

文章编号: 1006-2467(2024)11-1735-10

DOI: 10.16183/j.cnki.jsjtu.2023.109

基于人工神经网络的深基坑支护结构侧移预测

徐长节^{1,2,3}, 李欣雨^{1,2}

(1. 浙江大学 滨海和城市岩土工程研究中心, 杭州 310058; 2. 浙江大学 平衡建筑研究中心, 杭州 310028; 3. 华东交通大学 江西省岩土工程基础设施安全与控制重点实验室, 南昌 330013)

摘要: 为了更精准地预测基坑开挖引起的支护结构侧移, 采用支持向量机模型、传统人工神经网络模型及2种考虑时序性输入的循环神经网络模型, 建立了不同基坑支护结构最大侧移、同一基坑不同工况支护结构侧移的预测模型。结果显示, 人工神经网络可以根据工程实测数据实时更新和预测支护结构变形, 有助于及时规划工程下一步施工工艺。在支护结构不同工况侧移的预测上, 考虑了时序性输入的循环神经网络模型效果优于传统人工神经网络模型。

关键词: 基坑工程; 变形预测; 机器学习; 支护结构

中图分类号: TU46+3

文献标志码: A

Lateral Deformation Prediction of Deep Foundation Retaining Structures Based on Artificial Neural Network

XU Changjie^{1,2,3}, LI Xinyu^{1,2}

(1. Research Center of Coastal and Urban Geotechnical Engineering, Zhejiang University, Hangzhou 310058, China; 2. Center for Balance Architecture, Zhejiang University, Hangzhou 310028, China; 3. Jiangxi Key Laboratory of Infrastructure Safety Control in Geotechnical Engineering, East China Jiaotong University, Nanchang 330013, China)

Abstract: In order to more accurately predict the lateral deformation of retaining structures caused by foundation pit excavation, this paper adopts support the vector machine model, traditional artificial neural network model, and two kinds of recurrent neural network models considering temporal inputs to establish a prediction model for the maximum lateral deformation of retaining structures in different foundation pits, and for the same foundation pit under different working conditions. The results show that the artificial neural network can update and predict the deformation of the retaining structure in real time based on the measured data of the project, which is helpful for timely planning of the next construction process of the project. In the prediction of lateral deformation of retaining structures under different working conditions, the cyclic neural network model considering temporal inputs is better than the traditional artificial neural network model.

Keywords: excavation; deformation prediction; machine learning; retaining structures

收稿日期: 2023-03-27 修回日期: 2023-07-11 录用日期: 2023-07-24

基金项目: 国家自然科学基金(51878276), 浙江省自然科学基金委员会-华东院联合基金(LHZ19E080001), 浙江大学平衡建筑研究中心配套资金(20203512-10C)资助项目

作者简介: 徐长节(1972—), 教授, 博士生导师, 主要从事基坑工程支护结构方面的研究; E-mail: xucj@zju.edu.cn.

近年来,随着我国经济建设高速发展,地面空间的开发日益饱和,对地上空间的开发利用已经不能满足需求,越来越多的工程项目开始向地下发展,如地铁站、地下车库、地下商场、地下综合体,因此出现了大量与深基坑相关的工程. 在进行基坑开挖过程中,保证基坑的稳定性及施工人员的安全性十分重要,支护结构及周围环境变形过大会引起很大的安全隐患,导致安全事故的发生. 随着城市建设的愈发完善,许多工程位于人口密集、周围建筑物或管线密集的区域. 基坑的支护结构除了满足自身要求外,还需要对支护结构及周围环境的变形进行控制,减小基坑工程对周围环境的影响. 因此,有必要对基坑开挖过程中的变形规律进行研究,进一步形成支护结构变形预测方法,为基坑工程的支护结构设计及施工提供指导性意见.

基坑开挖过程中支护结构及周围环境的变形问题受到许多学者关注. 文献[1-9]中通过对土层分布、施工方法、支护方案、开挖深度等实测和模拟数据的分析,发现支护结构的最大侧移与开挖深度的关系可以由地质条件、施工方法、支护方案等条件来确定,根据不同施工条件确定的支护结构最大侧移一般为开挖深度的 $0.1\% \sim 0.87\%$. 文献[10-11]中通过分析基坑开挖时桩锚支护结构的实测数据与数值模拟结果,归纳冠梁和支护桩的变形分布规律,并利用最小势能原理,推导出支护桩桩顶最大水平位移的计算公式. 文献[12-18]中通过数值模型模拟了不同的施工状态、地质条件和支护方式,动态分析了基坑开挖不同阶段支护结构、周围环境和结构的变形,为支护结构的设计选型、设计优化和施工提供了参考数据. 随着机器学习技术的快速发展与普及,部分学者开始将机器学习应用于岩土工程领域. 张雯超等^[19]基于 70 多个苏州地铁深基坑设计及监测数据,分析施工工法、地连墙厚度等 8 个输入参数与地连墙变形的相关性. Kung 等^[20]通过对 12 个历史基坑开挖工程进行数值模拟,并通过人工神经网络(ANN)对基坑开挖过程中各个工况下地下连续墙的水平位移进行预测,通过对比数值模拟与 ANN 预测结果说明 ANN 算法可以准确预测软土地区基坑开挖过程中地下连续墙的变形情况.

针对基坑工程中支护结构侧移预测问题,已经有学者通过有限元模拟、经验公式总结、理论解析解推导、机器学习等方法进行了研究,但仍存在一定的局限性. 经验公式和理论解析解难以准确计算基坑开挖过程中各时刻支护结构的变形,在基坑开挖过程中,可能会出现内部支护失效、周边环境变化等突

发情况,通过相关公式难以充分考虑这些突发情况对围护结构变形的影响. 与上述两种方法相比,数值模拟能更好地考虑围护结构周围环境参数对围护结构变形的影响,并能模拟施工各阶段的变形,但模型的建立耗时较长,建模所需的参数范围较大. 建模值具有一定的随机性,难以验证仿真结果的准确性和有效性. 机器学习算法在岩土工程领域得到了应用,可以有效地预测基坑开挖过程中围护结构的变形和地表沉降的情况,与上述 3 种预测方法相比,该方法具有学习速度快、预测模型实时更新、预测精度高等优点. 但现有的预测模型大多是针对同一工程项目中不同时间序列的支护结构变形,不能预测不同基坑设计方案下支护结构的最大侧移,而且大多数考虑时间序列的变形预测模型只考虑了时间因素对变形的影响,忽略了其他因素对变形的影响,造成特征数据的浪费.

基于此,引入支持向量机模型和人工神经网络中的多层感知器模型,对基坑开挖引发的支护结构最大侧移进行了预测;引入长短期记忆(LSTM)和门循环单元(GRU)两种改进后的循环神经网络模型,对基坑开挖不同阶段的支护结构侧移进行了预测;并进一步探究了这几种模型的预测效果及输入步长对预测效果的影响.

1 机器学习模型

1.1 支持向量机模型

支持向量机(SVM)由 Vapnik 等^[21-24]在 20 世纪 90 年代提出. 在早期,这种方法被用于统计学习领域,可以通过对小样本数据的学习,来探究样本数据之间的规律.

SVM 的核心理论是找到数据归类的最优超平面,如图 1 所示. 图中: $f(x)$ 为最优超平面函数; w 为最优超平面的法向量; 红色(I类)和蓝色(II类)的点依次代表两类数据样本; L 为分类线; L_1 为穿过 I 类样本数据、平行于分类线且与分类线距离最近的线; L_2 为穿过 II 类样本数据、平行于分类线且与分类线距离最近的线. 最优超平面应该满足, L 能将两组数据样本归类,且 L_1 与 L_2 的相距最大.

对于回归问题,可以使用 SVM 中的支持向量机回归(SVR)模型,如图 2 所示. 图中 L_1 、 L_2 为穿过样本数据、与回归线平行且相距最远的线. 最优超平面应该满足,所有样本数据都在 L_1 与 L_2 之间的区域,且 L_1 与 L_2 的相距最小.

对于线性可分样本,可以直接在原始空间中找到最优分类面;对于非线性可分样本,通过引入核函

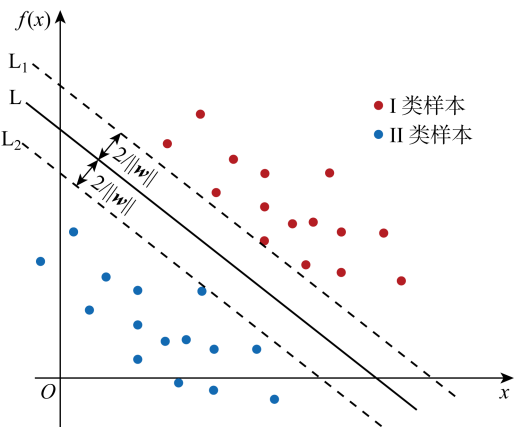


图 1 SVM 最优超平面示意图

Fig. 1 Schematic diagram of SVM optimal hyperplane

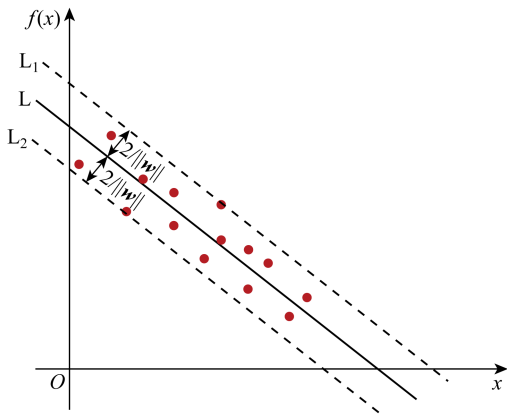


图 2 SVR 最优超平面示意图

Fig. 2 Schematic diagram of SVR optimal hyperplane

数,将输入特征量从低维空间变换到更高维的空间,并在新建的高维空间中找到最优分类面.最优超平面的计算公式如下:

$$f(x) = \text{sgn}(\mathbf{w}^* \cdot \Phi(x) + \mathbf{b}^*) =$$
$$\text{sgn}\left(\sum_{j=1}^l a_j^* y_j \Phi(x_i) \cdot \Phi(x_j) + \mathbf{b}^*\right),$$
$$\Phi(x) \in H$$

(1)

式中: y_j 为 x_j 的类标记; \mathbf{w}^* 为最优权值向量; \mathbf{b}^* 为最优偏置向量; a_j^* 为拉格朗日乘数最优解; l 为数据样本集的数量; $\Phi(x)$ 为特征向量; H 为特征空间.

1.2 多层感知器模型

多层感知器(MLP)模型是一种基本的 ANN 模型.最早由 Rosenblatt^[25] 在 20 世纪 60 年代提出. MLP 模仿人脑在处理问题时的工作机制,大脑由多个神经元连接在一起,当一个神经元接收到来自其他神经元的信号时,它会产生刺激,将信号经过处理后传输给其他神经元,神经元重复地传输信号,并且不断调整反馈输入和输出信号之间的关系.

在处理回归问题时,MLP 的基本结构如图 3 所示.神经网络由输入层、隐藏层和输出层组成.隐藏层的数量根据问题的情况而变化.特征数据通过输入层输入,输入数据由隐藏层神经元进行学习和处理,神经元的输入输出值通过决策函数进行变换,最后通过输出层输出预测结果.一般来说,隐层的层数越多,每层神经元的个数越多,模型的拟合能力越强,但过多的隐藏层和过多的神经元也可能导致过拟合.

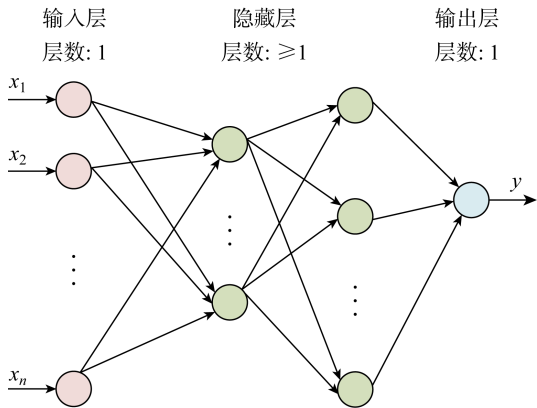


图 3 MLP 模型示意图

Fig. 3 Schematic diagram of MLP model

神经元上的决策函数模拟了人脑处理问题时激活神经元的过程,因此也叫激活函数.激活函数可以在神经元中引入非线性因素,使神经网络能够处理非线性问题.在本文中,使用 ReLU 激活函数,如下式所示:

$$f_{\max}(x) = \max\{0, x\}$$

(2)

当输入值小于或等于 0 时,下一层中的神经元将不被激活.当输入值大于 0 时,下一层的神经元将被激活.这种激活函数在运算过程中收敛速度快,在机器学习领域得到了广泛的应用.

1.3 长短期记忆网络模型

LSTM 最早由 Hochreiter 和 Schmidhuber^[26] 在 1997 年提出. LSTM 模型是一种考虑时序性输入参数的人工神经网络模型,与 MLP 相比,该模型考虑了不同时序输出量之间的影响,在预测时序性数据时更准确.与传统的循环神经网络(RNN)模型相比,LSTM 引入了门控单元,解决了传统 RNN 模型在输入数据过大时梯度可能消失的问题.

图 4 是传统 RNN 模型的示意图.图中: y_t, y_{t-1}, y_{t+1} 为时间 $t, t-1, t+1$ 的输出值; s_t, s_{t-1}, s_{t+1} 分别为时间 $t, t-1, t+1$ 的记忆; x_t, x_{t-1}, x_{t+1} 为时间 $t, t-1, t+1$ 的输入值; U, V, W 为权重参数.

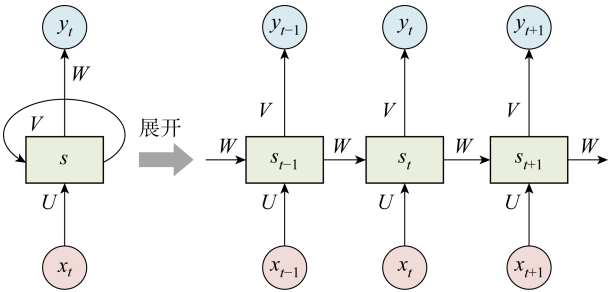


图 4 传统 RNN 模型示意图

Fig. 4 Schematic diagram of traditional RNN model

时间 t 的输出值计算如下：

$$y_t = G(Vs_t) \tag{3}$$

$$s_t = F(Ux_t + Ws_{t-1}) \tag{4}$$

式中： $F(x)$ 、 $G(x)$ 为激活函数，其中 $F(x)$ 通常使用 tanh、ReLU、Sigmoid 激活函数， $G(x)$ 通常使用 Softmax 激活函数。

如图 5 所示，LSTM 单元由遗忘门、输入门、输出门组成。遗忘门通过 Sigmoid 函数对 x_t 、 y_{t-1} 进行激活，Sigmoid 函数的输出值在 0~1 之间，只有当 $\sigma=1$ 时，对应的信息才会被完整的保存下来；输入门对输入值和前一个单元格的状态执行更新；输出门通过对输入值和当前单元状态的处理，得到时间 t 的输出值。

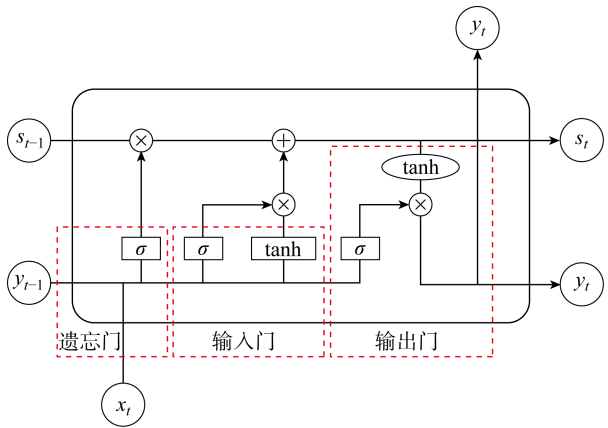


图 5 LSTM 单元示意图

Fig. 5 Schematic diagram of LSTM unit

1.4 门循环单元模型

GRU 模型是 LSTM 的简化模型，由 Cho 等^[27]于 2014 年首次提出。与 LSTM 模型相比，GRU 模型结构更简单，结构参数更少，因此在同等条件下，GRU 模型计算量更小，训练时间更短。

如图 6 所示，GRU 单元由重置门、更新门组成。重置门对输入值和上个单元的隐藏状态进行激活。

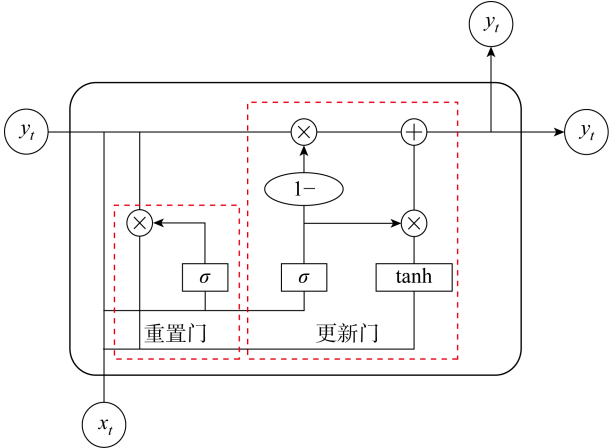


图 6 GRU 单元示意图

Fig. 6 Schematic diagram of GRU unit

2 支护结构最大侧移预测模型

2.1 数据集构建

支护结构最大水平位移预测模型的研究资料取自徐中华^[28]根据岩土工程领域已出版的专著、各类期刊和论文编制的上海地区基坑工程变形实测案例数据库，数据集包括上海地区的 89 个基坑工程的实测数据。该数据集包括水文地质参数、施工参数、基坑设计参数和测量的最大水平位移。在本节中，通过 SVM 和 MLP 模型来预测支护结构的最大侧移。随机选取总数据集 70% 的数据（即 62 个）作为模型的训练集，剩余 30% 的数据（即 27 个）作为模型的测试集。根据国内外对基坑支护结构变形影响因素的研究和数据库涉及实际工程的情况，主要关注基坑设计参数对支护结构侧移的影响，选取基坑开挖深度 x_1 、地下连续墙厚度 x_2 、支护结构长度 x_3 、支护结构单位长度抗弯刚度 x_4 、内支撑数量 x_5 、首道内支撑距离地面深度 x_6 、内支撑平均间隔 x_7 、支护类型 x_8 这 8 个主要因素作为模型的输入量，支护结构最大侧移 y 为输出量。 $x_8 = 1$ 时对应支护结构与地下主体结构相结合的支护方式； $x_8 = 2$ 时对应顺作法的地下连续墙支护方式。

由于数据集的参数范围和单位的差异很大，所以需要数据集的参数进行预处理，以加快神经网络模型的收敛速度。对于参数 $x_1 \sim x_7$ 和 y 使用归一化方法对数据进行预处理，使这些参数的数值分布在 $[0, 1]$ 的范围内。具体计算方法如下式：

$$x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{5}$$

式中： x_{norm} 为归一化处理后的参数； x 为归一化处理前的参数； x_{\max} 为归一化处理前的最大参数； x_{\min} 为

归一化处理前的最小参数。

对于 x_8 该分类特征类型的参数,通过独热编码对数据进行预处理,并通过 2 位状态寄存器对 2 个支持类型的状态进行编码。

2.2 模型建立及训练

采用 SVM 和 MLP 模型支护结构最大侧移的预测进行训练,并随机选取样本数据的 10% 作为验证集,以验证训练模型的准确性。

通过随机搜索算法和参数优化经验调整模型所需的超参数。在优化过程中,通过损失函数监督神经网络的学习过程,计算预测模型在当前迭代次数下的损失。对于回归问题,经常使用均方误差损失函数,计算公式如下:

$$E_{MSE}(g, g') = \frac{1}{m} \sum_{i=1}^m (g_i - g'_i)^2 \tag{6}$$

式中: m 为样本数据数量; g 为实测的输出值; g' 为预测的输出值。对于同一模型, $E_{MSE}(g, g')$ 取值越小,模型准确度越高

在训练过程中,为了防止训练模型过拟合,增加 L2 正则化,并对原损失函数项和正则项比例进行加权,优化原损失函数。计算公式如下:

$$f'_{loss}(x) = f_{loss}(x) + \frac{\lambda}{2m} \sum w_i^2 \tag{7}$$

式中: $f_{loss}(x)$ 为原有的损失函数; $f'_{loss}(x)$ 为 L2 正则化后的损失函数; λ 为正则项的权重。

最终确定 SVM 模型参数取值:惩罚系数为 1、核函数为 Gauss 径向基核函数、回归精度 epsilon 为 0.1。MLP 模型参数取值:隐藏层层数为 3、神经元数量为 64、学习率为 0.000 5、激活函数为 ReLU 函

数、训练步数为 5 000。

3 不同阶段支护结构侧移预测模型

3.1 数据集构建

不同阶段支护结构水平位移预测模型数据取自杭州市某高铁站基坑工程实测数据,该车站为地下三层双柱三岛式车站,基坑总长约 457 m,开挖宽度 22 m,开挖深度 26.05 m,基坑底部主要位于粉砂层。基坑标准段采用明挖顺作法施工,支护结构由 1 m 厚的地下连续墙、7 道内支撑及 2 个换钢支撑组成。数据集选择 CX23 监测点的支护结构侧移数据进行预测和分析。总共收集 80 期数据,因用于预测的模型考虑了时序性输入,所以按照测量时间的顺序对训练集和测试集进行分组更合理,前 64 期数据作为训练集,后 16 期数据作为测试集。由于施工过程中各监测点周围的水文地质条件及基坑设计参数保持不变,对支护结构变形的影响可以忽略不计,本节预测模型的影响因素主要考虑施工参数。根据国内外对基坑支护结构变形影响因素的研究,结合工程实际情况,选取基坑开挖深度 X_1 、开挖阶段 X_2 、第 1 道混凝土钢结构支撑轴力 X_3 、第 5 道混凝土钢结构支撑轴力 X_4 、第 2 道钢支撑轴力 X_5 、第 3 道钢支撑轴力 X_6 、第 4 道钢支撑轴力 X_7 、距离开工时间 X_8 、侧移变化速率 X_9 这 9 个主要因素作为模型的输入,支护结构侧移 Y 为输出量。

部分样本数据如表 1 所示。对于参数 $X_1 \sim X_9$ 和 Y ,使用归一化方法对数据进行预处理,使这些参数的数值分布在 $[0, 1]$ 的范围内。具体计算方法见式(5)。

表 1 不同阶段支护结构侧移预测模型部分样本数据表

Tab. 1 Partial sample data for predicting lateral displacement of retaining structures in different stages										
编号	X_1/m	X_2	X_3/kN	X_4/kN	X_5/kN	X_6/kN	X_7/kN	X_8/d	$X_9/(\text{mm} \cdot \text{h}^{-1})$	Y/mm
1	5	1	1 913.59	0	0	0	0	22	−0.22	−0.66
2	5	1	2 369.11	0	0	0	0	25	0.43	0.63
3	5	1	2 309.13	0	0	0	0	27	0.37	1.37
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
78	26.05	7	1 870.24	7 096.59	968.79	845.28	1 107.62	118	1.54	58.14
79	26.05	7	1 677.10	7 336.54	925.58	791.38	1 100.69	119	−1.55	56.59
80	26.05	7	2 028.95	7 431.26	978.13	870.61	1 152.58	120	−0.4	56.19

对于 X_2 该分类特征类型的参数,通过独热编码对数据进行预处理,并通过 7 位状态寄存器对 7

个施工阶段进行编码, X_2 取值及对应施工阶段如表 2 所示。表中: h 为基坑开挖深度。

表 2 X_2 预处理后的参数取值及对应施工阶段

Tab. 2 Pre-processed parameter values of X_2 and corresponding construction stages

独热编码前的参数	独热编码后的参数	对应的施工阶段
1	1000000	$h \leq 5$ m
2	0100000	$5 \text{ m} < h \leq 10$ m
3	0010000	$10 \text{ m} < h \leq 15$ m
4	0001000	$15 \text{ m} < h \leq 26.05$ m
5	0000100	$h = 26.05$ m
6	0000010	架设第 6 道支撑
7	0000001	架设第 7、8 道支撑

3.2 模型建立及训练

采用 MLP、LSTM、GRU 模型不同阶段支护结构侧移的预测进行训练,并随机选取样本数据的 10%作为验证集,以验证训练模型的准确性.

通过随机搜索算法和参数优化经验调整模型所需的超参数.在优化过程中,利用均方误差损失函数,计算公式见式(6),来监督神经网络的学习过程,计算预测模型在当前迭代次数下的损失.为了训练模型过拟合,向损失函数添加 L2 正则化项以形成新的损失函数,计算公式见式(7).

最终确定各模型参数取值如表 3 所示,在该参数组合下的损失函数如图 7 所示.由图 7 可知,随着训练步数的增加,训练集的损失逐渐减少. MLP 模

型在训练步数达到约4 500时逐渐收缩,LSTM 模型在训练步数达到约13 000时逐渐收敛,GRU 模型在训练步数达到约8 000时逐渐收敛,损失减小到最小值附近. LSTM 和 GRU 模型有两个收敛平台,这可能是由于训练数据集分布不均匀,导致训练过程中出现局部最优解.在第 1 个收敛平台阶段,模型参数更新速度较快,但随着训练步长的增加,模型仍处于局部最优解状态,损失函数值基本不变.在第 2 个收敛平台中,模型参数的更新速度变慢,但随着训练步长的增加,模型跳出了局部最优解,模型的表达能力得到了提高.此时,模型具有较好的泛化能力,可用于预测新的数据.

表 3 不同阶段支护结构侧移预测模型参数取值

Tab. 3 Parameter values of lateral displacement prediction models for support structures in different stages

模型	隐藏层 层数	神经元 数量	步长	学习率	激活函数	训练 步数
MLP	3	64	—	0.000 1	ReLU	4 500
LSTM	1	64	4	0.000 1	ReLU	13 000
					Sigmoid	13 000
					tanh	13 000
GRU	1	64	6	0.000 5	ReLU	8 000
					Sigmoid	8 000
					tanh	8 000

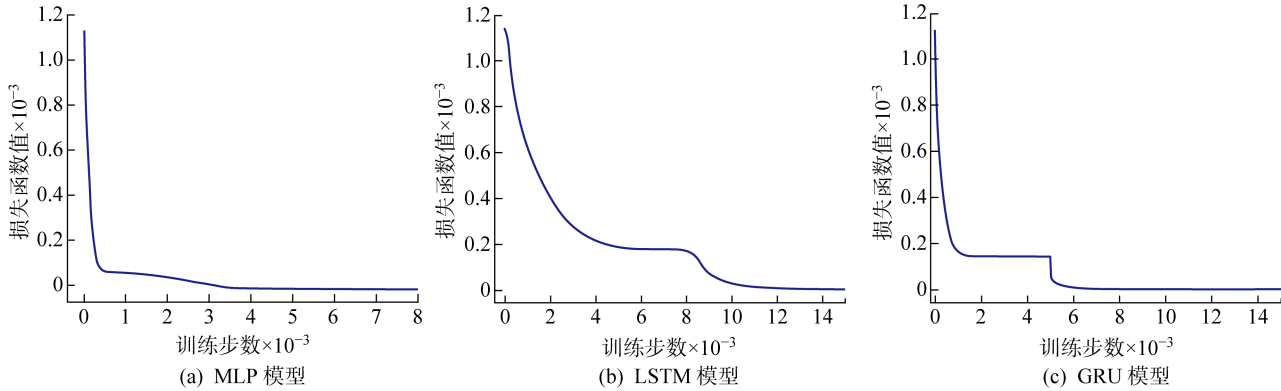


图 7 损失函数与训练步数关系

Fig. 7 Relation between loss function and training steps

4 预测结果分析

4.1 支护结构最大侧移预测

基于 2.2 节中的最优参数值,分别采用 SVM 和 MLP 模型对不同基坑支护结构的最大侧移进行预测.训练数据来自文献[28]中随机选择的 62 组数据,经过训练和交叉验证,得到上述两个模型对应的

预测模型,并对样本数据剩余的 27 组数据进行预测,图 8 为各模型最优参数组合下测试集回归效果对比图.

从图 8 可以看出,SVM 模型和 MLP 模型在预测支护结构最大侧移时表现较好,其中 MLP 模型预测值与实测值的误差主要在 10%以内,SVM 模型的误差范围主要在 20%以内.相比之下,MLP 模

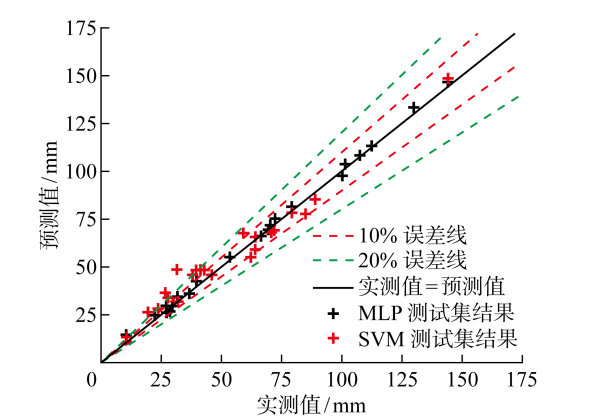


图 8 各模型测试集回归效果

Fig.8 Regression effect of each model test set

型在预测包络线最大位移时具有更高的准确度和精度.这是因为与 SVM 模型相比,MLP 模型具有更好的泛化能力和更精确的预测模式,它的神经元结构具有更强的学习能力,可以学习更多的特征信息,从而提高预测精度.

表 4 给出了各模型预测结果的评价指标值.表中: R^2 为决定系数,

$$R^2(g, g') = 1 - \frac{\sum_{i=1}^m (g_i - g'_i)^2}{\sum_{i=1}^m (g_i - \bar{g})^2} \tag{8}$$

式中: \bar{g} 为实际输出值平均值. $R^2(g, g')$ 取值在 $[0, 1]$ 的范围内,取值大小可以反映输入参数对输出参数的解释程度,值越大,模型拟合效果越好.通过比较两个模型的 R^2 值,可以看出 MLP 的整体预测效果优于 SVM.

表 4 支护结构最大侧移预测模型评价指标

Tab.4 Evaluation indicators of maximum lateral displacement prediction model for retaining structures

模型	E_{MSE}	R^2
SVM	4.99	0.99
MLP	36.96	0.95

综上所述,MLP 模型在支护结构最大侧移预测的表现更好,是一个更可靠和准确的预测模型.在实际基坑工程中,可优先考虑采用基于机器学习的 MLP 模型,为工程的稳定性、安全性分析提供可靠的参考数据.

4.2 不同阶段支护结构侧移预测

基于表 3 中的最优参数值,分别使用 MLP、LSTM 和 GRU 模型预测水平位移不同阶段的包络结构.训练数据来自对应于表 1 中编号 1~64 的数

据.通过训练和交叉验证得到上述 3 个模型对应的预测模型,并对样本数据的最后 16 组数据进行预测.预测结果与实测对比如图 9 所示.

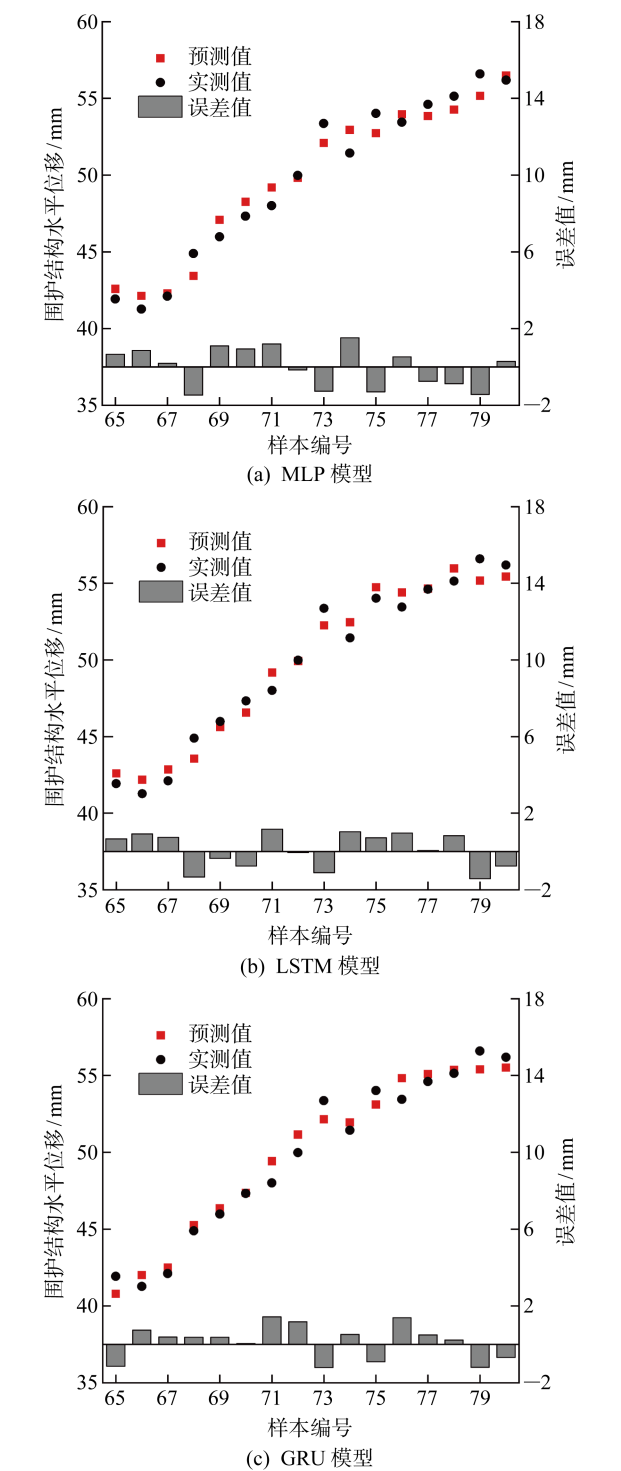


图 9 预测结果与实测对比

Fig.9 Comparison of predicted results and actual measurements

由图 9(a)可以看出,MLP 模型的预测误差范围约为 $-1.46 \sim 1.50$ mm,其中最大绝对误差值为 1.50 mm,最小绝对误差值为 0.15 mm.由图 9(b)

可以看出, LSTM 模型的预测误差范围约为 $-1.42\sim 1.17\text{ mm}$, 其中最大误差的绝对值为 1.42 mm , 最小误差的绝对值为 0.05 mm . 由图 9(c) 可以看出, GRU 模型的预测误差范围约为 $-1.21\sim 1.43\text{ mm}$, 其中最大绝对误差值为 1.43 mm , 最小绝对误差值为 0.04 mm . 根据 3 种模型的预测结果, MLP、LSTM 和 GRU 模型的支护结构侧移的预测值与实测值一致, 预测误差范围在 1.5 mm 以内, 满足实际工程对支护结构侧移测量误差的要求. 这 3 种模型对基坑开挖过程中不同阶段的支护结构侧移预测具有较高的精度.

表 5 给出了每个模型的均方误差、决定系数和预测时间. 通过比较 3 种模型的 R^2 值和预测用时, 可以看出总体预测效果 $\text{GRU} > \text{LSTM} > \text{MLP}$. LSTM 和 GRU 模型考虑了时间序列对产出结果的影响. 与 MLP 模型相比, 输入参数的特征更接近基坑开挖的真实情况, 预测结果更准确. 通过比较 3 种模型的预测时间, 可看出预测速度 $\text{MLP} > \text{GRU} > \text{LSTM}$, 与其他两个 RNN 模型相比, MLP 的神经元数量较少, 每个神经元之间只有一个连接, 没有回路结构, 在处理数据时可以并行计算每个神经元的输出. 与 LSTM 模型相比, GRU 的单元结构更简单, 这使得 GRU 的计算量小于 LSTM, 此外, GRU 的参数数量也比 LSTM 少, 减少了内存消耗和大模型存储, 加快了数据处理速度.

表 5 不同阶段支护结构侧移预测模型评价指标
Tab.5 Evaluation indicators of lateral displacement prediction models for support structures in different stages

模型	E_{MSE}	R^2	预测用时/min
MLP	1.01	0.96	1.7
LSTM	0.77	0.97	4.7
GRU	0.75	0.97	4.3

4.3 输入步长分析

LSTM 和 GRU 考虑了输入参数按时间序列顺序输入的影响, 与 MLP 模型相比增加了步长参数的概念. 例如, 在预测时间序列 t 的支护结构侧移时, 需要考虑从 $t-n$ 到 $t-1$ 时间序列的隐藏状态对时间序列 t 侧移的影响, 其中 n 是输入步长参数. 为了探索步长参数对预测效果的影响, 在保持其他参数不变的情况下, 改变模型的步长, 得到不同取值的损失函数, 如图 10 所示.

从图 10(a) 可以看出, 对于 LSTM 模型, 步长为 4 的预测模型具有最小的损失函数值; 当步长小于 4 时, 预测模型的损失函数随着步长的增大而减小, 向

预测模型输入更多的时间序列较小的序列数据, 可以为模型带来更多的用于学习的历史数据, 使得预测模型更加准确; 当步长大于 4 时, 预测模型的损失函数随着步长的增大而增大, 用训练集训练的模型比用步长为 4 训练的模型更接近过拟合现象. 从图 10(b) 可以看出, 对于 GRU 模型, 步长为 6 预测模型具有最小的损失函数值; 当步长小于 6 时, 预测模型的损失函数随着步长的增大而减小; 当步长大于 6 时, 预测模型的损失函数随步长的增大而增大.

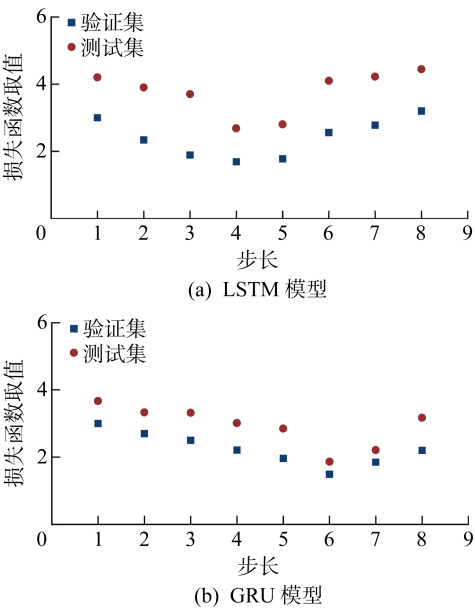


图 10 不同步长对应的损失函数值
Fig. 10 Values of loss function corresponding to different steps

综上所述, 对于考虑时间序列的支护结构侧移预测模型, 增加输入步长可以在一定程度上提高模型预测精度, 但步长过大也会导致模型过拟合. 对于不同的模型, 最优输入步长的取值可能不同, 这是因为不同的模型设置了不同的超参数, 比如学习率、隐层数、神经元数等, 这些参数的不同设置可能导致需要不同的输入序列长度以获得更好的性能.

5 结论

- (1) SVM 和 MLP 模型在预测不同基坑支护结构最大侧移时有较好表现, 预测误差大部分在 $\pm 20\%$ 范围内. 与 SVM 模型相比, MLP 神经元结构具有更强的学习能力, 可以学习到更多的特征信息, 预测精度更高.
- (2) MLP、LSTM 和 GRU 模型对不同阶段支护结构变形的预测结果误差范围均在 1.5 mm 以内, 3 种模型均能准确预测基坑开挖过程中支护结

构的变形,满足实际工程测量精度要求.与传统的神神经网络 MLP 模型相比,LSTM 和 GRU 两个 RNN 模型考虑了时间输入对输出的影响,能够捕捉到时间输入之间的依赖关系,更接近基坑开挖过程中支护结构变形的实际情况,预测效果更好.

(3) 对于考虑时间序列输入的支护结构侧移预测模型,增大输入步长可以在一定程度上提高模型的预测精度,但步长过大也会导致模型过拟合.对于不同的模型,由于模型参数的不同,最优输入步长可能不同.

(4) 提出的基于机器学习的变形预测方法与传统的有限元模拟、公式计算等变形预测方法相比,计算速度更快,并可根据工程实测数据实时更新和预测,有助于及时规划工程下一步施工工艺.

参考文献:

[1] 徐中华,王建华,王卫东.上海地区深基坑工程中地下连续墙的变形性状[J]. *土木工程学报*, 2008(8): 81-86.
XU Zhonghua, WANG Jianhua, WANG Weidong. Deformation behavior of diaphragm walls in deep excavations in Shanghai[J]. *China Civil Engineering Journal*, 2008(8): 81-86.

[2] 陈保国,闫腾飞,王程鹏,等.深基坑地连墙支护体系协调变形规律试验研究[J]. *岩土力学*, 2020, 41(10): 3289-3299.
CHENG Baoguo, YAN Tengfei, WANG Chengpeng, *et al.* Experimental study on compatible deformation of diaphragm wall support system for deep foundation pit[J]. *Rock and Soil Mechanics*, 2020, 41(10): 3289-3299.

[3] PENG X, KIANOOSH H. Sliding stability and lateral displacement analysis of reinforced soil retaining walls[J]. *Geotextiles and Geomembranes*, 2019, 47: 483-492.

[4] 孙小力,孙铁成,张旭,等.地铁基坑开挖数值模拟及变形特征研究[J]. *施工技术*, 2020, 49(7): 41-44.
SUN Xiaoli, SUN Tiecheng, ZHANG Xu, *et al.* Research on numerical simulation and deformation characteristics of subway foundation excavation[J]. *Construction Technology*, 2020, 49(7): 41-44.

[5] RITWIK N, DEEPANKAR C. Displacement-controlled approach for the analysis of embedded cantilever retaining walls with a distanced strip surcharge[J]. *Computers and Geotechnics*, 2022, 151: 104970.

[6] 乔世范,蔡子勇,张震,等.南沙港区软土狭长深基坑围护体系性状[J]. *浙江大学学报(工学版)*, 2022,

56(8): 1473-1484.

QIAO Shifan, CAI Ziyong, ZHANG Zhen, *et al.* Behavior of retaining system of narrow-long deep foundation pit in soft in Nansha Port Area[J]. *Journal of Zhejiang University (Engineering Science)*, 2022, 56(8): 1473-1484.

[7] 秦会来,黄俊,李奇志,等.深厚淤泥地层深基坑变形影响因素分析[J]. *岩土工程学报*, 2021, 43(Sup. 2): 23-26.

QIN Huilai, HUANG Jun, LI Qizhi, *et al.* Influencing factors for deformation of deep foundation pits in thick mud stratum[J]. *Chinese Journal of Geotechnical Engineering*, 2021, 43(Sup. 2): 23-26.

[8] 刘新荣,王林枫,陈峰,等.内撑式地连墙变形特征及其参数优化研究[J]. *地下空间与工程学报*, 2021, 17(3): 727-738.

LIU Xinrong, WANG Linfeng, CHEN Feng, *et al.* Research on deformation characteristics and parameter optimization of diaphragm wall with interior bracing structure[J]. *Chinese Journal of Underground Space and Engineering*, 2021, 17(3): 727-738.

[9] 钟俊辉,尹培林,滕超,等.福州软土地区地铁车站基坑墙体变形特性分析[J]. *地下空间与工程学报*, 2018, 14(Sup. 2): 821-827.

ZHONG Junhui, YIN Peilin, TENG Chao, *et al.* The diaphragm walls of deformation properties of excavation for metro stations in Fuzhou soft soil deposits[J]. *Chinese Journal of Underground Space and Engineering*, 2018, 14(Sup. 2): 821-827.

[10] 申正洋,黄华,李新生,等.深大基坑中冠梁的变形特征及一种新型冠梁[J]. *工程地质学报*, 2023, 31(2): 661-670.

HUANG Zhengyang, HUANG Hua, LI Xinsheng, *et al.* Deformation characteristics of conventional and new top beams in deep and large foundation pit with numerical simulations[J]. *Journal of Engineering Geology*, 2023, 31(2): 661-670.

[11] 居玥辰,宫全美,赵昱,等.软土基坑开挖引起地下连续墙水平变形的能量法[J]. *地下空间与工程学报*, 2021, 17(6): 1762-1774.

JU Yuechen, GONG Quanmei, ZHAO Yu, *et al.* Energy method for horizontal deformation of diaphragm wall due to soft soil foundation pit excavation[J]. *Chinese Journal of Underground Space and Engineering*, 2021, 17(6): 1762-1774.

[12] 张福海,周玉石,刘学港,等.基于斜桩支护的框架结构基坑变形分析[J]. *山东大学学报(工学版)*, 2022, 52(1): 58-65.

ZHANG Fuhai, ZHOU Yushi, LIU Xuegang, *et al.*

- Deformation analysis of frame structure foundation pit based on inclined pile support[J]. **Journal of Shandong University (Engineering Science)**, 2022, 52(1): 58-65.
- [13] 涂芬芬. 深基坑开挖对邻近既有建筑变形影响研究[J]. **土工基础**, 2022, 36(1): 5-8.
- TU Fenfen. Influence of deep excavation induced deformation on the adjacent buildings[J]. **Soil Engineering and Foundation**, 2022, 36(1): 5-8.
- [14] 尹利洁, 李宇杰, 朱彦鹏, 等. 兰州地铁雁园路站基坑支护监测与数值模拟分析[J]. **岩土工程学报**, 2021, 43(Sup. 1): 111-116.
- YIN Lijie, LI Yujie, ZHU Yanpeng, *et al.* Monitoring and numerical simulation of support for foundation pit at Yanyuan Road Station of Lanzhou Metro [J]. **Chinese Journal of Geotechnical Engineering**, 2021, 43(Sup. 1): 111-116.
- [15] 周勇, 王旭日, 朱彦鹏, 等. 强风化软硬互层岩质高边坡监测与数值模拟[J]. **岩土力学**, 2018, 39(6): 2249-2258.
- ZHOU Yong, WANG Xuri, ZHU Yanpeng, *et al.* Monitoring and numerical simulation of an interbedding high slope composed of soft and hard strong-weathered rock[J]. **Rock and Soil Mechanics**, 2018, 39(6): 2249-2258.
- [16] 张超翔, 张志强. 深基坑桩锚支护结构位移分析及数值模拟[J]. **科学技术与工程**, 2022, 22(18): 8022-8029.
- ZHANG Chaoxiang, ZHANG Zhiqiang. Displacement analysis and numerical simulation of pile-anchor retaining structure in deep foundation pit[J]. **Science Technology and Engineering**, 2022, 22(18): 8022-8029.
- [17] 蔡建军, 谢臻, 李树忱, 等. 复杂条件下深基坑多层支护方法及数值模拟研究[J]. **工程力学**, 2018, 35(2): 188-194.
- CAI Jianjun, XIE Can, LI Shuchen, *et al.* Multi-larger supporting method and numerical simulation for deep foundation pit under complex condition[J]. **Engineering Mechanics**, 2018, 35(2): 188-194.
- [18] 王恩钰, 周海祚, 郑刚, 等. 基坑倾斜桩支护的变形数值分析[J]. **岩土工程学报**, 2019, 41(Sup. 1): 73-76.
- WANG Enyu, ZHOU Haizuo, ZHENG Gang, *et al.* Numerical analyses of deformation of inclined pile-retained excavations[J]. **Chinese Journal of Geotechnical Engineering**, 2019, 41(Sup. 1): 73-76.
- [19] 张雯超, 史培新, 刘维, 等. 基于改进 KNN 与基坑参数对地连墙变形预测研究[J]. **华中科技大学学报(自然科学版)**, 2021, 49(9): 101-106.
- ZHANG Wenchao, SHI Peixin, LIU Wei, *et al.* Research on deformation prediction of diaphragm wall based on improved KNN and parameters of subway deep excavation[J]. **Journal of Huazhong University of science and Technology (Natural Science Edition)**, 2021, 49(9): 101-106.
- [20] KUNG G T C, HSIAO E C L, SCHUSTER M, *et al.* A neural network approach to estimating deflection of diaphragm walls caused by excavation in clays [J]. **Computers & Geotechnics**, 2007, 34(5): 385-396.
- [21] VAPNIK V N. The nature of statistical learning theory[M]. New York, USA: Springer-Verlag, 1995.
- [22] BOSER B, GUYONG I, VAPNIK V. A training algorithm for optional margin classifiers[C] // **Proceedings of The Fifth Annual Workshop on Computational Learning Theory**. New York, USA: ACM Press, 1992: 144-152.
- [23] CORTES C, VAPNIK V. Support-vector networks [J]. **Machine Learning**, 1995, 20: 273-297.
- [24] SCHOLKOPF B, BURGESS C, VAPNIK V. Extracting support data for a given task[C] // **Proceedings of First International Conference on Knowledge Discovery & Data Mining**. California, USA: AAAI Press, 1995: 252-257.
- [25] ROSENBLATT F. The perceptron: A probabilistic model for information storage and organization in the brain[J]. **Psychological Review**, 1958, 65(6): 386-408.
- [26] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. **Neural Computation**, 1997, 9(8): 1735-1780.
- [27] CHO K, VAN M B, GULCEHRE C, *et al.* Learning phrase representations using RNN encoder-decoder for statistical machine translation[C] // **Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing**. Maryland, USA: ACL, 2014: 1724-1734.
- [28] 徐中华. 上海地区支护结构与主体地下结构相结合的深基坑变形性状研究[D]. 上海: 上海交通大学, 2007.
- XU Zhonghua. Deformation behavior of deep excavations supported by permanent structure in Shanghai soft deposit[D]. Shanghai: Shanghai Jiao Tong University, 2007.

(本文编辑: 王一凡)