基于 AE-TCN-Transformer 并行模型的巷道变形多步预测研究*

王龙飞[1],陈俊智[1],任春芳[1],蔡川川[1],徐晓文[1],刘洋[2],何昌耀[2]

(1.昆明理工大学 国土资源工程学院, 云南 昆明 650093;
2.昆明理工大学 公共安全与应急管理学院, 云南 昆明 650093)

摘要:为实现巷道变形的精准预测,提出一种基于阿尔法进化算法(AE)优化的时间卷积网络(TCN)和 Transformer 并行预测模型。该模型通过并行架构融合 TCN 与 Transformer 模型,TCN 利用膨胀因果卷积挖掘局部时序特征,Transformer 通过多头自注意力机制提升长时序模式的捕捉效率,最后通过时序注意力引导的多模型动态特征融合机制输出结果。同时,引入 AE 算法优化超参数,凭借自适应进化路径与双步长机制,增强模型在复杂时序数据中的鲁棒性。以云南某铁矿三个不同中段监测点的巷道位移监测数据为研究对象进行试验验证,结果表明:AE-TCN-Transformer 模型在预测步长为 6～48 步时,其均方根误差控制在 0.013 2 mm 以下,对称平均绝对百分比误差低于 0.09%,决定系数高于 0.97,显著优于 TCN、Transformer、LSTM 等单一模型,充分验证了 AE-TCN-Transformer 模型在巷道变形预测中的精确性与可靠性,为巷道变形精准预测提供了新路径。

关键词:巷道变形预测;AE-TCN-Transformer 并行模型;时间卷积网络;阿尔法进化算法;多尺度特征融合

中图分类号:TD325 **文献标识码**:A **文章编号**:1005-2763(2026)04-0084-10

Research on Multi-Step Prediction of Roadway Deformation Based on AE-TCN-Transformer Parallel Model

WANG Longfei[1], *CHEN Junzhi*[1], *REN Chunfang*[1], *CAI Chuanchuan*[1], *XU Xiaowen*[1], *LIU Yang*[2], *HE Changyao*[2]

(1.Faculty of Land Resources Engineering, Kunming University of Science and Technology, Kunming, Yunnan 650093, China;
2.Faculty of Public Safety and Emergency Management, Kunming University of Science and Technology, Kunming, Yunnan 650093, China)

Abstract: To achieve accurate prediction of roadway deformation, a parallel prediction model that combines a Temporal Convolutional Network (TCN) and a Transformer, optimized using the Alpha evolution (AE) algorithm was proposed. This model integrated TCN and Transformer architectures through a parallel framework. The TCN utilized dilated causal convolutions to extract local temporal features, while the Transformer employed a multi-head self-attention mechanism to enhance the capture of long-term temporal patterns. Finally, the model output results through a dynamic multi-model feature fusion mechanism guided by temporal attention. At the same time, the AE algorithm was introduced to optimize hyper parameters, leveraging adaptive evolutionary paths and a two-step mechanism to enhance the model's robustness as handling complex time-series data. Experimental validation was conducted using roadway displacement monitoring data from three different monitoring points in the middle levels of an iron ore in Yunnan. The results show that when the prediction step size ranges from 6 to 48, the AE-TCN-Transformer model maintains a root mean square error (*RMSE*) below 0.013 2 mm, a symmetric mean absolute percentage error (*SMAPE*) below 0.09%, and a coefficient of determination (R^2) above 0.97. These performance metrics significantly outperform those of single models such as TCN, Transformer, LSTM, and other single-model approaches. This fully validates the accuracy and reliability of the AE-TCN-Transformer model in roadway deformation prediction and provides a new approach for precise prediction of roadway deformation.

Key words: Prediction of roadway deformation, AE-TCN-Transformer parallel model, Temporal convolutional network, Alpha evolution algorithm, Multi-scale feature fusion

---

* **收稿日期**:2025-05-14
基金项目:国家自然科学基金项目(U1602232)
作者简介:王龙飞(2000—),男,陕西宝鸡人,硕士研究生,主要从事矿山岩体力学方面的研究。E-mail:wanglongfei0322@foxmail.com
通信作者:陈俊智(1974—),男,陕西榆林人,博士,副教授,主要从事岩土工程方面的研究。E-mail:465424873@qq.com

0　引言

巷道作为矿山运输和生产安全的核心通道，其稳定性受到复杂力学环境的影响。在开采过程中，巷道围岩不仅要承受原岩应力场作用，还需承受周期性爆破振动、持续开采扰动等外部荷载影响[1-2]。同时，受岩体内部节理裂隙发育、岩体力学参数各向异性的影响，巷道围岩破裂、变形等失稳风险显著加剧[3]。巷道变形失稳不仅会造成设备损毁、生产中断等经济损失，更直接威胁着井下人员安全。因此，开展巷道变形的实时监测与准确预测，对矿山安全生产至关重要[4]。

传统的巷道稳定性预测方法主要建立在连续介质力学框架下，通过 FLAC3D 等数值模拟软件，基于弹塑性本构模型求解应力平衡微分方程[5]。这类方法通过 Mohr-Coulomb 强度准则等屈服条件，可推导理想工况下的围岩变形规律，但在实际工程应用中精度不足。张小军等[6]通过数值模拟方法预测巷道位移，顶板位移预测效果较好，但两帮位移预测值与现场监测结果存在约 30% 的误差。误差成因可归纳为以下两方面：一方面，岩体参数具有非均匀分布特性，其力学特性受到多种因素共同影响，同时受限于现场取样偏差，获取的本构模型参数难以还原复杂地质条件；另一方面，模型将动态荷载、非线性流变等复杂过程进行了简化[7]，且基于连续介质力学框架难以精确反映实际工程中的非连续性变形，最终导致预测结果和实际变形存在较大误差。

部分学者尝试通过弹性地基梁理论[8]、Verhulst 生物生长模型[9]、灰色预测模型[10]及贝叶斯更新算法[11]等数学建模方法求解巷道变形问题，并取得了阶段性进展。但部分模型仍存在预测精度不足的问题，如弹性地基梁理论预测误差达 8%～26%[8]；Verhulst 模型平均误差达 10%[9]。此外，数学模型在工程应用中仍存在以下局限性：其理论假设大多沿用了传统预测方法中的理想化条件，与实际巷道围岩存在差异；部分模型求解涉及高阶微分方程或复杂迭代算法，对计算资源要求高，难以满足现场实时预测要求。

近年来，随着计算机和人工智能技术的突破，深度学习模型研究在时间序列预测领域取得了显著进展，众多学者开始尝试通过随机森林算法[12]、模糊逻辑神经网络[13-14]、BP 神经网络[15]、长短期记忆网络[16]及卷积神经网络[17]等方法，构建巷道变形的非线性映射关系。但仍存在以下局限：其一，现有研究多采用单一网络结构进行预测，难以协同捕捉时序数据中局部突变特征与长期演化趋势之间的多尺度耦合规律，限制了模型对复杂工况的表征能力，例如随机森林算法预测结果相对误差达 11.18%[12]，神经网络预测结果最大偏差达 4.3 cm[17]；其二，传统优化算法虽广泛应用于模型参数调优，但其群体易过早聚集于参数空间局部区域，导致模型陷入局部最优陷阱，最终限制了深度神经网络的特征挖掘能力。

针对上述问题，本研究提出了一种基于阿尔法进化算法（AE）优化的时间卷积网络（TCN）和 Transformer 并行预测模型。该模型将 TCN 和 Transformer 模型并行设计，通过独立分支提取时序数据的多尺度特征。TCN 分支利用膨胀因果卷积聚焦围岩变形的局部时序规律，Transformer 分支借助多头自注意力机制解析围岩变形的长期演化趋势，两分支通过时序注意力引导的多模型动态特征融合机制输出预测结果。同时引入 AE 算法，通过自适应进化路径和双步长策略，实现模型超参数的高效搜索，提升模型在复杂环境下的预测精度。

1　理论基础

1.1　时间卷积网络(TCN)

时间卷积网络是一种专为时序数据建模设计的深度卷积神经网络架构，通过因果约束、膨胀卷积和残差学习，高效捕捉时间序列数据中的时序关系[18]。其核心结构主要由膨胀因果卷积、残差连接组成。

1.1.1　膨胀因果卷积

膨胀因果卷积是时间卷积网络中的核心组件，其通过左填充的方式，确保模型输出仅依赖历史数据，避免未来信息泄露；同时通过指数级的膨胀因子扩展感受野，在不增加参数量的前提下捕捉更多序列关系，膨胀因果卷积的公式如式(1)所示。

$$y_t = \sum_{i=0}^{k-1} w_i x_{t-di} \tag{1}$$

式中：y_t 为时间步 t 的输出；w_i 为卷积核的第 i 个权重参数；x_{t-di} 为输入序列中时间步 $t-di$ 的值；d 为膨胀率，用来控制卷积核的间隔采样；k 为卷积核大小，决定覆盖时间步数量。

1.1.2　残差连接

在时间卷积网络中，残差连接能够解决深层网络训练难题。其通过跳跃连接将浅层信息直接传递到深层，有效缓解了梯度消失的问题，并加快了收敛

速度。残差连接的表达式如式(2)所示。

$$Output = Activation(F(x) + Conv_{1\times1}(x)) \quad (2)$$

式中:$Output$ 为残差连接模块的最终输出特征;$Activation(\cdot)$ 为激活函数;$F(x)$ 为经过卷积层和激活函数后的特征;$Conv_{1\times1}$ 为 1×1 卷积调整通道数,使残差与主路径维度相匹配。

1.2 Transformer

Transformer 模型是一种基于自注意力机制的深度学习架构,其打破了传统循环神经网络的逐步序列处理限制,通过并行化计算和全局依赖建模能力,显著提升了长时序模式的捕捉效率[19]。Transformer 模型的核心结构由嵌入层、位置编码、编码器堆叠层组成,嵌入层与位置编码将输入序列映射至隐藏空间并与位置信息融合,编码器堆叠层由多头自注意力机制和前馈网络构成,每层中均包含残差连接和层归一化。

1.2.1 多头自注意力机制

多头自注意力机制是 Transformer 模型的核心组件,通过并行多组注意力头捕捉序列中的多样化关系,从而增强模型的表达能力,具体流程如下。

(1) 输入序列的每个位置通过三组线性变换生成查询矩阵 $\boldsymbol{Q}$、键矩阵 $\boldsymbol{K}$、值矩阵 $\boldsymbol{V}$,计算公式如式(3)所示:

$$\begin{cases} \boldsymbol{Q}_i = \boldsymbol{Q}\boldsymbol{W}_i^{Q} \\ \boldsymbol{K}_i = \boldsymbol{K}\boldsymbol{W}_i^{K} \\ \boldsymbol{V}_i = \boldsymbol{V}\boldsymbol{W}_i^{V} \end{cases} \quad (3)$$

式中:$\boldsymbol{W}_i^{Q}, \boldsymbol{W}_i^{K}, \boldsymbol{W}_i^{V} \in \boldsymbol{R}^{d_{\text{model}}\times d_k}$,为可学习的参数矩阵;$i = 1,2,\cdots,h$,$h$ 为注意力头数量。

(2) 将三个矩阵沿特征维度均匀切分成 h 个子头,每个子头在各自的子空间内计算自注意力权重,计算公式如式(4)所示:

$$Head_i = Attention(\boldsymbol{Q}_i, \boldsymbol{K}_i, \boldsymbol{V}_i) = \text{Softmax}\left(\frac{\boldsymbol{Q}_i \boldsymbol{K}_i^{\mathrm{T}}}{\sqrt{d_k}}\right)\boldsymbol{V}_i \quad (4)$$

式中:$Head_i$ 为第 i 个注意力头的输出特征;$Attention(\boldsymbol{Q}_i, \boldsymbol{K}_i, \boldsymbol{V}_i)$ 为第 i 个注意力头的自注意力计算函数,Softmax(·)为激活函数;$\sqrt{d_k}$ 为缩放点积结果,防止梯度消失。

(3) 将 h 个子头的拼接输出后,通过线性变换映射为原始维度,计算公式如式(5)所示:

$$MultiHead = (\boldsymbol{Q}, \boldsymbol{K}, \boldsymbol{V}) = \text{Concat}(Head_1, Head_2, \cdots, Head_h)\boldsymbol{W}^{O} \quad (5)$$

式中:$MultiHead$ 表示多头自注意力模块的最终输出;$Head$ 表示注意力头;Concat 表示拼接操作;$\boldsymbol{W}^{O} \in R^{h\cdot d_k \times d_{\text{model}}}$,为融合权重矩阵。

1.2.2 前馈网络

Transformer 模型中的前馈网络位于多头自注意力模块之后,通过非线性变换进一步增强序列特征的表达能力。其结构包含两个线性层:首先,将输入扩展至更高的中间维度,通过激活函数引入非线性;然后,通过第二层线性层降回原始维度。前馈网络通过参数共享处理所有位置特征,并通过残差连接和层归一化协同工作,缓解梯度消失并稳定训练。前馈网络通过与自注意力机制互补,共同提升了 Transformer 模型对长序列的建模效率。

1.3 阿尔法进化算法

阿尔法进化算法(AE)作为一种新型非隐喻驱动的元启发式算法[20],针对传统算法因过度依赖自然隐喻导致的计算模糊性和参数冗余的问题,提出了两项优化策略:其一,提出基于进化路径的自适应基向量更新策略,利用历史信息累计增强搜索起点的全局相关性,避免局部最优陷阱;其二,融合复合差分操作与衰减因子 α,构建随机与自适应双步长机制,动态平衡搜索与开发能力。阿尔法进化算法的运行流程如下。

在 D 维搜索空间中,对候选解 $\boldsymbol{X}_i$ 进行均匀初始化,确保能够均匀覆盖搜索空间,如式(6)所示:

$$\begin{cases} \boldsymbol{X}_i = b_1 + (b_u - b_1)\cdot \text{rand}(0,1,[1,D]) \\ i = 1,2,\cdots,N \end{cases} \quad (6)$$

式中:b_u 和 b_1 为搜索空间的上限和下限;N 代表解的数量;rand(0,1,[1,D]) 为生成一个包含 D 个均匀分布在[0,1]范围内随机数的向量。

自适应向量 $\boldsymbol{P}$ 通过两种策略进行初始化,如式(7)和式(8)所示:

$$\boldsymbol{P} = \begin{cases} diagonal(\boldsymbol{A}), \text{if rand}(0,1) < 0.5 \\ \boldsymbol{\omega}\boldsymbol{B}, \text{otherwise} \end{cases} \quad (7)$$

$$\omega_{i;k} = \frac{f(X_{i;k})}{\sum_{i=1}^{K} f(X_{i;k})} \quad (8)$$

式中:$diagonal(\boldsymbol{A})$ 表示取矩阵 $\boldsymbol{A}$ 的主对角线元素,构成一个向量;$\boldsymbol{A}$ 为从候选矩阵 $\boldsymbol{X}$ 中进行有放回采样 D 次生成的 D 阶方阵;$\boldsymbol{B}$ 为从候选矩阵 $\boldsymbol{X}$ 中进行无放回采样 K 次生成的 K 行 D 列矩阵;K 为[N · rand(0,1)];$\boldsymbol{\omega}$ 为权重向量;$f(\cdot)$ 为特征映射函数;i;k 表示第 i 行第 k 列。

算法通过阿尔法算子更新解,其核心公式如(9)至式(12)所示:

$$\boldsymbol{E}_i^{t+1} = \boldsymbol{P}' + \alpha\boldsymbol{\Delta}\boldsymbol{r}_i + \boldsymbol{\theta}\cdot(\boldsymbol{W}_i + \boldsymbol{E}_i^{t} - \boldsymbol{P}' - \boldsymbol{L}_i) \quad (9)$$

$$\alpha = \exp\left(\ln\left(1-\frac{FEs}{\max FEs}\right)-\left(\frac{FEs}{\max FEs}\right)^{2}\right) \tag{10}$$

$$\boldsymbol{\Delta r} = (b_u - b_l)\cdot(2R_1\cdot R_2 - R_2)\cdot S \tag{11}$$

$$\boldsymbol{\theta} = \begin{cases} \text{rand}(0,2)\cdot \text{ones}(1,D), \text{if rand}(0,1) < 0.5 \\ \text{rand}(0,1,[1,D]), \text{otherwise} \end{cases} \tag{12}$$

式中：$\boldsymbol{E}_i^{t+1}$ 为进化矩阵 $\boldsymbol{E}$ 中的第 i 个在 $t+1$ 迭代时的进化解；$\boldsymbol{P}'$ 为基向量，决定进化起点；α 为衰减因子，控制算法的探索和利用；$\boldsymbol{\Delta r}_i$ 为第 i 个随机步长；$\boldsymbol{\theta}$ 为控制参数，控制自适应步长；$\boldsymbol{W}_i$ 为采样中得到的优于当前解 $\boldsymbol{E}_i^t$ 的解；$\boldsymbol{L}_i$ 为采样中得么的劣于当前解 $\boldsymbol{E}_i^t$ 的解；FEs 为适应度评估次数；$\max FEs$ 为最大适应度评估次数；R_1、R_2 表示由均匀分布生成的随机量；S 为缩放因子。

为了增加连续迭代步骤之间的相关性，同时为 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 生成两个独立的进化路径 $\boldsymbol{P}_a$ 和 $\boldsymbol{P}_b$，计算公式如(13) 和式(14) 所示。

$$P = \begin{cases} c_a \boldsymbol{P}_a^t + (1-c_a)\cdot diagonal(\boldsymbol{A}), \\ \quad \text{if rand}(0,1) < 0.5 \\ c_b \boldsymbol{P}_b^t + (1-c_b)\cdot \boldsymbol{\omega B}, \text{otherwise} \end{cases} \tag{13}$$

$$c_a = c_b = 1 - \frac{FEs}{\max FEs} \tag{14}$$

式中：c_a 和 c_b 分别为 $\boldsymbol{P}_a$ 和 $\boldsymbol{P}_b$ 的学习率；$\boldsymbol{P}_a^t$、$\boldsymbol{P}_b^t$ 分别为上一代进行路径 $\boldsymbol{P}_a$、$\boldsymbol{P}_b$ 的历史值。

边界通过反射边界处理，通过公式(15) 将越界值拉回边界中点，以平衡探索与开发。

$$\boldsymbol{E}_{i,j} = \begin{cases} \dfrac{\boldsymbol{E}_{i,j}+b_u}{2}, \boldsymbol{E}_{i,j} > b_u \\ \dfrac{\boldsymbol{E}_{i,j}+b_l}{2}, \boldsymbol{E}_{i,j} < b_l \\ \boldsymbol{E}_{i,j}, \text{otherwise} \end{cases} \tag{15}$$

式中：$\boldsymbol{E}_{i,j}$ 为进化解 $\boldsymbol{E}_i$ 的第 j 个维度的值；b_u、b_l 分别为第 j 维决策变量的允许最大值、最小值。

同时采用贪心选择策略，计算公式如式(16) 所示。

$$\boldsymbol{X}_k^{t+1} = \begin{cases} \boldsymbol{E}_i^{t+1}, f(\boldsymbol{E}_i^{t+1}) \leqslant f(\boldsymbol{E}_i^t) \\ \boldsymbol{X}_k^t, \text{otherwise} \end{cases} \tag{16}$$

通过上述操作，阿尔法进化算法实现了基向量自适应、探索与开发平衡及高效搜索的目标。

1.4　时序注意力引导的多模型动态特征融合机制

时序注意力是一种专为序列数据处理设计的注意力机制，其核心原理是通过计算各时间步与序列全局的关联度，自适应调节注意力强度，从而有效突出关键时序信息[21]。作为多尺度特征融合的核心组件，该机制通过对时间步的自适应加权实现关键时序模式的强化捕捉：以 TCN 提取的局部时序特征与 Transformer 建模的全局依赖建模为双输入，通过动态权重模块为每个时间步生成差异化权重参数，依据序列的短期波动或长期趋势特性，自适应调节不同输入之间的贡献比例。同时，通过引入非线性变换层挖掘不同特征间的潜在关联，构建交互权重矩阵以增强特征融合的非线性表达能力，从而避免简单特征叠加引发的信息协同不足问题，其数学形式如下。

首先对 TCN 与 Transformer 输出进行特征维度上的拼接，并使用 MLP 多层感知机对拼接特征进行处理，随后通过 Sigmoid 函数生成时序注意力权重向量，如式(17)所示：

$$\boldsymbol{\delta} = \sigma(\text{MLP}([TCN(x); Transformer(x)])) \tag{17}$$

式中：$\boldsymbol{\delta}$ 为时序注意力权重向量；σ 为 Sigmoid 函数；[;]表示特征拼接；x 表示输入变量。

融合过程中，通过逐元素乘法实现对 TCN 和 Transformer 特征的动态加权，使得模型能够依据输入时序特性灵活调整双分支贡献度，同时引入残差连接，在维持较低计算复杂度的同时，增强特征流动性，如式(18)和式(19)所示：

$$Fused = (1-\boldsymbol{\delta})\otimes TCN(x) + \boldsymbol{\delta}\otimes Transformer(x) \tag{18}$$

$$\hat{y} = LayerNorm(Fused + \lambda\cdot(TCN(x) + Transformer(x))) \tag{19}$$

式中：$Fused$ 为 TCN 与 Transformer 双分支输出的加权组合；$\otimes$表示逐元素乘法；λ 为残差系数，控制原始特征和融合特征之间的平衡；$LayerNorm(\cdot)$ 表示层归一化操作；$\hat{y}$ 表示模型的最终输出。

1.5　AE-TCN-Transformer 并行模型

综合上述方法，将 TCN 模型和 Transformer 模型并行设计，依托 TCN 模型的局部特征提取和 Transformer 模型的全局依赖建模能力，充分发挥两种模型架构的互补优势，实现对时间序列数据的多尺度特征提取。整个模型运行步骤如下：

(1) 对获取的原始位移及应力数据进行异常值剔除、缺失值插补，并进一步对位移数据进行自适应噪声完备集合经验模态分解(CEEMDAN)去噪；

(2) 将处理后数据划分为训练集、验证集、测试集，并对数据进行一阶差分及归一化处理；

(3) 将数据分别输入使用 AE 算法优化超参数后的 TCN 模型和 Transformer 模型，分别捕捉局部时序模式和长期依赖关系；

(4) 将 TCN 模型和 Transformer 模型的输出通过多模型动态特征融合,输出最终预测结果。

AE-TCN-Transformer 并行模型预测流程如图 1 所示。

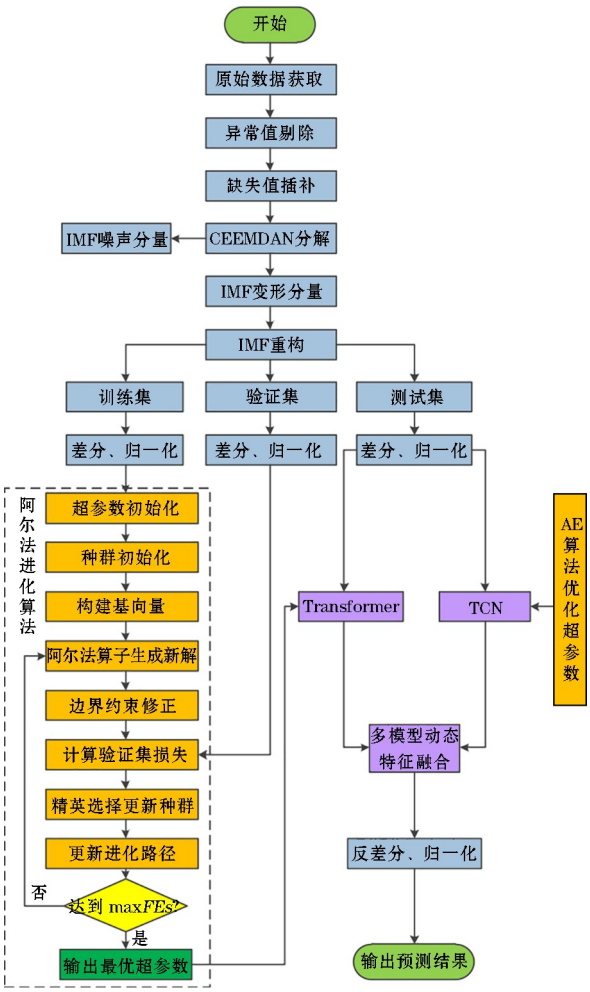


图 1 AE-TCN-Transformer 并行模型预测流程

Fig.1 Prediction flowchart of AE-TCN-Transformer parallel model

2 工程应用

2.1 工程概况

云南某铁矿矿区地处川滇南北向构造带南段,位于青杨断裂与通海山字型构造前缘石屏弧西翼的复合部位,处于断裂带区域,节理裂隙较为发育。区域内岩体主要为沉积-变质的碳酸盐类岩石,揭露岩石以灰岩为主。矿床为以溶蚀裂隙为主的岩溶充水矿床,最低开采标高高于最低侵蚀基准面,地下水补给条件差,巷道围岩含水情况较好,对围岩稳定性无较大影响。选取 1 430 m 中段巷道 W1430-4 监测点为研究对象,采用 NL4410 型位移计和 NL5407 型应力计对围岩变形及应力进行连续监测,数据采集周期覆盖 2023 年 3 月 21 日至 2023 年 9 月 12 日,采样间隔为 1 h,共获取 4 224 组时序数据。受现场机械振动、爆破振动等复杂动态扰动影响,原始监测数据呈现出非平衡不规则波动特征,并伴随有高频噪声和突发干扰,数据质量难以满足后续分析要求,因此需要对原始数据进行预处理。

2.2 原始数据处理

针对原始监测数据中存在缺失值、异常值及高频噪声干扰的问题,首先采用箱型图法识别并剔除异常数据,然后利用三次样条插值法填补缺失值,以确保时序完整性,最后基于 CEEMDAN 分解重构剔除噪声成分。

2.2.1 异常值剔除

通过对现场位移数据分析发现,监测数据中存在离散型异常值,因此采用箱型图法对异常值进行剔除。对数据计算第一四分位数和第三四分位数,确定四分位距和上下限,并对箱型图法筛选出的异常值进行直接剔除。箱型图法的计算公式如式(20)至式(22)所示。

$$IQR = Q_3 - Q_1 \tag{20}$$

$$B_L = Q_1 - k \times IQR \tag{21}$$

$$B_U = Q_3 + k \times IQR \tag{22}$$

式中:Q_1 为数据从小到大排列后 25%位置的值;Q_3 为数据从小到大排列后 75%位置的值;IQR 为数据的四分位距;k 为默认系数,本次计算取值为 $k=3$;B_L、B_U 分别为箱型图法判定的异常值下限和上限。W1430-4 监测点位移数据箱型图如图 2 所示。

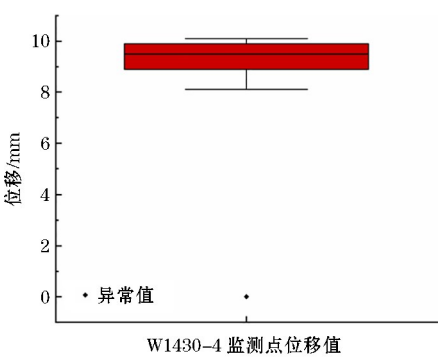


图 2 W1430-4 监测点数据箱型图

Fig.2 Box diagram of the data from the W1430-4 monitoring point

2.2.2 缺失值插补

三次样条插值法通过构造分段三次多项式 $S(x)$ 实现插值,其特点是保证了 $S(x)$ 以及其一阶和二阶导数在节点数的连续性,其具体过程如下:

在区间$[x_i, x_{i+1}]$上，三次样条 $S_i(x)$函数如式(23)所示：

$$S_i(x)=y_i+[\frac{y_{i+1}-y_i}{h_i}-\frac{h_i}{6}(2M_i+M_{i+1})](x-x_i)+\frac{M_i}{2}(x-x_i)^2+\frac{M_{i+1}-M_i}{6h_i}(x-x_i)^3 \quad (23)$$

式中：h_i 为区间步长，$h_i=x_{i+1}-x_i$；x_i、y_i 为第 i 个节点的横、纵坐标；M_i 为节点 x_i 处的二阶导数，$M_i=S''(x_i)$。

为保证二阶导数连续，需满足三对角方程组，如式(24)所示：

$$h_{i-1}M_{i-1}+2(h_{i-1}+h_i)M_i+h_iM_{i+1}=6(\frac{y_{i+1}-y_i}{h_i}-\frac{y_i-y_{i-1}}{h_{i-1}}) \quad (24)$$

对于自然边界条件，令其二阶导数 $M_0=M_n=0$；若已知端点一阶导数 f'_0 和 f'_n，则采用固定边界条件，如式(25)和式(26)所示：

$$2h_0M_0+h_0M_1=6(\frac{y_1-y_0}{h_0}-f'_0) \quad (25)$$

$$h_{n-1}M_{n-1}+2h_{n-1}M_n=6(f'_n-\frac{y_n-y_{n-1}}{h_{n-1}}) \quad (26)$$

联立上述公式，利用追赶法求解二阶导数序列，可求得各分段多项式系数，从而实现插值曲线整体光滑并避免产生龙格现象。

2.2.3　数据去噪

通过 CEEMDAN 对原始数据进行分解，剔除噪声分量后进行重构，CEEMDAN 是由经验模态分解(EMD)发展而来，其分解时首先对原始序列添加不同自适应高斯白噪声以抑制 EMD 算法的模态混叠现象，再将原始时间序列分解为不同频率的子序列，即本征模态分量(IMF)，具体分解过程如下。

向原始数据 $S(t)$ 中加入不同自适应的高斯白噪声 $\omega_i(t)$，得到添加噪声的数据 $S_i(t)$，如式(27)所示：

$$S_i(t)=S(t)+\varepsilon\omega_i(t) \quad (27)$$

式中，ε 为噪声强度系数。

对加入的高斯白噪声进行 EMD 分解，得到 IMF_1 分量，如式(28) 所示：

$$r_1(t)=S_1(t)-IMF_1 \quad (28)$$

式中，$r_1(t)$ 为首次分解后的残差。

重复以上步骤，依次得到第 i 个残余量及第 $i-1$ 个 IMF 分量，如式(29) 所示：

$$r_i(t)=S_{i-1}(t)-IMF_i \quad (29)$$

当残余量极值点不多于 2 个时，停止分解，得到最终残余量 $R(t)$ 和 i 个 IMF 分量，如式(30) 所示：

$$R(t)=S(t)-\sum_{i=1}^{i}IMF_i \quad (30)$$

CEEMDAN 分解参数设置如下：噪声试验次数为 100，噪声添加的标准差为 0.1，本征模态分量分解上限为 10，CEEMDAN 分解停止阈值为 0.005，CEEMDAN 分解后得到 9 个 IMF 分量，各分量如图 3 所示。

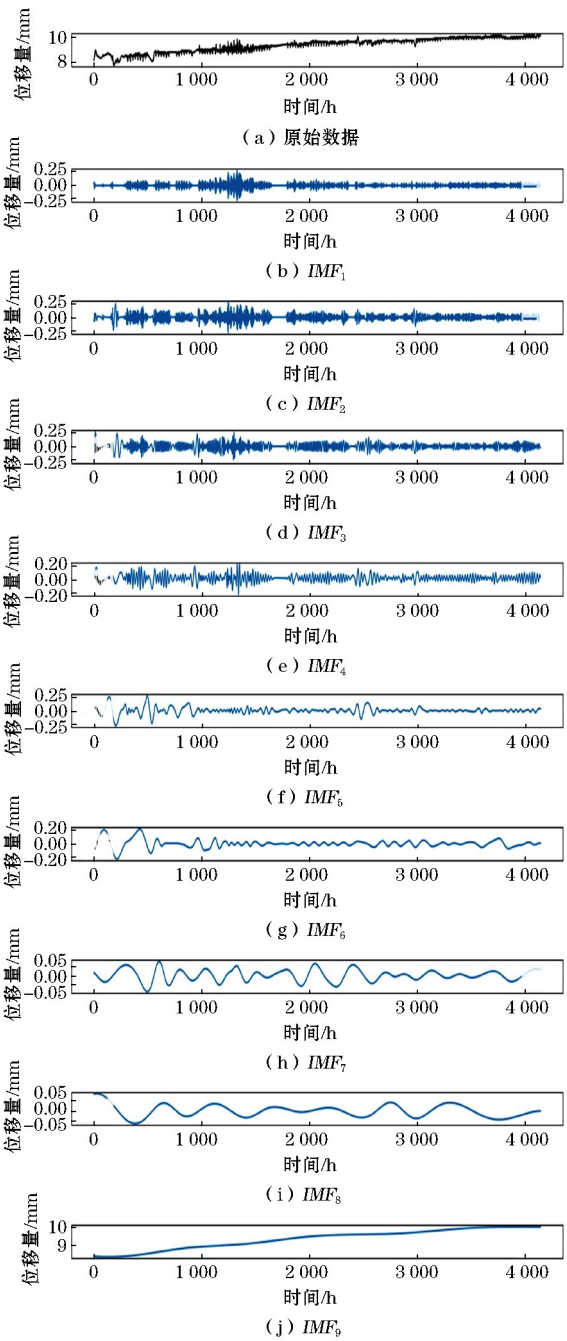


图 3　位移监测数据 CEEMDAN 分解结果

Fig.3　Decomposition results of displacement monitoring data by CEEMDAN

为剔除无关噪声数据,通过各 *IMF* 分量的样本熵、峭度及皮尔逊相关系数开展噪声识别。结果显示:IMF_1～IMF_3 分量的样本熵为 0.262,0.531,0.602,皮尔逊相关系数为 0.048,0.077,0.085,均低于 0.1。其中 IMF_2、IMF_3 符合随机噪声特性,对 IMF_1 分量进一步分析,其峭度高达 10.65,符合高频噪声特性。因此,认定 IMF_1～IMF_3 为噪声分量并将其直接剔除,将其余分量重构得到去噪后的位移数据,如图 4 所示。

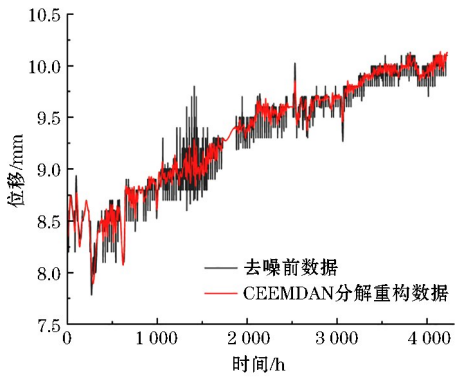


图 4 W1430-4 位移数据处理结果

Fig.4 Processing results of W1430-4 displacement data

由于应力数据变化具有更强的瞬时性和突发性,其信号中可能包含与位移变化密切相关的短时高频成分,为防止数据中关键特征损失,不对应力数据进行去噪处理,处理后的应力数据如图 5 所示。

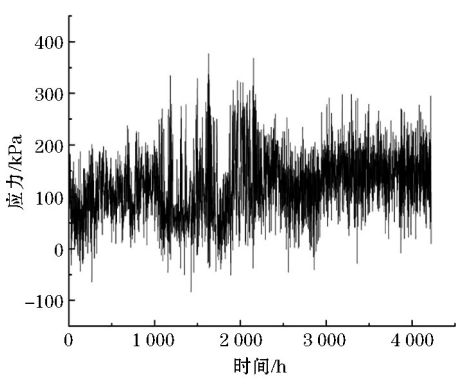


图 5 W1430-4 应力数据处理结果

Fig.5 Processing results of W1430-4 stress data

将处理后的应力及位移数据按照 8∶1∶1 的比例划分为训练集、测试集、验证集,进行一阶差分和 min-max 归一化处理后输入 AE-TCN-Transformer 并行模型中进行预测。

2.3 超参数优化

使用阿尔法进化算法(AE)分别对 TCN 和 Transformer 模型的超参数进行寻优,AE 参数预设如下:种群规模为 50,搜索维度为 3,最大迭代次数为 100。通过 AE 算法对 TCN 的学习率、批量大小和通道数进行寻优,其中学习率范围为[0.000 01,0.001],批量大小范围为[16,128],通道数范围为[16,64]。同时,使用 AE 算法对 Transformer 的学习率、d_model 和编码器层数进行寻优,学习率范围为[0.000 01,0.000 1],d_model 范围为[16,256],编码器层数范围为[2,6]。经优化后的各模型超参数见表 1。

表 1 模型超参数

Table 1 Hyper parameters of the models

模型	训练参数	参数设置
TCN	训练轮数	200
	滑动窗口	48
	Dropout 率	0.4
	卷积核尺寸	3
	优化器	AdamW
	激活函数	ReLU
	学习率	0.001
	批量大小	26
	通道数	[16,16]
	损失函数	Huber
Transformer	训练轮数	200
	滑动窗口	48
	Dropout 率	0.2
	注意力头数	4
	前馈网络维度	128
	批量大小	32
	优化器	AdamW
	学习率	5.3069×10^{-5}
	d_model	64
	编码器层数	2
	损失函数	Huber

为验证 AE 算法在超参数优化中的有效性,选取常用的鲸鱼优化算法(WOA)作为对照,分别在 TCN 和 Transformer 模型中进行超参数寻优对比试验,结果见表 2。由表 2 可知,AE 优化后的模型在数据验证集的损失上均显著优于 WOA,表明 AE 的进化路径动态调整策略和双步长机制有效平衡了全局搜索和局部开发能力,避免了传统算法易陷入局部最优的缺陷。

表 2 超参数优化结果

Table 2 Optimization results of hyper parameters

优化算法	不同模型超参数优化结果	
	TCN 模型	Transformer 模型
AE	$3.528\ 3\times10^{-4}$	$8.237\ 2\times10^{-4}$
WOA	$1.822\ 5\times10^{-3}$	$4.296\ 4\times10^{-3}$

3　试验结果及分析

为验证 AE-TCN-Transformer 并行模型的稳定性及泛化性能，以均方根误差（*RMSE*）、决定系数（R^2）、对称平均绝对百分比误差（*SMAPE*）作为评价指标，将该模型与 TCN 模型、Transformer 模型和 LSTM 模型在 W1430-4、W1480-7、W1530-1 三个不同中段监测点及不同预测步长条件下进行对比分析。

SMAPE 通过修正传统平均绝对百分比误差（*MAPE*）的分母不对称性，将分母替换为真实值和预测值的绝对值和均值，从而均衡正负偏差的惩罚强度，避免模型因系统性低估或高估产生指标失真，确保误差评估的公平性和方向无关性[22]。

RMSE、R^2 和 *SMAPE* 指标计算公式如式(31)至式(33)所示：

$$RMSE=\sqrt{\frac{1}{N}\sum_{i=1}^{N}[Y_i-\hat{Y}_i]^2} \tag{31}$$

$$R^2=1-\frac{\sum_{i=1}^{N}[\hat{Y}_i-Y_i]^2}{\sum_{i=1}^{N}[\bar{Y}_i-Y_i]^2} \tag{32}$$

$$SMAPE=\frac{100\%}{N}\sum_{i=1}^{N}\frac{|\hat{Y}_i-Y_i|}{(|\hat{Y}_i|+|Y_i|)/2} \tag{33}$$

式中：Y_i 为实际值；$\hat{Y}_i$ 为模型预测值；N 为样本数量。

RMSE、*SMAPE* 指标越小，表明建立的神经网络模型对数据的预测精度越高；R^2 越接近 1，说明该模型对数据的拟合效果更好。

监测点 W1430-4、W1480-7 及 W1530-1 的预测结果评价指标见表 3。在预测步长为 6 和 12 时，相较于 Transformer 模型和 LSTM 模型，单一 TCN 模型展现出显著优势，在预测步长为 6 时，其 *RMSE* 降低 6.72%～40.94%，*SMAPE* 降低 6.05%～51.75%；预测步长为 12 时，其 *RMSE* 降低 0.83%～46.15%，*SMAPE* 降低 4.93%～46.24%，充分体现出 TCN 模型的膨胀因果卷积结构在局部特征提取方面的优势。随着预测步长的增加，数据集的长期时序关系逐渐显著，在预测步长为 48 时，Transformer 模型展现出更好的预测效果，相较于 TCN 模型和 LSTM 模型，Transformer 模型的 *RMSE* 降低 10.47%～33.49%，*SMAPE* 降低 13.55%～43.20%，验证了该模型在长期预测中捕捉全局关系的能力。而 LSTM 模型受序列迭代处理机制限制，长程捕捉效率较低，在当前预测任务中的预测性能较差。

表 3　模型评价指标对比

Table 3　Comparison of model evaluation indicators

预测模型	预测步长	监测点 W1430-4			监测点 W1480-7			监测点 W1530-1		
		RMSE/mm	*SMAPE*/%	R^2	*RMSE*/mm	*SMAPE*/%	R^2	*RMSE*/mm	*SMAPE*/%	R^2
AE-TCN-Transformer	6	0.006 4	0.027 1	0.992 0	0.008 3	0.059 0	0.990 2	0.010 5	0.016 6	0.993 9
	12	0.008 0	0.032 2	0.987 1	0.010 6	0.079 2	0.983 9	0.011 2	0.017 3	0.992 8
	24	0.008 6	0.035 0	0.985 0	0.012 0	0.080 8	0.978 4	0.012 8	0.022 3	0.990 1
	48	0.008 7	0.034 9	0.984 7	0.012 8	0.089 3	0.973 1	0.013 2	0.022 7	0.988 4
TCN	6	0.007 5	0.031 9	0.989 0	0.010 7	0.080 7	0.984 0	0.011 1	0.017 9	0.993 1
	12	0.008 4	0.034 2	0.986 1	0.013 0	0.086 7	0.975 8	0.012 0	0.019 3	0.991 8
	24	0.010 1	0.042 3	0.979 3	0.013 9	0.091 2	0.970 8	0.014 7	0.030 5	0.987 1
	48	0.014 4	0.057 9	0.957 5	0.017 2	0.128 4	0.951 9	0.016 1	0.034 3	0.982 8
Transformer	6	0.010 8	0.046 2	0.977 4	0.011 8	0.085 9	0.980 5	0.011 9	0.019 8	0.992 1
	12	0.011 6	0.046 2	0.973 4	0.014 0	0.109 6	0.972 1	0.012 1	0.020 3	0.991 2
	24	0.011 7	0.047 5	0.972 2	0.014 9	0.100 0	0.966 3	0.013 5	0.024 3	0.989 2
	48	0.012 2	0.049 4	0.969 6	0.015 4	0.111 0	0.961 2	0.014 1	0.024 2	0.986 8
LSTM	6	0.012 7	0.051 1	0.968 8	0.014 9	0.113 5	0.968 9	0.017 1	0.037 1	0.983 6
	12	0.015 6	0.063 2	0.951 7	0.016 5	0.113 1	0.961 1	0.018 5	0.035 9	0.980 5
	24	0.016 5	0.065 3	0.945 0	0.017 1	0.126 7	0.955 9	0.020 5	0.044 1	0.975 0
	48	0.016 7	0.068 1	0.942 9	0.018 4	0.135 3	0.944 5	0.021 2	0.042 6	0.970 1

AE-TCN-Transformer 模型在三个监测点及不同步长的预测任务中均展现出显著优势，*RMSE*、*SMAPE*、R^2 指标均优于基准模型。与 TCN、Transformer、LSTM 模型相比，AE-TCN-Transformer 模型的 *RMSE* 降低 5.41%～49.6%，*SMAPE* 降低 7.26%～55.26%，且 R^2 均保持在 0.97 以上，充分验证了融合 TCN 局部特征与 Transformer 全局依赖建模的协同效应。同时，随着预测步长由 6 增加至 48，AE-TCN-Transformer 模型的抗衰减特性显著，其 *RMSE* 较 TCN 模型平

均增幅为52.19%、*SMAPE* 平均增幅为43.57%,有效缓解了长期预测中的误差累积问题。综上,AE-TCN-Transformer模型在各监测点的预测任务中均取得了最优结果,进一步证实了该模型良好的泛化性能。

4 结论

本研究针对矿山巷道变形具备多尺度非线性的特征,提出了一种基于阿尔法进化算法优化的AE-TCN-Transformer并行巷道预测模型,并通过云南某铁矿巷道位移监测数据验证其性能,得出以下结论。

(1) AE-TCN-Transformer并行模型通过并行分支设计及多模型动态特征融合机制,有效融合了TCN的膨胀因果卷积局部特征提取能力与Transformer多头自注意力机制的全局依赖建模优势。试验结果表明,该架构在预测任务中优于单一模型,验证了多尺度特征融合模型对复杂时序数据中非线性规律的高效捕捉能力。

(2) 引入阿尔法进化算法自适应优化模型超参数,通过进化路径动态调整与双步长机制,克服了传统启发式优化算法易陷入局部最优的缺陷,使本文模型在预测稳定性和泛化性能获得更优表现。

(3) 在云南某铁矿W1430-4、W1480-7、W1530-1三个不同中段监测点实测数据中,AE-TCN-Transformer并行模型在预测步长为6~48步时均保持高预测精度,其中,*RMSE* 均低于0.013 2 mm,*SMAPE* 均低于0.09%,R^2 均高于0.97,验证了该模型在矿山巷道变形预测中的有效性和实用性。

参考文献(References):

[1] 崔建军.深部沿空掘巷变形破坏机理及控制技术研究[J].煤炭科学技术,2017,45(7):12-17.
CUI Jianjun. Study on deformation and failure mechanism and control technology of deeply gob-side entry driving[J]. Coal Science and Technology,2017,45(7):12-17.

[2] 马斌文,谢和平,张修峰,等.动载扰动下巷道围岩冲击破坏与能量释放规律研究[J].采矿与岩层控制工程学报,2024,6(4):5-22.
MA Binwen, XIE Heping, ZHANG Xiufeng, et al. Study on the impact failure and energy release of surrounding rock of roadway under dynamic disturbance[J]. Journal of Mining and Strata Control Engineering,2024,6(4):5-22.

[3] 赵卫东,张德飞,王雨,等.节理岩体中双巷道稳定性的边界元分析[J].煤炭科学技术,2020,48(增刊2):118-123.
ZHAO Weidong, ZHANG Defei, WANG Yu, et al. Boundary element analysis of double roadway in jointed rock mass author[J]. Coal Science and Technology, 2020, 48(Suppl.2):118-123.

[4] 芦楠楠,陈永珩,马占国,等.深部地下巷道围岩结构监测信息化技术研究[J].工矿自动化,2019,45(6):1-9.
LU Nannan, CHEN Yongheng, MA Zhanguo, et al. Research on informatization technology for roadway surrounding rock structure monitoring in deep underground [J]. Industry and Mine Automation,2019,45(6):1-9.

[5] 李俊平,胡文强,张浩.断层切割围岩的巷道稳定性数值分析[J].地下空间与工程学报,2018,14(增刊2):925-929+979.
LI Junping, HU Wenqiang, ZHANG Hao. Numerical analysis of roadway stability by fault cutting surrounding rock[J]. Chinese Journal of Underground Space and Engineering,2018,14(Suppl.2):925-929+979.

[6] 张小军,孙佳瑞,马扬.近距离煤层采掘关系对下位巷道围岩变形规律影响研究[J].工矿自动化,2024,50(8):61-68.
ZHANG Xiaojun, SUN Jiarui, MA Yang. Study on the influence of close range coal seam mining relationship on the deformation law of surrounding rock in lower roadway[J]. Journal of Mine Automation,2024,50(8):61-68.

[7] 罗明坤,宋丽强,李思齐,等.动力扰动条件下巷道围岩变形规律研究[J].煤矿安全,2025,56(3):157-166.
LUO Mingkun, SONG Liqiang, LI Siqi, et al. Study on deformation law of roadway surrounding rock under dynamic disturbance[J]. Safety in Coal Mines,2025,56(3):157-166.

[8] 张智敏,康天合.基于弹性地基梁理论的巷道顶板变形预测[J].矿业研究与开发,2021,41(10):104-108.
ZHANG Zhimin, KANG Tianhe. Roof deformation prediction of roadway based on elastic foundation beam theory[J]. Mining Research and Development, 2021, 41(10):104-108.

[9] 杨家俊,陈星明.Verhulst模型在巷道围岩变形预测中的应用[J].化工矿物与加工,2017,46(12):41-43+49.
YANG Jiajun, CHEN Xingming. Application of Verhulst model inprediction of deformation of surrounding rock in roadway[J]. Industrial Minerals and Processing, 2017, 46(12):41-43+49.

[10] 邓康宇,薛欣然,诸葛连兵,等.基于非等间距GM(1,1)模型群对深井巷道围岩变形的预测[J].现代隧道技术,2018,55(2):91-95.
DENG Kangyu, XUE Xinran, ZHUGE Lianbing, et al. Deformation prediction for surrounding rock of a deep roadway based on non-equidistant GM(1,1) model group [J]. Modern Tunnelling Technology,2018,55(2):91-95.

[11] 张斌,贾海宾,李阿涛,等.基于贝叶斯更新的煤矿回采巷道采动变形预测方法[J].隧道与地下工程灾害防治,2025,7(1):57-67.
ZHANG Bin, JIA Haibin, LI Atao, et al. Prediction method for mining-induced deformation in coal mine roadways based on Bayesian updating[J]. Hazard Control in Tunnelling and Underground Engineering, 2025, 7(1): 57-67.

[12] 陈攀，马鑫民，向俊杰，等.基于随机森林算法的煤巷顶板位移预测与应用[J].煤炭工程，2024，56(7)：127-135.
CHEN Pan，MA Xinmin，XIANG Junjie，et al. Roof displacement prediction of coal roadways based on random forest algorithm and the application[J]. Coal Engineering，2024，56(7)：127-135.

[13] 蔡舒凌，李二兵，陈亮，等.基于 FA-NAR 动态神经网络的隧洞围岩变形时序预测研究[J].岩石力学与工程学报，2019，38(增刊 2)：3346-3353.
CAI Shuling，LI Erbing，CHEN Liang，et al. The time series prediction of tunnel surrounding rock deformation based on FA-NAR dynamic neural network[J]. Chinese Journal of Rock Mechanics and Engineering，2019，38(Suppl. 2)：3346-3353.

[14] XIE Chengyu，NGUYEN Hoang，BUI Xuan Nam，et al. Predicting roof displacement of roadways in underground coal mines using adaptive neuro-fuzzyinference system optimized by various physics-based optimization algorithms[J]. Journal of Rock Mechanics and Geotechnical Engineering，2021，13(6)：1452-1465.

[15] 吴泽鑫，张成良，张华超，等.基于 WOA 优化神经网络的斜坡道拱顶沉降预测研究[J].有色金属工程，2024，14(4)：150-160＋174.
WU Zexin，ZHANG Chengliang，ZHANG Huachao，et al. Study optimization of neural network based on whale algorithm for prediction of vault subsidence in slope[J]. Nonferrous Metals Engineering，2024，14(4)：150-160＋174.

[16] 崔靖奇，吴顺川，程海勇，等.滇中引水软岩隧洞围岩位移时序预测[J].清华大学学报(自然科学版)，2024，64(7)：1215-1225.
CUI Jingqi，WU Shunchuan，CHENG Haiyong，et al. Time series prediction of the surrounding rock displacement of a soft rock tunnel in the Central Yunnan Water Diversion Project[J]. Journal of Tsinghua University (Science and Technology)，2024，64(7)：1215-1225.

[17] 刘辉，李国强，朱晓峻，等.基于深度学习的井筒变形预测模型与应用[J].煤炭学报，2025，50(2)：732-747.
LIU Hui，LI Guoqiang，ZHU Xiaojun，et al. Exploration and application of deep learning based wellbore de-formation forecasting mode[J]. Journal of China Coal Society，2025，50(2)：732-747.

[18] LIU Zaobao，WANG Yongchen，LI Yong，et al. Realtime prediction of hard rock TBM advance rate using temporal convolutional network (TCN) with tunnel construction big data[J]. Frontiers of Structural and Civil Engineering，2022，16(4)：401-413.

[19] WU Xiaofeng，YANG Song，ZHANG Di，et al. Transformer based neuralnetwork for daily ground settlement prediction of foundation pit considering spatial correlation[J]. PloS One，2023，18(11)：e0294501.

[20] GAO Hao，ZHANG Qingke. Alpha evolution：an efficient evolutionary algorithm with evolution path adaptation and matrix generation[J]. Engineering Applications of Artificial Intelligence，2024，137：109202.

[21] SHIH Shunyao，SUN Fankeng，LEE Hungyi. Temporal pattern attention formultivariate time series forecasting[J]. Machine Learning，2019，108：1421-1441.

[22] LIU Liu，ZHANG Jianqiu，XUE Shibei. Photovoltaic power forecasting：using wavelet threshold denoising combined with VMD[J]. Renewable Energy，2025，249：123152.