

基于 RF - GWO 的水利工程地质 渗透系数智能反演分析

雷艳¹, 温立峰², 赵明仓¹, 殷乔刚²

(1. 中国电建集团西北勘测设计研究院有限公司, 陕西 西安 710065; 2. 西安理工大学 省部共建
西北旱区生态水利国家重点实验室, 陕西 西安 710048)

摘要: 地质渗透系数是准确分析水利工程渗流的关键参数。针对传统反演方法计算效率低、精度差的问题,采用有限元正演模型和正交试验设计构建渗透系数反演样本集,建立了基于随机森林(RF)算法的渗流计算代理模型;在此基础上,引入灰狼优化(GWO)算法,提出了基于 RF - GWO 的渗透系数智能反演方法,并以 Z 抽水蓄能电站为研究案例进行了验证。结果表明:RF 模型对各钻孔水位预测结果均接近实测值,性能优于 CART 和 BP 模型;GWO 可搜寻到地质最佳渗透系数,钻孔水位反演结果合理,相对误差最大为 0.42%,精度满足工程要求,计算的天然渗流场分布形态也符合一般山体渗流场分布规律。建立的反演模型能够快速准确地推断工程区地层渗透系数,具有实际工程应用价值。

关键词: 地质渗透系数; 反演分析; 正交试验设计; 随机森林; 灰狼优化

中图分类号:TV139.1

文献标识码:A

文章编号:1672-643X(2024)02-0139-010

Intelligent inversion analysis of geological permeability coefficients in hydraulic engineering based on RF - GWO

LEI Yan¹, WEN Lifeng², ZHAO Mingcang¹, YIN Qiaogang²

(1. PowerChina Northwest Engineering Corporation Limited, Xi'an 710065, China; 2. State Key Laboratory of
Eco-hydraulics in Northwest Arid Region of China, Xi'an University of Technology, Xi'an 710048, China)

Abstract: The geological permeability coefficients are the crucial parameters for accurately analyzing the seepage in hydraulic engineering. To address the issues of low efficiency and accuracy of conventional inversion approaches, this research utilized a combination of the finite element forward model and orthogonal experimental design to construct a sample set for the permeability coefficient inversion. Then, a permeability calculation surrogate model based on the random forest (RF) algorithm was developed. Subsequently, the grey wolf optimization (GWO) algorithm was introduced to develop an intelligent inversion method based on RF - GWO. Taking the Z pumped storage power station as the case study, it is found that the water level prediction outcomes of the RF model are closely aligned with the actual measured values at each borehole, with the performance surpassing CART and BP models. The optimal geological permeability coefficient calculated by GWO performs excellently in the borehole water level inversion, with a maximum relative error of 0.42%, thereby satisfying the required accuracy for engineering applications. The calculated natural seepage field distribution conforms to the general distribution patterns of mountainous seepage fields. Therefore, the proposed inversion model can rapidly and precisely predict the geological permeability coefficient in the project area, demonstrating significant engineering practicality and value.

Key words: geological permeability coefficient; inversion analysis; orthogonal experimental design; random forest (RF); grey wolf optimization (GWO)

收稿日期:2023-06-06; 修回日期:2023-12-07

基金项目:国家自然科学基金项目(51909215)

作者简介:雷艳(1982—),女,陕西渭南人,硕士,正高级工程师,主要从事水电站与抽水蓄能电站水工结构设计研究。

通讯作者:温立峰(1989—),男,江西赣州人,博士,副教授,硕士生导师,研究方向为水工结构工程。

1 研究背景

目前,我国新建了一大批高坝大库和抽水蓄能电站等水利工程,有效地促进了国家经济社会的持续健康发展^[1-2]。为确保工程的安全建设和正常运行,必须深入了解工程区域内的渗流分布情况。通常使用数值模拟方法分析和判断渗流情况,但这一过程实现的前提是获得可靠的天然地层渗透系数^[3-4]。相比于室内试验和现场原位试验,以现场观测资料为基础的反演方法能够充分利用已知数据信息,更为准确地获取渗透参数^[5],目前已成为确定地层渗透参数的主要手段之一^[6]。随着科学理论的发展,许多方法被相继提出,包括正反算法^[7]、展开水头势函数法^[8]、序列二次规划法^[9]、融合渗透系数张量理论和序列二次规划反演^[10]等。然而,这些传统的反演方法需要大量调用渗流分析正演模型,计算耗时且效率极低,反演的参数精度也相对较低。因此,有必要采取有效措施来提高参数反演方法的计算效率和精度。

近年来,基于智能算法的代理模型作为一种优化反演效率的新方法,得到了广泛应用。基于支持向量机(support vector machine, SVM)^[11]、BP(back propagation)神经网络^[12-13]、极限学习机(extreme learning machine, ELM)^[14]、粒子群算法(particle swarm optimization, PSO)^[15]和多元自适应回归样条(multivariate adaptive regression splines, MARS)^[16]等算法的代理模型相继问世。代理模型方法通过有限元正演模型生成学习样本,利用算法强大的非线性拟合能力构建渗透系数(输入变量)与钻孔水位(输出变量)之间的映射关系,从而代替渗流正演模型快速求解。代理模型方法提高了反演计算效率,但其本质仍是传统的“正问题反算”参数反演方法。在代理模型的基础上,结合优化算法反演最佳参数已成为参数反演的新途径,众多学者对此进行了研究。李雅琦等^[17]基于相关向量机与布谷鸟算法建立了反演分析模型,成功确定了研究区域内各地层的渗透系数;Shu等^[18]利用 ELM 和优化算法对大坝防渗帷幕渗透系数进行了反演分析,确定其最佳取值;李皓璇等^[19]建立了基于广义回归神经网络的代理模型,并结合粒子群优化算法搜索到了某混凝土面板堆石坝待反演分区的最佳渗透系数。这些研究表明利用优化算法并结合代理模型去反演渗流参数的最佳取值是可行的,且提高了反演精度。

然而,在对水利工程地质渗透系数进行非线性

建模过程中发现,上述代理模型存在预测精度低、鲁棒性差、收敛速度慢、泛化能力一般和容易陷入局部最小值等问题。鉴于此,本文引入随机森林(random forest, RF)算法^[20],构建渗流分析代理模型,结合群智能优化算法——灰狼优化^[21](grey wolf optimization, GWO),建立渗透系数智能反演模型,并以 Z 抽水蓄能电站为例,对所建模型的可行性与有效性进行验证。

2 RF 和 GWO 算法基本原理

2.1 RF 算法原理

随机森林(RF)是一种广泛应用于非线性回归问题的集成学习算法^[22]。该算法的核心是利用集成学习方法——Bootstrap aggregating (Bagging),有放回地抽取(Bootstrap)样本集进行决策树建模,将存在过度拟合和局部收敛问题的单个回归器(决策树)组合成多个回归器“森林”去提高模型的预测性能。其基本思路如图 1 所示,具体按照以下 4 个步骤执行:

(1) 设训练样本集合为 $S = \{x_1, x_2, \dots, x_n\}$,利用 Bootstrap 方法有放回地从 S 中抽取 n 次样本,随机地产生 T 个新训练集 $S_i (i = 1, 2, \dots, T)$;新集合中的样本总数与 S 相同,但仅包含了 S 中约 62.3% 的样本;

(2) 根据训练集 $S_i (i = 1, 2, \dots, T)$ 生成相应的决策树模型 $C_1, C_2, \dots, C_T$;

(3) T 个决策树模型分别对测试集样本 X 进行测试,得到相应的结果为 $C_1(X), C_2(X), \dots, C_T(X)$;

(4) 对 T 个决策树模型的输出结果求均值,作为 X 的最终计算结果。

RF 模型具有 3 个重要参数,分别为决策树的个数 N_{tree} 、节点分裂的随机特征数 M_{try} 和叶节点的最小样本数 $Node_{\text{size}}$,其中 $Node_{\text{size}}$ 对 RF 模型的性能和效率影响较小,通常使用默认值。

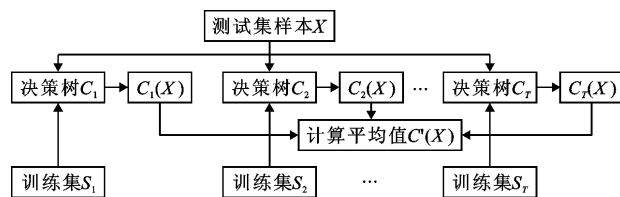


图 1 Bagging 算法基本思路

2.2 GWO 算法原理

GWO 算法是由 Mirjalili 等^[21]通过模拟灰狼群

体捕食行为构建的一种新型全局搜索算法。该算法具有调节参数少、随机搜索路径优、收敛快和求解精度高等特点,目前已被应用于凝汽器真空预测^[23]、城市需水预测^[24]、燃气轮机参数整定^[25]和铅锌烧结反粉料水分预测^[26]等领域。

灰狼种群具有 α 、 β 、 δ 和 ω 等4个社会等级,其中, α 为头狼,对狼群起决定性作用; β 为第二级的狼,负责协助 α 狼; δ 为第三级的狼,服从于 α 狼和 β 狼的决定; ω 为最后一层的狼,严格服从于 α 、 β 和 δ 狼。它们的位置参数分别为优化目标的最优解、次优解、第三优解和候补解。灰狼包围猎物行为的数学定义见公式(1)。

$$\begin{cases} \vec{D} = |\vec{C} \cdot \vec{X}_p(t) - \vec{X}(t)| \\ \vec{X}(t+1) = \vec{X}_p(t) - \vec{A} \cdot \vec{D} \end{cases} \quad (1)$$

式中: t 为当前迭代次数; $\vec{D}$ 为灰狼与猎物间的距离向量; $\vec{X}_p$ 和 $\vec{X}$ 分别为猎物和灰狼的位置向量; $\vec{A}$ 和 $\vec{C}$ 为系数向量,具体计算见公式(2)。

$$\begin{cases} \vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a} \\ \vec{C} = 2\vec{r}_2 \end{cases} \quad (2)$$

式中: $\vec{a}$ 为收敛因子,在迭代过程中从二线性递减到0; $\vec{r}_1$ 和 $\vec{r}_2$ 为 $[0,1]$ 间的随机向量。狼群追踪和击杀猎物时,其位置参数的更新策略详见文献[23]~[26],最终对所有狼的位置求平均值获得位置更新,见公式(3)。

$$\vec{X}(t+1) = \frac{\vec{X}_\alpha(t+1) + \vec{X}_\beta(t+1) + \vec{X}_\delta(t+1)}{3} \quad (3)$$

式中: $\vec{X}(t+1)$ 为下一代候选狼的位置向量; $\vec{X}_\alpha(t+1)$ 、 $\vec{X}_\beta(t+1)$ 和 $\vec{X}_\delta(t+1)$ 分别为候选狼向 α 、 β 和 δ 移动的矢量。

3 基于RF-GWO的渗透系数反演模型

3.1 三维渗流场计算原理

根据水流连续性方程,稳定渗流的基本微分方程可表示为公式(4)^[6]。

$$\frac{\partial}{\partial x} \left(k_x \frac{\partial H}{\partial x} \right) + \frac{\partial}{\partial y} \left(k_y \frac{\partial H}{\partial y} \right) + \frac{\partial}{\partial z} \left(k_z \frac{\partial H}{\partial z} \right) = 0 \quad (4)$$

式中: H 为水头函数,m; k_x 、 k_y 、 k_z 分别为 x 、 y 、 z 方向的渗透系数,m/s。

根据常见的边界条件^[3]和变分原理,上述渗流微分方程的求解可归结为求公式(5)泛函的极小值。

$$I(H) = \iiint_{\Omega} \frac{1}{2} \left[k_x \left(\frac{\partial H}{\partial x} \right)^2 + k_y \left(\frac{\partial H}{\partial y} \right)^2 + \right.$$

$$\left. k_z \left(\frac{\partial H}{\partial z} \right)^2 \right] dx dy dz - \iint_{S_2} q H ds \Rightarrow \min \quad (5)$$

式中: Ω 为计算区域。

将渗流计算区域离散化,由公式(5)可得求解三维渗流场的有限元基本方程^[27]。

$$[K] \cdot \{H\} = \{F\} \quad (6)$$

式中: $[K]$ 为整体渗透矩阵; $\{H\}$ 为节点水头列阵; $\{F\}$ 为对应流量矩阵。

3.2 GWO算法的适应度函数

渗流参数反演的主要目标是在其取值范围内寻找最优值,因此它本质上是一个优化问题求解的过程。本文采用GWO算法搜寻工程区地质最佳渗透系数,建立基于RF-GWO的反演模型。在反演过程中,以钻孔水位实测值与计算值的最小均方误差(mean squared error, MSE)作为GWO的适应度函数,计算公式见公式(7)。

$$f = \min MSE = \min \left[\frac{1}{q} \sum_{k=1}^q (H_k - H'_k)^2 \right] \quad (7)$$

式中: H_k 为第 k 个勘探钻孔的水位实际勘测值,m; H'_k 为第 k 个勘探钻孔的水位有限元计算值,m; q 为钻孔的总数。

目标函数的约束条件见公式(8)。

$$\text{s. t. } x_m^d \leq x_m \leq x_m^u \quad (m = 1, 2, \dots, M) \quad (8)$$

式中: x_m 为待反演渗透系数,m/s; x_m^u 和 x_m^d 分别为渗透系数的上、下限,m/s; M 为待反演渗透系数总数。

3.3 RF-GWO反演模型的建立

本文基于RF和GWO算法,建立工程区地质渗透系数反演模型,智能搜寻其最佳值,以反映工程区的天然渗流性态。反演模型的执行流程如图2所示,主要由7个步骤完成。

(1)采用正交试验方法生成渗透系数组合样本,通过有限元正演模型计算各组合下的钻孔水位,建立渗透系数反演样本集。

(2)RF的输入变量为渗透系数组合,输出变量为钻孔水位有限元计算值;选取渗透系数样本集的前80%为训练集,剩余20%为测试集;同时,使用归一化公式(9)对样本数据进行预处理,将变量数据限定在 $[0,1]$ 内,以消除变量量纲对RF模型的影响, x_m' 为归一化后的渗透系数值。

$$x_m' = \frac{x_m - x_m^d}{x_m^u - x_m^d} \quad (9)$$

(3)采用十折交叉验证确定RF的最佳参数 M_{try} 和 N_{tree} ,并训练得到RF(代理模型)。

(4)验证步骤(3)建立的代理模型训练精度是

否满足误差阈值;若精度达到要求则进行步骤(5),否则重复步骤(3)~(4)。

(5) 对输出结果反归一化,利用统计指标^[28]——平均绝对误差(MAE)、平均绝对百分误差(MAPE)、均方根误差(RMSE)和拟合优度(R^2)来评估代理模型的性能。

(6) 代理模型计算 GWO 搜索时不断构造新渗透系数组合下的钻孔水位,若该组合下 GWO 的适应度优于已有最佳适应度,则更新为当前最优渗透系数组合和适应度,否则进入到步骤(7)。

(7) 若 GWO 的迭代次数尚未达到设定的最大值,重复步骤(6),否则将当前渗透系数组合作为最终反演结果并输出。

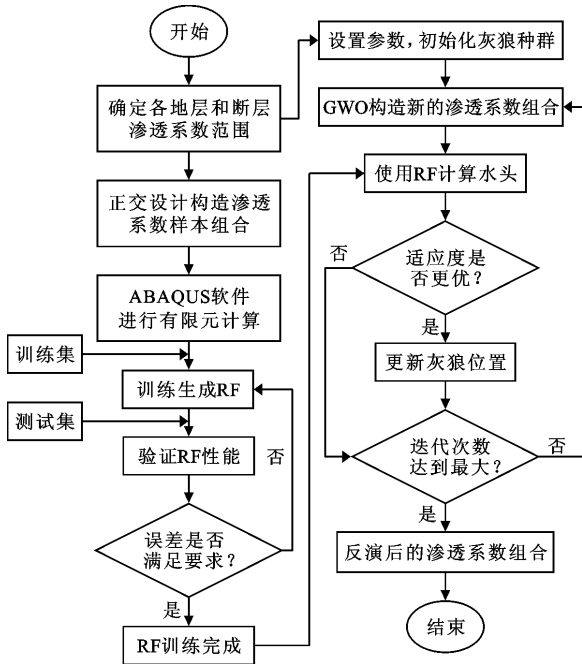


图2 工程区地质渗透系数反演模型的执行过程

4 工程实例

4.1 工程概况

Z 抽水蓄能电站为一等大(1)型工程,枢纽建筑物主要由上、下库沥青混凝土面板堆石坝(最大坝高分别为 40.0 和 38.9 m)、下水库泄洪排沙洞和输水发电系统(上、下游电站进/出水口、尾水明渠、输水隧洞和地下厂房)等组成。上水库正常蓄水位为 2 332.00 m,死水位为 2 300.00 m,调节库容为 $680 \times 10^4 \text{ m}^3$;下水库正常蓄水位为 1 745.00 m,死水位为 1 718.00 m,调节库容为 $679 \times 10^4 \text{ m}^3$;输水发电系统沿线山体雄厚,地下厂房上覆岩体厚度为 478~533 m。

工程区基岩埋深大,主要由微新风化泥质粉砂岩组成,总体上具有弱透水性。地下水主要存在于基岩裂隙中,并且较不发育。压水试验表明,强风化和弱风化岩体的透水率差异不大,微风化岩体的透水率相对较低。下水库址区覆盖有深厚的第四系松散堆积物,主要为全新统坡洪积碎石土(Q_4^{pl+dl})和上更新统冲洪积物(Q_3^{al+pl}),局部夹有粉土层。在上、下库址区,地下水位和相对隔水层($q \leq 1 \text{ Lu}$)的顶板埋深均低于正常蓄水位,勘探钻孔均未揭示地下水的存在,埋深分别大于 200.0 和 110.5 m。输水发电系统的勘测钻孔也揭示了其地下水位埋深相对较深。工程区内裂隙和断层发育不明显。

4.2 有限元模型的建立

为确保建立的数值模型计算边界的合理性,本文根据渗流场情况和场区地形确定了模型计算域边界。四周边界设置在山体分水岭位置,上、下库和输水系统及地下厂房洞室群周围岩体的厚度在 6 倍坝高或开挖跨度以上。模型计算范围及其钻孔测点的分布如图 3 所示。

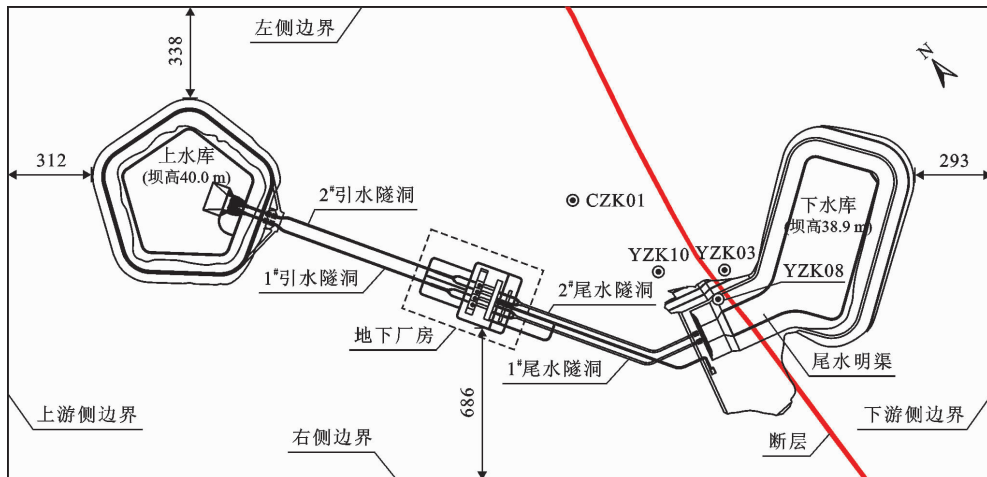


图3 工程区渗流场反演分析模型计算范围及钻孔测点分布(单位:m)

根据工程区地质剖面(图 4)及水文地质资料,通过 ABAQUS 建立了工程区渗流场反演计算模型(图 5),精细化模拟下库址区覆盖层(厚度约为 178 m)、强风化层(厚度约为 25 m)、弱风化层(厚度约为 20 m)、微新岩层和断层破碎带(宽度为 10~30 m,断层穿越尾水明渠,总体产状为 NW320°~350°SW∠60°~75°)。模型 X 轴为正东方向,最大距离

为 3 679. 41 m; Y 轴为正北方向,最大距离为 1 745. 68 m; Z 为高程方向,底高程为 1 405. 50 m,顶高程为对应的地表高程,最高高程为 2 326. 14 m。采用 4 节点等参单元(C3D4P)对模型进行网格剖分,节点数量为 171 742,单元数量为 949 015。计算时模型四周侧面边界为定水头边界,其值为推测或调整的地下水位;模型底部为不透水边界。

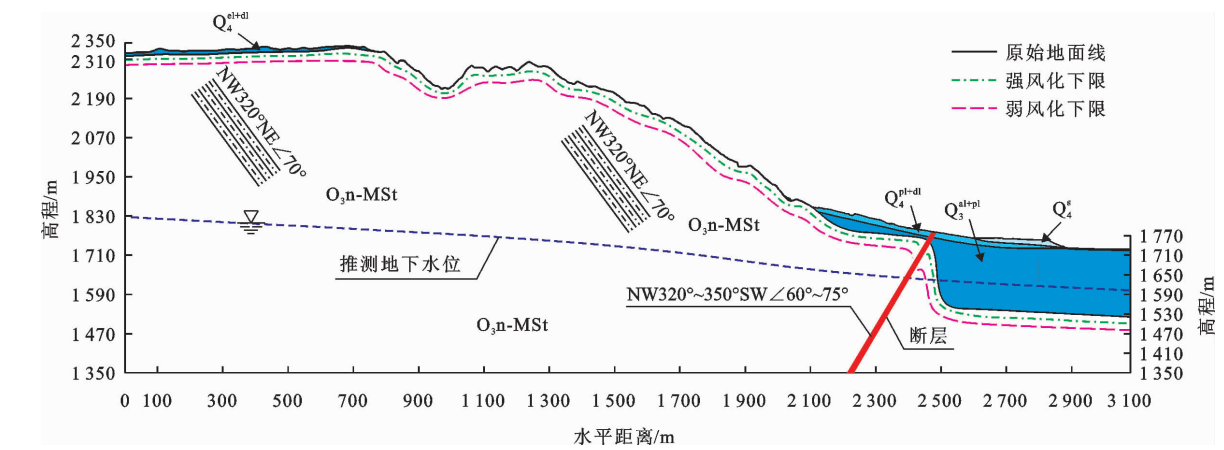


图 4 工程区地质剖面图

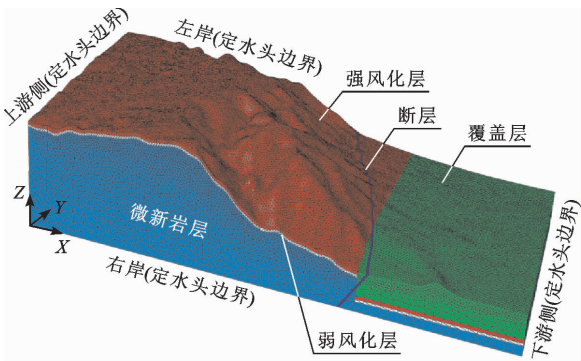


图 5 渗流场反演计算模型

本文根据天然渗流场 4 个钻孔(YZK03、YZK08、YZK10 和 CZK01)的实测水位,基于所提出的反演模型对 Z 抽水蓄能电站地质渗透系数开展反演研究。

4.3 基于正交设计的样本构造

在反演分析时,各地层渗透系数按各向同性考虑,即 k_1 、 k_2 和 k_3 取值相等,其中 k_1 和 k_2 为切向渗透系数, k_3 为法向渗透系数。断层渗透系数根据文献[29]的建议按各向异性考虑,其中 k_1 和 k_2 的值相等,为实际测量值; k_3 为周围岩体(微新岩层)的渗透系数值。

根据地质勘测报告并结合一般工程经验,确定的本研究地质渗透系数取值范围见表 1。

表 1 各地层和断层渗透系数取值范围 m/s

岩层	k_1 、 k_2 实测范围	k_3 实测范围
覆盖层	$3.85 \times 10^{-5} \sim 3.00 \times 10^{-4}$	$3.85 \times 10^{-5} \sim 3.00 \times 10^{-4}$
强风化层	$8.00 \times 10^{-7} \sim 2.96 \times 10^{-5}$	$8.00 \times 10^{-7} \sim 2.96 \times 10^{-5}$
弱风化层	$6.50 \times 10^{-7} \sim 7.90 \times 10^{-6}$	$6.50 \times 10^{-7} \sim 7.90 \times 10^{-6}$
微新岩层	$7.00 \times 10^{-9} \sim 4.00 \times 10^{-7}$	$7.00 \times 10^{-9} \sim 4.00 \times 10^{-7}$
断层	$4.48 \times 10^{-7} \sim 1.24 \times 10^{-6}$	$7.00 \times 10^{-9} \sim 4.00 \times 10^{-7}$

为了在保证 RF 模型精度的同时,不显著增加有限元正分析的工作量,本文采用多因素多水平的正交试验设计方法^[30]构造具有代表性的渗透系数样本集。正交表是正交试验设计的关键,通常表示为 $L_n(t^c)$,其中 L 代表正交表; n 表示总的试验次数; t 表示因素的水平数; c 表示可以安排的最多因素个数。强风化层、弱风化层、微新岩层和断层的渗透系数为正交试验的因素,各因素取 9 个水平数,如表 2 所示。因此,试验样本按正交表 $L_{81}(9^5)$ 设计,并使用 SPSS (statistical product and service solutions) 软件生成相应的 81 组渗透系数组合方案。在此基础上,通过有限元正演模型计算各渗透系数组合下的水位值,构造 81 组渗透系数反演样本。其中,选取样本集的 65 组样本(数据的前 80%)为训练集,用于学习生成 RF 模型;剩余的 16 组样本(数据的后 20%)为测试集,用于验证训练生成的 RF 模型的性能。

表 2 基于正交设计的不同岩层渗透系数各水平取值

水平	覆盖层	强风化层	弱风化层	微新岩层	断层	
	$k_1 = k_2 = k_3$	$k_1 = k_2 = k_3$	$k_1 = k_2 = k_3$	$k_1 = k_2 = k_3$	$k_1 = k_2$	k_3
1	3.85×10^{-5}	8.00×10^{-7}	6.50×10^{-7}	7.00×10^{-9}	4.48×10^{-7}	7.00×10^{-9}
2	7.12×10^{-5}	4.40×10^{-6}	1.56×10^{-6}	5.61×10^{-8}	5.47×10^{-7}	5.61×10^{-8}
3	1.04×10^{-4}	8.00×10^{-6}	2.46×10^{-6}	1.05×10^{-7}	6.46×10^{-7}	1.05×10^{-7}
4	1.37×10^{-4}	1.16×10^{-5}	3.37×10^{-6}	1.54×10^{-7}	7.45×10^{-7}	1.54×10^{-7}
5	1.69×10^{-4}	1.52×10^{-5}	4.28×10^{-6}	2.04×10^{-7}	8.44×10^{-7}	2.04×10^{-7}
6	2.02×10^{-4}	1.88×10^{-5}	5.18×10^{-6}	2.53×10^{-7}	9.43×10^{-7}	2.53×10^{-7}
7	2.35×10^{-4}	2.24×10^{-5}	6.09×10^{-6}	3.02×10^{-7}	1.04×10^{-6}	3.02×10^{-7}
8	2.67×10^{-4}	2.60×10^{-5}	6.99×10^{-6}	3.51×10^{-7}	1.14×10^{-6}	3.51×10^{-7}
9	3.00×10^{-4}	2.96×10^{-5}	7.90×10^{-6}	4.00×10^{-7}	1.24×10^{-6}	4.00×10^{-7}

5 结果分析与讨论

5.1 RF 模型性能的验证

本研究以钻孔 YZK03 为例验证 RF 模型在渗透参数反演中的有效性,并与经典的分类回归树 CART(classification & regression tree)模型和 BP 神经网络模型进行对比分析。通过十折交叉验证搜寻到 RF 参数 N_{tree} 和 M_{try} 的最佳取值分别为 300 和 4。同时,设定 CART 模型参数 minLeafSize 为 6,smoothingK 为 20;BP 模型的隐含层神经元个数为 7,最大迭代次数为 200。RF、CART 和 BP 模型对钻孔水位

的预测结果如图 6 所示。从图 6(a)可以看出,3 种模型的预测曲线与有限元计算值曲线较为接近,预测误差较小。然而,从图 6(b)可以看出,整体上 RF 模型的预测残差值小于 CART 和 BP 模型。

通过评估指标定量评估 3 种模型的预测精度(表 3)。从表 3 可以看出,RF 模型的各项评估指标均优于 CART 和 BP 模型。其中 R^2 在训练集和测试集上分别达到了 0.954 和 0.847,并且其他指标也处于较低水平,这表明 RF 模型对钻孔 YZK03 水位计算值的拟合效果十分显著,其预测精度比 CART 和 BP 模型更高。

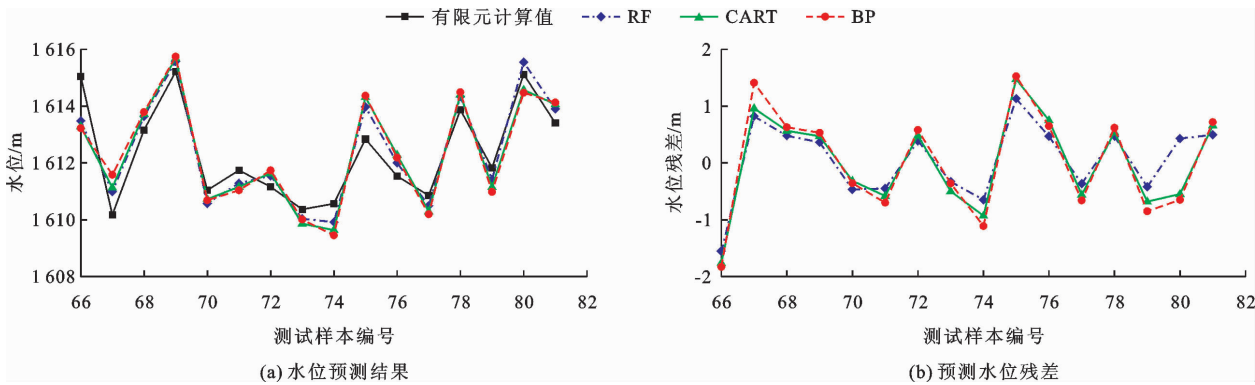


图 6 RF、CART 和 BP 模型对钻孔 YZK03 水位预测结果

表 3 RF、CART 和 BP 模型对钻孔 YZK03 水位预测结果的评价指标值

模型	模型训练				模型验证			
	MAE /m	MAPE /%	RMSE /m	R^2	MAE /m	MAPE /%	RMSE /m	R^2
RF	0.342	0.021	0.391	0.954	0.581	0.036	0.664	0.847
CART	0.422	0.026	0.483	0.929	0.736	0.046	0.825	0.764
BP	0.468	0.029	0.518	0.919	0.824	0.051	0.920	0.705

为评估 RF 模型的泛化能力,将钻孔 YZK08、YZK10 和 CZK01 的测试样本水位值用于验证。RF、CART 和 BP 模型在此 3 个钻孔处的水位预测曲线如图 7 所示。其中 RF 模型在 3 个钻孔处的水位预测误差均最小,预测曲线的变化趋势也与水位计算值曲线更为一致,RF 相较于其他 2 种模型表现出更好的抗干扰和泛化能力。

表 4 为 RF、CART 和 BP 模型对钻孔 YZK08、YZK10 和 CZK01 水位预测结果的评价指标值。由表 4 可以看出,RF 模型在 3 个钻孔测试集上的 R^2 分别为 0.859、0.883 和 0.833,比 CART (R^2 分别为 0.768、0.799 和 0.740)和 BP 模型 (R^2 分别为 0.725、0.776 和 0.698)更优,证明了其卓越的泛化能力。

综上,本文建立的 RF 模型在 3 个钻孔处均表现出了更好的性能,准确地预测了钻孔水位。因此,RF 模型具有良好的预测精度以及泛化性能和鲁棒性,可作为渗流参数反演的代理模型。

5.2 基于 GWO 算法的渗透系数反演

本文在 RF 代理模型的基础上,利用 GWO 算法对渗透系数进行全局寻优。设初始灰狼数量 $N = 100$,最大迭代次数 $T = 200$ 。为验证 GWO 算法的寻优能力,引入经典的优化算法 GA (genetic algorithm)^[31]和粒子群优化 (particle swarm optimization, PSO) 算法^[32]与之对比。在计算时,PSO 算法的参数设置为:粒子群数量为 200,个体学习因子为 1.5,社会学习因子为 2,最大惯性因子为 1,最小惯性因子为 0.8,最大迭代次数为 200。GA 算法的参数设置为:种群规模为 200,代沟为 0.85,交叉概率为 0.6,变异概率为 0.01,最大遗传代数 为 200。

图 8 为 GWO、PSO 和 GA 3 种算法搜索最优解的收敛过程。图 8 显示,GWO、PSO 和 GA 算法分别迭代至 76、101 和 116 次搜索到最优值。反演得到的各地层和断层的最佳渗透系数如表 5 所示。

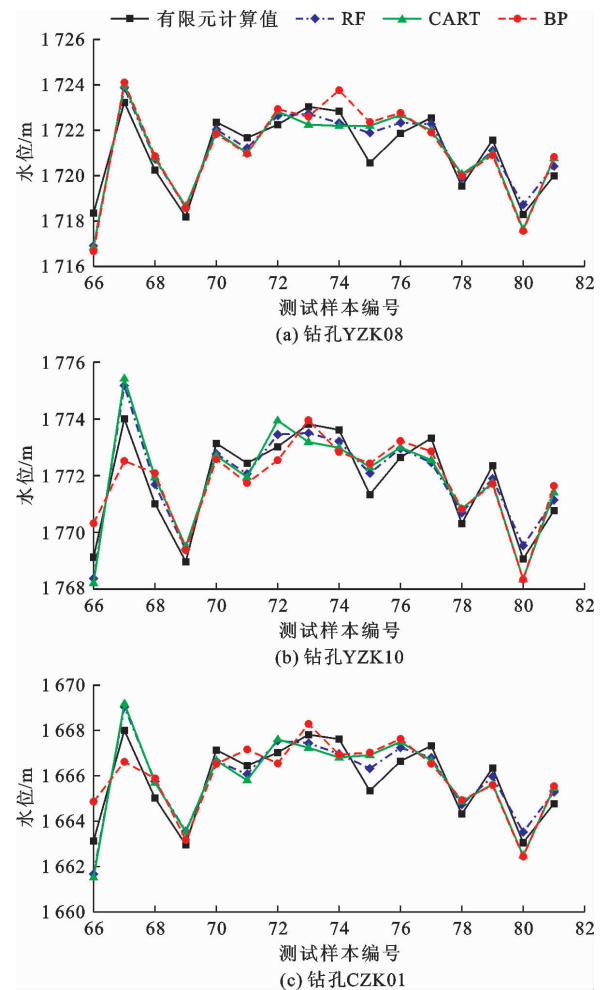


图 7 RF、CART 和 BP 模型对不同钻孔水位的预测结果

表 4 RF、CART 和 BP 模型对不同钻孔水位预测结果的评价指标值

钻孔	模型	模型训练				模型验证			
		MAE /m	MAPE /%	RMSE /m	R^2	MAE /m	MAPE /%	RMSE /m	R^2
YZK08	RF	0.343	0.020	0.398	0.952	0.547	0.032	0.637	0.859
	CART	0.409	0.024	0.469	0.934	0.746	0.043	0.816	0.768
	BP	0.448	0.026	0.511	0.921	0.800	0.046	0.890	0.725
YZK10	RF	0.354	0.020	0.431	0.944	0.532	0.030	0.581	0.883
	CART	0.375	0.021	0.439	0.942	0.715	0.040	0.760	0.799
	BP	0.388	0.022	0.449	0.939	0.731	0.041	0.802	0.776
CZK01	RF	0.381	0.023	0.422	0.946	0.631	0.038	0.693	0.833
	CART	0.405	0.024	0.464	0.935	0.796	0.048	0.865	0.740
	BP	0.470	0.028	0.525	0.917	0.836	0.050	0.931	0.698

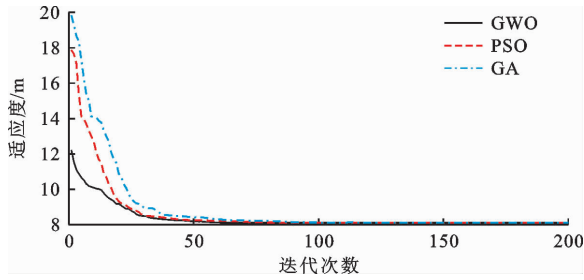


图8 GWO、PSO 和 GA 算法的收敛过程

表5 各地层和裂隙渗透系数反演结果 m/s

岩层	$k_1、k_2$ 计算值	k_3 计算值
覆盖层	2.38×10^{-4}	2.38×10^{-4}
强风化层	1.42×10^{-5}	1.42×10^{-5}
弱风化层	4.54×10^{-6}	4.54×10^{-6}
微新岩层	1.67×10^{-7}	1.67×10^{-7}
断层	1.02×10^{-6}	1.67×10^{-7}

为验证反演得到的渗透系数的合理性,将其代入有限元正演模型计算出各钻孔水位值,并与钻孔实测水位值进行对比,对比结果见表6。表6中绝对误差 e 的计算见公式(10),相对误差 e_r 的计算见

公式(11)。

$$e = H_c - H_m \quad (10)$$

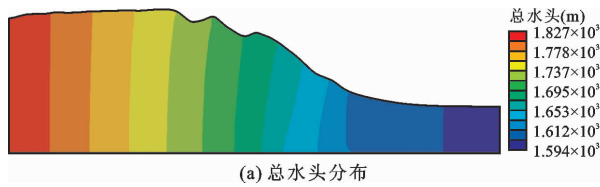
$$e_r = [(H_c - H_m)/H_m] \times 100\% \quad (11)$$

式中: H_c 为钻孔水位的计算值,m; H_m 为钻孔水位的实测值,m。

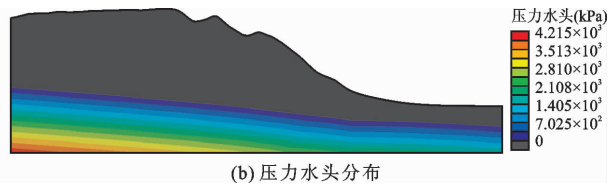
表6 各钻孔水位的实测值与反演计算值对比

钻孔编号	计算水位/m	实测水位/m	绝对误差 e /m	相对误差 e_r /%
YZK03	1 614.13	1 608.56	5.57	0.35
YZK08	1 721.89	1 718.59	3.30	0.19
YZK10	1 768.52	1 771.13	2.61	0.15
CZK01	1 666.26	1 659.32	6.94	0.42

将计算模型中 $X、Y$ 方向的中间位置($Y = 872.840$ m和 $X = 1\,839.705$ m)作为渗流计算结果分析的典型剖面。计算所得的 $Y = 872.840$ m和 $X = 1\,839.705$ m剖面总水头和压力水头分布云图如图(9)~(10)所示。其中, X 剖面左侧和右侧分别对应模型的右岸和左岸; Y 剖面左侧和右侧分别对应模型的上游和下游。

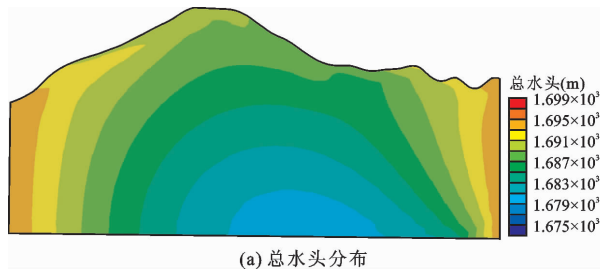


(a) 总水头分布

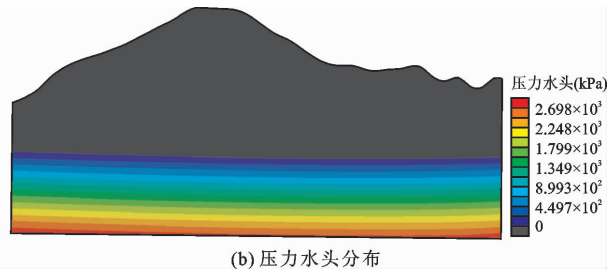


(b) 压力水头分布

图9 $Y = 872.840$ m剖面总水头及压力水头计算结果



(a) 总水头分布



(b) 压力水头分布

图10 $X = 1\,839.705$ m剖面总水头及压力水头计算结果

5.3 讨论

文献[12]~[14]、[17]和[19]仅对反演得到的渗透系数合理性进行了分析,而缺少对代理模型性能的分析,本文对建立的渗流代理模型性能进行验证,弥补了上述文献的不足。根据不同钻孔水位预测结果,RF模型的性能评估指标(MAE、MAPE、RMSE和 R^2)均优于CART和BP模型,论证了RF模型良好的泛化性能和鲁棒性,可代替有限元模型对钻孔水位值进行有效预测。

从图8可以看出,随着GWO、PSO和GA3种种群的不断进化,其收敛速度均趋于平缓,最终收敛于同一适应度,但GWO算法的收敛曲线更为平滑,初始寻优结果更好,并且收敛速度也更快。由表6可以看出,所有钻孔的反演计算水位绝对误差最大值为6.94 m,相对误差最大值为0.42%。根据文献[6]、[17]、[33]~[34]的研究,当反演结果中各钻孔水位的绝对误差 $e \leq 10$ m、相对误差 $e_r \leq 5\%$ 时,获得的渗透系数较为合理,精度能够满足一般工程

需要。此外,由图 9、10 可知,天然渗流场上游库区侧水头值较高,下游侧水头值较低,呈现由上游侧向下游侧渗流补给的规律,上游水头线分布相对密集,水力坡降相对较大。天然渗流场模型左侧和右侧水头值基本相等,呈现由山体两侧向山体中部渗流补给的规律,两侧水头线分布相对密集,水力坡降相对较大。计算得到的天然渗流场分布形态与钻孔揭示的规律一致,符合一般山体渗流场分布规律,总水头和孔隙水压力分布基本合理。因此,本文的反演方法和结果是合理可靠的。

然而,本文仅使用工程区单一钻孔数据建模去反演地质的最佳渗透系数,建模时如何利用不同钻孔之间的联系还有待进一步研究。

6 结 论

地质渗透系数是水利工程三维渗流场安全分析的重要参数,但由于受多种因素相互作用,其实际值往往难以快速准确地确定。因此,建立快速获取渗透系数“真值”的高精度模型具有实际意义。本研究建立了基于 RF-GWO 的水利工程地质渗透系数智能反演模型,成功解决了 Z 抽水蓄能电站地质渗透系数反演的问题,获得了其天然渗流场整体分布特征。

(1)相较于 CART 和 BP 模型,RF 模型对各钻孔水位值的预测更接近有限元计算值,性能评价指标值也最小,其预测精度和泛化能力更强,可代替有限元正演模型进行渗流计算,提高反演效率,在工程区渗透参数反演中具有很大的应用前景。

(2)GWO 算法比 PSO 和 GA 算法的全局搜索能力强,寻优效率更高,可以有效地搜寻到地质最佳渗透系数;各钻孔水位反演结果合理,最大绝对误差和相对误差分别为 6.94 m 和 0.42%,反演精度满足一般工程要求,计算得到的天然渗流场分布形态也符合一般山体渗流场分布规律,建立的反演模型可以为工程区天然渗流场数值分析奠定良好基础。

参考文献:

[1] 李明超,任秋兵,孔 锐,等. 多维复杂关联因素下的大坝变形动态建模与预测分析[J]. 水利学报,2019,50(6): 687-698.

[2] 杨金孟,赵兰浩,沈振中,等. 某抽水蓄能电站上水库防渗帷幕深度优选研究[J]. 水资源与水工程学报,2021, 32(2):209-216.

[3] 冯亚新,王润英,孙 丹,等. 深厚覆盖层上面板堆石坝渗流敏感性分析[J]. 中国农村水利水电,2023(3):191-198.

[4] 余红玲,王晓玲,王 成,等. 贝叶斯框架下大坝渗流参数反演组合代理模型研究[J]. 水利学报,2022,53(3): 306-315 + 324.

[5] 吕 鹏,王晓玲,吴斌平,等. 基于熵-盲数的贝叶斯渗流参数反演分析研究[J]. 水力发电学报,2019,38(4):108-118.

[6] 许增光,曹 成,李康宏,等. 某抽水蓄能电站上水库局部防渗控分析[J]. 应用力学学报,2018,35(2):417-422 + 459.

[7] 毛昶熙. 渗流计算分析与控制[M]. 北京:水利电力出版社,1990.

[8] 朱伯芳. 渗流场反分析的一种新的数学解法[J]. 水利学报,1994(9):42-46.

[9] 盛金昌,速宝玉,詹美礼. 三维天然渗流场反演分析及工程应用[J]. 岩石力学与工程学报,2003,22(2):203-207.

[10] 段 斌,张 林,何江达,等. 复杂裂隙岩体天然渗流场反演分析[J]. 水力发电学报,2012,31(3):188-193 + 198.

[11] 倪沙沙,迟世春. 基于粒子群支持向量机的高心墙堆石坝渗透系数反演[J]. 岩土工程学报,2017,39(4):727-734.

[12] 唐少龙,熊 威,万小强,等. 基于 GA-BP 的大坝渗透系数多目标反演分析方法[J]. 中国农村水利水电, 2020(9):213-216.

[13] 沈宇扬,沈振中,李佳杰,等. 基于思维进化算法优化 BP 神经网络的渗透参数反演研究[J]. 水电能源科学, 2020,38(2):102-105 + 121.

[14] 徐 丽,沈振中. 基于 ELM-GA 的复杂土石坝渗透系数反演模型及其应用[J]. 水电能源科学,2021,39(9): 86-90.

[15] TAN Jiacheng, XU Liqun, ZHANG Kailai, et al. A biological immune mechanism-based quantum PSO algorithm and its application in back analysis for seepage parameters [J]. Mathematical Problems in Engineering, 2020, 8: 2191079. DOI:10.1155/2020/2191079.

[16] YU Hongling, WANG Xiaoling, REN Bingyu, et al. An efficient Bayesian inversion method for seepage parameters using a data-driven error model and an ensemble of surrogates considering the interactions between prediction performance indicators [J]. Journal of Hydrology, 2022, 604: 127235.

[17] 李雅琦,杨 杰,程 琳,等. 基于 RVM-CS 的工程区地层渗透系数反演分析[J]. 长江科学院院报,2020,37 (11):121-127 + 135.

[18] SHU Yongkang, SHEN Zhenzhong, XU Liqun, et al. Inversion analysis of impervious curtain permeability coefficient using calcium leaching model, extreme learning machine, and optimization algorithms [J]. Applied Sciences, 2022, 12(7): 3272.

[19] 李皓璇,沈振中,张文兵. 基于 PSO-GRNN 的混凝土面板堆石坝渗透系数反演方法及应用[J]. 水电能源科

- 学,2023,41(5):67-70+80.
- [20] BREIMAN L. Random forests[J]. Machine Learning, 2001, 45: 5-32.
- [21] MIRJALILI S, MIRJALILI S M, LEWIS A. Grey wolf optimizer[J]. Advances in Engineering Software, 2014, 69: 46-61.
- [22] DAI Bo, GU Chongshi, ZHAO Erfeng, et al. Statistical model optimized random forest regression model for concrete dam deformation monitoring[J]. Structural Control and Health Monitoring, 2018, 25(6): e2170.
- [23] 李汪繁,丁先,方晶剑. 基于 GWO-RF 的凝汽器真空预测方法[J]. 动力工程学报,2023,43(4):436-442.
- [24] 陈永政. 灰狼优化算法在城市需水预测模型中的应用[J]. 人民黄河,2023,45(2):97-100+126.
- [25] 单正,王元龙,白若忱,等. 基于灰狼优化算法的燃气轮机自抗扰控制参数整定研究[J]. 热能动力工程, 2023,38(4):1-6.
- [26] 刘再涛,王震,贺建军,等. 基于灰狼算法优化的多隐层径向基神经网络铅锌烧结返粉料水分预测[J]. 中南大学学报(自然科学版),2023,54(2):754-764.
- [27] 许凯,雷学文,孟庆山. 渗流反演分析在某水电工程中的应用[J]. 武汉大学学报(工学版),2011,44(1):37-39+43.
- [28] LI Yanlong, YIN Qiaogang, ZHANG Ye, et al. Prediction of long-term maximum settlement deformation of concrete face rockfill dams using hybrid support vector regression optimized with HHO algorithm[J]. Journal of Civil Structural Health Monitoring, 2023, 13(2-3): 371-386.
- [29] 张巍,肖明. 地下工程渗流断层数值模拟的隐式复合材料单元法研究[J]. 岩土工程学报,2005,27(10): 1203-1206.
- [30] 李炎隆,李守义,丁占峰,等. 基于正交试验法的邓肯-张 E-B 模型参数敏感性分析研究[J]. 水利学报, 2013,44(7):873-879.
- [31] HAUPT R L, HAUPT S E. Practical genetic algorithms[M]. Hoboken: John Wiley & Sons, 2003.
- [32] MIRJALILI S, DONG Jinsong, LEWIS A, et al. Nature-inspired optimizers[M]. Berlin/Heidelberg: Springer, 2020.
- [33] 孙超伟,柴军瑞,许增光,等. 金川水电站地下厂房裂隙围岩渗控效应数值模拟与评价[J]. 岩土工程学报, 2016,38(5):786-797.
- [34] XU Zengguang, LIU Yang, HUANG Jian, et al. Performance assessment of the complex seepage control system at the Lu Dila Hydropower Station in China[J]. International Journal of Geomechanics, 2019, 19(3): 05019001.

(上接第 138 页)

- [31] PAERL H W, HUISMAN J. Climate change: a catalyst for global expansion of harmful cyanobacterial blooms[J]. Environmental Microbiology Reports, 2009, 1(1): 27-37.
- [32] 秦伯强,杨桂军,马健荣,等. 太湖蓝藻水华“暴发”的动态特征及其机制[J]. 科学通报,2016,61(7):759-770.
- [33] 练继建,王海军,熊浩. 一种长距离输水渠道消冰、抑藻的装置:202020335394.7[P]. 2021-01-05.
- [34] 杨忠勇,钱门亮,纪道斌,等. 香溪河初冬降温过程中垂向混合结构特征分析[J]. 中国环境科学,2021,41(6):2862-2870.
- [35] 李衍庆,黄廷林,张海涵,等. 水源水库藻类功能群落演替特征及水质评价[J]. 环境科学,2020,41(5):2158-2165.
- [36] ZHANG Yuyang, YU Shuangen, WANG Wenlei, et al. Temperature sensitivity of marine macroalgae for aquaculture in China[J]. Aquaculture, 2023, 567: 739262.
- [37] 沈英嘉,陈德辉. 不同光照周期对铜绿微囊藻和绿色微囊藻生长的影响[J]. 湖泊科学,2004(3):285-288.
- [38] 刘鑫,王超,王沛芳,等. 营养盐比例对硅藻水华优势种小环藻生长和生理的影响[J]. 环境科学研究,2021, 34(5):1196-1204.
- [39] YANG Yuangen, HE Zhenli, LIN Youjian, et al. Temporal and spatial variations of nutrients in the Ten Mile Creek of South Florida, USA and effects on phytoplankton biomass[J]. Journal of Environmental Monitoring, 2008, 10(4): 508-516.
- [40] 刘春光,金相灿,孙凌,等. pH 值对淡水藻类生长和种类变化的影响[J]. 农业环境科学学报,2005(2):294-298.
- [41] 王文超,徐绍峰,徐梦瑶,等. 水体扰动对铜绿微囊藻生长和酶活性的影响[J]. 军事医学,2019,43(6):448-453.
- [42] SONG Yang, ZHANG Linglei, LI Jia, et al. Mechanism of the influence of hydrodynamics on *Microcystis aeruginosa*, a dominant bloom species in reservoirs[J]. Science of the Total Environment, 2018, 636: 230-239.
- [43] 于汇洋,徐志嫻,龙怡静,等. 再生水补水的景观水体优势藻种水华爆发阈值条件的确定[J]. 水资源与水工程学报,2021,32(6):102-108.
- [44] 蒋聘凤,童思陈,黄国鲜,等. 气候变化和人类活动对河流营养盐通量过程影响研究综述[J]. 水资源与水工程学报,2023,34(4):23-34+44.