

基于 GRU 神经网络的软土地区公路 路基沉降预测方法

刘晓剑, 金可云, 温勇兵

(中国铁建投资集团有限公司, 广东 珠海 519000)

摘要: 采用门控递归单元 (GRU) 神经网络生成预测模型, 依据 Nesterov-accelerated Adam (Nadam) 算法学习软土地区公路路基沉降数据的演变规律, 采用沉降监测历史数据和实时数据进行训练, 并在训练结束后再次分析实时监测数据, 预测下一阶段的沉降变形; 结合湖南某高速公路工程路基沉降数据, 利用该模型对软土地区高速公路路基不均匀沉降进行预测, 并与沉降实测值进行对比, 分析预测模型的准确性和可靠性。结果表明, 模型预测值与路基沉降实测值相吻合, 基于 GRU 的路基沉降预测方法可靠。

关键词: 公路; 路基; 沉降预测; 软土地区; 门控递归单元 (GRU) 神经网络

中图分类号: U416.1

文献标志码: A

文章编号: 1671-2668(2023)02-0062-06

路基填筑过程中, 软土地基的沉降会对路基和上部结构产生破坏^[1-2], 在下一层路基填筑前, 对现有路基结构进行沉降监测和预测至关重要。保守的沉降预测会增加施工成本, 低估路基沉降则会增加施工风险, 甚至导致事故, 基于现场沉降监测数据进行阶段性更新预测的方法更具实用价值。如何计算因填土高度增加而引起的路基沉降是阶段性更新预测方法的核心。预测沉降的常用方法为有限元法^[3], 但其预测准确性受很多因素影响, 如地下水水位 (不同路段的路基水位不同, 预测中难以计算地下水水位)、土体构成模型的选择^[4-5]、合理参数的确定^[6]、土体的空间变化^[7]。将有限元模拟与贝叶斯方法相结合, 可实现阶段性更新预测, 但该方法需进行大量数值计算, 实用性差。另一种预测沉降的方法为半经验模型^[8], 其缺点是只能预测地表最大沉降。机器学习方法可分析现场沉降监测数据的演变趋势, 与有限元模拟和半经验模型相比, 该方法能学习现场沉降数据的特征, 进而预测后续施工阶段的沉降变形^[3,9]。该方法不需要进行大量数值计算^[10-11], 但其准确性在很大程度上依赖于所用算法的预先构成 (preformation), 传统的机器学习方法不能满足路基沉降预测需要, 需采用先进的深度学习算法预测路基沉降。门控递归单元 (GRU) 属于递归神经网络 (RNN)。理论上, RNN 更适用于捕捉连续数据类型之间的关系, 但由于 (随机) 梯度的

消失或爆炸, 简单的 RNN 难以捕获长期的依存性。为解决梯度消失或爆炸的问题, Chung J. Y. 等提出长短时记忆 (LSTM) 单元 GRU-RNN^[12]。在 GRU 中, LSTM 的输入门和遗忘门合并为一个门操作, 不仅使 GRU 维持了与 LSTM 相似的计算性能, 还降低了 GRU 的权重, 因而计算成本降低。本文采用基于 GRU 神经网络的路基沉降预测方法, 利用现场沉降监测数据对软土地区公路路基沉降进行预测。

1 基于 GRU 的更新程序框架

1.1 GRU 网络

如图 1 所示, 在 t 时间步, 在 GRU 的一个隐藏节点有两种门操作, 分别为更新门 z_t 和重置门 r_t 。与 LSTM 类似, 当前隐藏输出 h_t 为当前输入 x_t 和前一个隐藏输出 h_{t-1} 的计算结果。

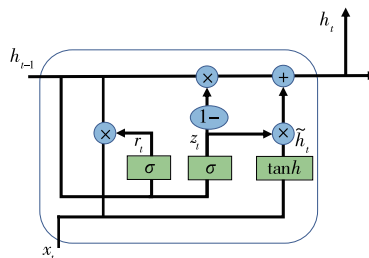


图 1 GRU 网络示意图

重置门公式如下:

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (1)$$

更新门公式如下:

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \quad (2)$$

隐藏状态(记忆)可表示为:

$$\begin{cases} h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \\ \tilde{h}_t = \tanh[Wx_t + U(r_t \odot h_{t-1}) + b] \end{cases} \quad (3)$$

全连接层到输出层的输出可表示为:

$$y_t = \text{ELU}(W_o h_t + b_o) \quad (4)$$

式中: σ 表示sigmoid函数; W_r 、 W_z 、 W 、 W_o 为与输入 x_t 相关的GRU网络的权重矩阵; x_t 为输入向量; U_r 、 U_z 、 U 为与隐藏状态 h_{t-1} 相关的GRU网络的权重矩阵; b_r 、 b_z 、 b 、 b_o 为偏差项;操作 \odot 代表Hadamard乘积; \tanh 表示双曲函数; y_t 为输出向量;ELU表示指数线性单元。

Sigmoid函数 σ 和双曲函数 \tanh 定义如下:

$$\sigma(x) = 1/(1 + e^{-x}) \quad (5)$$

$$\tanh(x) = (e^x - e^{-x})/(e^x + e^{-x}) \quad (6)$$

指数线性单元ELU定义为:

$$f(x) = \begin{cases} x & (x > 0) \\ \alpha(e^x - 1) & (x \leq 0) \end{cases} \quad (7)$$

式中: α 为超参数,本文取 $\alpha=0.01$ 。

ELU通过使平均激活更接近于零加快学习速度,ELU可获得比整流线性单元(ReLU)更高的精度:

$$f(x) = \begin{cases} x & (x > 0) \\ 0 & (x \leq 0) \end{cases} \quad (8)$$

门操作以外的激活函数可根据实际问题进行更改。

1.2 通过现场沉降监测数据更新程序

基于GRU的路基沉降预测模型见图2,更新预测流程见图3。在路基填筑第1、第2阶段监测路基沉降,将监测数据作为训练数据集训练GRU预测模型,并学习这两个阶段的路基变形模式;在后续填

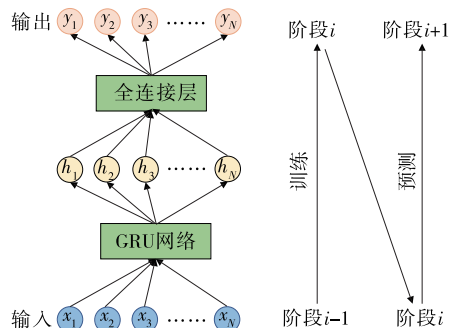


图2 基于GRU的沉降预测模型

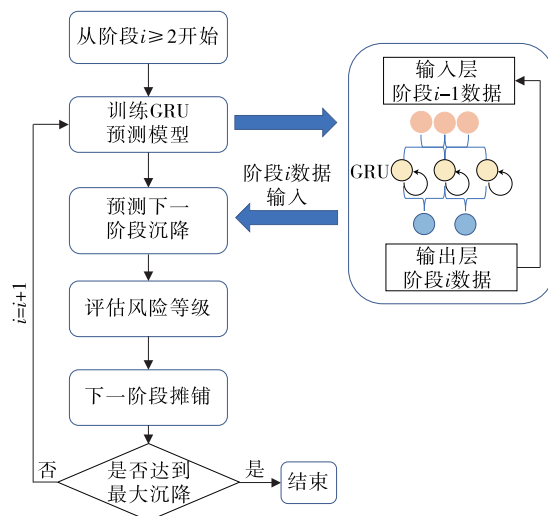


图3 路基沉降预测模型训练与更新预测流程

筑之前,通过已训练的GRU预测模型预测沉降响应,如果预测沉降过大,则采取预防措施保护相邻结构;在输入第1阶段数据进行训练后,当输出的第2阶段沉降预测结果与实测数据接近、满足精度要求时,将第2阶段监测数据作为新的训练数据集重新训练一个新的基于GRU的模型预测第3阶段数据,并比较预测值和实测值,评估模型预测精度。随着施工的进行,上述训练和预测过程不断重复,直到施工结束。

基于GRU的沉降预测模型的训练通过时间反向传播算法和批量梯度下降来实现,所有训练实例的梯度平均值用于更新神经网络的权重。采用文献[13]中的Nesterov-accelerated Adam(Nadam)优化算法进行模型优化训练。

过拟合是当机器学习模型的预测与训练数据点过于接近却远离测试数据点时发生的统计建模错误,会使模型过于解释训练数据中的个性特质。过拟合是数据驱动方法的一个常见问题。路基沉降预测中,由于GRU模型的参数远多于训练数据数量,过拟合也可能存在,采用系数 $l=0.001$ 的L2正则化来避免过拟合问题。为消除输入变量大小对模型性能的影响并降低计算成本,将所有训练数据归一化为 $[0,1]$ 。损失函数定义如下:

$$L_{\text{Loss}} = \frac{1}{N} \sum_{i=1}^N \left(\frac{U_{\text{obs}}^i - U_{\text{GRU}}^i}{U_{\text{obs}}^i} \right)^2 + \lambda \|W\| \quad (9)$$

式中: U_{obs}^i 为归一化后监测数据的第 i 个点; U_{GRU}^i 为归一化后预测数据的第 i 个点; N 为实测数据总数; λ 为权重系数; W 为权重。

L_{Loss} 的值小于 1×10^{-4} 或训练时长 e_{poch} 大于 $1 \times$

10^4 时训练停止。

多目标优化更新程序与贝叶斯更新程序类似,两者的区别如下:贝叶斯更新中的参数通过马尔可夫链蒙特卡罗算法模拟获得的后验分布和样本统计量表示,前向预测通过更新土壤参数间接实现,而基于深度学习的多目标优化更新过程可直接从监测数据得到预测数据,不需要土壤参数。

2 验证案例

采用 GRU 模型预测湖南某高速公路 K21+450—650 路段的路基沉降,验证模型的准确性和可靠性。该路段路肩宽 26 m,厚 5 m,图 4 为其路基横断面。路基采用分层填筑,共分 7 个阶段进行摊铺。根据电子式压力计、地下水压力传感器、沉降仪等监测数据,该路段的土体以软土为主,地下水位于地表以下 1 m。预测中不考虑土壤的蠕变行为。

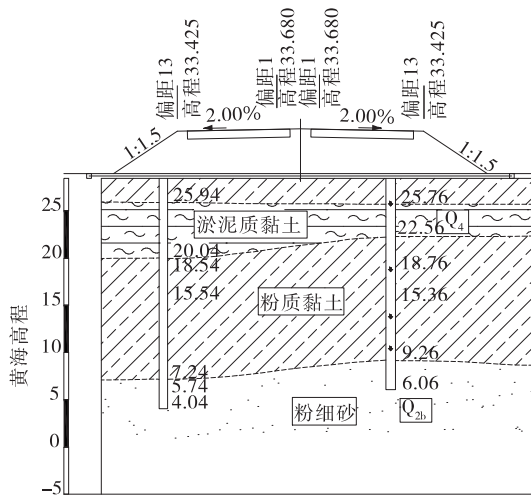


图 4 某高速公路 K21+450—650 路段路基横断面(单位:m)

2.1 路基沉降预测

为排除监测点数量对 GRU 模型预测准确性的影响,在第 1 阶段路基沉降预测的基础上进行 4 次测试,其中监测点数量分别设置为 45 个、23 个、16 个和 10 个。监测点数量对 GRU 模型预测准确性的影响见图 5。由图 5 可知:训练 GRU 模型至少需要 23 个监测点,监测点数量少于 23 个时,模型会出现过拟合的问题(训练误差小,预测误差大)。

采用 Nadam 算法进行模型优化训练。Nadam 算法的优点是收敛快速,但 Nadam 优化训练的最优解在很大程度上取决于初始试验结果,且最优解只能搜索局部最小值。为解决该问题,从不同初始位置开始搜索。为避免随机性并保证预测模型的可靠

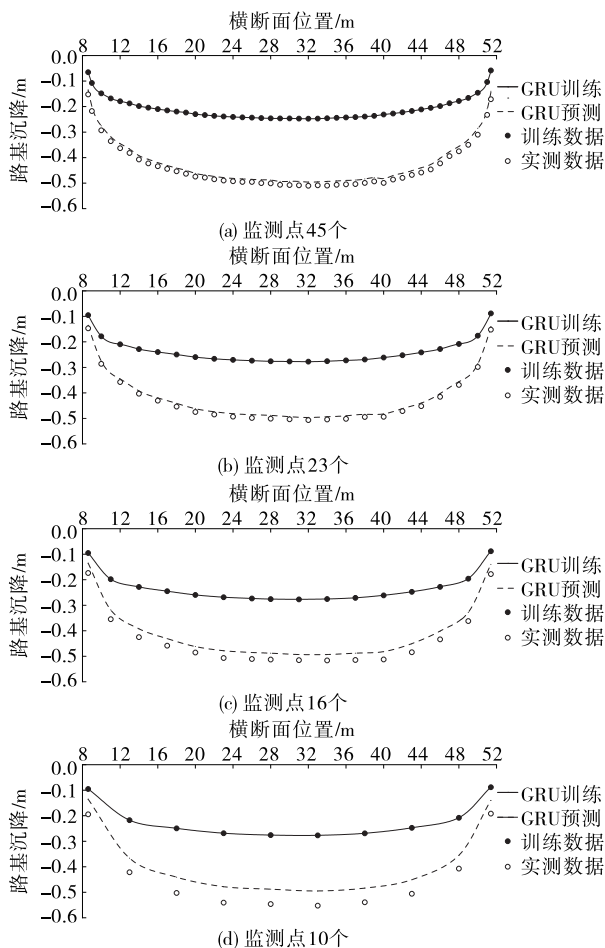


图 5 监测点数量对 GRU 模型预测准确性的影响

性,GRU 模型的每个训练和预测过程都独立开展 1 000 次试验,根据试验结果进行统计分析。图 6 为路基沉降预测结果。由图 6 可知:在路基摊铺第 3 阶段,沉降实测值和预测值之间误差较大,这是由于沉降变形模式发生了变化。为保证施工车辆通过,路基一侧留有车道,路基摊铺两侧所受荷载类型不相同,导致利用 GRU 模型从第 1、第 2 阶段的沉降值训练得到的最优解预测第 3 阶段沉降失效。此外,在路基摊铺早期,路基沉降受荷载影响较大,容易出现测量误差,GRU 预测模型的可靠性受到影响。第 3 阶段以后,预测结果与实测结果越来越吻合,表明 GRU 预测模型准确学习、分析了路基沉降的演变规律。

2.2 沉降预测模型精度评价

采用绝对误差 E_{AE} 、相对误差 E_{RE} 和相对均方根误差 E_{RRMSE} 评估 GRU 沉降预测模型的准确性和可靠性,其定义见式(10)~(12)。

$$E_{AE} = Y_{obs} - Y_{GRU} \quad (10)$$

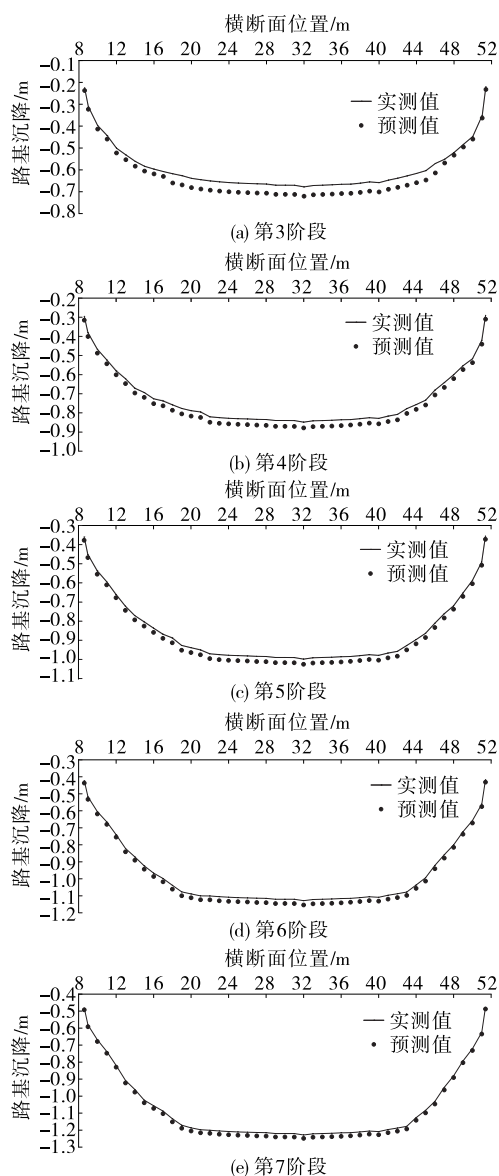


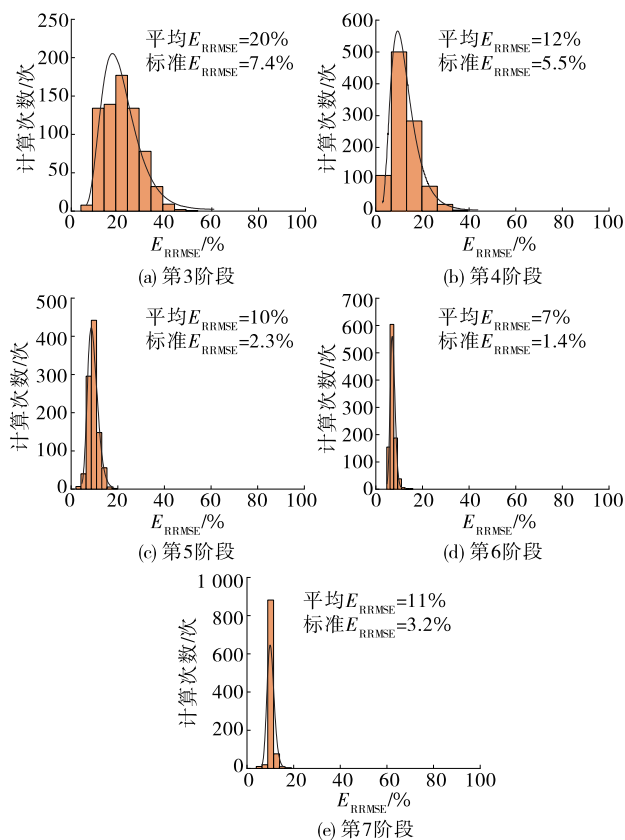
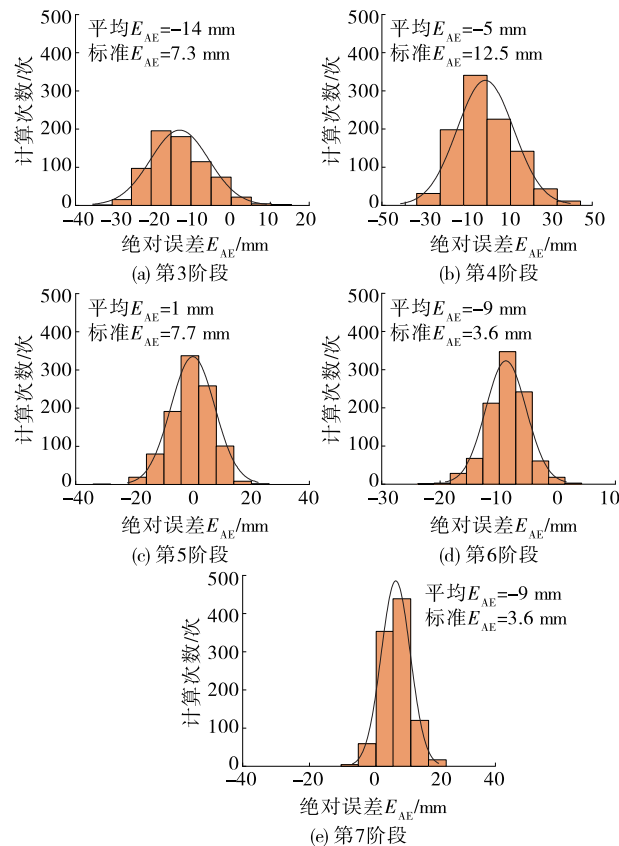
图6 K21+450—650路段路基沉降预测值与实测值对比

$$E_{RE} = \frac{Y_{obs} - Y_{GRU}}{Y_{obs}} \times 100 \quad (11)$$

$$E_{RRMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(\frac{Y_{obs}^i - Y_{GRU}^i}{Y_{obs}^i} \right)^2} \times 100 \quad (12)$$

式中: Y_{obs} 为沉降实测数据; Y_{GRU} 为 GRU 预测模型的预测数据。

不同填筑阶段路基沉降预测值的 E_{RRMSE} 概率分布见图 7, 最大路基沉降时 E_{AE} 和 E_{RE} 的概率分布分别见图 8、图 9。由图 7~9 可知: 1) 所有摊铺阶段的路基沉降 E_{RRMSE} 都接近于对数正态分布, 且除第 3 阶段外, 其他阶段的预测误差都控制在 10% 左右。2) 最大沉降点的 E_{AE} 和 E_{RE} 接近于正态分布。除第 3、第 4 阶段外, 沉降实测值和预测值之间的

图7 不同阶段模型预测结果的相对均方根误差 E_{RRMSE} 图8 不同阶段模型预测结果中最大沉降点的绝对误差 E_{AE}

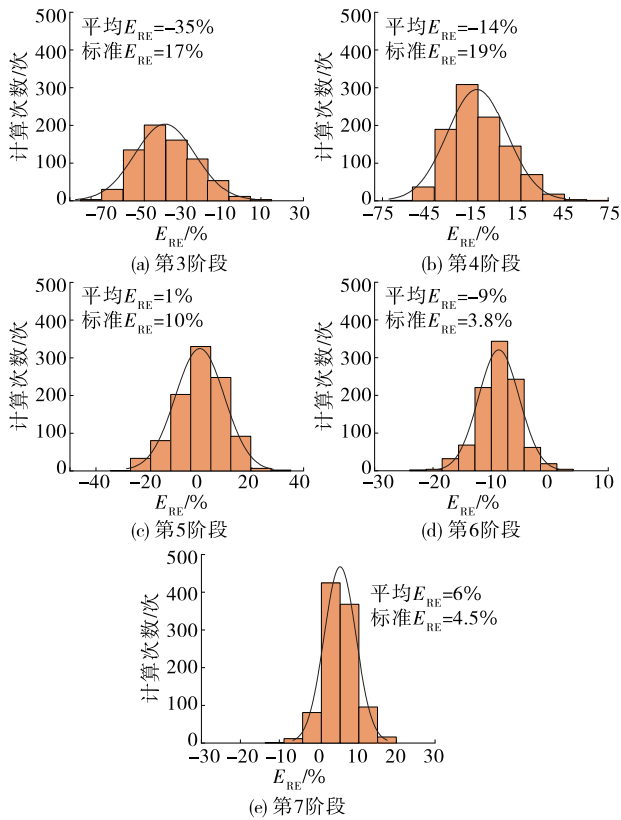


图9 不同阶段模型预测结果中最大沉降点的相对误差 E_{RE}

E_{AE} 大多在 15 mm 内;第3阶段由于变形模式变化, E_{RE} 小于 20%, 其他阶段的 E_{RE} 都随着摊铺的进行呈下降趋势。基于 GRU 的路基沉降预测模型的准确性和可靠性满足要求。

为进一步验证基于 GRU 的预测模型的准确性,采用该模型对该高速公路 K21+307 处的路基沉降进行预测。该处路基宽度为 26.2 m, 厚度为 6 m, 采用分层填筑方法,分 7 个阶段进行摊铺,路基与土体横断面见图 10。采用上述模型对该处路基沉降进行预测,同时采用 K21+450—650 段路基沉降监测仪对路基沉降进行监测,图 11 为第 3~7 阶段路基沉降预测值和实测值对比。由于前两个阶段之间的荷载类型变化较大,模型不能有效学习沉降的演化规律,导致第 3 阶段的沉降预测值和实测值之间存在较大误差,其他各阶段的沉降预测值和实测值基本吻合,基于 GRU 的路基沉降预测模型成功预测了该处路基的沉降变形。

2.3 与贝叶斯和优化更新的对比

路基沉降预测结果初步证明了基于 GRU 的路基沉降预测模型的有效性。但在施工环境不断改变的工程建设中,预测模型的学习效率同样重要。与贝叶斯更新方法相比,基于 GRU 的预测模型不仅

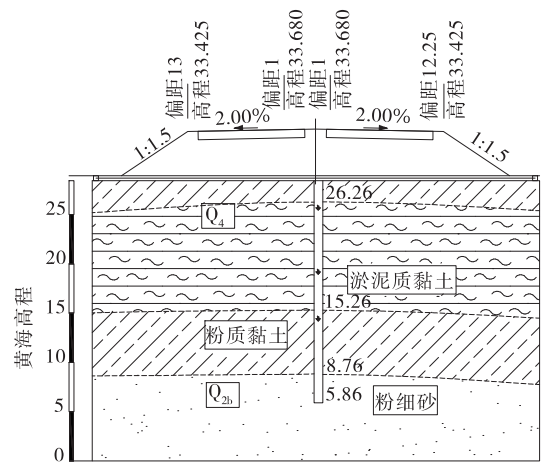


图10 某高速公路 K21+370 处路基横断面(单位:m)

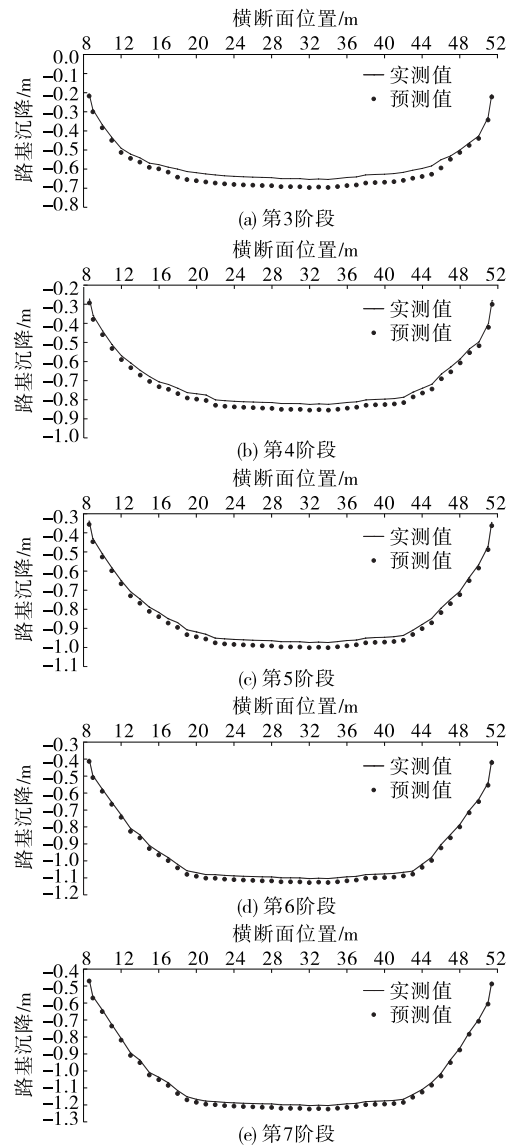


图11 某高速公路 K21+370 处路基沉降预测值与实测值对比

能准确预测地面沉降最大值,还可通过GRU神经网络的更新方法准确学习和预测路基沉降的演变规律。但分层填筑涉及的不确定性不能被准确量化。尽管采用半经验模型代替贝叶斯更新方法中有限元模型可显著减少计算量,但要得出试验相关参数的后验分布,仍需要数千次的计算。与基于优化的更新相比,GRU模型可在2 h甚至几分钟内实现一次算法更新,而优化更新的计算时间通常超过3 d,并且基于优化的更新结果的有效性取决于选择合适的土壤模型,同时需要更新的参数数量越多,所需实地监测越多。基于GRU的预测模型可避免上述问题。综上,基于GRU的预测模型在准确性和效率方面都优于贝叶斯和优化更新。

2.4 预测模型的局限性

尽管基于GRU的预测模型具有显著优势,但无论是贝叶斯或优化更新,还是基于GRU的预测模型,都无法预测沉降的突然变化。此外,每个阶段的路基铺设深度应尽可能相同,否则从前一阶段数据学习到的变形趋势可能不适合预测下一阶段的沉降反应;路基铺设过程中需制定合理、详细的施工计划,施工人员应严格遵循施工规律,避免人为因素的影响。更重要的是,应对整个路基铺设阶段的变形和地面反应进行实时监测,并将监测结果作为训练预测模型的数据来源。

此外,本文没有考虑监测数据的不确定性对预测模型的影响。在未来的研究中,可运用贝叶斯深度学习^[14]对基于GRU的沉降预测模型进行改进。

3 结论

本文提出一种基于深度学习算法与现场沉降监测数据的软土地地区公路路基沉降预测方法,该方法采用GRU神经网络生成预测模型。主要结论如下:

(1) 基于GRU的预测模型可以避免机器学习模型的过拟合问题。

(2) 基于GRU的路基沉降预测模型的预测准确性和效率都优于贝叶斯和优化更新,训练后的预测误差在10%左右,且算法更新迅速,可避免土壤模型的选择与更新算法所需数据监测量大的问题。

参考文献:

- [1] 许飞.公路软土路基加固处理及沉降分析[D].合肥:安徽建筑大学,2020.
- [2] 唐予.公路软土路基沉降规律及施工控制[J].广东公路

交通,2021,47(5):22—25+32.

- [3] ZHANG P, CHEN R P, WU H N, et al. Ground settlement induced by tunneling crossing interface of water-bearing mixed ground: A lesson from Changsha, China[J]. Tunnelling and Underground Space Technology, 2020, 96: 103224.
- [4] JIANG M, YIN Z Y. Influence of soil conditioning on ground deformation during longitudinal tunneling[J]. Comptes Rendus Mécanique, 2014, 342(3): 189—197.
- [5] 刘军, 杨志军, 韩俊青, 等. 几种软土路基沉降预测方法的对比分析[J]. 测绘工程, 2016, 25(2): 52—56.
- [6] 蒋鑫, 蒋怡, 梁雪娇, 等. 软土地基高速公路路基拓宽改建全过程变形特性数值模拟[J]. 铁道科学与工程学报, 2015, 12(5): 1039—1046.
- [7] JIN Y F, YIN Z Y, ZHOU W H, et al. Identifying parameters of advanced soil models using an enhanced transitional Markov chain Monte Carlo method[J]. Acta Geotechnica, 2019, 14(2): 1925—1947.
- [8] KUNG G T, JUANG C H, HSIAO E C, et al. Simplified model for wall deflection and ground-surface settlement caused by braced excavation in clays[J]. Journal of Geotechnical and Geoenvironmental Engineering, American Society of Civil Engineers, 2007, 133(6): 731—747.
- [9] AL-ANI M, FATTAH M Y, AL-LAMY M T A. Artificial neural networks analysis of treatment process of gypseous soils [J]. Engineering and Technology Journal, 2009, 27(9): 1811—1832.
- [10] 张军辉, 周剑坤, 周平. 软土地区高速公路路基沉降监测及预测分析[J]. 公路与汽运, 2020(3): 64—68.
- [11] 张潭, 刘展瑞, 蔡宇鹏, 等. 片石处治干线公路软土路基沉降规律及预测分析[J]. 公路与汽运, 2021(3): 90—93+97.
- [12] CHUNG J Y, GULCEHRE C, CHO K, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling [EB/OL]. <https://doi.org/10.48550/arXiv.1412.3555>, 11 Dec 2014.
- [13] DOZAT T. Incorporating nesterov momentum into adam[EB/OL]. OpenReview.net, 19 Feb 2016.
- [14] ZHANG P, JIN Y F, YIN Z Y. Machine learning-based uncertainty modelling of mechanical properties of soft clays relating to time-dependent behavior and its application[J]. International Journal for Numerical and Analytical Methods in Geomechanics, 2021, 45(11): 1588—1602.

收稿日期:2022-03-15