

曾永军, 张俊文, 曹登刚, 等. RS-RF 模型在混凝土坝变形预测中的应用[J]. 水利水电技术(中英文), 2021, 52(5): 82-88.  
ZENG Yongjun, ZHANG Junwen, CAO Denggang, et al. Application of RS-RF model in deformation prediction of concrete dam [J]. Water Resources and Hydropower Engineering, 2021, 52(5): 82-88.

# RS-RF 模型在混凝土坝变形预测中的应用

曾永军<sup>1</sup>, 张俊文<sup>2</sup>, 曹登刚<sup>3</sup>, 武红科<sup>4</sup>

(1. 贵州省水利投资(集团)有限责任公司, 贵州 贵阳 550081; 2. 长江水利委员会陆水试验枢纽管理局, 湖北 咸宁 437300; 3. 黔南州水利水电勘测设计研究院, 贵州 黔南 558000; 4. 山东大学 土建与水利学院, 山东 济南 250061)

**摘要:** 高性能的混凝土坝变形预测模型作为结构安全性态诊断、预警和科学决策制定的重要参考, 同时也是工程效益得以充分发挥的保障措​​施之一。针对混凝土坝变形监控模型中因子选取的主观性、因子间的多重共线性和预测模型泛化性差等问题, 结合粗糙集和随机森林理论在特征属性约简、重要性评价和高精度预测等方面的优势, 建立了基于 RS-RF 的混凝土坝变形预测模型。通过工程实例应用表明: 基于 RS-RF 的混凝土坝变形监控模型可以对影响因子集进行约简并给出各因子的重要性, 且预测精度优于常用的 SVM 模型和 RF 模型。由此可知, 基于 RS-RF 的混凝土坝变形预测模型实现了影响因子优选, 弥补了智能预测模型在定量化分析、预测泛化性等方面的不足, 具有较强的工程实用性。

**关键词:** 粗糙集; 随机森林; 预测模型; 变形监测

doi: 10.13928/j.cnki.wrahe.2021.05.009

开放科学(资源服务)标志码(OSID):

中图分类号: TV698.1

文献标志码: A

文章编号: 1000-0860(2021)05-0082-07



听语音  
聊科研  
与作者互动

## Application of RS-RF model in deformation prediction of concrete dam

ZENG Yongjun<sup>1</sup>, ZHANG Junwen<sup>2</sup>, CAO Denggang<sup>3</sup>, WU Hongke<sup>4</sup>

(1. Guizhou Water Conservancy Investment Group, Co., Ltd., Guiyang 550081, Guizhou, China; 2. Bureau of Lushui Experiment Hydropower Complex Management, Changjiang Water Resources Commission, Xianning 437300, Hubei, China; 3. Qiannan Water Conservancy and Hydropower Research Institute, Qiannan 558000, Guizhou, China; 4. School of Civil Engineering, Shandong University, Jinan 250061, Shandong, China)

**Abstract:** As an important reference for structural safety diagnosis, early warning and scientific decision-making, the deformation prediction model of high-performance concrete dam is also one of the safeguard measures to give full play to the engineering benefits. In this paper, considering the subjectivity of factor selection, multicollinearity among factors and poor generalization of prediction model in concrete dam deformation monitoring model, combined with the advantages of rough set and random forest theory in feature attribute reduction, importance evaluation and high-precision prediction, a concrete dam deformation prediction model based on RS-RF is established. The engineering application shows that the deformation monitoring model of concrete dam based on RS-RF can reduce the set of influencing factors and give the importance of each factor, and SVM model and RF model are commonly used for prediction accuracy. Therefore, the deformation prediction model of concrete dam based on RS-RF has

收稿日期: 2020-01-08

基金项目: 山东省自然科学基金项目(ZR2017MEE501)

作者简介: 曾永军(1977—), 男, 高级工程师, 学士, 主要从事水利工程建设及行营管理。E-mail: zhaolunpi807528@163.com

realized the optimization of influencing factors, made up for the deficiency of intelligent prediction model in quantitative analysis and generalization of prediction, and has strong engineering practicability.

**Keywords:** rough set; random forest; prediction model; deformation monitoring

## 0 引言

混凝土坝是目前坝工建设中优先采用的坝型, 其具有适应能力强, 安全系数高、运行维护方便等优点, 但混凝土坝工作性态影响因素众多, 且服役环境复杂多变, 其服役安全性态是一个受多因素协同作用下材料与结构交互影响的非线性动态演化过程<sup>[1]</sup>。变形作为直观反映混凝土坝安全性态的综合效应量, 可以作为结构性态发生趋势性变化甚至发生转异的重要指标, 因此在此基础上加强混凝土坝变形安全监控模型的研究, 是确保其长效服役安全性的重要手段之一<sup>[2]</sup>。

混凝土坝设计、施工和运行管理等多方面因素综合作用其工作性态, 反映变形效应量和影响因子间关系的预测模型是一个复杂的非线性函数, 采用完整、可靠的监测数据进行定长序列预测时, 影响因子、建模方法的选取对于预测结果具有较大影响。影响因子选取方面: 变形预测模型影响因子集以简化的大坝与坝基物理模型、监测设备建设情况、原型监测资料等为基础, 采用统计模型中的影响因子集。模型选取的水压分量因子、温度分量因子和时效分量因子之间存在多重共线性, 这种特征在建模时有可能降低模型的精度, 对预测结果造成不利影响。同时, 在实际应用过程中, 预测模型未考虑渗流量、裂缝开度、扬压力等可量化影响因子和筑坝材料、施工质量、地质条件等难以量化的影响因子对变形的影响, 同时也未对影响因子重要度加以区分<sup>[3]</sup>。目前, 混凝土坝变形预测模型影响因子最常用的优选方法包括传统线性回归方法、降维分析方法、模糊数学分析法等<sup>[4-8]</sup>。但在实际应用过程中, 先验知识法过于依赖经验, 误差较大; 降维分析法中的主成分分析法要求变量具有较高的线性相关性, 在此基础上分析处理才能得到精度较高的分析结果, 但实际中影响混凝土坝变形的因子之间作用关系复杂, 具有较强的非线性; 模糊数学分析法中的灰色关联度分析法没有给出因子选择的评判标准, 仅仅只能给出关联度排序。因此, 混凝土坝变形预测模型影响因子选取过程中存在选入准则不完善, 影响因子重要性未加区分等问题, 严重影响模型预测性能。预测模型建立方面: 混凝土坝变形预测模型目前主要包括统计模型、确定性模型、混合模型、组合模型及智能

算法模型。随着坝工理论、有限元理论、人工智能技术等快速发展, 各学者所建立的监控模型在科研生产中的了丰硕的实践成果。但由于水利工程的复杂性、结构的波动性和工况的不确定性, 目前所有预测模型研究和应用均没有趋近完善, 存在稳健性差、计算量大、预测精度不高和过度拟合情况等不足之处<sup>[9-11]</sup>, 影响了模型的可靠性和实用性。

近年来, 人工智能领域成熟的粗糙集理论和随机森林算法在进行特征属性约简、消除冗余信息和学习复杂非线性关系、提高模型泛化性等方面取得了丰硕的研究成果。因此, 本文针对混凝土坝变形监控模型中存在的缺点和粗糙集、随机森林的特有优势, 建立了基于 RS-RF 的混凝土坝变形监控模型, 实现了影响因子集的优选和重要性度量, 提高了模型的拟合效果及预测精度, 更加接近真实地反映大坝的工作性态。

## 1 粗糙集理论

粗糙集理论<sup>[12-15]</sup>以“在保持信息等价下实现知识的挖掘、提炼”为核心思想, 最终实现对不确定性信息系统的属性约简, 重要性评价等目标。

### 1.1 信息系统

在粗糙集理论中, 为了对论域内已具有必要的信息和知识的对象(或称元素、样本)进行描述, 建立了信息系统  $S$

$$S = \{U, R, V, f\} \quad (1)$$

式中,  $U = \{s_1, s_2, \dots, s_n\}$ , 为所有对象构成的非空有限集合;  $R = \{\alpha \mid \alpha \in C \cup D\}$  为属性集, 其中  $C$  为条件属性集和,  $D$  为决策属性集;  $V$  为属性值域构成的集合;  $f$  为样本条件属性与值域之间映射关系的信息函数, 对于  $\forall x \in U, \alpha \in A, f(x, \alpha) \in V_\alpha$ 。

### 1.2 属性约简

对于任意的一个集合  $P$ , 若  $P$  满足:  $P \subseteq R$  且  $B \neq \phi$ , 则定义  $B$  与  $U$  的不可分辨关系为

$$IND(P) = \{(x, y) \in U^2 \mid \forall \alpha \in P, \alpha(x) = \alpha(y)\} \quad (2)$$

任意一个对象集合  $X \subseteq U$  以及属性集合  $B \subseteq C$ , 则

$X$  的  $B$  下近似定义为:  $\underline{B}X = \{x \mid [x]_B \subseteq X\}$ ;  $X$  的  $B$  上近似定义为:  $\overline{B}X = \{x \mid [x]_B \cap X \neq \phi\}$ ;  $X$  的

边界区定义为:  $BN_B(X) = \overline{BX} - \underline{BX}$ , 其中  $[X]B$  为  $B$  对于  $U$  的不可分辨关系集合, 若  $BN_B(X)$  非空, 则称  $X$  是  $B$  的粗糙集。

定义  $B$  相对于  $D$  的正区域为

$$POS_B(D) = \{ \underline{BX} \mid X \in U/IND(D) \} \quad (3)$$

设  $a \in C$ , 若  $SIM = POS_C(D) - POS_{C-\{a\}}(D) = 0$ , 则  $C$  中的  $a$  可以省略; 当  $C$  中每个元素都不可省略时, 称  $C$  为  $D$  独立。当  $C' = C - C^*$  为  $D$  独立, 且  $D$  中与  $C^*$  中完全相同的所有元素均可省略, 则称为  $C'$  为  $C$  的相对约简。

### 1.3 重要性评价

在属性约简中, 利用属性集合之间的相互依赖程度, 属性集  $B$  对  $D$  的依赖程度定义如下

$$\gamma_B(D) = \frac{|POS_B(D)|}{|U|} \quad (4)$$

式中,  $|\cdot|$  为与其对应的集合的基数。

综上所述, 考虑属性之间的依赖关系, 决策属性  $D$  中的属性  $a$  的重要度如下

$$Sig(\alpha, B, D) = \gamma_B(D) - \gamma_{B-\{\alpha\}}(D) \quad (5)$$

## 2 随机森林原理

RF 使用了 CART 决策树<sup>[16]</sup> 作为弱学习器, 并在此基础上进行集成学习, 即采用 Bootstrap<sup>[17]</sup> 方法对原始样本进行重采样, 基于重采样样本进行训练, 最终以各决策树分析结果为基础, 采用投票的形式确定最终结果。

随机森林模型在保证模型预测精度的前提下, 可以有效地提高模型运算效率, 且其对于变量间存在的多元共线性具有较强的鲁棒性, 可以实现多变量综合作用下的预测<sup>[18]</sup>。

### 2.1 样本集选择

模型原始样本集总共包括  $N$  个样本,  $M$  个特征, 这些数据蕴含了自变量和因变量之间的复杂的作用关系, 可以是线性关系, 也可以是非线性关系。首先通过 Bootstrap 方式对原始数据集进行重采样, 从原始样本集中采用随机有放回的方式抽取  $N$  个样本集作为模型的训练集, 总共抽取  $n$  次。每个样本中均可能含有重复的样本, 同时又不包含全部样本, 避免了将全部样本输入模型而造成过拟合现象。抽样过程中, 我们把没有被采样到的数据称为袋外数据 (Out of bag, OOB), 考虑到其在模型训练中未出现, 因此可以其评估自变量对因变量的影响程度和检验模型性能。

随机森林模型抽样次数  $n$  (既决策树个数) 和特征

分裂节点数  $m$  决定着模型的预测能力。决策树个数可以通过其与 OOB 误差的关系实验来确定; 特征分裂节点数  $m$  按照推荐值选取为  $M/3$ <sup>[19]</sup>。

### 2.2 模型训练和预测

分别采取独立同分布的训练样本对每颗决策树进行训练, 基于所有决策树预测结果投票决定 RF 最终的预测结果。RF 无需专门设置交叉验证, 通常采用带外数据样本输入到训练好的最优参数模型中进行模型测试<sup>[20]</sup>。

## 3 基于 RS-RF 的混凝土坝变形预测模型

传统的混凝土坝变形影响因子集具有冗余性、共线性等特征, 其不仅会导致模型预测精度的降低, 同时还会导致模型病态或者错误。因此, 针对 RF 算法对样本属性不加区分、模型定量化解释模糊等方面的不足, 借助 RS 理论在约简属性关系, 重要性评价方面的优势, 建立基于 RS-RF 的混凝土坝变形预测模型, 并对各影响因子进行重要性评价。具体建模步骤如下。

### 3.1 数据预处理

采用统计学方法对混凝土坝监测数据进行粗差处理, 为预测模型的建立提供可靠的数据基础。采用粗糙集理论对水位、气温、渗流、裂缝、扬压力等变量进行属性约简, 准确获取影响混凝土坝变形的代表性因子。将代表性影响因子及其衍生变量样本、若干测点对应的位移监测样本作为模型数据集, 采用独立标准化公式对环境影响因子样本进行处理。

### 3.2 模型训练

将经预处理的标准化训练集样本作为模型输入, 通过梯度下降算法进行误差反向传播驱使模型损失函数收敛, 获得最优参数模型。

### 3.3 模型预测

将测试集自变量因子数据输入到训练好的最优参数预测模型, 获得相应的变形预测结果。

### 3.4 模型性能评价

为了准确衡量模型预测性能, 结合统计多元回归理论, 考虑到评价指标的重叠性, 采用均方差 RMSE 和平均百分比误差 MAPE 作为模型预测效果评价指标

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (6)$$

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (7)$$

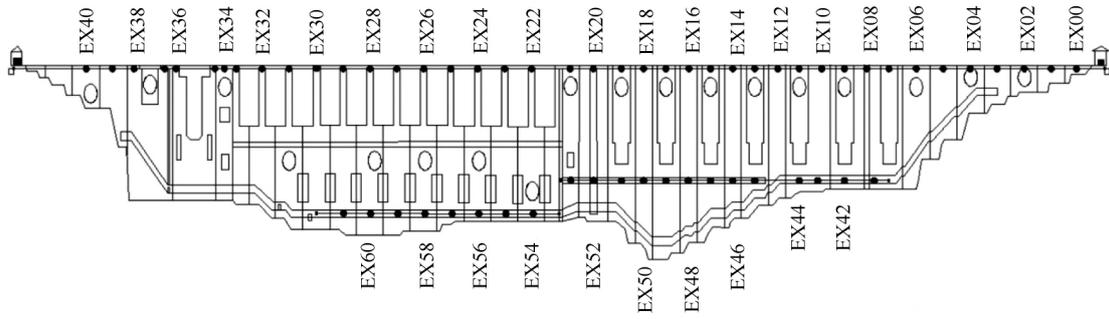


图 1 混凝土坝坝顶引张线测点平面布置

Fig. 1 Plane layout of measuring points of concrete dam crest extension

## 4 工程实例应用

### 4.1 工程简介

某混凝土坝位于福建省岷江干流的中游, 工程主要由大坝、电厂、船闸、升船机等建筑物组成。坝型为混凝土重力坝, 坝轴线总长 783 m, 最大坝高为 101 m。工程的变形监测主要包括水平位移、垂直位移、接缝开度等监测项目。图 1 为混凝土坝坝顶引张线测点平面布置图。

### 4.2 预测模型影响因子优选

根据理论知识、监测数据、专家经验等资料, 初始选取的经验影响因子集为

$$\{(H - H_0)^1, (H - H_0)^2, (H - H_0)^3, (T_5 - \bar{T}), (T_{20} - \bar{T}), (T_{60} - \bar{T}), (T_{90} - \bar{T}), \theta, \ln(1 + \theta), \bar{J}, \bar{Q}, \bar{U}\} \quad (8)$$

式中,  $H$  为观测日水深;  $H_0$  为基准日水深;  $T_i$  为前  $i$  天的气温的均值, 重力坝  $i = 5, 20, 60, 90$ ;  $\bar{T}$  为年平均气温;  $\theta = (t - t_0) / 100$ , 其中,  $t$  为观测日期,  $t_0$  为始测日期;  $\bar{J}$  为各裂缝开度测点平均值;  $\bar{Q}$  为渗流量;  $\bar{U}$  为各扬压力监测点平均值。

初始选取的经验影响因子集为条件属性  $X$ ; 选定该工程坝顶引张线 EX5 测点水平位移(向下游为正, 反之为负)为决策属性  $Y$ , 选取 864 个监测样本为样本集  $U$ 。为了消除输入变量中不相关或弱相关的成分, 准确获取混凝土坝变形核心影响因子, 采用粗糙集理论对所有监测数据组成的初始信息表  $S = \{U, X \cup Y, V, F\}$  进行属性约简和重要性评价, 其结果如表 1 所列。

根据属性约简和重要性评价结果: 确定预测模型影响因子集为  $\{(H - H_0)^1, (H - H_0)^2, (H - H_0)^3, (T_5 - \bar{T}), (T_{20} - \bar{T}), \theta, \ln(1 + \theta), \bar{J}, \bar{Q}, \bar{U}\}$ , 相对于 EX5 测点水平位移的重要性分别为: 0.14, 0.10,

表 1 属性约简和重要性评价结果

Table 1 Attribute reduction and importance evaluation results

序号	经验影响因子	分量名称	SIM	是否约简	重要性评价 Sig( $\alpha, B, D$ )
1	$H - H_0$	水 压	-4	否	0.14
2	$(H - H_0)^2$		-2	否	0.10
3	$(H - H_0)^3$		-4	否	0.05
4	$T_5 - \bar{T}$	温 度	-7	否	0.17
5	$T_{20} - \bar{T}$		-2	否	0.34
6	$T_{60} - \bar{T}$		0	是	0.00
7	$T_{90} - \bar{T}$		0	是	0.00
8	$\theta$	时 效	-2	否	0.04
9	$\ln(1 + \theta)$		-3	否	0.05
10	$\bar{J}$	裂 缝	-2	否	0.03
11	$\bar{Q}$	渗 流	-4	否	0.05
12	$\bar{U}$	扬压力	-3	否	0.03

0.05, 0.17, 0.31, 0.00, 0.00, 0.04, 0.05, 0.03, 0.05, 0.03。因此, 引张线水平位移受水位变动、温度变化影响较大。考虑到  $T_{20} - \bar{T}$  这一影响因子对 EX5 测点影响较大, 因此水压滞后期约为 20 d。

### 4.3 预测模型样本选取

首先对原始水位、气温、渗流、裂缝、扬压力等环境量监测数据进行预处理(粗差剔除和标准化处理)。根据影响因子优选结果, 选取 2006 年 1 月 1 日 2008 年 12 月 31 日影响因子集监测数据为自变量样本, 坝顶引张线 EX5 测点水平位移(向下游为正, 反之为负)为因变量样本, 以此为基础进行基于 SVM、RF 和 RS-RF 的混凝土坝变形预测模型的应用研究。图 2 为环境量过程线。

### 4.4 预测模型训练与预测

#### 4.4.1 模型参数设置

随机森林模型决策树个数是其最重要的参数, 其选取结果对于模型的拟合预测能力起着关键性作用。为了获取最优的模型参数, 通过设置决策树个数取值范围为 [1, 25], 分别计算每个决策树个数参数下

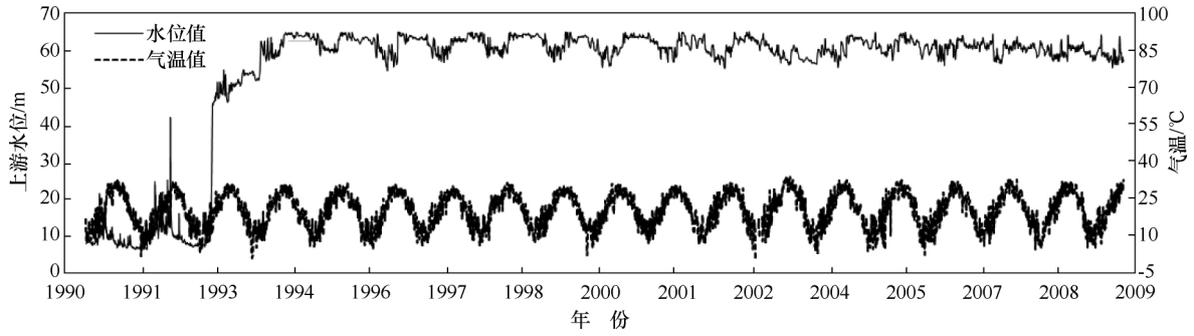


图2 环境量过程线

Fig. 2 Environmental quantity process line

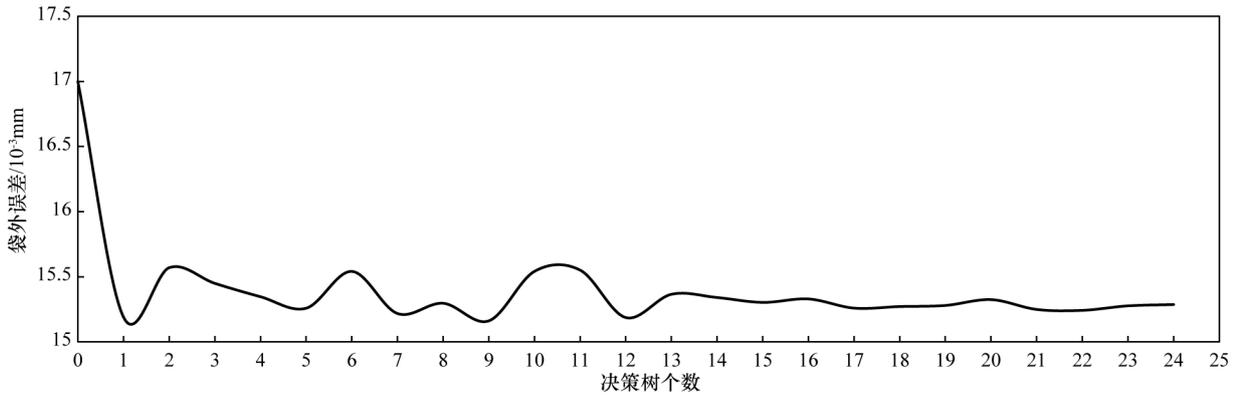


图3 决策树个数与带外误差的关系

Fig. 3 Relationship between the number of decision trees and out-of-band errors

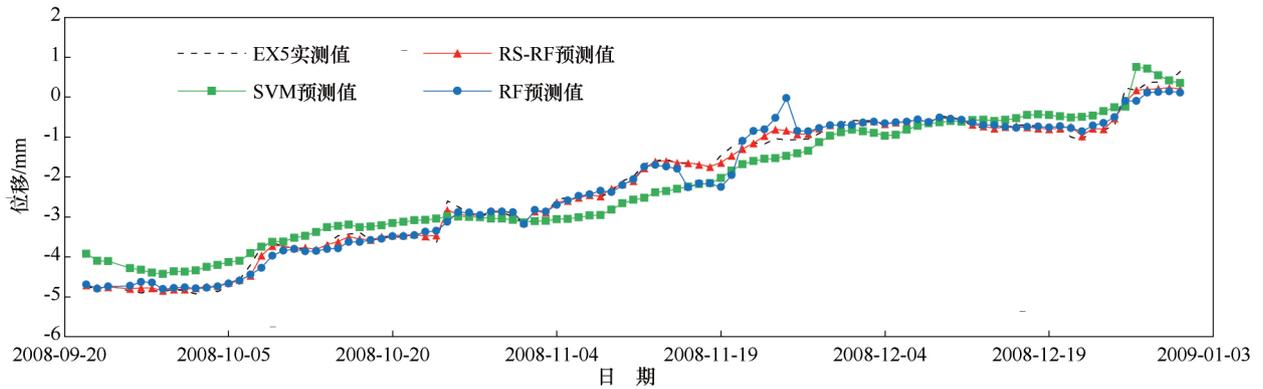


图4 各模型混凝土坝变形预测值和实测值过程线

Fig. 4 The process line of model horizontal displacement predicted value and measured value

随机森林模型的带外误差, 决策树个数与带外误差关系曲线见图3。基于最优参数下模型预测误差最小的准则, 确定本数据集下决策树个数取值为20。

#### 4. 4. 2 模型预测分析

以预处理的标准化监测数据为基础, 分别建立基于SVM、RF及RS-RF的混凝土坝变形预测模型。图4为各模型混凝土坝变形预测模型实测值和预测值过程线。

通过对图4分析可知, 基于RS-RF的混凝土坝变形预测模型曲线吻合度最高, 目标损失函数最小, 模型训练结果较优, 同时其预测性能也明显优于基于SVM、RF的混凝土坝变形预测模型。同时也证明了RS理论能较为准确地捕捉EX5测点水平位移的核心影响因子, 更好地挖掘监测信息的内部特征以反映结构的真实性态, 具有良好的实际参考价值。

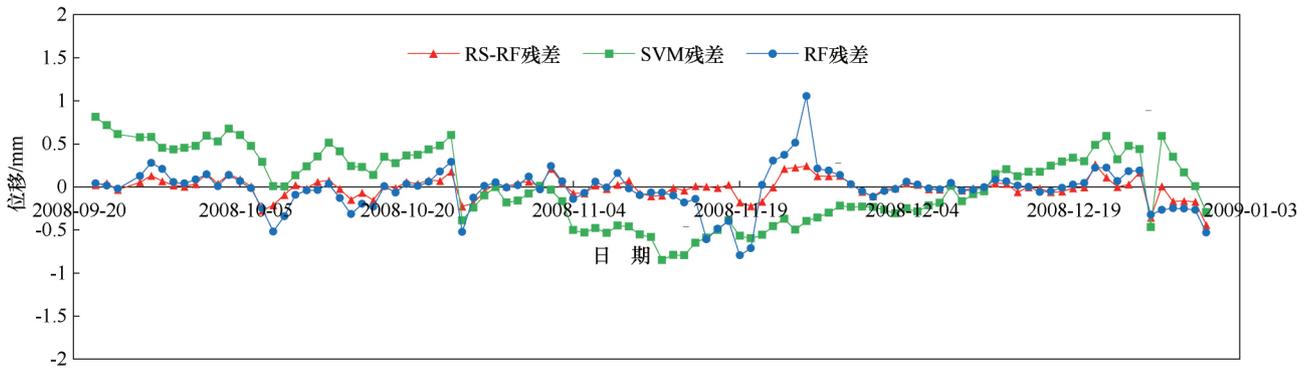


图5 模型水平位移残差

Fig. 5 Model horizontal displacement residual diagram

#### 4.4.3 模型评价

为了验证基于 RS-RF 的混凝土坝变形预测模型的性能,选取常用的预测值与实测值残差图、均方误差和平均百分比误差等指标评价预测模型精确性,并对基于 SVM、RF 及 RS-RF 的混凝土坝变形预测模型评价指标计算结果进行对比分析。各模型的水平位移残差如图 5 所示,各模型的均方误差和平均百分比误差计算结果如表 2 所列。

表 2 预测模型精确性指标

Table 2 Prediction model accuracy index

预测模型	SVM	RF	RS-RF
RMSE	0.55	0.46	0.13
MAPE	20.17	19.20	5.12

分析图 5 和表 2 可知,常用的 SVM、RF 模型能够有效地提高了模型的精确性,但基于 RS-RF 的混凝土坝变形预测模型预测效果最好,具有较高的精准度。相较于基于 SVM、RF 混凝土坝变形预测模型各项指标,基于 RS-RF 的混凝土坝变形预测模型水平位移残差最小, RMSE 低于 0.15, MAPE 低于 10,均处于较低的区间。因此,基于 RS-RF 的混凝土坝变形预测模型精准度性能较佳,预测结果更接近真实数据。

## 5 结论

基于粗糙集理论,将改进后的随机森林模型引入到大坝安全监测领域,建立了基于 RS-RF 的混凝土坝变形预测模型,通过工程实例验证可得以下结论。

(1) 机器学习技术在大坝安全预测模型建立中的成功应用,能够有效地提高模型预测性能,更加准确地预测大坝安全性态的发展趋势。

(2) 基于 RS-RF 的混凝土坝变形预测模型能够提

取影响变形的核心影响因子,提高运算效率,同时能有效地避免过拟合,具有较高的预测精度,并且兼有较强的外延性和泛化性,这些良好的性能使得混凝土坝变形安全预警具有较高的灵敏度。

(3) 随着特大型混凝土坝建设的发展,将人工智能深度学习理论、在线动态学习与时空变形预测模型相结合,建立全面的混凝土坝变形整体实时监测系统,实现监测智能化将是混凝土坝变形安全监控发展的必然趋势。

#### 参考文献 (References):

- [1] 徐镇凯,王锋,魏博文,等.多因素协同作用下混凝土坝性能退化机理研究进展[J].水利水电科技进展,2016,36(4):80-88.  
XU Zhenkai, WANG Feng, WEI Bowen, et al. Research progress on performance degradation mechanism of concrete dams under multi-factor synergy [J]. Advances in Science and Technology of Water Resource, 2016, 36(4): 80-88.
- [2] 顾冲时,苏怀智,王少伟.高混凝土坝长期变形特性计算模型及监控方法研究进展[J].水力发电学报,2016,35(5):1-14.  
GU Chongshi, SU Huaizhi, WANG Shaowei. Research progress on calculation models and monitoring methods of long-term deformation characteristics of high concrete dams [J]. Journal of Hydroelectric Engineering, 2016, 35(5): 1-14.
- [3] 仲琳.碾压混凝土坝安全监控模型及其应用研究[D].南京:河海大学,2005.  
ZHONG Lin. Study on safety monitoring model of RCC dam and its application [D]. Nanjing: Hohai University, 2005.
- [4] 卢骏,戴吾蛟,章浙涛.大坝变形变系数回归建模[J].武汉大学学报(信息科学版),2015,40(1):139-142.  
LU Jun, DAI Wujiao, ZHANG Zhetao. Regression modeling of dam deformation with variable coefficients [J]. Journal of Wuhan University(Information Science), 2015, 40(1): 139-142.
- [5] 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.  
QIU Liting, SHEN Zhenzhong, NIE Baisong. Dam deformation prediction based on stepwise regression analysis-Markov chain model

- [J]. *Hydropower and Energy Science*, 2014, 32(5): 51-55.
- [6] 吕开云. 黄河小浪底水利枢纽大坝变形预测方法研究与分析 [D]. 北京: 中国矿业大学(北京), 2012.  
LYU Kaiyun. Research and Analysis on Dam Deformation Prediction Method of Xiaolangdi Water Control Project on the Yellow River [D]. Beijing: China University of Mining and Technology (Beijing), 2012.
- [7] 孙玮玮, 李雷, 郑昊尧. 基于灰色关联度法的大坝风险后果综合评价[J]. *人民黄河*, 2012, 34(4): 102-104.  
SUN Weiwei, LI Lei, ZHENG Haoyao. Comprehensive evaluation of dam risk consequences based on grey relational grade method [J]. *Yellow River*, 2012, 34(4): 102-104.
- [8] 李广春, 戴吾蛟, 杨国祥, 等. 时空自回归模型在大坝变形分析中的应用[J]. *武汉大学学报(信息科学版)*, 2015, 40(7): 877-881.  
LI Guangchun, DAI Wujiao, YANG Guoxiang, et al. Application of spatio-temporal autoregressive model in dam deformation analysis [J]. *Journal of Wuhan University (Information Science)*, 2015, 40(7): 877-881.
- [9] 杨杰, 胡德秀, 吴中如. 大坝安全监控模型因子相关性及其不确定性研究[J]. *水利学报*, 2004, 35(12): 99-105.  
YANG Jie, HU Dexiu, WU Zhongru. Study on factor correlation and uncertainty of dam safety monitoring model [J]. *Journal of Hydraulic Engineering*, 2004, 35(12): 99-105.
- [10] 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(技术科学)*, 2011, 41(12): 1574-1579.  
GU Chongshi, WANG Yachao, PENG Yan, et al. Pathological problems of dam safety monitoring model and their treatment methods [J]. *Science of China (Science of Technology)*, 2011, 41(12): 1574-1579.
- [11] 高澜. 刘家峡主坝位移确定性模型研究[J]. *大坝与安全*, 1993(4): 51-55.  
GAO Lan. Study on deterministic model of displacement of liujiaxia main dam [J]. *Dam and Safety*, 1993(4): 51-55.
- [12] 于洪, 王国胤, 姚一豫. 决策粗糙集理论研究现状与展望[J]. *计算机学报*, 2015, 38(8): 1628-1639.  
YU Hong, WANG Guoyin, YAO Yiyu. Research status and prospect of decision rough set theory [J]. *Chinese Journal of Computers*, 2015, 38(8): 1628-1639.
- [13] 叶回春, 张世文, 黄元仿, 等. 粗糙集理论在土壤肥力评价指标权重确定中的应用[J]. *中国农业科学*, 2014, 47(4): 710-717.  
YE Huichun, ZHANG Shiwen, HUANG Yuanfang, et al. Application of rough set theory in determining the weight of soil fertility evaluation index [J]. *Chinese Academy of Agricultural Sciences*, 2014, 47(4): 710-717.
- [14] 张明. 粗糙集理论中的知识获取与约简方法的研究[D]. 南京: 南京理工大学, 2012.  
ZHANG Ming. Research on Knowledge Acquisition and Reduction Method in Rough Set Theory [D]. Nanjing: Nanjing University of Science and Technology, 2012.
- [15] 张乐文, 张德永, 李术才, 等. 基于粗糙集理论的遗传-RBF神经网络在岩爆预测中的应用[J]. *岩土力学*, 2012, 33(S1): 270-276.  
ZHANG Lewen, ZHANG Deyong, LI Shucan, et al. Application of genetic-RBF neural network based on rough set theory in rock burst prediction [J]. *Rock and Soil Mechanics*, 2012, 33(S1): 270-276.
- [16] 柳林, 刘文娟, 廖薇薇, 等. 基于随机森林和时空核密度方法的不同周期犯罪热点预测对比[J]. *地理科学进展*, 2018, 37(6): 761-771.  
LIU Lin, LIU Wenjuan, LIAO Weiwei, et al. Comparison of different cycle crime hotspots prediction based on random forest and spatio-temporal kernel density method [J]. *Advances in Geographical Sciences*, 2018, 37(6): 761-771.
- [17] 任才溶, 谢刚. 基于随机森林和气象参数的PM<sub>2.5</sub>浓度等级预测[J]. *计算机工程与应用*, 2019, 55(2): 213-220.  
REN Cairong, XIE Gang. Prediction of PM<sub>2.5</sub> concentration grade based on random forest and meteorological parameters [J]. *Computer Engineering and Application*, 2019, 55(2): 213-220.
- [18] 周泽虹. 基于机器学习的风电场功率预测研究[D]. 南京: 南京师范大学, 2017.  
ZHOU Zehong. Research on Wind Farm Power Prediction Based on Machine Learning [D]. Nanjing: Nanjing Normal University, 2017.
- [19] 王鹏新, 齐璇, 李俐, 等. 基于随机森林回归的玉米单产估测[J]. *农业机械学报*, 2019, 50(7): 237-245.  
WANG Pengxin, QI Xuan, LI Li, et al. Corn yield estimation based on stochastic forest regression [J]. *Journal of Agricultural Machinery*, 2019, 50(7): 237-245.
- [20] 陈丹玲, 卢新海, 匡兵. 基于随机森林的耕地利用效率测度模型构建及其应用[J]. *自然资源学报*, 2019, 34(6): 1331-1344.  
CHEN Danling, LU Xinhai, KUANG Bing. Construction and application of cultivated land use efficiency measurement model based on random forest [J]. *Journal of Natural Resources*, 2019, 34(6): 1331-1344.

(责任编辑 郭利娜)