

DOI: 10.20040/j.cnki.1000-7709.2025.20241840

# 基于 Attention-LSTM 施工期面板堆石坝 坝体沉降预测模型研究

周子玉, 肖亚子, 鄢显昆, 许爱平

(中国电建集团中南勘测设计研究院有限公司, 湖南 长沙 410014)

**摘要:** 构建高精度的坝体沉降预测模型, 对于保障施工期大坝安全与风险管控意义重大。以筑坝高度、降雨、时效等作为施工期坝体沉降变形的影响因子, 引入长短时记忆神经网络 LSTM 算法, 并在其基础上嵌入注意力机制 Attention, 提出了适用于施工期面板堆石坝坝体沉降的预测模型。工程应用表明, Attention-LSTM 模型弥补了 LSTM 在网络层不能动态调整权重系数的缺陷, 提升了模型的计算效率与精度, 具有较好的非线性数据处理能力, 能够更加精准反映施工期监测数据在时间维度上的变化趋势, 相关经验可供类似工程参考。

**关键词:** 面板堆石坝; 施工期; 坝体沉降预测; 长短时记忆神经网络; 注意力机制

中图分类号: TV698.1

文献标志码: A

文章编号: 1000-7709(2025)09-0146-04

## 1 引言

面板堆石坝坝体沉降是影响防渗结构变形甚至开裂的主要原因<sup>[1]</sup>。根据大坝原观监测数据, 利用统计模型、机器学习等方法, 对施工期面板堆石坝坝体沉降构建高精度预测模型, 对于保障施工期大坝安全与风险管控意义重大<sup>[2]</sup>。目前, 通常采用统计模型对堆石坝坝体沉降进行研究, 建模时常以水压、降雨、时效分量等作为运行期堆石坝的影响因子, 而施工期则考虑筑坝高度、降雨、时效分量等。该方法有效提高了模型预测精度, 但均将效应量视为影响因素的线性显示函数, 难以准确表达因子与变量之间的非线性与不确定性, 存在精度不高、模型稳定性较差等问题<sup>[3]</sup>。随着人工智能的发展, 机器学习算法依靠其强大的非线性数据挖掘能力, 逐渐应用于大坝安全监控领域, 并取得了显著成效<sup>[4]</sup>。黄灵芝等<sup>[5]</sup>通过小波去噪对原始数据进行分解, 运用高斯回归 (GPR) 建立堆石坝沉降变形的预测模型; 燕乔等<sup>[6]</sup>运用极限学习机 (ELM), 结合改进粒子群算法 (IPSO) 构建面板堆石坝运行期沉降预测模型; 王亚坤<sup>[7]</sup>运用 LSTM 方法建立高斜心墙堆石坝沉降预测模型。由此可见, 机器学习算法的应用研究取得了较好的成果, 但仍存在模型计算效率低、易陷入局部最优解等问题, 且针对施工期面板

堆石坝的研究成果较少。为此, 融合 LSTM、Attention 的各自特征, 提出基于 Attention-LSTM 施工期堆石坝坝体沉降预测模型, 并以某面板堆石坝为例, 验证了模型的有效性。

## 2 基本原理

### 2.1 长短时记忆神经网络 (LSTM) 模型

长短时记忆神经网络 (LSTM) 是一种改进的循环神经网络, 在网络中增加遗忘门、输入门、输出门, 调节控制三扇门的学习率, 避免了模型在训练中产生梯度消失和爆炸等现象, 相较于循环神经网络 RNN 模型, LSTM 模型可凭借特殊的神经元对长时间跨度的时间序列保持良好记忆, 从而提高信息的利用率与时间序列预测的准确性<sup>[8]</sup>。其神经网络结构见图 1。

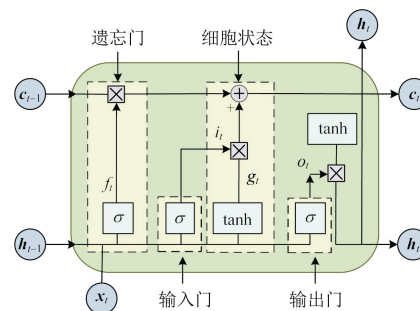


图 1 LSTM 神经网络结构

Fig. 1 LSTM neural network structure

LSTM 信息传递过程如下。

收稿日期: 2024-09-27, 修回日期: 2024-11-05

作者简介: 周子玉 (1998-), 男, 硕士, 研究方向为水工结构工程与安全监测, E-mail: 810998668@qq.com

(1)首先由遗忘门  $f_t$  确定从细胞中丢弃的信息,计算公式为:

$$f_t = \sigma(\mathbf{W}_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

式中,  $\sigma$  为 sigmoid 激活函数;  $\mathbf{W}_f$  为遗忘门的权重矩阵;  $[h_{t-1}, x_t]$  为输入向量;  $b_f$  为遗忘门的偏置项。

(2)输入门确定更新哪些新的信息到单元中,计算公式为:

$$i_t = \sigma(\mathbf{W}_i \cdot [h_{t-1}, x_t] + b_i) \quad (3)$$

$$\tilde{C}_t = \tanh(\mathbf{W}_c \cdot [h_{t-1}, x_t] + b_c) \quad (4)$$

式中,  $i_t, \tilde{C}_t$  分别为输入门、候选值;  $\mathbf{W}, b$  分别为权重和偏置。

(3)细胞更新,需要将旧的细胞状态  $C_{t-1}$  更新为新的细胞状态  $C_t$ ,计算公式为:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (5)$$

(4)最后输出门决定输出哪些信息,计算公式为:

$$o_t = \sigma(\mathbf{W}_o \cdot [h_{t-1}, x_t] + b_o) \quad (6)$$

$$o_t = \sigma(\mathbf{W}_{oh} h_{t-1} + \mathbf{W}_{ox} x_t + b_o) \quad (7)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (8)$$

式中,  $o_t$  为输出门值;  $\mathbf{W}_o$  为输出门的权重矩阵;  $b_o$  为输出门的偏置项。

### 2.2 注意力机制(Attention)模型

Attention 模型通过对 LSTM 层的输出进行加权融合,选择对模型影响较大的因子加权处理,从而高效准确处理时序数据。其结构见图 2。

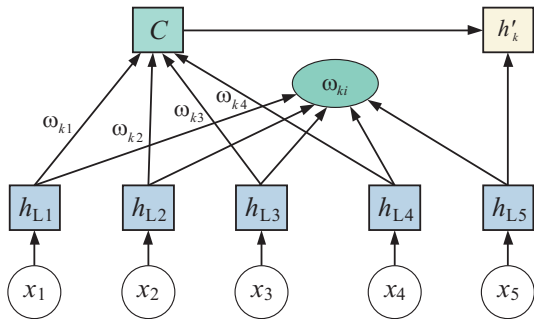


图 2 Attention 机制结构

Fig. 2 Attention mechanism structure

图 2 中  $x_1, x_2, \dots, x_5$  均为输入的影响因子,例如筑坝高度、降雨、蠕变等;  $h_{L1}, h_{L2}, \dots, h_{L5}$  均为输入影响因子的隐藏层状态值;  $\omega_{ki}$  为注意力权重;  $h'_k$  为最终的隐藏层状态值。

对于输入  $x$  的序列中的每个向量  $x_i$ ,注意力权重  $\alpha_i$  为:

$$\alpha_i = \text{softmax}(f(x_i)) = \frac{\exp(f(x_i))}{\sum_{j=1}^m \exp(f(x_j))} \quad (9)$$

注意力层的输出  $\text{attention}_x$  为:

$$\text{attention}_x = \sum_{i=1}^m \alpha_i x_i \quad (10)$$

式中,  $f(x_i)$  为评分函数。

在传统的 LSTM 模型中,由于输入序列的每个时间步都是固定不变的,未考虑不同时间步的重要性差异,可能造成信息过载、解释模糊、计算冗余等问题。通过引入注意力机制,将其嵌入到 LSTM 网络中,增强模型对输入序列中不同因子的关注程度,根据输入因子的重要性动态调整模型的权重,从而提升模型的拟合预测精度与泛化能力。其网络结构见图 3,  $\mathbf{H}_m = [h_{L1} \dots h_{Lm} \dots h_{Lm}]^T$  为 LSTM 层长度  $m$  的输出向量。

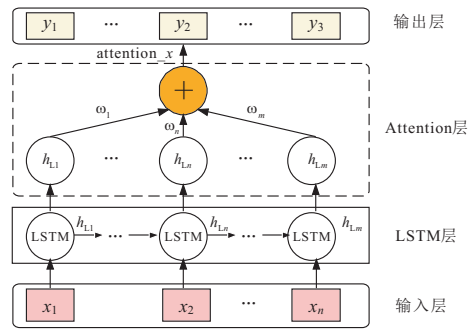


图 3 Attention-LSTM 神经网络结构

Fig. 3 Attention-LSTM neural network structure

## 3 基于 Attention-LSTM 的堆石坝坝体沉降预测模型

### 3.1 模型构建步骤

结合施工期混凝土面板堆石坝坝体沉降变形数据,分析坝体沉降影响因子,引入 LSTM 算法,深度嵌入注意力机制 Attention,构建施工期混凝土面板堆石坝坝体沉降的 Attention-LSTM 预测模型,其建模步骤如下。

**步骤 1** 确定坝体沉降影响因子。根据施工期坝体沉降监测成果分析可知,施工期堆石坝的沉降量随着筑坝高度的提升而增加;同时降雨会增加堆石体的含水率,导致水土压力发生变化与渗流变形,尤其雨季较为明显<sup>[9]</sup>;此外,考虑到坝体及坝基材料流变性,时效分量同样重要。由于施工期上库大坝未蓄水,且温度对堆石体的影响较小,不考虑水位、温度的影响。因此,以坝体填筑高度  $\delta_H$ 、降雨  $\delta_P$ 、坝体坝基蠕变  $\delta_\theta$  等因素作为沉降因子,构建施工期坝体沉降预测模型,其表达式为:

$$\delta = \delta_H + \delta_P + \delta_\theta \quad (11)$$

$$\begin{cases} \delta_H = \sum_{i=1}^3 a_i h_d \\ \delta_P = \sum_{i=1}^5 b_i p_i \\ \delta_\theta = c_1 \theta + c_2 \ln(\theta) \quad \theta = t/100 \end{cases} \quad (12)$$

式中,  $\delta$  为沉降拟合预测值;  $a_i$  为填筑分量系数;  $h_d$  为测点以上填筑增加高度;  $b_i$  为降雨分量系数;  $p_i$  为当地降雨情况;  $c_1, c_2$  均为时效分量系数;  $t$  为起测日至今天的天数。

**步骤 2** 数据标准化处理。将面板堆石坝沉降的影响因子设置为  $X = (X_1, X_2, \dots, X_m)$ , 坝体沉降量设为  $Y = (Y_1, Y_2, \dots, Y_m)$ , 进行归一化处理, 计算公式为:

$$X^* = (X - X_{\min}) / (X_{\max} - X_{\min}) \quad (13)$$

式中,  $X^*$  为归一化后的任一测值;  $X$  为监测序列中任一测值;  $X_{\min}$  为单个监测序列中的最小值;  $X_{\max}$  为单个监测序列中的最大值。

**步骤 3** 设置评价函数。采用平均平方误差 (MSE) 量化模型计算过程中实测值与拟合值之间的差异, 将其作为 Attention-LSTM 模型的损失函数, 通过模型不断迭代更新, 使损失函数最小化, 从而得到本次训练计算的结果。

**步骤 4** 计算与输出。计算获取面板堆石坝坝体沉降数据序列的拟合预测数据集, 并将其进行反归一化处理, 输出相应的拟合预测数据  $\hat{y} = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_m)$ , 并与实测值进行对比分析, 构建 Attention-LSTM 预测模型。

### 3.2 模型评价指标

为检验模型的拟合预测效果, 选用可决系数 ( $R^2$ )、均方根误差 ( $R_{RMSE}$ )、平均绝对误差 ( $M_{MAE}$ ) 进行量化表达。

## 4 工程应用

### 4.1 工程概况

河南某电站大坝为混凝土面板堆石坝, 坝顶高程 351.00 m, 顶宽 10.00 m, 最大坝高 128.20 m, 上游坝坡坡比为 1 : 1.40, 下游坝坡坡比为 1 : 1.35, 主副坝坝顶中心线总长 1 238.04 m, 电站总装机容量为 1 000 MW。坝体沉降采用水管式沉降仪进行监测, VSA23-5 测点位于坝右 0+003.350 坝纵 0+000.000EL. 266.514 0 m, VSA32-4 测点位于坝右 0+122.660 坝纵 0+000.000EL. 295.509 0 m, 均位于坝轴线处, 具有较好的代表性, 故作为典型测点进行建模分析。

### 4.2 模型网络参数设置

为验证 Attention-LSTM 模型的性能, 将其与 BP、SVR、LSTM 模型进行对比, 并统一各模型的训练集、测试集。设置 BP、LSTM、Attention-LSTM 模型学习率为 0.001、最大迭代次数为 1 000、训练步长为 10 等。构建 SVR 模型时, 将惩罚因子  $C$ 、核函数  $\sigma$  分别设置为 20、0.3。将输入层的输入神经单元数 feature Dimension 数据集代入 LSTM 网络层中, 经分析, 选取坝体填筑高度、降雨、坝体及坝基的蠕变等因素作为输入神经单元数。LSTM、Attention-LSTM 模型的训练过程中, 为提升模型机器学习效率, 占用更少的资源, 采用 Adam 优化器进行优化, 排序函数 Shuffle 采用 every-epoch。同时, 为避免模型在训练过程中出现过拟合现象, 在模型中引入丢弃率 Dropout 层搭建, 通过在训练过程中按照一定概率随机将神经网络单元从网络结构中移除, 从而提高模型的泛化能力, 设置丢弃率 Dropout 为 0.2。

### 4.3 坝体沉降变形预测

为验证模型的可行性与有效性, 比较不同模型之间的拟合预测精度、稳定性及计算效率, 选用水管式沉降仪 VSA23-5 测点 2022 年 4 月至 2024 年 1 月的沉降数据、VSA32-4 测点 2022 年 7 月至 2024 年 1 月的沉降数据, 将数据集进行分割, 75% 为训练集、25% 为测试集, 即 VSA23-5 测点 2022 年 4 月~2023 年 7 月数据为训练样本, 2023 年 8 月~2024 年 1 月数据为测试样本, VSA32-4 测点 2022 年 4 月~2023 年 9 月数据为训练样本, 2023 年 10 月~2024 年 1 月数据为测试样本, 并以相应测点以上填筑增加高度、降雨量、时效量、沉降实测值作为输入, 模型计算的沉降拟合预测结果作为输出, 并将其与 BP、SVR、LSTM 模型进行对比分析。

经计算, 各模型评价指标见表 1, 拟合预测对比见图 4, 残差箱型分布见图 5。通过图 4、5 和表

表 1 各模型拟合预测性能对比

Tab. 1 Comparison of fitting prediction performance of each model

测点	模型	$R^2$	$R_{RMSE}$	$M_{MAE}$	时间
			/mm	/mm	/s
VSA23-5	BP	0.953 6	42.697 9	20.271 9	68
	SVR	0.980 9	39.944 6	14.848 7	87
	LSTM	0.988 7	37.312 1	9.079 3	72
	Attention-LSTM	0.994 2	36.456 7	6.059 0	51
VSA32-4	BP	0.904 2	46.457 7	19.839 3	65
	SVR	0.938 8	43.290 7	15.465 5	82
	LSTM	0.975 8	40.031 8	9.971 7	69
	Attention-LSTM	0.996 5	37.605 4	4.362 5	47

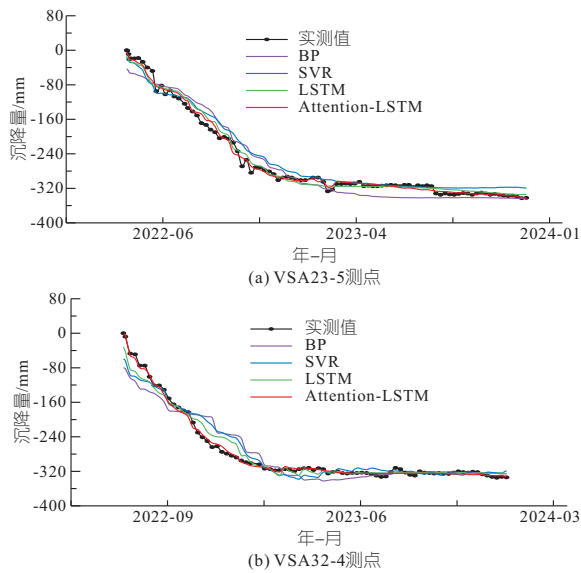


图 4 各模型沉降拟合预测值与实测值

Fig. 4 The settlement fitting predicted value and measured value of each model

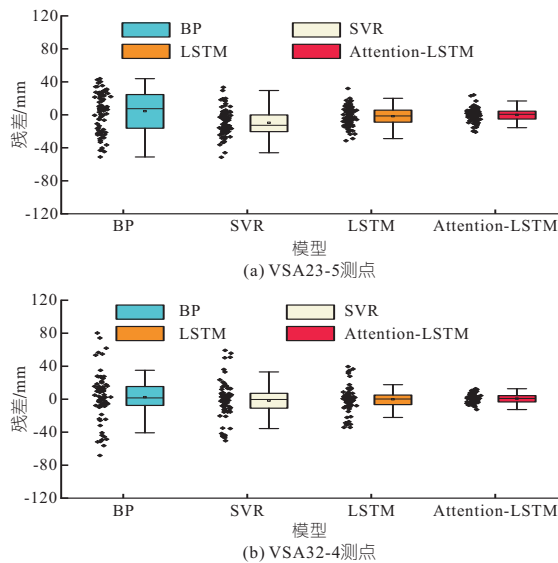


图 5 各模型拟合预测残差箱型分布

Fig. 5 The box-shaped distribution of the fitting prediction residuals for each model

1 可知,Attention-LSTM 模型各评价指标均优于其他模型, $R^2$  在 0.99 以上,曲线贴合程度最接近实测值,拟合预测效果最佳,残差分布更为稳定。

施工期沉降变形机理复杂,影响因素众多,坝体沉降变形与影响因子之间存在较强非线性与不确定性。随着大坝的填筑,坝体沉降值前期增加明显,后期趋于平稳,传统的 BP、SVR 机器学习模型,在沉降变化较大时段,拟合效果欠佳,误差较大,而 LSTM 模型凭借其具有较强的非线性数据处理能力,拟合预测精度提升显著,曲线更加贴合真实值,均方根误差  $R_{RMSE}$ 、平均绝对误差  $M_{MAE}$  分别降低约 10.25%、45.95%。Attention-LSTM 模型因其融入 Attention 机制,能够高效、

准确提取影响坝体沉降的重要信息,自适应调整权重系数,更好挖掘堆石坝坝体沉降变形与筑坝高度、降雨、蠕变等影响因子之间的潜在联系,建模时间下降明显,预测性能得到进一步提升, $R_{RMSE}$ 、 $M_{MAE}$  分别降低约 4.20%、44.76%。

### 5 结论

a. 嵌入 Attention 机制的 LSTM 模型能够更好挖掘堆石坝坝体沉降变形与筑坝高度、降雨、蠕变等影响因子之间的潜在联系,有效处理数据集中的非线性特征,并提升计算效率。

b. Attention-LSTM 模型较 BP、SVR、LSTM 模型在拟合预测精度上提升显著,曲线更贴合真实值,残差更为稳定,能更好地描述施工期堆石坝坝体沉降变形规律与特征。

### 参考文献:

[1] 温立峰,李炎隆,刘云贺,等.考虑门限效应的面板堆石坝变形特性改进支持向量机预测模型研究[J].应用基础与工程科学学报,2023,31(4):876-893.

[2] 王晓玲,梁羽翎,王佳俊,等.耦合注意力机制大坝变形改进 LSTM 序列到序列预测模型[J].天津大学学报(自然科学与工程技术版),2023,56(7):702-712.

[3] 林川,桂星煜,朱律运,等.基于因子融合的混凝土面板堆石坝变形预测模型[J].水力发电学报,2023,42(10):139-152.

[4] 隋旭鹏,王少伟,朱圣辉,等.不同机器学习模型对混凝土坝位移预测性能的影响[J].水电能源科学,2022,40(1):107-111.

[5] 黄灵芝,陈思琦,李成宇,等.基于小波去噪的高斯过程回归模型在面板堆石坝沉降预测中的应用研究[J].水资源与水工程学报,2023,34(3):144-150.

[6] 燕乔,高名杨,梁明浩,等.改进粒子群-极限学习机模型在面板堆石坝运行期沉降预测中的应用[J].水电能源科学,2021,39(10):110-113.

[7] 王亚坤,傅志敏,苏正洋.基于长短期记忆网络的黄河小浪底水库高斜心墙堆石坝沉降预测模型及其预测精度分析[J].水电能源科学,2022,40(2):110-113.

[8] LIU W J, PAN J W, REN Y S, et al. Coupling prediction model for long-term displacements of arch dams based on long short-term memory network[J]. Structural control and health monitoring, 2020, 27(7): e2548.

[9] 梁彬彬.高心墙堆石坝施工期沉降监控模型研究[D].南昌:南昌工程学院,2019.

- tions; A case study in Anhui Province, China[J]. Hydrogeology journal, 2021, 29(5): 1831-1855.
- [6] 吴忠康, 陈文理. 乌江构皮滩水利枢纽下坝址岩溶发育特征及工程地质条件[J]. 中国岩溶, 1994, 13(2): 141-152.
- [7] YANG J M, ZHAO L H, SHEN Z Z, et al. An efficient procedure for optimization design of anti-seepage curtains; A case study[J]. Bulletin of engineering geology and the environment, 2021, 80(3): 2671-2685.
- [8] 任智丽, 路明, 孙小双. 会仙湿地岩溶地下水数值模拟[J]. 南水北调与水利科技(中英文), 2020, 18(5): 157-164.
- [9] 赵瑞, 张强, 许模, 等. 基于数值模拟的复杂岩溶库区渗漏研究[J]. 南水北调与水利科技, 2016, 14(3): 150-155.
- [10] 冯树荣, 蒋中明. 工程渗流理论研究与实践[M]. 北京: 中国水利水电出版社, 2018.
- [11] 纪伟, 赵坚, 沈振中, 等. 水布垭水利枢纽岩溶体内防渗帷幕优化布置研究[J]. 水电能源科学, 2005, 23(1): 36-39, 91.
- [12] 沈振中, 岑威钧, 徐力群, 等. 工程渗流分析与控制[M]. 北京: 科学出版社, 2020.
- [13] 王硕, 沈振中, 徐力群, 等. 某抽水蓄能电站下水库混凝土面板坝防渗系统优选研究[J]. 水电能源科学, 2023, 41(2): 90-94.

## Optimization of Seepage Control Scheme of Lower Reservoir of Pumped Storage Power Station in Karst Area

MA Ze-kai<sup>1a,2</sup>, SHEN Zhen-zhong<sup>1a,2,3</sup>, LI Dong-ze<sup>1a,2</sup>, LIU Yue-chi<sup>4</sup>, ZHANG Hong-wei<sup>1a,1b,2,3</sup>

(1a. College of Water Conservancy and Hydropower Engineering; 1b. College of Civil and Transportation Engineering, Hohai University, Nanjing 210098, China; 2. State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210000, China; 3. Key Laboratory of Transport Industry of Comprehensive Transportation Theory (Nanjing Modern Multimodal Transportation Laboratory), Ministry of Transport, Nanjing 211100, China; 4. School of Electrical and Power Engineering, Hohai University, Nanjing 211100, China)

**Abstract:** Aiming at the complex karst environment of the lower reservoir of a pumped storage power station, a three-dimensional finite element model for seepage analysis was established to simulate the main buildings and karst passages in the reservoir area. "The depth of seepage control curtain at the base of the dam is 0.5 times of the pre-dam head, the depth of seepage control curtain on both sides of the dam is 3 m below the 3 Lu line, the length of seepage control curtain on the right side of the dam is 200 m, and a single-row curtain is set up" is used as the preliminary seepage control scheme. The seepage field, infiltration slope and infiltration volume of the reservoir area were calculated. The preliminary seepage control scheme met the specification requirements. But the seepage rate and infiltration slope were close to the critical value. The four indexes of curtain depth on both sides of the bank, curtain depth at the base of the dam, length of curtain on the right bank, and double-row curtain were changed to optimize seepage control scheme. The impact of the changes of the indexes on the infiltration volume of the reservoir area was investigated so that the optimization of the seepage control scheme was put forward. The analysis results show that on the basis of the preliminary scheme, the double-row curtain is set up at the dam base, the infiltration flow at the dam base is reduced by 267.3 m<sup>3</sup>/d, and the effect of seepage control in the reservoir area is remarkable.

**Key words:** karst area; pumped storage power station; gravity dam; seepage control system; anti-seepage curtain; permeation gradient

\*\*\*\*\*

(上接第 149 页)

## Research on Settlement Prediction Model of Face Rockfill Dam During Construction Period Based on Attention-LSTM

ZHOU Zi-yu, XIAO Ya-zi, WU Yu-kun, XU Ai-ping

(PowerChina Zhongnan Engineering Corporation Limited, Changsha 410014, China)

**Abstract:** Constructing a high-precision dam settlement prediction model is of great significance for ensuring the safety and risk control of dam during the construction period. Taking dam height, rainfall and aging as the influencing factors of dam settlement deformation during construction period, the long-term and short-term memory neural network LSTM algorithm is introduced, and the attention mechanism is embedded. Thus, a prediction model suitable for dam settlement of concrete face rockfill dam during construction period is proposed. The engineering application shows that the attention-LSTM model makes up for the defect that the LSTM cannot dynamically adjust the weight coefficient at the network layer, improves the computational efficiency and accuracy of the model, and has better nonlinear data processing ability, which can more accurately reflect the change trend of monitoring data in the time dimension during the construction period. The relevant experience can be used as a reference for similar projects.

**Key words:** face rockfill dam; construction period; dam settlement prediction; long short-term memory neural network; attention mechanism