doi:10.11835/j.issn.1000-582X.2025.02.004

融合EEMD与BiLSTM深度网络的结构监测缺失数据重构

何盈盈[1a,1b],黄正洪[1a],张利凯[2],赵智航[2],关腾达[2]
(1. 重庆人文科技学院 a. 计算机工程学院; b. 大数据与网络信息安全工程技术研究中心,重庆 401524;
2. 重庆大学 土木工程学院,重庆 400045)

摘要:在长期监测过程中,由于传感器设备故障、供能中断、网络传输问题等诸多因素,导致结构健康监测系统采集的数据存在不完整性。针对这一问题,结合集合经验模态分解(ensemble empirical mode decomposition, EEMD)与双向长短期记忆网络(bi-directional long short-term memory,BiLSTM)在时序处理方面的优势,提出一种基于EEMD-BiLSTM的结构监测缺失数据重构方法。该方法利用EEMD自适应分解监测时序数据为1组代表不同时间尺度的本征模态分量(intrinsic mode function,IMF),使非线性、非平稳序的时序信号平稳化。然后,将IMF分量输入到BiLSTM网络中进行缺失数据重构,提高BiLSTM预测精度。针对6层框架结构缩尺模型和Benchmark有限元仿真模型进行分析,试验结果表明,相比EEMD-LSTM、BiLSTM、LSTM主流方法,提出的EEMD-BiLSTM具有最高预测精度,在5%、10%、15%缺失数据情况下,其R^2指标保持在0.8以上。因此,采用EEMD方法对非稳态结构加速响应数据进行预处理,可有效提高BiLSTM的预测精度,对于结构监测缺失数据问题,提供更具适应性的方法。

关键词:结构健康监测;数据重构;集合经验模态分解;双向长短期记忆网络
中图分类号:TPU446.2 **文献标志码:**A **文章编号:**1000-582X(2025)02-035-15

Reconstructing missing health monitoring data using a deep network integrating EEMD and BiLSTM

HE Yingying[1a,1b], *HUANG Zhenghong*[1a], *ZHANG Likai*[2], *ZHAO Zhihang*[2], *GUAN Tengda*[2]
(1a. School of Computer Engineering; 1b. Big Data and Network Information Security Engineering Technology Research Center, Chongqing College of Huamnitics Science & Technology, Chongqing 401524, P. R. China;
2. College of Civil Engineering, Chongqing University, Chongqing 400045, P. R. China)

收稿日期:2023-08-12
基金项目:重庆市教委科学技术研究项目(KJQN202201805,KJQN202301801);重庆市合川区科技计划项目(HCKJ-2024-110);重庆人文科技学院科学技术研究项目(CRKZK2023007,JSJGC202201,JSJGC202205)。
Supported by the Science and Technology Research Program of Chongqing Municipal Education Commission (KJQN202201805, KJQN202301801), Science and Technology Program of Hechuan District of Chongqing(HCKJ-2024-110), and Science and Technology Research Program of Chongqing College of Humanities Science & Technology (CRKZK2023007, JSJGC202201, JSJGC202205).
作者简介:何盈盈(1994—),女,硕士,主要从事大数据分析与挖掘、结构健康监测、人工智能方向研究,(E-mail) heyingyg@163.com。
通信作者:张利凯(1993—),男,(E-mail)zhanglikai@cqu.edu.cn。

Abstract: In long-term monitoring processes, the structural health monitoring(SHM) system often encounters data incompleteness due to various factors, including sensor malfunctions, power interruptions, and network transmission issues. To address this challenge, this study proposes a missing data reconstruction method for structural monitoring based on ensemble empirical mode decomposition(EEMD) and bidirectional long short-term memory(BiLSTM) networks, leveraging their advantages in time-series processing. The proposed approach utilizes EEMD to adaptively decompose the monitoring time-series data into a set of intrinsic mode functions (IMFs), each representing different time scales. This decomposition effectively transforms the nonlinear and non-stationary time-series signals into stationary components. The IMFs are then input into a BiLSTM network for missing data reconstruction, enhancing the prediction accuracy of the BiLSTM model. Analysis is conducted on a six-story scaled structural model and a benchmark finite element simulation model. Experimental results demonstrate that, compared to the mainstream methods such as EEMD-LSTM, BiLSTM, and LSTM, the proposed EEMD-BiLSTM approach achieves the highest prediction accuracy. In cases of 5%, 10% and 15% missing data, the R^2 value remains above 0.8. Therefore, the use of the EEMD method for preprocessing non-stationary structural acceleration response data significantly improves the prediction accuracy of BiLSTM, providing a more adaptive solution to the problem of missing data in structural monitoring.

Keywords: structural health monitoring; data reconstruction; ensemble empirical mode decomposition; bi-directional long short-term memory network

近年来,结构工程领域取得了令人瞩目的成就,其中包括:港珠澳大桥、珠海中心大厦、杭州湾跨海大桥等。然而,随着这些建筑物的规模和复杂性增加,安全性和可靠性面临日益显著的问题。结构健康监测(structural health monitoring,SHM)是一种状态监测、特征识别和状态评估的自动化系统,为结构的管理和养护提供决策支撑,对结构进行实时、准确监测,保证结构的安全性和可靠性。

中国在许多大跨桥梁上安装了结构健康监测系统[1-3],这些系统不仅能够实时监测结构的运行状态,还能为基础科学问题提供大量数据支持。然而,由于恶劣的环境条件、仪器故障、断电等因素,采集的数据可能存在各种异常和缺失,这将影响数据的有效性和准确性,进而影响结构状态识别的可靠性。因此,如何利用有效的方法重构结构监测中的缺失数据,是当前结构健康监测领域待解决的关键问题。

近年来,许多研究人员对SHM系统不完整的数据集进行建模并重构。一类方法试图从可用数据中重建完整信号。Wan等[4]提出一种基于Bayesian多任务学习的结构健康监测数据恢复方法,该方法通过多任务高斯过程先验联合建模提取任务间信息,实现更准确的数据恢复。结果表明,与单任务学习相比,该方法可以更好地恢复大块缺失监测数据。Li等[5]提出一种基于多任务高斯过程回归的缺失数据重建方法,该方法通过建立多个输出的高斯过程模型,充分利用不同传感器信号之间的内在相关性,实现对缺失传感器数据的估计和重建。但上述提及的数据重构方法往往针对特定的场景具有较好效果,在SHM不具有普遍适用的规律。

另一类方法试图采用估计值对数据进行修复,如简单的插值方法(利用空间、时间相关性对缺失数据进行补全)、统计方法(差分自回归移动平均模型)、机器学习等方法。Chen等[6]提出一种非参数的联合分布方法,该方法可以考虑不同传感器应变之间的相关性,并对缺失数据进行有效处理。Zhang等[7]提出一种基于监测点间相关性分析的缺失数据恢复方法,该方法通过构建监测点间的相关性矩阵,采用奇异值分解技术分析矩阵的主分量,根据主分量对缺失监测点进行重构,进而实现对缺失监测数据的有效重构。由于上述方法,大多是根据SHM系统采集的传感器数据之间相关性对数据进行修复,在实际应用过程中,具有一定局限性。考虑到SHM采集数据具有非线性、非平稳性特征,传统方法处理非线性数据能力较弱,且机器学习方法需要手动构建特征,模型算法精度还有待提升。

随着人工智能的崛起,采用深度学习方法进行数据重构受到来自学术界的广泛关注。由于SHM系统采集的传感器数据属于时间序列数据,现有的深度学习方法如卷积神经网络(convolutional neural networks,

CNN)[8]、长短期记忆(long short-term memory, LSTM)[9]、双向长短期记忆(bidirectional LSTM, BiLSTM)[10]、门控循环单元(gated recurrent unit, GRU)[11]、双向门控循环单元(bidirectional GRU, BiGRU)[12]对时序数据都有很好的处理能力，可以有效对缺失数据进行重构。Fan等[13]提出一种基于CNN的结构监测缺失数据恢复方法，该方法采用CNN捕捉缺失数据和真实数据之间的非线性关系，通过压缩层提取不完整的数据特征，并在重构层中逐步扩展这些特征，恢复丢失的数据。Liu等[14]提出了基于LSTM网络的结构监测温度缺失数据恢复方法，该方法利用LSTM网络学习温度时间序列的长短期依赖性，预测填补缺失值，并在真实桥梁温度数据上测试，结果表明LSTM方法可以更准确地重构缺失的温度数据。Oh等[15]提出一种基于CNN的结构响应数据恢复方法，该方法使用传感器发生故障前稳定测量的应变监测数据训练卷积神经网络模型，通过数值、实验研究及实际桥梁的应用案例验证了该方法的有效性。Du等[16]提出基于多模态融合自动编码器的异构结构响应恢复方法，该方法利用多模态输入训练深度神经网络模型进行结构响应预测，克服单一模态的限制，该方法相比单模态方法提高了平均预测精度，为异构结构响应恢复提供了有效解决方案。

尽管深度学习方法具有很好的时序特征抽取与建模能力，并在其他领域具有广泛应用，但现有方法仍存在一定改进空间。从现有的文献发现，基于混合数据重构方法，可以结合深度学习方法的优点，形成适应性更强的网络模型，有效解决缺失数据重构问题。Lei等[17]提出一种基于深度卷积对抗生成网络的结构健康监测缺失数据重建方法，该方法使用编码器-解码器结构的生成器网络提取信号特征并重建缺失数据，使用判别器网络区分真实数据和重建数据，准确重建失效信号中的低频和高频特征。该方法在数字和实际案例中表现良好。Li等[18]提出一种混合的经验模态分解和长短期记忆网络方法，用于结构健康监测系统中丢失测量信号数据的修复，该方法利用EMD处理非稳定信号的能力及LSTM记忆长期时间相关性的优势，性能优于单一的ARIMA、ANN或SVR方法。

因此，针对结构监测数据缺失问题，笔者提出一种混合集合经验模态分解(ensemble empirical mode decomposition, EEMD)与BiLSTM深度学习模型，进行缺失数据重构。相比已有的传统方法和单一深度学习模型，提出的混合EEMD-BiLSTM深度学习方法具有显著优势，其中EEMD算法可以显著提高信号的稀疏性，增加重构可行性，同时BiLSTM网络能够学习时序前后依赖特征，实现对缺失数据的准确重构。本研究的创新与贡献如下：

1)在结构健康监测领域，引入混合EEMD与BiLSTM深度学习模型，实现对监测缺失数据的自动重构；

2)设计一种有效的网络结构，使BiLSTM网络输入各IMF分量，采用误差反向传播算法训练模型参数；

3)通过实验分析不同方法（如EEMD-BiLSTM，EEMD-LSTM，BiLSTM，LSTM）的重构效果，提出EEMD-BiLSTM方法具有最佳精度。

1　结构监测缺失数据重构方法

1.1　EEMD-BiLSTM数据重构方法

研究提出一种基于EEMD-BiLSTM的监测缺失数据重构方法，如图1所示。首先，采用传感器获取结构关键位置历史监测数据，并对其进行数据预处理（数据清洗、滑动窗口）。考虑到EEMD可以有效处理非平稳、非线性监测数据，能够将其按照自身特点分解时间序列为若干本征模态函数(intrinsic mode functions, IMF)和残差Res。然后，将EEMD分解得到的IMF分量进行特征融合，划分为6:2:2，依次为训练集、验证集、测试集。最后，调整BiLSTM的输入神经元的个数，将IMF分量输入到BiLSTM深度网络，其中BiLSTM采用2个独立的隐藏层对特征序列数据的正向和反向进行特征提取，得到之前和之后的信息，作为当前时刻时间序列数据的基础，输入到全连接层，预测缺失数据，具体实现流程如下：

某一传感器的实时监测结构状态数据作为原始时间序列$s(t)$，通过EEMD方法分解得到n个IMF分量和1个残差序列$r(t)$。选取IMF分量Seq作为特征输入，则可表示为

$$\mathrm{Seq}=\{\mathrm{IMF}_1(t),\mathrm{IMF}_2(t),\cdots,\mathrm{IMF}_n(t)\},\tag{1}$$

将数据划分为训练数据、验证数据、测试数据，利用训练数据训练BiLSTM模型参数，通过验证数据进行BiLSTM模型参数的调整，最后将测试数据输入BiLSTM预测模型中进行缺失数据预测。

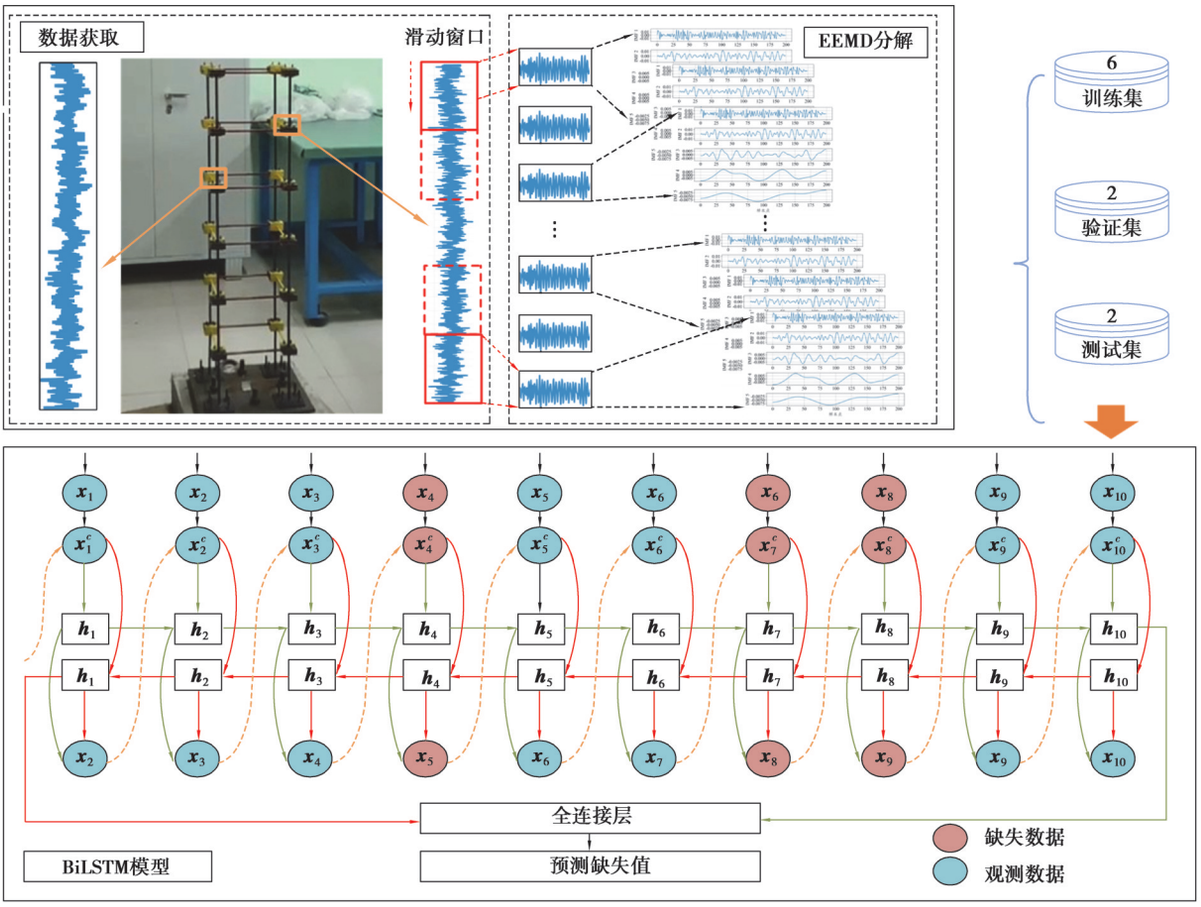


图 1 基于EEMD-BiLSTM的缺失数据重构流程图

Fig. 1 Flowchart of missing data reconstruction based on EEMD-BiLSTM

1.2 集合经验模态分解

集合经验模态分解(ensemble empirical mode decomposition,EEMD)是Flandrin等[19]在经验模态分解(empirical mode decomposition,EMD)[20]基础上提出的改进信号处理方法。EEMD是向原始信号中添加白噪声进行处理,利用白噪声的频谱均匀分布特性,有效消除EMD中由于信号极值点间隔不均匀导致的模式混合问题。同时也减小对信号包络线拟合误差的敏感性,再对每个新序列进行EMD分解,得到一组IMF分量,最后对所有分解结果取平均,得到EEMD的最终结果。因此,EEMD可以更准确地分解出信号的各种本征模态函数IMF,更适用于处理结构振动响应等复杂非线性时序信号。EEMD方法的过程如下

1)向原始序列中添加白噪声序列,得到新序列$x_j(t)$

$$x_j(t)=s(t)+o_j(t), \tag{2}$$

其中,$o_j(t)$表示第j次添加白噪声序列;

2)将$x_j(t)$通过EMD方法进行分解,得到1个残差序列$m_j(t)$以及i个IMF分量$g_{ij}(t)$;

3)重复1)2)步骤N次,并每次向原始序列中加入新的白噪声序列,得到N组IMF分量$g_{ij}(t)$和残差序列$m_j(t)$;

4)分别计算经分解得到N组$g_{ij}(t)$和$m_j(t)$平均值

$$g_i(t)=\frac{1}{N}\sum_{j=1}^{N}g_{ij}(t), \tag{3}$$

$$m(t)=\frac{1}{N}\sum_{j=1}^{N}m_j(t), \tag{4}$$

经上述将N组IMF分量和残差序列求平均值操作,消除原始序列加入均值为零的白噪声。通过EEMD

分解原始序列得到的结果与其序列本身相同。

5)最终得到 $s(t)$ 的 EEMD 分解结果为

$$s(t)=\sum_{i=1}^{n} g_i(t)+m(t), \tag{5}$$

式中：n 表示 IMF 分量的个数。

1.3　双向长短期记忆网络

双向长短期记忆(Bi-directional long short-term memory，BiLSTM)网络是对长短期记忆(long short-term memory，LSTM)网络[21]的一种改进，其克服了 LSTM 结构易出现的忽略未来信息的问题。BiLSTM 神经网络结构如图 2 所示，由 1 个输入-输出与 2 个 LSTM 层构成，其中 2 个 LSTM 层包括前向 LSTM 层和反向 LSTM 层，输出层既能获取输入的原始时间序列的历史信息，也能获取输入时序的未来信息。因此，BiLSTM 网络在利用原始时序的历史信息预测未来的基础上，也考虑未来信息对当前状态的影响，该网络的更新公式如下

$$\vec{\boldsymbol{h}}_t=H\left(\boldsymbol{W}_1\boldsymbol{x}_t+\boldsymbol{W}_2\vec{\boldsymbol{h}}_{t-1}+\vec{\boldsymbol{b}}\right), \tag{6}$$

$$\overleftarrow{\boldsymbol{h}}_t=H\left(\boldsymbol{W}_3\boldsymbol{x}_t+\boldsymbol{W}_5\overleftarrow{\boldsymbol{h}}_{t-1}+\overleftarrow{\boldsymbol{b}}\right), \tag{7}$$

$$\boldsymbol{y}_t=\boldsymbol{W}_4\vec{\boldsymbol{h}}_t+\boldsymbol{W}_6\overleftarrow{\boldsymbol{h}}_t+\boldsymbol{b}_y, \tag{8}$$

式中：$\boldsymbol{W}_1$，$\boldsymbol{W}_2$，$\boldsymbol{W}_3$，$\boldsymbol{W}_4$，$\boldsymbol{W}_5$，$\boldsymbol{W}_6$ 表示权重系数；$\vec{\boldsymbol{b}}$，$\overleftarrow{\boldsymbol{b}}$，$\boldsymbol{b}_y$ 表示偏置向量，H 表示 LSTM 层的激活函数，$\vec{\boldsymbol{h}}_t$，$\overleftarrow{\boldsymbol{h}}_t$，$\boldsymbol{y}_t$ 分别表示前向传播向量、反向传播向量及输出向量。

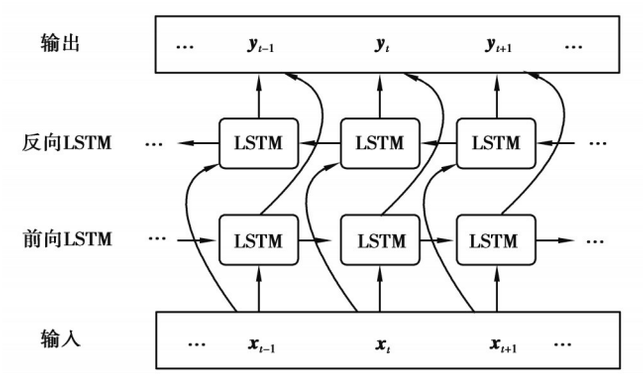


图 2　BiLSTM 网络结构示意图

Fig. 2　Schematic diagram of BiLSTM network structure

2　试验验证与评估

为验证提出的 EEMD-BiLSTM 结构监测缺失数据重构方法的有效性，在 6 层框架结构缩尺模型数据集上进行测试，并与当前主流的深度学习算法，如 EEMD-LSTM、BiLSTM 和 LSTM 进行对比。

2.1　数据获取介绍

如图 3 所示，在实验室环境下对 6 层框架结构缩尺模型开展分析[22]，缩尺模型的几何比例为 1:20。该结构采用钢材料制成，结构总高度为 1 125 mm，层间距为 187.5 mm，结构平面尺寸为 250 mm×250 mm，结构底部通过地脚螺栓固定在混凝土基块上。楼层梁截面由 6 mm×6 mm 矩形截面构成，楼层柱的截面由 7.8 mm×7.8 mm 矩形截面构成，底层柱的截面尺寸比楼层柱宽，使结构具有更好的稳定性，其宽度为 15 mm。每个楼层的楼板由多个 250 mm×250 mm 的小钢板拼接组成。

为了确保结构的抗侧力性能及刚度，在预先制作好的混凝土基础上安装底层的基础钢板，用螺栓进行固定连接。然后，将底层钢柱与基础钢板进行螺栓连接，吊装第 1 层的楼板连接到底层钢柱上，并用高强螺栓将柱和梁连接起来。依次类推，完成 6 层框架结构缩尺模型的整体吊装。

为获取 6 层框架结构缩尺模型动力响应，在缩尺模型每个楼层的中心位置布置了 6 个加速度传感器，所

有加速度计都安装在同1个方向，与结构的主轴平行，形成1个L形截面，捕捉振动方向的数据。首先，整个模型被拉出一定的初始位移，然后释放并允许其自由振动，同时用Pulse Labshop和Pulse reflex软件记录每个楼层的加速度时程数据，其中，采样频率为4 167 Hz。

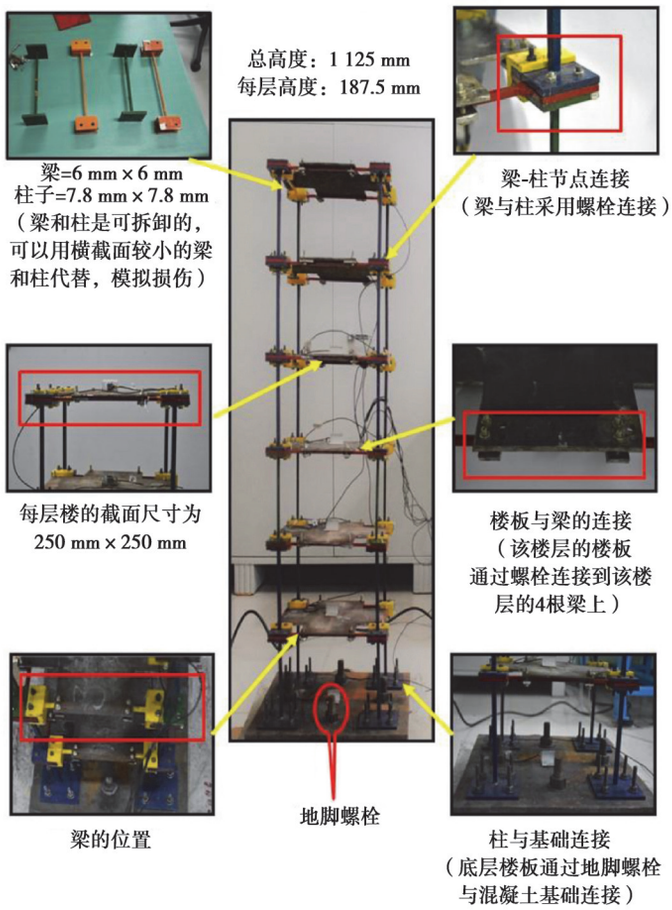


图 3　6层框架结构缩尺模型示意图

Fig. 3　Schematic diagram of the six-storey shear building

2.2　实验流程设置

图4展示了基于EEMD方法对加速度时程曲线进行的多尺度特征分析结果。其中，红色线代表原始的加速度时程曲线，蓝色线代表EEMD分解所得到的多个IMF分量。从图中可以直观看出，EEMD方法将原始非平稳、包含随机性的非线性加速度时程数据，分解为多个振荡模式独立、幅值稳定的IMF分量，表明EEMD能够有效提取信号中的不同尺度特征。通过这种分解，原始复杂的加速度响应信号被拆分为多个不同频率范围的稳定信号，每个信号都代表了结构在特定频带下的本征振动模式，这种分解明确揭示了信号所包含的不同频率成分，有利于进一步分析信号的频率特性。这种分解产生的稳定信号时序也更利于后续的BiLSTM算法进行时序特征的提取及缺失数据预测。

为获取更多训练样本，采用固定大小的滑动窗口对原始加速度数据进行划分，以采样频率为依据，将窗口大小设定为采样频率的1/20。经滑窗处理后获得约9 500个样本，按6∶2∶2的比例将样本数据集划分为训练集、验证集和测试集。将训练集输入EEMD模型进行信号分解，得到多个IMF分量和残差Res。将分解后的IMF各分量作为输入特征，输入BiLSTM模型进行训练。在模型训练过程中，BiLSTM能够综合考虑当前时间步的输入特征序列以及过去和未来的上下文信息，完成序列数据的建模。为提升模型泛化能力，加入Dropout层随机屏蔽部分神经元，模型训练完成后，利用测试集数据进行模型预测，并与真实结果比较，以评价模型的预测性能。

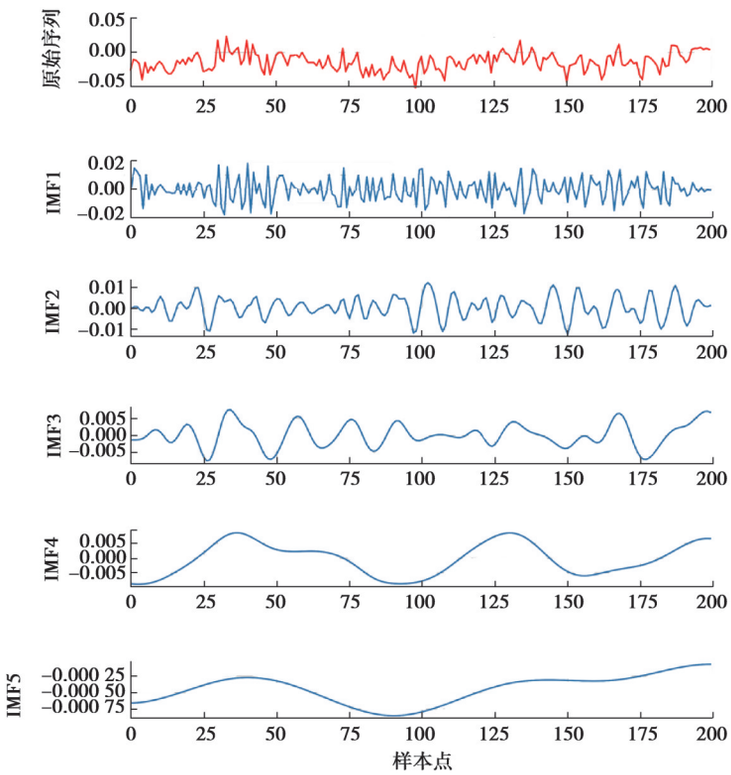


图 4　基于 EEMD 的加速度时程曲线多尺度特征分析图

Fig. 4　Multi-scale feature analysis of acceleration time-history curve based on EEMD

2.3　评价指标

为进一步评估缺失值预测效果，共选取 4 个常用的统计评估标准，包括拟合优度（R-squared）[23]，均方误差（mean squared error，MSE）[24]，均方根误差（root mean square error，RMSE）[25]，均值绝对误差（mean absolute error，MAE）[26]。其中 R^2 衡量了回归模型预测的平方和与总平方和之间的比例，越接近 1 表示模型拟合越好。MSE、RMSE 和 MAE 用于表示真实值与预测值之间的误差程度，对于不同的预测模型，MSE、RMSE 和 MAE 的值越低，预测性能越好。这些评估标准公式如下

$$R^2 = 1 - \sum_{i=1}^{n}(\hat{y}_i - y_i)^2 / \sum_{i=1}^{n}(y_i - \bar{y}_i)^2, \tag{9}$$

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2, \tag{10}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}, \tag{11}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|\hat{y}_i - y_i\right|。 \tag{12}$$

3　实验结果及对比分析

3.1　6 层框架结构缩尺模型

实验开发环境主要是采用 Keras 框架，参数设置的合理性直接影响模型性能，融合 EEMD 与 BiLSTM 深度网络模型具体参数设置如表 1 所示，神经网络模型的参数调优目前还没有成熟系统的方法，主要依据人工经验、网格搜索、遗传算法等确定，例如训练集、验证集和测试集的样本比例分配依据人工经验，其他参数如学习率、迭代次数、批尺寸主要通过网格搜索来确定具体取值。

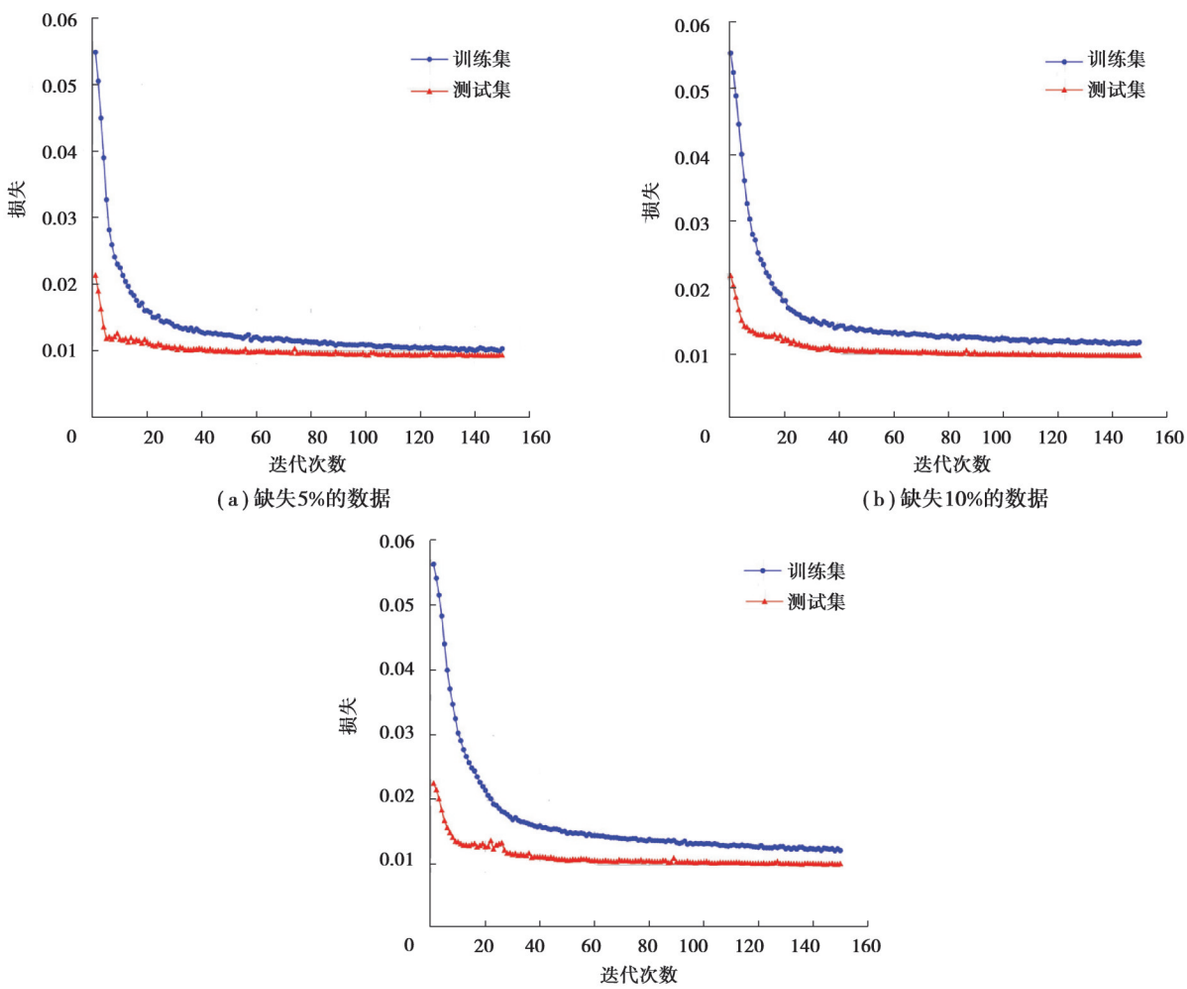
设定 3 种时间序列缺失状态，分别是原始加速度时间序列数据随机缺失 5%、10%、15%。针对 3 种数据缺失状态，在训练集和测试集上，对 EEMD-BiLSTM 方法进行实验分析，并绘制损失函数（Loss）随迭代次数变化情况，结果如图 5 所示，从图中可以观察到，经过多次训练迭代后，EEMD-BiLSTM 方法的预测值与真实值

之间的误差逐渐减小，表明EEMD-BiLSTM方法在预测任务中具有优越性。另外，在经过大约100次的迭代训练后，训练集和验证集上的损失函数值均呈现稳定趋势，验证了EEMD-BiLSTM方法在捕捉时间序列数据中的复杂依赖关系方面的有效性，实现了缺失数据的精确建模。

表 1　EEMD-BiLSTM模型参数设置

Table 1　Parameter settings of the EEMD-BiLSTM model

实验参数	值
优化器	Adam
学习率	0.001
损失函数	MAE
批尺寸	64
迭代次数	150
训练集占比/%	60
验证集占比/%	20
测试集占比/%	20



(a) 缺失5%的数据

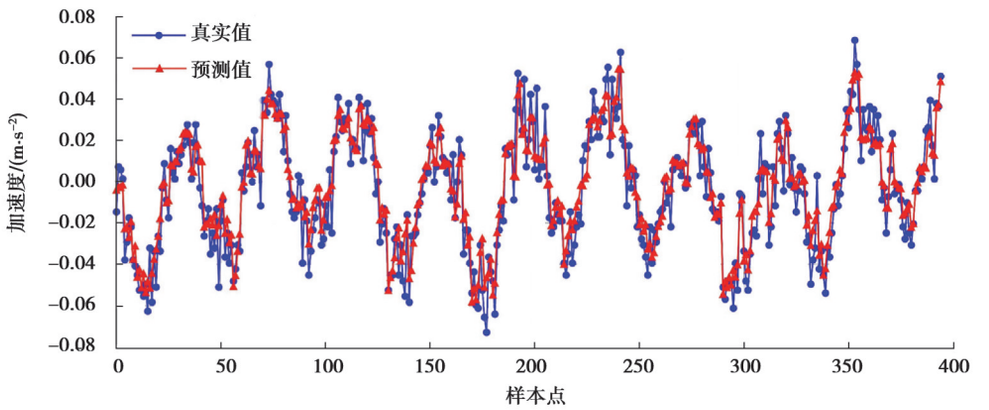
(b) 缺失10%的数据

(c) 缺失20%的数据

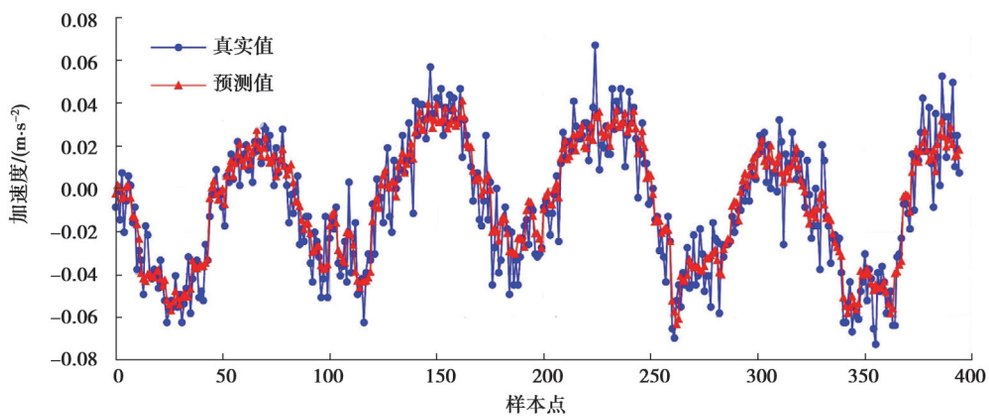
图 5　不同缺失比例数据的训练过程

Fig. 5　Training process of different missing data ratio

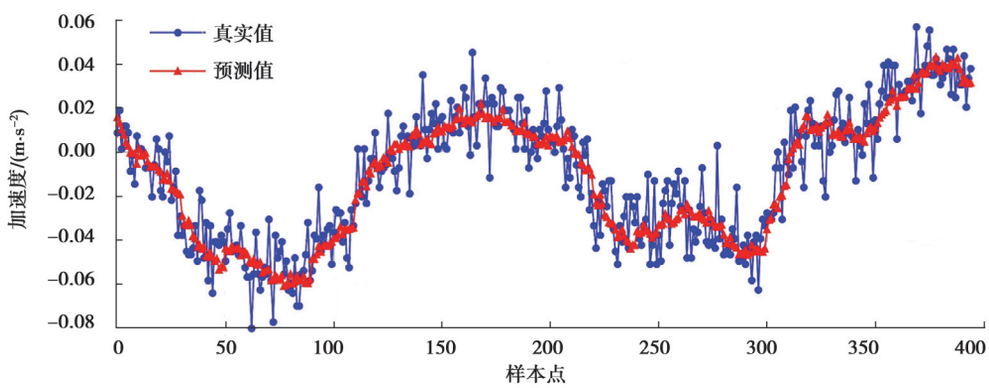
图 6 直观显示了所提出的 EEMD-BiLSTM 方法在不同缺失比率下的真实值与预测值之间的差异，其中蓝色曲线代表真实值，红色曲线代表预测值。图 6(a)展示了在数据缺失比率为 5% 的情况下，预测值能够较为准确地捕捉真实值的变化趋势和具体数值，2 者之间基本重合，误差较小。图 6(b)展示了在数据缺失比率提高到 10% 的情况下，预测值与真实值之间的误差有一定增加，但预测值整体仍能够较好反映真实值的变化趋势，重构效果较好。图 6(c)展示了在数据缺失比率达到 15% 时，预测值与真实值之间的误差进一步增加，但是预测曲线的走势还是能够基本捕捉真实值的主要变化，重建结果还算可以接受。随着数据缺失比率提高，EEMD-BiLSTM 方法的预测误差也有所增加，但在缺失比率较高情况下，预测结果仍能基本反映真实值的变化趋势，预测曲线结果表明融合 EEMD 与 BiLSTM 深度网络进行结构监测缺失数据重构仍是有效方法。



(a) 在5%数据缺失比例下预测结果



(b) 在10%数据缺失比例下预测结果



(c) 在15%数据缺失比例下预测结果

图 6　不同缺失比率下真实值与预测值对比

Fig. 6　Comparison of true and predicted values for different missing ratios

表2展示了不同方法在各评价指标下的结果。可以看出，随着数据缺失比率增加，各方法的预测性能呈现下降趋势，但EEMD-BiLSTM的预测效果始终优于其他对比方法。

表2 不同预测方法的评价指标

Table 2 Evaluation indicators for different prediction methods

方法	指标	缺失比率/%		
		5	10	15
EEMD-BiLSTM	R^2	**0.854 78**	**0.830 91**	**0.812 52**
	MSE	**0.000 12**	**0.000 14**	**0.000 15**
	RMSE	**0.010 83**	**0.011 68**	**0.012 30**
	MAE	**0.008 58**	**0.009 28**	**0.009 78**
EEMD-LSTM	R^2	0.811 05	0.809 74	0.796 10
	MSE	0.000 15	0.000 15	0.000 16
	RMSE	0.012 35	0.012 39	0.012 83
	MAE	0.009 83	0.009 85	0.010 21
BiLSTM	R^2	0.840 07	0.821 56	0.804 57
	MSE	0.000 13	0.000 14	0.000 16
	RMSE	0.011 33	0.012 00	0.012 56
	MAE	0.009 01	0.009 50	0.009 95
LSTM	R^2	0.835 15	0.816 56	0.787 07
	MSE	0.000 13	0.000 15	0.000 17
	RMSE	0.011 54	0.012 17	0.013 11
	MAE	0.009 16	0.009 65	0.010 45

注：加粗数据表示最好的结果。

具体来看，图7(a)的R^2指标比较可直观得到，在不同缺失比率下EEMD-BiLSTM的预测效果最佳，其次是BiLSTM，LSTM次之，EEMD-LSTM最差。在5%、10%、15%缺失比率下，EEMD-BiLSTM的R^2值比BiLSTM提高了1.47%、0.94%、0.80%。从图7(b)的MSE指标可直观得到EEMD-BiLSTM方法的预测效果最好，其次是BiLSTM，LSTM次之，EEMD-LSTM效果最差。在5%、10%、15%缺失比率下，EEMD-BiLSTM相比BiLSTM在MSE值上的降低幅度并不显著。从图7(c)的MAE指标比较可直观看出，EEMD-BiLSTM的预测绝对误差最低，BiLSTM次之，LSTM和EEMD-LSTM相对较高。在5%、10%、15%缺失比率下，EEMD-BiLSTM的MAE值比BiLSTM降低了0.04%、0.02%、0.02%。说明EEMD-BiLSTM模型对缺失数据具有较强的鲁棒性。最后，图7(d)的RMSE结果也验证了EEMD-BiLSTM的预测均方根误差最小，优于其他方法。在5%、10%、15%缺失比率下，EEMD-BiLSTM的RMSE值比BiLSTM降低了0.05%、0.03%、0.03%。这表明EEMD-BiLSTM方法对缺失数据处理具有较强鲁棒性。

综上所述，随着缺失比率持续增加，各方法的预测性能下降还是不可避免的，相比主流的方法EEMD-LSTM、BiLSTM、LSTM，提出的EEMD-BiLSTM结合EEMD分解和BiLSTM的双向学习的优势，使其成为处理时间序列缺失数据的最佳方法。

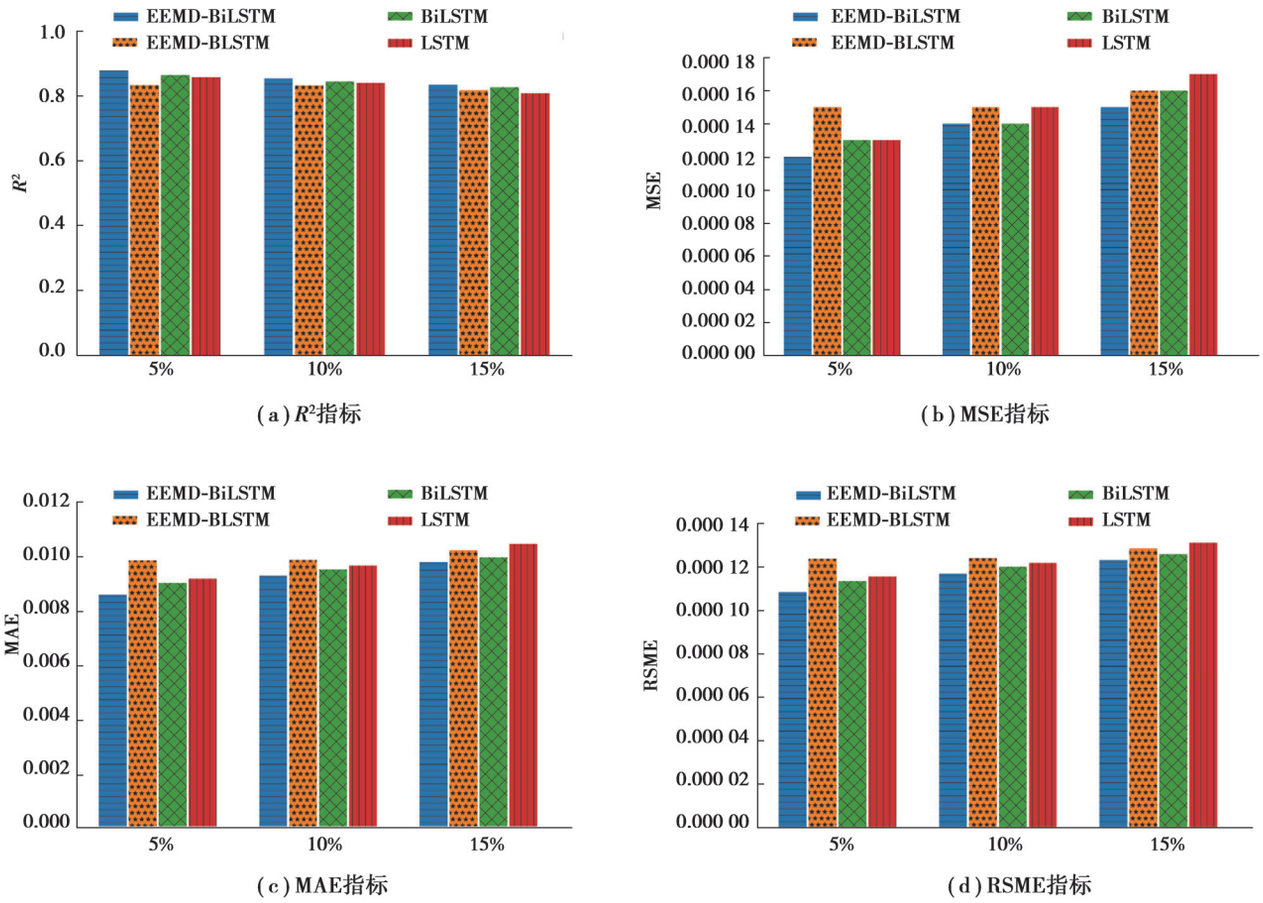


图 7　不同缺失数据比例下的预测误差

Fig. 7　Prediction error of different algorithms under missing ratio

3.2　Benchmark 有限元仿真模型

为进一步验证所提 EEMD-BiLSTM 算法的有效性，选取了 Benchmark 有限元仿真模型[27]进行分析，如图 8 所示。该模型为 4 层 2×2 的钢框架结构，每层的长度为 2.5 m，宽度为 2.5 m，高为 0.9 m，一共 4 层共 3.6 m，每层各有 8 根斜撑。结构上每层安装 4 个加速度传感器，共安装了 16 个，其中 8 个 x 向，8 个 y 向。然后，选取 Benchmark 有限元仿真模型在 120 个自由度非对称质量布置，噪声水平为 10% 下，顶层斜对角振动激励进行实验研究，并对第 3 层测点 A 的结构振动响应信息进行缺失值重构。

图 9 展示了 EEMD-BiLSTM 方法在不同缺失数据比例下的预测结果，其中，蓝色线条代表真实值，红色线条代表预测值。在图 9(a) 中，当数据丢失率为 5% 时，EEMD-BiLSTM 的 R^2 值为 0.876 39，预测值与真实值表现出高度的一致性，误差极小。图 9(b) 显示，数据丢失率增至 10% 时，尽管预测值与真实值之间的偏差有所增加，但预测值仍能有效反映真实值的趋势，表现出良好的重构性能，其 R^2 值为 0.835 02。图 9(c) 中，当数据丢失率达到 15% 时，其 R^2 值为 0.817 84，预测误差进一步增加，然而预测曲线仍大致跟随真实值的主要动态，重构结果仍能接受。总体来看，随着数据丢失率的增加，EEMD-BiLSTM 方法的预测误差有所上升，但

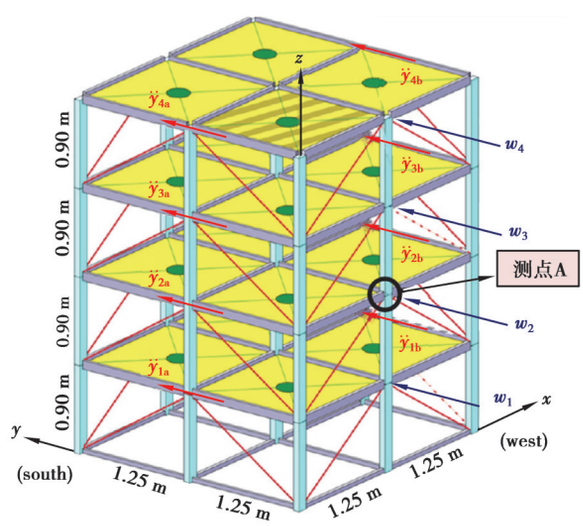
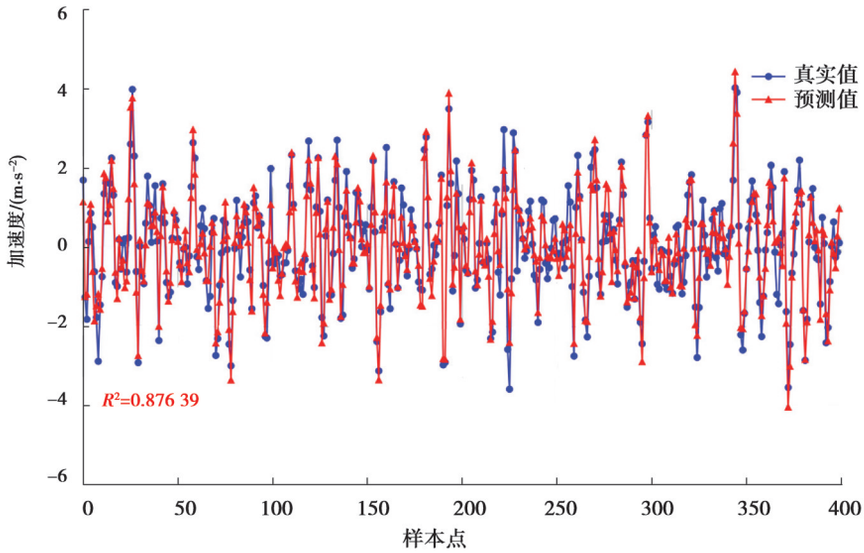


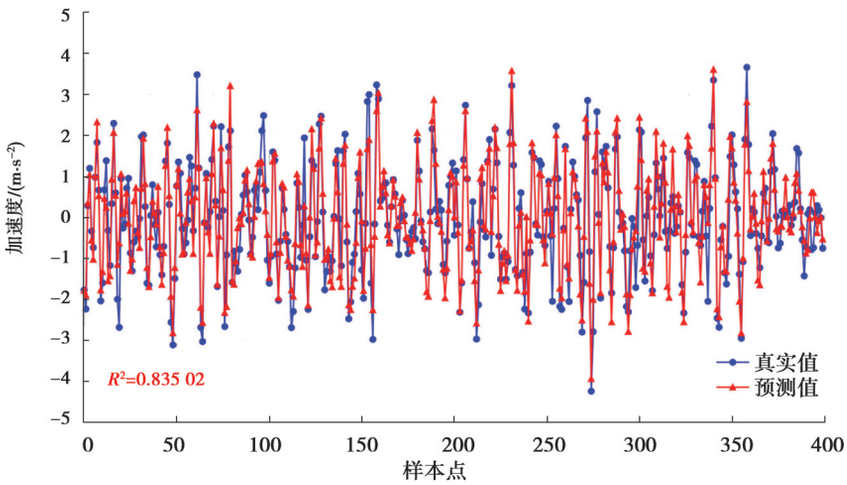
图 8　Benchmark 有限元仿真模型

Fig. 8　Benchmark finite element simulation model

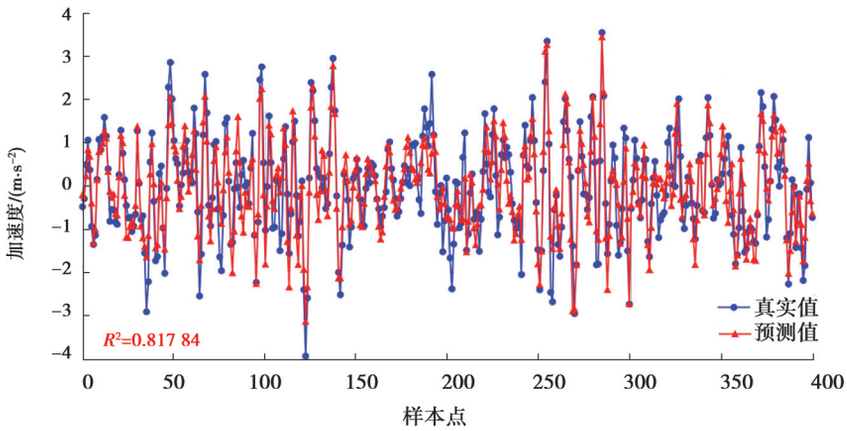
即便在高丢失率下，预测结果仍能基本捕捉到真实值的变化趋势。这表明，EEMD-BiLSTM 方法在处理结构监测缺失数据重构问题上仍然是有效的。



(a) 在5%数据缺失比例下预测结果



(b) 在10%数据缺失比例下预测结果



(c) 在15%数据缺失比例下预测结果

图 9 基于 EEMD-BiLSTM 的不同缺失数据比率下预测结果分析

Fig. 9 Analysis of prediction results under different missing data ratios based on EEMD-BiLSTM

图 10 和表 3 展示了 EEMD-BiLSTM、EEMD-LSTM、BiLSTM 及 LSTM 方法在不同缺失数据比率下的预测性能指标，包括 R^2、MSE、RMSE 和 MAE。通过分析这些指标，可以评估模型的拟合优度和预测精度。随着缺失比率的增加，各个预测方法的 R^2 值普遍下降，说明数据缺失对各个方法性能有显著影响，较高的缺失比率会导致更多的信息损失，从而降低预测能力。

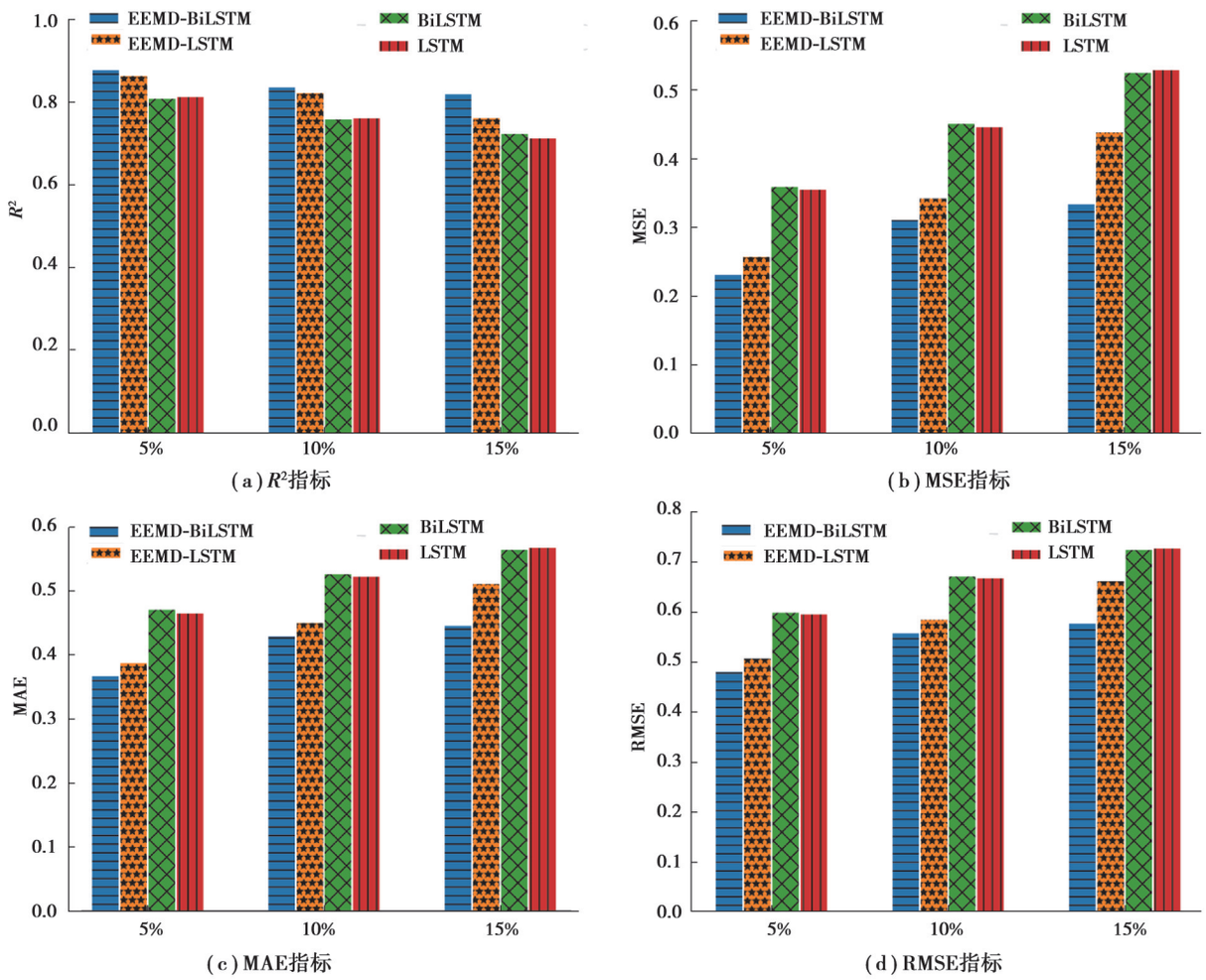


图 10　在不同缺失数据比例下的预测误差

Fig. 10　Prediction error of different algorithms under missing ratio

表 3　基于不同方法的评价指标

Table 3　Evaluation indicators base on different prediction methods

方法	指标	缺失比率/%		
		5	10	15
EEMD-BiLSTM	R^2	**0.876 39**	**0.835 02**	**0.817 84**
	MSE	**0.230 59**	**0.310 64**	**0.333 05**
	RMSE	**0.480 20**	**0.557 35**	**0.577 10**
	MAE	**0.366 77**	**0.428 91**	**0.445 64**
EEMD-LSTM	R^2	0.862 45	0.821 20	0.760 33
	MSE	0.256 79	0.341 63	0.437 76
	RMSE	0.506 74	0.584 49	0.661 63
	MAE	0.387 10	0.450 02	0.510 60
BiLSTM	R^2	0.807 32	0.757 31	0.722 65
	MSE	0.358 56	0.450 38	0.524 59

续表 3

方法	指标	缺失比率/%		
		5	10	15
	RMSE	0.598 80	0.671 10	0.724 29
	MAE	0.470 42	0.525 89	0.564 02
LSTM	R^2	0.811 14	0.759 85	0.711 85
	MSE	0.354 33	0.445 50	0.528 50
	RMSE	0.595 25	0.667 46	0.726 98
	MAE	0.464 77	0.522 09	0.567 22

注:加粗数据表示最好的结果。

EEMD-BiLSTM方法在所有缺失比率下均表现最佳,其R^2值最高,MSE、RMSE和MAE值最低。例如,在缺失比率为5%时,该方法的R^2值为0.876 39,MSE为0.230 59,RMSE为0.480 20,MAE为0.366 77。这些指标表明,EEMD-BiLSTM能够实现较高的预测精度。EEMD-LSTM的性能次于EEMD-BiLSTM,在缺失比率为10%时,EEMD-LSTM的R^2值为0.821 20,而单一的BiLSTM和LSTM的R^2值分别为0.757 31和0.759 85。相比之下,BiLSTM的表现普遍优于LSTM,这可能是由于BiLSTM能够同时捕捉序列的前向和后向信息,从而获得更好的建模能力。在缺失比率为15%时,BiLSTM的R^2值为0.722 65,略高于LSTM的0.711 85。此外,BiLSTM的MSE、RMSE和MAE值也均优于LSTM。

综上所述,EEMD-BiLSTM在所有缺失比率下均表现最优,在时间序列预测方法的性能方面具有显著优势。

4 结 论

在结构健康监测系统中,针对数据采集过程中存在的数据缺失问题,提出EEMD-BiLSTM数据重构方法,并对6层框架结构缩尺模型和Benchmark有限元仿真模型采集的加速度响应数据进行验证。首先,基于数据增强方法获取训练样本,利用EEMD将加速度时间序列分解为若干IMF分量,从中提取加速度数据的动力模态特征,输入到BiLSTM进行缺失数据重构。从实验结果可以得出以下结论:

1)随着数据缺失比率增加,EEMD-BiLSTM指标衰减较缓慢,这主要由于EEMD分解后,降低了原始数据波动造成的影响,有效提取加速度的数据动力模态特征,BiLSTM利用过去和未来的上下文动力模态信息进行建模,实现序列中的每个时间点准确分析,提高数据缺失预测的精度。

2) BiLSTM具有2个相向传递的LSTM结构,可以更好捕捉时间序列数据中的长期依赖关系,但是非平稳、非线性的数据会影响预测缺失数据结果。在预测缺失数据之前,对数据进行EEMD分解,将平稳的数据带入BiLSTM。EEMD-BiLSTM的组合模型比单一BiLSTM模型的预测精度高。

3) 针对6层框架结构缩尺模型和Benchmark有限元仿真模型,在5%、10%和15%缺失数据比例下,本研究对比EEMD-BiLSTM、EEMD-LSTM、BiLSTM和LSTM方法。结果表明,提出的EEMD-BiLSTM方法具有最高预测精度,在6层框架结构缩尺模型,5%的缺失数据R^2精度分别为0.854 78,10%的缺失数据R^2精度0.830 91,15%的缺失数据R^2精度81 252。在Benchmark有限元仿真模型,R^2精度分别为0.876 39(5%缺失),0.835 02(10%缺失),0.817 84(15%缺失)。

参考文献

[1] Bao Y Q, Li H. Machine learning paradigm for structural health monitoring[J]. Structural Health Monitoring, 2021, 20(4): 1353-1372.

[2] Chen Z C, Bao Y Q, Li H, et al. A novel distribution regression approach for data loss compensation in structural health monitoring[J]. Structural Health Monitoring, 2018, 17(6): 1473-1490.

[3] Ou J P, Li H. Structural health monitoring in China: review and future trends[J]. Structural Health Monitoring, 2010, 9(3): 219-231.

[4] Wan H P, Ni Y Q. Bayesian multi-task learning methodology for reconstruction of structural health monitoring data[J]. Structural Health Monitoring, 2019, 18(4): 1282-1309.

[5] Li Y T, Bao T F, Chen Z X, et al. A missing sensor measurement data reconstruction framework powered by multi-task Gaussian process regression for dam structural health monitoring systems[J]. Measurement, 2021, 186: 110085.

[6] Chen Z C, Li H, Bao Y Q. Analyzing and modeling inter-sensor relationships for strain monitoring data and missing data imputation: a copula and functional data-analytic approach[J]. Structural Health Monitoring, 2019, 18(4): 1168-1188.

[7] Zhang Z Y, Luo Y Z. Restoring method for missing data of spatial structural stress monitoring based on correlation[J]. Mechanical Systems and Signal Processing, 2017, 91: 266-277.

[8] Rosafalco L, Manzoni A, Mariani S, et al. Fully convolutional networks for structural health monitoring through multivariate time series classification[J]. Advanced Modeling and Simulation in Engineering Sciences, 2020, 7(1): 38.

[9] Hua Y X, Zhao Z F, Li R P, et al. Deep learning with long short-term memory for time series prediction[J]. IEEE Communications Magazine, 2019, 57(6): 114-119.

[10] Peng T, Zhang C, Zhou J Z, et al. An integrated framework of bi-directional long-short term memory(BiLSTM) based on sine cosine algorithm for hourly solar radiation forecasting[J]. Energy, 2021, 221: 119887.

[11] Li X, Zhuang W, Zhang H. Short-term power load forecasting based on gate recurrent unit network and cloud computing platform[C]//4th International Conference on Computer Science and Application Engineering. Sanya China: ACM, 2020: 1-6.

[12] Deng Y, Jia H, Li P, et al. A deep learning methodology based on bidirectional gated recurrent unit for wind power prediction[C]//14th IEEE Conference on Industrial Electronics and Applications (ICIEA). Xi'an, China: IEEE, 2019: 591-595.

[13] Fan G, Li J, Hao H. Lost data recovery for structural health monitoring based on convolutional neural networks[J]. Structural Control and Health Monitoring, 2019, 26(10): e2433.

[14] Liu H, Ding Y L, Zhao H W, et al. Deep learning-based recovery method for missing structural temperature data using LSTM network[J]. Structural Monitoring and Maintenance, 2020, 7(2): 109-124.

[15] Oh B K, Glisic B, Kim Y, et al. Convolutional neural network-based data recovery method for structural health monitoring[J]. Structural Health Monitoring, 2020, 19(6): 1821-1838.

[16] Du B W, Wu L Y, Sun L L, et al. Heterogeneous structural responses recovery based on multi-modal deep learning[J]. Structural Health Monitoring, 2023, 22(2): 799-813.

[17] Lei X M, Sun L M, Xia Y. Lost data reconstruction for structural health monitoring using deep convolutional generative adversarial networks[J]. Structural Health Monitoring, 2021, 20(4): 2069-2087.

[18] Li L C, Zhou H J, Liu H L, et al. A hybrid method coupling empirical mode decomposition and a long short-term memory network to predict missing measured signal data of SHM systems[J]. Structural Health Monitoring, 2021, 20(4): 1778-1793.

[19] Flandrin P, Gonçalvès P, Rilling G. EMD equivalent filter banks, from interpretation to applications[M]//Interdisciplinary Mathematical Sciences: 5. Singapore: World Scientific, 2005: 57-74.

[20] Huang N E, Shen Z, Long S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[J]. Proceedings of the Royal Society of London Series A: Mathematical, Physical and Engineering Sciences, 1998, 454(1971): 903-995.

[21] Yang K, Ding Y L, Geng F F, et al. A multi-sensor mapping Bi-LSTM model of bridge monitoring data based on spatial-temporal attention mechanism[J]. Measurement, 2023, 217: 113053.

[22] Anjneya K, Roy K. Acceleration time history data for experimental shear building model (Kumar Anjneya & Koushik Roy)[EB/OL]. (2021-01-04)[2023-08-23]. https://data.mendeley.com/datasets/snmz587nvb/2.

[23] Saidin S S, Kudus S A, Jamadin A, et al. Vibration-based approach for structural health monitoring of ultra-high-performance concrete bridge[J]. Case Studies in Construction Materials, 2023, 18: e01752.

[24] Kurian B, Liyanapathirana R. Machine learning techniques for structural health monitoring[M]//13th International Conference on Damage Assessment of Structures. Singapore: Springer Singapore, 2020: 3-24.

[25] Nasrollahi A, Deng W, Ma Z Y, et al. Multimodal structural health monitoring based on active and passive sensing[J]. Structural Health Monitoring, 2018, 17(2): 395-409.

[26] Bao Y Q, Beck J L, Li H. Compressive sampling for accelerometer signals in structural health monitoring[J]. Structural Health Monitoring, 2011, 10(3): 235-246.

[27] He Y Y, Zhang L K, Chen Z S, et al. A framework of structural damage detection for civil structures using a combined multi-scale convolutional neural network and echo state network[J]. Engineering with Computers, 2023, 39(3): 1771-1789.

（编辑　侯　湘）