



基于智能优化算法和Bi-LSTM神经网络的盾构隧道刀盘扭矩智能预测方法

孔凡超¹, 刘懿晗², 刘士吉¹, 路德春^{2*}, 杜修力²

1. 华北电力大学水利与水电工程学院, 北京 102206

2. 北京工业大学岩土与地下工程研究所, 北京 100124

* E-mail: dechun@bjut.edu.cn

收稿日期: 2024-06-11; 接受日期: 2024-10-30; 网络版发表日期: 2025-01-07

国家自然科学基金(批准号: 52025084, 52408420)、国家重点研发计划(编号: 2022YFC3800901)、北京市自然科学基金(批准号: 8244058)和中央高校基本科研业务费专项(编号: 2024MS066)资助项目

摘要 智能预测合理的刀盘扭矩可有效降低盾构隧道开挖对地层的扰动并减小刀盘磨损, 对于盾构安全高效施工具有重要意义. 本文基于智能优化算法和深度学习神经网络, 提出了盾构隧道刀盘扭矩时序智能预测方法. 采用Bi-LSTM神经网络作为深度学习预测模型, 分别联合粒子群优化算法(PSO)、遗传优化算法(GA)、饥饿游戏搜索算法(HGS)和麻雀搜索算法(SSA)智能确定Bi-LSTM神经网络的3个超参数, 即学习率、神经元数量和丢弃率, 进而获得了盾构隧道刀盘扭矩智能预测方法. 智能模型预测值与实测值的均方根误差作为4种优化算法的适应度函数, 数据预处理采用归一化方法降低不同参数间的量级差异影响. 智能预测模型的输入考虑了15个参数对刀盘扭矩的影响, 包括1个隧道几何参数、6个地质力学参数和8个盾构掘进参数. 结合盾构隧道现场实测数据, 从精度、效率和稳定性3个角度评价4种优化算法寻优Bi-LSTM神经网络超参数的性能, 结果表明相比于其他3种优化算法, HGS方法具有较好的稳定性、较快的效率和较高的计算精度, 是确定Bi-LSTM超参数的适用性方法.

关键词 Bi-LSTM神经网络, 超参数, 启发式智能优化算法, 刀盘扭矩, 性能分析

1 引言

随着交通强国战略目标的明确提出, 我国综合交通基础设施进入了建设黄金期, 截至2023年底, 我国城市轨道交通运营总里程约10165 km. 盾构法施工具有安全、高效、机械化程度高、对城市地面交通干扰小等优势, 是目前大部分城市轨道交通的首选施工方法^[1]. 刀盘扭矩是盾构机掘进过程中刀盘切削土体时

需要刀盘驱动系统提供的作用力^[2], 合理的刀盘扭矩可有效减小刀盘磨损, 提高盾构机掘进速度, 同时降低对周围地层的扰动^[3]. 盾构隧道开挖涉及多系统, 包括围岩土体、临近结构、施工人员和盾构机. 岩土体具有非连续性和不确定性的特点, 同时伴有地下水、断层等不良地质条件, 使得盾构机处于不完全确定的施工环境中, 需要不断调整掘进参数, 优化施工方案^[4]; 随着地下工程的蓬勃发展, 新建盾构隧道近接既

引用格式: 孔凡超, 刘懿晗, 刘士吉, 等. 基于智能优化算法和Bi-LSTM神经网络的盾构隧道刀盘扭矩智能预测方法. 中国科学: 技术科学, 2025, 55: 170–186
Kong F C, Liu Y H, Liu S J, et al. Intelligent prediction method for cutterhead torque of shield tunnel based on intelligent optimization algorithm and Bi-LSTM neural network (in Chinese). Sci Sin Tech, 2025, 55: 170–186, doi: [10.1360/SST-2024-0145](https://doi.org/10.1360/SST-2024-0145)

有结构的工程逐渐增多, 施工环境的复杂性对掘进参数的选取提出了更高的要求^[5~7]; 盾构工程施工工序繁多, 需要人与机器协同工作, 操作人员的能力、经验和判断各不相同, 又在一定程度上增加了盾构刀盘扭矩的不确定性。研究学者以往大多采用经验法、模型试验法及现场测试法研究盾构掘进过程中刀盘扭矩变化规律及定量分析各类因素对刀盘扭矩的影响机理。卢浩等人^[8]提出了考虑刀具贯入度和土仓压力的刀盘扭矩计算模型, 通过与现场数据的对比, 验证了经验公式的合理性。胡光静等人^[9]建立了1:10盾构隧道开挖模型试验, 分析了不同土仓出土率以及刀盘开口率对刀盘扭矩的影响机理。刘明月等人^[10]基于盾构隧道掘进现场实测数据, 给出了空推、硬岩掘进以及软岩掘进条件下的盾构机扭矩统计值。上述传统方法是揭示刀盘扭矩变化机理的适用性方法, 然而影响刀盘扭矩的因素众多, 传统方法很难综合考虑地质力学参数、隧道几何参数及其他盾构掘进参数对刀盘扭矩的耦合影响。人工智能方法充分利用隧道工程大数据, 综合考虑盾构掘进过程中各类复杂因素的耦合影响, 可建立高维参数的非线性关系, 科学合理地对复杂条件下盾构刀盘扭矩进行智能预测与优化决策。

近些年, 大数据人工智能技术的涌现, 推动了盾构隧道工程建设向数字化、智能化方向发展^[11,12]。基于隧道工程大数据库, 采用智能方法对盾构工程中地质条件、施工环境、掘进参数等因素进行汇总, 进而解析其中的关联性, 为盾构隧道设计、施工与运维提供理论指导^[13,14]。盾构刀盘扭矩的智能预测方法包括机器学习方法以及深度学习方法。机器学习方法采用离散数据组表示连续的开挖过程, 每个数据组表示这一开挖环的所有信息, 不考虑盾构各环之间的连续性影响。Kong等人^[15]采用随机森林方法预测了盾构刀盘扭矩; Yu等人^[16]考虑了推力与刀盘扭矩的相关性和差异性, 设计了多通道解耦神经网络模型预测推力和刀盘扭矩; 尹志清等人^[17]基于符号回归算法, 分析了推力、土仓压力以及贯入度变化及换刀对刀盘扭矩的影响规律。由于盾构掘进的连续性, 盾构隧道的掘进参数是随时间连续变化的时序数据, 这表明当前开挖环的掘进参数受邻近已开挖隧道环的影响^[18~20]。为了克服传统的非时序机器学习方法无法考虑掘进参数时序特性影响, 时序深度学习方法可用于建立更合理的预测模型。在循环神经网络(RNN)的基础上, 长短期记忆神

经网络(LSTM)添加了门结构与细胞状态, 减小了时序网络中梯度异常的可能性。Xu等人^[21]采用了LSTM方法预测了盾构隧道的掘进速率以及扭矩, 并分析了不同网络参数对预测结果的影响规律。Lin等人^[22]基于LSTM和GRU算法预测了盾构机掘进参数, 结果表明时序深度学习神经网络的预测性能优于机器学习方法。Zhang等人^[23]对比了LSTM神经网络和随机森林的预测性能, 结果表明LSTM在预测盾构运行参数的精度和泛化能力方面均优于随机森林。Qin等人^[24]提出了卷积神经网络与LSTM相结合的预测模型, 提出的方法比机器学习方法XGBoost和随机森林在预测刀盘扭矩方面有较高的精度。相比于LSTM网络, 通过对时序数据的双向分析, 双向长短期记忆神经网络(Bi-LSTM)可以更加合理地捕获盾构掘进参数的时间序列特征。Wang等人^[25]利用Bi-LSTM准确预测了泥水平衡盾构机的推进速度与扭矩, Bi-LSTM相较于LSTM可以利用更少的参数实现更加准确的掘进参数预测。综上, 采用时序型深度学习算法可合理地预测盾构掘进参数。

深度学习神经网络的超参数是在训练模型前人工设置的参数, 选择合理的超参数是建立高精度预测模型的关键因素。Bi-LSTM网络中的超参数包括学习率、神经元数量和丢弃率等, 基于经验确定的深度学习超参数难以保证神经网络达到最佳预测性能。启发式优化算法能够准确合理地确定全局最优解, 这使得启发式优化方法成为了确定深度学习超参数的优势方法。常用的启发式优化算法包括粒子群优化算法(PSO)、遗传优化算法(GA)与麻雀搜索算法(SSA)等^[26~28]。Ye等人^[29]采用PSO优化了LSTM与BP神经网络中的神经元数量, 结果表明神经元的数量对模型的预测性能有很大影响; Mahmoodzadeh等人^[30]采用灰狼搜索算法确定了TBM掘进速度预测模型的迭代次数和神经元数量的取值; Lin等人^[31]采用PSO优化了LSTM与GRU网络的学习率, 提高了混合模型预测盾构机掘进速度的精度。饥饿游戏搜索算法(HGS)基于群居动物的饥饿和觅食行为进行目标优化, 强调了动物根据感知信息和计算规则与环境交互, 优先选择和追求食物来源, 这种机制使得算法在寻优方面表现出色^[32], 可为深度学习网络(如Bi-LSTM)提供更加合理的超参数组合。Shaiba等人^[33]采用HGS优化超参数实现了网络钓鱼攻击的检测与分类。Zhang等人^[34]联合

HGS与Bi-LSTM实现了致密砂岩储层的孔隙度、黏土含量等参数的反演预测. HGS在上述研究中有效地提高了模型预测性能, 然而目前研究尚未将HGS应用于盾构掘进参数的模型优化. 在上述研究中, 启发式智能优化算法基于特定机制自动搜索参数空间中的超参数, 无需手动调整或设置参数, 降低了时间成本的同时提高了预测精度. 上述研究均是采用一种优化算法结合深度学习网络预测掘进参数, 然而在确定深度学习方法的超参数时, 采用不同优化算法的预测效率、精度和稳定性差异很大, 并未系统深入研究不同优化算法的预测性能; 同时目前鲜有联合智能优化算法和深度学习方法的刀盘扭矩预测研究.

考虑盾构掘进参数、隧道几何参数以及围岩力学参数对刀盘扭矩的影响, 基于Bi-LSTM神经网络和启发式优化算法, 本文建立盾构隧道开挖刀盘扭矩智能预测框架. 采用粒子群优化算法、遗传优化算法、饥饿游戏搜索算法和麻雀搜索算法4种优化算法对Bi-LSTM的3个超参数即学习率、神经元数量和丢弃率进行智能优化决策, 提升深度学习神经网络的预测性能. 结合盾构隧道工程现场实测数据, 研究4种优化算法对提出的智能预测方法精度、效率和稳定性的影响规律.

2 盾构机刀盘扭矩智能预测方法框架

本研究分别采用PSO, GA, SSA和HGS优化Bi-LSTM的3个超参数, 即学习率、神经元数量和丢弃率, 进而建立联合智能优化算法和Bi-LSTM的盾构隧道刀盘扭矩混合智能预测方法.

2.1 Bi-LSTM神经网络计算原理

Bi-LSTM是LSTM的一种变体形式. LSTM通过单元状态和门控机制实现了时序特征的传输, 如图1所示. 门结构实现了时序特征的添加和遗忘, 分为遗忘门(f_t)、输入门(i_t)和输出门(o_t). 输入门控制每个时间步向网络传递的信息, 其公式为

$$i_t = \sigma(W_i x_t + W_i h_{t-1} + b_i), \quad (1)$$

式中, i_t 是时间 t 的输入门, σ 是激活函数, W_i 和 b_i 是输入门中的权重和偏置, x_t 是时间 t 的输入, h_{t-1} 是时间 $t-1$ 的隐藏状态.

遗忘门读取先前的输出和当前的输入, 决定了信

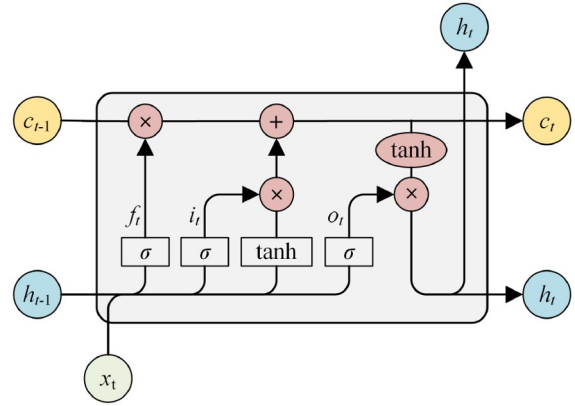


图1 (网络版彩图) Bi-LSTM细胞结构

Figure 1 (Color online) Cell structures of Bi-LSTM.

息的保留与删除. 遗忘门表示为

$$f_t = \sigma(W_f x_t + W_f h_{t-1} + b_f), \quad (2)$$

式中, f_t 是时间 t 的遗忘门, W_f 和 b_f 是遗忘门中的权重和偏置.

输出门控制输出信息, 表示为

$$o_t = \sigma(W_o x_t + W_o h_{t-1} + b_o), \quad (3)$$

式中, o_t 是时间 t 的输出门, W_o 和 b_o 是输出门的权重和偏置.

细胞状态代表了网络对数据的长期记忆, 并将传递到后续的细胞结构中. 细胞状态表示为

$$c_t = f_t c_{t-1} + i_t \tanh(W_c x_t + W_c h_{t-1} + b_c), \quad (4)$$

式中, c_t 是时间 t 的细胞状态, W_c 和 b_c 分别是输入数据单元状态中的权重和偏差.

隐藏状态是LSTM提取的时序特征, 并将作为短期记忆传递到后续的细胞结构中. 隐藏状态 h_t 的表达式为

$$h_t = o_t \tanh(c_t). \quad (5)$$

Bi-LSTM由前向LSTM和后向LSTM组成, 可以拼接正向和反向的时序信息. 时序信息的双向传播使Bi-LSTM可更准确地提取时序信息特征. Bi-LSTM的计算框架如图2所示.

2.2 Bi-LSTM神经网络中的超参数

神经网络预测模型中的超参数分为结构超参数和训练超参数. 在本文中优化的Bi-LSTM神经网络结构

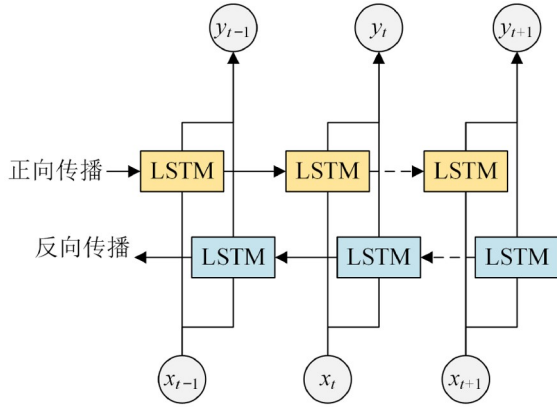


图2 (网络版彩图) Bi-LSTM网络结构

Figure 2 (Color online) Network structures of Bi-LSTM.

超参数为神经元数量与丢弃率, 优化的训练超参数是学习率。

神经元是LSTM神经网络的计算单元, 接收输入并产生输出。由神经元组成的隐含层是所有神经网络进行特征提取与分析的基础。神经元数量过少会导致网络无法合理表达样本数据的一般性规律, 进而严重影响模型的泛化能力。网络中的神经元数量过多, 训练过程数据的拟合程度较高, 但会使网络过于重视训练样本独有的特殊性规律, 从而无法适用于新的数据样本, 即过拟合现象。因此需要确定合理的神经元数量。

丢弃机制是一种常见的神经网络正则化方法, 即每一次迭代过程中随机丢弃一部分神经元的训练结果, 如图3所示。神经元的丢弃和保留是概率随机确定的, 因此训练过程中每次迭代神经网络所能利用的神经元组成均不相同。丢弃机制可以避免模型在训练集时对某些节点间的关系产生过强的相互依赖, 提高模型的鲁棒性与泛化能力, 丢弃机制对提高神经网络的性能具有重要意义。

神经网络的训练过程是网络参数根据预测值与真实值间的差异不断向最优解移动的过程。移动的步长就是训练网络所采用的学习率, 因此学习率控制着网络的收敛进度。采用较大的学习率可以加快网络的收敛速度, 但过大的学习率可能导致参数在最优解附近振荡甚至错过最优解。相反, 学习率过小会导致网络的收敛速度过慢或陷入局部最优。确定合理的学习率可以训练出效果最优的深度学习模型。

2.3 基于智能优化算法的超参数优化

采用4种启发式智能优化算法, 即GA, PSO, SSA和HGS确定Bi-LSTM模型的3个超参数。模型预测值与实测值的均方根误差(RMSE)作为适应度函数, 用于确定最优的超参数组合。

2.3.1 粒子群优化算法

PSO是一种群体智能优化算法, 如图4所示, 在参数空间中随机初始化一组粒子, 通过特定的规则迭代来找到最优的超参数。第 t 次迭代中第 i 个粒子的属性可以用速度 V_i^t 和参数空间位置 X_i^t 来表示, 其中 X_i^t 的坐标表示超参数的值, 即Bi-LSTM神经网络中的学习率、神经元数量和丢弃率。PSO优化算法的主要过程是计算适应度函数并找到当前迭代步的全局最佳适应度函数位置。第 i 个粒子的 V_i^{t+1} 和 X_i^{t+1} 的表达式为

$$V_i^{t+1} = w_t \cdot V_i^t + c_1 r_1 (P_i^t - X_i^t) + c_2 r_2 (P_g^t - X_i^t), \quad (6)$$

$$X_i^{t+1} = X_i^t + V_i^{t+1}, \quad (7)$$

$$w_t = w_u - \frac{t(w_u - w_d)}{n_{\max}}, \quad (8)$$

式中, c_1 是个体认知系数, 表示个体粒子当前与历史最优位置对优化结果的影响; c_2 是群体认知系数, 表示当前与历史上的全局最优位置对粒子优化结果的影响。认知系数越大表示影响权重越大, 本文中均取2。 r_1 和 r_2 是[0, 1]范围内的随机数; P_i^t 表示第 i 个粒子的个体最优位置; P_g^t 表示全局最优位置; w 表示惯性权重; w_u 表示最大惯性权重, 本文取0.9; w_d 表示最小惯性权重, 本文取0.4; $n_{\max}$ 表示最大迭代次数。

2.3.2 遗传优化算法

GA是一种模拟自然选择和遗传进化的群体智能优化算法, 优化过程如图5所示。算法的初始种群是随机生成的, 种群中的每个个体拥有不同的Bi-LSTM学习率、神经元数量和丢弃率。对这3个参数进行二进制编码就构成了种群中每个个体的基因。

遗传优化算法每次迭代计算不同超参数组合下的适应度值, 选择适应度值较小的超参数生成新样本, 淘汰适应度值较大的超参数。本研究中遗传优化算法每轮迭代随机选择20%种群比较适应度值, 从中选择适

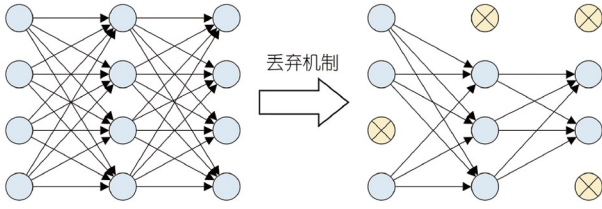


图 3 (网络版彩图) Bi-LSTM神经网络的丢弃机制

Figure 3 (Color online) Discard mechanism of Bi-LSTM neural network.

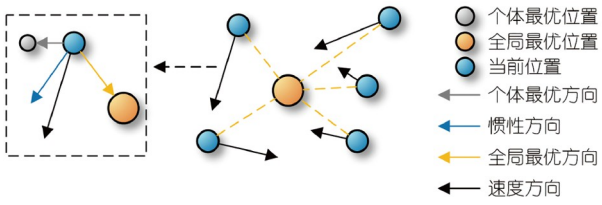


图 4 (网络版彩图) PSO算法计算原理

Figure 4 (Color online) Calculation principle of PSO algorithm.

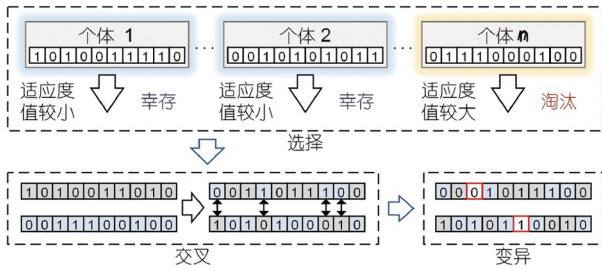


图 5 (网络版彩图) GA算法计算原理

Figure 5 (Color online) Calculation principle of GA algorithm.

应度函数最小的两个个体作为超参数父代。重复上述过程直到所选择的父代总数等于初始种群数量。在交叉过程中, 两组超参数父代个体交换代码片段生成新的超参数, 直到种群个数等于初始种群个数。本文中代码片段的交叉概率设置为0.95。在交叉过程后, 个体的基因会在随机位置发生变异, 从而产生新的基因, 变异避免算法陷入局部最优, 本文的变异概率为0.025。完成选择、交叉和变异过程的群体继续进行下一次迭代, 直到达到终止条件。

2.3.3 饥饿游戏搜索算法

HGS算法是一种基于群居动物的饥饿和觅食行为的优化算法, 流程如图6所示。HGS分为靠近食物和饥

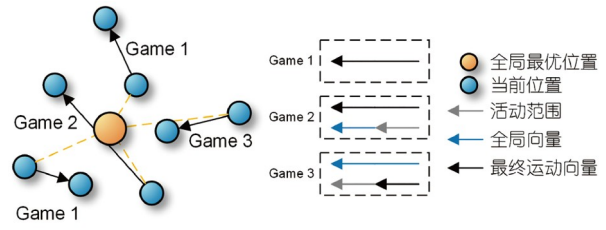


图 6 (网络版彩图) HGS算法计算原理

Figure 6 (Color online) Calculation principle of HGS algorithm.

饿角色两部分。在靠近食物时, 第 $(t+1)$ 次迭代中动物个体的位置更新表达式为

$$X(t+1) = \begin{cases} \text{Game 1: } X(t) \cdot [1 + \text{randn}(1)], & r_1 < l, \\ \text{Game 2: } W_1 \cdot X_b + R \cdot W_2 \cdot |X_b - X(t)|, & r_1 > l, r_2 > E, \\ \text{Game 3: } W_1 \cdot X_b - R \cdot W_2 \cdot |X_b - X(t)|, & r_1 > l, r_2 < E, \end{cases} \quad (9)$$

$$E = \text{sech}(|F(i) - BF|), \quad (10)$$

$$R = 4 \times \left(1 - \frac{t}{n_{\max}}\right) \times \text{rand} - 2 \times \left(1 - \frac{t}{n_{\max}}\right), \quad (11)$$

式中, r_1 和 r_2 是 $[0, 1]$ 范围内的随机数, $\text{randn}(1)$ 是满足标准正态分布的随机数, W_1 和 W_2 表示饥饿权重, X_b 表示全局最优位置; $X(t)$ 表示当前个体位置, 即Bi-LSTM神经网络中的学习率、神经元数量和丢弃率; l 为常数, 本文中取0.08; $F(i)$ 是第 i 个动物的适应度值, BF 是当前最优适应度值, rand 是 $[0, 1]$ 范围内的随机数, $n_{\max}$ 表示最大迭代次数。

饥饿游戏搜索算法通过Game 1模拟动物独立觅食, Game 2和Game 3模拟动物群体合作觅食。动物接近食物的行为受个体的饥饿特征影响, 表示为

$$W_1^i = \begin{cases} \text{hungry}(i) \cdot \frac{N}{SHungry} \times r_4, & r_3 < 1, \\ 1, & r_3 > 1, \end{cases} \quad (12)$$

$$W_2^i = (1 - \exp(-|\text{hungry}(i) - SHungry|)) \times r_5 \times 2, \quad (13)$$

式中, r_3 , r_4 和 r_5 均为 $[0, 1]$ 范围内的随机数, N 表示种群总数, $\text{hungry}(i)$ 表示第 i 个个体的饥饿程度, $SHungry$ 表示所有个体饥饿程度的总和, $\text{hungry}(i)$ 的表达式为

$$\text{hungry}(i) = \begin{cases} 0, & \text{AllF}(i) = BF, \\ \text{hungry}(i) + H, & \text{AllF}(i) \neq BF, \end{cases} \quad (14)$$

$$H = \begin{cases} LH \times (1 + r), & TH < LH, \\ TH, & TH \geq LH, \end{cases} \quad (15)$$

$$TH = \frac{F(i) - BF}{WF - BF} \times r_6 \times 2 \times (UB - LB), \quad (16)$$

其中, $AllF(i)$ 表示当前迭代中每个个体的适应度值, LH 为 H 的下界, 本研究中设置为10000, WF 为当前最差适应度值, r_6 为[0, 1]范围内的随机数, UB 和 LB 表示特征空间的上限和下限。

2.3.4 麻雀搜索算法

SSA是一种受麻雀捕食行为启发的群体智能优化算法, 过程如图7所示。在SSA的优化过程中, 麻雀分为两类, 即生产者和跟随者。不同麻雀的坐标位置即反映了Bi-LSTM神经网络中的学习率、神经元数量和丢弃率。生产者是麻雀种群中适应函数较小的个体并为所有跟随者提供方向, 本文中生产者占比20%。第 $(t+1)$ 次迭代中生产者位置更新如式(17)所示:

$$X_{ij}^{t+1} = \begin{cases} X_{ij}^t \cdot \exp\left(\frac{-i}{a \cdot iter_{max}}\right), & \text{if } R_2 < ST, \\ X_{ij}^t \cdot Q, & \text{if } R_2 \geq ST, \end{cases} \quad (17)$$

式中, a 是[0, 1]范围内的随机数, Q 是标准正态分布随机数, R_2 是[0, 1]范围内的随机数, ST 为警戒阈值, 本文中取0.8。

跟随者为种群中除生产者以外的个体。在生产者完成位置更新后, 跟随者会向食物最丰富的位置靠拢, 即向适应度函数最低的生产者方向移动。第 $(t+1)$ 次迭代中跟随者的位置更新为

$$X_{ij}^{t+1} = \begin{cases} Q \cdot \exp\left(\frac{X_{wj}^t - X_{ij}^t}{i^2}\right), & \text{if } i > n/2, \\ X_{pj}^t + \frac{1}{J} \sum_{j=1}^J L_j |X_{ij}^t - X_{pj}^t|, & \text{if } i \leq n/2, \end{cases} \quad (18)$$

式中, n 是麻雀种群的大小, J 是超参数的数量, L_j 是一个随机数, 取0或1。

全部麻雀的位置更新完成后, 会有一部分麻雀发现危险成为警戒者, 警戒者既可以是生产者也可以是跟随者, 放弃当前位置并移动到一个新位置, 本研究中警戒者占比10%。警戒者的存在可以避免麻雀种群陷入局部最优解, 并帮助种群维持多样性。位于种群边缘的警戒者会迅速前往种群中的当前最优位置, 位

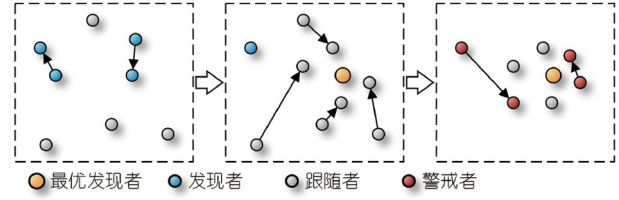


图7 (网络版彩图) SSA算法计算原理

Figure 7 (Color online) Calculation principle of SSA algorithm.

于种群中央的警戒者会随机移动。警戒者的位置变化表达式为

$$X_{ij}^{t+1} = \begin{cases} X_{bj}^t + \beta \cdot |X_{ij}^t - X_{bj}^t|, & \text{if } f_i > f_g, \\ X_{ij}^t + K \cdot \left(\frac{|X_{ij}^t - X_{wj}^t|}{(f_i - f_w) + \theta} \right), & \text{if } f_i = f_g, \end{cases} \quad (19)$$

式中, β 是符合标准正态分布随机数, K 是[-1, 1]范围内的随机数, θ 是一个较小的数字, 防止分母为0。适应度值最小的麻雀的位置坐标代表最优的超参数组合。

2.4 预测方法框架

Bi-LSTM模型在不同启发式优化算法下的实现过程如图8所示。考虑隧道几何参数、地质力学参数和盾构掘进参数的影响, 构建盾构刀盘扭矩预测的数据集。本文将数据集分为训练集和测试集, 训练集样本占总数据集的80%, 剩余20%的数据样本构成测试集。在训练模型之前, 须对数据集进行归一化处理降低不同参数间的量级差异对模型性能的影响。根据式(20)将数据归一化至[-1, 1]区间范围内。

$$P_n = \frac{(P_{max} - P_{min})(X - X_{min})}{(X_{max} - X_{min})} + P_{min}, \quad (20)$$

式中, P_n 是归一化值, P_{max} 和 P_{min} 是归一化数据的最大值和最小值, X_{max} 和 X_{min} 是原始数据的最大值和最小值。

Bi-LSTM的时间步设置为20, 即临近已开挖20环盾构隧道参数会对被预测环产生影响。优化算法的适应度函数是Bi-LSTM模型在训练过程中预测值与真实值的均方根误差, 优化对象是Bi-LSTM模型的学习率、神经元个数与丢弃率, 优化目标是优化次数达到最大迭代步数且适应度值达到稳定。每种优化算法在迭代过程中达到最小适应度函数值时对应的学习率、

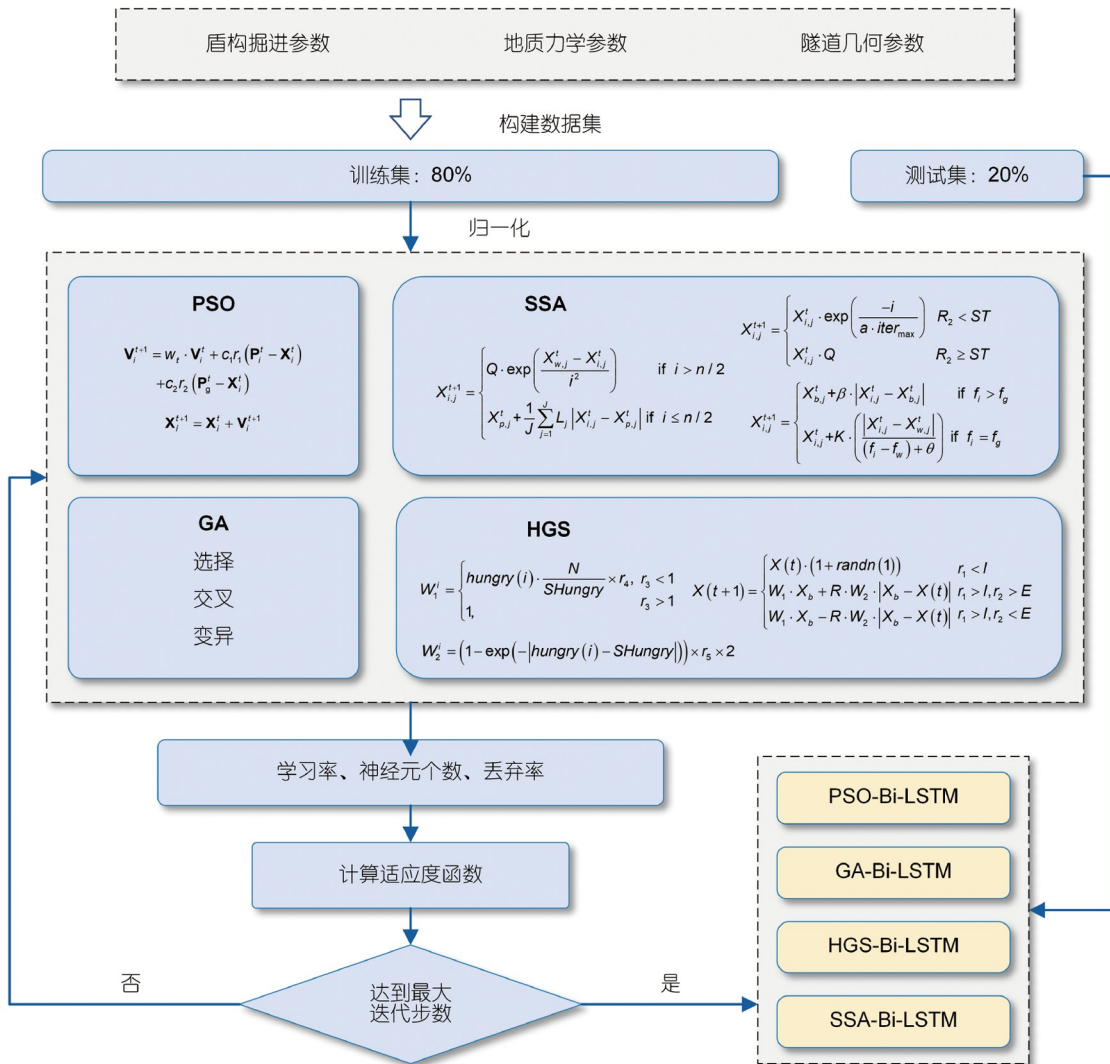


图 8 (网络版彩图)基于智能优化算法的Bi-LSTM模型预测流程图

Figure 8 (Color online) Flow chart of Bi-LSTM model prediction based on intelligent optimization algorithm.

神经元个数与丢弃率取为该种优化算法的最终优化结果. 本文建立了4种混合智能预测模型, 即PSO-Bi-LSTM, GA-Bi-LSTM, SSA-Bi-LSTM与HGS-Bi-LSTM.

3 工程案例及数据分布

刀盘扭矩是地层阻力与切削反力相互作用的结果, 其他盾构掘进参数、地质力学参数和隧道几何参数联合影响着盾构刀盘扭矩值^[35,36]. 地质力学参数反映当前环的地层条件, 若当前环的地层土体难以开挖, 刀盘所受的切削反力就会增大, 刀盘扭矩随之增

大^[22,37]. 隧道埋深越大, 盾构机刀盘所受的土压力越大, 刀盘扭矩会相应增加^[23,38]. 基于某盾构隧道工程, 考虑15个参数对刀盘扭矩预测的影响, 包括1个几何参数, 即隧道埋深(h); 8个盾构掘进参数, 即推力(TH)、掘进速度(ES)、土仓压力(CP)、进尺(EL)、出土量(EG)、同步注浆量(GV)、同步注浆压力(GP)以及同步注浆系数(GC); 6个地质力学参数, 即天然密度(γ)、等效内摩擦角(φ)、变形模量(G)、泊松比(ν)、渗透系数(PC)以及岩石饱和单轴极限抗压强度(U). 本文盾构掘进参数选取盾构机在每一环稳定掘进段的平均值, 综合反映掘进参数在当前环的工作状态. 案例隧道

工程的地层条件为由粉质黏土、中粗砂、卵石土和风化花岗岩组成的复合地层。考虑不同地层的厚度和深度, 对不同地层的地质力学参数进行加权平均, 从而确定当前地层的地质力学参数, 如式(21)所示:

$$X = \sum_{i=1}^N \frac{t_i}{h} \cdot \frac{h_i}{h} \cdot X_i, \quad (21)$$

式中, X_i 为地质力学参数, N 为地层数, t_i 为第*i*层土的厚度, h_i 为第*i*层土的埋深, h 为隧道埋深。每个参数的最大、最小以及平均值如表1所示。15个输入值和1个输出值的分布如图9所示。需要注意的是, 在软土地层中开挖时, 岩石饱和单轴极限抗压强度值取0。

本文采用的盾构隧道工程共采集得到了500组数据。其中, 每环数据包含上述的16个值(15个输入值, 1个输出值)。为满足深度神经网络的训练需求, 对整理完成的数据样本进行划分, 400个样本为训练集, 100个样本为测试集。

4 预测结果分析

从训练过程的适应度函数值收敛情况研究4种混合预测模型的稳定性, 从适应度函数达到稳定值所需

计算时间的角度研究4种混合预测方法的计算效率, 基于测试集数据研究4种混合方法的预测精度。

4.1 智能优化算法稳定性分析

Bi-LSTM神经网络和优化算法中存在随机参数, 导致混合预测模型在训练过程中适应度函数有随机性。不同优化算法均在不同种群数量下分别计算10次来获得适应度函数包络线。对于每一个迭代步, 将10次优化中的最大和最小适应度函数值视为上下边界点。通过连接每个迭代步的上下边界点即可获得包含完整上下边界的适应度函数包络线。多次计算获得的适应度函数差异越小, 即适应度函数包络线的上下边界值越接近, 表明优化算法的稳定性越强。反之, 适应度函数值的上下边界差距越大表明优化算法的稳定性越差。当种群数分别为10、30和50时, 4种启发式优化算法的适应度函数包络线如图10所示。4种优化算法在不同种群数量下, 适应度函数达到稳定状态时对应的迭代次数均小于70次。因此最大迭代次数设置为100满足优化算法寻优Bi-LSTM网络超参数的需求。随着种群数量的增加, 4种优化算法的适应度函数值上下边界的差值有明显减小的趋势, 表明增加种群数量能够提高优化算法的稳定性。

表1 参数最大值、最小值、平均值及中位数

Table 1 Maximum, minimum, mean, and median values of parameters

参数	最大值	最小值	平均值	中位数
h (m)	17.974	7.140	10.467	9.417
γ (g/cm ³)	2.211	1.120	1.366	1.345
φ (°)	53.338	14.001	20.253	16.531
G (MPa)	18.103	9.941	12.546	12.322
ν	0.367	0.148	0.251	0.259
PC (m/d)	22.279	1.041	6.598	5.641
U' (MPa)	50.349	0.000	5.966	0.000
TH (kN)	19000.000	5800.000	9952.655	9000.000
ES (mm/min)	85.000	3.000	41.548	49.000
CP (bar)	2.000	0.100	1.325	1.500
EL (mm)	1698.000	1112.000	1487.843	1491.000
EG (m ³)	72.000	41.000	54.462	51.000
GV (m ³)	6.600	5.000	6.083	6.000
GP (bar)	4.400	1.100	2.270	2.100
GC	1.650	1.250	1.516	1.500
CHT (kN m)	3800.000	840.000	1591.506	1320.000

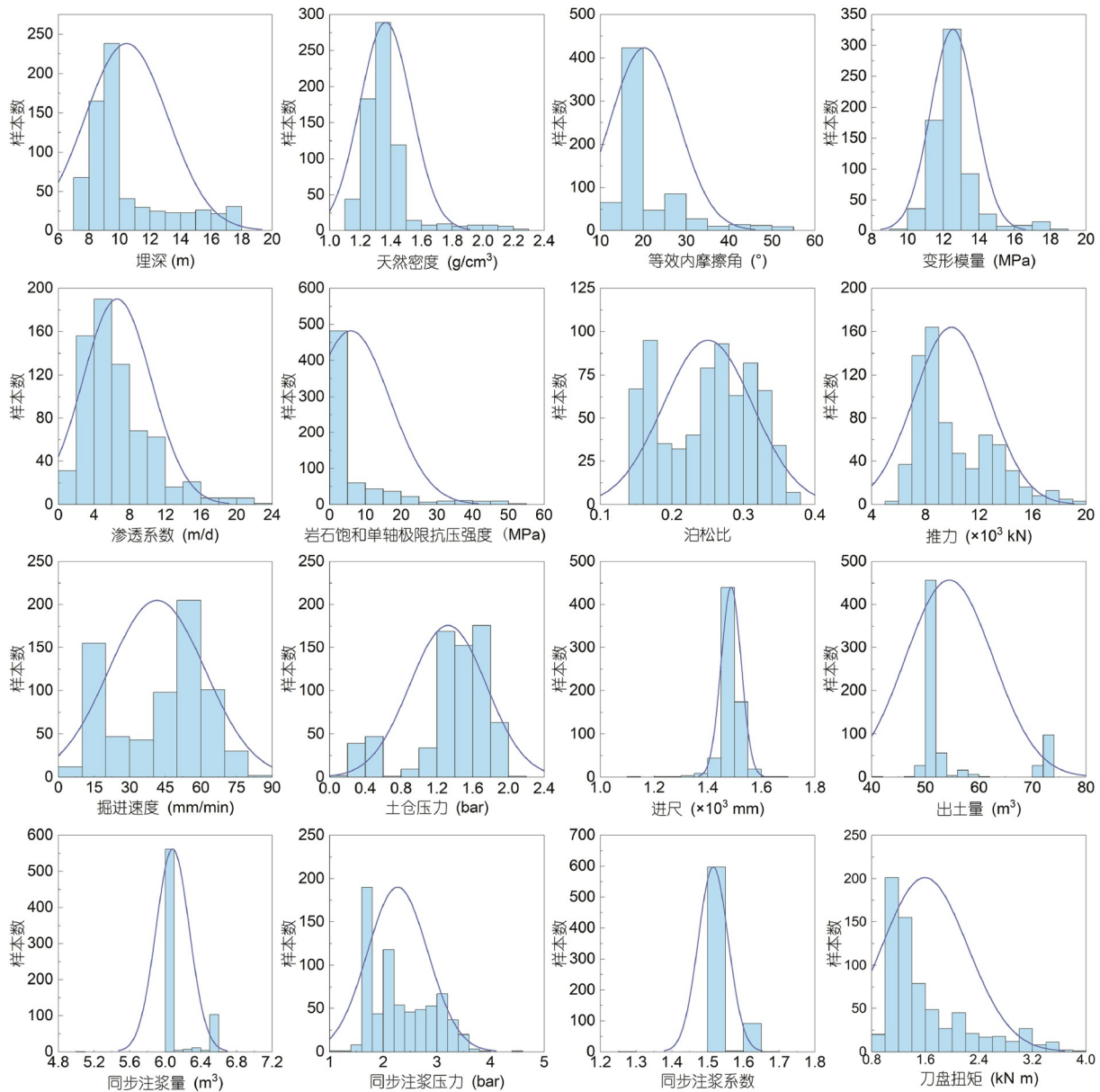


图 9 (网络版彩图)输入和输出参数分布图

Figure 9 (Color online) Input and output parameter distribution diagram.

由于4种优化算法有各自的优化策论和寻优方式, 对于相同的数据集和种群数量, 也将表现出不同的预测效果, 因此获得的最优适应度值不相同. 当种群数量为50时, GA和HGS的适应度函数下边界较小, SSA在种群数量为30时的适应度函数下边界较小. 当种群数量为50并且迭代次数达到第100次时, 在4种方法中, HGS的上下边界差值最小, 表明相比于其他3种方法, HGS优化算法的稳定性较好.

4.2 预测效率分析

采用适应度函数达到稳定值时对应的迭代次数衡量启发式优化算法寻优超参数的效率. 图11为4种优化算法在10次计算中适应度函数达到稳定所需的迭代次数及其平均值.

当种群数量为10次时, GA和HGS的平均稳定迭代次数较高, 分别为52.2次以及47.6次; PSO和SSA的平

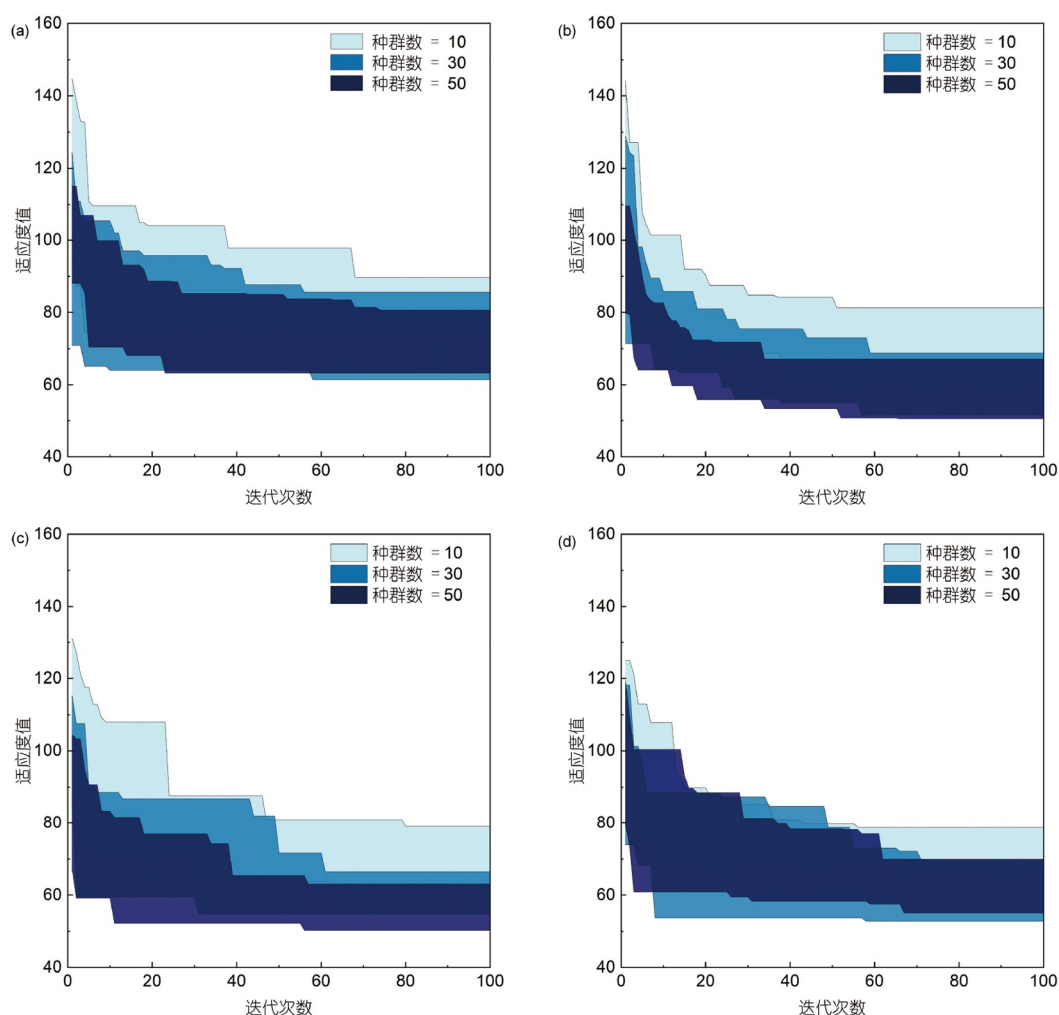


图 10 (网络版彩图)4种智能优化算法的适应度值. (a) PSO方法; (b) GA方法; (c) HGS方法; (d) SSA方法
Figure 10 (Color online) Fitness values of four intelligent optimization algorithms. (a) PSO; (b) GA; (c) HGS; (d) SSA.

均稳定迭代次数较少, 分别为45.5次及43.0次. 当种群数量为30时, 4种优化算法的平均稳定迭代次数持平. 当种群数量为50次时, HGS平均稳定迭代次数最小, GA和SSA次之, PSO的迭代次数最大.

基于4种优化算法的适应度值达到稳定的平均迭代次数计算了优化算法达到稳定所消耗的时间. 当种群数量为10, 30和50时, 4种启发式优化算法的适应度值达到稳定的平均时间如图12所示. PSO, GA和HGS的优化时间相差较小, 而SSA的优化时间较长. SSA在全部麻雀的位置更新完成后, 一部分麻雀随机成为警戒者并移动到一个新位置. 每一轮迭代中, SSA中都有一部分粒子需要更新两次. 因此, SSA优化相较于其他

启发式优化算法要花费更长的时间. 当种群数量为50时, SSA所消耗的时间长达547.722 min, 是HGS运行时间的3倍之多. 以HGS优化算法为例, 其平均优化时间在种群数量为10, 30和50时, 分别为28.56, 74.456和166.872 min, 优化时间与种群数量基本成正比. 如图12所示, 在4种优化算法中, HGS优化算法在不同种群条件下计算时间均较小, 表明HGS在优化Bi-LSTM超参数时具有较高的效率.

4.3 预测精度分析

在优化稳定性方面, PSO, GA, SSA和HGS在种群数量为30和50时, 适应度函数达到稳定时差异较小, 分

别为2.123%, 2.364%, 2.092%和6.462%。4种启发式优化算法的适应度值达到稳定的平均计算时间如图12所示。优化算法在种群数量为50时所花费的时间要明显

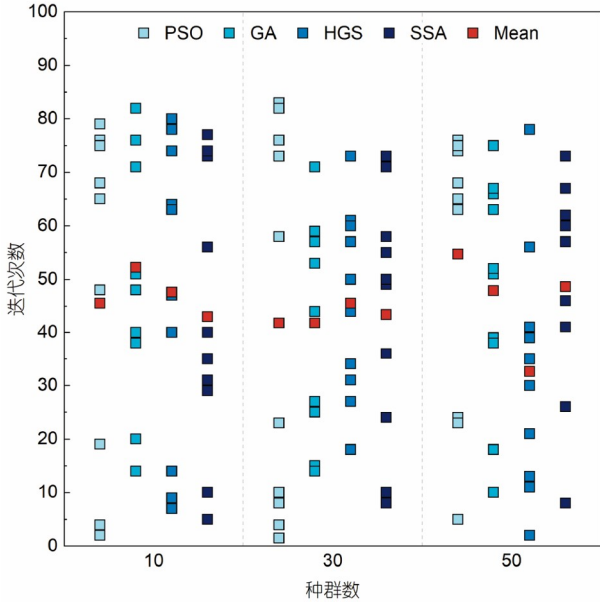


图 11 (网络版彩图)适应度值达到稳定时的迭代次数
Figure 11 (Color online) Iteration times when fitness values reach stability.

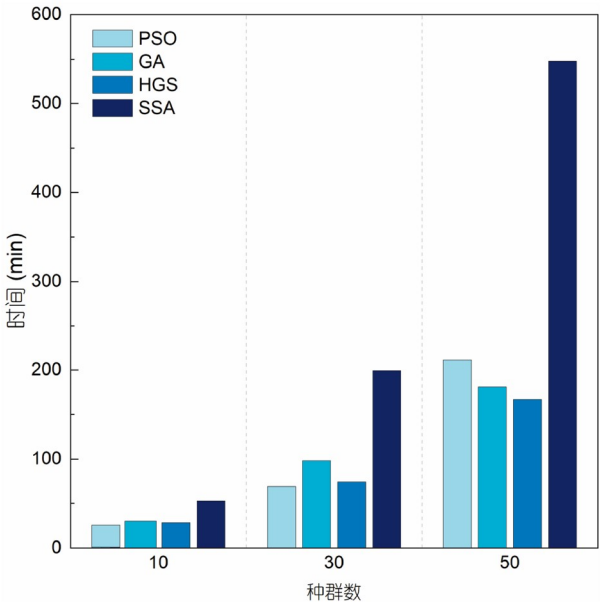


图 12 (网络版彩图)适应度值达到稳定时的计算时间
Figure 12 (Color online) Calculation time when the fitness value reaches stability.

高于种群数量为30的情况。以SSA为例,在种群数量为50时,采用一张NVIDIA RTX A5000显卡优化10次的时间超过100 h。综合考虑优化稳定性与计算效率,选取种群数量为30获得的适应度函数最小时的超参数和神经网络参数作为Bi-LSTM的训练参数。

将已训练的4种混合智能模型用于预测相同的测试集,预测结果如图13所示。基于测试集计算结果进行了线性拟合,并绘制了对应的95%置信区间和95%预测区间。由图13可知,与PSO-Bi-LSTM和SSA-Bi-LSTM相比,GA-Bi-LSTM和HGS-Bi-LSTM的预测结果中更多的样本在95%预测区间与置信区间内,表明GA-Bi-LSTM以及HGS-Bi-LSTM的预测结果更接近于实测值。

采用4种性能评价指标,即决定系数(R^2),均方根误差(RMSE),平均误差(MAE)以及平均百分比误差(MAPE)分析4种混合预测模型的性能,性能评价指标的表达式如式(22)~(25)所示:

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{q}_i - q_i)^2}{\sum_{i=1}^n (q_i - \bar{q})^2}, \quad (22)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (q_i - \hat{q}_i)^2}, \quad (23)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{q}_i - q_i|, \quad (24)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{q}_i - q_i}{q_i} \right|, \quad (25)$$

式中, $\hat{q}_i$ 为预测值, q_i 为实测值, $\bar{q}$ 是实测值的平均值, n 为测试集样本总数。

4种混合预测模型的评价指标如表2所示。由表2可知, HGS-Bi-LSTM的预测性能最好, 其 R^2 为0.786, RMSE, MAE和MAPE分别为321.941 kN m, 263.922 kN m和14.9%。其余3种模型的性能评价指标均差于HGS-Bi-LSTM。基于4种混合预测模型在测试集的预测结果和刀盘扭矩实测值, 计算了绝对误差及误差分布, 如图14所示。绝对误差平均值从小到大依次是HGS-Bi-LSTM方法、SSA-Bi-LSTM方法、PSO-Bi-LSTM方法、GA-Bi-LSTM方法, 结果表达采用HGS优化算法可获得Bi-LSTM模型合理的超参数, 具有较高的预测精度。

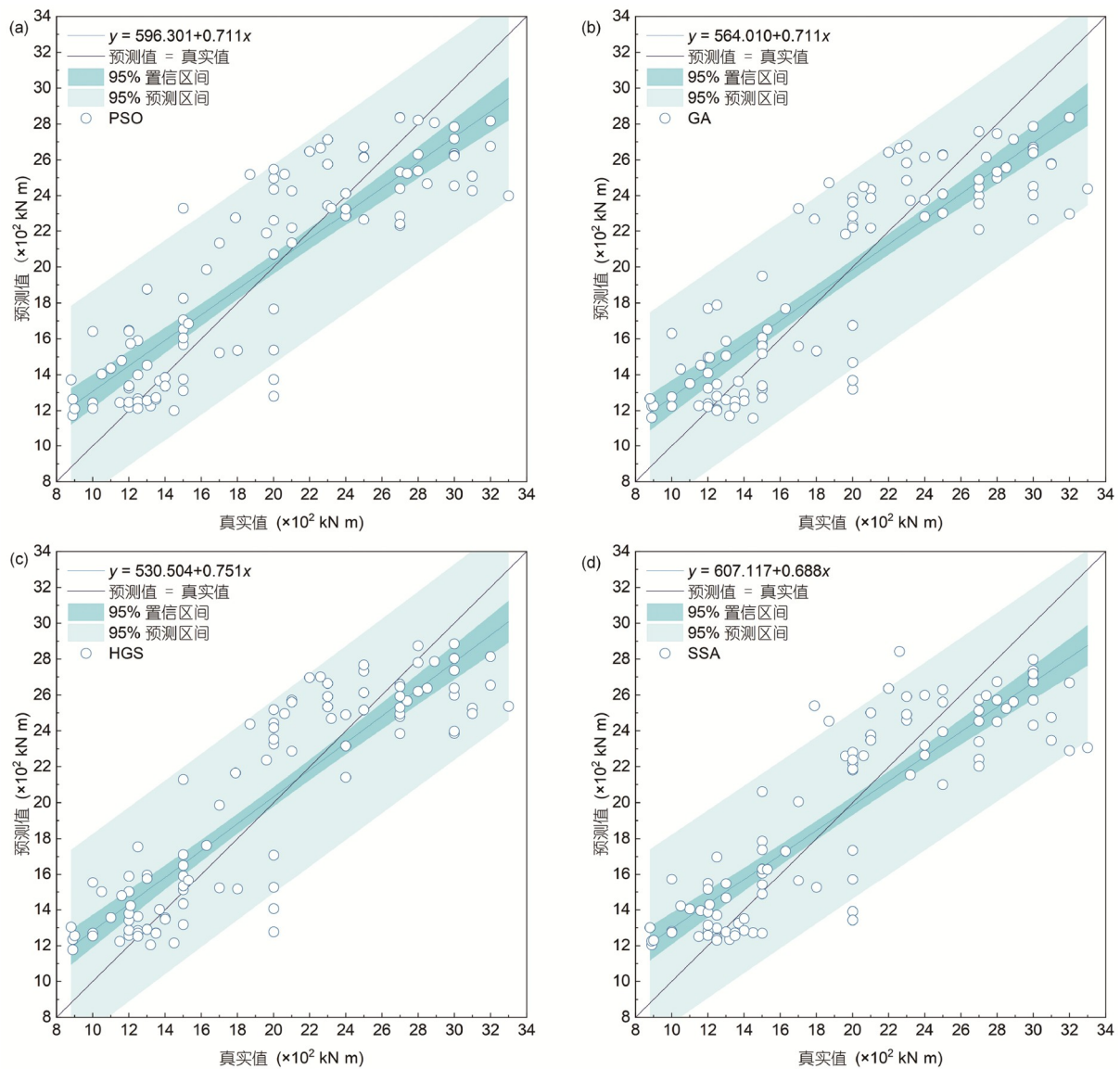


图 13 (网络版彩图)4种方法预测盾构机刀盘扭矩散点图. (a) PSO-Bi-LSTM方法; (b) GA-Bi-LSTM方法; (c) HGS-Bi-LSTM方法; (d) SSA-Bi-LSTM方法

Figure 13 (Color online) Scatter plot of predicting shield machine cutterhead torque for four methods. (a) PSO-Bi-LSTM; (b) GA-Bi-LSTM; (c) HGS-Bi-LSTM; (d) SSA-Bi-LSTM.

表 2 不同方法预测盾构机刀盘扭矩的评价指标对比

Table 2 Comparison of evaluation indicators for predicting shield machine cutterhead torque using different methods

方法	RMSE (kN m)	MAE (kN m)	MAPE (%)	R^2
PSO-Bi-LSTM	343.038	279.206	15.8	0.757
GA-Bi-LSTM	339.975	280.439	15.5	0.761
HGS-Bi-LSTM	321.941	263.922	14.9	0.786
SSA-Bi-LSTM	337.971	275.718	15.1	0.764

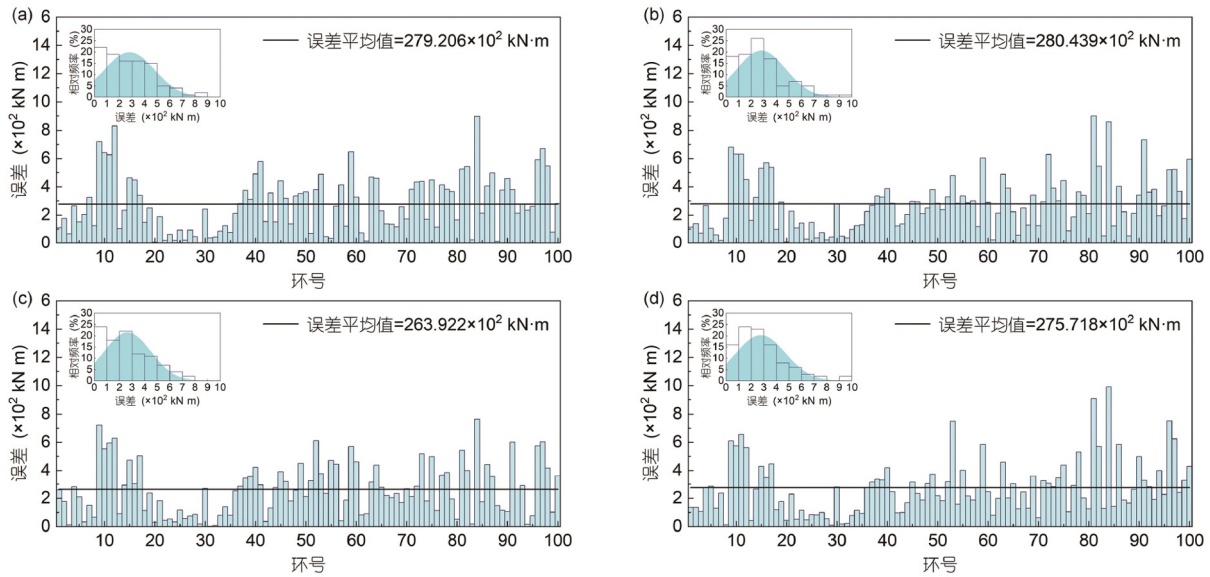


图 14 (网络版彩图)4种方法预测盾构机刀盘扭矩误差图. (a) PSO-Bi-LSTM方法; (b) GA-Bi-LSTM方法; (c) HGS-Bi-LSTM方法; (d) SSA-Bi-LSTM方法

Figure 14 (Color online) Error diagram of predicting the torque of shield machine cutterhead for four methods. (a) PSO-Bi-LSTM; (b) GA-Bi-LSTM; (c) HGS-Bi-LSTM; (d) SSA-Bi-LSTM.

5 结论

本文采用启发式智能优化算法确定Bi-LSTM神经网络的超参数, 提出盾构隧道刀盘扭矩混合智能预测方法. 本文获得的结论如下.

(1) 提出了盾构隧道刀盘扭矩智能预测方法. 深度学习预测模型采用了Bi-LSTM神经网络, 联合PSO, GA, HGS和SSA等4种优化算法确定了Bi-LSTM中的3个超参数, 考虑了隧道几何参数、地质力学参数和部分掘进参数对刀盘扭矩的耦合影响, 提出了从精度、效率和稳定性三方面评价混合预测方法的指标体系, 最终确定预测刀盘扭矩最优的混合模型.

(2) 研究了PSO, GA, HGS和SSA寻优Bi-LSTM超参数的稳定性和计算效率. 基于本文的数据集, 对于4

种优化算法, 增加种群数量可有效提升寻优超参数的稳定性, HGS算法稳定性最好, PSO算法的稳定性最差. 4种优化算法寻优超参数的计算时间与种群数量成正比关系, HGS算法计算效率最快, SSA算法计算效率最慢.

(3) 研究了4种混合智能模型预测刀盘扭矩的精度. 在4种方法中, HGS-Bi-LSTM预测模型的 R^2 最大, $RMSE$, MAE 和 $MAPE$ 最小. 基于4种评价指标, 预测精度从高到低依次是HGS-Bi-LSTM, SSA-Bi-LSTM, GA-Bi-LSTM和PSO-Bi-LSTM.

基于本文提出的刀盘扭矩智能预测方法和所搜集的数据集, HGS算法是优化Bi-LSTM神经网络超参数的适用方法, 建立的HGS-Bi-LSTM预测模型对于刀盘扭矩有较高的预测精度.

参考文献

- Han J, Zhao W, Jia P, et al. Risk analysis of the opening of shield-tunnel circumferential joints induced by adjacent deep excavation. *J Perform Constr Facil*, 2018, 32: 04017123
- Zhou X P, Zhai S F. Estimation of the cutterhead torque for earth pressure balance TBM under mixed-face conditions. *Tunn Undergr Sp Tech*, 2018, 74: 217–229
- Ates U, Bilgin N, Copur H. Estimating torque, thrust and other design parameters of different type TBMs with some criticism to TBMs used in Turkish tunneling projects. *Tunn Undergr Sp Tech*, 2014, 40: 46–63

- 4 Zhang P, Chen R P, Wu H N. Real-time analysis and regulation of EPB shield steering using Random Forest. *Automat Constr*, 2019, 106: 102860
- 5 Do N A, Dias D, Oreste P, et al. Three-dimensional numerical simulation of a mechanized twin tunnels in soft ground. *Tunn Undergr Sp Tech*, 2014, 42: 40–51
- 6 Zhang Z X, Liu C, Huang X, et al. Three-dimensional finite-element analysis on ground responses during twin-tunnel construction using the URUP method. *Tunn Undergr Sp Tech*, 2016, 58: 133–146
- 7 Soomro M A, Ng C W W, Liu K, et al. Pile responses to side-by-side twin tunnelling in stiff clay: Effects of different tunnel depths relative to pile. *Comput Geotech*, 2017, 84: 101–116
- 8 Lu H, Wang M Y, Xia Y P, et al. Calculation model of cutterhead torque for earth pressure balance shield (in Chinese). *J Zhejiang Univ (Eng Sci)*, 2014, 48: 1640–1645 [卢浩, 王明洋, 夏沅谱等. 土压平衡盾构刀盘扭矩计算模型. 浙江大学学报(工学版), 2014, 48: 1640–1645]
- 9 Hu G J, Yuan Y F, Yu K, et al. Experimental study on the interaction between excavation face and cutter head of shield tunneling machine with different opening rate (in Chinese). *J East China Jiaotong Univ*, 2023, 40: 16–21 [胡光静, 袁岳峰, 余昆. 不同开口率下盾构机刀盘与开挖面相互作用试验研究. 华东交通大学学报, 2023, 40: 16–21]
- 10 Liu Y M, Zhao W G, Zhang Z P, et al. Digging and analysis of implicit relationship between working parameters of tunnel boring machine (TBM) (in Chinese). *J Shijiahuang Railw Inst*, 2006, 3: 55–59 [刘明月, 赵维刚, 张筑平. 全断面隧道掘进机工作参数匹配规律的挖掘与利用. 石家庄铁道学院学报, 2006, 3: 55–59]
- 11 Kong F, Lu D, Ma Y, et al. Analysis and intelligent prediction for displacement of stratum and tunnel lining by shield tunnel excavation in complex geological conditions: A case study. *IEEE Trans Intell Transp Syst*, 2022, 23: 22206–22216
- 12 Chen X S, Zeng S Q, Han W L, et al. Review and prospect of machine learning method in shield tunnel construction (in Chinese). *J Civ Environ Eng*, 2022, 46: 1–13 [陈湘生, 曾仕琪, 韩文龙, 等. 机器学习方法在盾构隧道工程中的应用研究现状与展望. 土木与环境工程学报, 2022, 46: 1–13]
- 13 Kong F C, Lu D C, Ma Y D, et al. Novel hybrid method to predict the ground-displacement field caused by shallow tunnel excavation. *Sci China Tech Sci*, 2023, 66: 101–114
- 14 Zhang Q, Liu Z, Tan J. Prediction of geological conditions for a tunnel boring machine using big operational data. *Automat Constr*, 2019, 100: 73–83
- 15 Kong X, Ling X, Tang L, et al. Random forest-based predictors for driving forces of earth pressure balance (EPB) shield tunnel boring machine (TBM). *Tunn Undergr Sp Tech*, 2022, 122: 104373
- 16 Yu H, Qin C, Tao J, et al. A multi-channel decoupled deep neural network for tunnel boring machine torque and thrust prediction. *Tunn Undergr Sp Tech*, 2023, 133: 104949
- 17 Yin Z Q, Zhai W J, Han A M, et al. Predictive research of cutterhead torque for metro shield boring machine based on symbolic regression algorithm (in Chinese). *Urban Mass Transit*, 2021, 24: 127–131 [尹志清, 翟维骏, 韩爱民, 等. 基于符号回归算法的地铁盾构刀盘扭矩预测研究. 城市轨道交通研究, 2021, 24: 127–131]
- 18 Zhou C, Ding L, Zhou Y, et al. Visibility graph analysis on time series of shield tunneling parameters based on complex network theory. *Tunn Undergr Sp Tech*, 2019, 89: 10–24
- 19 Gao B, Wang R R, Lin C, et al. TBM penetration rate prediction based on the long short-term memory neural network. *Undergr Space*, 2021, 6: 718–731
- 20 Wang K, Zhang L, Fu X. Time series prediction of tunnel boring machine (TBM) performance during excavation using causal explainable artificial intelligence (CX-AI). *Automat Constr*, 2023, 147: 104730
- 21 Xu Q, Huang X, Zhang B, et al. TBM performance prediction using LSTM-based hybrid neural network model: Case study of Baimang River tunnel project in Shenzhen, China. *Undergr Space*, 2023, 11: 130–152
- 22 Lin S S, Shen S L, Zhang N, et al. Modelling the performance of EPB shield tunnelling using machine and deep learning algorithms. *Geosci Front*, 2021, 12: 101177
- 23 Zhang P, Wu H N, Chen R P, et al. A critical evaluation of machine learning and deep learning in shield-ground interaction prediction. *Tunn Undergr Sp Tech*, 2020, 106: 103593
- 24 Qin C, Shi G, Tao J, et al. Precise cutterhead torque prediction for shield tunneling machines using a novel hybrid deep neural network. *Mech Syst Signal Process*, 2021, 151: 107386
- 25 Wang R, Li D, Chen E J, et al. Dynamic prediction of mechanized shield tunneling performance. *Automat Constr*, 2021, 132: 103958

- 26 Lu D, Ma Y, Kong F, et al. Support vector regression with heuristic optimization algorithms for predicting the ground surface displacement induced by EPB shield tunneling. [Gondwana Res](#), 2023, 123: 3–15
- 27 Wang K, Hua Y, Huang L, et al. A novel GA-LSTM-based prediction method of ship energy usage based on the characteristics analysis of operational data. [Energy](#), 2023, 282: 128910
- 28 Li X, Ma X, Xiao F, et al. Time-series production forecasting method based on the integration of Bidirectional Gated Recurrent Unit (Bi-GRU) network and Sparrow Search Algorithm (SSA). [J Pet Sci Eng](#), 2022, 208: 109309
- 29 Ye X W, Wei Y J, Zhang X L, et al. Confining pressure forecasting of shield tunnel lining during construction based on LSTM-PSO models combined with the multi-output recursive strategy. [Tunn Undergr Sp Tech](#), 2024, 143: 105499
- 30 Mahmoodzadeh A, Nejati H R, Mohammadi M, et al. Forecasting tunnel boring machine penetration rate using LSTM deep neural network optimized by grey wolf optimization algorithm. [Expert Syst Appl](#), 2022, 209: 118303
- 31 Lin S S, Zhou A, Shen S L. A coupled deep learning approach for shield moving performance prediction of underground tunnel construction. [Gondwana Res](#), 2023, 123: 125–139
- 32 Mehta P, Yildiz B S, Sait S M, et al. Hunger games search algorithm for global optimization of engineering design problems. [Mater Testing](#), 2022, 64: 524–532
- 33 Shaiba H, S. Alzahrani J, M. Eltahir M, et al. Hunger search optimization with hybrid deep learning enabled phishing detection and classification model. [Comput Mater Continua](#), 2022, 73: 6425–6441
- 34 Zhang J, Liu Z, Zhou Y, et al. Joint inversion method of rock physics based on hunger games search correction and bidirectional long-short-term memory network. [IEEE Trans Geosci Remote Sens](#), 2024, 62: 1–10
- 35 Shen S L, Elbaz K, Shaban W M, et al. Real-time prediction of shield moving trajectory during tunnelling. [Acta Geotech](#), 2022, 17: 1533–1549
- 36 Elbaz K, Zhou A, Shen S L. Deep reinforcement learning approach to optimize the driving performance of shield tunnelling machines. [Tunn Undergr Sp Tech](#), 2023, 136: 105104
- 37 Qin C, Liu M, Zhang Z, et al. An adaptive operating parameters decision-making method for shield machine considering geological environment. [Tunn Undergr Sp Tech](#), 2023, 141: 105372
- 38 Ye X W, Zhang X L, Zhang H Q, et al. Prediction of lining upward movement during shield tunneling using machine learning algorithms and field monitoring data. [Transp Geotech](#), 2023, 41: 101002

Intelligent prediction method for cutterhead torque of shield tunnel based on intelligent optimization algorithm and Bi-LSTM neural network

KONG FanChao¹, LIU YiHan², LIU ShiJi¹, LU DeChun² & DU XiuLi²

¹ School of Water Resources and Hydroelectric Engineering, North China Electric Power University, Beijing 102206, China

² Institute of Geotechnical and Underground Engineering, Beijing University of Technology, Beijing 100124, China

Intelligent prediction of reasonable cutterhead torque can effectively reduce the stratum disturbance caused by shield tunneling and reduce cutterhead wear, which is of great significance for the safe and efficient construction of shield tunnels. A time-series intelligent prediction method for shield tunnel cutterhead torque is proposed based on intelligent optimization algorithms and deep-learning neural networks. Bi-LSTM neural network is used as the deep learning prediction model. The particle swarm optimization algorithm (PSO), genetic optimization algorithm (GA), sparrow search algorithm (SSA), and hunger game search algorithm (HGS) are jointly used to intelligently determine the three hyperparameters of the Bi-LSTM neural network, respectively, namely that, learning rate, the number of neurons, and dropout rate. Furthermore, an intelligent prediction method for the cutterhead torque of the shield tunnel is obtained. The root mean square error between the predicted values of intelligent models and measured values is used as fitness functions for four optimization algorithms. Normalization method is used for data preprocessing to reduce the impact of magnitude differences between different parameters. The input of the intelligent prediction model considers the influence of 15 parameters on the cutterhead torque, including 1 tunnel geometric parameter, 6 geological mechanical parameters, and 8 shield tunneling parameters. The performances of four optimization algorithms for determining hyperparameters of Bi-LSTM neural network are evaluated from three perspectives: accuracy, efficiency, and stability, based on on-site measured data of shield tunneling. The results indicate that compared with the other three optimization algorithms, the HGS method has better stability, faster efficiency, and higher computational accuracy, which is an appropriate method to determine the hyperparameters of Bi-LSTM.

Bi-LSTM neural network, hyperparameters, heuristic intelligent optimization algorithm, cutterhead torque, performance analysis

doi: [10.1360/SST-2024-0145](https://doi.org/10.1360/SST-2024-0145)