

DOI: 10.7672/sgjs2026020119

# 改进 LSTM 沉降预测模型在既有线铁路 结构沉降预测中的应用\*

严申华<sup>1</sup>, 张家奇<sup>1</sup>, 张拯<sup>2</sup>, 褚韶波<sup>1</sup>, 钱华<sup>1</sup>, 陈愿<sup>1</sup>

(1. 中国铁路上海局集团有限公司杭州铁路枢纽工程建设指挥部, 浙江 杭州 310000;

2. 成都西南交大技术转移中心有限公司, 四川 成都 610036)

[摘要] 以沪昆铁路义乌高架站房建设工程杭长道工程为例, 构建了新型耦合残差连接与多头注意力机制 LSTM, 即既有线铁路结构变形预测网络(Res-MHA-LSTM), 实现了准确地既有线铁路结构沉降预测。对比不同测点下的 5 种常见时序预测网络的预测结果趋势和取值, 并通过多种指标对结果进行了量化分析, 验证了 Res-MHA-LSTM 网络在多个测点数据下的预测准确性和可行性。在 HC-2 测点数据下, Res-MHA-LSTM 方法在模型预测结果的平均绝对误差、均方根误差和平均绝对百分比误差分别是 0.042, 0.079, 0.45, 相比 LSTM 预测模型, 分别降低了 23.09%, 23.51% 和 87.69%。决定系数为 0.91, 提升了 7.4%, 相对其他试验模型也有提升。

[关键词] 铁路; 沉降; 变形; 预测

[中图分类号] U213.2

[文献标识码] A

[文章编号] 2097-0897(2026)02-0119-07

## Application of Improved LSTM Settlement Prediction Model in Structural Settlement Prediction of Existing Railway Line

YAN Shenhua<sup>1</sup>, ZHANG Jiaqi<sup>1</sup>, ZHANG Zheng<sup>2</sup>, CHU Shaobo<sup>1</sup>, QIAN Hua<sup>1</sup>, CHEN Yuan<sup>1</sup>

(1. China Railway Shanghai Group Co., Ltd. Hangzhou Railway Hub Project Construction Command Headquarters,

Hangzhou, Zhejiang 310000, China; 2. Chengdu Southwest Jiaotong University Technology Transfer Center

Co., Ltd., Chengdu, Sichuan 610036, China)

**Abstract:** Taking the Hangchang Road project of Yiwu Elevated Station building on Shanghai-Kunming Railway as an example, a new type of coupled residual connection and multi head attention mechanism LSTM, namely the Res-MHA-LSTM, was constructed to accurately predict the structural settlement of existing railway lines. Comparing the prediction results trends and values of five common time-series prediction networks under different measurement points, and quantitatively analyzing the results through multiple indicators, the accuracy and feasibility of the Res-MHA-LSTM network in predicting data from multiple measurement points were verified. Under the HC-2 measurement point data, the average absolute error, root mean square error, and average absolute percentage error of the Res-MHA-LSTM method in the model prediction results were 0.042, 0.079 and 0.45, respectively. Compared with the LSTM prediction model, they decreased by 23.09%, 23.51% and 87.69%, respectively. The coefficient of determination is 0.91, an increase of 7.4%, which is also an improvement compared to other experimental models.

**Keywords:** railroads; settlement; deformation; prediction

## 0 引言

目前针对铁路变形的研究方法大致为有限元仿真模型变形预测法和基于机器学习的变形预测法。Tucho 等<sup>[1]</sup>研究了运动轮载荷对轨道动力响应的影响, 以此评估钢轨下部结构的应力-变形关系。

\* 中国铁路上海局集团有限公司科技项目: 铁路高架站房改造多源扰动下既有线铁路设施设备及钢混基础结构智能监测预警研究 (E2024120)

[作者简介] 严申华, 工程师, E-mail: 2822174313@qq.com

[收稿日期] 2025-09-05

Varandas 等<sup>[2]</sup>考虑三维应力条件与三维模型,估计了道砟和底砟的长期变形。Chawla 等<sup>[3]</sup>研究了软路基加固铁路的变形和强度。Chen<sup>[4]</sup>提出了精确的列车-轨道-桥梁动力学模型和桥梁长期变形的确定方法,并对桥梁长期变形和列车运行的纵向连通轨道进行了评价。Zambrano 等<sup>[5]</sup>研究了轮轨相互作用引起严重塑性变形后,铁路钢轨的失效情况。基于有限元仿真模型变形预测方法结合了真实的边界条件,具有良好的物理解释性。但在实际建模中,边界条件和物理模型往往会得到简化,有限元模型需要真实数据进行修正,无法达到变形预测的要求,并且需要极大的计算资源。

基于机器学习的变形预测方法通过训练时间序列与未来变形量的映射关系,可利用历史数据预测未来的变形量。Xing 等<sup>[6]</sup>利用支持向量机(SVM)实现水坝的变形预测。Wan 等<sup>[7]</sup>和 Zhou 等<sup>[8]</sup>使用粒子群算法(PSO)结合 BP 神经网络模型分别实现了不同合金的变形。Zhang 等<sup>[9]</sup>使用水循环算法优化 BP 神经网络,实现了大坝的变形预测。Lin 等<sup>[10]</sup>利用麻雀优化算法优化卷积神经网络(CNN)与门控循环单元(GRU),实现了变形的时序预测。长短期记忆神经网络(LSTM)通过引入记忆单元和门控机制在变形序列预测应用中取得了极大的发展。夏超鹏等<sup>[11]</sup>基于 LSTM 网络对电力负荷数据进行了预测。谭茜元等<sup>[12]</sup>利用灰狼优化算法进一步优化了算法的参数选择。基于启发式算法有助于模型达到最佳的泛化能力,但因 LSTM 难以有效捕捉长时间序列中的长程依赖关系,从而造成了模型的准确率下降。Chen 等<sup>[13]</sup>利用人工神经网络(ANN)和 LSTM 结合实现了路基土体的不均匀冻胀预测。Wang 等<sup>[14]</sup>结合 CNN 和 LSTM 预测了铁路轨道几何形状。袁利伟等<sup>[15]</sup>结合 CNN 和 LSTM 模型在尾矿坝浸润线的预测上取得了不错的效果。这些研究利用卷积神经网络对特征进行初步提取,使 LSTM 可以识别更易学习的特征,从而取得更好的预测效果,但卷积神经网络只能提取局部

特征信息,无法有效关联上下文之间的联系,并且无法忽视多层连接的网络退化问题。Transformer 利用自注意力机制,可以更有效地捕捉长距离依赖关系,并且具有更好的并行化能力<sup>[16]</sup>。Wang 等<sup>[17]</sup>利用 Transformer 模型实现了准确的时序预测。周康敏等<sup>[18]</sup>将残差连接融入 LSTM,提升深层网络训练的可行性。

在铁路沉降监测中,由于长期作业、不定期检修及不确定车流等因素的影响,沉降监测数据往往表现出局部非线性。为提升预测的准确性和稳定性,增强监测的可靠性,本文提出了改进 LSTM 沉降预测模型(Res-MHA-LSTM),用于实现铁路变形的精准预测。该模型通过残差连接模块对特征进行初步提取,随后利用多头注意力机制分配注意力权重,最后通过 LSTM 网络对经过处理后的特征进行深入学习,以更好地拟合铁路变形的时间序列关系。

## 1 Res-MHA-LSTM 模型概况

### 1.1 整体模型

本文提出了基于 Res-MHA-LSTM 网络的既有有线铁路结构变形预测方法,其整体流程如图 1 所示。其中提出的 Res-MHA-LSTM 网络主要由残差特征粗提取模块、特征连接多头注意力机制和时序映射拟合网络组成,主要作用如下。

1) 残差特征粗提取模块。对需要的特征进行初步提取,将其转换成更易学习的特征,并利用其包含的跳跃连接操作,使梯度在长时间序列的训练过程中能够更有效地流动,从而提高 LSTM 训练的稳定性。

2) 特征连接多头注意力机制。动态地对残差特征粗提取后的特征分配注意力权重,提取重要的时序信息,并允许模型在整个序列中直接交互信息。

3) 时序映射拟合网络。直接利用处理后的特征,可更有效地捕捉时间序列的动态变化。

### 1.2 残差特征粗提取模块

残差特征粗提取模块主要由 3 层残差连接块组成,相较于传统卷积神经网络(CNN),残差连接

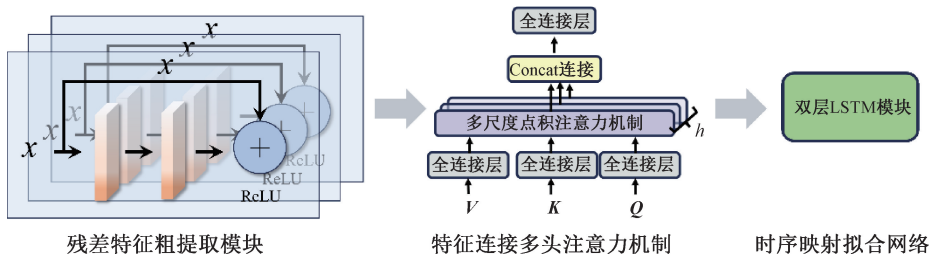


图 1 基于 Res-MHA-LSTM 网络的既有有线铁路结构变形预测流程

Fig. 1 Process of structural deformation prediction of existing line railroad based on Res-MHA-LSTM network

(residual connection)在特征提取中的优势在于更好的特征表达能力。对于单层的残差连接,假设  $x$  是该层的输入,  $H(x)$  是该层经过系列变换后的输出,残差  $F(x)$  连接的表达式为:

$$F(x) = H(x) - x \quad (1)$$

因此经由该层得到的输出可以表示为:

$$H(x) = F(x) + x \quad (2)$$

通过残差连接块,网络的实际学习任务变为对残差  $F(x)$  的学习,理论上如果网络的目标函数是  $H(x)$ ,那么网络学习的对象由  $H(x)$  的最佳参数变为  $F(x)$ 。当  $F(x) = 0$  时,  $H(x) = 0$ ,这意味着如果某些层不需要学习复杂的特征,它们可以直接将输入传递到输出,从而避免梯度消失问题,同时保留输入信息。

残差连接的另一个优势在于梯度求解时,梯度的传播不再完全依赖于  $\partial H(x)/\partial x$ ,而是直接依赖于  $\partial F(x)/\partial \theta$  与  $\partial x/\partial \theta$  的和,其反向传播的表示式为:

$$\frac{\partial L}{\partial \theta} = \frac{\partial L}{\partial H} \cdot \frac{\partial H}{\partial \theta} = \frac{\partial L}{\partial H} \cdot \left( \frac{\partial F}{\partial \theta} + \frac{\partial x}{\partial \theta} \right) \quad (3)$$

式中:  $L$  代表损失函数;  $\theta$  代表该层的模型参数。因此梯度可通过残差路径直接流动,减少了梯度消失的风险。

### 1.3 特征连接多头注意力机制

多头注意力机制是现代深度学习模型,特别是 Transformer 架构中的核心<sup>[19]</sup>,它允许模型在多个不同的子空间中同时关注输入序列的不同部分。假设给予模型输入序列  $\mathbf{X}$ ,其中  $\mathbf{X}$  是  $n \times d$  矩阵( $n$  代表序列长度,  $d$  代表序列的维度)。

对于位置  $i$  和查询  $Q_i$  的特征向量,需要计算该位置与其他所有位置相关性。首先,定义 3 个矩阵:查询矩阵  $\mathbf{Q}$ 、键矩阵  $\mathbf{K}$  和值矩阵  $\mathbf{V}$ 。这些矩阵是从输入序列  $\mathbf{X}$  通过线性变换得到,具体计算公式为:

$$\begin{aligned} \mathbf{Q}_i &= \mathbf{X}\mathbf{W}_{Q_i} \\ \mathbf{K}_i &= \mathbf{X}\mathbf{W}_{K_i} \\ \mathbf{V}_i &= \mathbf{X}\mathbf{W}_{V_i} \end{aligned} \quad (4)$$

式中:  $\mathbf{W}_{Q_i}$ ,  $\mathbf{W}_{K_i}$ ,  $\mathbf{W}_{V_i}$  为权重矩阵,分别对应查询、键和值;  $i$  为第  $i$  个头。

注意力分数  $AS$  (即相似度分数)具体表达式为:

$$AS = \frac{\mathbf{Q}_i \mathbf{K}_i^T}{\sqrt{d_k}} \quad (5)$$

式中:  $d_k$  为键的维度。

进一步对注意力分数  $AS$  应用 softmax 函数将其

转换为概率分布,使每个位置的注意力权重介于 0 和 1 之间,得到注意力权重  $AW$ ,其具体计算表达式为:

$$AW = \text{softmax}(AS) \quad (6)$$

其最终的输出  $AO$  为注意力权重  $AW$  与值矩阵  $\mathbf{V}$  的乘积,即有:

$$AO = AW \cdot \mathbf{V} \quad (7)$$

将所有注意力头的输出拼接起来,并通过线性变换将其转换为最终输出:

$$\begin{aligned} \text{MultiHead}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) &= \\ \text{Concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_h) \mathbf{W}_o \end{aligned} \quad (8)$$

式中:  $\text{head}_i$  为  $AO_i$ ;  $\mathbf{W}_o$  为输出权重矩阵。

### 1.4 时序映射拟合网络

时序映射拟合网络由双层 LSTM 网络构成。LSTM 是改进的递归神经网络,通过门控机制对信息进行选择性地保留、更新和输出,有效地捕捉和记忆序列中的重要信息,其核心单元包括遗忘门、输入门和输出门<sup>[20]</sup>。

#### 1.4.1 遗忘门

遗忘门决定了当前单元状态中有多少信息需被遗忘。假设给予模型输入序列为  $x_t$ ,上一个时间步的隐藏状态为  $h_{t-1}$ ,则遗忘门输出  $f_t$  的计算公式为:

$$f_t = \sigma(\mathbf{W}^f \cdot [h_{t-1}, x_t] + b^f) \quad (9)$$

式中:  $\mathbf{W}^f$  为遗忘门的权重矩阵;  $b^f$  为遗忘门的偏执项;  $\sigma$  为 sigmoid 函数。

#### 1.4.2 输入门

输入门决定了当前输入信息有多少需要被写入单元状态中。输入门由激活和候选记忆单元组成。输入门输出  $i_t$  和  $c'_t$  的计算公式为:

$$\begin{aligned} i_t &= \sigma(\mathbf{W}^i \cdot [h_{t-1}, x_t] + b^i) \\ c'_t &= \tanh(\mathbf{W}^c \cdot [h_{t-1}, x_t] + b^c) \end{aligned} \quad (10)$$

式中:  $\mathbf{W}^i$  为输入门的权重矩阵;  $b^i$  为输入门的偏执项;  $\mathbf{W}^c$  为候选记忆单元的权重矩阵;  $b^c$  为候选记忆单元的偏执项。

#### 1.4.3 更新单元状态

单元状态的更新由遗忘门和输入门共同决定。更新公式为:

$$c_t = f_t \odot c_{t-1} + i_t \odot c'_t \quad (11)$$

式中:  $\odot$  代表元素级乘法;  $c_t$  代表当前时间步状态;  $c_{t-1}$  代表当前上一时间步状态。

#### 1.4.4 输出门

输出门决定了当前单元状态有多少信息需要传递到隐藏状态中,计算公式为:

$$o_i = \sigma(W^o \cdot [h_{i-1}, x_i] + b^o) \tag{12}$$

$$h(t) = o_i \odot \tanh(c_i)$$

式中： $W^o$  为输入门的权重矩阵； $b^o$  为输入门的偏执项。

### 2 工程概况

本研究依托沪昆铁路义乌高架站房建设工程、信息及相关工程 YWZFTJ-1 标段中邻近杭长铁路 G6 道的基础工程,铁路监测范围对应里程为 K268+489—K268+799,以改建接触网对应的路基、轨道沉降及水平位移为主要监测目标,且采用以全自动全站仪为核心的智能化自动监测系统,通过该数据采集系统智能化一体化测控终端设备将监测数据无线传输到服务器和监测管理平台,再由专业监测管理人员通过监测管理平台查看监测数据并汇总,工作原理如图 2 所示。

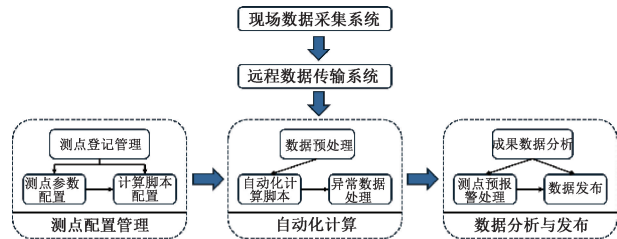


图 2 全自动监测系统工作原理

Fig. 2 Working principle of automatic monitoring system

根据影响杭长铁路的对应里程 K268+489—K268+939,每 10m 布设 1 个路基监测点,且每根接触网立柱布设 2 个监测点,上部监测点位置距离接触网  $\geq 2m$ ,每个观测墩布设  $\geq 3$  个棱镜作为基准点,如图 3 所示。

### 3 试验概况

本文所用数据集来源于高速铁路义乌高架站房杭长场 G6 道,涵盖了从 2023 年 7 月 19 日—2024 年 1 月 18 日的采集结果,采集周期为 2h/d,并对数据进行去噪处理,随机选择 3 个测点 (HC-2, HC-3 和 HC-6) 对所提模型和常见模型 (LSTM, BiLSTM, GRU 和 Transformer) 的泛化能力进行对比,试验设置的训练集和验证集的比值为 50%。所有模型均采用目标值与预测值之间的均方误差 (MSE) 作为损失函数,并使用 Adam 优化器进行参数优化。训练过程共进行 5 000 次迭代,时间步长设置为 8。各模型所使用的超参数如表 1 所示。试验运行环境为 Intel(R) Core(TM) i7-14700F 处理器,配备 NVIDIA GeForce RTX 4060GPU 显卡和 8GB 显存。试验结果的评价指标包括平均绝对误差 (MAE)、均方根误差 (RMSE)、平均绝对百分比误差 (MAPE) 和决定

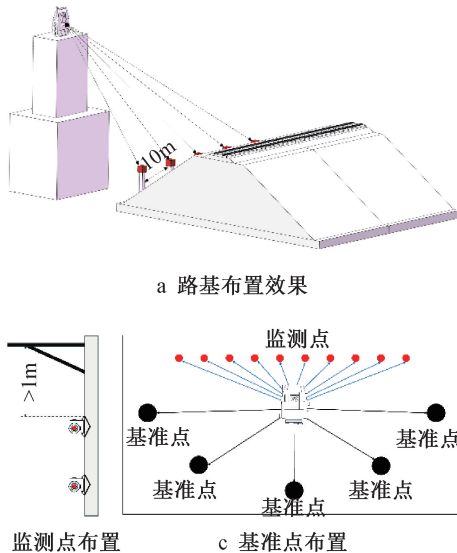


图 3 测点布置

Fig. 3 Arrangement of measuring points

系数 ( $R^2$ )。

表 1 超参数设置

Table 1 Hyperparameter settings

参数名称	LSTM	BiLSTM	GRU	Transformer	Res-MHA-LSTM
批处理个数	64	64	64	64	64
隐藏层数	128	128	50	无	128
时间步长	8	8	8	8	8
学习率	0.001	0.001	0.001	0.001	0.001
丢弃	0.1	0.1	0.1	0.1	0.1
多头注意力头数	无	无	无	16	16
嵌入向量维度	无	无	无	16	无

### 4 试验结果对比

各模型在 HC-2, HC-3 和 HC-6 测点数据下的预测结果如图 4 所示。

在测点集为 HC-2 时,Res-MHA-LSTM 模型不仅能够精准预测数据的总体趋势,还在数据的非线性特征和细节准确性方面表现得尤为突出,特别是在处理数据的复杂非线性变化时,Res-MHA-LSTM 能保持较高的预测精度。相较之下,LSTM, BiLSTM 和 GRU 模型在 100~200 样本点区间内的预测结果明显偏高,表明这些模型可能存在对数据非线性波动的过度响应或缺乏对实际趋势的准确把握。在 600~700 样本点区间,LSTM 和 BiLSTM 则出现了不应有的扰动现象,这可能是由于它们在捕捉数据的长期依赖关系时存在困难,导致预测不稳定。此外,Transformer 模型虽在一些情况下能捕捉到数据的全局信息,但在多个关键点上未能准确反映数据细节,表明 Transformer 在处理细粒度数据时能力有限。相比之下,Res-MHA-LSTM 通过其残差连接与

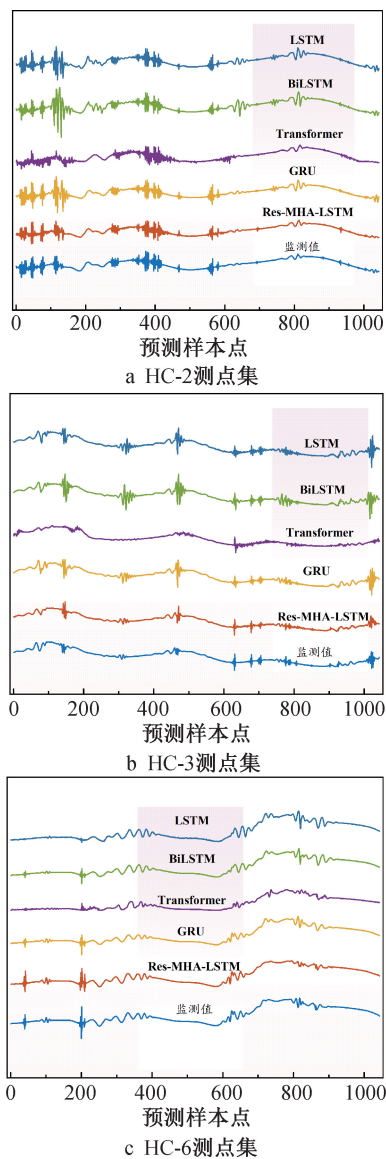


图4 各模型在不同测点测试集上的预测结果对比

Fig. 4 Comparison of prediction results of various models on different test sets at different measurement points

多头自注意力机制的深度结合,能够充分学习数据中的复杂模式与非线性特征,从而对所有显著波动进行精准捕捉,充分体现了 Res-MHA-LSTM 在非线性数据适应性和特征提取能力方面的显著优势。综上所述,Res-MHA-LSTM 在各种数据条件下均表现出色,显示了其在复杂预测任务中的强大能力。

在测点集为 HC-3 时,Res-MHA-LSTM 模型在捕捉数据变化趋势和预测取值方面表现出色,准确性明显优于其他模型。LSTM, BiLSTM 和 GRU 模型在预测过程中对非线性微小扰动的响应不够准确,甚至出现了过度放大或减弱的现象,这种异常性增强可能导致实际监测中的误导性判定,从而引发不必要的资源浪费。尤其是在处理数据的细微波动时,这些模型表现较差,未能有效地捕捉到数据的真实

特征。Transformer 模型虽然能够较好地描述数据的总体变化趋势,但在捕捉局部急促峰值变化时表现出一定的无效特性。这意味着虽 Transformer 可处理数据的全局信息,但在细节方面仍存在不足。相比之下,Res-MHA-LSTM 依托其独特的多头自注意力机制,能同时关注数据的局部与全局非线性特性,并通过残差连接实现对复杂模式的高效学习,从而提供了更可靠的预测结果。这不仅提升了数据预测的准确性,也为实际应用过程中的监测和维护提供了更加有效的支持。因此可以看出 Res-MHA-LSTM 在处理 HC-3 测点数据时显示出显著优势,验证了其在复杂数据预测任务中的强大能力。

在测点集为 HC-6 时,在样本点约为 600 的位置,其他模型在捕捉数据局部急促变化方面存在困难。LSTM 和 BiLSTM 对于这种局部变化的描述过于简单,未能准确反映数据的实际变化,而 Transformer 模型未能精确描述相应的非线性峰值关系,表现出一定的局限性。在样本点约为 800 的位置,LSTM 和 BiLSTM 的预测结果普遍偏高,可能是由于对数据变化的过度反应。而 Transformer 的预测则普遍偏低,未能充分捕捉到数据的实际变化趋势。GRU 模型在此处同样未能准确描述细节变化,导致预测结果的准确性不足。而 Res-MHA-LSTM 能够精确捕捉这些关键点上的非线性变化与细节特征,通过对复杂数据模式的全面学习,提供更为准确和可靠的预测结果。这表明 Res-MHA-LSTM 在处理复杂数据模式时具备明显优势,能为实际监测应用中的数据分析和决策提供更加精确的支持。

本文对反归一化后的结果进行了量化评估。各模型在水平方向不同训练集占比下,预测结果的评价指标如图 5 所示。Res-MHA-LSTM 模型在所有测点下均表现出更优的预测能力,其 MAE, RMSE, MAPE 均最小,  $R^2$  则最高。具体而言,以 HC-2 测点数据为例,相较于 LSTM 模型, Res-MHA-LSTM 的 MAE, RMSE 和 MAPE 分别降低了 23.09%, 23.51% 和 87.69%, 同时  $R^2$  提升了 7.4%。与 BiLSTM 模型相比, Res-MHA-LSTM 的 MAE, RMSE 和 MAPE 分别降低了 19.22%, 24.25% 和 87.07%,  $R^2$  则提升了 7.8%。与 Transformer 模型相比, Res-MHA-LSTM 的 MAE, RMSE 和 MAPE 分别降低了 53.94%, 38.78% 和 80.46%,  $R^2$  提升了 19.5%。与 GRU 模型相比, Res-MHA-LSTM 的 MAE, RMSE 和 MAPE 分别降低了 20.67%, 22.54% 和 87.36%,  $R^2$  提升了 6.9%。值得注意的是,在 MAPE 指标  $< 1$  的情况下,通常被判定为高质量模型。HC-2 测点数据下,其他模型均被

判定为劣质模型,而本文提出的 Res-MHA-LSTM 模型则被判定为可行模型,体现了所提模型在数据适应性和泛化性能方面的优势。此外,在 HC-3 和 HC-6 测点的数据预测结果中,Res-MHA-LSTM 模型也显示了类似的优势,进一步验证了其在各种测试条件下的优越性。

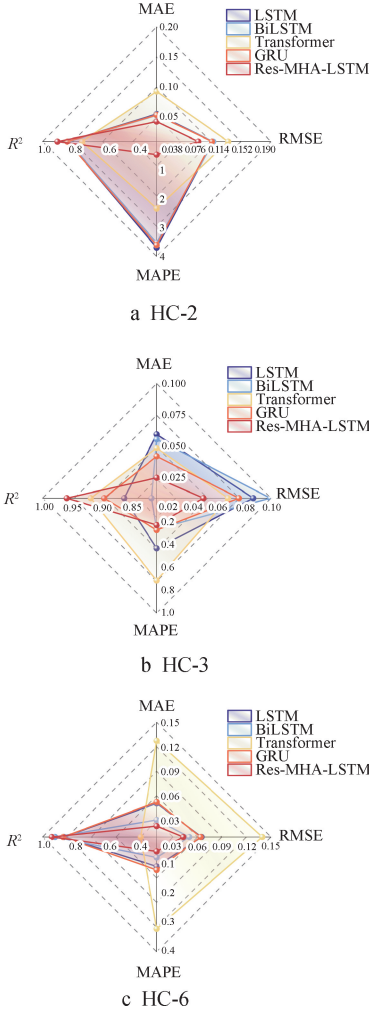


图5 各模型在不同测点下预测结果的评价指标对比

Fig.5 Comparison of evaluation indicators for prediction results of each model at different monitoring points

决定系数反映了模型在解释因变量变异方面的有效性。较高的  $R^2$  值表明模型在捕捉数据趋势和规律方面表现优异,能够更好地解释因变量的变化。为了进一步验证模型的性能,以 HC-2 测点为例,本文对各模型的拟合结果进行了可视化分析。预测值与真实值之间的拟合关系如图 6 所示,并可视化了 95% 的置信区间。置信区间越窄,意味着模型的预测结果越稳定。由图 6 可知,Res-MHA-LSTM 模型表现最佳,其置信区间是所有模型中最窄的,表明其预测结果的稳定性和可信度最高。此外,Res-MHA-LSTM 的预测值与真实值之间的拟合

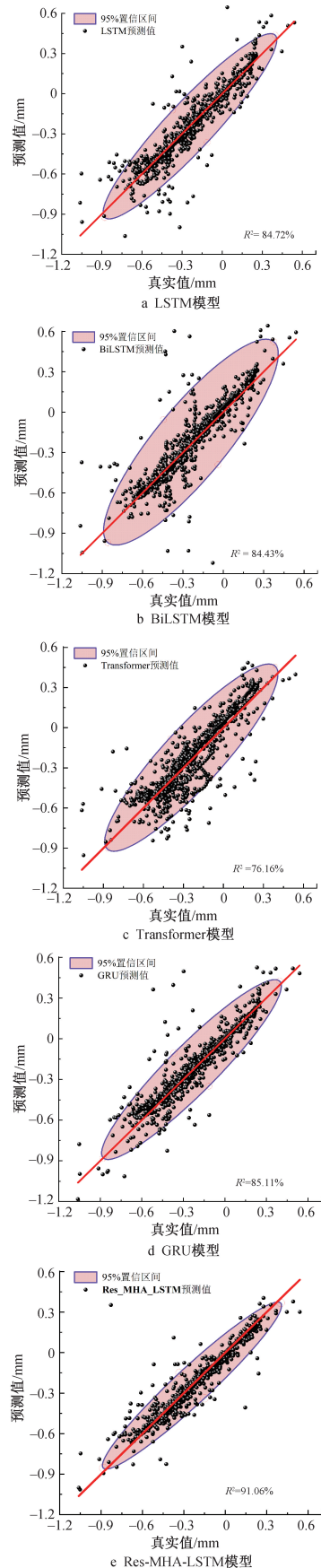


图6 各模型拟合结果可视化分析(HC-2测点)

Fig.6 Visual analysis of fitting results of various models(measuring point HC-2)

关系也最为紧密,说明该模型在精确捕捉数据趋势方面的能力最强。相较之下,BiLSTM 和 Transformer 模型的表现较差。它们的置信区间较宽,反映出较大的预测不确定性,特别是 Transformer 模型,其预测结果与真实值偏差较大,显示出较低的预测精度和可靠性。LSTM 和 GRU 模型的预测结果在可信度和准确性方面虽优于 BiLSTM 和 Transformer,但仍不如 Res-MHA-LSTM 模型。

综上所述,Res-MHA-LSTM 模型显现出优越性,其稳定性和准确性均优于其他模型,进一步证明 Res-MHA-LSTM 在复杂数据预测任务中的显著优势,并为其在实际中广泛应用提供了有力支持。

## 5 结语

1)提出了改进 LSTM 沉降预测模型(Res-MHA-LSTM 模型),该模型由残差连接、多头注意力机制和 LSTM 构成,具有更强的训练稳定性和泛化能力,实现了准确的沉降预测。

2)在沪昆铁路义乌高架站房建设工程杭长道工程的应用中,通过对比 3 个测点数据下的 5 种常见时序预测网络,利用平均绝对误差和均方根误差量化分析了模型的预测效果。试验结果表明,Res-MHA-LSTM 网络在不同训练样本下具有较高的预测准确性和可行性。

## 参考文献:

[ 1 ] TUCHO A, INDRARATNA B, NGO T. Stress-deformation analysis of rail substructure under moving wheel load [ J ]. *Transportation geotechnics*, 2022, 36: 100805.

[ 2 ] VARANDAS J N, PAIXÃO A, FORTUNATO E, et al. Long-term deformation of railway tracks considering train-track interaction and non-linear resilient behaviour of aggregates—a 3D FEM implementation [ J ]. *Computers and geotechnics*, 2020, 126: 103712.

[ 3 ] CHAWLA S, SHAHU J T, KUMAR S. Analysis of cyclic deformation and post-cyclic strength of reinforced railway tracks on soft subgrade [ J ]. *Transportation geotechnics*, 2021, 28: 100535.

[ 4 ] CHEN Z W. Evaluation of longitudinal connected track under combined action of running train and long-term bridge deformation [ J ]. *Journal of vibration and control*, 2020, 26(7/8): 599-609.

[ 5 ] ZAMBRANO O A, TRESSIA G, SOUZA R M. Failure analysis of a crossing rail made of Hadfield steel after severe plastic deformation induced by wheel-rail interaction [ J ]. *Engineering failure analysis*, 2020, 115: 104621.

[ 6 ] XING Y, CHEN Y, HUANG S P, et al. Research on dam deformation prediction model based on optimized SVM [ J ]. *Processes*, 2022, 10(9): 1842.

[ 7 ] WAN P, ZOU H, WANG K L, et al. Research on hot deformation behavior of Zr-4 alloy based on PSO-BP artificial neural network

[ J ]. *Journal of alloys and compounds*, 2020, 826: 154047.

[ 8 ] ZHOU Z W, GONG H Y, YOU J, et al. Research on compression deformation behavior of aging AA6082 aluminum alloy based on strain compensation constitutive equation and PSO-BP network model [ J ]. *Materials today communications*, 2021, 28: 102507.

[ 9 ] ZHANG Y G, TANG J, LIAO R P, et al. Application of an enhanced BP neural network model with water cycle algorithm on landslide prediction [ J ]. *Stochastic environmental research and risk assessment*, 2021, 35(6): 1273-1291.

[ 10 ] LIN C, WENG K L, LIN Y L, et al. Time series prediction of dam deformation using a hybrid STL-CNN-GRU model based on sparrow search algorithm optimization [ J ]. *Applied sciences*, 2022, 12(23): 11951.

[ 11 ] 夏超鹏, 洪宇, 刘新山, 等. 基于 LSTM 深度学习网络的中长期电力负荷预测 [ J ]. *自动化与仪器仪表*, 2024(7): 343-346.

XIA C P, HONG Y, LIU X S, et al. Based on LSTM network for medium-term and long-term power prediction based on pytouch [ J ]. *Automation & instrumentation*, 2024(7): 343-346.

[ 12 ] 谭茜元, 谷佩奇, 张悦悦. 基于灰狼优化算法优化 LSTM 网络的短期负荷预测 [ J ]. *现代工业经济和信息化*, 2024, 14(7): 267-268, 271.

TAN X Y, GU P Q, ZHANG Y Y. Optimising LSTM networks for short-term load forecasting based on the grey wolf optimisation algorithm [ J ]. *Modern industrial economy and informationization*, 2024, 14(7): 267-268, 271.

[ 13 ] CHEN J, LI A Y, BAO C Y, et al. A deep learning forecasting method for frost heave deformation of high-speed railway subgrade [ J ]. *Cold regions science and technology*, 2021, 185: 103265.

[ 14 ] WANG X, BAI Y, LIU X. Prediction of railroad track geometry change using a hybrid CNN-LSTM spatial-temporal model [ J ]. *Advanced engineering informatics*, 2023, 58: 102235.

[ 15 ] 袁利伟, 杨柳, 何涛, 等. 结合 CNN-LSTM 优化模型在尾矿坝浸润线预测中的应用 [ J ]. *安全与环境学报*, 2024, 24(10): 3770-3779.

YUAN L W, YANG L, HE T, et al. Application of CNN-LSTM optimization model in predicting tailing dam saturation lines [ J ]. *Journal of safety and environment*, 2024, 24(10): 3770-3779.

[ 16 ] HAN K, XIAO A, WU E, et al. Transformer in transformer [ J ]. *Advances in neural information processing systems*, 2021, 34: 15908-15919.

[ 17 ] WANG J, LI C, LI L, et al. InSAR time-series deformation forecasting surrounding Salt Lake using deep transformer models [ J ]. *Science of the total environment*, 2023, 858: 159744.

[ 18 ] 周康敏, 程康, 曾少翔, 等. 基于深度残差 LSTM 的盾构姿态预测 [ J ]. *隧道建设(中英文)*, 2024, 44(8): 1643-1651.

ZHOU K M, CHENG K, ZENG S X, et al. Prediction of shield attitude using deep residual long short-term memory model [ J ]. *Tunnel construction*, 2024, 44(8): 1643-1651.

[ 19 ] LI J, WANG X, TU Z P, et al. On the diversity of multi-head attention [ J ]. *Neurocomputing*, 2021, 454: 14-24.

[ 20 ] YU Y, SI X S, HU C H, et al. A review of recurrent neural networks; LSTM cells and network architectures [ J ]. *Neural computation*, 2019, 31(7): 1235-1270.