



基于深度学习的日本强震远场触发中国东北地震预测

高峰, 李美

引用本文:

高峰, 李美. 基于深度学习的日本强震远场触发中国东北地震预测[J]. 大地测量与地球动力学, 2026, 46(6): 710–717.

Gao Feng, Li Mei. Prediction on Far-Field Triggering Earthquakes in Northeast China Induced by Japanese Strong Earthquakes Based on Deep Learning[J]. Journal of Geodesy and Geodynamics, 2026, 46(6): 710–717.

在线阅读 View online: <https://doi.org/10.14075/j.jgg.2025.09.317>

您可能感兴趣的其他文章

Articles you may be interested in

山东长清4.1级地震序列的遗漏地震检测与发震断层

Detection of Missing Earthquakes and Seismogenic Fault of Changqing 4.1 Earthquake Sequence in Shandong Province

大地测量与地球动力学. 2023, 43(12): 1211–1217 <https://doi.org/10.14075/j.jgg.2023.12.001>

基于小波分析对形变扰动信号的触发检测方法

Trigger Detection Method of Deformation Disturbance Signal Based on Wavelet Analysis

大地测量与地球动力学. 2023, 43(12): 1229–1234 <https://doi.org/10.14075/j.jgg.2023.03.121>

2022年门源6.9级地震前后 b 值时空特征研究

Research on Spatio-Temporal Characteristics of b -Value before and after the 2022 Menyuan M_S 6.9 Earthquake

大地测量与地球动力学. 2023, 43(11): 1155–1161 <https://doi.org/10.14075/j.jgg.2023.11.010>

InSAR约束下的2021年西藏比如 M_W 5.7地震断层几何参数及滑动分布

Fault Geometry Parameters and Slip Distribution of the 2021 Biru, Tibet M_W 5.7 Earthquake Constraint by InSAR

大地测量与地球动力学. 2023, 43(10): 1086–1089,1100 <https://doi.org/10.14075/j.jgg.2023.10.017>

基于KF-LSTM组合模型的短期电离层TEC预测

Short-Term Ionospheric TEC Prediction Based on KF-LSTM Combination Model

大地测量与地球动力学. 2023, 43(10): 1020–1025 <https://doi.org/10.14075/j.jgg.2023.10.006>

基于深度学习的日本强震远场触发 中国东北地震预测

高峰¹ 李美²

¹ 黑龙江省地震局, 哈尔滨, 150090

² 中国地震局地震预测研究所, 北京, 100036

摘要: 远场地震触发机制是地震预测研究的重要前沿方向。为探究日本强震活动与中国东北地区中强地震的时空关联, 本文构建了一种数据驱动的预测模型。基于 1980—2024 年 USGS 地震目录, 以日本 $M \geq 6.0$ 地震为潜在触发源, 预测未来 60 d 内中国东北地区发生 $M \geq 4$ 地震的概率。构建一个涵盖地震统计、空间分布、能量释放及余震序列等 32 维特征的日分辨率时间序列数据集, 系统对比长短期记忆神经网络 (LSTM)、注意力机制增强 LSTM (Attention-LSTM)、Transformer 等深度学习模型与逻辑回归 (logistic regression, LR)、随机森林 (random forest, RF) 等传统机器学习方法。结果显示, Attention-LSTM 模型表现最优, F1 分数达 0.915, AUC 值为 0.661, 在地震预测专用的 Molchan 误差图分析中亦展现显著优势。该模型不仅可实现二元分类预测, 还可生成 $1^\circ \times 1^\circ$ 网格分辨率的空间概率分布图。研究验证了深度学习在揭示跨区域地震远场触发机制中的潜力, 可为中短期区域地震危险性评估提供新思路。

关键词: 深度学习; 注意力机制长短期记忆网络; 远场地震触发; 地震时空预测

中图分类号: P315

文献标识码: A

地震本质上是地球内部应力长期积累后的瞬时释放现象。地震预测作为地球科学领域最富挑战性的前沿课题之一, 长期依赖区域前兆观测与统计分析方法, 如 Gutenberg-Richter 关系与 Omori 定律等经典模型等^[1]。然而, 近年研究揭示, 大地震可通过远程传播的地震波及其动态应力扰动, 在远场区域诱发触发性地震活动, 短期内显著提升其他区域的地震发生率^[2-5]。典型案例如 2011 年日本 $M9.0$ 地震后, 全球多地观测到远场地震活动增强及时序统计特征的系统性变化^[6]。鉴于日本地处全球最活跃的板块边界, 其强震事件是否对中国东北地区产生远场触发效应, 构成一项兼具科学价值与现实意义的研究课题。

传统的地震预测统计模型在刻画地震过程的强非线性与复杂动力学行为方面存在固有局限。伴随计算能力的提升和观测数据的累积, 机器学习 (machine learning, ML), 特别是深度学习 (deep learning, DL), 为革新地震预测范式开辟了新路径。研究表明, 深度神经网络可有效学

习大地震后的余震空间分布模式, 其预测效能超越传统 Coulomb 应力模型^[7]; 基于实验室断层声发射信号的 ML 模型能够识别微米级前兆信息, 实现地震成核时间的精准预测^[8]。在时间序列建模方面, 循环神经网络 (recurrent neural network, RNN) 及其变体长短期记忆网络 (long short-term memory, LSTM) 凭借捕捉长程依赖关系的能力, 可从复杂时序数据中挖掘与地震活动密切相关的深度特征。此外, 注意力 (Attention) 机制与 Transformer 架构的引入, 使模型能够自适应地聚焦于关键时空片段, 在提升预测精度的同时, 显著增强了决策过程的物理可解释性。

尽管如此, 针对跨区域地震远场触发的深度学习建模仍处探索初期, 面临多重挑战: 远场触发信号强度微弱, 易被区域背景噪声淹没; 中强震预测属于典型的小样本不平衡问题, 正例样本稀缺导致模型训练困难; 深度模型固有的“黑箱”特性制约了物理机制的揭示; 评估体系缺乏统一标准, 尤其欠缺 Molchan 误差图等地震预测专用

收稿日期: 2025-09-17

项目来源: 2025 年度黑龙江省地震局黑龙江及邻区地震数据综合应用创新团队。

第一作者简介: 高峰, 高级工程师, 主要研究方向为地震学和地电地磁, E-mail: 1.92@163.com。

评估指标,难以实现模型性能的客观比较^[9-10]。上述瓶颈亟需通过数据增强、算法优化与评估框架完善等系统性方法加以突破。

为应对上述挑战,本文开展如下研究工作:1)构建日本主震→中国东北远场触发的完整预测框架,涵盖数据预处理、特征工程、模型训练与多维度评估;2)在统一数据集与评价体系下,系统对比 LSTM、Attention-LSTM、Transformer 与逻辑回归、随机森林等基线模型的预测性能;3)突破常规分类指标局限,引入 Molchan 错误图作为核心评估工具,量化模型在实际地震预测场景中的时空预报效能;4)实现时空联合建模,生成高分辨率空间概率预测图,并融合 SHAP 值与注意力权重分析,提升模型的可解释性与物理一致性。

1 数据及数据预处理

1.1 数据与预处理

本研究使用的数据来源于美国地质调查局(United States Geological Survey, USGS)提供的全球地震目录,该目录具备较高的完整性和时空精度。为构建地震触发模型,选取了2个关键区域的数据。

1)触发区:日本及其周边地区($120^{\circ}\sim 150^{\circ}\text{E}$, $20^{\circ}\sim 50^{\circ}\text{N}$),时间范围为1980-01-01—2024-12-31,震级 $M\geq 5.0$,共3871次地震事件。

2)响应区:中国东北地区($116^{\circ}\sim 134^{\circ}\text{E}$, $38^{\circ}\sim 54^{\circ}\text{N}$),时间范围与触发区相同,震级 $M\geq 4.0$,共374次地震事件。

选择1980年作为研究起点,旨在确保地震目录记录的均一性与可靠性。所有地震时间均统一为世界标准时(UTC),为后续数据处理和模型构建奠定一致的时间基准。

1.2 特征工程

为了全面刻画日本地区的地震活动状态,以天为单位构建一个包含32个维度的时间序列特征集。这些特征旨在从不同角度捕捉可能与远场触发有关的物理量。特征体系分为4类。

1)基础地震统计特征(8维):包括每日地震次数、最大震级、总能量释放(根据 Gutenberg-Richter (G-R)能量-震级关系计算)等。

2)空间分布特征(6维):包括经度差、纬度差、距离加权震级等,用于描述震源区的空间迁移和聚集模式。

3)地震学特征(6维):包括主震标识、震源

深度标准差、浅源地震数等,旨在捕捉地震序列的演化特性。

4)时序与余震序列特征(12维):包括能量释放率、距上次 $M\geq 6.0$ 主震的时间间隔、G-R 关系 b 值等,用于刻画地震活动的时间演化趋势。

所有特征在输入模型前均进行了标准化处理,以消除量纲差异,提高模型训练的稳定性和收敛速度。

1.3 样本构建与不平衡处理

本部分的核心任务是,利用过去14d日本地区的32维特征序列,预测未来60d内中国东北地区是否会发生 $M\geq 4.0$ 地震。这是一个典型的二元分类问题。

采用滑动窗口法构建样本,窗口长度为14d,步长为7d。对于每个窗口,如果其后60d内响应区发生了至少一次 $M\geq 4.0$ 地震,则该样本标签为1(正样本),否则为0(负样本)。

由于目标地震事件是稀疏的,直接构建的样本集存在严重的类别不平衡问题(正样本比例远低于10%)。为解决此问题,采用了基于严格时间关联的平衡化采样策略,并结合 SMOTE (synthetic minority over-sampling technique) 算法对训练集中的少数类样本进行过采样,最终使训练集中的正负样本比例达到一个相对均衡的水平(如3:7),同时确保测试集保持原始的数据分布,客观评估模型的泛化能力。

2 数据处理方法

2.1 模型架构

为系统评估不同模型在远场触发地震预测任务中的表现,本文构建以下3类先进的深度学习模型。

1)长短期记忆网络(LSTM)。作为处理时间序列问题的基准深度学习模型,LSTM 通过其独特的门控机制(输入门、遗忘门、输出门)有效解决了传统 RNN 中的梯度消失和爆炸问题,能够学习到序列中的长期依赖关系。

2)注意力机制长短期记忆网络(Attention-LSTM)。在标准 LSTM 的基础上引入注意力机制。该机制允许模型在生成输出时,对输入序列的不同时间步赋予不同的权重。在本地震预测任务中,这意味着模型可以自动识别出对未来地震发生最具指示意义的某几天或某个关键事件。

3)Transformer 模型。该模型摒弃了循环结构,完全基于自注意力机制来捕捉序列内的依赖

关系。通过其多头注意力(multi-head attention)和位置编码(positional encoding)机制,Transformer 能够并行处理序列数据,并从全局视角建立输入与输出之间的关系。

此外,本文还训练了逻辑回归和随机森林作为传统机器学习方法的基线,以便与先进机器学习方法进行性能比较。

2.2 特征构建

本研究聚焦“源区(日本)→响应区(中国东北)”的远场触发效应,输入特征主要基于日本侧地震活动序列,用于表征外部动力扰动;同时引入日本地震事件到东北主要构造带中心的距离与方位角,以及目标区域的历史背景地震率,分别刻画可能的应力传播路径和响应区的长期活动水平。

经典远场触发研究多采用 $M \geq 6$ 强震作为触发源。为兼顾样本量与时间连续性,本文选取 $M \geq 5$ 事件作为输入震级下限,以提高序列的时间分辨率和覆盖度;实际起主导动态扰动作用的仍为序列中的 $M \geq 6$ 事件,注意力可视化结果也显示强震后的关键时间窗权重增大。由于预测目标为“未来是否发生 $M \geq 4$ 地震”,采用 $M \geq 5$ 输入有利于保留更多前兆型时间演化信息,增强模型的灵敏度。

2.3 空间概率预测

除了进行二元分类预测,本文还将响应区(中国东北)划分为 $1^\circ \times 1^\circ$ 的空间网格。修改深度学习模型的输出层,使其能够为每个网格预测一个独立的发生概率。这使得模型不仅能回答“是否会发生地震”,还能指示“哪个区域最可能发生地震”,为地震危险性的空间评估提供了更精细化的信息。

2.4 空间命中率

空间命中率(spatial hit rate, SHR)用于衡量模型对地震发生位置的空间定位能力。对于每一个实际发生了地震的时间样本,模型会给出各网格的预测概率,并取其中概率最高的网格作为预测位置。如果该网格恰好与真实地震发生的网格一致,则视为一次“命中”。空间命中率即所有发生过地震的样本中,被正确命中的比例,取值范围为 $0 \sim 1$,数值越大表示空间定位越准确。

2.5 训练与评估

将构建好的数据集按 8:2 的比例划分为训练集和测试集,确保划分过程按时间顺序进行,

避免数据泄漏。模型训练采用 Adam 优化器和 Focal Loss 损失函数。Focal Loss 是在交叉熵损失基础上引入动态调制因子的改进方法,通过降低易分类样本对整体损失的贡献,使模型在训练过程中更加关注难分类样本,从而有效缓解类别不平衡问题。

模型性能的评估采用了一系列综合指标。

1) 分类指标:准确率(Accuracy)、精确率(Precision)、召回率(Recall)、F1 分数(F1-Score)和 ROC-AUC。准确率定义为:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}$$

式中,TP 表示真阳性:实际发生地震,模型也预测会发生(判为 1);TN 表示真阴性:实际未发生地震,模型也预测不会发生(判为 0);FP 表示假阳性:实际未发生地震,但模型误报为会发生,为虚警;FN 表示假阴性:实际发生地震,但模型漏判为不会发生,为漏报。准确率在未参与训练的 20% 独立测试集上计算得到。

ROC(receiver operating characteristic)曲线是在 $[0, 1]$ 范围内扫描判决阈值,分别计算每个阈值下的真阳性率 $\text{TPR} = \text{TP} / (\text{TP} + \text{FN})$ 与假阳性率 $\text{FPR} = \text{FP} / (\text{FP} + \text{TN})$,并以 FPR 为横轴、TPR 为纵轴绘制而成;AUC(area under curve)为该曲线下的面积,取值 0.5 (随机)~ 1.0 (理想),可理解为从正负样本各随机抽取一个时,模型给正样本更高分的概率。

2) 地震预测专业指标:Molchan 错误图及其对应的错误面积(Molchan error area, MEA)。Molchan 图通过绘制漏报率随警报时空占用率的变化关系,直观地评估预测模型在不同警报策略下的效能^[11]。一个理想的预测模型应在尽可能低的警报率下实现尽可能低的漏报率,其曲线应尽可能偏向左下角。

3 结 果

3.1 模型性能对比

图 1 展示了 5 种模型在 6 个评价指标下的对比结果。从结果看,LSTM 在 Accuracy(0.850)与 F1-Score(0.916)上表现最佳;Transformer 在 AUC 上最高(0.741),体现出优异的阈值无关分类判别能力;random forest 的 Recall 最高(0.995),但其 Precision 较低(0.842),显示以高召回为代价的取舍。对于地震预测专用指标 Molchan 错误面积,各模型均明显优于随机基线

(约 0.5),其中 LSTM 和 Attention-LSTM 的错误面积更小,表明其在漏报-报警覆盖率平衡方面更具优势。综合来看,擅长捕捉时间依赖特征的深度学习模型(LSTM、Attention-LSTM、Transformer)在 F1 分数与 Molchan 错误面积等关键指标上整体表现更优。

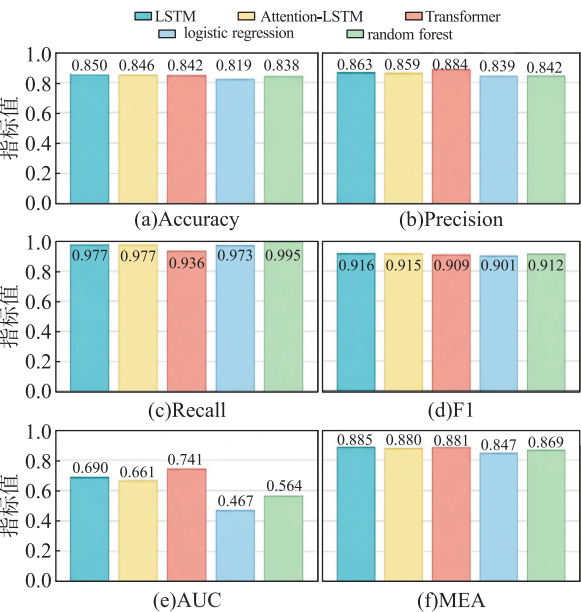


图 1 不同模型在 6 项指标上的性能对比
Fig.1 Performance comparison of different models across six metrics

表 1 汇总了不同模型在分类预测任务中的主要评价指标(Accuracy、Precision、Recall、F1、AUC 以及 MEA)。由表可见:1)LSTM 的 F1 最高(0.916),且在 Accuracy(0.850)与 Recall(0.977)上表现靠前,综合性能最佳;2)Attention-LSTM 的 Precision 最高(0.859),且 AUC 表现稳定(0.661),说明注意力机制有助于提升正例判定的精确性;3)Transformer 的 AUC 最高(0.741),体现出更强的阈值无关分类判别能力;4)传统模型(LR、RF 等)在个别指标上虽有竞争力,但整体处于次优梯队。需要指出,Molchan 错误面积(MEA)面向地震异常检出的专业指标,进一步揭示模型在时空告警效率上的差异,提示部分模型仍有优化空间。

表 1 不同模型在主要分类指标上的性能对比
Tab.1 Performance comparison of different models on major classification metrics

模型名称	准确率	精确率	召回率	F1 分数	AUC	MEA
LSTM	0.850	0.863	0.977	0.916	0.690	0.470
Attention-LSTM	0.846	0.859	0.977	0.915	0.661	0.474
Transformer	0.842	0.884	0.936	0.909	0.741	0.462
LR	0.819	0.839	0.973	0.901	0.467	0.504
RF	0.777	0.843	0.904	0.872	0.560	0.490

可以看出,LSTM 模型在分类预测任务中整体表现最优,尤其在召回率与 F1 分数上优势明显,适合用于对中强震等小概率事件的识别。Attention-LSTM 和 Transformer 则在提高模型鲁棒性与可解释性方面具有潜力。

综合上述评价指标,尽管 LSTM 在 F1 与召回率上略占优势,但 Attention-LSTM 在 Molchan 错误面积(MEA)、空间命中率以及注意力权重所带来的可解释性方面表现更优。基于其更强的空间化能力与解释性,本文选择 Attention-LSTM 作为后续空间预测与解释分析的核心模型。

为评估模型在地震场景中的实际可用性,以 Attention-LSTM 为例,绘制了模型的 Molchan 错误图(图 2)。该图以告警占比 α (横轴,0~1)与漏报率 β (纵轴,0~1)为坐标,随机猜测基准线为 $\beta=1-\alpha$ 。曲线越靠近原点、越远离基准线,表示在相同漏报率下所需的告警时空占比更小,模型越有效。其曲线整体位于基准线下方,Molchan 错误面积(MEA)=0.474(随机预测为 0.5,越小越好),表明在保持同等漏报率时,该模型所需的告警时空占比小于随机指派,具有实际应用价值。

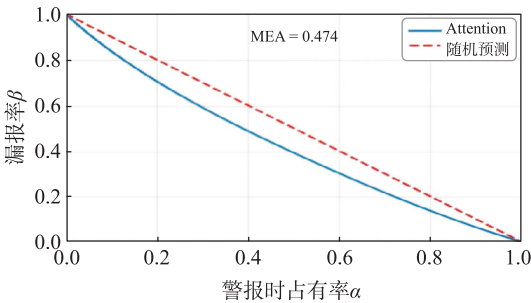


图 2 Attention-LSTM 模型的 Molchan 错误图
Fig.2 Molchan error diagram for the Attention-LSTM model

3.2 空间预测结果

依据表 1 的整体表现与可用性,仅保留 3 个代表性深度模型(LSTM-Spatial、Attention-Spatial 和 Transformer-Spatial)开展空间分析。为定量比较其空间预测能力,表 2 给出了各模型的空间命中率(取值 0~1,数值越大表示空间定位越准确)。结果显示,Attention-Spatial 的空间命中率最高(0.545),优于其他模型,说明注意力机制能够有效提升空间依赖建模与关键区域对齐能力;LSTM-Spatial 表现居中;Transformer-Spatial 具备一定的全局依赖建模能力,但总体命中率仍不及 Attention-Spatial。空间命中率与

Molchan 错误面积(MEA)在不同维度上反映模型的空间告警效率,两者可形成互补:空间命中率衡量定位精度,而 MEA 则衡量漏报与告警覆盖率之间的平衡。二者的一致性进一步支持 Attention-Spatial 在空间预测任务中的优势。

表 2 不同模型在空间预测任务中的性能对比
Tab.2 Performance comparison of different models in spatial prediction tasks

模型名称	空间命中率	模型特点
LSTM_Spatial	0.357	基础 LSTM 空间建模
Attention_Spatial	0.545	注意力增强空间预测
Transformer_Spatial	0.403	全局依赖捕捉能力

表 2 的结果表明,Attention-Spatial 模型在空间定位能力上最为突出,能够为区域地震危险性评估提供更为可靠的参考。

图 3 展示 Attention-Spatial (即 Attention-LSTM 模型的空间版)对中国东北地区未来 60 d 地震发生概率场的空间预测结果。由图可见,高概率带与区域主要活动构造带(如郯庐断裂带、燕山-渤海地震带)对应性良好,表明模型能够刻画该区地震活动的空间规律;同时,部分高值区与近年中强震位置相邻或重合,从独立证据侧面印证了其时空联合预测的有效性。

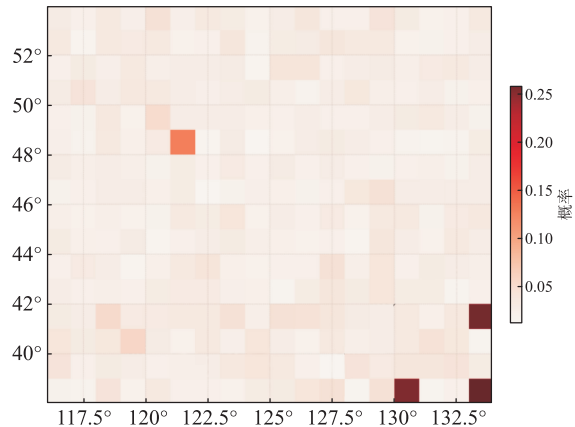


图 3 Attention-LSTM 模型预测的中国东北地区未来 60 d 地震发生概率空间分布

Fig.3 Spatial distribution of earthquake occurrence probability in northeast China for the next 60 days predicted by the Attention-LSTM model

上述空间化结果既可为地震危险性分区提供定量依据,也有助于监测资源优化配置(如台网加密与重点监测区划定),并对应急管理中长期风险评估具有潜在应用价值。

3.3 模型可解释性分析

为了理解模型的决策机制,利用 SHAP 方法对 Attention-LSTM 模型进行特征重要性分析。图 4 展示了对模型预测影响最大的前 15 个特征。

结果显示,日最大震级、地震能量总和以及到目标区域的最小距离是影响预测的最关键因素。这与地震学的普遍认知相符,即源区发生大震级、高能量释放的事件,且距离目标区越近,其产生影响的可能性越大。此外,一些时序特征如多天滑动平均地震数和余震相关特征也显示出较高的重要性,表明模型成功学习到了地震活动的时间演化模式。

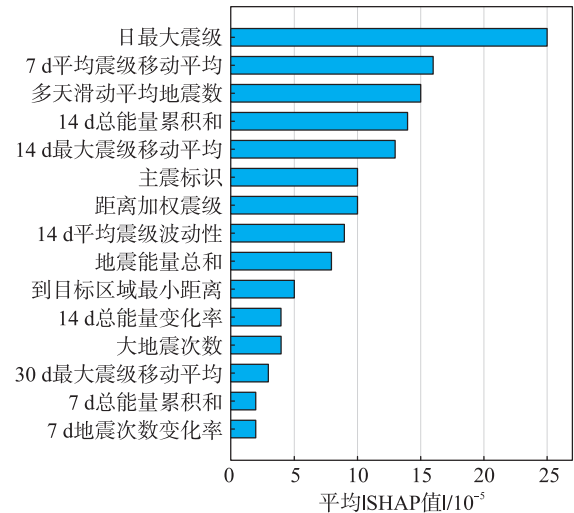


图 4 基于 SHAP 值的 Attention-LSTM 模型特征重要性排序

Fig.4 Feature importance ranking for the Attention-LSTM model based on SHAP values

图 5 展示了 Attention-LSTM 模型在分类任务中的整体表现。图 5(a) ROC 曲线(AUC = 0.661)表明,模型对正负样本具有较好的区分能力;图 5(b)P-R 曲线(AUC = 0.909)显示,在样本严重不平衡的情况下仍能保持较高的精确率和召回率;图 5(c)混淆矩阵表明,大部分正样本被成功识别,漏报和误报均较为有限;图 5(d) Molchan 错误图表明,模型 MEA = 0.474 略优于随机基线(约 0.5),提示仍有降低虚警率、优化阈值的改进空间。总体来看,Attention-LSTM 在不平衡样本条件下表现稳健,可作为跨区域地震预测的有效基线模型。

4 讨 论

本研究从多个独立维度验证了深度学习模型在捕捉远场触发信号方面的有效性:在分类任务中,LSTM 与 Attention-LSTM 在测试集上均表现出优于随机预测的性能(图 1);在地震预测专用的 Molchan 错误图中,两者的表现明显优于随机基线(图 2);在空间概率分布中,模型预测的高概率区与东北主要活动构造带具有良好一致

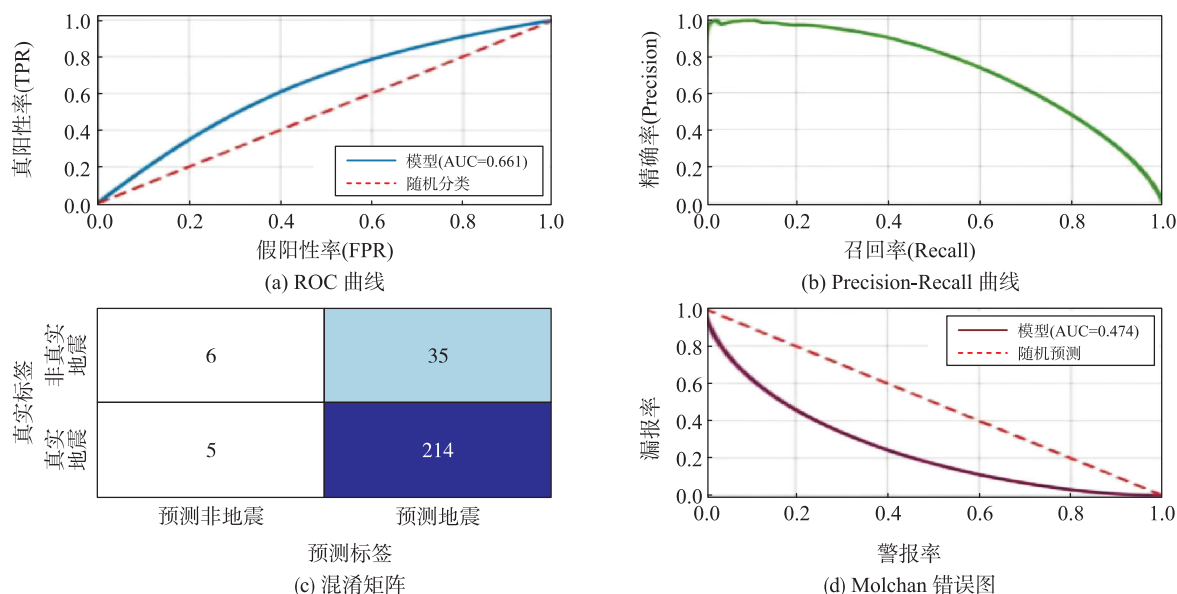


图 5 Attention-LSTM 模型的综合分类评估

Fig.5 Comprehensive classification evaluation of the Attention-LSTM model

性(图 3)。此外,注意力可视化结果显示,模型在强主震后的关键时间窗增强关注度(图 6)。这些来自分类性能、预警效能、空间分布与时间依赖性的多重证据共同表明,深度学习模型能够识别日本强震对中国东北地震活动的潜在远场触发效应。

本研究的结果初步证实了利用深度学习模型挖掘日本强震与中国东北地震活动远场触发信号的可行性。Attention-LSTM 模型之所以表现突出,可能主要有两个方面的原因:首先,LSTM 结构有效捕捉了地震活动特征的时间依赖性;其次,注意力机制使模型能够聚焦于输入序列中信息量最大的部分,例如一次强主震发生后的几天,这对于识别微弱的远场触发信号至关重要。

图 6 的注意力可视化分析揭示了模型显著的“末端聚焦”特征:如图 6(a)所示,注意力权重随时间步推进而持续攀升,在序列末端(第 12—13 天,紧邻第 14 天预测时刻前)达到峰值;图 6(b)的直方图则呈现左偏分布,绝大多数权重集中于 0.03~0.06 区间,仅少数时间步权重超过 0.10。进一步观察发现,这些高权重时间点往往对应着临近预测时刻的显著地震活动或短时聚集事件。这表明模型并未对 14 d 特征进行简单平均,而是自主地捕捉并强调了关键的触发信息,该行为模式与远场触发的物理机制假设高度一致。

相较于传统统计地震学模型(如 ETAS),本研究采用的数据驱动方法规避了对 Omori 定律等特定物理假设的依赖,转而从观测数据中直接学习复杂的非线性映射关系,这一特性在远场触

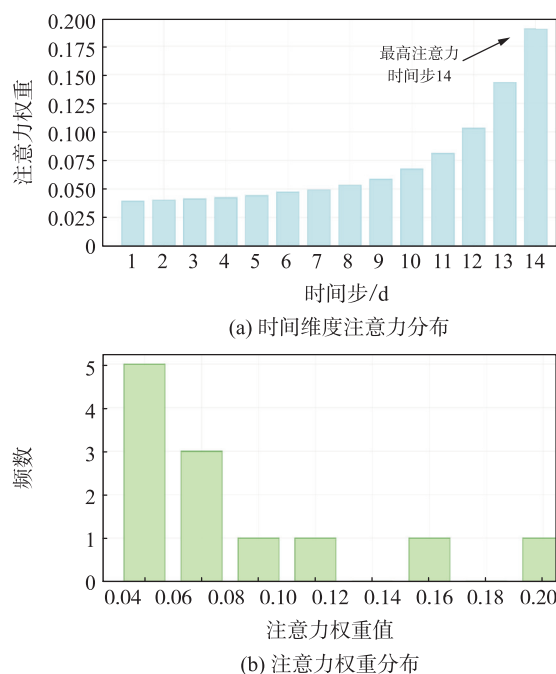


图 6 Attention-LSTM 模型在输入时间序列上的注意力权重分布

Fig.6 Attention weight distribution of the Attention-LSTM model on the input time series

发机制尚未完全厘清的背景下展现出独特优势。然而,该方法也引发了可解释性不足的问题。通过整合 SHAP 值与注意力权重分析,得以部分解析模型的内部决策机理:其结果表明,模型所关注的关键特征与地震学先验认知基本吻合(图 4),从而有效提升了对预测结果可靠性的置信度。

本研究尚存若干局限亟待完善。首先,地震目录固有的定位误差及其随时间演化的不完备性(尤其在早期观测阶段)可能引入系统性偏差。

其次,现有特征体系主要依托震例统计构建,尚未充分融合地壳形变、GNSS 观测、地下流体等孕震物理过程中的多源地球物理信息,而多模态数据融合将是提升模型预测能力的重要方向。此外,当前模型仅在“日本→中国东北”这一特定传播路径上完成验证,其不同区域组合间的可迁移性与泛化性能仍需开展系统性评估。

5 结 语

本文基于深度学习方法,开展日本强震对中国东北地区中强地震远场触发的预测建模研究,主要结论如下。

1)模型有效性验证。基于 32 维日尺度地震活动数据集构建的深度学习模型(尤其是 Attention-LSTM 模型),可有效预测未来 60 d 内中国东北地区 $M \geq 4$ 地震的发生风险,其预测性能显著优于传统机器学习方法。

2)专业评估指标佐证。在地震预测领域专用的 Molchan 错误面积(MEA)评估中,Attention-LSTM 模型表现优异,其 MEA 值显著低于随机基线(约 0.5,该指标值越小越好)。这表明该模型能够在漏报风险与虚警概率之间实现更优平衡,预测结果具备实际应用价值,而非随机猜测。

3)时空联合预测突破。模型成功实现从“是否发生地震”的定性判断到“地震在哪里发生”的定位预测跨越,可生成 $1^\circ \times 1^\circ$ 网格化的地震发生概率空间分布图,为区域地震危险性评估提供了更精细化的决策支撑信息。

4)模型可解释性解析。借助 SHAP 分析与注意力权重可视化方法,明确模型决策主要依赖于震源区震级、地震能量、震源距等物理意义明确的核心特征,以及关键触发地震事件。这一发现显著提升了模型的可信度与可解释性。

本研究构建了数据驱动方法用于跨区域地震远场触发研究的完整技术框架,并证实了该框架在该领域的应用潜力。未来研究将重点围绕 3

个方面展开:融合多源地球物理数据以丰富特征维度,探索更先进的深度学习模型架构以提升预测精度,将该方法推广应用于全球更多潜在的远场触发区域对。

致谢:感谢 USGS 提供公开地震目录数据(USGS earthquake catalog)。

参考文献

- [1] Scholz C H. The Mechanics of Earthquakes and Faulting [M]. Cambridge: Cambridge University Press, 2019
- [2] Hill D P, Reasenber P A, Michael A, et al. Seismicity Remotely Triggered by the Magnitude 7.3 Landers, California, Earthquake [J]. Science, 1993, 260 (5114): 1617-1623
- [3] Freed A M. Earthquake Triggering by Static, Dynamic, and Postseismic Stress Transfer [J]. Annual Review of Earth and Planetary Sciences, 2005, 33: 335-367
- [4] Parsons T. Global Omori Law Decay of Triggered Earthquakes: Large Aftershocks Outside the Classical Aftershock Zone [J]. Journal of Geophysical Research: Solid Earth, 2002, 107(B9)
- [5] Velasco A A, Hernandez S, Parsons T, et al. Global Ubiquity of Dynamic Earthquake Triggering [J]. Nature Geoscience, 2008, 1(6): 375-379
- [6] Brodsky E E, van der Elst N J. The Uses of Dynamic Earthquake Triggering [J]. Annual Review of Earth and Planetary Sciences, 2014, 42: 317-339
- [7] Beroza G C. Machine Learning Improves Forecasts of Aftershock Locations [J]. Nature, 2018, 560(7720): 556-557
- [8] Li D K, Xie F, Wang Q Y, et al. An Ensemble Deep Learning-Based Acoustic Emission Picking Model Reveals Migratory Foreshocks on Large-Scale Laboratory Fault [J]. Journal of Geophysical Research: Solid Earth, 2025, 130(7)
- [9] De Vries P M R, Viégas F, Wattenberg M, et al. Deep Learning of Aftershock Patterns Following Large Earthquakes [J]. Nature, 2018, 560(7720): 632-634
- [10] Zhang Y, Wen C C, Didier S, et al. Integrating Artificial Intelligence and Geophysical Insights for Earthquake Forecasting: A Cross-Disciplinary Review [EB/OL]. <https://arxiv.org/abs/2502.12161>, 2025
- [11] Molchan G. Space-Time Earthquake Prediction: The Error Diagrams [J]. Pure and Applied Geophysics, 2010, 167 (8): 907-917

Prediction on Far-Field Triggering Earthquakes in Northeast China Induced by Japanese Strong Earthquakes Based on Deep Learning

GAO Feng¹ LI Mei²

¹ Heilongjiang Earthquake Agency, Harbin 150090, China

² Institute of Earthquake Forecasting, CEA, Beijing 100036, China

Abstract: Far-field earthquake triggering mechanisms represent a critical frontier in earthquake predic-

tion research. To explore the spatiotemporal relationship between strong seismic activity in Japan and moderate-to-strong earthquakes in northeast China, this study develops a data-driven prediction model. Based on the USGS earthquake catalog from 1980 to 2024, and using $M \geq 6.0$ earthquakes in Japan as potential triggers, the model predicts the probability of earthquakes with $M \geq 4$ occurring in northeast China within the next 60 days. A daily-resolution time series dataset is constructed, incorporating 32-dimensional features covering earthquake statistics, spatial distribution, energy release, and aftershock sequences. Deep learning models including long short-term memory (LSTM) neural networks, attention mechanism-enhanced LSTM (Attention-LSTM), and Transformer are systematically compared against traditional machine learning methods such as logistic regression and random forest. Results indicate that the Attention-LSTM model performs optimally, achieving an F1 score of 0.915 and an AUC value of 0.661, and shows significant advantages in Molchan error diagram analysis, a method specific to earthquake prediction evaluation. The model enables both binary classification prediction and generation of spatial probability distribution maps at $1^\circ \times 1^\circ$ grid resolution. This study demonstrates the potential of deep learning for revealing cross-regional far-field earthquake triggering mechanisms and provides new insights for short-to-medium-term regional seismic hazard assessment.

Key words: deep learning; attention-based long short-term memory network (Attention-LSTM); far-field earthquake triggering; spatiotemporal earthquake forecasting

(上接第 667 页)

[18] Husár L, Švaral P, Janák J. About the Geometry of the Earth Geodetic Reference Surfaces[J]. Journal of Geometry

and Physics, 2017, 120: 192-207

Coordinate Inverse Transformation for the Triaxial Geoid Model

ZHANG Hanwei¹ XU Pengfei¹ WANG Ke¹

¹ School of Civil Engineering and Geomatics, Shandong University of Technology, Zibo 255000, China

Abstract: By employing fundamental methods from the theory of curves and surfaces, the concepts of different types of latitude and longitude (reduced latitude and longitude, geometric latitude and longitude, and geodetic latitude and longitude) at an arbitrary point on a triaxial ellipsoidal surface are proposed. Theoretical relationship models are established between Cartesian coordinates (spatial rectangular coordinates) and geodetic coordinates (geodetic latitude, longitude, and height), as well as between Cartesian coordinates and geometric coordinates (geometric latitude, longitude, and height), for any point in space both inside and outside the ellipsoid. Based on the theoretical relationships among the different latitude and longitude types, approximate geodetic coordinates of the study point are derived. Using these approximations as initial values, a novel method for the inverse transformation of coordinates is presented via Newton's iterative approach. Extensive numerical calculations demonstrate that, for any arbitrary point in space, the proposed method accomplishes the transformation from Cartesian to geodetic coordinates nearly instantaneously; even for points near the Earth's surface, only three iterations are required to achieve convergence.

Key words: Earth ellipsoid; reduced latitude and longitude; geometric latitude and longitude; geodetic latitude and longitude; geodetic coordinates