

基于TCN-SMHA-LSTM-TPA混合神经网络的复杂地层盾构土仓压力预测新方法

樊秋轶¹, 贾连辉², 张子新¹

¹同济大学土木工程学院, 上海

²中铁工程装备集团有限公司, 河南 郑州

收稿日期: 2024年9月12日; 录用日期: 2024年10月7日; 发布日期: 2024年10月18日

摘要

文章针对传统方法在复杂地层盾构掘进过程中关于土仓压力超前预测精度不高的问题, 提出了一种融合时间卷积神经网络(TCN)、多头稀疏自注意力(SMHA)、长短时记忆网络(LSTM)和时间感知注意力(TPA)的混合神经网络模型, 用于精准预测复杂地层盾构施工中的土仓压力。TCN模块通过多层卷积捕捉时间序列的局部特征, 而多头稀疏自注意力机制进一步增强了模型对长距离依赖关系的建模能力, LSTM层则专注于捕获时间序列的长期依赖信息, TPA模块自适应地调整注意力权重, 从而提高预测精度。该模型在盾构施工中的应用表明, 其在复杂地质条件下对盾构土仓压力的超前预测具有较高的准确性和适应性, 有助于盾构施工的实时监控与决策。

关键词

盾构隧道, 混合神经网络, 超前预测, 土仓压力

Novel Method for Predicting Chamber Pressure in Complex Geological Conditions Based on TCN-SMHA-LSTM-TPA Hybrid Neural Network

Qiuyi Fan¹, Lianhui Jia², Zixin Zhang¹

¹School of Civil Engineering, Tongji University, Shanghai

²China Railway Engineering Equipment Group Co., Ltd., Zhengzhou Henan

Received: Sep. 12th, 2024; accepted: Oct. 7th, 2024; published: Oct. 18th, 2024

文章引用: 樊秋轶, 贾连辉, 张子新. 基于 TCN-SMHA-LSTM-TPA 混合神经网络的复杂地层盾构土仓压力预测新方法[J]. 地球科学前沿, 2024, 14(10): 1256-1266. DOI: 10.12677/ag.2024.1410116

Abstract

To address the issue of low prediction accuracy in traditional methods for forecasting chamber pressure during shield tunneling in complex geological conditions, this article proposes a hybrid neural network model based on Temporal Convolutional Network, Sparse Multi-Head Attention, Long Short-Term Memory, and Temporal Pattern Attention. The model is designed for precise prediction of chamber pressure in shield construction. The TCN module captures local features of time series through multi-layer convolution, while the SMHA mechanism further strengthens the model's ability to model long-range dependencies. The LSTM layer focuses on extracting long-term dependencies in time series, and the TPA module adaptively adjusts attention weights to improve prediction accuracy. The application of this model in shield tunneling demonstrates its high accuracy and adaptability in forecasting chamber pressure under complex geological conditions, contributing to real-time monitoring and decision-making in tunneling operations.

Keywords

Shield Tunnel, Hybrid Neural Network, Advance Prediction, Chamber Pressure

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在复杂地层隧道施工领域，盾构法作为一种成熟且广泛采用的技术手段，其应用日益普遍。其中，土压平衡盾构具有适用面广、占地面积小、对环境影响小的特点，往往成为首选设备。在众多掘进参数中，土仓压力是保证盾构正常掘进的一项非常重要的参数。合理控制土仓压力于预定范围内，对于预防开挖面失稳、减轻地表沉降或隆起等不利地质响应具有决定性作用。值得注意的是，土仓压力作为一个间接参数，其值通过精密传感器监测获取，且无法直接进行调节，而是受到多种复杂变量的综合影响。因此，复杂地层中盾构掘进时土仓压力的控制过程显得尤为复杂，需要基于掘进过程中的实时数据，如盾构机进出土速率等，动态调整土仓内的土体存量，以实现压力的稳定控制。如何基于现有的掘进参数体系，精准预测并有效控制复杂地层中盾构掘进的土仓压力，已成为盾构隧道开挖技术研究中亟待解决的重要课题。

在以往的研究中，很多研究人员[1][2]根据土仓开挖和排出的土量守恒和胡克定律，通过分析土压平衡盾构机运行中可能涉及的各类方程关系，建立了土压平衡盾构掘进过程的数学物理模型，即解析法，并通过调控螺旋输送机的转速等多种策略，实现土压的动态平衡。在此基础上，Yu等[3]着眼于土仓内土体质量守恒的核心，综合考量了推进速率、添加剂注入速率、螺旋输送机排土速率等多维度因素，开发了一种适用于工作与非工作状态的土仓压力预测模型。但是，这类基于理想化假设和固定参数设定的理论模型，在实际工程应用中往往面临挑战。它们往往对开挖土体的物理力学特性及开挖条件的复杂性进行了较大程度的简化，忽略了各变量间的非线性耦合效应，因此难以完全反映实际工程环境的动态变化。

近年来，随着大数据和人工智能技术的快速发展，盾构掘进过程中产生的海量数据得以高效收集与利用。众多学者通过大数据统计分析[4]、遗传算法[5]、支持向量机[6]、随机森林[7]、粒子群[8]、数字孪生[9]、神经网络[10][11]、混合神经网络[12][13]等机器学习智能方法建立盾构掘进参数预测模型，实现

了从复杂数据中自动挖掘并构建目标变量与多维度影响因素间复杂关系的目标[14][15]。研究表明,相较于传统的数学物理模型,这些智能算法构建的预测模型在准确性与适应性方面展现出显著优势,对于指导不同地质条件下土压平衡盾构掘进参数的精准预测与控制具有重要研究价值和应用前景。然而,现有的土仓压力智能模型研究仍存在以下问题:1) 尽管现有模型能够基于实时盾构参数变化辅助决策,确保掘进过程的安全稳定,但其缺乏前瞻性,即无法提前预知前方地质条件变化对掘进参数的影响,进而无法实现超前预警与预控,限制了其在复杂地质环境中的全面应用。2) 多数研究在构建模型时,过于依赖数学力学理论框架,倾向于仅考虑推进系统与排土系统的参数,而忽视了渣土改良系统参数的关键作用。改良剂的加入深刻改变了土体的物理力学特性,直接影响理论分析中的核心参数。因此,未将渣土改良系统纳入考量范围的模型在完整性上存在明显缺陷,难以全面反映实际掘进过程中的复杂动态变化。

针对以上不足之处,本文提出一种基于双重注意力机制的 TCN-SMHA-LSTM-TPA 网络结构对复杂地层中盾构掘进过程的土仓压力进行超前预测。从上述研究内容来看,在诸多机器学习方法中,以神经网络为代表的深度学习方法因其强大的高维、非线性特征提取能力,预测性能相对较优,应用范围也相对更广。因此,本研究在传统的 LSTM 模型的基础上,进行了系列优化。首先采用时间卷积网络(TCN)感受并初步提取掘进数据序列中不同时间尺度的时序信息,再引入多头稀疏自注意力模块(SMHA)聚焦输入信息的多个不同部分,捕捉更为丰富和全面的特征,增强模型的表征能力。同时保留长短期记忆网络(LSTM)的核心优势,即保留并传递更长期的依赖关系。最后加入时间模式注意力机制(TPA)学习并识别不同时间步长的重要性,同时考虑不同变量之间的内在联系。双重注意力机制的加入使得整个网络结构在面对多样化的掘进任务时,能够自动调整其内部权重,展现出高度的自适应性和自调节能力。此外,本文在工程应用时也充分考虑了渣土改良系统相关参数的影响。

2. 模型构建

2.1. 时间卷积层

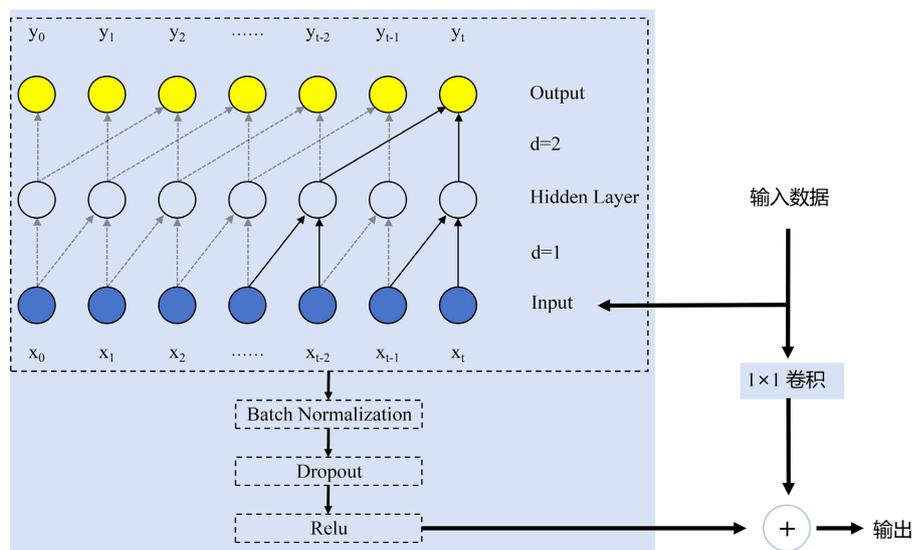


Figure 1. Unit structure of TCN
图 1. TCN 单元结构

时间卷积网络能够在序列数据中有效学习复杂的非线性特征,并精准捕捉跨越不同时间尺度的信息。TCN 主要包括因果卷积、扩张卷积和残差连接,这使得其在捕捉长时间依赖关系的同时保持了序列的时

间顺序,从而深入挖掘土仓压力与特征变量之间潜在的长期依赖关系。其关键结构如图1所示。

图1中,各层间的箭头表示TCN的残差处理结构, d 为网络层膨胀因子。输入序 $\{x_1, x_2, \dots, x_n\}$ 通过两个残差结构得到输出序列 $\{y_1, y_2, \dots, y_n\}$ 。TCN的核心思想在于使用因果卷积和扩张卷积。因果卷积确保输出仅依赖于当前及之前的时间步,从而保持时间顺序,避免了全连接层中输出的前一时刻的神经元与输入的后一时刻的神经元产生了连接的问题。扩张卷积通过在卷积核中引入间隔,使网络能够在不增加参数数量的情况下扩展感受野(通过改变网络层膨胀因子来实现)。具体公式为:

$$\text{Output}_t = \sum_{i=0}^{k-1} w_i \cdot \text{Input}_{t-d \cdot i} \quad (1)$$

式中, Output_t 为时间步 t 的输出,表示在当前位置 t 上的卷积操作结果; w_i 为卷积核的权重,表示第 i 个卷积核元素的值,每一个 w_i 都是模型学习到的参数; $\text{Input}_{t-d \cdot i}$ 为输入序列在时间步 $t-d \cdot i$ 上的值; k 是卷积核的大小(长度); d 是膨胀因子,表示卷积核在时间序列上的扩张程度。

在此基础上,TCN采用残差连接,避免了随网络层增加而导致的梯度消失或梯度爆炸等问题。

2.2. 多头稀疏自注意力机制

多头稀疏自注意力机制(SMHA)是对传统多头自注意力机制的一种改进。它通过引入稀疏性降低计算复杂度,并通过多头机制来捕捉输入序列中不同部分之间的依赖关系。SMHA保留了自注意力机制的优势,同时有效减少了计算资源的消耗。

自注意力机制的核心思想是根据输入序列中的每个元素与其他元素的关联程度(即注意力权重)对该元素进行加权求和。公式如下:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2)$$

式中, Q 是查询矩阵,表示要在其他向量中查找的信息; K 是键矩阵,表示与查询向量进行匹配的参考信息; V 是值矩阵,表示最终输出的内容;这三项由输入序列 $\{x_1, x_2, \dots, x_n\}$ 通过不同的线性变换得到; K^T 是键矩阵的转置,用于计算查询向量和键向量之间的相关性; d_k 是键向量的维度,用于缩放因子; softmax 用于将注意力权重归一化为概率分布,确保所有的权重和为1; $\text{Attention}(Q, K, V)$ 表示通过注意力机制加权求和值矩阵 V 的结果。

对于多头注意力,多头机制将输入分解为 h 个独立的头,每个头在一个子空间内执行注意力操作:

$$\text{Multihead}(Q, K, V) = \text{Concat}(\text{Head}_1, \dots, \text{Head}_h)W^0 \quad (3)$$

其中, W 为线性变换权重矩阵,第 i 个头的计算为:

$$\text{Head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \quad (4)$$

在稀疏自注意力机制中,引入了稀疏性,通过稀疏矩阵 S 来约束注意力的计算,使得某些输入对之间的注意力权重为零,从而减少不必要的计算。稀疏自注意力机制的公式可以表示为:

$$\text{SparseAttention}(Q, K, V) = \text{softmax}\left(\frac{QK^T \odot S}{\sqrt{d_k}}\right)V \quad (5)$$

其中, S 为稀疏矩阵,用于指示哪些位置的注意力权重应被计算。

2.3. 长短期记忆网络

LSTM是一种广泛使用的循环神经网络(Recurrent Neural Network, RNN)变种,专门用于解决长期依

赖问题。LSTM 通过引入遗忘门、输入门和输出门来控制信息的流动，从而在长序列中保留有效信息，有效地缓解了传统 RNN 的梯度消失问题。

LSTM 的工作原理基于三个门控机制。遗忘门 f_t ，控制遗忘的速度：

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (6)$$

输入门 i_t 和候选记忆单元 C'_t ：决定将多少新信息加入记忆单元：

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (7)$$

$$C'_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (8)$$

输出门 o_t 和新的隐藏状态 h_t ：

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (9)$$

$$h_t = o_t \tanh(C'_t) \quad (10)$$

式中， f_t 为遗忘门的输出，决定了当前记忆单元的多少信息应该被遗忘； i_t 为输入门的输出，决定了当前输入的信息应该添加到记忆单元中； C'_t 为候选记忆单元，表示当前输入的信息经过激活函数处理后可以被写入记忆单元的内容； o_t 为输出门的输出，决定了记忆单元中的信息有多少会被输出到隐藏状态； h_t 为隐藏状态，表示当前时间步的输出，结合了输出门的控制和记忆单元的内容； W 为权重矩阵， b 为偏置项。

2.4. 时间模式注意力机制

时间模式注意力机制的核心思想是根据时间信息对输入数据的各时间步进行加权，从而使模型能够重点关注在时间上相关性较强的部分。因此，TPA 在处理时间序列任务时具备独特的优势。TPA 通常与 LSTM 等序列模型结合使用，对输出的各时间步长加权求和，生成最终的输出。

在这一过程中，TPA 接收 LSTM 的隐藏状态向量序列 H 作为输入，对这些隐藏状态进行加权，以突出与当前任务最相关的时间模式。实现过程为：

$$Q = HW^Q \quad (11)$$

$$K = HW^K \quad (12)$$

$$V = HW^V \quad (13)$$

式中， W^Q 、 W^K 、 W^V 为线性变换矩阵， Q 、 K 、 V 的维度均为 $R^{T \times d_a}$ ， d_a 是注意力机制的维度。

然后，TPA 通过计算查询与键的点积，生成注意力分数矩阵，并通过 softmax 函数将其转换为注意力权重。

$$A = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} + T \right) \quad (14)$$

其中， T 是可选的时间相关性矩阵，用于引入时间步之间的额外依赖关系， A 是注意力权重矩阵。

最后，使用注意力权重 A 对值矩阵 V 进行加权求和，生成时间模式注意力的输出 O 。

$$O = AV \quad (15)$$

为了保证模型的稳定性和优化效果，本研究中将 TPA 的输出与 LSTM 的隐藏状态进行残差连接。

2.5. 模型集成

将上述模型融合，最后使用全连接层进行输出。由于该模型网络层较多，为了避免出现梯度消失、

保证模型稳定性, 文本采用残差连接的方式构建模型, 最终建立得到 TCN-SMHA-LSTM-TPA 混合模型, 对输入的多变量时间序列数据进行预测, 如图 2 所示。同时, 为了防止模型过拟合, 模型中加入了早停机制, 用于监控模型的验证损失, 在验证损失不再下降时提前停止训练。

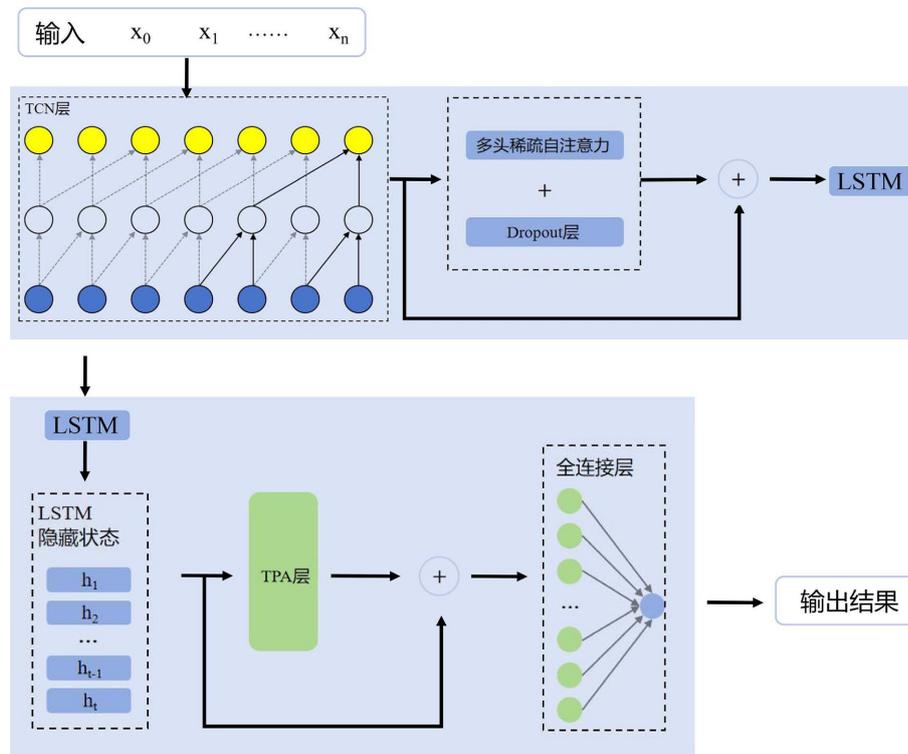


Figure 2. Network architecture of the model
图 2. 模型网络结构图

3. 模型训练

基于上述构建的 TCN-SMHA-LSTM-TPA 模型, 将选取的各特征(包括要预测的土仓压力值)过去 5 个时间步的盾构掘进序列数据作为模型输入, 输出则是下一个时间步的土仓压力。

3.1. 数据预处理及特征选取

按照黄鸿宇[13]的方法对数据进行预处理, 并通过相关性分析进行变量选择。最终本研究的模型输入变量确定为 9 个特征参数(刀盘转速、推进压力、D 组推进压力、螺机转速、推进速度、刀盘扭矩、螺机扭矩、总推进力、泡沫混合液当前环累计量)和 1 个标签参数(土仓压力)。

3.2. 数据集划分

将预处理后的数据集按照 8:1:1 的比例划分为训练集、验证集和预测集。训练集用于模型的训练, 验证集用于优化模型的超参数, 而预测集用于评估最终优化后的模型性能。

3.3. 模型训练

基于 python 的 pytorch 库构建网络模型, 所有的训练及优化过程均在同一硬件平台完成。该平台的 CPU 为 12th Gen Inter(R) Core(TM)i7-12700H 2.30 GHz, RAM 为 16.0 GB, GPU 为 NVIDIA GeForce RTX

3060 Laptop GPU。

在模型中，神经网络训练的目标是优化各个神经元的权重参数和偏置项，使得整个神经网络对学习目标的拟合度达到最大化。这些参数在训练开始时具有初始值，并随着训练过程的推进逐步进行优化。

经过试验，各模型最优超参数组合如表 1 所示。

Table 1. Main parameters used in TCN-SMHA-LSTM-TPA

表 1. TCN-SMHA-LSTM-TPA 的主要参数

模型	超参数	参数值
TCN	膨胀因子	[1, 2]
	膨胀层	2
	隐藏层单元数	64
多头稀疏自注意力	注意力	8
	注意力因子	5
	隐藏单元数	64
LSTM	层数	2
	隐藏单元数	50
TPA	隐藏单元数	50

3.4. 模型评价指标

为评估模型的预测效果，选取了平均绝对误差(MAE)、平均绝对百分比误差(MAPE)、均方根误差(RMSE)和决定系数(R²)四项指标作为评价标准。这些指标能够从不同角度衡量模型的预测性能，其计算方式如下。

平均绝对误差(MAE):

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \tilde{y}_i| \quad (16)$$

式中， y_i 为真实值， \tilde{y}_i 为预测值。MAE 衡量预测值与真实值之间的平均绝对差值，即在所有样本中，预测值与真实值之间的平均差异程度。MAE 反映了预测误差的整体大小，但不区分误差的正负。MAE 值越小，表明模型的预测结果与真实值之间的误差越小，模型的拟合精度越高。

平均绝对百分比误差(MAPE):

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \tilde{y}_i}{y_i} \right| \times 100\% \quad (17)$$

式中， y_i 为真实值， \tilde{y}_i 为预测值。MAPE 衡量预测误差相对于真实值的百分比，其优势在于该指标不受数据尺度的影响。MAPE 值越小，表明模型的预测误差相对于真实值的比例越小，模型的预测精度越高。

均方根误差(RMSE):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \tilde{y}_i)^2} \quad (18)$$

式中， y_i 为真实值， \tilde{y}_i 为预测值。RMSE 反映了预测误差的标准差，它通过对误差进行平方后再开平方，

赋予较大误差更高的权重。由于 RMSE 具有数据尺度依赖性，因此在数据值较大时，RMSE 的值也可能相应增大。RMSE 值越小，表明预测值与真实值之间的偏差越小，模型的拟合效果越好。

决定系数(R^2):

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \tilde{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (19)$$

式中， y_i 为真实值， \tilde{y}_i 为预测值， \bar{y}_i 为真实值的平均值。 R^2 是衡量模型对数据拟合优度的指标，表示模型能够解释数据变异的比率。 R^2 的取值范围在 0 到 1 之间，其中 1 表示模型能够完全解释数据的变异，0 则表示模型未能解释数据中的任何变异。 R^2 值越接近 1，说明模型对数据的拟合程度越高。 R^2 值为 1 意味着预测值完全等同于真实值，而 R^2 值为 0 则表明模型的预测毫无解释力。

综上所述，这些指标相结合，可以从多个角度综合评估模型的性能。MAE 和 MAPE 分别提供了平均误差和相对误差的直接衡量，RMSE 则强调了较大误差的影响，而 R^2 则评估了模型的整体拟合优度。在实际应用中，通常希望 MAE、MAPE、RMSE 值越小越好，而 R^2 值越接近 1 越理想。

4. 实例分析

4.1. 工程背景

本文研究基于中铁工程装备集团有限公司承担的上海市轨道交通 18 号线工程 1 标段，选取该工程长江南路站至通南路站区间的现场施工数据作为研究对象。拟建区间为地下隧道，沿线环境较为复杂。该区段全长 664 米，盾构隧道的外径为 6.6 米，环宽 1.2 米，共计 550 环。截至本研究开展时，该区段盾构已推进 110 环。

4.2. 数据集构建

在该工程中，盾构机以 60Hz 的频率对运行参数进行实时记录。本研究采用了近期推进的 91 至 110 环相关掘进数据，总计记录了 346,304 个数据样本。

由于后续研究需要实现渣土改良智能决策相关内容，必须对土仓压力等掘进参数进行超前预测。然而，在实际工程中，从盾构司机作出反应到渣土改良剂输送至土体的时间约为 8 秒。因此，预测 1/60 秒后的土仓压力在工程应用中并无实际意义。为此，在原始数据集的基础上首先以 10 秒为时间步长进行采样，经过前述的数据预处理和特征参数选择后，共形成 20 个数据集，并以 8:1:1 的比例划分训练集、验证集和预测集。模型的输入参数包括前述的 9 个特征参数以及 1 个标签参数。

4.3. 实验结果分析

使用训练好的模型在预测集上对盾构的土仓压力进行超前预测，每一个时间步长为 10 秒，用前 5 个时间步的历史数据预测下一个时间步的土仓压力值，各神经网络模型的预测效果见图 3。各模型在预测集上的评价指标值(MAE, MAPE, RMSE, R^2)如表 2 所示。从图 3 中可以看出，引入了双重注意力机制和时间卷积网络的 TCN-SMHA-LSTM-TPA 混合模型对土仓压力的预测效果最好，在局部波动的区段，其预测表现也要优于 LSTM 和 GRU。

从表 2 所示的评价指标值来看，TCN-SMHA-LSTM-TPA 混合模型的精度也是最高的。

4.4. 模型应用

4.3 节说明所提出的 TCN-SMHA-LSTM-TPA 混合模型在土仓压力预测方面相较于其他预测模型具有更高的精度。然而，在实际应用过程中，由于该模型在完成预测、输出结果到盾构司机查看结果和作

出决策等一系列步骤时同样需要几秒钟的时间，仅预测 10 秒后的土仓压力的实际意义有限。因此，采用 TCN-SMHA-LSTM-TPA 模型进行更长期的土仓压力预测具有更大的实际应用价值。

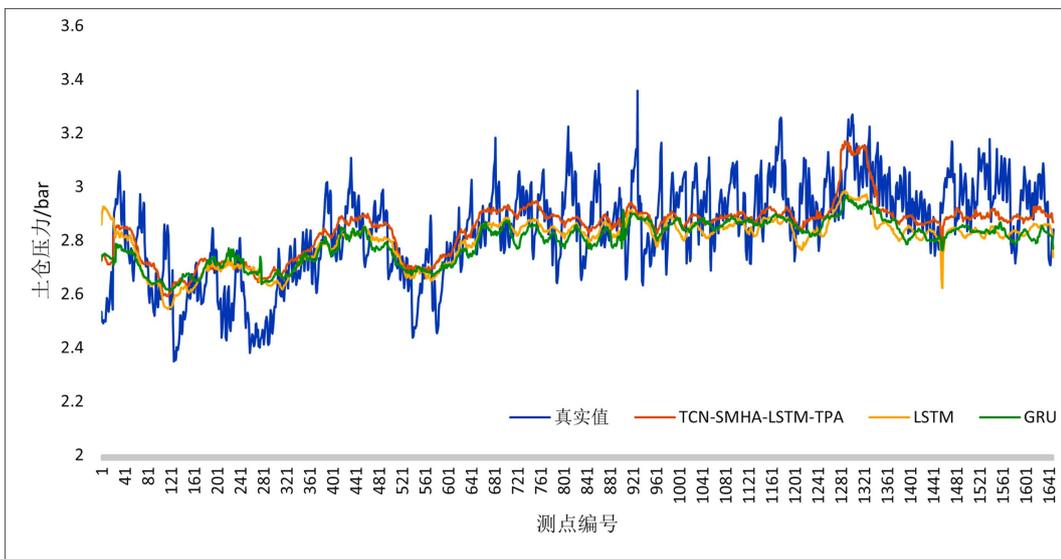


Figure 3. Prediction results of the models

图 3. 各模型预测效果

Table 2. Values of evaluation index

表 2. 各模型评价指标值

模型	评价指标			
	MAE	MAPE	RMSE	R ²
TCN-SMHA-LSTM-TPA	0.097	3.484	0.121	0.530
LSTM	0.112	3.937	0.141	0.368
GRU	0.111	3.904	1.137	0.400

采用递归多步预测方法，将时间步长调整为 5 秒，通过 TCN-SMHA-LSTM-TPA 模型对盾构机未来 60 个时间步长的土仓压力进行预测。具体而言，通过模型预测 $t+1$ 秒的土仓压力后，将预测值作为新的输入特征代入模型，继续预测 $t+2$ 秒的目标值，并重复这一过程。TCN-SMHA-LSTM-TPA 模型基于前 120 个时间步长的掘进数据对未来 60 个时间步长的土仓压力预测结果的评价指标值如图 4 所示。

从图 4 中可以观察到，TCN-SMHA-LSTM-TPA 模型在前 24 个时间步长(即 2 分钟)的预测准确度较高，能够较好地捕捉土仓压力的变化趋势。然而，在第 25 个时间步长之后，误差显著增加。这一现象主要是由于递归预测过程中误差会随着预测时间步的累积而传播，导致四个评价指标值迅速恶化。这表明，模型能有效地预测未来一段时间内盾构的相关参数变化情况，但若预测周期过长(超过 2 分钟)，模型的准确性将显著下降。

5. 结论

本文构建了一种基于 TCN-SMHA-LSTM-TPA 的混合神经网络模型，用于实现复杂地层盾构掘进过程中土仓压力的超前预测，并依托上海市轨道交通 18 号线工程进行了相关应用。得到以下结论：

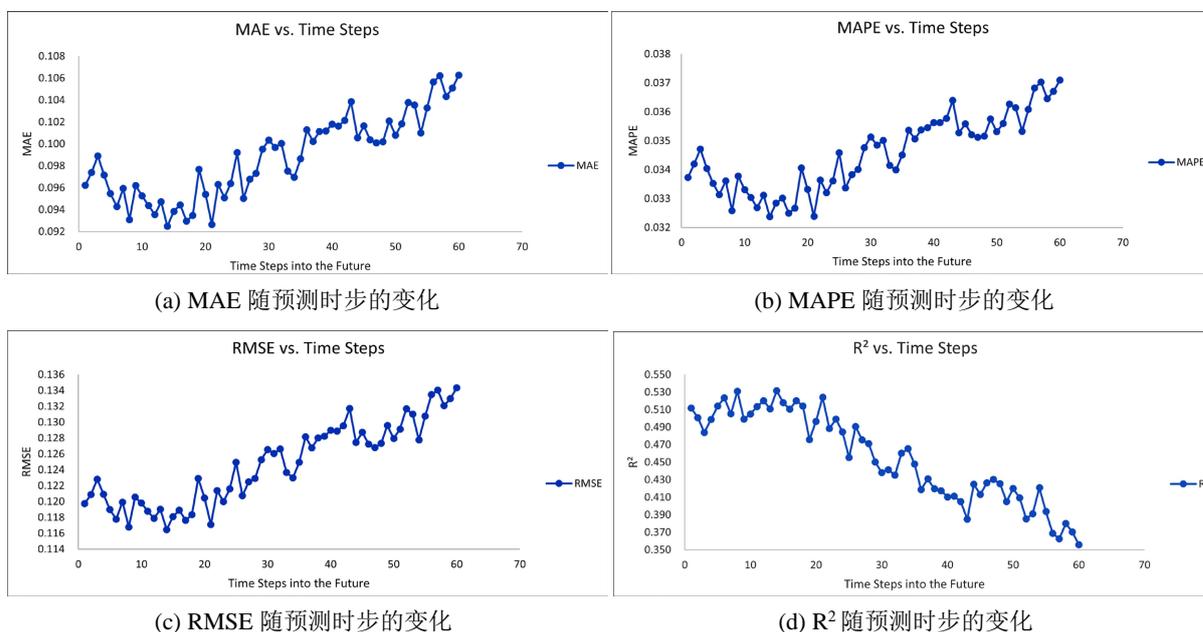


Figure 4. Changes of the values of evaluation index
图 4. 各评价指标值的变化

1) 短期预测任务(预测未来 10 s 的土仓压力)中, TCN-SMHA-LSTM-TPA 混合模型的预测精度远高于传统的 LSTM 和 GRU 模型, 具体表现为四项评价指标值(MAE, MAPE, RMSE, R²)均最优。

2) 短期预测任务(预测未来 10 s 的土仓压力)中, 在局部波动的区段, TCN-SMHA-LSTM-TPA 混合模型的预测表现相较于传统的 LSTM 和 GRU 模型更贴合实际。说明双重注意力机制和时间卷积网络的引入增强了模型对土仓压力变化的捕捉能力, 并提高了模型对复杂掘进过程的适应性。

3) 长期预测任务(预测未来 300 s 的土仓压力)中, TCN-SMHA-LSTM-TPA 混合模型在前 2 分钟内表现出高精度(MAE, MAPE, RMSE, R² 均在理想范围), 显示出该模型具备一定的长期有效预测能力。但若预测周期过长, 模型的准确性会随着递归预测过程中误差的累积而下降。

参考文献

- [1] 王洪新, 傅德明. 土压平衡盾构掘进的数学物理模型及各参数间关系研究[J]. 土木工程学报, 2006, 39(9): 86-90.
- [2] 李守巨, 屈福政, 曹丽娟, 等. 土压平衡盾构机密封舱压力控制实验研究[J]. 煤炭学报, 2011, 36(6): 934-937.
- [3] Yu, H., Mooney, M. and Bezuijen, A. (2020) A Simplified Excavation Chamber Pressure Model for EPBM Tunneling. *Tunnelling and Underground Space Technology*, **103**, Article 103457. <https://doi.org/10.1016/j.tust.2020.103457>
- [4] 沈翔, 袁大军, 吴俊, 等. 复杂地层条件下盾构掘进参数分析及预测[J]. 现代隧道技术, 2020, 57(5): 160-166.
- [5] Ahangari, K., Moeinossadat, S.R. and Behnia, D. (2015) Estimation of Tunnelling-Induced Settlement by Modern Intelligent Methods. *Soils and Foundations*, **55**, 737-748. <https://doi.org/10.1016/j.sandf.2015.06.006>
- [6] 吴惠明, 常佳奇, 李刚, 等. 基于支持向量机的盾构掘进姿态预测与施工参数优化方法[J]. 隧道建设(中英文), 2021, 41(S1): 11-18.
- [7] Zhou, J., Shi, X., Du, K., Qiu, X., Li, X. and Mitri, H.S. (2017) Feasibility of Random-Forest Approach for Prediction of Ground Settlements Induced by the Construction of a Shield-Driven Tunnel. *International Journal of Geomechanics*, **17**, Article 817. [https://doi.org/10.1061/\(asce\)gm.1943-5622.0000817](https://doi.org/10.1061/(asce)gm.1943-5622.0000817)
- [8] 林春金, 杨晓达, 龚英杰, 等. 基于 PSO-BP 的土压盾构土仓压力预测模型及掘进参数敏感性分析[J]. 应用基础与工程科学学报, 2021, 29(5): 1220-1233.
- [9] 吴贤国, 刘俊, 陈虹宇, 等. 盾构掘进系统的数字孪生构架体系研究[J]. 土木建筑工程信息技术, 2023, 15(4):

105-110.

- [10] 邱道宏, 傅康, 薛翊国, 等. 深埋隧道 TBM 掘进参数 LSTM 时序预测模型及应用研究[J]. 中南大学学报(自然科学版), 2021, 52(8): 2646-2660.
- [11] 肖浩汉, 陈祖煜, 徐国鑫, 等. 基于 GRU 算法的盾构掘进参数预测——以成都地铁 19 号线为例[J]. 长江科学院院报, 2023, 40(1): 123-131.
- [12] 高昆, 于思溟, 许维青, 等. 基于 Attention-ResNet-LSTM 混合神经网络的盾构掘进速度预测新方法[J]. 隧道建设(中英文), 2023, 43(4): 592-601.
- [13] 黄鸿宇. 土压平衡盾构机土仓压力影响规律与预测优化研究[D]: [硕士学位论文]. 杭州: 浙江大学, 2023.
- [14] 宋曙光. 渗流作用下复合地层盾构隧道施工开挖面稳定性及控制研究[D]: [博士学位论文]. 济南: 山东大学, 2016.
- [15] Lu, P., Yuan, D., Chen, J., Jin, D., Wu, J. and Luo, W. (2021) Face Stability Analysis of Slurry Shield Tunnels in Rock-Soil Interface Mixed Ground. *KSCE Journal of Civil Engineering*, **25**, 2250-2260. <https://doi.org/10.1007/s12205-021-1254-8>