

# 基于 ISVM—逐步回归组合的混凝土坝变形监控模型

王 锋, 孙 杨

(江西水利职业学院, 江西 南昌 330013)

**摘 要:**针对混凝土坝监控模型中监测数据较少、维数和非线性程度较高以及逐步回归方法预报值波动大,部分测值失真致拟合效果不佳等问题,鉴于支持向量机(SVM)能有效解决小样本、高维数以及非线性问题的能力,为此,基于改进 SVM 理论(ISVM)对大坝监控模型中的水压分量进行模拟,并采用逐步回归方法将模拟后的水压分量、温度分量与时效分量进行结合并模拟,建立了基于 ISVM—逐步回归组合的大坝监控模型,将拟合和预报结果与逐步回归—逐步回归、ISVM—ISVM 及逐步回归—ISVM 组合的预报结果进行对比.结果表明,基于 ISVM—逐步回归组合的混凝土坝安全监控模型拟合精度高,预测效果好,将更适于大坝监控模型的建立.

**关键词:**混凝土坝; ISVM; 逐步回归方法; 监控模型; 水压分量; 预测; 精度

**中图分类号:** TV698.1      **文献标识码:** A      **文章编号:** 1004-4701(2018)02-0085-06

## 0 引 言

大坝变形监控模型受多因素的综合影响,其表现为一个复杂的非线性函数关系,对大坝工作性态的模拟和预测,也一直是大坝安全监控研究的重点。目前,常用的安全监控模型有统计模型、确定性模型以及混合模型<sup>[1]</sup>;研究方法有逐步回归方法<sup>[2]</sup>和时间序列方法<sup>[3]</sup>等。近年来,随着一些新型学科不断发展,神经网络方法<sup>[4]</sup>、小波分析<sup>[5]</sup>、支持向量机<sup>[6]</sup>等方法已应用于混凝土坝的安全监测之中,这些方法为解决大坝工作性态拟合和预报起到了重要作用,但也存在诸多不足。逐步回归方法预报值波动大,部分测值失真致拟合效果不佳<sup>[7]</sup>;时间序列方法易受环境量相关性的影响<sup>[8]</sup>;神经网络方法易陷入局部最优,对初始权值和阈值较敏感<sup>[9]</sup>;小波分析去噪时易造成大方差,且所得信号可能出现伪吉布斯现象<sup>[10]</sup>等不足。

针对传统监控模型的不足,依据 SVM 在处理小样本、高维数、非线性等问题的计算优势,以及其良好的泛化能力<sup>[11]</sup>,采用改进 SVM 算法(ISVM)与逐步回归方法结合,有效的弥补了逐步回归方法预报值波动大,部分测值失真致拟合效果不佳的缺点,利用 ISVM 算法对影响大坝变形的水压分量进行拟合和预报,并将其预

报值与温度分量、时效分量相结合,采用逐步回归方法建立基于 ISVM—逐步回归组合方法的大坝变形监控模型,可有效提高模型预报精度。为探究该模型拟合及预测效果,将其结果与逐步回归—逐步回归、ISVM—ISVM 及逐步回归—ISVM 组合的预报结果进行对比分析。

## 1 混合预报模型的建立

### 1.1 大坝位移预报模型中各因子的选取

大坝中任一点的位移  $\delta$  可表示为水压分量  $\delta_H$ 、温度分量  $\delta_T$  和时效分量  $\delta_\theta$  三者的叠加,形如<sup>[12]</sup>:

$$\delta = \delta_H + \delta_T + \delta_\theta \quad (1)$$

(1) 水压分量  $\delta_H$ 。通常采用逐步回归方法模拟水压分量  $\delta_H$  与水位之间的函数关系,其多项式拟合表达式形如:

$$\delta_H = \sum_{i=1}^{m_1} a_i H^i \quad (2)$$

式中:  $a_i$  为回归系数;对于重力坝,常采用三次式,即  $m_1 = 3$ ;对于拱坝,常采用四次式,即  $m_1 = 4$ 。

(2) 温度分量  $\delta_T$ 。对于服役时间较长,但大坝工作正常且只有气温观测资料时,坝体内部任一点的温度变化可用周期函数来表示,考虑到位移温度分量与混凝土温度呈线性关系,故选取多周期的谐波作因子:

$$\delta_T = \sum_{i=1}^n \left( b_{1i} \sin \frac{2\pi it}{365} + b_{2i} \cos \frac{2\pi it}{365} \right) \quad (3)$$

式中： $i = 1$  时为年周期， $i = 2$  时为半年周期，……。 $n$  一般取 1, 2； $b_{1i}$  和  $b_{2i}$  为回归系数； $t$  为监测值至始测值累计天数。

(3) 时效分量  $\delta_\theta$ 。时效分量综合反映了坝体混凝土及基岩的徐变、塑性变形以及坝体混凝土裂缝导致的不可逆位移和自身体积变形。其一般采用下式表示：

$$\delta_\theta = c_1(t - t_0)/100 + c_2 \ln [(t - t_0)/100 + 1] \quad (4)$$

式中： $c_1$  和  $c_2$  为回归系数； $t_0$  为基准日期； $t$  为监测日期。

### 1.2 改进 SVM 建模原理

改进的 SVM 模型 (ISVM) 是在标准 SVM 模型上进行了改进，通过参数  $\nu$  实现对经验误差  $\varepsilon$  的自动计算<sup>[13]</sup>。设存在一组训练样本集：

$$T = \{ (x_i, y_i) \mid i = 1, \dots, n \}, x_i \in R^n, y_i \in R \quad (5)$$

式中： $x_i$  为输入数据； $y_i$  为目标输出数据。在回归问题中，就是采用一个非线性映射函数  $\phi(x)$  将样本从原空间映射到高维的特征空间，即从输入空间中  $R^n$  到输出空间  $R$  上的映射  $f: R^n \rightarrow R$ 。设回归函数如下：

$$f(x) = (\omega \cdot \phi(x)) + b \quad (6)$$

式中： $\omega$  为全矢量； $b$  为偏置量； $x \in R^n$ ； $b \in R$ 。

将回归问题视为风险函数最小化问题，回归函数可用风险泛函表示如下：

$$A = \frac{1}{2} \omega^T \omega + \frac{1}{n} C \sum_{i=1}^n L_\varepsilon(y_i, f(x_i)) \quad (7)$$

其中：

$$L_\varepsilon = (y_i, f(x_i)) = \max \{0, |y_i - f(x_i) - \varepsilon|\} \quad (8)$$

式中：式(7)右侧第一项为规则化项，第二项为经验风险泛化函数； $C$  惩罚参数，通常为常数； $L_\varepsilon(y_i, f(x_i))$  为损失函数，常采用不敏感损失函数  $\varepsilon$ ，其取值为估计输出  $f(x_i)$  与期望输出  $y_i$  差的绝对值与  $\varepsilon$  中较大者。

在规则化的风险泛函中引入  $\xi_i, \xi_i^*$  两个非负松弛变量，如下式所示：

$$A' = \min_{\omega, b, \xi, \xi^*, \varepsilon} \left[ \frac{1}{2} \omega^T \omega + C[\nu\varepsilon + \frac{1}{n} \sum_{i=1}^n L_\varepsilon(\xi + \xi^*)] \right] \quad (9)$$

采用如下约束：

$$\begin{cases} y_i - (\omega, \phi(x_i)) - b \leq \varepsilon + \xi_i \\ (\omega, \phi(x_i)) + b - y_i \leq \varepsilon + \xi_i^* \end{cases} \quad (10)$$

该二次优化问题的对偶问题如下：

$$B = \min_{\alpha, \alpha^*} \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)(\phi(x_i), \phi(x_j)) + \sum_{i=1}^n y_i(\alpha_i - \alpha_i^*) \quad (11)$$

式中： $\alpha_i, \alpha_i^*, \alpha_j, \alpha_j^*$  均为方程待定系数。其约束条件如下：

$$\begin{cases} \sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0 \\ \sum_{i=1}^n (\alpha_i + \alpha_i^*) \leq C\nu \quad i = 1, 2, \dots, n \\ 0 \leq \alpha_i \\ \alpha_i^* \leq C/n \end{cases} \quad (12)$$

求解二次优化问题，得  $\alpha$  值及  $\omega$  的表达式为：

$$\omega = \sum_{i=1}^n (\alpha_i, \alpha_i^*) \phi(x_i) \quad (13)$$

最终可得回归方程近似解为：

$$f(x) = \sum_{i=1}^n (\alpha_i^* - \alpha_i)(\phi(x_i) \cdot \phi(x_i^*)) + b \quad (14)$$

非线性回归问题求解过程中，由于  $\phi(x_i)$  未知，且  $\omega$  无法显式表达，常采用定义核函数  $K(x_i, x_j)$ <sup>[14]</sup> 来解决非线性映射函数  $\phi(x)$  具体形式难以确定的问题，核函数形式如下：

$$K(x_i, x_j) = (\phi(x_i)^T \cdot \phi(x_j)) \quad (15)$$

核函数是通过特征变换的定义使训练样本在特征空间实现某种特征变化。其注重的是变换后的结果，并非变换的具体形式。常用的核函数主要有径向基核函数、多项式核函数及感知器核函数等<sup>[11]</sup>，不同的核函数对 ISVM 的影响不大，本文采用较常用的径向基核函数，其具体形式如下：

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

式中： $\sigma$  为核函数参数，其代表方差。

建立 ISVM 模型的关键对惩罚参数  $C$  和核函数参数  $\sigma$  的选取，本文将采用粒子群算法对参数  $C$  和  $\sigma$  进行寻优，其中一个粒子即代表一个 ISVM 模型。

### 1.3 逐步回归方法

在混凝土坝安全监测中，采用逐步回归分析方法进行大坝变形预测，通过对具有相关关系的荷载集（自变量）及荷载效应（因变量）进行数理统计分析，找出各预报因子对大坝变形的影响，并建立相应的数学表达式，由此推求某一荷载集作用下大坝位移预报值，并与原型观测数据进行比对与残差分析，以此完成大坝的的监测预报。

逐步回归的基本思想是:在建模过程中,将各自变量根据自身对因变量的影响程度大小,以从大到小的顺序依次引入回归方程,为使模型所建立的回归方程具有良好的精度,需随时对回归方程当时所含的全部变量进行显著性检验,若不显著则将其剔除,直到方程中所含的所有自变量对因变量的作用显著时,方可引入新的变量。再在剩下的未选自变量中,选出对因变量影响最大者,检验其是显著性,若显著,则引入方程,否则不引入。直至最后既无显著自变量可以引入,也无不显著自变量需要剔除。

#### 1.4 ISVM—逐步回归组合预报模型的建立

ISVM 方法是基于统计学理论的一种新的机器学习算法,该方法可实现由低维空间到高维空间的非线性映射,能较好的解决小样本、非线性、高维数和局部极小值等实际问题,借助 ISVM 算法可有效解决水压分量拟合过程中的非线性问题<sup>[15]</sup>,本文采用 ISVM 算法进行水压分量的拟合和预报,将预报值与温度分量、时效分量进行结合,并采用逐步回归方法建立大坝变形监控模型,基于 ISVM—逐步回归组合建立的大坝变形监控模型能够有效的提高模型的预报精度,具体步骤如下:

(1) 通过 ANSYS 软件建立大坝有限元模型,计算各特征水位下大坝某测点的位移值以及各特征水位  $H$  对应的  $H^2$  和  $H^3$  值;

(2) 将各特征水位  $H$  值与  $H^2$ 、 $H^3$  值进行归一化以及降维处理;

(3) 利用 MATLAB 软件,采用粒子群算法对惩罚参数  $C$  和核函数参数  $\sigma$  进行参数寻优<sup>[16]</sup>;

(4) 将参数寻优结果结合 ISVM 算法,建立该大坝测点相应的水压分量预报模型,并对该测点水压分量进行拟合和预报;

(5) 将该测点的水压分量预报值与温度分量、时效分量进行结合,采用逐步回归方法建立该测点的大坝位移监控模型;

(6) 通过对上个步骤建立的大坝监控模型进行拟合和预报,将预报值与实测值进行对比并展开预报模型预测精度分析。

综上所述,基于 ISVM—逐步回归组合方法的大坝变形监控模型的建立流程如图 1 所示。

## 2 工程算例

### 2.1 某混凝土坝有限元模型

某混凝土重力坝坝顶高程 179.00 m,最大坝高

113.00 m,坝顶总长 308.50 m,坝顶宽 7.00 m,最大坝底宽 84.50 m。水库正常蓄水位 173.00 m,调节库容 11.22 亿  $m^3$ ,校核洪水位 177.80 m,相应总库容 20.35 亿  $m^3$ 。2#坝段有限元模型以六面体八节点单元为主,其中坝体部分共有单元 4 463 个,节点 5 393 个;基岩部分共有单元 12 440 个,节点 13 933 个,假设有限元模型中弹性模量取 24.0 GPa。该大坝共分 6 个坝段,选取 2#典型坝段正垂线测点 PL5 (179.00 m 高程) 建立模型,2#坝段有限元模型如图 2 所示,大坝各特征水位及相应特征水位下的 ANSYS 计算值如表 1 所示。

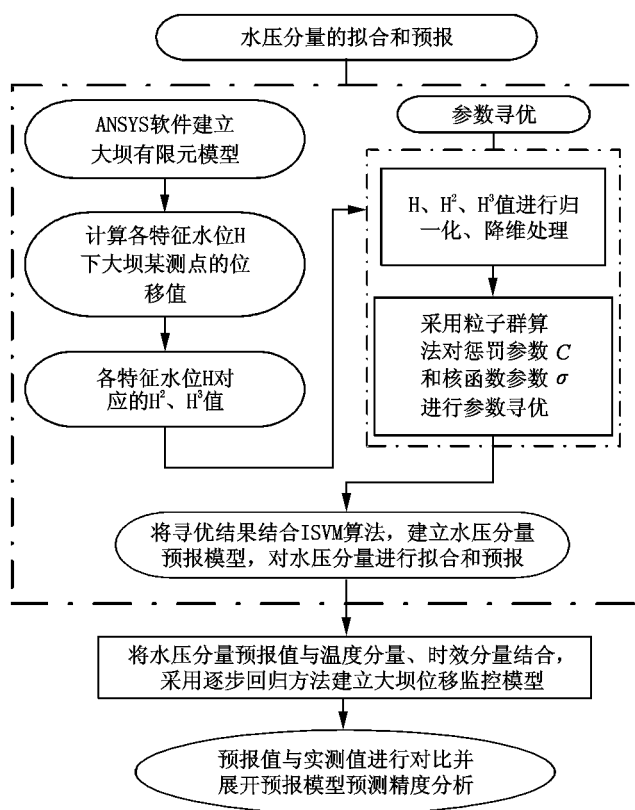


图1 基于 ISVM—逐步回归组合的大坝变形监控模型的建立流程图

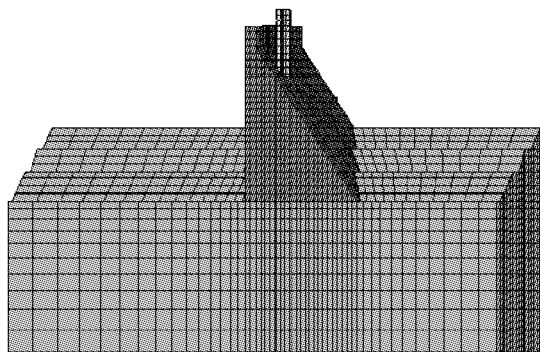


图2 2#坝段有限元模型

表1 大坝各特征水位 H 相应 ANSYS 计算值

H/m	75.60	76.10	76.60	79.10	80.10	81.60	83.60	84.10	86.60	87.60	88.10
ANSYS 计算值	0.94	1.10	1.30	2.30	2.90	3.60	4.90	5.20	6.70	7.60	8.01

2.2 计算结果对比与分析

选取 2<sup>#</sup>典型坝段正垂线测点 PL5(179.00 m 高程) 建立模型,将该测点 2006 年 1 月 1 日~2007 年 11 月 1 日间的 590 组监测数据作为训练样本,2007 年 11 月 2 日~2007 年 11 月 11 日间的 10 组监测数据作为测试

样本。通过计算得 PL5 测点在改 ISVM—逐步回归组合模型下的拟合和预报结果,并与基于逐步回归—逐步回归、ISVM—ISVM 及逐步回归—ISVM 组合方法的模型预报结果进行对比分析,结果如图 3 所示。

由图 3 可看出,采用 4 种组合模型的拟合值与实测

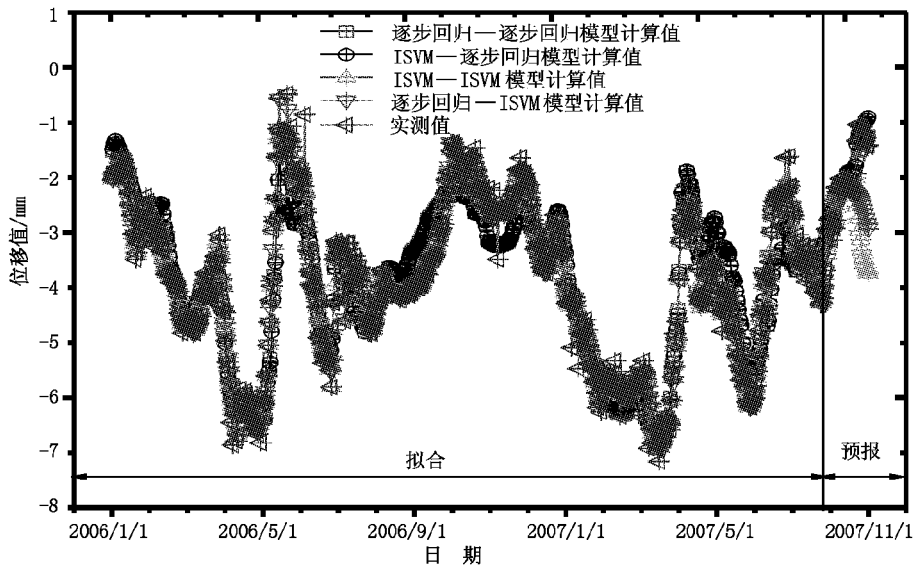


图3 PL5 测点实测值与各组合模型计算值

表2 PL5 测点测试样本下各组合模型预测结果比较

年/月/日	实测位移	逐步回归—逐步回归			ISVM—逐步回归			ISVM—ISVM			逐步回归—ISVM		
		预测值	绝对误差	相对误差	预测值	绝对误差	相对误差	预测值	绝对误差	相对误差	预测值	绝对误差	相对误差
2007/11/2	-1.33	-1.72	0.39	-0.29	-1.40	0.07	-0.05	-2.53	1.20	-0.90	-2.30	0.97	-0.73
2007/11/3	-1.04	-1.61	0.57	-0.54	-1.28	0.24	-0.23	-2.75	1.71	-1.64	-2.33	1.29	-1.24
2007/11/4	-1.02	-1.53	0.51	-0.50	-1.20	0.18	-0.17	-2.97	1.95	-1.91	-2.37	1.35	-1.33
2007/11/5	-1.08	-1.50	0.42	-0.39	-1.17	0.09	-0.08	-3.13	2.05	-1.90	-2.43	1.35	-1.25
2007/11/6	-1.08	-1.42	0.34	-0.32	-1.09	0.01	-0.01	-3.32	2.24	-2.08	-2.49	1.41	-1.31
2007/11/7	-1.17	-1.38	0.21	-0.18	-1.04	0.13	-0.11	-3.46	2.29	-1.95	-2.57	1.40	-1.20
2007/11/8	-1.34	-1.34	0.00	0.00	-1.01	0.33	-0.25	-3.56	2.22	-1.66	-2.65	1.31	-0.98
2007/11/9	-1.28	-1.33	0.05	-0.04	-0.99	0.29	-0.22	-3.63	2.35	-1.84	-2.74	1.46	-1.14
2007/11/10	-1.44	-1.29	0.15	-0.11	-0.95	0.49	-0.34	-3.69	2.25	-1.57	-2.84	1.40	-0.97
2007/11/11	-1.41	-1.24	0.17	-0.12	-0.90	0.51	-0.36	-3.74	2.33	-1.65	-2.93	1.52	-1.08

值的变化趋势相近,表明该 4 种模型均可用于预报分析,其计算结果存在可比性,模型的精度主要通过预报

阶段的预报值体现,为比较 4 种组合模型的计算精度,将 PL5 测点在四种组合模型下,于 2007 年 11 月 2 日~

2007年11月11日间的10组预报值予以列出,并与实测值进行比对,计算出绝对误差和相对误差,结果见表2。

由表2可知,在预报阶段,该测点在逐步回归—逐步回归组合模型和ISVM—逐步回归组合模型下的误差指标总体上要低于ISVM—ISVM组合模型和逐步回归—ISVM组合模型,而ISVM—逐步回归组合模型下的误差指标总体上又要低于逐步回归—逐步回归组合模型,由此可得,此模型预报精度较其他模型有所提高。在采用粒子群算法进行寻优过程中,各参数分别为 $c_1 = 1.5, c_2 = 1.7$ ,进化代数为60,种群规模为30,寻优速度为3,惩罚系数 $C$ 范围为 $[0.1, 400]$ ,核函数参数 $\sigma$ 的范围为 $[0.001, 0.1]$ ,最终寻优结果为 $C = 400, \sigma = 0.029141$ 。

为将各组合模型的预报精度表现更为清楚,对该测点在四种模型下的拟合值及预报值的均方误差(MSE)、平均绝对百分比误差(MAPE)等误差指标进行统计,其中MSE可以反映预测值相对实测值的整体变化程度,MAPE可用于衡量预测值与实测值误差的整体水平,计算公式如下:

$$MSE = \frac{\sum_{i=1}^n (\delta_i - \hat{\delta}_i)^2}{n} \quad (17)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\delta_i - \hat{\delta}_i}{\delta_i} \right| \quad (18)$$

式中: $n$ 为预测样本总数; $\delta_i$ 为实测值; $\hat{\delta}_i$ 为模型预测值。该测点各种模型误差指标计算结果如表3。

表3 PL5测点各组合模型误差指标对比

误差指标	逐步回归— 逐步回归	ISVM— 逐步回归	ISVM— ISVM	逐步回归— ISVM
MSE	0.11	0.08	4.36	1.83
MAPE	0.25	0.18	1.71	1.12

由表3可知,在预报阶段,逐步回归—逐步回归模型的精度已经很高,且均高于ISVM—ISVM及逐步回归—ISVM组合建立的模型,而采用ISVM—逐步回归组合建立的模型得到的预报结果可从两项误差指标得以反映,其两项误差指标分别为 $MSE = 0.08, MAPE = 0.18$ ,均好于逐步回归—逐步回归模型的两项误差指标 $MSE = 0.11, MAPE = 0.25$ ,由此可见,采用ISVM—逐步回归组合可以有效的提高模型预报的精度,这是由于模型中的水压分量是采用ISVM算法进行模拟,该算

法是基于结构风险最小原理,能有效的解决小样本、高维数、非线性的问题,能从一定程度上顾及水位与坝体位移间的部分非线性关系,使水压分量的精度得以改善,进而提高监控模型的预报精度,故对于大坝位移监控,采用ISVM—逐步回归方法进行建模将更为合理。

### 3 结 论

本文通过采用SVM算法对水压分量进行模拟,并结合逐步回归方法建立混凝土坝变形监控模型,有效地将结构计算成果与数据训练方法相结合,体现了大坝在水压力作用下的力学特征,通过采用组合方法构建大坝变形监控模型,可以得出以下结论:

(1)基于ISVM—逐步回归组合建立的混凝土坝安全监控模型与基于逐步回归方法建立的监控模型相比,拟合精度更高,预测效果更好,将更加适合于大坝监控模型的建立。

(2)在采用ISVM算法进行建模时,研究发现寻优过程中惩罚系数 $C$ 的范围对模型预报精度的影响较大,其次是核函数参数 $\sigma$ 的影响。

(3)在结合水压分量、温度分量及时效分量对大坝建立ISVM模型的过程中,采用粒子群算法对惩罚系数和核函数参数进行寻优时,受高维数的影响,寻优过程较为缓慢,如何提高寻优效率将是以后研究的方向。

### 参考文献:

- [1] 肖浩波,谷艳昌.混凝土坝安全监控最小二乘支持向量机模型[J].长江科学院院报,2013,30(5):34~37.
- [2] 许昌,岳东杰,董育烦,等.基于主成分和半参数的大坝变形监测回归模型[J].岩土力学,2011,32(12):3738~3742.
- [3] 文富勇.基于混沌时间序列的大坝变形监测自适应预测模型[J].水力发电,2013,39(11):72~76.
- [4] 张帆,胡伍生.神经网络融合模型在大坝安全监控中的应用[J].测绘工程,2015,24(1):53~56.
- [5] 刘博智,唐娟.小波分析—回归模型在混凝土坝变形分析中的应用[J].水电能源科学,2012,30(7):80~82.
- [6] 丁世飞,齐丙娟,谭红艳.支持向量机理论与算法研究综述[J].电子科技大学学报,2011,40(1):2~10.
- [7] 邱莉婷,沈振中,聂柏松.基于逐步回归分析—马尔可夫链模型的大坝变形预测[J].水电能源科学,2014,32(5):51~55.
- [8] 苏怀智,温志萍,吴中如.基于SVM理论的大坝安全预警模型研究[J].应用基础与工程科学学报,2009,17(1):40~48.
- [9] 王雪红,刘晓青,陶海龙,等.优化BP神经网络的位移预测模型[J].水利水运工程学报,2014,2(2):38~42.
- [10] 袁开明,舒乃秋,孙云莲,等.基于阈值寻优法的小波去噪分析[J].武汉大学学报(工科版),2015,48(1):74~80.

- [11] SU Huai-zhi, HU Jiang, WU Zhong-ru. A Study of Safety Evaluation and Early-warning Method for Dam Global Behavior [J]. Structural Health Monitoring, 2012, 11(3): 269 ~ 279.
- [12] 顾冲时, 汪亚超, 彭妍, 等. 大坝安全监控模型的病态问题及其处理方法[J]. 中国科学技术, 2011, 41(12): 1574 ~ 1579.
- [13] 丛康林, 岳建平, 李希灿.  $\nu$ -SVR 改进的滚动预测模型在大坝沉降监测中的应用[J]. 大地测量与地球动力学, 2014, 34(4): 92 ~ 95.
- [14] 张倩, 杨耀权. 基于支持向量机核函数的研究[J]. 电力科学与工程, 2010, 28(5): 42 ~ 45.
- [15] 杜传阳, 郑东健, 张毅, 等. 基于动态 SVM-MC 的大坝变形监测模型及应用[J]. 水电能源科学, 2015, 33(1): 71 ~ 74.
- [16] P. J. García Nieto, E. García-Gonzalo, J. R. Alonso Fernández, et al. Hybrid PSO-MARS-based model for forecasting a successful growth cycle of the *Spirulina platensis* from experimental data in open raceway ponds[J]. Ecological Engineering, 2015, 81(1): 534 ~ 542.

编辑: 张绍付

## Concrete dam deformation monitoring mixed model built on the combination method of ISVM-stepwise regression method

WANG Feng, SUN Yang

(Jiangxi Water Resources Institute, Nanchang 330013, China)

**Abstract:** In view of the concrete dam monitoring data in monitoring model is less, high dimension and nonlinear degree as well as the method of stepwise regression forecast value volatility, part of the measuring value loss can cause such problems as poor fitting effect, support vector machine (SVM) can effectively solve the problem of small sample, high dimension and nonlinear, so, support vector machine was used in the simulation of the concrete dam monitoring model of hydraulic component, and stepwise regression method is used to simulate the water pressure component with a combination and simulation of the temperature component and aging component to build the dam monitoring model based on the combination method of SVM-stepwise regression, then give a comparison and analysis between stepwise regression-stepwise regression model, SVM-SVM model and stepwise regression-SVM model, results show that the SVM-concrete dam safety monitoring model of combination of stepwise regression fitting precision is high, the forecast effect is good, this model will be more suitable for dam monitoring model.

**Key words:** Concrete dam; Support vector machine; Stepwise regression method; Monitoring model

翻译: 王 锋