文章编号:1671-6833(2026)02-0122-06

基于 TSNE-NGO-RF 算法的混凝土坝变形预测模型

郑东健[1], 赵　宇[1], 冉　成[1], 林英浩[1], 陈林泽[2]

(1. 河海大学 水利水电学院,江苏 南京 210098;2. 河海大学 河海里尔学院,江苏 南京 210098)

摘　要: 对混凝土坝变形监测资料进行合理的数据分析和准确的预测是确保大坝安全长效运行的关键手段,针对影响大坝变形的环境量具有周期性和非线性的特点,以及传统随机森林模型参数寻优方法适用性差和计算效率低等问题,提出了一种新型的大坝变形预测模型。该模型采用 t-分布式随机邻域嵌入对特征值进行降维,提高模型的分类性能,并运用北方苍鹰优化算法对传统随机森林模型进行了改进,提高了随机森林模型参数的择优选取效率。运用北方苍鹰优化算法在第 80 次迭代时即可确定随机森林模型的参数,且适应度函数为 0.249 3,相较麻雀搜索算法和粒子群优化算法取得了较好的结果。选取某混凝土坝第 18# 坝段和第 26# 坝段进行实例分析,结果表明:所提融合模型预测结果的平均绝对误差分别为 0.501 93 和 0.173 02 mm,均方误差分别为 0.359 71 和 0.043 87 mm^2,平均绝对百分比误差分别为 0.819 59%,0.113 62%,决定系数分别为 0.914 56 和 0.892 74,相较于其他模型,该模型在预测准确性和模型稳定性方面表现最优,为混凝土坝变形的精准预测开辟了新的可能性。

关键词: 混凝土坝; 变形预测; 降维; 北方苍鹰优化算法; 随机森林算法

中图分类号: TV698;TU528　**文献标志码:** A　**doi:** 10.13705/j.issn.1671-6833.2025.05.023

大坝安全监测是指对已有的观测资料进行采集、分析,并结合已有的历史资料,构建高效的预报模型,用来指导水库的科学调度和安全生产。因此,建立稳定且准确的混凝土坝变形预测模型至关重要[1-2]。近年来,研究人员已经开发了大量的数学模型,例如统计模型、确定性模型以及混合模型,这些模型通过分析混凝土坝的实时监测数据来描述和评估混凝土坝行为,是 3 种经典方法[3-4]。这些模型基于实时监测数据,主要分析了静水压力、时间和温度的影响,并能够指出监测值和影响量之间的关系。然而,由于这些特征具有周期性和非线性的特点,维度高、冗余信息较多,会影响机器学习模型的性能和效果[5-6]。对此,研究者引入随机森林(random forest, RF)算法,RF 算法凭借其独特的决策树模型特征,基于纯度和信息增益的原理,可以清晰地区分特征的重要程度,从而选择最具代表性和相关性的特征。但是,如果特征的重要程度相似或特征过多,选择特征时会损失过多的有效信息,降低模型的预测效果。TSNE (t-distributed stochastic neighbor embedding)降维可以通过仿射变换将原始数据映射到新的特征空间,以改善数据的分布和可分性,提高模型的分类性能。该算法的优点是能够在保持样本之间的局部相似性的同时,捕捉非线性的数据结构,在可视化和聚类分析等领域有广泛应用[7]。然而,TSNE 降维无法清晰地比较不同特征的重要性程度。为快速找到需要降到的目标维度,将 RF 算法和 TSNE 降维相结合,可以在不损失特征的情况下,有效地挖掘特征中的信息来提高模型的性能和效果[8]。

随机森林算法不仅可以清晰地分析特征的重要程度,还具有准确度高和训练速度快等优点,并且能够有效地处理多个参数之间存在的非线性关系,在水利水电开发中的运用日渐增多[9]。近年来对于 RF 算法的研究主要集中在提升其可解释能力与参数优化方面,黄海燕等[10]利用随机森林算法建立兼顾预测能力和解释能力的大坝变形监控模型,成功地对各个特征开展重要性分析;潘宇等[11]通过基于改进的随机森林算法建立了大坝安全监测历史数据质量评价算法,实例分析结果明显优于原始算法;曾永军等[12]采用经验法进行参数寻优,以构建基于随

收稿日期: 2025-08-10;**修订日期:** 2025-10-26

基金项目: 国家自然科学基金资助项目(52179128)

作者简介: 郑东健(1965—),男,福建闽清人,河海大学教授,博士,博士生导师,主要从事水工结构安全监控研究,E-mail:zhengdj@hhu.edu.cn。

引用本文: 郑东健,赵宇,冉成,等. 基于 TSNE-NGO-RF 算法的混凝土坝变形预测模型[J]. 郑州大学学报(工学版),2026,47(2):122-127,135. (ZHENG D J, ZHAO Y, RAN C, et al. Concrete dam deformation prediction model based on TSNE-NGO-RF algorithm [J]. Journal of Zhengzhou University (Engineering Science), 2026,47(2):122-127,135.)

机森林的混凝土坝变形预测模型;仝晓哲等[13]采用网格搜索法寻找随机森林模型的最优参数组合。但是经验法和网格搜索法这两种传统方法都有很大的局限性。经验法主观性强,有局部极值倾向,算法预测性能差;网格搜索法可获得良好的优化结果,但是受到网络结构的影响,计算效率较低。因此,需要更好的方法来提高优化性能。群体智能算法具有准确性高、运算速度快等优势,例如麻雀搜索算法(sparrow search algorithm, SSA)、粒子群算法(particle swarm optimization, PSO)、灰狼优化算法(grey wolf optimizer, GWO)等,常用于求解优化问题[14-15]。北方苍鹰优化算法(northern gannet optimization,NGO)能够在解空间中搜索最优解或接近最优解的解决方案,相较于其他优化算法具有良好的稳定性以及较高的收敛精度等优点,可以解决复杂优化问题[16]。因此,本文通过 TSNE 降维提取特征中的有效信息,采用 NGO 算法优化随机森林的超参数,提出了一种基于 TSNE-NGO-RF 的大坝变形预测模型。

1　算法原理

1.1　TSNE 算法

TSNE 算法是一种用于将复杂的高维数据转换为更容易解释的低维形式的非线性算法。该算法的核心在于保留高维数据样本之间的局部相似性关系,并在低维空间中重建这种相似性关系来实现降维。公式为

$$P_{i|j}=\frac{\exp(-\|x_i-x_j\|^2\cdot(2\sigma_i^2)^{-1})}{\sum\limits_{k\neq i}\exp(-\|x_i-x_j\|^2\cdot(2\sigma_i^2)^{-1})};\tag{1}$$

$$Q_{i|j}=\frac{\exp(-\|y_i-y_j\|^2\cdot(2\sigma_i^2)^{-1})}{\sum\limits_{k\neq i}\exp(-\|y_i-y_j\|^2\cdot(2\sigma_i^2)^{-1})};\tag{2}$$

$$\mathrm{KL}(P\|Q)=\sum_i\sum_j p_{i|j}\log\frac{P_{ij}}{Q_{ij}}。\tag{3}$$

式中:$P_{i|j}$、$Q_{i|j}$为相似度概率;σ_i 表示以 x_i 为中心点的高斯分布的方差;P_{ij} 为高维空间中的联合分布概率;Q_{ij} 为低维空间中计算样本概率密度;$\mathrm{KL}(P\|Q)$ 为 P 和 Q 之间分布的 KL 散度[17]。

1.2　RF 算法

RF 算法是一种强大且有效的并行式集成分类器,它通过随机取样构建多个决策树,在所有的决策树训练完成后,将各个决策树合并成一个决策树森林,对于新数据的分类或预测任务,森林中的每棵树都会给出自己的判断,最终通过各个决策树投票方式做出决策[18]。RF 这种集成学习方法,通过将多个决策树的预测结果进行集成,可以降低单个决策树对噪声和异常值的敏感度,从而提高模型的鲁棒性[19]。RF 算法流程如图 1 所示。

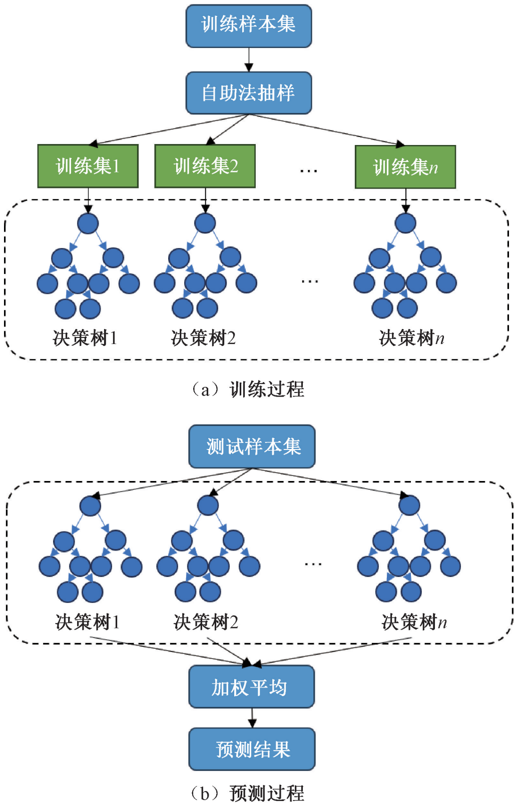


图 1　RF 算法流程图

Figure 1　Flow chart of RF algorithm

1.3　NGO 算法

本文使用 NGO 算法的主要目的是通过迭代寻优找出当前初始信号的最佳分解参数,其数学模型如下。

北方苍鹰对猎物的选择在全区搜索空间中随机进行全局搜索,这一过程会有效增加算法的搜索能力,从而确定猎物位置的最优区域。在这一阶段,北方苍鹰进行猎物选择和攻击的行为,用公式描述为

$$P_i=X_k, i=1,2,\cdots,N, k=1,2,\cdots,i-1,\ i+1,\cdots,N。\tag{4}$$

$$x_{i,j}^{\mathrm{new},P_1}=\begin{cases}x_{i,j}+r(p_{i,j}-Ix_{i,j}), & F_{P_i}<F_i;\\ x_{i,j}+r(x_{i,j}-p_{i,j}), & F_{P_i}\geqslant F_i。\end{cases}\tag{5}$$

$$X_i=\begin{cases}X_i^{\mathrm{new},P_1}, & F_i^{\mathrm{new},P_1}<F_i;\\ X_i, & F_i^{\mathrm{new},P_1}\geqslant F_i。\end{cases}\tag{6}$$

式中:P_i 为第 i 只猎物位置;F_{P_i} 为目标函数;X_i^{new,P_1} 为第 i 个苍鹰的新位置;$x_{i,j}^{\mathrm{nwe},P_1}$ 为第 i 只苍鹰的第 j 维度的新位置;F_i^{new,P_1} 为对应的目标函数值;r 为(0,1)的随机数;I 为 1 或 2 的随机整数。

北方苍鹰持续追踪逃跑中的猎物直至完成捕

猎,通过模拟这一行为,可以有效提升模型的搜索能力。假设这种狩猎活动接近于一个半径为 R 的攻击位置,这一阶段可以用以下公式来表示:

$$x_{i,j}^{\mathrm{new},P_2} = x_{i,j} + R(2r-1)x_{i,j}; \tag{7}$$

$$R = 0.02\left(1-\frac{t}{T}\right); \tag{8}$$

$$X_i = \begin{cases} X_i^{\mathrm{new},P_2},\ F_i^{\mathrm{new},P_2} < F_i; \\ X_i,\ F_i^{\mathrm{new},P_2} \geqslant F_i。\end{cases} \tag{9}$$

式中:t 为迭代次数;T 为最大迭代次数;$x_{i,j}$ 为当前迭代时的种群位置;X_i^{new,P_2} 为第 i 个苍鹰的新位置;$x_{i,j}^{\mathrm{nwe},P_2}$ 为第 i 只苍鹰在第 j 维的新位置;F_i^{new,P_2} 为目标函数[20]。

1.4 TSNE-NGO-RF 算法

TSNE-NGO-RF 算法流程如下:

步骤 1 输入原始数据,对数据进行预处理,包括异常值修正、缺失值填补以及归一化等操作。将处理后的数据按时间先后顺序进行划分,将数据的前 90%设为训练集,后 10%设为测试集。

步骤 2 对数据中的特征值进行 TSNE 降维处理,得到最适用于模型的特征值组合并用于后续的模型训练。

步骤 3 输入训练集数据进行模型训练,初始化北方苍鹰优化算法中的种群数和随机森林算法中的参数,生成模型迭代次数并将需要进行寻优的决策树个数 n 和最小叶子节点数 m 输入 NGO 算法中,使用 NGO 优化算法进行迭代寻优,设定适应度函数为平均绝对误差(MAE)并以此评估模型性能,当适应度函数最小时得到随机森林模型最优模型参数。

步骤 4 将测试集数据输入训练好的预测模型中,得到模型的预测值的图像和评价指标等结果,最后通过比较预测值和真实值的图像拟合效果以及计算评价指标的差异来量化分析模型的预测效果。

图 2 为基于 TSNE-NGO-RF 算法的预测模型流程图。

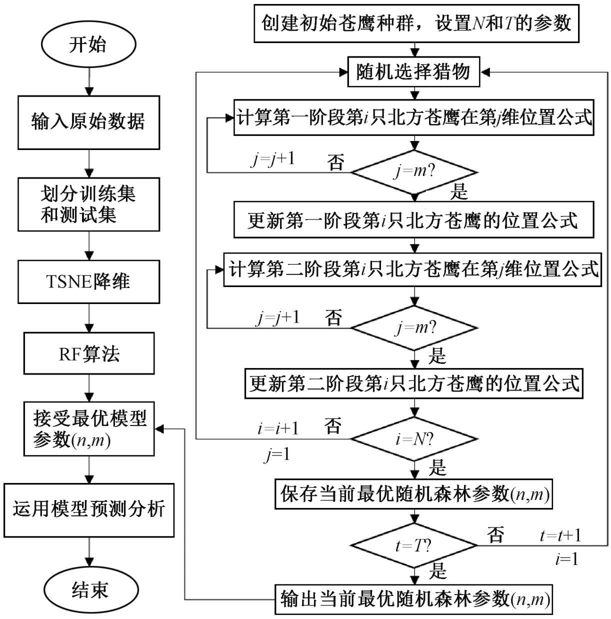


图 2 基于 TSNE-NGO-RF 算法的预测模型流程图

Figure 2 Flow chart of prediction model based on TSNE-NGO-RF algorithm

2 实例分析

2.1 某大坝位移监测概况

某大坝为重力拱坝,共分 30 个坝段,布置有正垂线测点 6 个,倒垂线测点 10 个,正垂布置在 8#、18#、26# 3 个代表性坝段。坝区岩层除原生的单斜层状构造外,较大的断层有 F11、F31、F32、F35 等,微细裂隙发育,地质条件复杂,倒垂主要监测坝基和 F_{11}、F_{32} 大断层及两坝肩岩体变形。监测系统更新改造时,新增倒垂测点 6 条、正垂测点 1 条[21]。大坝垂线测点分布图见图 3。

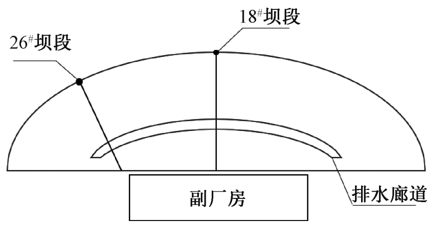


图 3 大坝垂线测点分布图

Figure 3 Distribution map of vertical measurement points for the dam

大坝水平位移受复杂环境的综合影响,对评价大坝安全稳定运行有重要意义,在该大坝监测数据中,正垂线测点数据较为完整,用于监测典型坝段的位移,因此选取 1973 年 1 月至 2007 年 7 月期间第 18、第 26 坝段坝顶水平位移监测数据建立数据集,对数据集进行预处理后总共得到 3 854 组数据,作为预测模型的数据输入。图 4 为两个坝段的测点位移。

2.2 模型参数设定

在混凝土坝的运行过程中,坝顶水平位移受到静水水压、温度变化及时间效应的显著影响。为了模拟这些影响,采用以下模型因子:水压因子 H、H^2、H^3、H^4,温度因子 $\sin\frac{2\pi it}{365}$、$\cos\frac{2\pi it}{365}$、$\sin\frac{2\pi it}{365}\cos\frac{2\pi it}{365}$,以及时效因子 $\ln\theta$。其中,H 为监测日上下游水位差;t 为从开始测量到当前日期的天数;θ 为 $0.01t$[22]。对这些模型因子进行重要性评估,结果

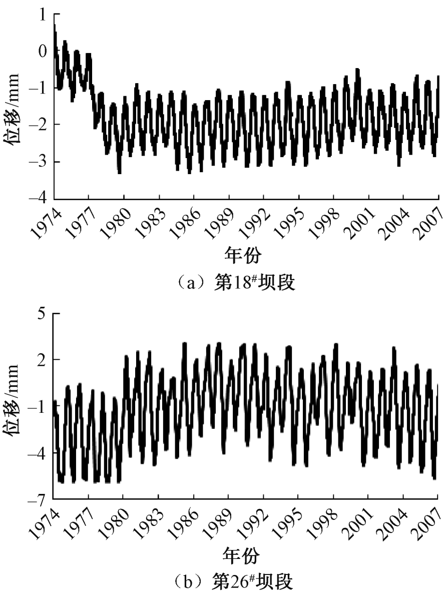


图 4　坝段测点位移

Figure 4　Displacement of measuring points in the dam section

如图 5 所示。从图 5 可以看出,温度因子 $\sin\dfrac{2\pi it}{365}$、$\cos\dfrac{2\pi it}{365}$ 的重要性较高,因此本文将 8 维模型因子通过 TSNE 算法降至 2 维,可以很好地保留其关键信息和数据完整性,然后将降维后的数据输入 TSNE-NGO-RF 组合模型中,用于后续的位移预测。

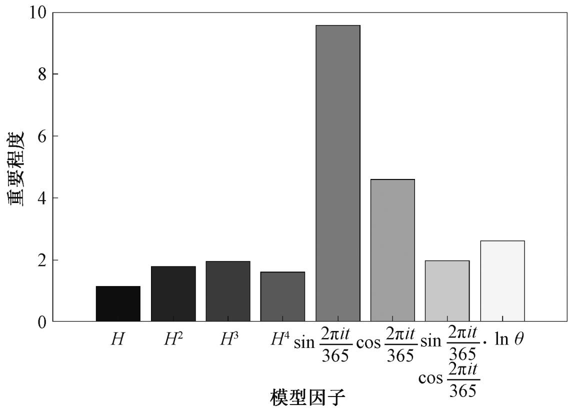


图 5　模型因子的重要程度

Figure 5　Importance of model factors

为寻找随机森林模型的最优参数,进行以下参数设置:设定最大迭代次数为 200 次,同时初始种群数量为 5,规定随机森林主要超参数决策树个数 n 和最小叶子节点数 m 的优化范围分别为[100, 1 000]、[1,10]。设定适应度函数为采用平均绝对误差(MAE),当适应度函数为最小值时,选择其所对应的 n 和 m 作为随机森林模型参数的最优组合[23]。为比较不同模型的寻优性能,使用不同优化算法(SSA、PSO、NGO)寻找随机森林的最优参数,图 6 为不同优化算法的迭代收敛过程。由图 6 可以看出,NGO 算法的收敛速度最快,第 80 次迭代完成,且收敛精度最高,MAE 最小为 0.249 3,此时决策树个数 n 和最小叶子节点数 m 分别为 556 和 5。由于 NGO 算法结合了启发式搜索和进化算法的特点,这种结合不仅提升了算法的优化精度,还加快了优化过程的速度,在处理复杂优化问题时具有较好的性能,尤其适用于处理多峰值和离散类型的优化问题。

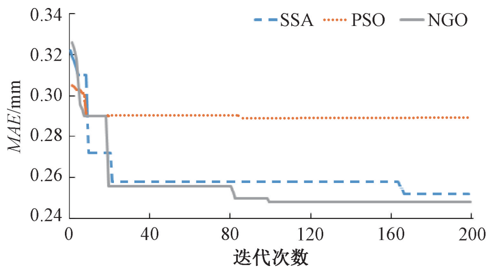


图 6　SSA、PSO、NGO 算法迭代收敛过程

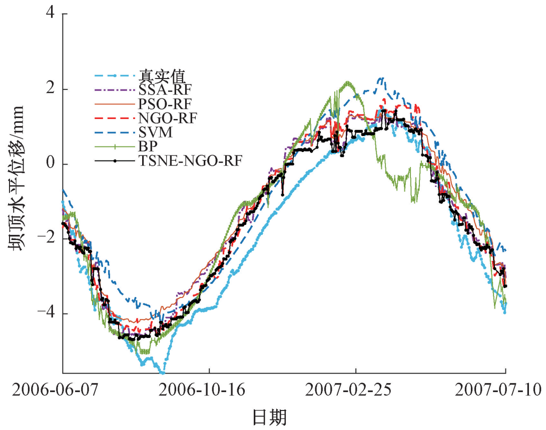
Figure 6　Iterative convergence process of SSA, PSO, NGO algorithms

2.3　模型预测结果与分析

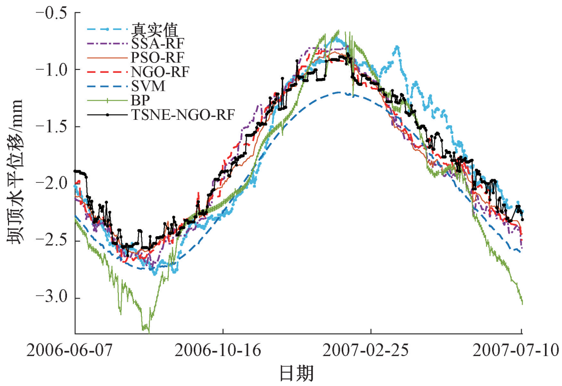
为了更好地对比和评价模型的准确性,对比 TSNE-NGO-RF 与 SSA-RF、PSO-RF、NGO-RF 模型的预测结果,同时为了验证 RF 算法具有集成学习的优势,引入支持向量机(support vector machine, SVM)模型和反向传播神经网络(back propagation, BP)模型进行预测,该仿真实验的硬件环境:CPU 为 Intel(R) Core(TM) i7-9750H; 显卡为 NVIDIA GeForce GTX 1660 Ti with Max-Q。所有仿真实验及模型均在 MATLAB R2024b 软件中建立。每个大坝变形预测模型的预测曲线如图 7 所示。为评估模型的预测准确性,采用平均绝对误差(MAE)、平均绝对百分比误差($MAPE$)、均方误差(MSE)和复相关系数 R^2 进行评估。表 1 总结了图 7 所示 6 个模型对 $18^{\#}$ 坝段和 $26^{\#}$ 坝段的预测评估结果。分析结果可知,TSNE-NGO-RF 模型的预测评价指标在这 6 个模型中均表现最好,评价指标相较于其他模型有明显的提升。尤其是 R^2 增加了 7.20%~17.07%。SVM 和 BP 模型预测精度较差,这是因为这两种模型在超参数取值方面存在过拟合现象,而随机森林模型具有集成学习的优势,能有效避免过拟合现象,得到了较好的预测结果。同时,TSNE-NGO-RF 模型相较于 SSA-RF、PSO-RF、NGO-RF 模型具有更高的寻优速度和精度,体现出该融合模型的优越性。

为验证本文提出模型的鲁棒性,使用生成对抗网络(generative adversarial network, GAN)来生成对抗样本并进行测试。对 $18^{\#}$ 坝段共 3 854 组原始数据进行生成对抗网络后得到 5 564 组数据,并使用 TSNE-

NGO-RF 模型进行预测,结果见表 2。



(a) $18^{\#}$坝段



(b) $26^{\#}$坝段

图 7 大坝变形预测模型的预测曲线

Figure 7 Prediction curves of dam deformation prediction models

表 1 预测模型对 $18^{\#}$ 和 $26^{\#}$ 坝段的预测性能比较

Table 1 Comparison of predictive performance of prediction models for the $18^{\#}$th and $26^{\#}$th dam sections

坝段	模型	MSE/mm^2	MAE/mm	$MAPE$/%	R^2
$18^{\#}$坝段	SSA-RF	0.576 17	0.646 63	1.093 21	0.853 15
	PSO-RF	0.735 99	0.758 42	1.137 23	0.825 19
	NGO-RF	0.483 34	0.613 08	1.039 96	0.885 20
	SVM	0.887 82	0.890 71	1.614 11	0.789 13
	BP	0.921 08	0.824 88	1.504 03	0.781 23
	TSNE-NGO-RF	0.359 71	0.501 93	0.819 59	0.914 56
$26^{\#}$坝段	SSA-RF	0.073 26	0.205 70	0.139 43	0.820 89
	PSO-RF	0.058 49	0.184 29	0.126 85	0.856 99
	NGO-RF	0.059 44	0.192 34	0.132 01	0.854 67
	SVM	0.080 77	0.24217	0.189 32	0.802 54
	BP	0.092 62	0.24516	0.150 85	0.773 55
	TSNE-NGO-RF	0.043 87	0.173 02	0.113 62	0.892 74

表 2 原始样本和生成对抗样本测试结果对比

Table 2 Comparison of adversarial samples test results

样本	MSE/mm^2	MAE/mm	$MAPE$/%	R^2
原始样本	0.359 71	0.501 93	0.819 59	0.914 56
生成对抗样本	0.462 21	0.586 07	0.825 99	0.890 22

综上所述,TSNE-NGO-RF 模型不仅在 $18^{\#}$ 坝段和 $26^{\#}$ 坝段的预测结果较好,适用于多测点预测和长期监测数据预测,而且在进行对抗样本噪声输入后,模型在对抗样本上的准确率与原样本仅有 2.7%的差距,说明模型不仅具有一定的泛化能力,而且对数据扰动的接受能力强,具有一定的鲁棒性。

3 结论

(1)TSNE 降维方法可以处理维度高、冗余信息多的特征值,有效改善数据的分布和可分性,提高模型的分类性能。NGO 算法结合了群体智能算法和进化算法的优势,具有很高的收敛精度。与其他优化算法相比,NGO 算法减少了迭代次数并且能快速达到最佳适应度值,更加适合解决复杂的优化问题。

(2)TSNE-NGO-RF 混凝土坝变形预测模型展现出集成学习方法的显著优点,并有效规避了过拟合的风险,其预测结果的准确性和可靠性均表现优异。相较于其他优化模型,TSNE-NGO-RF 模型预测结果取得了较好的结果,具有较高的预测精度和较强的稳健性。

参考文献:

[1] PRAKASH G, DUGALAM R, BARBOSH M, et al. Recent advancement of concrete dam health monitoring technology: a systematic literature review[J]. Structures, 2022, 44: 766-784.

[2] 樊小永, 刘嘉. 水库大坝运行期安全监测分析[J]. 工程技术研究, 2025, 10(9): 223-225.
FAN X Y, LIU J. Analysis of safety monitoring of reservoir dam during operation period[J]. Engineering and Technological Research, 2025, 10(9): 223-225.

[3] 王健, 王士军. 全国水库大坝安全监测现状调研与对策思考[J]. 中国水利, 2018(20): 15-19.
WANG J, WANG S J. Investigations on reservoir dam safety monitoring in China and measure study[J]. China Water Resources, 2018(20): 15-19.

[4] 龙耿文, 袁树才, 罗家良, 等. 水利工程大坝安全监测技术与发展研究[J]. 水上安全, 2025(20): 148-150.
LONG G W, YUAN S C, LUO J L, et al. Research on dam safety monitoring technology and development of water conservancy projects[J]. Maritime Safety, 2025(20): 148-150.

[5] 吴中如, 吉肇泰. 坝前水库水温的变化规律和预测研究[J]. 水力发电, 1984, 10(4): 33-41.
WU Z R, JI Z T. Research on the changes and prediction of water temperature in reservoir[J]. Water Power, 1984, 10(4): 33-41.

[6] 刘健, 王继敏, 杨强, 等. 混凝土大坝变形监测异常值识别方法及应用[J]. 水利水电科技进展, 2025, 45(6): 77-84.
LIU J, WANG J M, YANG Q, et al. Method for identification of outliers in concrete dam deformation monitoring and its application[J]. Advances in Science and Technology of Water Resources, 2025, 45(6): 77-84.
[7] MAATEN L V D, HINTON G. Visualizing data using t-SNE[J]. Journal of Machine Learning Research, 2008, 9: 2579-2605.
[8] BREIMAN L. Random forests[J]. Machine Learning, 2001,45:5-32.
[9] SU H Y, SHEN W J, WANG J R, et al. Machine learning and geostatistical approaches for estimating aboveground biomass in Chinese subtropical forests[J]. Forest Ecosystems, 2020, 7(1): 64.
[10] 黄海燕, 艾星星, 刘兴阳, 等. 基于可解释性分析的大坝变形监控模型对比研究[J]. 人民长江, 2024, 55(9): 203-209.
HUANG H Y, AI X X, LIU X Y, et al. Comparison of monitoring model for dam deformation based on interpretability analysis[J]. Yangtze River, 2024, 55(9): 203-209.
[11] 潘宇, 李登华, 丁勇. 基于改进随机森林的大坝监测数据质量评价算法[J]. 人民长江, 2024, 55(2): 231-237.
PAN Y, LI D H, DING Y. Data quality evaluation algorithm on dam monitoring based on improved random forest[J]. Yangtze River, 2024, 55(2): 231-237.
[12] 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, 2021, 52(5): 82-88.
ZENG Y J, ZHANG J W, CAO D G, et al. Application of RS-RF model in deformation prediction of concrete dam[J]. Water Resources and Hydropower Engineering, 2021, 52(5): 82-88.
[13] 仝晓哲, 赵黎晨, 王佳明. 随机森林回归在大坝变形预测中的应用研究[C]//2019 年江苏省测绘地理信息学会学术年会论文集. 南京:江苏省测绘地理信息学会, 2019: 49-51.
TONG X Z, ZHAO L C, WANG J M. Research on the application of random forest regression in dam deformation prediction[C]//Proceedings of the 2019 Academic Annual Conference of the Jiangsu Society for Geodesy Photogrammetry and Cartography. Nanjing: Jiangsu Society for Geodesy Photogrammetry and Cartography, 2019: 49-51.
[14] 宋宝钢, 包腾飞, 向镇洋, 等. 基于小波的 SSA-ELM 大坝变形时空预测模型[J]. 长江科学院院报, 2023, 40(8): 145-151.
SONG B G, BAO T F, XIANG Z Y, et al. Wavelet-based SSA-ELM spatio-temporal prediction model for dam deformation[J]. Journal of Changjiang River Scientific Research Institute, 2023, 40(8): 145-151.
[15] 张文煜, 马可可, 郭振海, 等. 基于灰狼算法和极限学习机的风速多步预测[J]. 郑州大学学报(工学版), 2024, 45(2): 89-96.
ZHANG W Y, MA K K, GUO Z H, et al. Multistep prediction of wind speed based on grey wolf algorithm and extreme learning machine[J]. Journal of Zhengzhou University (Engineering Science), 2024, 45(2): 89-96.
[16] DEHGHANI M, HUBÁLOVSKY Š, TROJOVSKY P. Northern goshawk optimization: a new swarm-based algorithm for solving optimization problems[J]. IEEE Access, 2021, 9: 162059-162080.
[17] 谭建所, 吴兴华, 徐文光, 等. 基于 tSNE-LSTM 算法的工业预测模型[J]. 现代电子技术, 2024, 47(12): 81-85.
TAN J S, WU X H, XU W G, et al. Industrial prediction model based on tSNE-LSTM algorithm[J]. Modern Electronics Technique, 2024, 47(12): 81-85.
[18] BARAL P, HAQ M A. Spatial prediction of permafrost occurrence in Sikkim Himalayas using logistic regression, random forests, support vector machines and neural networks[J]. Geomorphology, 2020, 371: 107331.
[19] SHANNON C E. A mathematical theory of communication [J]. Bell System Technical Journal, 1948, 27(3): 379-423.
[20] 詹明强, 陈波, 袁志颖. 基于 KPCA-NGO-LSSVM 的混凝土坝变形预测模型[J]. 水电能源科学, 2024, 42(8): 127-131.
ZHAN M Q, CHEN B, YUAN Z Y. Deformation prediction method of concrete dam based on KPCA-NGO-LSSVM[J]. Water Resources and Power, 2024, 42(8): 127-131.
[21] 秦俊德. 陈村大坝坝基渗压问题的探讨[J]. 大坝与安全, 1991(4): 28-33.
QIN J D. Discussion on the seepage pressure problem of Chencun Dam foundation [J]. Large Dam and Safety, 1991(4): 28-33.
[22] 吴中如. 水工建筑物安全监控理论及其应用[M]. 北京: 高等教育出版社, 2003: 157-170.
WU Z R. Safety monitoring theory & its application of hydraulic structures[M]. Beijing: Higher Education Press, 2003: 157-170.
[23] 张石, 郑东健, 陈卓研. 基于改进 PSO-RF 算法的大坝变形预测模型[J]. 水利水电科技进展, 2022, 42(6): 39-44.
ZHANG S, ZHENG D J, CHEN Z Y. Dam deformation prediction model based on improved PSO-RF algorithm [J]. Advances in Science and Technology of Water Resources, 2022, 42(6): 39-44.

(下转第 135 页)

Zonal Inversion of Underwater Topography of Large Water Bodies Based on Landsat Images

DOU Ming[1,2], SHI Yuxian[1], QU Lingbo[2], WANG Jihua[3], XING Aoqi[2]

(1. School of Water Conservancy and Transportation, Zhengzhou University, Zhengzhou 450001, China; 2. School of Ecology and Environment, Zhengzhou University, Zhengzhou 450001, China; 3. Henan Provincial Natural Resources Monitoring and Land Improvement Institute, Zhengzhou 450016, China)

Abstract: To address the difficulty of obtaining underwater topography data for large water bodies with insufficient data, Danjiangkou Reservoir was selected as the study area, and a retrieval method based on Landsat remote sensing imagery and water depth zoning was proposed. The underwater topography of the shallow and deep water areas of the reservoir was reconstructed using the waterline kriging interpolation method and four water depth inversion models (single-band, dual-band ratio, BP neural network, and multi-band random forest), and the inversion accuracy was evaluated. The results showed that the underwater topography inversion in the shallow water area performed well (Root Mean Square Error, $RMSE = 2.553$ m). In the deep water area, the multi-band random forest model performed best in the Han Reservoir area ($RMSE = 2.428$ m), while the BP neural network model performed best in the Dan Reservoir area ($RMSE = 1.599$ m). The accuracy of different inversion models varied across different depths and regions, with the multi-band random forest model demonstrating advantages in deep-water topography inversion.

Keywords: underwater topography inversion; Landsat remote sensing imagery; BP neural network model; multi-band random forest model; Danjiangkou Reservoir

(上接第 127 页)

Concrete Dam Deformation Prediction Model Based on TSNE-NGO-RF Algorithm

ZHENG Dongjian[1], ZHAO Yu[1], RAN Cheng[1], LIN Yinghao[1], CHEN Linze[2]

(1. School of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210098, China; 2. Institute Hohai-Lille, Hohai University, Nanjing 210098, China)

Abstract: Reasonable data analysis and accurate prediction of deformation monitoring data for concrete dams are key means to ensure the safe and long-term operation of dams. In response to the periodic and nonlinear characteristics of environmental variables that could affect dam deformation, as well as the shortcomings of traditional random forest model parameter optimization methods such as poor applicability and low computational efficiency, a new type of dam deformation prediction model was proposed. The model uses t-distributed random neighborhood embedding to reduce the dimensionality of eigenvalues and improve the classification performance of the model. The traditional random forest model was improved using the northern eagle optimization algorithm, which enhanced the efficiency of selecting optimal parameters for the random forest model. The parameters of the random forest model could be determined using the northern eagle optimization algorithm in the 80th iteration, and the fitness function was 0.249 3, which achieves better results compared to the Sparrow Search Algorithm and Particle Swarm Optimization Algorithm. The analysis of the $18^{\#}$th and $26^{\#}$th sections of a concrete dam showed that the fusion model proposed in this study had average absolute errors of 0.501 93 and 0.173 02 mm, mean square errors of 0.359 71 and 0.043 87 mm^2, average absolute percentage errors of 0.819 59% and 0.113 62%, and determination coefficients of 0.914 56 and 0.892 74, respectively. Compared with other models, this model performed better in prediction accuracy and model stability, opening up new possibilities for accurate prediction of concrete dam deformation.

Keywords: concrete dam; deformation prediction; dimensionality reduction; northern eagle optimization algorithm; random forest algorithm