{i=1}^n (\xi_i + \xi_i^*), \quad (8)$$

式(8)中 $C > 0$ 表示对超出误差 ϵ 的样本的惩罚程度. 求解式(7)和式(8)可看出, 这是一个凸二次优化问题, 所以引入 Lagrange 函数:

$$L = \frac{1}{2} \omega \cdot \omega + C \sum_{i=1}^n (\xi_i + \xi_i^*) - \sum_{i=1}^n \alpha_i [\xi_i + \epsilon - y_i + f(x_i)] - \sum_{i=1}^n \alpha_i^* [\xi_i^* + \epsilon - y_i + f(x_i)] - \sum_{i=1}^n (\xi_i \gamma_i + \xi_i^* \gamma_i^*), \quad (9)$$

式中： $\alpha, \alpha_i^* \geq 0, \gamma_i, \gamma_i^* \geq 0$, 为 Lagrange 乘数， $i=1, 2, \dots, n$. 求函数 L 对 $\omega, b, \xi_i, \xi_i^*$ 的最小化，对 $\alpha_i, \alpha_i^*, \gamma_i, \gamma_i^*$ 的最大化，代入 Lagrange 函数得到对偶形式，最大化函数：

$$W(\alpha, \alpha^*) = \frac{1}{2} \sum_{i=1}^n (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) (x_i \cdot x_j) + \sum_{i=1}^n (\alpha_i - \alpha_i^*) y_i - \sum_{i=1}^n (\alpha_i + \alpha_i^*) \epsilon,$$

及约束条件
$$\begin{cases} \sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0, \\ 0 \leq \alpha_i, \alpha_i^* \leq C, \end{cases} \quad (10)$$

求解式(10)其实也是一个求解二次规划问题，由 Kuhn-Tucker 定理，在鞍点处有：

$$\begin{aligned} \alpha_i [\epsilon + \xi_i - y_i + f(x_i)] &= 0, \\ \xi_i \cdot r_i &= 0, \\ \alpha_i^* [\epsilon + \xi_i^* - y_i + f(x_i)] &= 0, \\ \xi_i^* \cdot r_i^* &= 0. \end{aligned} \quad (11)$$

得出 $\alpha_i \cdot \alpha_i^* = 0$, 表明 α_i, α_i^* 不能同时为零，还可以得出：

$$\begin{aligned} (C - \alpha_i) \xi_i &= 0, \\ (C - \alpha_i^*) \xi_i^* &= 0. \end{aligned} \quad (12)$$

从式(12)可得出，当 $\alpha_i = C$ 或 $\alpha_i^* = C$ 时， $|f(x_i) - y_i|$ 可能大于 ϵ , 与其对应的 x_i 称为边界支持向量 (boundary support vector, 简称 BSV); 当 $\alpha_i^* \in (0, C)$ 时， $|f(x_i) - y_i| = \epsilon$, 即 $\xi_i = 0, \xi_i^* = 0$, 与其对应的 x_i 称为标准支持向量 (normal support vector, 简称 NSV); 当 $\alpha_i = 0$ 且 $\alpha_i^* = 0$ 时，与其对应的 x_i 为非支持向量，它们对 w 没有贡献。因此 ϵ 越大，支持向量数越少。对于标准支持向量，如果 $0 < \alpha_i < C (\alpha_i^* = 0)$, 此时 $\xi_i = 0$, 由式(11)可以求出参数 b . 同样，对于满足 $0 < \alpha_i^* < C (\alpha_i = 0)$ 的标准支持向量，也可以计算出对应的 b . 因此，最后可根据样本点 (x_i, y_i) 求得的线性拟合函数为：

$$f(x) = \omega \cdot x + b = \sum_{i=1}^n (\alpha_i - \alpha_i^*) x_i \cdot x + b. \quad (13)$$

非线性 SVM 的基本思想是通过事先确定的非线性映射将输入向量映射的一个高维特征空间 (Hilbert 空间) 中，然后在此高维空间中再进行线性回归，从而取得在原空间非线性回归的效果。

首先将输入量 x 通过映射 $\Phi: R^n \rightarrow H$ 映射到高维特征空间 H 中用函数 $f(x) = \omega \cdot \Phi(x) + b$ 拟合数据 $(x_i, y_i), i=1, 2, \dots, n$. 则二次规划目标函数

式(10)变为：

$$W = (\alpha, \alpha^*) = -\frac{1}{2} \sum_{i=1, j=1}^n (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) \cdot (\Phi(x_i) \cdot \Phi(x_j)) + \sum_{i=1}^n (\alpha_i - \alpha_i^*) y_i - \sum_{i=1}^n (\alpha_i + \alpha_i^*) \epsilon, \quad (14)$$

式(14)中涉及到高维特征空间点积运算 $\Phi(x_i) \cdot \Phi(x_j)$, 且函数 Φ 未知. SVM 只考虑高维特征空间的点积运算 $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$, 用这个点积运算替代函数 Φ . 则可求的非线性拟合函数的表示式为：

$$f(x) = \omega \cdot \Phi(x) + b = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(x, x_i) + b. \quad (15)$$

本文采用默认的 RBF 核函数，如下式所示：

$$k(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right), \quad (16)$$

式中： x' 为核函数中心， σ 为函数的宽度参数，控制了函数的径向作用范围。

2 三舟溪滑坡概况

2.1 滑坡地质概况

三舟溪滑坡位于重庆市万州区三州村长江左岸，两侧边界为季节性冲沟，前缘至长江岸边，后缘基岩出露，剪出口位于长江水位以下。其前缘高程约为 125 m，后缘高程约 270 m，滑体长约 380 m，平均宽度约为 360 m；滑坡坡度为 $15^\circ \sim 24^\circ$ ，主滑方向为 245° ，滑坡平面和地下水位监测孔布置如图 3 所示。

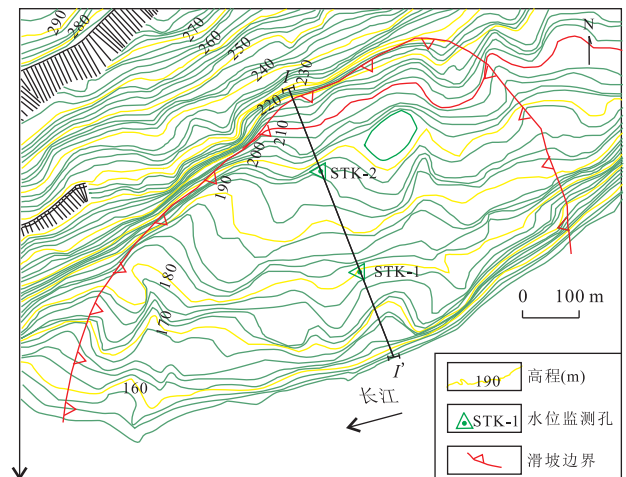


图 3 三舟溪滑坡地下水位监测点布置

Fig.3 Locations of groundwater monitoring points of the Sanzhouxi landslide

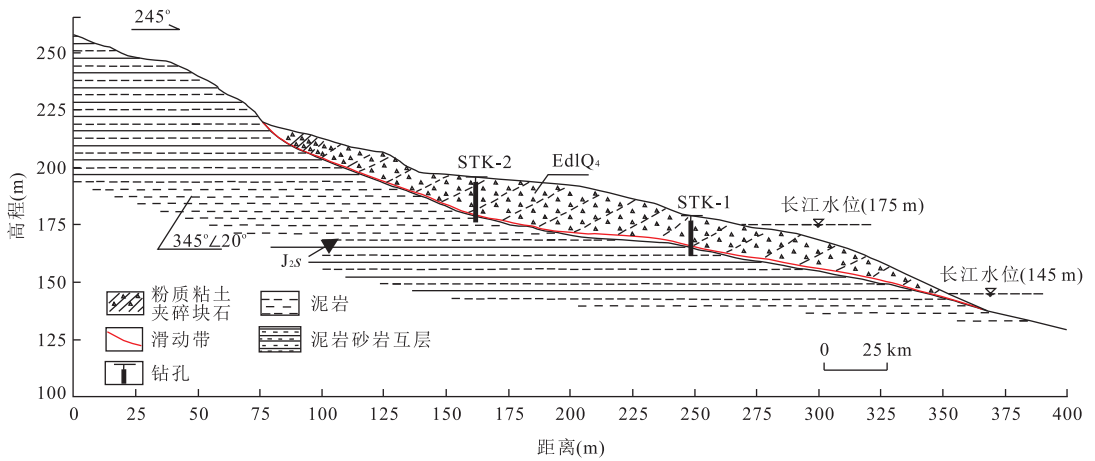


图 4 三舟溪滑坡 I-I' 工程地质剖面

Fig.4 I-I' cross-section of the Sanzhouxi landslide

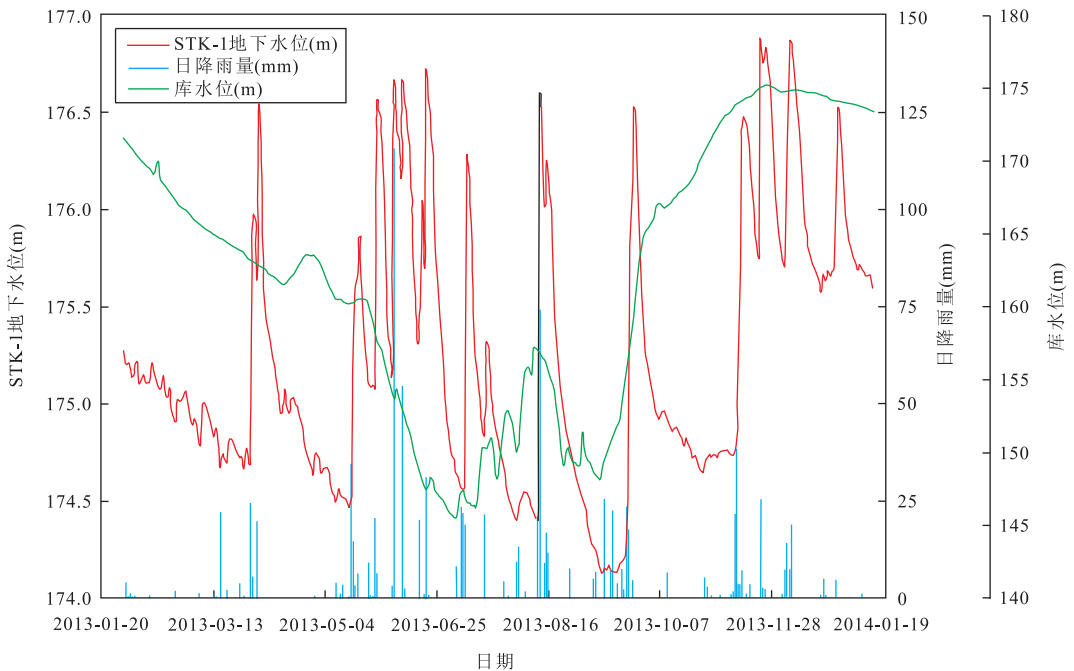


图 5 STK-1 水文孔地下水位值与降雨、库水位相关性分析

Fig.5 Correlation analysis between STK-1 groundwater level and rainfall, reservoir water level

该滑坡上覆第四系残积、坡积的粉质粘土夹碎块石,下伏侏罗系中统上沙溪庙组砂岩、泥岩互层,岩层产状为 $345^{\circ} \angle 20^{\circ}$ 。

滑坡范围内地表水系不发育,没有泉水出露,有季节性冲沟,大气降雨直接排入长江,地下水类型以松散层孔隙水为主,上部主要接受大气降雨补给,经短途径流后向长江排泄,水量贫乏,该滑坡 I-I' 工程地质剖面图如图 4 所示,当长江水位上升至 175 m 时,坡体的三分之一将会被淹没,因此高程为 177 m 的前缘 STK-1 水文孔地下水位受降雨和库水位综合作用影响,高程为 194 m 的中部 STK-1 水文孔地

下水位主要受降雨控制,该滑坡地下水位变化幅度较大,导致其稳定性差,易发生局部或整体变形破坏,因此非常有必要对该滑坡地下水位动态变化过程进行分析预测。

2003 年以来,三舟溪滑坡地表已出现多条裂缝,滑坡前缘出现局部坍塌现象,为了监测滑坡 I-I' 剖面上的地下水位动态变化过程,在滑坡前缘 177 m 和中部 194 m 高程处设置 STK-1 和 STK-2 水文孔。

2.2 STK-1 地下水位影响因素分析

本文以 STK-1 水文孔从 2013 年 1 月 30 日至

2014 年 1 月 14 日共 340 个地下水位时间序列数据用于实例分析,STK-1 水文孔地下水位及其影响因素相关性如图 5 所示,从图 5 中可知,滑坡地下水位的短时剧烈变化主要是由降雨引起的,一次强降雨过程将导致滑坡地下水位值大幅上升;当降雨结束时,地下水位又会快速恢复初始状态,整个过程中地下水位值波动幅度较大.此外,库岸滑坡前缘地下水位值主要受库水位影响,当库水位上升时,滑坡前缘地下水位值呈现上升趋势;当库水位较低时,滑坡前缘地下水位值在未受降雨影响的情况下也处于较低水平.从图 5 中可以看出,降雨和库水位对滑坡前缘地下水的影响均存在延迟效应,比如一次强降雨过后,滑坡体能维持几天较高地下水位值,之后就快速回落;同样滑坡前缘平均地下水位也是在库水位上升之后才经历一个逐步上升的过程.

综上所述,正是在随机性较强的短时降雨序列、周期性波动的库水位序列等因素影响下,滑坡前缘地下水位才表现为一个复杂的非线性非平稳系统.WA-PSVM 模型将地下水位时间序列的趋势性、周期性和随机性特征分离出来,再分别对各特征项进行预测,能有效解决滑坡地下水位非线性非平稳问题.

3 STK-1 水文孔地下水位预测

本文以 STK-1 水文孔 2013 年 1 月 30 日至 2013 年 11 月 15 日共 280 个数据用于模型训练,2013 年 11 月 16 日至 2014 年 1 月 14 日共 60 个地下水位数据用于模型测试.

3.1 地下水位序列小波分解

首先对原始 STK-1 地下水位时间序列 $X_i (i = 1, 2, \dots, N)$ 进行归一化处理,如式(17)所示,式中 X_{\min} 和 X_{\max} 分别为地下水位序列中的最小值和最大值, Z_i 为归一化后的 STK-1 地下水位序列.

$$Z_i = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}}. \quad (17)$$

选择 db_4 小波函数将归一化后的 STK-1 地下水位序列 $Z_i (i = 1, 2, \dots, N)$ 分解为 4 层,得到低频子序列 a_4 和高频子序列 d_1, d_2, d_3, d_4 ,分解结果如图 6 所示.

3.2 相空间重构及 PSVM 模型预测

各子序列延迟时间均选取为 1,经假邻近点法计算得出,原始序列 Z_i 最佳嵌入维为 3,低频子序列 a_4 和高频子序列 d_1, d_2, d_3, d_4 的最佳嵌入维值分别为 3 和 7,6,4,4.再进行相空间重构,即可得到

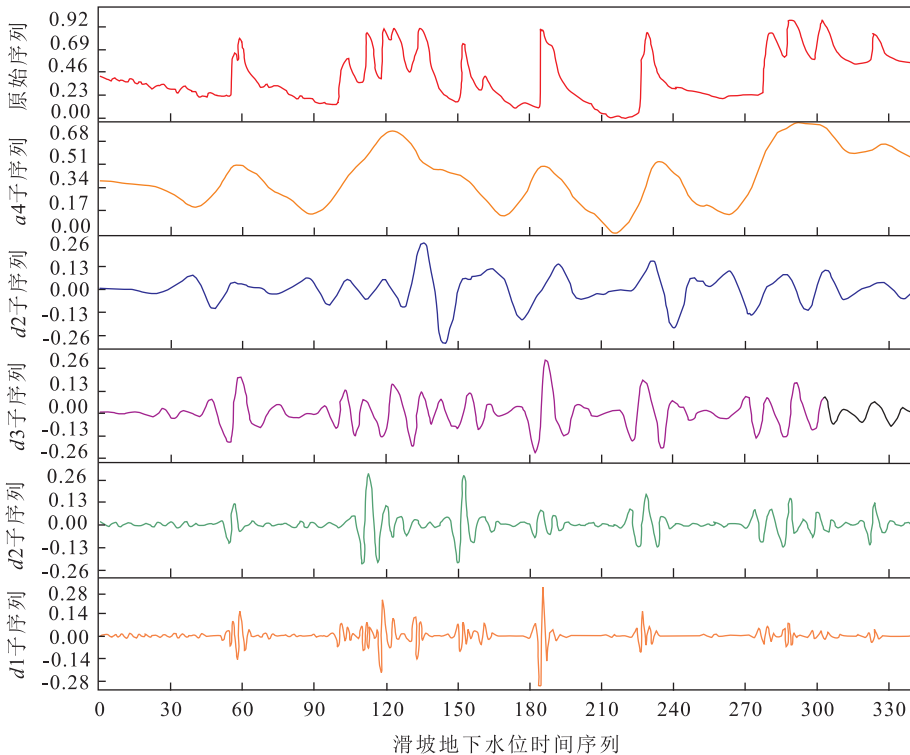


图 6 STK-1 水文孔小波分解后的各个子序列

Fig.6 Wavelet decomposition of STK-1 groundwater series

表 1 WA-PSVM 模型、WA-BP 模型和单独 PSVM 模型的输入输出变量

Table 1 Input variables and output variables for WA-PSVM, WA-BP and single PSVM model

序列	输出变量	输入变量
原始序列	(Z_i)	$(Z_{i-1}, Z_{i-2}, Z_{i-3})$
$a4$	$(a4_i)$	$(a4_{i-1}, a4_{i-2}, a4_{i-3})$
$d1$	$(d1_i)$	$(d1_{i-1}, d1_{i-2}, d1_{i-3}, d1_{i-4}, d1_{i-5}, d1_{i-6}, d1_{i-7})$
$d2$	$(d2_i)$	$(d2_{i-1}, d2_{i-2}, d2_{i-3}, d2_{i-4}, d2_{i-5}, d2_{i-6})$
$d3$	$(d3_i)$	$(d3_{i-1}, d3_{i-2}, d3_{i-3}, d3_{i-4})$
$d4$	$(d4_i)$	$(d4_{i-1}, d4_{i-2}, d4_{i-3}, d4_{i-4})$

表 2 SVM 模型对原始序列和各个子序列进行预测时的最佳参数组合

Table 2 Parameter combinations of SVM model for original groundwater level time series and each component

序列	SVM 模型参数组合
原始序列	$c = 1\ 047.85, \epsilon = 0.012, \varphi = 0.173$
$a4$	$c = 1\ 446.33, \epsilon = 0.011, \varphi = 0.187$
$d1$	$c = 2\ 477.26, \epsilon = 0.078, \varphi = 0.092$
$d2$	$c = 1\ 764.97, \epsilon = 0.009, \varphi = 0.136$
$d3$	$c = 869.33, \epsilon = 0.012, \varphi = 0.045$
$d4$	$c = 2\ 879.52, \epsilon = 0.011, \varphi = 0.038$

各序列用于 WA-PSVM 模型, WA-BP(全称 wavelet analysis-back propagation)模型和单独 PSVM 模型预测的输入输出变量.如下表 1 所示.

本文采用 PSO 算法搜寻 SVM 的参数 c 、 ϵ 和 φ 的最佳组合,其中设定 PSO 算法的适应度函数为 SVM 模型预测结果的均方根误差,每个粒子的维数为 3,种群大小为 100,迭代次数为 100.经过训练得到 SVM 模型对原始序列和各个子序列进行预测时的最佳参数组合,如表 2 所示.最终,WA-PSVM 模型对各个子序列的预测值如图 7 所示.

3.3 WA-PSVM 模型预测结果

另外为了验证本文模型的有效性,将 WA-PSVM 模型和单独 PSVM 模型、WA-BP 神经网络模型作比较.其中单独 PSVM 模型未对 STK-1 地下水位序列进行小波分解,只对归一化之后的原始序列进行相空间重构,再输入 PSVM 模型中进行预测.本文实验所采用的单独 PSVM 模型的最佳嵌入维数值为 3,输入输出变量如表 1 第 1 行所示,最佳参数组合如表 2 第 1 行所示.

本文同时采用 WA-BP 模型对 STK-1 地下水位进行预测时,BP 神经网络隐含层和输出层神经元传递函数分别为 tansig 函数和 purelin 函数,网络训练采用梯度下降算法.BP 网络隐含层节点数也采用

最小预测误差法确定,对各子序列 $a4$ 和 $d1, d2, d3, d4$ 进行预测时网络节点数分别 11 和 20, 22, 15, 12.3 组模型最终预测结果如图 8 所示.

从图 8 中可知,由于滑坡地下水位序列存在非线性非平稳性特征,导致单独 PSVM 模型预测值与实测值存在较大差距.且单独 PSVM 模型预测出的地下水位值与实测值相比往往存在延迟效应,即难以及时捕捉地下水位的动态变化信息,这也是由短时降雨的随机性引起的.而本文提出的 WA-PSVM 模型则能有效地预测出滑坡地下水位的快速变化特征,图 8 中可以看出,WA-PSVM 模型预测值与实测值拟合效果良好.另外 WA-BP 网络模型预测效果尚可,但比 WA-PSVM 模型预测误差更大.

3.4 模型预测精度评定

本文另外采用均方根误差(root mean squared error,简称 RMSE)和拟合优度(R^2)这 2 个指标作为模型预测性能评定标准.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}}, \quad (18)$$

$$R^2 = \frac{\left(\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})\right)^2}{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}, \quad (19)$$

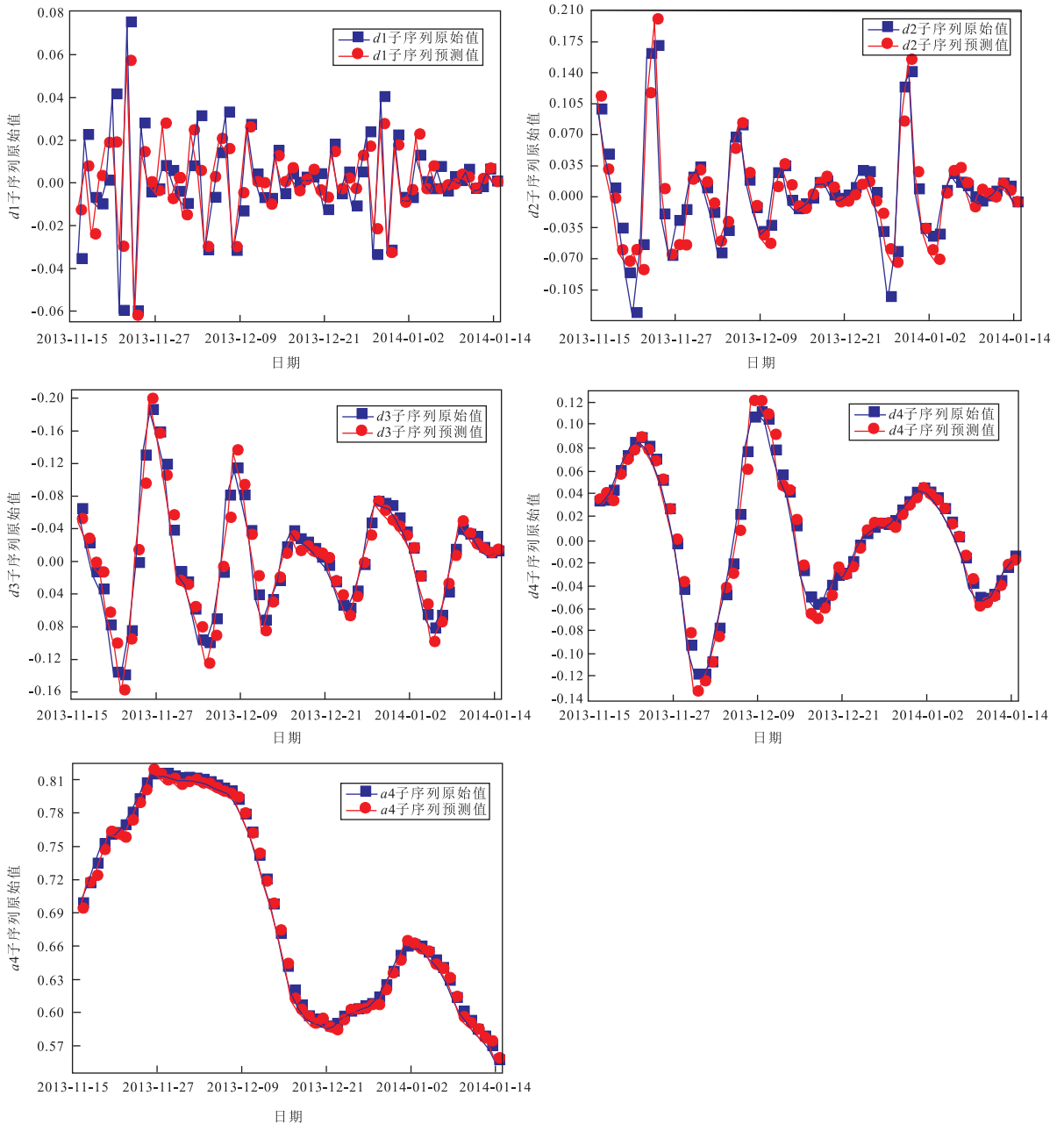
式中: x_i 为真实值, y_i 为预测值, \bar{x} 为 x_i 平均值, \bar{y} 为 y_i 平均值, n 为预测值个数.3 组模型对 STK-1 地下水位预测精度如表 3 所示.

从表 3 可知,运用 WA-PSVM 模型对 STK-1 地下水位预测结果的均方根误差为 0.073 m,拟合优度为 0.966.3 组预测模型对比结果表明,WA-PSVM 模型预测效果优于其他两组模型,且 WA-BP 模型预测精度也高于单独 PSVM 模型.

4 模型讨论

本文在将 WA-PSVM 模型应用于三舟溪滑坡地下水位预测的实例分析中发现,通过小波变换将非平稳的地下水位序列分解成 4 层较为适宜.其中 $a4$ 子序列反映了地下水位序列发展演化的趋势性特征, $d3$ 和 $d4$ 子序列反映了地下水位序列周期性变化的特征,而 $d1$ 和 $d2$ 子序列主要包含的是随机性信息.

相空间重构是 WA-PSVM 模型的理论基础,在对各个子序列 $a4, d1, d2, d3$ 和 $d4$ 进行相空间重

图 7 子序列 $a4$ 和 $d1, d2, d3, d4$ 预测值Fig.7 Predicted values of each frequency series $a4$ and $d1, d2, d3, d4$

构时,延迟时间设置为 1.本文采用假邻近点法计算各子序列的最佳嵌入维,能有效克服最佳嵌入维选取的盲目性.但不足的是,必须对原始地下水水位序列和各个子序列分别进行计算,因为各个子序列的最佳嵌入维数值并不一致.

本文将 WA-PSVM 模型、单独 PSVM 和 WA-BP 网络模型应用 STK-1 水文孔地下水水位序列预测,其中 WA-PSVM 模型预测的精度高于单独 PSVM 模型,表明通过小波变换多尺度分析,非线性非平稳的库岸滑坡地下水水位序列转变成了更加平

稳的多个地下水水位子序列,提高了各个子序列的信噪比,从而使得 PSVM 模型预测效果更优.在同等条件下,WA-PSVM 模型预测效果优于 WA-BP 模型,表明 PSVM 模型的时间序列预测性能优于传统的 BP 网络模型,PSO 算法能有效获取 SVM 模型的参数,减少了人为因素对建模过程的干扰,更能反映地下水水位序列发展演化的趋势.而在传统 BP 模型外推性能不具有优势的条件下,WA-BP 模型预测精度也高于单独 PSVM 模型,更进一步表明小波分析能有效提高模型预测精度.

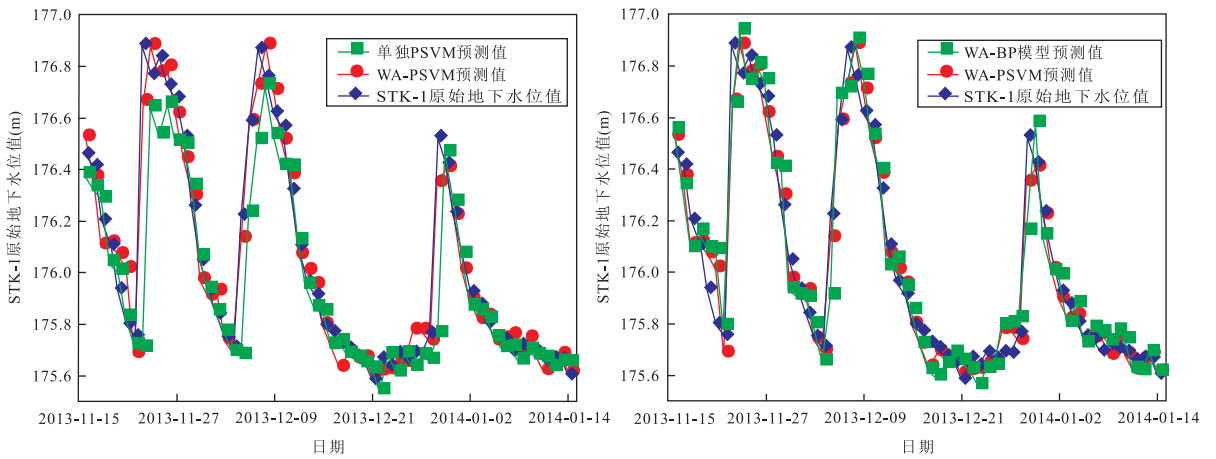


图 8 WA-PSVM 模型、WA-BP 模型和单独 PSVM 模型滑坡 STK-1 地下水水位一步预测结果对比

Fig.8 Final prediction results comparison of WA-PSVM model, WA-BP model and single PSVM model

表 3 STK-1 地下水水位不同模型预测结果对比

Table 3 Comparison of different models for one day ahead forecasting of STK-1 groundwater level series

模型	RMSE(m)	R ²
WA-PSVM	0.073	0.966
WA-BP	0.112	0.931
单独 PSVM	0.216	0.741

虽然 WA-PSVM 模型预测地下水水位效果较好,但预测值与实际监测值之间仍然存在一定的差距,这主要是由以下几个方面的因素引起的:

(1)库岸滑坡地下水水位序列是一个混沌序列,难以准确预测.

(2)采用 WA-PSVM 模型对地下水水位进行预测时,需要确定小波变换函数、分解层数、相空间重构参数、SVM 模型参数等诸多变量,这些建模细节都会影响模型预测精度.

(3)原始地下水水位序列由于仪器监测等原因含有一定的噪声,而混沌时间序列模型预测精度对噪声非常敏感,这也影响了模型预测精度.

(4)本文从监测获得的地下水水位时间序列本身出发进行建模预测,充分挖掘了地下水水位序列包含的丰富信息.但是诸多影响因素未予考虑,这也是导致模型预测精度有限的重要原因.

针对 WA-PSVM 模型预测地下水水位时存在的以上问题,下一步研究中可以在如下几个方面进行改进.首先需要深入研究滑坡地下水水位变化的机理,考虑将滑坡地下水水位的影响因素加入预测模型中,使得预测模型具有更明确的物理意义,模型信息也将更加丰富.再者需要对 WA-PSVM 模型的诸多参数的设置问题进行改进,尽量减少需要设置的模型

参数,使得预测模型更为稳健可靠.最后还可以尝试其他类型的预测模型,比如极端学习机模型、混沌时间序列局域预测法等.

5 结论

本文以三峡库区三舟溪滑坡为例,通过对滑坡前缘地下水水位影响因素分析,得出滑坡地下水水位呈现出非线性特征的外部原因为降雨和库水位波动中存在的趋势性、周期性和随机性特征.正是这种复杂的非线性特征,使得利用单独 PSVM 模型难以进行有效预测.

针对地下水位的非线性特征,本文基于混沌理论,从滑坡地下水位的单变量时间序列出发,利用小波变换法分离出地下水水位时间序列的趋势性、周期性和随机性等特征,克服了地下水水位序列的非线性非平稳性对 PSVM 模型预测性能的干扰.在未考虑外部影响因素的情况下,较好地预测出了滑坡地下水水位序列的细节特征和趋势特征,大幅提高了地下水水位的预测精度,为滑坡地下水水位预测提供了一种新的思路.

References

An, X., Jiang, D., Liu, C., et al., 2011. Wind Farm Power Prediction Based on Wavelet Decomposition and Chaotic Time Series. *Expert Systems with Applications*, 38 (9): 11280—11285.

Cui, W. Z., Zhu, C. C., Bao, W. X., et al., 2004. Prediction of the Chaotic Time Series Using Support Vector Machines. *Acta Physica Sinica*, 53(10): 3303—3310 (in Chinese)

- with English abstract).
- Gu, W. C., Chai, B. R., Teng, Y. P., 2014. Research on Support Vector Machine Based on Particle Swarm Optimization. *Transactions of Beijing Institute of Technology*, 34 (7): 705—709 (in Chinese with English abstract).
- Han, M., 2007. Prediction Theory and Method of Chaotic Time Series. China Water Power Press, Beijing, 28—30 (in Chinese).
- Huang, F. M., Tian, Y. G., 2014. WA-VOLTERRA Coupling Model Based on Chaos Theory for Monthly Precipitation Forecasting. *Earth Science—Journal of China University of Geosciences*, 39 (3): 368—374 (in Chinese with English abstract). doi:10.3799/dqkx.2014.035
- Kennedy, J., Eberhart, R., 1995. Particle Swarm Optimization. In: Proceedings of the 4th IEEE International Conference on Neural Networks. IEEE Service Center, Piscataway, 1942—1948.
- Li, C. D., Tang, H. M., Hu, B., et al., 2008. Research on Application of Wavelet Analysis and RBF Neural Network to Prediction of Foundation Settlement. *Rock and Soil Mechanics*, 29 (7): 1917—1922 (in Chinese with English abstract).
- Li, C. D., Tang, H. M., Hu, X. L., et al., 2009. Landslide Prediction Based on Wavelet Analysis and Cusp Catastrophe. *Journal of Earth Science*, 20 (6): 971—977.
- Liu, H., Tian, H. Q., Chen, C., et al., 2011. Short-Term Forecasting Optimization Algorithm for Unsteady Wind Speed Signal Based on Wavelet Analysis Method and Neural Networks Method. *Journal of Central South University (Science and Technology)*, 42 (9): 2704—2711 (in Chinese with English abstract).
- Liu, X. X., Xia, Y. Y., Zhang, X. S., et al., 2005. Effects of Drawdown of Reservoir Water Level on Landslide Stability. *Chinese Journal of Rock Mechanics and Engineering*, 24 (8): 1439—1444 (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—Journal of China University of Geosciences*, 33 (5): 687—692 (in Chinese with English abstract).
- Maiti, S., Tiwari, R. K., 2014. A Comparative Study of Artificial Neural Networks, Bayesian Neural Networks and Adaptive Neuro-Fuzzy Inference System in Groundwater Level Prediction. *Environmental Earth Sciences*, 71 (7): 3147—3160.
- Peng, L., Niu, R. Q., Wu, T., 2013. Time Series Analysis and Support Vector Machine for Landslide Displacement Prediction. *Journal of Zhejiang University (Engineering Science)*, 47 (9): 1672—1679 (in Chinese with English abstract).
- Peng, L., Niu, R. Q., Ye, R. Q., et al., 2012. Prediction of Ground Water Level in Landslides Based on Genetic-Support Vector Machine. *Journal of Central South University (Science and Technology)*, 43 (12): 4788—4795 (in Chinese with English abstract).
- Qiu, J. P., Xing, J., Jiang, A. N., et al., 2012. Supersonic Forecast of the Strength of Ore and Rock Based on Particle Swarm Support Vector Machine. *Journal of Northeastern University (Natural Science)*, 33 (5): 731—734 (in Chinese with English abstract).
- Ren, J., Huang, D. D., Li, Z. N., 2005. Character Recognition Based on Support Vector Machines and Wavelet. *Journal of Zhejiang University (Engineering Science)*, 39 (12): 2016—2020 (in Chinese with English abstract).
- Ren, L., Xiang, X. Y., Ni, J. J., 2011. Forecast Modeling of Monthly Runoff with Adaptive Neural Fuzzy Inference System and Wavelet Analysis. *Journal of Hydrologic Engineering*, 18 (9): 1133—1139.
- Siek, M., Solomatine, D. P., 2010. Nonlinear Chaotic Model for Predicting Storm Surges. *Nonlinear Processes in Geophysics*, 17 (5): 405—420.
- Tang, M., Ma, G. W., Xu, G., 2007. Analysis and Forecasting SMP Using Wavelet Transform. *Journal of Sichuan University (Engineering Science Edition)*, 39 (4): 12—15 (in Chinese with English abstract).
- Wang, Z. L., Sun, H. Y., Liu, Y. L., et al., 2011. Time Series Analysis about Groundwater Level in Slope and Rainfall. *Journal of Zhejiang University (Engineering Science)*, 45 (7): 1031—1037 (in Chinese with English abstract). doi:10.3785/j.issn.1008—973X.2011.07.027
- Wolf, A., Swift, J. B., Swinney, H. L., 1986. Determining Lyapunov Exponents from Time Series. *Physica D*, 16 (2): 285—371.
- Yang, Z. P., Lu, W. X., Li, P., 2005. Application of Time Series Model to Predict Groundwater Regime. *Journal of Hydraulic Engineering*, 32 (12): 1475—1479 (in Chinese with English abstract). doi:0559—9350(2005)12—1475—05
- Zhang, S. Q., Jia, J., Gao, M., 2010. Study on the Parameters Determination for Reconstructing Phase-Space in Chaos Time Series. *Acta Physica Sinica*, 59 (3): 1576—1582 (in Chinese with English abstract).
- Zhou, C. Y., Zhang, L., Huang, X. Y., 2005. Classification of Rocks Surrounding Tunnel Based on Improved BP Network Algorithm. *Earth Science—Journal of China*

University of Geosciences, 30(4): 480—486 (in Chinese with English abstract).

附中文参考文献

崔万照,朱长纯,保文星,等,2004.混沌时间序列的支持向量机预测.物理学报,53(10):3303—3310.

谷文成,柴宝仁,滕艳平,2014.基于粒子群优化算法的支持向量机研究.北京理工大学学报,34(7):705—709.

韩敏,2007.混沌时间序列预测理论与方法.北京:中国水利水电出版社,28—30.

黄发明,田玉刚,2014.混沌序列 WA-VOLTERRA 耦合模型在月降水量预测中的应用.地球科学——中国地质大学学报,39(3):368—374.

李长冬,唐辉明,胡斌,等,2008.小波分析和 RBF 神经网络在地基沉降预测中的应用研究.岩土力学,29(7):1917—1922.

刘辉,田红旗,陈超,等,2011.基于小波分析法与神经网络法的非平稳风速信号短期预测优化算法.中南大学学报(自然科学版),42(9):2704—2711.

刘新喜,夏元友,张显书,等,2005.库水位下降对滑坡稳定性的影响.岩石力学与工程学报,24(8):1439—1444.

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——中国地质大学学报,33(5):687—692.

彭令,牛瑞卿,吴婷,2013.时间序列分析与支持向量机的滑坡位移预测.浙江大学学报(工学版),47(9):1672—1679.

彭令,牛瑞卿,叶润清,等,2012.基于进化支持向量机的滑坡地下水位动态预测.中南大学学报(自然科学版),43(12):4788—4795.

邱景平,邢军,姜谔男,等,2012.基于粒子群支持向量机的矿岩强度指标的超声预测.东北大学学报(自然科学版),33(5):731—734.

任俊,黄丹丹,李志能,2005.基于支撑向量机和小波的字符识别.浙江大学学报(工学版),39(12):2016—2020.

唐明,马光文,徐刚,2007.基于小波变换的系统边际电价分析与预测.四川大学学报(工程科学版),39(4):12—15.

王智磊,孙红月,刘永莉,等,2011.降雨与边坡地下水位关系的时间序列分析.浙江大学学报(工学版),45(7):1031—1037.

杨忠平,卢文喜,李平,2005.时间序列模型在吉林西部地下水动态变化预测中的应用.水利学报,32(12):1475—1479.

张淑清,贾健,高敏,2010.混沌时间序列重构相空间参数选取研究.物理学报,59(3):1576—1582.

周翠英,张亮,黄显艺,2005.基于改进 BP 网络算法的隧洞围岩分类.地球科学——中国地质大学学报,30(4):480—486.

《地球科学——中国地质大学学报》

2015 年 8 月 第 40 卷 第 8 期 要目预告

高光谱遥感油气探测技术	汪大明等
基于 Hymap 模拟的星载高光谱数据质量评价	周萍等
高光谱遥感技术及资源勘查应用进展	李志忠等
基于植被胁迫光谱的高光谱遥感植被元素富集信息提取	帅琴等
基于光谱特征空间的苏必利尔湖型铁矿高光谱遥感蚀变信息提取方法	黄爽等
一种新的遥感蚀变信息编图方法	李建国等
基于特征导向主成分分析的遥感蚀变异常提取方法	谢明辉等
植被覆盖区羟基和碳酸盐矿物光谱吸收深度校正模型	周超等