



基于长短时记忆网络的结构动态载荷预测方法

樊昱玮 郭腾博 李哲 洪良友 刘超 蒋东翔

Structural dynamic load prediction method based on long short-term memory network

FAN Yuwei, GUO Tengbo, LI Zhe, HONG Liangyou, LIU Chao, JIANG Dongxiang

在线阅读 View online: <https://doi.org/10.19693/j.issn.1673-3185.03463>

您可能感兴趣的其他文章

Articles you may be interested in

基于混合深度学习的燃气轮机动态过程关键参数在线辨识

Hybrid deep learning-based online identification method for key parameters of gas turbine dynamic process
中国舰船研究. 2023, 18(3): 222-230 <https://doi.org/10.19693/j.issn.1673-3185.02914>

长峰与短峰不规则波对船舶运动与波浪载荷的影响分析

Difference in ship motion and wave load in long and short crested irregular waves
中国舰船研究. 2023, 18(1): 107-115 <https://doi.org/10.19693/j.issn.1673-3185.02679>

基于混合深度学习的压气机喘振快速诊断及自抗扰控制方法

Rapid diagnosis and active disturbance rejection control of compressor surge based on hybrid deep learning
中国舰船研究. 2024, 19(2): 187-196 <https://doi.org/10.19693/j.issn.1673-3185.03259>

基于深度学习的多级曲线加筋壁板布局优化设计

Layout optimization design of hierarchical curvilinearly stiffened panels based on deep learning
中国舰船研究. 2021, 16(4): 86-95 <https://doi.org/10.19693/j.issn.1673-3185.02188>

基于SMOTE算法和动态代理模型的船舶结构可靠性优化

Reliability-based design optimization of ship structure using SMOTE algorithm and dynamic surrogate model
中国舰船研究. 2020, 15(5): 114-123 <https://doi.org/10.19693/j.issn.1673-3185.01657>

基于兴趣子域动态代理模型的船舶结构可靠性优化

Reliability-based optimization of ship structure based on interest subdomain dynamic surrogate model
中国舰船研究. 2021, 16(4): 96-107 <https://doi.org/10.19693/j.issn.1673-3185.02043>



扫码关注微信公众号，获得更多资讯信息

本文网址: <http://www.ship-research.com/cn/article/doi/10.19693/j.issn.1673-3185.03463>

期刊网址: www.ship-research.com

引用格式: 樊昱玮, 郭腾博, 李哲, 等. 基于长短时记忆网络的结构动态载荷预测方法 [J]. 中国舰船研究, 2024, 19(6): 228–236.

FAN Y W, GUO T B, LI Z, et al. Structural dynamic load prediction method based on long short-term memory network[J]. Chinese Journal of Ship Research, 2024, 19(6): 228–236 (in both Chinese and English).

基于长短时记忆网络的结构动态 载荷预测方法



扫码阅读全文

樊昱玮¹, 郭腾博¹, 李哲², 洪良友², 刘超^{*1}, 蒋东翔^{*1}

1 清华大学 能源与动力工程系, 北京 100084

2 北京强度环境研究所, 北京 100076

摘要: [目的] 针对传统代理模型无法处理具有时间依赖性的动态过程和异构数据的问题, 提出一种基于长短时记忆网络(LSTM)的动态载荷代理模型方法。[方法] 代理模型包含载荷特征编码和载荷响应解码2个模块。首先, 通过载荷特征编码模块的LSTM对动态外载荷时间序列进行特征提取; 然后, 将外载荷时序特征与结构参数特征进行融合, 由载荷解码模块的LSTM进一步进行特征提取并生成最终输出, 从而综合考虑动态外载荷时间序列和结构参数一维特征的异构数据输入, 预测结构内力响应时间历程; 最后, 在有限元仿真数据集上对模型进行精度评估, 并与其他代理模型方法进行对比。[结果] 结果显示, 该动态载荷代理模型的平均精度可达98%, 高于其他对比方法, 且计算速度相较于有限元方法更快。[结论] 所提方法可解决时序-非时序异构数据的代理模型问题, 具有精度高、效率高的优点, 在快速迭代计算场景下能够发挥较大作用。**关键词:** 结构优化; 动态载荷; 人工智能; 代理模型; 深度学习; 长短时记忆网络

中图分类号: U662.9

文献标志码: A

DOI: 10.19693/j.issn.1673-3185.03463

Structural dynamic load prediction method based on long short-term memory network

FAN Yuwei¹, GUO Tengbo¹, LI Zhe², HONG Liangyou², LIU Chao^{*1}, JIANG Dongxiang^{*1}

1 Department of Energy and Power Engineering, Tsinghua university, Beijing 100084, China

2 Beijing Institute of Structure and Environment Engineering, Beijing 100076, China

Abstract: [Objective] To address the limitations of traditional surrogate models in handling time-dependent dynamic processes and heterogeneous data, this paper proposes a dynamic load surrogate model method based on a long short-term memory (LSTM) network. [Methods] The surrogate model is comprised of two modules: the load feature encoder and load response decoder. First, the LSTM in the load feature encoder performs feature extraction on the time series of dynamic external loads. Next, the extracted load features are combined with the structural parameter features. The LSTM in the load decoder conducts further feature extraction and finally generates output while comprehensively considering the heterogeneous data input of the dynamic external load time series and one-dimensional structural parameter features in order to predict the time history of internal force responses. Finally, the model's accuracy is evaluated using a finite element simulation dataset and compared with other surrogate model methods. [Results] The results show that the average accuracy of the dynamic load surrogate model can reach 98%, which is higher than that of other methods, and its calculation speed is faster than that of the finite element method. [Conclusions] The proposed method addresses the issue of heterogeneous data involving both time-series and non-time-series features, and offers advantages such as high accuracy and efficiency, making it effective for fast iterative computation tasks.

Key words: structural optimization; dynamic loads; artificial intelligence; surrogate model; deep learning; long short-term memory (LSTM) network

收稿日期: 2023-07-15 修回日期: 2023-12-05 网络首发时间: 2024-10-10 11:12

基金项目: 航空发动机及燃气轮机基础科学中心资助项目(P2022-C-I-002-001)

作者简介: 樊昱玮, 男, 2000年生, 博士生。研究方向: 代理模型与物理信息网络。E-mail: fan-yw22@mails.tsinghua.edu.cn

刘超, 男, 1986年生, 博士, 副研究员。研究方向: 动力系统预测诊断与健康。E-mail: cliu5@tsinghua.edu.cn

蒋东翔, 男, 1963年生, 博士, 教授。研究方向: 动力系统故障诊断。E-mail: jiangdx@tsinghua.edu.cn

*通信作者: 刘超, 蒋东翔

0 引言

在船舶、空天、建筑等诸多领域,通过结构优化获得良好的动力学或静力学特征并同时实现安全、可靠、经济等目标是非常有价值的。结构优化涉及到的影响因素众多,结构在不同受力情况和结构参数下的力学响应均需考虑在内,然后再通过设计准则、数学规划或其他合适的算法对参数进行优化调整。

对结构优化而言,工作载荷预测是影响其效果的重要因素。一方面,精准预测工作载荷是优化的基础,对载荷过高或过低的估计均会带来安全性或是性能上的缺失。在结构设计的不同阶段,通常采用实验方法或是以有限元为主的数值计算方法。这些方法的有效性已得到广泛验证,并能保证工作载荷预测的精确度。另一方面,载荷预测的效率决定了优化算法的实现成本。通常而言,需要通过大量的实验或计算获取数据来实现对目标函数的估计,且数据量越大,对目标函数的估计越准确,这就意味着需要对满足约束条件的所有参数取值进行遍历。在设计参数量较大时,维度灾难的存在将导致这种遍历需要进行上万次甚至更多。出于时间上的考虑,想通过实验进行如此大量的验证不太可能,而对于以有限元为主的数值计算方法来说,如此巨大的算例数量也难以实现。除了庞大的计算量和计算时间外,网格划分也带来了巨大的人力成本。这就使得这些方法无法在较短的开发周期内达到精准的结构优化效果。可见,一种高效、快速且准确的载荷预测方法对于结构优化而言意义重大。

代理模型是解决这一问题的高效替代方法,被越来越多地应用于结构优化领域^[1-2]。在通过有限元计算或实验获取一定量的数据后,可以根据数据用黑箱数学模型近似模拟真实物理模型,从而描述参数与载荷之间的关系。在之后的参数遍历过程中,只需通过函数的计算就可以获得载荷值,从而显著降低遍历成本^[3]。例如,罗文俊等^[4]的研究表明,引入代理模型的结构优化方法可大幅降低计算成本,提高计算效率,这对于缩短产品开发周期而言非常有利。代理模型方法包括响应面法(response surface)^[5]、克里金法(Kriging)^[6]、径向基函数法(radial basis function, RBF)^[7]以及各种机器学习(machine learning)方法。响应面法、克里金法和RBF法是较为经典的代理模型构建方法,而其中克里金法的应用最为广泛^[8-12]。这些经典方法是通过对一些简单函数的线性组合来完

成对复杂函数的近似,例如多项式、指数函数等,因此处理非线性和高维问题的能力略显不足。近年来,随着人工智能特别是机器学习方法的迅速发展,随机森林(random forest)、支持向量回归(support vector regression, SVR)以及人工神经网络(artificial neural networks)等机器学习方法被引入代理模型领域。机器学习具有强大的非线性拟合能力,这是其相对于传统代理模型方法的一个巨大优势。此外,机器学习的高维特征处理能力也强于传统方法,这使得机器学习方法在代理模型领域的适用性很高,基于决策树^[13-14]和SVR^[15-16]的代理模型均取得了一些研究成果。特别值得一提的是深度学习方法,通过模型结构的设计和模型容量的加大,深度学习模型具有强大的特征提取能力,不仅使代理模型的非线性能力更强,也使得端到端的代理模型成为可能,可以直接使用图像或时间序列作为输入输出^[17-19]。

在载荷预测问题上,深度学习强大的特征提取能力和灵活的架构使得建立动态外载荷输入与结构载荷响应之间的代理模型成为可能,而受制于方法的局限性,传统的代理模型往往不能将外载荷和结构参数的异构数据作为代理模型的输入。由于外载荷为时间序列的二维结构,而结构参数为非时间序列的一维结构数据,故传统代理模型方法通常只针对其中一种类型的数据。因此,有部分^[14]传统代理模型只针对静力学进行分析。然而,在实际情况下,静力学分析的适用范围有限,且无法为复杂的动态过程提供足够的载荷预测信息。另有部分代理模型^[20-21]是针对动力学进行分析,能够处理动态载荷预测问题,但不能考虑结构参数对载荷响应的影响,无法满足结构优化的需要。

因此,利用深度学习端到端的学习能力,本文将对载荷预测代理模型的输入输出进行重新定义,构建一种基于长短时记忆网络(long short-term memory, LSTM)的动态载荷代理模型,以动态外载荷的时间序列数据和分析对象结构参数的非时序数据作为输入,预测结构内部多个截面的内力响应时间历程。该模型以编码器-解码器框架为基础,通过编码器对动态外载荷输入进行特征提取,通过解码器融合外载荷特征和结构参数,然后输出最终结果。在实验验证方面,通过有限元分析计算同一梁结构在不同结构参数和外载荷激励下不同截面的内力响应,并以计算结果作为数据集,训练代理模型并验证其精度。最后,将以结构参数相关性分析为例,将所构建的代理模型应用到大量算例的快速计算场景,以展示代理模

型较高的计算效率。

1 代理模型构建方法

1.1 问题描述

本文主要研究不同结构参数的梁模型在动态外载荷条件下的内力响应,研究对象如图1所示。该模型为一维空心梁模型,共有6段,每段梁之间通过连接段连接。梁和连接段虽材料属性不一样,但均为线弹性,其外径相同但厚度不同。为简化计算,将在梁的特定位置处赋予集中质量,而其他梁段则均无质量。

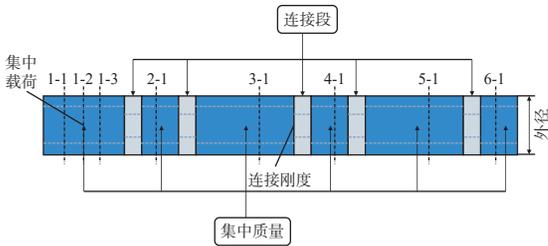


图1 梁模型示意图

Fig. 1 Schematic diagram of beam

在最左端集中质量点处,该梁模型受到 x 方向和 y 方向的动态集中载荷 F_x 和 F_y 。在不同梁段分别选取截面,监测其内力响应时间历程。所选取的截面分别为左端梁段3个截面、其他梁段各1个截面,这是因为在左端受到集中载荷的梁段的内力响应更受关注。所监测的内力包括 x 方向轴力和 y 方向弯矩。这意味着,代理模型关注的输出共有16个不同的内力。

1.2 代理模型构建

假设梁模型的结构参数为 P , 受到的外载荷为 L , 内部各截面的内力响应为 R , 三者之间的关系为:

$$R = f(P, L) \quad (1)$$

代理模型使用数据的方法,以一个黑箱数学模型估计 f 建立近似关系 \hat{f} 来代替数值计算或实验方法的计算过程,从而显著减少计算时间。其关系式如下:

$$R = \hat{f}(P, L, \theta) + \varepsilon \quad (2)$$

式中: θ 为代理模型参数; ε 为误差。

代理模型是一种基于数据而非物理规律的方法,其参数需要通过已有数据估计获得,而这些已有数据可以通过数值计算或是实验获得。因此,代理模型并不能取代数值计算或实验,必须先通

过这两者获取一部分数据完成对参数的估计后,才能在之后的迭代优化计算中起到加速的作用。

不同于传统代理模型,本文构建的代理模型将外载荷时间历程直接以时间序列二维数据的形式输入,通过一定的深度神经网络设计融合结构参数非时序一维输入,最终输出梁结构上多个截面的内力响应历程时间序列。外载荷输入包括给定位置外载荷沿 x, y 方向的分量 F_x 和 F_y , 结构参数输入包括梁模型梁段与连接段的连接刚度 K_c 、梁外径 R_b 、梁材料弹性模量 E_b 和连接段材料弹性模量 E_c , 输出参数包括各给定截面的轴力 F_s 和 y 方向弯矩 M_s 。

1.3 模型结构

本文设计的代理模型深度网络主要采用编码-解码结构将时序外载荷历程和非时序结构参数融合。首先,对于外载荷,通过 LSTM 网络对时间序列的处理能力,对输入的外载荷时间历程进行编码以提取特征;而对于结构参数,则通过线性层提取其特征。在解码器中将结构参数特征复制升维,再将外载荷每个时间步的特征与结构参数特征融合,之后通过另一个 LSTM 网络解码器对融合特征进行进一步的回归,最终输出与外载荷时间步相对应的内力响应时间序列。该模型结构示意图如图2所示,其中 \hat{L} 为外载荷特征, \hat{P} 为结构参数特征,下标 t 为时间步。

1.3.1 载荷特征编码

在涉及时间依赖性的动态载荷问题中,外载荷输入随时间变化,因此在代理模型中需以时间序列的形式表达,才能提取其隐藏于时间中的潜在模式。鉴此,将采用循环神经网络(recurrent neural network, RNN)类模型对载荷时间特征进行编码,RNN类模型的普遍结构如图3所示。图中, X_t 为时间步 t 的输入, H_t 为隐藏状态。值得一提的是,在动态结构载荷响应问题中,某一时刻载荷响应的值只由当前时间点及之前时间点决定,而 RNN 类模型包括很多沿时间方向连接排列的单元,每个单元接受当前时刻的输入和前一个时间点的隐藏状态并产生输出,其中隐藏状态包括前一个时间点及之前所有时间步单元提取的信息,这与载荷响应问题的特点非常契合。

具体地,针对动态载荷长序列的特点,载荷编码器选取 LSTM 网络。普通的 RNN 在处理长序列时存在梯度消失或爆炸的问题,导致性能不理想。为此, Hochreiter 等^[22] 提出以 LSTM 作为解决方案,在 RNN 单元中增加3个门,即输入门、输出门和遗忘门,门选择性地传输包含在单元状态

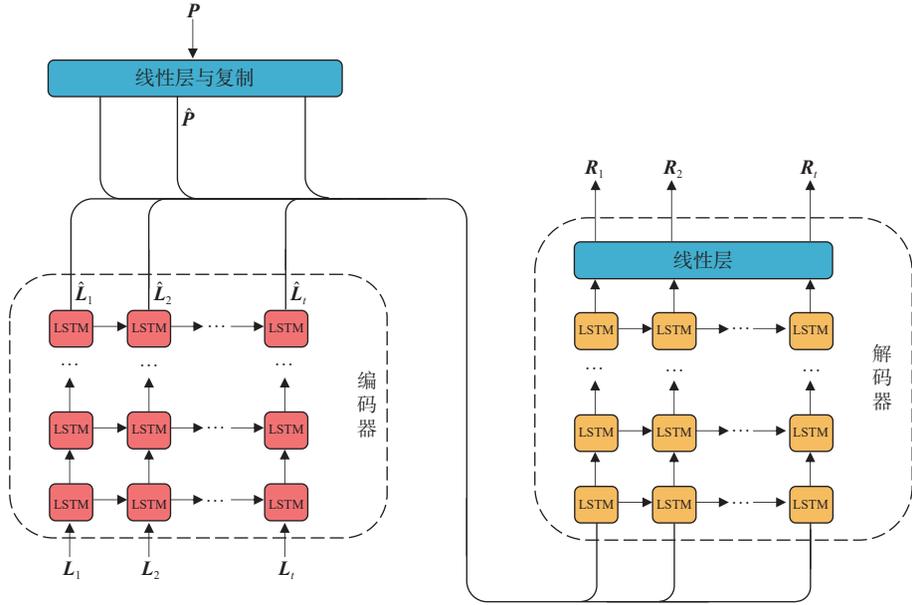


图 2 代理模型结构

Fig. 2 Structure of surrogate model

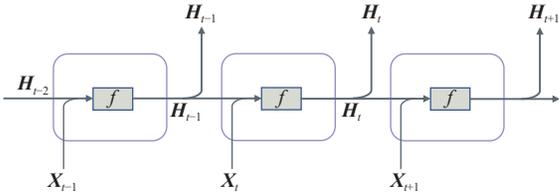


图 3 RNN 类模型结构

Fig. 3 Structure of RNN class model

中的信息,解决了梯度消失和梯度爆炸的问题,使得 LSTM 在处理长序列问题上相比传统 RNN 更具优势。在载荷预测问题中,为了获得高精度的结果,离散的时间步长非常短,有时不足 1 ms,这就导致短时间的仿真产生的时间序列也有成百上千的时间步,属于长序列的范畴,因此,比较适合采用 LSTM 进行处理。

LSTM 中单个记忆单元的结构如图 4 所示。图中, C_t 表示单元状态,其值受 3 个 σ 函数的影响,从左到右,第 1 个 σ 函数是遗忘门,它决定在上一个时间步的单元状态中哪些信息将被丢弃;第 2 个 σ 函数是输入门,其将输入的新信息添加到单元状态;最后一个 σ 函数是输出门,选择哪些信息应该输出到隐藏状态 H_t 。LSTM 单元可以表示如下:

$$\begin{aligned}
 I_t &= \sigma(X_t W_{xi} + H_{t-1} W_{hi} + b_i) \\
 F_t &= \sigma(X_t W_{xf} + H_{t-1} W_{hf} + b_f) \\
 O_t &= \sigma(X_t W_{xo} + H_{t-1} W_{ho} + b_o) \\
 \tilde{C}_t &= \tanh(X_t W_{xc} + H_{t-1} W_{hc} + b_c) \\
 C_t &= F_t \odot C_{t-1} + I_t \odot \tilde{C}_t \\
 H_t &= O_t \odot \tanh(C_t)
 \end{aligned} \tag{3}$$

式中: I_t 为输入门输出; F_t 为遗忘门输出; O_t 为输

出门输出; \tilde{C}_t 为候选单元状态; W 为权重; b 为偏差; σ 和 \tanh 分别为 sigmoid 函数以及双曲正切函数;下标 x, h, i, f, o, c 分别表示该权重或偏置与 $X_t, H_t, I_t, F_t, O_t, \tilde{C}_t$ 有关。

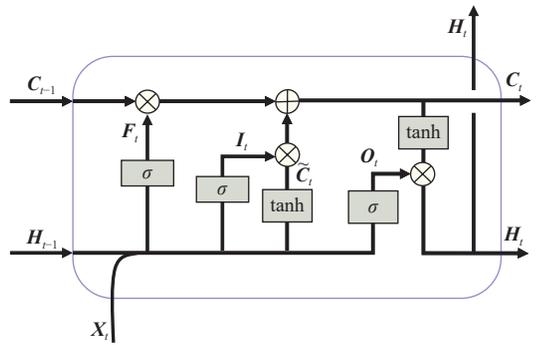


图 4 LSTM 单元结构

Fig. 4 Structure of LSTM cell

LSTM 编码网络可以提取时间序列的时间依赖性信息,完成对时间序列的特征提取,其表达式如下:

$$\hat{L}_t = \text{enc}(L_1, \dots, L_t), \quad t = 1, 2, \dots, T \tag{4}$$

式中, T 为时间序列总长度。

1.3.2 载荷响应解码

编码器完成对动态外载荷的特征提取后,由解码器融合外载荷特征和结构参数特征,然后解码输出结构内部动态内力响应。在此之前,首先需对结构参数进行简单的特征提取。使用一个不带任何非线性激活函数的线性全连接层,目的是对特征维度进行变换,然后根据超参数搜索的结果选择将结构参数投影到更高维度以获取高维特征,或者将其压缩到更低维度以排除与输出相关

性相对较低的参数,公式如下:

$$\hat{P} = W_x P + b \quad (5)$$

经过线性层后,这些特征仍是一维数据,因此将其复制,复制次数与外载荷时间序列长度一致,随后将复制后的特征与外载荷特征进行组合,方法如图5所示。组合后的特征为二维数据,其长度等于外载荷时间序列长度,其特征维数等于结构参数特征数与外载荷特征数之和。

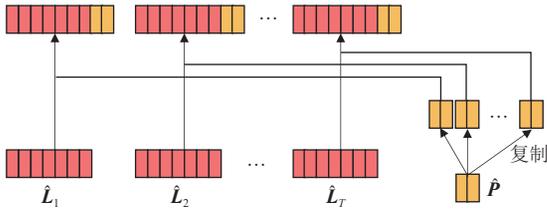


图5 复制与特征组合

Fig. 5 Repeat and feature combination

获得组合特征后,通过 LSTM 网络对其进行解码。该 LSTM 网络与编码器中的 LSTM 网络是相互独立的,这就意味着编码器和解码器的 LSTM 拥有不同的参数和超参数。在该 LSTM 网络后,使用一个线性全连接层,对 LSTM 网络的输出进行特征维度变换,输出 8 个截面共 16 个内力,避免 LSTM 网络的超参数受到输出维度的限制。

结构参数特征提取、复制和特征组合,以及组合特征解码共同构成解码器,可用下式表示:

$$R_t = \text{dec}(P, \hat{L}_1, \dots, \hat{L}_t) \quad (6)$$

1.4 评价指标

本文共使用 2 种指标对模型进行评价:均方误差 E_1 和相对误差 E_2 ,其中均方误差是神经网络处理回归问题时最常用的指标,本文将作为代理模型的损失函数用于模型参数更新,其公式为

$$E_1 = \frac{1}{NTD} \sum_{i=1}^N \sum_{j=1}^T \sum_{k=1}^D (y_{ijk} - \hat{y}_{ijk})^2 \quad (7)$$

式中: y_{ijk} 为真实值; \hat{y}_{ijk} 为模型输出值; N 为样本数量; D 为输出变量数。

为了更直观地表示代理模型对内力响应的预测精度,便于对代理模型进行评价,定义相对误差如下:

$$E_2 = \frac{1}{NTD} \sum_{i=1}^N \sum_{j=1}^T \sum_{k=1}^D \frac{|y_{ijk} - \hat{y}_{ijk}|}{\max_j |y_{ijk}|} \quad (8)$$

之所以不逐项计算相对误差,而要分母在时间维度上取最大绝对值,这是因为输出存在一些值为 0 的时间点,会影响相对误差的计算。不同

方法的评价和方法间的对比将主要基于此定义下的相对误差来进行。

2 数值实验

对于本文构建的动态载荷代理模型,在通过有限元计算获得的仿真数据集上对其进行验证。数值实验的主要内容包括数据集构造与预处理、模型训练与验证、结构参数相关性分析案例。下面,将分步对数值实验进行介绍。

2.1 数据集构造与预处理

通过 ABAQUS 软件进行有限元分析,本文构造了包含 10 000 个算例的仿真数据集。研究对象模型前文已介绍。在仿真数据集中,不同算例的外载荷和结构参数不同。对于外载荷,根据实际情况,假设其以集中载荷的形式施加在梁最左端的集中质量处。为尽量提高载荷复杂度,模拟梁段在工作过程中随机的受力情况,施加的载荷为随机载荷,其功率谱密度在 20~200 Hz 之间为常数,这覆盖了梁模型的前 4 阶模态。在其他频段,功率谱密度为 0。每个算例均随机生成 x 方向和 y 方向的随机载荷,这就意味着每个算例、每个方向的载荷分量均不同。

对于结构参数,为确保参数分布的均匀性,本文采用拉丁超立方方法在参数空间中随机采样 10 000 次,每个算例采用一组结构参数。改变的结构参数包括梁与连接段的连接刚度 K_c 、梁外径 R_b 、梁的弹性模量 E_b 和连接段材料的弹性模量 E_c 。不同参数的基准值和变化范围如表 1 所示。

表 1 结构参数的取值范围

Table 1 Range of structural parameters

	输入参数			
	连接刚度 $K_c / (\text{N} \cdot \text{m}^{-1})$	梁外径 R_b / m	梁的弹性 模量 E_b / Pa	连接段材料 弹性模量 E_c / Pa
基准	1×10^7	0.325	7.10×10^{10}	2.10×10^{11}
下限	1×10^7	0.309	7.10×10^{10}	2.10×10^{11}
上限	2×10^7	0.341	7.81×10^{10}	2.31×10^{11}

对于每个算例,施加时长为 0.2 s 的随机载荷,结构参数按照拉丁超立方采样结果设置,有限元仿真时输入载荷的时间间隔为 0.05 ms。仿真时间设置为 0.2 s,仿真步长为 0.5 ms,输出时间间隔为 0.001 s。在这样的设置下,载荷输入和输出的采样频率能够保证不会丢失随机载荷和载荷响应在 20~200 Hz 频段的信息。

在所有算例完成计算后,对算例进行数据处理。首先,对外载荷输入进行降采样。出于仿真

精度的要求, 计算时, 将外载荷输入的采样频率设置为输出频率的 20 倍。这就造成外载荷输入的时间序列长度极长, 使得代理模型的计算量巨大。考虑到输出的采样频率已达到随机载荷最高频率的 5 倍, 不会丢失外载荷及载荷响应的高频信息, 因此对代理模型的外载荷输入进行 20 倍降采样, 使其采样频率与输出的采样频率相同。经降采样后的外载荷时序长度将与内力响应输出一致, 均为 200, 这将使得代理模型的结构设计更为简单。

经降采样之后, 针对外载荷、结构参数与内力响应进行数据标准化, 用以加快神经网络训练速度, 提高训练精度。标准化的公式为:

$$\hat{X} = (X - X_{\text{mean}}) / X_{\text{std}} \quad (9)$$

式中: X 为单一特征变量; X_{mean} 为单一特征变量的平均值; X_{std} 为标准差。标准化将所有特征映射至平均值为 0、标准差为 1 的分布, 使每个特征都有相近的取值范围。对于结构参数, 在样本间计算平均值和标准差以进行标准化, 而对于外载荷和内力响应的时序而言, 则在样本和时间维度统一进行标准化。

完成数据预处理后, 按 60%, 20%, 20% 的比例将数据集划分为训练集、验证集和测试集, 即训练集包含 6 000 个样本, 验证集和测试集各包含 2 000 个样本。由于每个算例的载荷是随机生成的, 结构参数各不相同, 因此这 3 个数据集中的数据并不重复。

2.2 模型训练与验证

经过模型在训练集上的训练, 以及根据验证集进行的超参数选取, 确定采用的超参数如下: 编码 LSTM 的隐藏单元数 $d_{\text{enc}} = 8$, 层数 $n_{\text{enc}} = 2$; 结构参数投影线性层的隐藏单元数 $d_{\text{pram}} = 3$, 即对结

构参数进行了降维; 解码 LSTM 的隐藏单元数 $d_{\text{dec}} = 32$, 层数 $n_{\text{dec}} = 2$; 最后一个线性层的隐藏单元数 $d_{\text{lin}} = 16$ 。使用 Adam 对模型进行训练, 批量数 $B = 128$, 学习率 $\eta = 0.005$, 一阶矩估计的指数衰减率 $\beta_1 = 0.9$, 二阶矩估计的指数衰减率 $\beta_2 = 0.999$, 训练轮次 n_{epoch} 为 200 轮。表 2 列出了所有超参数的数值。

表 2 超参数设置
Table 2 Configuration of hyperparameters

超参数	数值	超参数	数值
n_{enc}	2	d_{enc}	8
n_{dec}	2	d_{dec}	32
d_{pram}	3	d_{lin}	16
B	128	η	0.005
β_1	0.9	β_2	0.999
n_{epoch}	200		

表 3 展示了训练后的模型在测试集数据上预测不同内力响应时的相对误差。该模型的平均相对误差为 1.5%。对于 x 方向轴力 F_s 和 y 方向弯矩 M_s , 相对误差的分布如图 6 所示。结果显示, F_s 的平均相对误差较小, 为 1.3%, 相对误差较小的区域概率密度较大, 而 M_s 的平均相对误差较大, 为 1.6%。图 7 以测试集误差最小的样本 1 和误差最大的样本 2 为例, 直观地展示了代理模型对 1-2 截面轴力 F_s 和弯矩 M_s 的预测结果与真实值的对比。

表 3 代理模型相对误差

Table 3 Relative error of surrogate model

参数	不同截面内力响应相对误差/%							
	1-1	1-2	1-3	2-1	3-1	4-1	5-1	6-1
轴力 F_s	1.2	1.4	1.2	1.2	0.9	1.2	1.6	1.7
弯矩 M_s	1.7	2.0	1.7	1.7	1.5	1.3	1.4	1.6

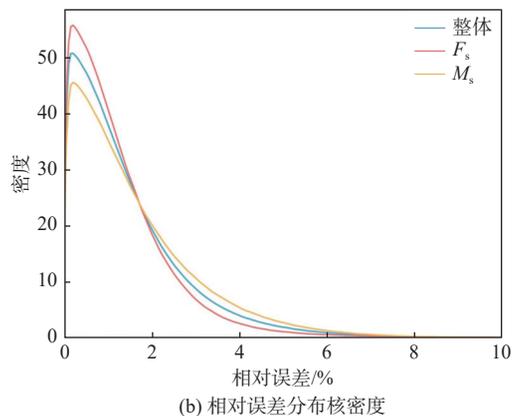
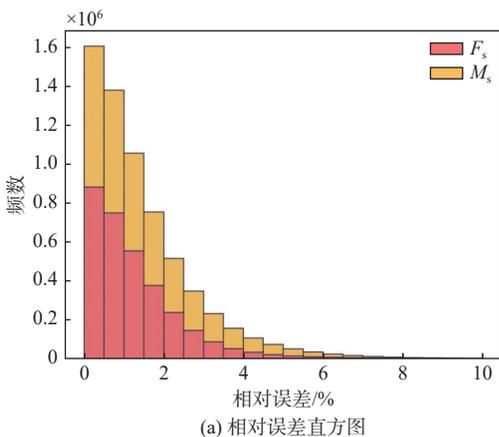


图 6 相对误差分布

Fig. 6 Distribution of relative error

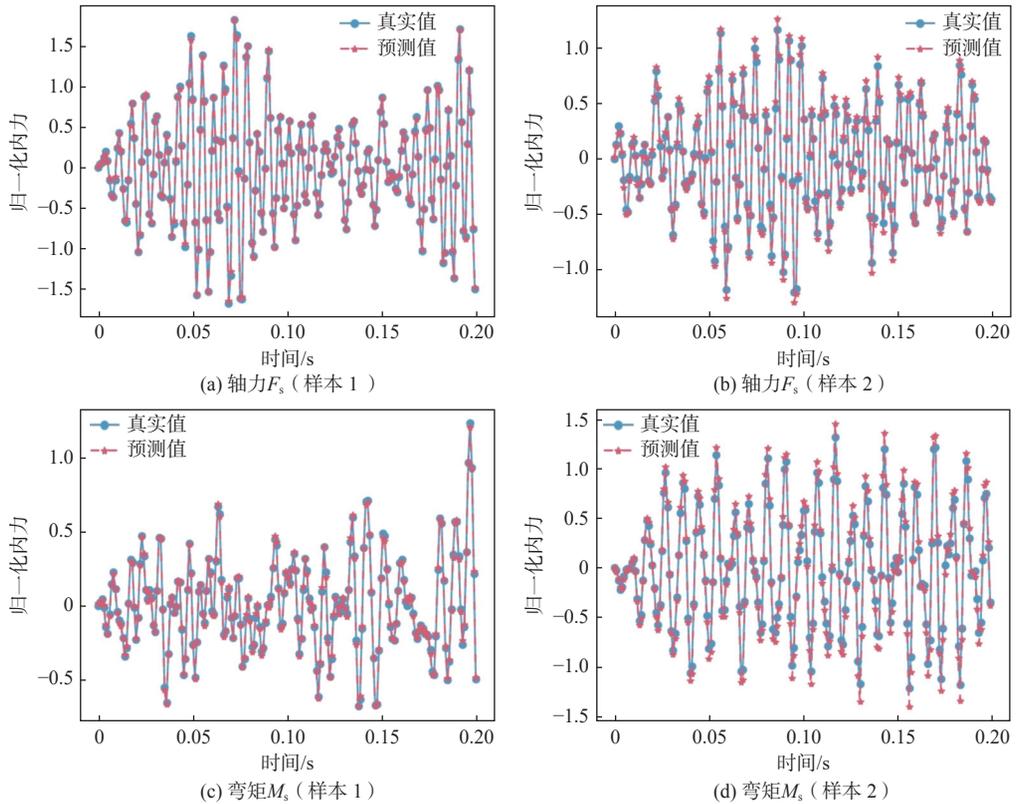


图7 模型预测结果

Fig. 7 Prediction results of model

为了验证本文所提模型对动态载荷的预测效果,与应用广泛的多层感知器和随机森林进行了精度比较。为了验证将结构参数融合于动态载荷代理模型的有效性,还构建了无结构参数输入的LSTM模型,以及其他文献中的动态载荷代理模型,包括文献[21]中的deep RNN和文献[22]中的ED-CRNN,并进行了对比。对比方法的结构如下:

1) 对于多层感知器,由于其只能接收一维输入并进行一维输出,因此将外载荷二维时间序列展开成一维特征并与结构参数合并作为输入,将内力响应时间序列展开为一维作为输出。除输入层和输出层外,该模型还包含3个隐藏层,每个隐藏层的单元为512个。其训练参数与本文所提模型一致。

2) 对于随机森林,由于其与多层感知器一样只能接收一维输入输出,因此其输入输出与多层感知器一致。结构上,其最大深度为10,集成决策树为100个。

3) 对于LSTM,仅以外载荷作为输入,以内力响应作为输出。结构上,相比本文所提模型,去掉了针对结构参数的线性层与复制,并且在编码器和解码器之间没有特征组合的步骤。其他参数均与本文所提模型一致。

4) 对于deep RNN^[21]和ED-CRNN^[22],由于其只适用于时间序列输入输出,无法考虑结构参数输

入,因此输入输出与LSTM一致。出于对超参数寻优结果和控制模型参数与本文基本一致的考虑,这2种模型的隐藏单元数均设置为32,其他参数的设置与原文献中相同。

每种方法在测试集上的误差分布如图8所示。结果表明,本文所提方法相比另外5种方法在准确性和稳定性上均有较大的优势。

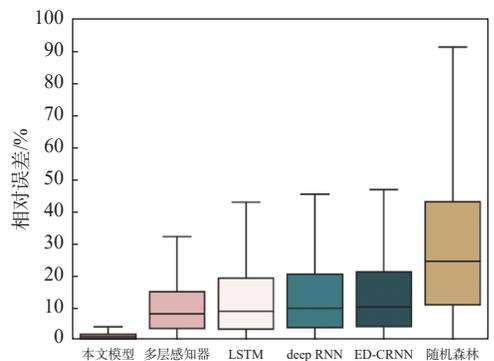


图8 不同方法相对误差

Fig. 8 Relative error of different methods

首先,对于多层感知器和随机森林,由于没有针对动态载荷问题所涉及的时序、非时序混合输入进行特别的设计,因此,结构参数与外载荷共同结合为一维特征,在这种设计下,外载荷由于时序长度较大,所占的特征数远远多于结构参数,导致结构参数特征在模型中的重要性很难被捕获。此外,从模型结构方面分析,多层感知器

对所有输入做了全连接, 外载荷时间序列的每个时间步都有对应的拟合参数, 导致网络整体参数量过大, 出现了比较严重的过拟合现象, 其在训练集上的相对误差能达到 5%, 但在测试集上的相对误差却只有 10% 左右。随机森林则是其每个单独的决策树都对输入参数进行了特征选择, 因此外载荷的时序信息会有所丢失, 无法给出精确的预测结果。

其次, 对于 LSTM, deep RNN 和 ED-CRNN, 尽管其模型能够较好地提取动态外载荷的时序特征, 但因为缺少结构参数的信息, 模型的内力响应无法被确定, 因此精度不高。这也说明将外载荷与结构参数的异构数据融合是有必要的, 本文提出的融合模型有效。

相比其他方法, 本文提出的模型是使用 LSTM 编码网络处理外载荷时序输入进行特征提取, 然后再将其与经过线性变化和复制的结构参数进行结合。这种方式使得 2 种输入的重要性在模型中是相似的。此外, 同一个 LSTM 层中每个单元的参数是共享的, 这就意味着对于每个时间步的输入, LSTM 单元都以相同的方式进行处理, 这种先验知识启发的结构设计使得其参数量小于多层感知器。在这 2 个因素的共同作用下, 模型能对本文所提问题进行较高精度的预测。

2.3 应用案例——结构参数相关性分析

相关性分析是代理模型的优势应用场景之一。在相关性分析中, 需改变参数并进行大量算例的计算, 然后判断参数对结果的影响。对有限元方法而言, 计算大量不同的算例非常耗时, 因此可以通过引入代理模型来提高计算效率。本节将运用本文构建的代理模型进行结构参数相关性分析, 以展示其快速计算能力。首先, 固定外载荷、改变结构参数, 然后利用代理模型对 2000 个算例进行计算, 最后, 根据计算结果画出了如图 9 所示的皮尔逊相关系数矩阵。图中, F 为轴力, M 为弯矩, 下标为截面编号。每 2 个变量 X , Y 之间的皮尔逊相关系数计算公式如下:

$$\text{corrcoef}(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)\text{var}(Y)}} \quad (10)$$

式中: $\text{cov}()$ 为协方差; $\text{var}()$ 为方差。皮尔逊相关系数的绝对值越接近于 1, 说明两个变量之间的相关性越强, 而正、负号则代表正相关或是负相关。值得一提的是, 皮尔逊相关系数只能在 2 个一维变量之间计算, 因此针对内力响应这一时序变量, 本文将采用其在时间维度上的标准差予以替代。由于所有内力响应在时间维度的均值都为 0,

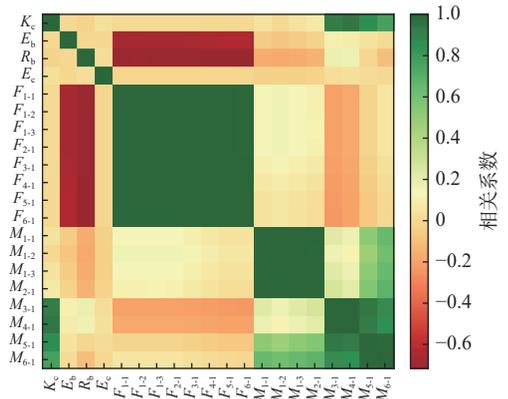


图9 皮尔逊相关系数矩阵

Fig. 9 Pearson correlation coefficient matrix

故标准差的大小即代表了内力响应的强度。

通过相关性矩阵可以得出, 不同结构参数之间相互独立, 这也表明了算例生成的正确性、截面轴力 F_s 与梁的弹性模量 E_b 和外径 R_b 具有较强的负相关性、截面弯矩 M_s 与连接刚度 K_c 具有较强的正相关性, 以及连接段弹性模量与内力响应没有明显的相关性。这主要是因为结构参数的改变影响了梁模型的固有频率。由于所施加的外载荷是固定的, 并不具备功率谱密度为常数的统计特性, 因此其在特定的频率下幅值较高, 当梁模型的固有频率与这些频率接近时, 容易引起这些频率上的共振, 从而使得轴力和弯矩的时间尺度标准差增大。在本案例的外载荷条件下, 相关性分析结果表明, 梁的弹性模量 E_b 和外径 R_b 越小、连接刚度 K_c 越大, 越容易引发共振。

另外, 值得注意的是, 代理模型超参数选择结果表明, 对结构参数进行特征提取的线性层只需提取出 3 个特征, 就能给出较好的预测, 这与相关性分析的结果是一致的, 也即代理模型在结构参数线性映射时将连接段弹性模量这一无关参数过滤, 完成了信息的压缩。

在 NVIDIA GeForce RTX 3070 8GB 上, 代理模型计算 2000 个算例的时间为 2.0 ms, 远少于有限元计算的时间, 这充分体现了代理模型相对于传统数值计算方法的优点, 即用极短的时间满足快速迭代计算的需求, 这在结构优化、敏感性分析、相关性分析等场景有着较强的应用潜力。

需要说明的是, 本节相关性分析的结果仅适用于所选定的外载荷条件, 因此在力学上的普适性较弱, 仅能作为应用案例用于验证代理模型在快速迭代计算上的效率优势。

3 结论

本文基于深度学习方法构建了一种动态载荷代理模型, 该模型利用编码器-解码器结构和

LSTM网络,通过融合动态外载荷时序输入和结构参数非时序输入,较好地预测了梁结构多个截面的内力响应时间历程。该模型在由10 000个样本构成的有限元仿真数据集上进行了精度评估、方法对比和相关性计算,主要得到如下结论:

1) 本文所提动态载荷代理模型拥有端到端的建模能力,能同时接收外载荷时序输入和结构参数非时序输入,并以时间序列的方式预测内力响应时间历程,可为载荷预测计算提供更多信息。

2) 所提的动态载荷代理模型在测试集上能达到较高的精度,对于所有截面内力响应其平均预测精度可达98%,整体平均相对误差为1.5%,高于其他代理模型方法,体现了该代理模型在处理动态载荷问题和异构数据上的优势。

3) 该代理模型可以用于替代有限元方法进行大量算例的快速计算。结构参数相关性分析案例表明,该模型的计算速度和有限元方法相比更快,能够快速计算大量不同参数条件的算例,并给出载荷预测结果,这对于需要快速迭代计算的应用场景有着重大意义。

参考文献:

- [1] 刘婧,王德禹.基于SMOTE算法和动态代理模型的船舶结构可靠性优化[J].*中国舰船研究*,2020,15(5):114-123.
LIU J, WANG D Y. Reliability-based design optimization of ship structure using SMOTE algorithm and dynamic surrogate model[J]. *Chinese Journal of Ship Research*, 2020, 15(5): 114-123 (in Chinese).
- [2] 韩忠华,许晨舟,乔建领,等.基于代理模型的高效全局气动优化设计方法研究进展[J].*航空学报*,2020,41(5):623344.
HAN Z H, XU C Z, QIAO J L, et al. Recent progress of efficient global aerodynamic shape optimization using surrogate-based approach[J]. *Acta Aeronautica et Astronautica Sinica*, 2020, 41(5): 623344 (in Chinese).
- [3] KUDELA J, MATOUSEK R. Recent advances and applications of surrogate models for finite element method computations: a review[J]. *Soft Computing*, 2022, 26(24): 13709-13733.
- [4] 罗文俊,王德禹.基于兴趣子域动态代理模型的船舶结构可靠性优化[J].*中国舰船研究*,2021,16(4):96-107.
LUO W J, WANG D Y. Reliability-based optimization of ship structure based on interest subdomain dynamic surrogate model[J]. *Chinese Journal of Ship Research*, 2021, 16(4): 96-107 (in both Chinese and English).
- [5] FREENY A. Empirical model building and response surfaces[J]. *Technometrics*, 1987, 30(2): 229-231.
- [6] KRIGE D G. A statistical approach to some basic mine valuation problems on the Witwatersrand[J]. *Journal of the South African Institute of Mining and Metallurgy*, 1951, 52(6): 201-203.
- [7] BROOMHEAD D S, LOWE D. Multivariable functional interpolation and adaptive networks[J]. *Complex Systems*, 1988, 2(3): 321-355.
- [8] FAN X X, WANG P F, HAO F F. Reliability-based design optimization of crane bridges using Kriging-based surrogate models[J]. *Structural and Multidisciplinary Optimization*, 2019, 59(3): 993-1005.
- [9] CHU L, SHI J J, DE CURSI E S. Kriging surrogate model for resonance frequency analysis of dental implants by a Latin hypercube-based finite element method[J]. *Applied Bionics and Biomechanics*, 2019, 2019: 3768695.
- [10] HASSAN A K S O, ETMAN A S, SOLIMAN E A. Optimization of a novel Nano antenna with two radiation modes using Kriging surrogate models[J]. *IEEE Photonics Journal*, 2018, 10(4): 4800807.
- [11] HAERI A, FADAEI M J. Efficient reliability analysis of laminated composites using advanced Kriging surrogate model[J]. *Composite Structures*, 2016, 149: 26-32.
- [12] SHI J J, CHU L, BRAUN R. A Kriging surrogate model for uncertainty analysis of graphene based on a finite element method[J]. *International Journal of Molecular Sciences*, 2019, 20(9): 2355.
- [13] LADICKÝ L, JEONG S, SOLENTHALER B, et al. Data-driven fluid simulations using regression forests[J]. *ACM Transactions on Graphics*, 2015, 34(6): 199.
- [14] 李海泉,陈小前,左林玄,等.基于随机森林的飞行载荷代理模型分析方法[J].*航空学报*,2022,43(3):225640.
LI H Q, CHEN X Q, ZUO L X, et al. Surrogate model for flight load analysis based on random forest[J]. *Acta Aeronautica et Astronautica Sinica*, 2022, 43(3): 225640 (in Chinese).
- [15] WANG Q, QIAN W Q, HE K F. Unsteady aerodynamic modeling at high angles of attack using support vector machines[J]. *Chinese Journal of Aeronautics*, 2015, 28(3): 659-668.
- [16] AL KAJBAF A, BENSI M. Application of surrogate models in estimation of storm surge: a comparative assessment[J]. *Applied Soft Computing*, 2020, 91: 106184.
- [17] BAI T, TAHMASEBI P. Characterization of groundwater contamination: a transformer-based deep learning model[J]. *Advances in Water Resources*, 2022, 164: 104217.
- [18] ZHAO X Y, GONG Z Q, ZHANG J, et al. A surrogate model with data augmentation and deep transfer learning for temperature field prediction of heat source layout[J]. *Structural and Multidisciplinary Optimization*, 2021, 64(4): 2287-2306.
- [19] SHIBATA R, OHIRA M, MA Z W. A novel convolutional-autoencoder based surrogate model for fast S-parameter calculation of planar BPFs[C]//2022 IEEE/MTT-S International Microwave Symposium - IMS 2022. Denver: IEEE, 2022: 498-501.
- [20] ZHOU J M, DONG L L, GUAN W, et al. Impact load identification of nonlinear structures using deep recurrent neural network[J]. *Mechanical Systems and Signal Processing*, 2019, 133: 106292.
- [21] CHEN T, GUO L, DUAN A D Z, et al. A feature learning-based method for impact load reconstruction and localization of the plate-rib assembled structure[J]. *Structural Health Monitoring*, 2022, 21(4): 1590-1607.
- [22] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. *Neural Computation*, 1997, 9(8): 1735-1780.