郭茂祖,崔正,赵玲玲,等.基于多尺度混合Kolmogorov-Arnold网络的结构地震响应预测[J].地震工程学报,2026,48(6):1259-1272.DOI:10.20000/j.1000-0844.20241220003
Guo Maozu, Cui Zheng, Zhao Lingling, et al. Seismic structural response prediction using multiscale hybrid Kolmogorov-Arnold networks[J].China Earthquake Engineering Journal,2026,48(6):1259-1272.DOI:10.20000/j.1000-0844.20241220003

基于多尺度混合Kolmogorov-Arnold网络的结构地震响应预测

郭茂祖[1,2], 崔　正[1,2], 赵玲玲[3], 张庆宇[1,2]
(1.北京建筑大学,北京 100044;
2.建筑大数据智能处理方法研究北京市重点实验室,北京 100044;
3.哈尔滨工业大学,黑龙江 哈尔滨 150001)

摘要:准确预测结构在地震作用下的响应对结构损伤和性能评价具有重要意义。针对现有结构地震响应预测方法对长时间序列数据预测能力、预测效率不足的问题,同时综合考虑地震数据稀缺的特点,结合Kolmogorov-Arnold神经网络与门控循环单元搭建耦合网络模型,配套提出多尺度混合方法,能够在地震样本稀缺场景下快速完成建筑结构多自由度响应时程实时预测。为验证所提出方法的准确性和效率,进行三个公开数据集上的案例研究。此外,还通过消融实验和对比实验进一步研究该方案的可行性。实验结果表明,该方法能够准确预测结构中多自由度对应的加速度、速度和位移时程,预测精度优于长短期记忆网络(LSTM)、卷积神经网络-长短期记忆网络(CNN-LSTM)等现有地震响应预测模型,可以在训练数据非常稀缺的情况下实现高效、准确的预测,从而达到工程应用需求。

关键词:结构地震响应预测;GRU;Kolmogorov-Arnold网络;多尺度混合;小样本学习
中图分类号:TU311.3; TU17　　**文献标志码**:A　　**文章编号**:1000-0844(2026)06-1259-14
DOI:10.20000/j.1000-0844.20241220003

Seismic structural response prediction using multiscale hybrid Kolmogorov-Arnold networks

Guo Maozu[1,2], Cui Zheng[1,2], Zhao Lingling[3], Zhang Qingyu[1,2]
(*1.Beijing University of Civil Engineering and Architecture, Beijing 100044, China*;
2.Beijing Key Laboratory of Intelligent Processing for Building Big Data, Beijing 100044, China;
3.Harbin Institute of Technology, Harbin 150001, Heilongjiang, China)

Abstract: Accurate prediction of structural dynamic responses under seismic loads is crucial for damage assessment and performance evaluation. Existing methods for predicting long-sequence time-series data face challenges such as insufficient predictive efficiency and an inherent scarcity of seismic

收稿日期:2024-12-20
基金项目:国家自然科学基金项目(62271036)
第一作者简介:郭茂祖(1966-),男,博士,教授,主要研究方向为机器学习、智能建造。E-mail:guomaozu@bucea.edu.com。

data. To overcome these limitations, this paper presents a neural network model that combines the Kolmogorov-Arnold network with gated recurrent units. Additionally, a multiscale hybrid approach is proposed to enable real-time prediction of multi-degree-of-freedom response time histories in structures with limited seismic data. The accuracy and efficiency of the proposed method are validated through case studies on three publicly available datasets. Furthermore, ablation experiments and comparative studies are conducted to evaluate the feasibility of the approach. The experimental results show that the proposed method accurately predicts acceleration, velocity, and displacement time histories for multiple degrees of freedom within the structure. It has higher predictive accuracy than existing models such as long short-term memory (LSTM) and convolutional neural network-LSTM. Moreover, it achieves efficient and accurate predictions even with severely limited training data, thereby meeting the requirements of practical engineering applications.

Keywords: seismic structural response prediction; gated recurrent units (GRU); Kolmogorov-Arnold network; multiscale hybrid; small-sample learning

0 引言

地震作为全球高发的自然灾害,严重威胁人类生命与财产安全。建筑结构在地震波激励下会产生显著的动力变形,强震引发的塑性累积变形极易造成结构损伤、局部破坏甚至整体倒塌,产生严重的灾害后果。因此,建筑结构地震响应的精准分析与预测是结构工程与抗震防灾领域的重要研究课题[1]。结构地震响应预测结果可支撑结构损伤评估[2]、结构健康监测[3]及抗震性能优化设计等工程应用,为提升建筑抗灾能力提供数据支撑,同时也能为震后灾害评估、建筑修复与应急救援提供重要指导依据。

传统研究多依托结构力学机理与物理建模开展非线性动力分析,通过有限元数值模拟实现结构响应预测[4-5]。有限元法(Finite Element Method, FEM)作为成熟的数值分析手段,已广泛应用于结构抗震计算与工程分析[6-8]。但该方法仍存在明显短板:复杂高层建筑的精细化建模流程复杂、计算代价极高,难以兼顾高精度与实时性,无法满足快速预测的工程需求。随着深度学习技术与算力水平的快速提升,数据驱动方法已逐步应用于结构设计优化[9-10]、结构健康监测[11-14]、结构系统识别[15-16]、材料性能模拟[17-19]及结构振动控制[20-22]等多个研究方向。在地震响应预测领域,深度学习可摆脱复杂物理建模与迭代计算的约束,众多研究已验证数据驱动方法在结构响应预测任务中的有效性与可行性。

近年来,基于深度学习的结构地震时程预测研究得到快速发展。多层感知器(Multilayer Perceptron, MLP)结构简单、搭建便捷,被初步用于建筑地震响应预测任务[23-24],但其静态拟合特性难以充分表征结构地震响应的复杂非线性动态演化规律。卷积神经网络(Convolution Neural Network, CNN)具备优异的空间特征提取能力,可实现单自由度与多自由度体系的动力响应预测[25-30],但在结构进入大塑性变形阶段后,CNN对非线性时序特征的捕捉能力显著下降,预测精度大幅衰减。相比而言,循环神经网络(Recurrent Neural Network, RNN)更适配时序数据建模,能够有效挖掘地震序列的时间依赖关系,在结构响应预测中得到广泛应用[31]。长短期记忆网络(Long Short-Term Memory, LSTM)[32]与门控循环单元(Gated Recurrent Unit, GRU)[33]作为RNN的经典改进结构,通过门控记忆机制缓解长序列梯度消失问题,适配地震长时序预测场景,各类改进策略也进一步提升了模型的综合性能[34-38]。与此同时,CNN与LSTM的跨结构融合方案、基于自注意力机制的全局特征增强方法,也被逐步引入结构预测任务,有效提升了模型的特征表征与端到端映射能力[39-44]。

现有深度学习预测方法仍普遍面临训练样本稀缺的瓶颈。真实结构地震响应数据采集难度大、有效样本数量有限,而依托有限元仿真批量构建数据集需耗费海量算力与时间成本,导致模型难以充分发挥大数据驱动的学习优势。针对该问题,本文提出多尺度GKAN-Mixer地震结构动力响应预测

框架。该模型依托多尺度混合策略对地震时序波动特征进行分解与重组，实现隐性数据增强，有效扩充特征空间、挖掘有限样本中的深层有效信息，显著降低模型对大规模标注数据的依赖，提升小样本场景下的学习效率与泛化性能。同时，多尺度融合机制能够同步捕获地震动序列的局部细节特征与全局演化规律，进一步提升复杂地震激励下的响应预测精度。

此外，本文通过融合改进 Kolmogorov-Arnold 网络（Kolmogorov-Arnold Network，KAN）与 GRU 网络，进一步提升模型的非线性拟合与时序预测能力。现有主流深度学习模型普遍采用 MLP 作为输出映射层，MLP 凭借通用性强、落地便捷的优势成为非线性拟合的基础模型[45]，但存在参数冗余度高、激活函数调试成本大、复杂映射拟合能力有限等固有缺陷[46]。基于 Kolmogorov-Arnold 表示定理构建的 KAN 网络为替代传统 MLP 提供了新的思路[47]。不同于 MLP 的权重乘法与固定激活机制，KAN 将可学习的单变量样条激活函数部署于网络端，通过训练自适应优化非线性映射关系。本文针对地震时序预测任务对 KAN 进行改进，并与擅长长时序建模的 GRU 深度融合，构建的 GKAN 模块兼具时序记忆能力与高精度非线性拟合优势，有效提升了模型的预测精度与鲁棒性。

本文的主要研究贡献总结如下：

（1）针对地震响应数据稀缺问题，引入多尺度混合数据增强策略，通过多尺度时序分解与特征重组充分挖掘有限样本信息，降低了模型对大规模训练数据的依赖，显著提升了小样本条件下预测模型的鲁棒性。

（2）构建 GRU-KAN 融合预测模块，结合 GRU 的长时序依赖捕捉能力与 KAN 的高精度非线性拟合优势，可精准建模高层建筑在复杂地震激励下的非线性动力时程响应。所提模型可适配多方向复杂地震波输入，同步实现建筑多自由度位移、速度、加速度响应的高精度预测，可满足工程实时预测需求。

（3）基于公开地震响应数据集开展多组定量对比实验，与多种主流先进预测方法进行多维度指标对比，实验结果充分验证了本文所提 GKAN-Mixer 模型在结构地震响应预测任务中的有效性。

1　结构地震响应预测方法

1.1　地震作用下结构响应预测的全过程

为实现地震作用下结构响应的精细化、全面化时程分析，需依托结构有限元模型构建训练数据集。本文整体研究流程如下：首先建立结构有限元数值模型，选取地震数据库中的实测地震波作为时程分析输入，并依据结构抗震设防烈度对地震波幅值进行调幅处理。其次，通过弹塑性有限元计算获取结构各节点的加速度、速度、位移响应时程，亦可直接采用公开实测数据集中的结构响应时程样本。随后对节点时程数据进行预处理，构建深度学习模型所需的输入样本与对应真值标签。在此基础上完成模型训练，通过反向传播算法迭代更新网络权重参数，最终得到优化后的预测模型。

训练完成的模型可在未知地震激励下，对结构部分或全部节点的动力响应时程进行高精度实时预测。本文整体技术流程如图 1 所示。

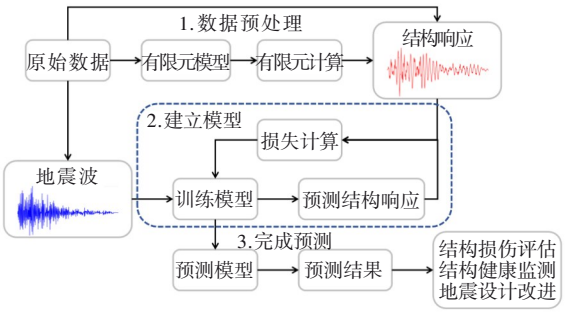


图 1　结构地震响应预测总体流程

Fig. 1　Workflow of structural seismic response prediction

由于结构弹塑性有限元数值计算耗时严重、算力成本高，难以批量生成大规模训练样本，导致深度学习模型面临数据稀缺问题。针对该问题，本文引入多尺度分解策略开展数据增强，充分挖掘有限样本的特征信息，实现小样本条件下的高精度、高效预测。

1.2　GKAN-Mixer 模型的结构

给定具有一个或多个观测变量的序列 X，X 可以表示为：

$$X=\{x_1,x_2,\ldots,x_i,\ldots,x_n\} \tag{1}$$

式中：每个 x_i 表示在时间点 $t(t=1,2,\ldots,n)_i$ 观测到的地面加速度值；n 是观测数据的总数。具体来说，x_i 可以是地震波在某一时刻的加速度测量值。

结构地震响应预测的核心目标为：基于地震动

观测输入序列 X，求解最优结构动力响应时序 Y（包含建筑各楼层的层间位移、速度、加速度），Y 的表达式如下：

$$Y=\{y_1,y_2,\ldots,y_j,\ldots,y_m\} \tag{2}$$

式中：每个 y_j 表示结构在不同位置的响应值；m 是结构响应数据的总数。具体来说，y_j 可以表示某一层的位移、速度或加速度的响应值。

为了更具体，可以进一步细化结构响应：

$$Y=\{Y_1,Y_2,\ldots,Y_i,\ldots,Y_n\} \tag{3}$$

式中：每个 Y_i 包括第 i 层（$i=1,2,\ldots,n$）的位移、速度和加速度：

$$Y_i=\{y_i^{\mathrm{disp}},y_i^{\mathrm{vel}},y_i^{\mathrm{acc}}\} \tag{4}$$

式中：y_i^{disp} 为位移响应；y_i^{vel} 为速度响应；y_i^{acc} 为加速度响应。

综上，结构地震响应精准预测的核心难点在于刻画地震动复杂的时变演化特征。为此，本文提出多尺度融合 GKAN-Mixer 模型，充分发挥多尺度时序分解分量的特征互补优势以提升预测效果。从模型架构来看，GKAN-Mixer 以多尺度混合模块为核心，依靠多尺度分解完成地震振动特征提取，再通过多尺度聚合完成多尺度预测信息融合；GRU 与改进 KAN 构成基础预测单元，用于建立地震动输入与结构动力响应间的非线性映射关系。GKAN-Mixer 完整模型架构如图 2 所示。

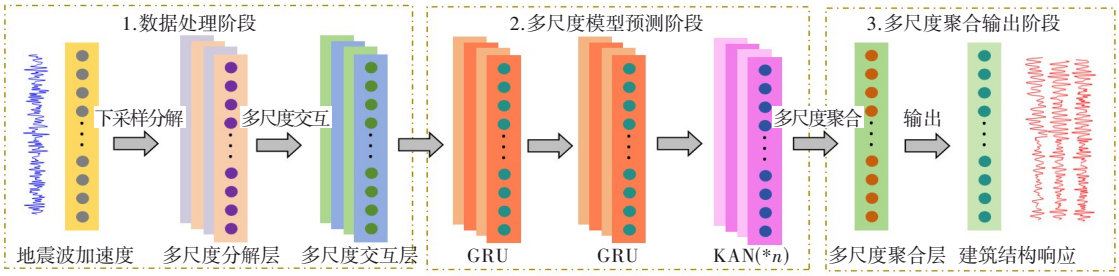


图 2 GKAN-Mixer 模型架构

Fig. 2 Architecture of the GKAN-Mixer model

1.2.1 多尺度混合方法

不同尺度时序数据具备差异化演化特征：细尺度分量可刻画振动局部细节变化，粗尺度分量能够反映地震动整体宏观演化规律[48]。针对时序预测任务，各尺度分量主导的时间演化模式存在明显区别，单独尺度分量的预测性能也各不相同[49]。

为了剖析地震波序列的复杂变化，首先将地震波的观测值 $X\in\mathbb{R}$ 下的采样，分为 L 个尺度，最终得到一组多尺度时间序列 $X'=\{X_1,\ldots,X_L\}$，其中，最低层次的输入序列 X_1 包含了最细微的时间变化；最高层次的序列 X_m 代表了宏观周期变化。然后，通过嵌入层将这些多尺度序列投影为高维特征 X^*，可以形式化为 $\hat{X}=Mixer(X')$。通过上述设计，我们得到了输入序列的多尺度表示。

季节时序分析相关研究表明[50]，长周期信号可拆解为若干短周期分量的叠加，局部细节信息对季节演化趋势的精准预测起到关键作用。基于该理论，本文采用自下而上的建模思路，从底层细尺度时序逐步向上传递特征，为粗尺度周期建模补充局部细节信息。具体实现上，针对任意多尺度时序，本文搭建层数为 L 的自下而上混合交互层 Bottom-Up-Mixing，以残差结构完成不同尺度周期特征的双向交互，该过程的表述列于式（5）：

$$X_i=X_i+\mathrm{Bottom\text{-}Up\text{-}Mixing}(X_{i-1}),\quad i=2,\ldots,L \tag{5}$$

式中：Bottom-Up-Mixing 模块可通过沿时间维度布置带中间激活函数的线性层或 KAN 层实现，模块结构如图 3 所示。

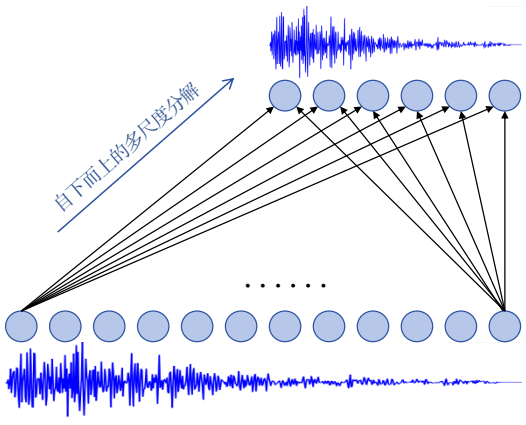


图 3 自下而上的多尺度分解

Fig. 3 Bottom-up multiscale decomposition

1.2.2　GRU 网络

本文选取 GRU 作为模型主干时序网络，核心原因在于其优越的计算效率。GRU 是 LSTM 的轻量化改进版本，在保留 LSTM 优异长序列预测性能的前提下，将原有的三门控结构简化合并为更新门 r_τ 与重置门 z_τ，有效降低网络复杂度与参数量。该结构可加快模型收敛速度、降低训练算力开销，更适配地震长时序响应预测场景。GRU 神经元内部结构如图 4 所示。

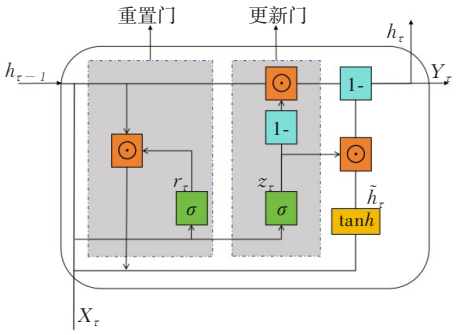


图 4　GRU 网络示意图

Fig. 4　Schematic diagram of the GRU network

更新门用于控制结构中信息的遗忘和更新，重置门用于控制保留当前状态中用于计算下一状态的信息。τ 时刻时更新门和重置门的计算公式如下：

$$z_\tau=\sigma(W^z X_\tau+U^z h_{\tau-1}) \tag{6}$$

$$r_\tau=\sigma(W^r X_\tau+U^r h_{\tau-1}) \tag{7}$$

式中：X_τ 为记忆神经元 τ 时刻时的输入；$h_{\tau-1}$ 为记忆神经元在 $\tau-1$ 时刻的状态；σ 为 sigmoid 函数；W^z 和 U^z、W^r 和 U^r 分别为更新门、重置门的权重系数。图 4 中：$\tilde{h}_\tau$ 为时刻细胞记忆；$h_{\tau-1}$ 为记忆神经元 τ 时刻的输出。表达式为：

$$\tilde{h}_\tau=\tanh(r_\tau\times U^h h_{\tau-1}+W^h X_\tau) \tag{8}$$

$$h_\tau=(1-z_\tau)\times\tilde{h}_\tau+z_\tau\times h_{\tau-1} \tag{9}$$

式中：W^h、U^h 为隐藏单元的权重系数；$\tanh(\theta)$ 为模拟门的功能及归一化函数。

1.2.3　Kolmogorov-Arnold 网络

KAN 的构建以 Kolmogorov-Arnold 表示定理为理论基础，该定理指出任意多元连续函数均可由有限个单变量连续函数与二元加法运算复合而成，即复杂多元映射可拆解为若干单变量函数叠加。但经典定理中的一维基函数通常具备非光滑特性，难以通过常规神经网络拟合学习，致使该定理长期难以落地于机器学习研究。而 KAN 网络相关理论成果有效解决了上述难题，挖掘出该定理在神经网络建模中的应用潜力[51]。样条函数是由若干控制点与节点约束生成的光滑连续曲线，多用于离散数据插值与拟合逼近。样条由阶数 k 界定，$k=3$ 为工程与建模常用取值，代表控制点间拟合所采用多项式的次数；参数 G 为区间数量，对应相邻控制点之间的分段子区间总数。开展样条插值时，各分段多项式相互衔接形成光滑曲线，整体共包含 $G+1$ 个节点。KAN 通过引入可学习单变量函数完成多元映射建模，仅需为各维度匹配合适的单变量拟合函数（$\varphi_{i,j}$ 和 Φ_i），即可实现对目标多元关系的逼近。

本文选用 B 样条对一维映射函数 $\varphi_{i,j}$ 进行拟合，B 样条属于特殊分段样条，依托 B 样条基函数完成曲线构造；这类基函数具备局部支撑特性，能够保证各控制点之间曲线过渡光滑连续。

将复杂高维映射拆解为若干简易子函数，能够提升模型对高维非线性关系的学习与逼近效率。KAN 的核心设计思路是将多元函数拆解为多个简单函数叠加的形式，该结构更适配神经网络的训练学习逻辑。具体而言，KAN 内的映射函数 $f(x)$ 表达式为：

$$f(x)=\sum_{i=1}^{2n+1}\Phi_i\left(\sum_{j=1}^{n}\varphi_{i,j}(x_j)\right) \tag{10}$$

式中：Φ_i 为定义域 $\mathbb{R}$ 映射至 $\mathbb{R}$ 的外函数；$\varphi_{i,j}$ 为定义域 $[0,1]$ 映射至 $\mathbb{R}$ 的内函数。

模型中全部待学习映射均为单变量函数，因此可采用 B 样条曲线对各一维函数进行参数化表征，局部 B 样条基函数对应的系数作为网络可训练参数。

传统 MLP 单层由线性变换与后置非线性激活构成，依靠堆叠网络层数提升拟合能力；而 KAN 的数学定义形式如下：

$$\Phi=\{\varphi_{i,j}\},\ p=1,2,\ldots,n_{\text{in}};\ q=1,2,\ldots,n_{\text{out}} \tag{11}$$

式中：n_{in} 和 n_{out} 均是可学习的参数化函数。

依据 Kolmogorov-Arnold 表示定理，定理中的内函数可构成一层输入维度 $n_{\text{in}}=n$、输出维度 $n_{\text{out}}=2n+1$ 的 KAN 网络层。式（10）给出的正是两层 KAN 层的基础组合形式。与 MLP 相同，KAN 支持多层堆叠构建深层网络，其网络结构维度可记作 $[n_0,n_1,\ldots,n_L]$，其中，n_L 为输入首层的神

经元数量。

定义第 l 层第 j 个神经元为 (l,j)，该神经元对应的激活输出记为 $x_{l,j}$；第 l 层与第 $l+1$ 层之间共包含 $l(l+1)$ 个单变量激活函数，用于建立两层神经元间的映射关系，连接神经元 (l,j) 和 $(l+1,j)$ 的激活函数表达式如下：

$$\varphi_{l,i,j},\ l=0,\ldots,L-1,\ i=1,\ldots,n_l,\ [\ j=1,\ldots,nl+1\ ] \tag{12}$$

将 $\varphi_{l,i,j}$ 的预激活值记为 $x_{l,i}$，激活 $\varphi_{l,i,j}$ 后变为 $\tilde{x}_{l,j,i}$。第 $(l+1,j)$ 神经元的激活值为所有的激活值的总和：

$$x_{l+1,j}=\sum_{i=1}^{n_l}\tilde{x}_{l,j,i}=\sum_{i=1}^{n_l}\varphi_{l,i,j}(x_{l,j}),j=1,\ldots,n_{l+1} \tag{13}$$

将其改写为矩阵将会得到：

$$\boldsymbol{x}_{l+1}=\underbrace{\begin{pmatrix}\varphi_{l,1,1}(\bullet) & \varphi_{l,1,2}(\bullet) & \cdots & \varphi_{l,1,n_l}(\bullet)\\ \varphi_{l,2,1}(\bullet) & \varphi_{l,2,2}(\bullet) & \cdots & \varphi_{l,2,n_l}(\bullet)\\ \vdots & \vdots & \ddots & \vdots\\ \varphi_{l,n_{l+1},1}(\bullet) & \varphi_{l,n_{l+1},2}(\bullet) & \cdots & \varphi_{l,n_{l+1},n_l}(\bullet)\end{pmatrix}}_{\boldsymbol{\Phi}_l}\boldsymbol{x}_l \tag{14}$$

式中：$\boldsymbol{\Phi}_l$ 是与第 l 个层 KAN 对应的函数矩阵。那么，一个通用的 L 层 KAN 就可以表示为：

$$KAN(x)=(\boldsymbol{\Phi}_{L-1}\circ\boldsymbol{\Phi}_{L-2}\circ\boldsymbol{\Phi}_{L-2}\circ\cdots\circ\boldsymbol{\Phi}_1\circ\boldsymbol{\Phi}_0)x \tag{15}$$

图 5 为任意数层的 KAN 网络的示意图。

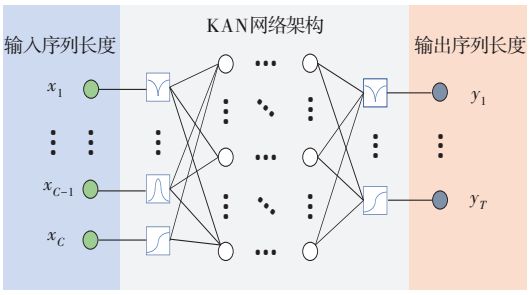


图 5 KAN 网络示意图

Fig. 5 Schematic diagram of the KAN network

对于带输入、输出的网络层，原始 KAN 需要先扩展输入张量的形状，才能执行激活运算；而其所有激活函数均由一组固定基函数线性组合得到，这类基函数运算会让 KAN 计算耗时达到 MLP 的 10 倍以上。本研究重构计算流程：先使用不同基函数对输入做激活处理，再开展线性组合。该计算方案可大幅降低内存开销，将运算转化为常规矩阵乘法，同时能够顺畅完成前向、反向传播。

该优化方案具备替代传统基于 MLP 架构深度学习模型的潜力，因此本文将 KAN 与 GRU 进行轻量化融合，把 KAN 作为建筑结构地震响应预测框架的核心模块之一。具体实现逻辑为：将 GRU 输出的时序隐特征作为 KAN 的输入，经由两层 KAN 网络完成特征降维与非线性映射，得到匹配任务需求维度的预测结果；最后将各尺度分支的输出送入多尺度聚合层完成信息融合。

1.2.4 多尺度聚合层

经中间预测层运算后，输出多尺度预测特征为 $T=\{T_1,\ldots,T_L\}$。不同尺度时序表征着不同层级的振动演化规律，各尺度分量的预测性能存在明显差异。为充分挖掘多尺度特征的互补优势，需对各尺度预测结果实施融合聚合，多尺度聚合模块的数学表达式定义如下：

$$Y=\sum_{i=1}^{L}\mathrm{Predictor}_i(T_i),i\in\{1,\ldots,L\} \tag{16}$$

式中：T_i 表示第 i 个尺度序列对结构响应的预测分量，最终输出为 Y；$\mathrm{Predictor}_i$ 表示第 i 个尺度序列的预测器，它将第 i 个尺度序列长度从 $L/2i$ 回归投影到序列原始长度 L，然后将投影后基于不同尺度的预测分量 $\mathrm{Predictor}_i$ 进行线性叠加，得到最终输出 Y，使多尺度方法具有互补预测能力。

2 实验与讨论

本节设置三组复杂度逐级递增的算例，分别在有限元数值模型与真实工程结构场景下，验证本文所提方法对结构地震响应时程的预测效果。同时依托算例将刚架模型扩充数据集，测试 GKAN-Mixer 在多点地震动输入工况下的多目标联合预测性能；并通过消融实验，定量验证模型内 KAN 模块对整体预测精度的增益作用。

2.1 数据集生成

为探究地震动激励下结构动力响应的建模与预测方法，本文整理构建 3 组数值模拟数据集，并采集工程强震动数据中心地震数据库中建筑实测传感器记录形成 1 组实测数据集，共计 5 份数据集用于验证所提框架的有效性，各数据集介绍如下：

（1）数据集 1(Data_Num)：选取太平洋地震工程研究中心（Pacific Earthquake Engineering Research Center, PEER）强震数据库[52]中满足 50 年

超越概率 10% 的地震动记录，用于单自由度非线性体系地震响应分析，该非线性体系在地震激励下的运动方程表达为：

$$m\ddot{x}+c\dot{x}+k_1x+k_2x^3=-m\boldsymbol{\Gamma}\ddot{x}_g \tag{17}$$

式中：$m=1$ kg 为质量；$c=1$ N·s/m 为阻尼系数；$k_1=20$ N/m 为线性刚度系数；$k_2=200$ N/m 为非线性刚度系数；$\boldsymbol{\Gamma}$ 代表力矢量分布；质量归一化恢复力为 $g=h/m$（h：体系的弹性恢复力；m：单自由度体系集中质量）。

对该单自由度非线性体系开展数值模拟，构建包含 100 组独立地震动序列样本的合成数据集。单次模拟时长 50 s，采样频率 20 Hz，每条时程记录共生成 1 001 个数据采样点。模型输入、输出数据维度分别整理为[10，1 001，1]与[10，1 001，3]。参照文献[27]的数据集划分方式，将全部样本分别按 64、21、15 的数量拆分为训练集、验证集与测试集。

（2）数据集 2（Data_SanBernardino）：该数据集实测对象为美国加州圣贝纳迪诺一栋 6 层酒店结构，传感器布设位置为一层、三层与屋面，数据由工程强震动数据库采集记录[53]。数据集共包含 21 条地震响应时程，采样频率 100 Hz，单条样本时程长度 7 200 个采样点；按样本划分，训练集、验证集、测试集分别包含 11 条、4 条、6 条数据。

（3）数据集 3（Data_MRFDBF）：该数据集原型为美国加州一栋足尺 3 层办公楼，建筑整体尺寸 45.7 m（长）×45.7 m（宽）×11.43 m（高）。结构体系包含抗侧力系统、阻尼系统与重力承载系统，其中抗侧力体系由 8 榀相同的单跨钢矩框架构成，结构阻尼效应显著，该算例体系命名为 MRFDBF。

依托文献[27]的研究方案，基于 MATLAB 平台调用 RT-Frame2D 软件对该三层单跨框架开展数值模拟，完成地震动幅值调幅与增量动力分析（Incremental Dynamic Analysis，IDA）。数据集共录入 100 组地震记录，经 IDA 计算得到 548 条响应时程样本。

本文沿用文献[27]的样本划分规则：选取 80 组地震对应的 481 条时程作为训练集，15 组地震对应的 47 条时程作为验证集，剩余 5 组地震生成的 20 条时程作为测试集。

同时设置单特征、多特征两种数据输入模式；其中经多特征扩充后的完整样本集单独记作数据集 4（Data_MRFDBF_final），用于开展全楼层结构动力响应预测。

2.2　实验环境与评价指标

本文全部对比模型统一采用固定超参数配置，且在相同硬件环境下完成训练。所有深度学习模型基于 Python 3.9、PyTorch 1.13.1 与 CUDA 11.7 编程环境搭建训练模型。模型超参数设置如下：批次大小（batch size）取 10，代表单次参数更新所输入样本总量；最大训练轮次设置为 5 000 轮，即训练集完整迭代上限；优化器选用 Adam，初始学习率设定为 0.001，完成网络参数迭代更新。为全面量化深度学习模型的预测精度，本文采用多类评价指标开展性能评估。训练阶段以均方误差（Mean Squared Error，MSE）作为核心损失函数，同时引入多项补充评价指标综合衡量预测效果[包含动态时间规划（Dnamic Time Warping，DTW）参数]，各指标具体计算公式如表 1 所列。

2.3　结果分析

本文在多组不同类型的结构地震响应预测数据集上完成模型测试，对所提 GKAN-Mixer 模型开展全方位性能评估，以此验证该预测框架的泛用性；同时通过对照实验深入剖析 KAN 模块对整体模型预测性能的增益作用。

2.3.1　对比模型

本文选取地震响应预测领域多种经典基线模型开展对比试验，以此评估 GKAN-Mixer 的预测性能，对比结果如图 6 所示。各基线模型介绍如下：

（1）RNN[31]：面向时序数据的经典深度学习预测模型，依靠内部循环结构存储历史信息（当前时刻输入、上一时刻隐状态），能够捕捉序列内部的时序依赖关系。

（2）LSTM-f[34]：现有建筑地震响应预测专用模型。将完整建筑地震动时程送入链式排布的 LSTM 单元，借助单元内部输入门、遗忘门、输出门调控信息流通，有效捕获长短期时序关联，解决传统 RNN 处理长序列时易出现的梯度消失、梯度爆炸缺陷。

（3）RestLSTM[36]：适用于桥梁地震响应预测的改进时序模型。在 LSTM 基础上引入残差连接机制，依靠堆叠残差结构缓解深层网络性能退化问题，降低深层模型的训练优化难度。

（4）GRU-p[42]：面向多类材料结构地震响应

表 1 评估深度学习算法的误差指标

Table 1 Error metrics for evaluating the proposed deep learning algorithm

误差评价指标	公式计算
均方误差	$MSE=\frac{1}{n}\sum_{i=1}^{n}(Actual_i-Predicted_i)^2$
平均绝对误差	$MAE=\frac{1}{n}\sum_{i=1}^{n}\lvert Actual_i-Predicted_i\rvert$
均方根误差	$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(Actual_i-Predicted_i)^2}$
决定系数	$R^2=1-\frac{\sum_{i=1}^{n}(Actual_i-Predicted_i)^2}{\sum_{i=1}^{n}(Actual_i-Actual_{avg})^2}$
动态时间规划	$C_{i,j}=\lvert Actual_i-Predicted_j\rvert+\min\{C_{i,j-1},C_{i-1,j},C_{i-1,j-1}\}$ $DTW=C_{n,n}/n$
5%置信区间	$CI_s=\frac{(Actual_i-Predicter_i)\in CI_{5\%}}{n}$

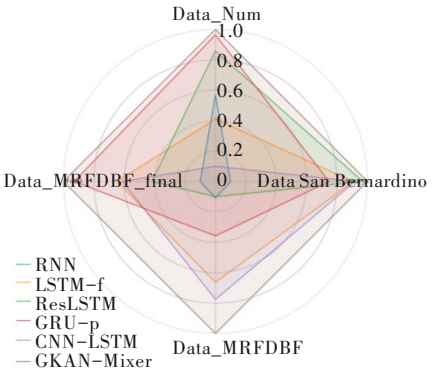


图 6 GKAN-Mixer模型平均结果(MSE)对比图

Fig. 6 Comparison of average results (MSE) of the GKAN-Mixer model

的通用预测模型。对门控机制进行轻量化简化以削减网络参数量,搭配多层残差结构降低模型对浅层特征的依赖。

(5) CNN-LSTM[39]:高速铁路结构地震响应预测混合模型。先通过CNN完成空间特征提取,再将提取后的特征构造为时序序列输入LSTM,最终由LSTM完成结构地震应变响应预测。

2.3.2 案例研究一

针对数据集1(Data_Num),表2试验结果显示:本文所提GKAN-Mixer在位移x、速度xd、加速度xdd全部预测任务上性能均优于其余对比模型。该模型各项预测均取得最低均方误差(Mean Squared Error, MSE):位移x的MSE为0.000 015,速度xd的MSE为0.000 305,加速度xdd的MSE为0.008 065;同时决定系数R^2均趋近于1。CI在±5%与±3%置信误差区间下,模型整体预测精度分别平均提升3.7%、15.7%,拟合精度高、预测误差小。动态时间规整指标同样以GKAN-Mixer最优,说明模型输出预测时程与真实响应时程的匹配对齐效果最优。GRU-p与ResLSTM在位移、速度预测任务中表现尚可,但加速度预测精度存在明显衰减;反观CNN-LSTM在本组数据集整体效果较差,尤其加速度预测的MSE显著偏高,时序曲线对齐效果差。

综上,在单自由度非线性体系数据集1中,GKAN-Mixer是结构地震响应时程预测的最优模型。

2.3.3 案例研究二

数据集2(Data_SanBernardino)为实测工程数据集,全部样本由建筑物内部布设的振动传感器采集得到,存在样本总量少、信号噪声水平高的特点,属于典型小样本预测工况。表3列出了各模型针对该六层建筑结构地震响应的预测结果。

由试验结果可知,GKAN-Mixer对建筑中间楼层响应y_1的预测性能最优,取得最小MSE与均方根误差(Root Mean Square Error, RMSE),对应数值分别为0.000 000 13、0.000 225;该楼层预测的决定系数R^2趋近于1,预测误差最低,拟合效果优异。

CNN-LSTM与ResLSTM在屋面顶层(y_2)的预测中表现突出,由此可推断:针对高噪声时序数据(y_2),LSTM类模型的适配性优于GRU类模型。

综合对比全部测点结果,GKAN-Mixer与CNN-LSTM为本数据集下结构地震响应预测的最优模型。

表 2　Peer 数据集的多变量响应预测结果

Table 2　Multivariate response prediction results on the Peer dataset

		MSE	*MAE*	*RMSE*	*DTW*	R_2	$CI_{5\%}$	$CI_{3\%}$
RNN	x	0.000 084	0.005 078	0.009 194	0.824 931	0.967 441	0.968 962	0.804 317
	xd	0.001 737	0.020 089	0.041 688	3.922 765	0.973 706	0.974 779	0.820 664
	xdd	0.044 246	0.068 015	0.210 347	21.301 823	0.987 492	0.999 999	0.996 668
LSTM-f	x	0.000 110	0.026 729	0.108 291	0.837 010	0.957 427	0.940 704	0.742 195
	xd	0.001 240	0.017 203	0.035 221	3.281 345	0.981 232	0.991 925	0.888 023
	xdd	0.033 830	0.057 140	0.183 930	19.641 019	0.981 232	0.999 999	0.888 023
ResLSTM	x	0.000 037	0.003 303	0.006 153	0.553 943	0.985 418	0.998 818	0.947 045
	xd	0.000 730	0.010 947	0.027 030	2.700 191	0.988 946	0.999 442	0.961 640
	xdd	0.024 234	0.055 961	0.155 673	16.337 446	0.993 149	0.999 999	0.999 927
GRU-p	x	0.000 022	0.002 245	0.004 698	0.493 183	0.991 498	0.999 977	0.988 735
	xd	0.000 410	0.008 318	0.020 251	2.194 194	0.993 795	0.999 995	0.994 294
	xdd	0.011 602	0.042 982	0.107 712	12.304 741	0.996 720	1.0	0.999 999
CNN-LSTM	x	0.000 160	0.006 095	0.012 656	1.014 172	0.938 305	0.882 810	0.652 786
	xd	0.002 745	0.021 258	0.052 397	4.074 675	0.958 462	0.925 021	0.714 632
	xdd	0.083 039	0.099 103	0.288 166	28.879 633	0.976 526	0.999 644	0.967 862
GKAN-Mixer	x	0.000 015	0.002 219	0.003 883	0.376 848	0.994 192	0.999 999	0.999 403
	xd	0.000 305	0.012 386	0.017 471	1.890 522	0.995 381	0.999 999	0.999 403
	xdd	0.008 065	0.037 756	0.089 807	10.217 274	0.995 381	1.0	0.999 999

表 3　圣贝纳迪诺数据集的多变量响应预测结果

Table 3　Multivariate response prediction results on the San Bernardino dataset

		MSE	*MAE*	*RMSE*	*DTW*	R_2	$CI_{5\%}$	$CI_{3\%}$
RNN	y_1	0.000 000 68	0.000 552	0.000 792	0.043 548	0.955 544	0.999 999	0.998 871
	y_2	0.000 001 60	0.000 652	0.001 192	0.072 631	0.945 336	0.999 988	0.991 620
LSTM-f	y_1	0.000 000 19	0.000 267	0.000 424	0.026 381	0.988 033	1.0	0.999 999
	y_2	0.000 000 90	0.000 543	0.000 901	0.051 729	0.968 443	0.999 999	0.999 560
ResLSTM	y_1	0.000 000 15	0.000 251	0.000 378	0.023 248	0.990 671	1.0	0.999 999
	y_2	0.000 000 77	0.000 445	0.000 816	0.047 211	0.977 692	0.999 999	0.999 852
GRU-p	y_1	0.000 000 29	0.000 396	0.000 521	0.028 111	0.980 404	1.0	0.999 999
	y_2	0.000 000 84	0.000 508	0.000 852	0.048 693	0.976 164	0.999 999	0.999 705
CNN-LSTM	y_1	0.000 000 18	0.000 301	0.000 420	0.024 881	0.987 359	1.0	0.999 999
	y_2	0.000 000 76	0.000 462	0.000 816	0.045 887	0.976 906	0.999 999	0.999 861
GKAN-Mixer	y_1	0.000 000 13	0.000 225	0.000 351	0.021 121	0.992 519	1.0	0.999 999
	y_2	0.000 000 81	0.000 448	0.000 833	0.046 287	0.977 718	0.999 999	0.999 785

2.3.4　案例研究三

数据集 3（Data_MRFDBF）的试验结果如表 4 所列。本文所提 GKAN-Mixer 在位移、速度、加速度的结构响应预测性能均优于对比模型。该模型各指标均取得最低 MSE：位移 x 的 MSE 为 0.000 129，速度 xd 的 MSE 为 0.000 111，加速度 xdd 的 MSE 为 0.000 149，同时拥有各组模型中最高的决定系数 R^2。在 ±5%、±3% 置信误差区间下，模型预测精度相较其他算法分别平均提升 21.23%、23.3%，性能优势显著。

综上，针对该三层钢框架数值模拟数据，GKAN-Mixer 是结构地震响应预测的最优模型。

依托数据集 3 对应的三层钢框架原型结构，本文额外构建一组多点位、多特征扩充样本作为补充数据集。为探究 GKAN-Mixer 对多特征输入工况的适配性能，将该扩充样本集记作数据集 4（Data_MRFDBF_final）开展对比试验，系统评估模型对高维混合时序数据的处理能力。该数据集整合结构不同楼层、不同测点、不同采集角度的传感器地震动记录，构建多维复合输入样本，试验重点

表 4 MRFDBF 数据集的多变量响应预测结果

Table 4 Multivariate response prediction results on the MRFDBF dataset

		MSE	*MAE*	*RMSE*	*DTW*	R_2	$CI_{5\%}$	$CI_{3\%}$
RNN	*x*	0.000 248	0.010 806	0.015 078	0.843 094	0.467 378	0.381 482	0.234 885
	xd	0.000 253	0.011 005	0.015 249	0.863 279	0.476 343	0.380 584	0.234 302
	xdd	0.000 293	0.011 872	0.016 506	0.956 775	0.425 208	0.356 312	0.218 625
LSTM-f	*x*	0.000 174	0.009 001	0.012 432	0.721 831	0.579 853	0.447 009	0.278 133
	xd	0.000 167	0.008 767	0.012 113	0.682 037	0.595 298	0.458 102	0.285 600
	xdd	0.000 250	0.010 587	0.014 898	0.892 730	0.440 996	0.382 101	0.235 272
ResLSTM	*x*	0.000 177	0.009 284	0.012 684	0.710 881	0.602 217	0.443 953	0.276 052
	xd	0.000 174	0.009 214	0.012 541	0.705 233	0.621 083	0.449 521	0.279 743
	xdd	0.000 237	0.010 740	0.014 781	0.866 818	0.504 297	0.392 063	0.241 748
GRU-p	*x*	0.005 528	0.048 636	0.069 520	3.909 966	0.694 527	0.463 447	0.289 228
	xd	0.005 202	0.046 703	0.067 111	3.789 375	0.724 012	0.449 521	0.279 743
	xdd	0.007 818	0.058 239	0.083 107	4.864 608	0.621 513	0.392 063	0.241 748
CNN-LSTM	*x*	0.000 207	0.009 053	0.013 061	0.710 191	0.487 146	0.413 460	0.255 808
	xd	0.000 192	0.008 768	0.012 606	0.689 043	0.504 544	0.430 520	0.267 113
	xdd	0.000 246	0.010 097	0.014 693	0.864 795	0.469 685	0.385 766	0.237 670
GKAN-Mixer	*x*	0.000 129	0.006 472	0.010 038	0.298 737	0.747 735	0.545 158	0.321 215
	xd	0.000 111	0.006 251	0.009 574	0.532 885	0.532 885	0.553 363	0.346 149
	xdd	0.000 149	0.007 355	0.011 327	0.674 767	0.679 243	0.482 327	0.302 104

分析多特征输入下模型的泛化性能、拟合效果与时程预测精度。

数据集 4 的试验结果如表 5 所列：面对高复杂度多维输入，GKAN-Mixer 仍可维持优异预测精度；在 ±5%、±3% 置信误差区间内，模型综合预测精度相较于单地震动加速度输入工况实现进一步提升，且各项指标全面优于所有对比基线模型。

上述结果证实，GKAN-Mixer 具备处理复杂多维工程数据的突出潜力。后续研究可拓展输入特征的类型与数量，探究模型在更高维数据集下的应

表 5 MRFDBF 多特征数据集的多变量响应预测结果

Table 5 Multivariate response prediction results on the MRFDBF multi-feature dataset

		MSE	*MAE*	*RMSE*	*DTW*	R_2	$CI_{5\%}$	$CI_{3\%}$
RNN	*x*	0.010 245	0.076 534	0.097 684	5.254 041	0.472 212	0.388 256	0.239 161
	xd	0.010 979	0.078 929	0.100 833	5.493 514	0.473 214	0.376 128	0.231 282
	xdd	0.012 931	0.085 231	0.109 516	6.006 086	0.396 986	0.353 590	0.216 780
LSTM-f	*x*	0.007 152	0.059 047	0.080 767	4.514 798	0.628 805	0.459 667	0.286 663
	xd	0.006 980	0.058 340	0.079 695	4.450 600	0.662 325	0.464 804	0.290 140
	xdd	0.010 070	0.069 647	0.096 594	5.677 898	0.527 852	0.399 632	0.246 719
ResLSTM	*x*	0.005 868	0.052 916	0.074 318	4.181 511	0.695 958	0.500 968	0.314 969
	xd	0.005 564	0.050 955	0.071 752	3.985 088	0.737 490	0.512 744	0.323 180
	xdd	0.008 451	0.061 969	0.089 202	5.196 837	0.605 752	0.432 655	0.268 536
GRU-p	*x*	0.005 528	0.048 636	0.069 520	3.909 966	0.694 527	0.513 831	0.323 938
	xd	0.005 202	0.046 703	0.067 111	3.789 375	0.724 012	0.527 415	0.333 502
	xdd	0.007 818	0.058 239	0.083 107	4.864 608	0.621 513	0.447 842	0.278 694
CNN-LSTM	*x*	0.007 337	0.056 304	0.081 154	4.432 211	0.590 165	0.453 580	0.282 556
	xd	0.007 195	0.055 727	0.080 047	4.420 954	0.609 586	0.457 802	0.285 405
	xdd	0.010 011	0.066 550	0.096 074	5.772 211	0.503 763	0.400 119	0.247 041
GKAN-Mixer	*x*	0.004 858	0.045 426	0.065 536	3.800 021	0.747 102	0.542 451	0.344 210
	xd	0.004 561	0.043 611	0.063 002	3.605 122	0.780 758	0.556 991	0.354 679
	xdd	0.006 827	0.054 502	0.078 309	4.684 503	0.684 023	0.475 359	0.297 334

用效果，并对网络架构进行轻量化优化，进一步提升模型运算效率与预测精度。

为进一步挖掘模型捕捉的结构地震响应内在规律，定量量化各输入特征对预测结果的贡献，并提升模型可解释性，本文采用噪声扰动法对数据集5的多特征样本开展特征敏感性分析。

试验设置10%、30%、50%三档强度的噪声扰动工况，具体实施方式为：针对单条样本内任意单一特征，10%扰动工况下以原始特征值为基准，叠加幅值为原值10%的随机噪声；遍历全部特征并依次施加不同强度扰动，根据模型输出变化幅度判定各特征对预测结果的影响权重。

由图7结果可得：施加10%轻度噪声扰动时，模型鲁棒性良好，预测性能维持稳定；随噪声扰动强度升高，整体预测误差逐步增大。即便在50%重度噪声干扰工况下，MAE波动幅度仍控制在0.01以内。该现象表明模型充分学习了全部输入特征的有效信息，不存在对单一特征的强依赖性，充分体现多特征联合输入方案的优越性。

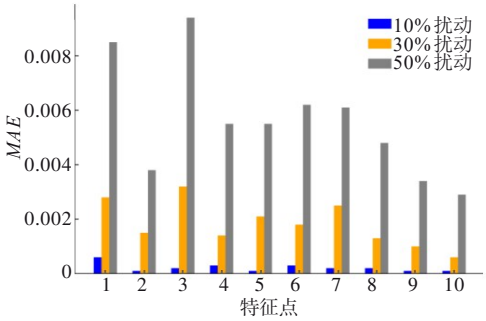


图7　多特征数据集扰动分析图

Fig. 7　Perturbation analysis on the multi-feature dataset

2.4　可视化分析

为直观展示GKAN-Mixer的拟合性能，绘制图8。由图8可知，GKAN-Mixer输出预测值与实测真实响应吻合度高，预测曲线几乎与理想参考曲线完全重合，证明该模型具备优异的数据拟合能力。

模型参数量直接决定训练耗时、收敛速度与预测精度，是衡量模型计算效率的核心评价指标。本文对比分析不同参数量规模对模型性能的影响规律，试验结果如图9所示。

2.5　消融实验

本文设置消融对照实验，将所提GKAN-Mixer与替换KAN模块为MLP的GRU-Mixer开展对比，定量验证KAN模块对模型预测性能的增益效果。以MSE作为评价指标可以发现：在全部数据集上，GKAN-Mixer均展现出明显性能优势，相较GRU-Mixer，MSE指标整体优化幅度约10%。该结果证明KAN模块的特征拟合能力优于传统MLP，可有效降低预测误差、提升模型整体精度。

综上，消融实验充分证实KAN模块是保障GKAN-Mixer预测性能的核心组件，具体试验数据如表6所列。

3　结论

本文融合机器学习与结构地震响应预测，构建一套地震动激励下结构动力响应高精度实时预测框架，所提出的GKAN-Mixer模型可完成建筑多测点动力响应时程同步预测。

首先，模型引入多尺度分解策略，降低网络对样本数量的依赖，缓解小样本工况对模型鲁棒性与泛化能力的不利影响；其次，选取GRU作为主干时序网络提升预测精度，计算速度显著优于传统有限元数值分析手段；再者，采用改进KAN模块替代传统MLP，使模型对多尺度、时变耦合时序数据具备更强适配能力。综合多组数值与实测算例试验分析，得到主要结论如下：

（1）本文所提GKAN-Mixer可通过学习地震动时程至结构动力响应的非线性映射关系，实现建筑多测点地震响应高精度实时预测。多类工程结构数据集试验充分证明了该方法的有效性、预测精度与计算高效性。

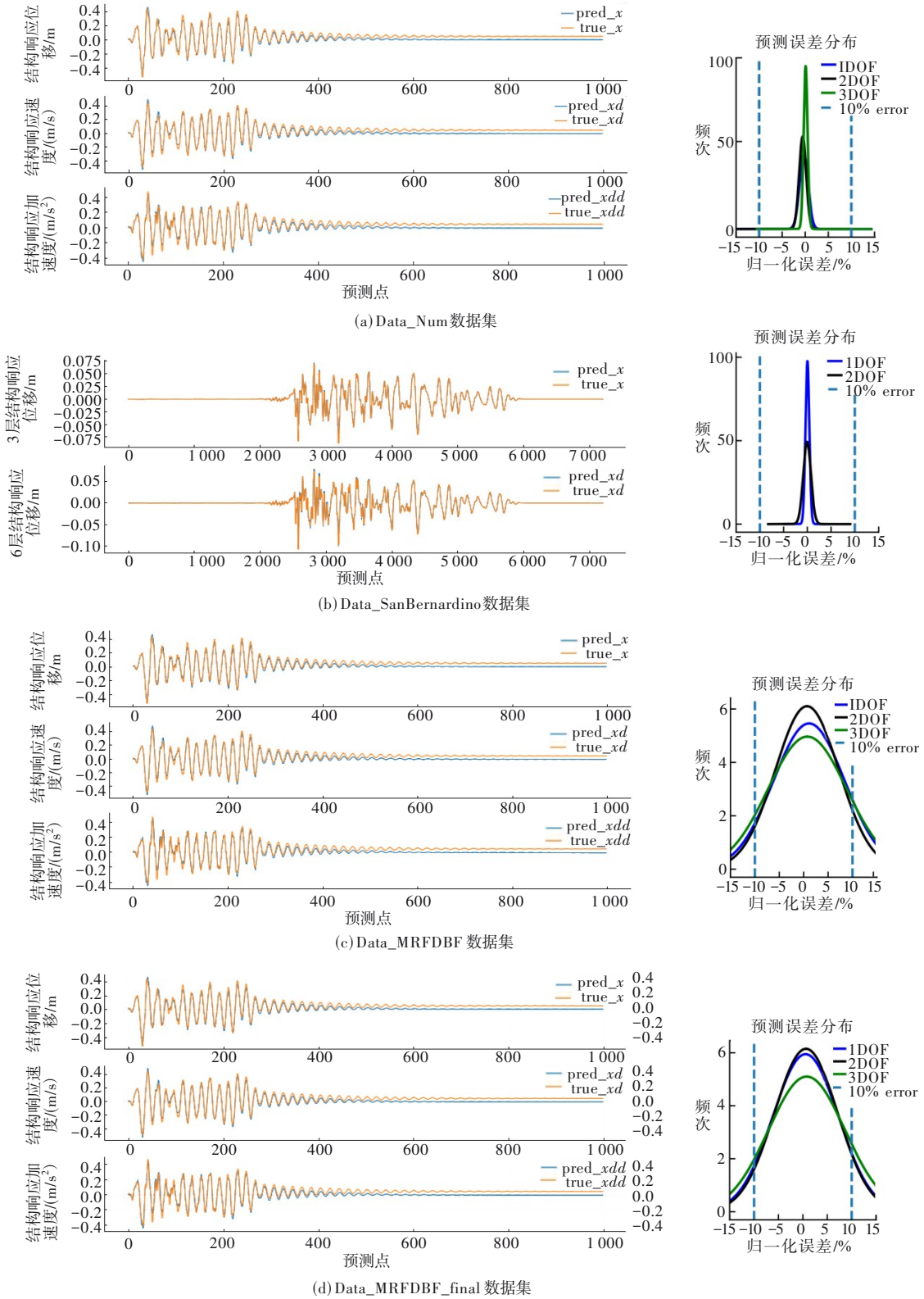
（2）在结构地震响应预测任务中，采用KAN模块替换传统MLP能够显著提升模型预测精度，该改进思路具备良好应用潜力。

（3）试验结果证实，该框架可有效融合多点位地震动输入与多维结构特征开展联合学习，提升模型训练稳定性，赋予模型优异的抗干扰鲁棒性。

同时，本文所提方法仍存在一定局限性，相关问题有待后续深入研究：

（1）算例二实测结果可见，模型对高层建筑顶层响应的预测精度低于低层测点，该特性会限制模型在超高层结构工程中的应用。

（2）当前模型属于纯数据驱动建模方式，若将建筑几何、刚度等先验结构信息编码后嵌入网络，有望进一步提升预测精度。



(a) Data_Num数据集

(b) Data_SanBernardino数据集

(c) Data_MRFDBF 数据集

(d) Data_MRFDBF_final 数据集

(1DOF:1自由度体系;2DOF:2自由度体系;3DOF:3自由度体系)

图8 GKAN-Mixer模型在Data_MRFDBF_final数据集的预测拟合曲线及回归分析

Fig. 8 Prediction fitting curves and regression analysis of the GKAN-Mixer model on the Data_MRFDBF_final dataset

(3) 现有结构响应预测模型普遍存在跨建筑迁移能力不足的缺陷,后续可将各类结构固有特征引入模型编码,提升模型通用性与多类结构适配精度。

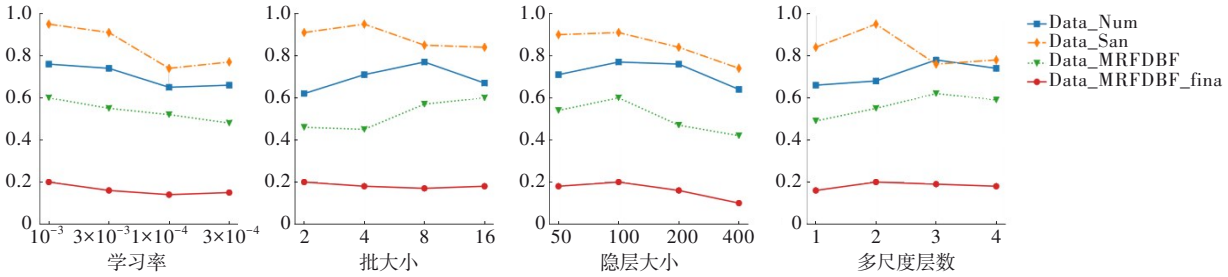


图 9　GKAN-Mixer 模型在不同超参数下的性能

Fig. 9　Performance of the GKAN-Mixer model under different hyperparameters

表 6　KAN 模块消融实验

Table 6　Ablation experiments on the KAN module

数据集	模型	*MSE*	*MAE*	*RMSE*	*DTW*	R_2	$CI_{5\%}$	$CI_{3\%}$
Data_Num	GKAN-Mixer	0.000 015	0.002 219	0.003 883	0.376 848	0.994 192	0.999 999	0.999 403
	GRU-Mixer	0.000 023	0.002 646	0.004 835	0.466 264	0.990 995	0.999 959	0.986 129
Data_San	GKAN-Mixer	0.000 000 13	0.000 225	0.000 351	0.021 121	0.992 519	1.0	0.999 999
	GRU-Mixer	0.000 000 15	0.000 240	0.000 396	0.048 429	0.997 060	1.0	0.999 999
Data_MRFDBF	GKAN-Mixer	0.000 025	0.003 498	0.005 088	0.298 737	0.941 286	0.501 260	0.315 168
	GRU-Mixer	0.000 026	0.003 618	0.005 161	0.334 812	0.939 572	0.519 388	0.327 844
Data_MRFDBF_final	GKAN-Mixer	0.004 858	0.045 426	0.065 536	3.800 021	0.747 102	0.542 451	0.344 210
	GRU-Mixer	0.005 357	0.045 333	0.066 424	3.746 959	0.729 541	0.520 761	0.328 810

参考文献（References）

[1] Luo H, Paal S G. Data-driven seismic response prediction of structural components[J]. Earthquake Spectra, 2022, 38(2): 1382-1416.

[2] Morfidis K, Kostinakis K. Rapid prediction of seismic incident angle's influence on the damage level of RC buildings using artificial neural networks[J]. Applied Sciences, 2022, 12(3): 1055-1070.

[3] Oh B K, Glisic B, Park S W, et al. Neural network-based seismic response prediction model for building structures using artificial earthquakes[J]. Journal of Sound and Vibration, 2020, 468: 115109.

[4] Bretas E M, Lemos J V, Lourenço P B. Seismic analysis of masonry gravity dams using the discrete element method: implementation and application[J]. Journal of Earthquake Engineering, 2016, 20(2): 157-184.

[5] Nie Y, Sheikh A, Griffith M, et al. A damage-plasticity based interface model for simulating in-plane/out-of-plane response of masonry structural panels[J]. Computers & Structures, 2022, 260: 106721.

[6] Suthar A, Suthar S M. Dynamic analysis of rcc building with shear wall[J]. International Journal of Earth Sciences and Engineering, 2011, 4(6): 659-662.

[7] Wilkinson S M, Hiley R A. A non-linear response history model for the seismic analysis of high-rise framed buildings[J]. Computers & Structures, 2006, 84(5/6): 318-329.

[8] Asgarieh E, Moaveni B, Stavridis A. Nonlinear finite element model updating of an infilled frame based on identified time-varying modal parameters during an earthquake[J]. Journal of Sound and Vibration, 2014, 333(23): 6057-6073.

[9] Saka M P, Geem Z W. Mathematical and metaheuristic applications in design optimization of steel frame structures: an extensive review[J]. Mathematical Problems in Engineering, 2013(16): a271031.

[10] Torky A A, Aburawwash A A. A deep learning approach to automated structural engineering of prestressed members[J]. International Journal of Structural and Civil Engineering Research, 2018: 347-352.

[11] Alves V, Cury A, Roitman N, et al. Structural modification assessment using supervised learning methods applied to vibration data[J]. Engineering Structures, 2015, 99: 439-448.

[12] Azimi M, Pekcan G. Structural health monitoring using extremely compressed data through deep learning[J]. Computer-Aided Civil and Infrastructure Engineering, 2020, 35(6): 597-614.

[13] Seventekidis P, Giagopoulos D, Arailopoulos A, et al. Structural health monitoring using deep learning with optimal finite element model generated data[J]. Mechanical Systems and Signal Processing, 2020, 145: a106972.

[14] Chi N W, Wang J P, Liao J H, et al. Machine learning-based seismic capability evaluation for school buildings[J]. Automation in Construction, 2020, 118: 103274.

[15] Alvin K F, Robertson A N, Reich G W, et al. Structural system identification: from reality to models[J]. Computers & Structures, 2003, 81(12): 1149-1176.

[16] Karimi I, Khaji N, Ahmadi M T, et al. System identification

of concrete gravity dams using artificial neural networks based on a hybrid finite element-boundary element approach [J]. Engineering Structures, 2010, 32(11): 3583-3591.

[17] Javadi A A, Tan T P, Elkassas A S I. Intelligent finite element method and application to simulation of behavior of soils under cyclic loading [M]//Abraham A, Hassanien A E, Snášel V. Foundations of Computational Intelligence Volume 5: Function Approximation and Classification. Berlin, Heidelberg: Springer, 2009: 317-338.

[18] Zopf C, Kaliske M. Numerical characterisation of uncured elastomers by a neural network based approach[J]. Computers & Structures, 2017, 182: 504-525.

[19] Su T H, Huang S J, Jean J G, et al. Multiscale computational solid mechanics: data and machine learning [J]. Journal of Mechanics, 2022, 38: 568-585.

[20] Wang Q, Wang J H, Huang X F, et al. Semiactive nonsmooth control for building structure with deep learning [J]. Complexity, 2017: 6406179.

[21] Kim H S. Development of seismic response simulation model for building structures with semi-active control devices using recurrent neural network [J]. Applied Sciences, 2020, 10 (11): 3915.

[22] Li T, Pan Y X, Tong K T, et al. Attention-based sequence-to-sequence learning for online structural response forecasting under seismic excitation [J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2022, 10(8): 2184-2200.

[23] Lagaros N D, Papadrakakis M. Neural network based prediction schemes of the non-linear seismic response of 3D buildings [J]. Advances in Engineering Software, 2012, 44 (1): 92-115.

[24] Huang C S, Hung S L, Wen C M, et al. A neural network approach for structural identification and diagnosis of a building from seismic response data [J]. Earthquake Engineering & Structural Dynamics, 2003, 32(2): 187-206.

[25] Le Cun Y, Bengio Y. Convolutional networks for images, speech, and time series[C]// ARBIB M A. The Handbook of Brain Theory and Neural Networks. Cambridge: MIT Press, 1998: 255-258.

[26] Wu R T, Jahanshahi M R. Deep convolutional neural network for structural dynamic response estimation and system identification [J]. Journal of Engineering Mechanics, 2019, 145: 04018125.

[27] Zhang R Y, Liu Y, Sun H. Physics-guided convolutional neural network (PhyCNN) for data-driven seismic response modeling[J]. Engineering Structures, 2020, 215: 110704.

[28] Sun S B, He Y Y, Zhou S D, et al. A data-driven response virtual sensor technique with partial vibration measurements using convolutional neural network [J]. Sensors, 2017, 17 (12): 288816.

[29] Kim T, Kwon O S, Song J. Response prediction of nonlinear hysteretic systems by deep neural networks [J]. Neural Networks, 2019, 111: 1-10.

[30] Messner M C. Convolutional neural network surrogate models for the mechanical properties of periodic structures[J]. Journal of Mechanical Design, 2020, 142(2): 024503.

[31] Williams R J, Zipser D. A learning algorithm for continually running fully recurrent neural networks [J]. Neural Computation, 1989, 1(2): 270-280.

[32] Graves A. Long short-term memory [M]. Heidelberg: Springer, 2012 : 37-45.

[33] Chung J, Gulcehre C, Cho K, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling [PP/OL]. arXiv (2014-12-11) [2026-08-06]. https://arxiv.org/abs/1412.3555.

[34] Zhang R Y, Chen Z, Chen S, et al. Deep long short-term memory networks for nonlinear structural seismic response prediction[J]. Computers & Structures, 2019, 220: 55-68.

[35] Zhang R Y, Liu Y, Sun H. Physics-informed multi-LSTM networks for metamodeling of nonlinear structures [J]. Computer Methods in Applied Mechanics and Engineering, 2020, 369: 113226.

[36] Liao Y C, Zhang R, Lin R, et al. Nonlinear seismic response prediction of bridges based on cascaded residual LSTM networks[J]. Engineering Mechanics. 2024, 41(4): 47-58.

[37] Xu Z K, Chen J, Shen J X, et al. Recursive long short-term memory network for predicting nonlinear structural seismic response[J]. Engineering Structures, 2022, 250: 113406.

[38] Ahmed B, Mangalathu S, Jeon J S. Generalized stacked LSTM for the seismic damage evaluation of ductile reinforced concrete buildings [J]. Earthquake Engineering & Structural Dynamics, 2023, 52(11): 3477-3503.

[39] Zhang X B, Xie X N, Tang S H, et al. High-speed railway seismic response prediction using CNN-LSTM hybrid neural network [J]. Journal of Civil Structural Health Monitoring, 2024, 14(5): 1125-1139.

[40] Vaswani A. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, 30:1623-1637.

[41] Li F Y, Zhou H L, Wang Z Y, et al. ADDCNN: an attention-based deep dilated convolutional neural network for seismic facies analysis with interpretable spatial - spectral maps [J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(2): 1733-1744.

[42] Xu Y J, Lu X Z, Fei Y F, et al. Hysteretic behavior simulation based on pyramid neural network: principle, network architecture, case study and explanation [J]. Advances in Structural Engineering, 2023, 26(13): 2359-2374.

[43] Meng S Q, Zhou Y, Gao Z Y. Refined self-attention mechanism based real-time structural response prediction method under seismic action[J]. Engineering Applications of Artificial Intelligence, 2024, 129: 107380.

[44] Zhang Q Y, Guo M Z, Zhao L L, et al. Transformer-based structural seismic response prediction [J]. Structures, 2024, 61: 105929.

(下转第 1357 页)