

基于 Stacking 模型融合策略的日本俯冲带板缘地震动预测

党浩天¹, 王自法^{2,3,4}, 赵登科^{2,3}, 位栋梁⁴, 王祥琪¹, WANG Jianming⁴, 李兆焱^{2,3}

(1. 河南大学 土木建筑学院, 河南 开封 475001; 2. 中国地震局工程力学研究所, 黑龙江 哈尔滨 150080;
3. 地震灾害防治应急管理部重点实验室, 黑龙江 哈尔滨 150080; 4. 中震科建(广东)防灾减灾研究院, 广东 韶关 512000)

摘要:高精度的地震动预测模型有助于提高地震灾害的预警和应对能力。传统回归方法构建地震动预测模型时提前设定了方程的形式,此种方法存在一定局限性,难以反映地震动传播过程中的复杂规律,因此越来越多的学者尝试应用机器学习方法构建地震动预测模型。但采用单一的机器学习算法,难以从数据中捕捉到更多规律,最终导致模型精度难以提升。本文基于日本 KiK-net 和 K-Net 强震台网收集到的俯冲带板缘地震动记录,使用 Stacking 模型融合策略,以 LightGBM、XGBoost 和 CatBoost 算法作为基学习器,线性回归算法作为元学习器,引入客观且高效的贝叶斯优化算法对模型进行超参数优化,最终训练并提出了一种适用于日本俯冲带板缘地震动预测的融合模型 Stacking-Interface。对比分析所提出模型、单一机器学习模型和传统模型,发现机器学习模型的精度普遍高于传统模型,且相较于单一的机器学习模型,融合模型的预测能力有一定的提升;通过与实际地震动记录的对比和特征参数敏感性分析,验证了所提模型的可靠性和泛化能力。研究方法和结果能够为地震风险分析提供参考。

关键词:地震动预测;Stacking;俯冲带板缘地震;部分依赖图

中图分类号:P315.9 **文献标识码:**A

Ground motion prediction based on the Stacking model fusion strategy for the Japanese subduction interface earthquake

DANG Haotian¹, WANG Zifa^{2,3,4}, ZHAO Dengke^{2,3}, WEI Dongliang⁴, WANG Xiangqi¹,
WANG Jianming⁴, LI Zhaoyan^{2,3}

(1. School of Civil Engineering and Architecture, Henan University, Kaifeng 475001, China; 2. Institute of Engineering Mechanics, China Earthquake Administration, Harbin 150080, China; 3. Key Laboratory of Earthquake Disaster Mitigation, Ministry of Emergency Management, Harbin 150080, China; 4. CEAJK ADPRHexa Inc., Shaoguan 512000, China)

Abstract: High-precision earthquake motion prediction models contribute to the improvement of earthquake disaster early warning and response capabilities. Traditional regression methods for constructing earthquake motion prediction models rely on predefined equations, which have certain limitations and struggle to capture the complex patterns of seismic wave propagation. As a result, more and more scholars have been exploring the application of machine learning methods in constructing earthquake motion prediction models. However, using a single machine learning algorithm makes it difficult to capture more patterns from the data, leading to limited improvement in model accuracy. In this study, we utilized ground motion records collected from the KiK-net and K-Net strong motion networks in Japan. This study employed a Stacking model fusion strategy, with LightGBM, XGBoost, and CatBoost algorithms as base estimators and linear regression as the meta-estimator. Additionally, an objective and efficient Bayesian

收稿日期:2023-06-02; 修回日期:2023-09-27

基金项目:中国地震局工程力学研究所基本科研业务费专项资助项目(2021B09);国家自然科学基金面上项目(51978634)

作者简介:党浩天(1998—),男,硕士研究生,主要从事地震动模型相关研究。E-mail:heroic98@163.com

通信作者:王自法(1965—),男,研究员,博士,主要从事巨灾风险相关研究。E-mail:zifa@iem.ac.cn

optimization algorithm was introduced to optimize the model's hyperparameters. The proposed approach resulted in a ground motion prediction fusion model called Stacking-Interface, specifically designed for subduction interface earthquakes in Japan. Comparative analysis of the proposed model, single machine learning models, and traditional models revealed that machine learning models generally outperformed traditional models in terms of accuracy. Furthermore, the fusion model demonstrated improved predictive capabilities compared to single machine learning models. The reliability and generalization ability of the proposed model was validated through comparisons with actual ground motion records and sensitivity analysis of feature parameters. The research methodology and results presented in this paper can serve as a reference for earthquake risk analysis.

Key words: ground motion prediction; Stacking; subduction interface earthquakes; partial dependence plot

0 引言

俯冲带地区的板块构造复杂,发震频繁^[1],给人们的生命和财产安全带来了巨大的威胁。限于科学技术发展,目前无法较为精确预测地震的发生,但能够通过地震动预测模型,对峰值加速度(peak ground acceleration, PGA),谱加速度(spectral acceleration, SA)等地震动强度参数进行估算,为地震安全性评估、地震危险性分析和工程设计提供一定参考,也可以为地震发生后的应急救援工作提供指导意见^[2-3],因此高精度的地震动预测模型至关重要。

地震动预测模型也称地震动衰减关系,传统模型构建方法是预先设定方程形式,通过历史地震动数据进行数学回归计算方程内的各个参数,最终得到显式方程。1961年,GUTENBERG等^[4]采用美国南加利福尼亚州的地震数据,研究了地面加速度与震级、距离的关系,开辟了地震动预测的先河。随着相关研究的不断深入,发现不同构造类别(浅层壳、上地幔、俯冲带板缘和俯冲带板内)地震的反应谱差异很大,采用统一的地震动预测模型是不可行的^[5-6]。因此地震动预测模型的研究逐渐朝着针对于某一构造类别地震的方向发展。美国下一代地震动衰减关系(next generation attenuation, NGA)项目开发出五组适用于浅层壳地震的地震动预测模型^[7-11],且随项目发展,NGA针对俯冲带开发了适用于全球多个国家和地区的地震动预测模型^[12]。赵兴权等^[13-15]基于日本地震动记录,开发了适用于不同地震构造类别的地震动预测模型 Zhao2016;2021年,肖亮等^[16]借鉴《中国地震动参数区划图》(GB 18306—2015)^[17]分段线性模型,针对我国海域实际地震环境,开发了重点关注俯冲带长周期地震动的地震动预测模型。然而,由于传统回归方法提前设定了方程的形式,在拟合大样本数据的非线性关系时存在一定局限性,难以反映地震动衰减的复杂规律,因此还有较大的改进空间。

机器学习是基于数据和统计学习理论的人工智能分支,旨在让计算机通过算法学习数据规律,从而实现自我优化。随着大量地震观测数据的积累,机器学习方法也被应用在地震动预测模型的开发中。VEMULA^[18]采用神经网络算法开发了适用于新西兰等地区的地震动预测模型,其预测精度在适用地区的误差允许范围内;SEO等^[19]基于韩国地震动数据,分析对比了人工神经网络算法、随机森林算法和梯度提升算法的预测能力,结果表明在该地区数据中,梯度提升算法的预测精度较高;籍多发等^[20]提出了一种修正的二阶神经元,改进了深度学习算法,使模型有更强的非线性拟合能力,从而达到更高的预测精度。

虽然机器学习方法的预测精度相较于传统方法有一定提升,但仍存在以下几个问题:

- 1) 单一机器学习算法所能捕捉到的数据规律有限,因此提升精度变得较为困难。
- 2) 机器学习模型中的超参数对预测结果起着决定性作用,根据人为经验调整超参数存在一定的主观性,且随着超参数的增多,调整变得更为困难。

3) 由于机器学习算法的复杂性,其可解释性较差,因此被称为黑箱模型,造成结果的认可度不高。

基于以上不足,本文采用日本 KiK-net 和 K-Net 强震台网 2005—2019 年在 2344 次俯冲带板缘地震事件中收集到的 122 352 条地震动记录,按照按 70%、15% 和 15% 的比例划分为训练集、验证集以及测试集,分别用于模型训练、超参数优化和模型测试。以日本气象厅震级(M_{JMA})、震源距(R_{hypo})、震源深度(D)和地下 30 m 平均剪切波速(V_{s30})四个地震动特征作为模型的输入特征参数,PGA 以及 SA 的自然对数作为模型的输出,首先对比分析多种单一机器学习算法的结果,从中选取结果较优的 LightGBM、XGBoost 和 CatBoost 三种算

法作为基学习器^[21-24],选用线性回归算法作为元学习器,采用客观且高效的贝叶斯优化算法对模型进行超参数优化,共同构建 Stacking 融合模型,最终训练并提出了适用于日本俯冲带板缘地震动预测的融合模型 Stacking-Interface。此外,对比了所提模型、单一机器学习模型和传统模型的精度,测试了模型的泛化能力,并引入部分依赖图解释了特征参数对模型输出结果的影响机理,分析所提出模型的合理性。

1 数据集设置

1.1 数据来源

KiK-net 和 K-Net 是由日本防灾科学研究所(National Research Institute for Earth Science and Disaster Resilience, NIED)(<https://www.kyoshin.bosai.go.jp>)建设和管理的强震观测网络,该网络中的每个台站提供三个方向地震动数据,分别为东西方向(EW)、南北方向(NS)和垂直方向(UD)。此外,日本防灾科学研究所提供每个台站详细的钻孔信息,包括土层结构和土层剪切波速图。

从 KiK-net 和 K-Net 获取到的地震动数据,其中包含了多种构造类型的地震数据,为获取本文所需要的俯冲带板缘地震动数据,按照以下条件进行数据筛选:

- 1) 对于 KiK-net 台站只选取井上记录数据(自由场地)。
- 2) 参考赵兴权等^[5,25]对地震构造的分类方法,挑选出震源位于海底且震源深度 $D \leq 50$ km 的俯冲带板缘地震。
- 3) 震级 $M_{JMA} \geq 4$ 级。
- 4) 当震级 $M_{JMA} < 5$ 级时,选取震源距 $R_{hypo} \leq 200$ km 的地震动数据;当震级 $M_{JMA} \geq 5$ 级时,选取震源距 $R_{hypo} \leq 300$ km 的数据。
- 5) 对 KiK-net 和 K-Net 获取到的加速度时程数据进行基线调整,使用低通截止频率 30 Hz 和高通截止频率 0.2 Hz 的四阶 Butterworth 滤波器对数据进行滤波处理,通过积分计算^[26],获取速度时程曲线和位移时程曲线,删除曲线出现基线漂移的地震动加速度记录。

基于以上条件,筛选出来自 2 344 次俯冲带板缘地震的 122 352 条地震动记录,数据集中地震事件的震级、震源深度以及台站震源距的分布见图 1。

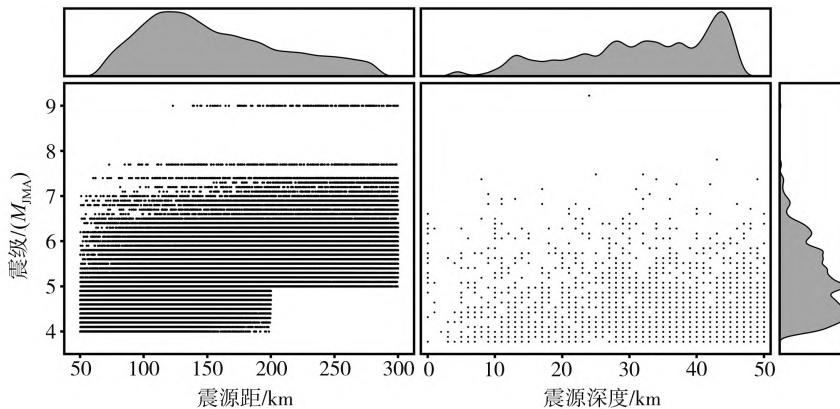


图 1 地震动数据分布情况,侧边图为数据概率密度分布情况

Fig. 1 Distribution of ground motion data, with the side plot showing the probability density distribution of the data

1.2 数据处理流程

为获取训练模型所需要的数据,经筛选后的地震动记录按照以下流程进行处理:

- 1) 根据日本防灾科学技术研究所提供的场地钻孔资料按照式(1)计算每个台站的 V_{s30} 值^[27]。

$$V_{s30} = 30 / \sum_i^n \frac{d_i}{V_{si}} \quad (1)$$

式中: i 为地下 30 m 内的土层数量; d_i 为第 i 层土厚度; V_{si} 为第 i 土层的剪切波速。当钻孔深度不足 30 m 的时候,按照外推公式去计算 V_{s30} 值^[28]。

- 2) 计算两个水平分量(南北分量和东西分量)的 PGA 和 SA 几何平均值,其中 SA 为在 0.01~5 s 之间的

36 个周期点,以 5% 的阻尼比计算谱加速度。

3) 本文选取传统方法模型 Zhao2016 俯冲带板缘地震动预测模型作为对比模型,由于 Zhao2016 模型中所采用的震级单位为矩震级 (M_w),因此本文依据如下原则进行震级转换。先在 GCMT (global centroid-moment-tensor) 获取本文所用地震事件对应的矩震级 M_w ,若本文所用地震事件在 GCMT 无记录,则根据震级转换模型进行转换^[29]。

$$M_w = 0.923 \times M_{JMA} + 0.370 \quad (2)$$

2 模型融合与超参数优化

2.1 Stacking 模型融合

Stacking 是一种模型融合策略^[30],通过将多个不同机器学习算法融合,形成具有多层结构的模型,其基本原理为选用预测精度较高且不同的机器学习算法作为基学习器,进行 K 折交叉验证,将基学习器的预测结果作为新的特征输入到元学习器中进行二次学习。Stacking 的关键优势在于其本身能够根据不同基学习器的优劣进行有效的组合,克服单个基学习器的局限性,从而达到各模型优势互补的目的,以此来提升预测精度。

本文初步选择的基学习器算法为 LightGBM、XGBoost、CatBoost、随机森林和深度神经网络 (deep neural net, DNN)^[31-32],通过对不同机器学习算法在验证集中的精度,选择精度较高的机器学习算法作为基学习器。此外,为了防止模型过拟合且考虑计算设备性能(本文设备配置:Intel(R) Core(TM) i9-10900K CPU @ 3.70GHz 3.70GHz, NVIDIA GeForce RTX 2080 Ti),本文选择线性回归算法为元学习器。

本文所用特征参数为震级 (M_{JMA})、震源深度 (D)、震源距 (R_{hypo}) 和 V_{s30} ,模型输出结果为 PGA 或 SA 的自然对数,单一模型的基本形式如下公式所示。

$$\ln(\hat{Y}) = f(M_{JMA}, R_{hypo}, D, V_{s30}) \quad (3)$$

式中: \hat{Y} 是 PGA 或 SA 的预测值; $f(\cdot)$ 代表机器学习算法,即 LightGBM、XGBoost、CatBoost、随机森林和 DNN。

2.2 超参数优化

机器学习算法中有众多超参数,不同的超参数对模型输出结果有不同的影响,如在 XGBoost 算法中,正则化系数指定了树节点分裂所需的最小损失减少量,设置过大导致模型欠拟合,因此如何找到一组最优超参数组合是提升模型精度的关键一步。根据人为经验的超参数调整较为主观,且随着模型超参数的增多,调整超参数的过程变得困难,虽然网格搜索可作为辅助方法进行调整,但其本质是在给定的超参数区间内对不同超参数组合进行枚举,所花费的时间成本较大。因此本文引入两种不依赖人为经验的自适应超参数调节方法:贝叶斯优化算法和遗传优化算法^[33-34],由于 LightGBM 算法具有计算速度快的特点,因此通过对 LightGBM 算法进行案例测试,综合考虑计算效率和计算精度,确定本文的超参数调整方法。

本文设定优化目标 MSE_{obj} ,即 $\ln(PGA)$ 和 $\ln(SA)$ 真实值与预测值的均方误差平均值,如公式(4)所示,优化的目标为搜寻到一组使验证集的 MSE_{obj} 最小的超参数组合。

$$MSE_{obj} = \frac{1}{37} (\text{MSE}_{\text{PGA}} + \sum_{i=1}^{36} \text{MSE}_{\text{SA}_i}) \quad (4)$$

式中: MSE 是均方误差,由式(5)计算, I 代表本文所选 36 个周期中的第 I 个周期。

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (5)$$

式中: y 和 \hat{y} 分别为 $\ln(PGA)$ 或 $\ln(SA)$ 的真实值和预测值; n 为数据总数。

表 1 为贝叶斯优化算法与遗传优化算法在 LightGBM 中优化所需时间与精度,结果表明:搜索次数均为 200 次时,两种超参数优化方法在搜索用时较为接近的情况下,贝叶斯优化算法能够达到较高的预测精度,原因是由于贝叶斯优化算法使用代理模型拟合真实目标函数,并根据拟合结果主动选择最有“潜力”的评估点进行评估,能够有效地利用完整历史信息来提高搜索效率,避免不必要的采样^[35]。遗传优化算法是模拟自然界生物进化过程的一类自组织和自适应算法,存在随机游走的特点^[36],所以相较贝叶斯算法,遗传算法在部分情况下可能达不到理想的效果,因此本文选择贝叶斯优化算法作为超参数优化方法。

采用贝叶斯优化对 LightGBM、XGBoost、CatBoost、随机森林和 DNN 五种机器学习算法进行超参数优化,优化后各算法的超参数及其 MSE_{obj} 见表 2,可以发现:在日本俯冲带板缘地震动验证数据集中,LightGBM 算法的预测精度最高, MSE_{obj} 为 0.3448,其次是 XGBoost、CatBoost、随机森林和 DNN,后 4 种方法相较 LightGBM, MSE_{obj} 分别增加了 7.83%、15.13%、32.27% 和 52.04%。分析结果表明:随机森林和 DNN 算法的误差较大,因此本文选择 LightGBM、XGBoost 和 CatBoost 作为基学习器。

表 1 遗传优化算法与贝叶斯优化算法耗时与 MSE_{obj} 对比结果

Table 1 Comparison of time consumption and MSE_{obj} between genetic optimization algorithm and Bayesian optimization algorithm

优化算法	搜索次数	耗时/h	MSE_{obj}
贝叶斯优化算法	200	10	0.3448
遗传优化算法	200	9.5	0.3721

表 2 单个机器学习算法超参数及 MSE_{obj}

Table 2 Hyperparameters and MSE_{obj} for individual machine learning algorithms

机器学习算法	超参数	超参数值	MSE_{obj}
LightGBM	learning_rate	0.1850150	0.3448
	max_depth	10.0000000	
	n_estimators	981.0000000	
	num_leaves	44.0000000	
	reg_alpha	2.4762810	
	reg_lambda	40.0251030	
XGBoost	gamma	0.0013330	0.3718
	learning_rate	0.6738010	
	max_depth	6.0000000	
	n_estimators	379.0000000	
	reg_alpha	10.6091520	
	reg_lambda	15.0452820	
CatBoost	depth	11.0000000	0.3976
	iterations	769.0000000	
	l2_leaf_reg	19.7868590	
	learning_rate	0.2567410	
随机森林	max_depth	88.0000000	0.4561
	min_samples_leaf	1.0000000	
	min_samples_split	6.0000000	
	n_estimators	934.0000000	
DNN	batch_size	1024.0000000	0.5241
	dense	6.0000000	
	activation	elu	
	learning_rate	0.0013844	

2.3 模型融合

通过不同机器学习算法对比,本文选择 LightGBM、XGBoost 和 CatBoost 作为基学习器,Stacking 融合模型结构如图 2 所示,首先将训练数据集随机划分成五份数量相等但不重复的子集进行五折交叉验证,其中四份子集作为训练子集用于基学习器训练,剩余一份子集作为验证子集,使用训练完毕的基学习器计算出验证子集预测值。三个基学习器分别经过五折交叉验证,即可获取训练集所对应的三份预测值,将预测值作为新特征,使用元学习器学习三份预测值与真实值之间的数据规律,最终训练出融合模型 Stacking-Interface。

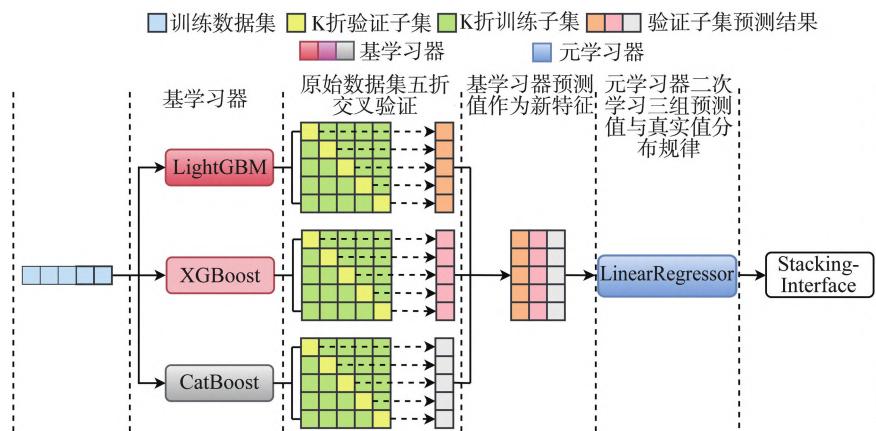


Fig. 2 Architecture of the Stacking Fusion Model

3 结果

由于机器学习隐性函数的特点,目前尚未有公开的适用于日本俯冲带板缘地区的机器学习地震动预测模型,因此本文仅引入传统模型 Zhao2016 作为对比目标^[15],随后将经贝叶斯优化后的超参数配置至各个算法中。对比 Zhao2016 模型,单个机器学习算法模型与 Stacking-Interface 模型在测试集中的精度,观察真实值与模型预测值分布情况、残差分别随震级与震源距变化的分布情况,并使用均方误差(MSE)、皮尔逊相关系数(R)与标准差(σ)作为评价指标,对比各模型的预测性能,残差按照式(6)计算,

$$\delta = \ln(y) - \ln(\hat{y}) \quad (6)$$

此外,参考传统模型所使用的随机效应回归法的评估方法^[37],按照式(7)将残差分解为事件间残差和事件内残差进行对比。

$$\ln y_{ij} = \ln \hat{y}_{ij} + \eta_i + \varepsilon_{ij} \quad (7)$$

式中: η_i 是地震事件 i 的事件间残差; ε_{ij} 是事件 i 中记录 j 的事件内残差。

为进一步分析场地效应 V_{s30} 对残差分布的影响,将事件内残差分解为场地间残差和场地内残差,即:

$$\varepsilon = \gamma_{m,n}^s + \eta_m^s \quad (8)$$

式中: $\gamma_{m,n}^s$ 为第 m 个场地内的第 n 个记录场地内残差; η_m^s 为第 m 个场地的场地间残差。

3.1 残差分布

图 3 为不同算法模型真实值与预测值的分布情况(PGA、 $T=0.3$ s 和 $T=1$ s),其中 Stacking-Interface、LightGBM、XGBoost 和 CatBoost 所预测结果与真实结果在对角线两侧均匀分布,没有较为明显的偏差趋势,预测结果表现一致且可靠。从图 3 残差概率密度曲线可以看出残差分布符合以 0 为均值的正态分布,符合残差的一般分布特征,而 Zhao2016 模型在不同周期中均出现了的“高估小值,低估大值”情况。从图 4~图 5 可以看出:机器学习方法的残差几乎不受震级或震源距的影响,说明在不同特征的输入下,机器学习方法都能够给出较为可靠的预测结果,而 Zhao2016 模型随震级或震源距的改变残差出现了一定变化,预测结果较不稳定。

为进一步分析残差分布情况,将总残差按照式(7)和式(8)分解为事件间残差、事件内残差、场地间残差和场地内残差。图 6 为事件间残差随震级变化的情况,可以看出:在震级较大时,不同周期下,各机器学习模型的事件间残差随震级的变化均匀地分布在 0 左右,没有产生明显的系统性偏差趋势,表明事件间残差与震级没有明显的关系,一定程度上说明在大震与小震下均有良好的适用性,而 Zhao2016 模型的事件间残差在震级较小时,出现了不均匀分布现象,原因可能是由于 Zhao2016 模型拟合模型时所用数据为 5 级以上,造成了此种偏差。图 7 为事件内残差随震源距分布情况,可以看出:各模型变化都较为平稳,但机器学习模型的残差分布更为收敛。图 8~图 9 为场地间残差与场地内残差随 V_{s30} 的变化情况,总体上各模型变化都较为均匀,未见明显偏差。

从残差数值分布来看:机器学习模型残差数值分布在 [-2, 2] 区间内的占比均在 99.56% 以上,Zhao2016 模型占比约为 96.54%,此外,对于残差数值分布在 [-1, 1] 内的数据占比,Stacking-Interface、LightGBM、XGBoost 和 CatBoost 分别约为 92.22%、91.94%、91.19% 和 90.06%,Zhao2016 占比约为 72.41%。

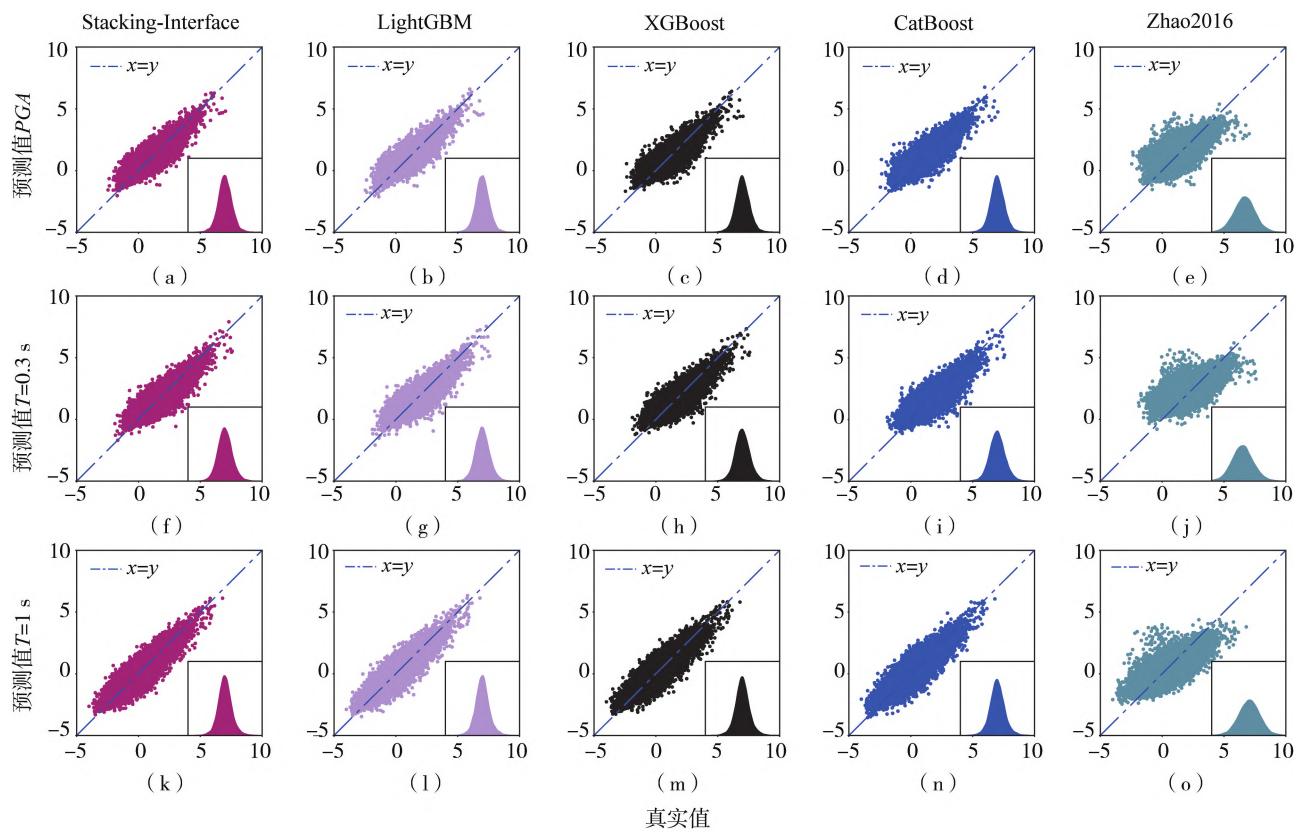


图 3 不同算法模型真实值与预测值分布情况, 图右下角为残差的概率密度曲线

Fig. 3 Distribution of the true and predicted values of different algorithms models, with the residual probability density curve in the lower right corner of the figure

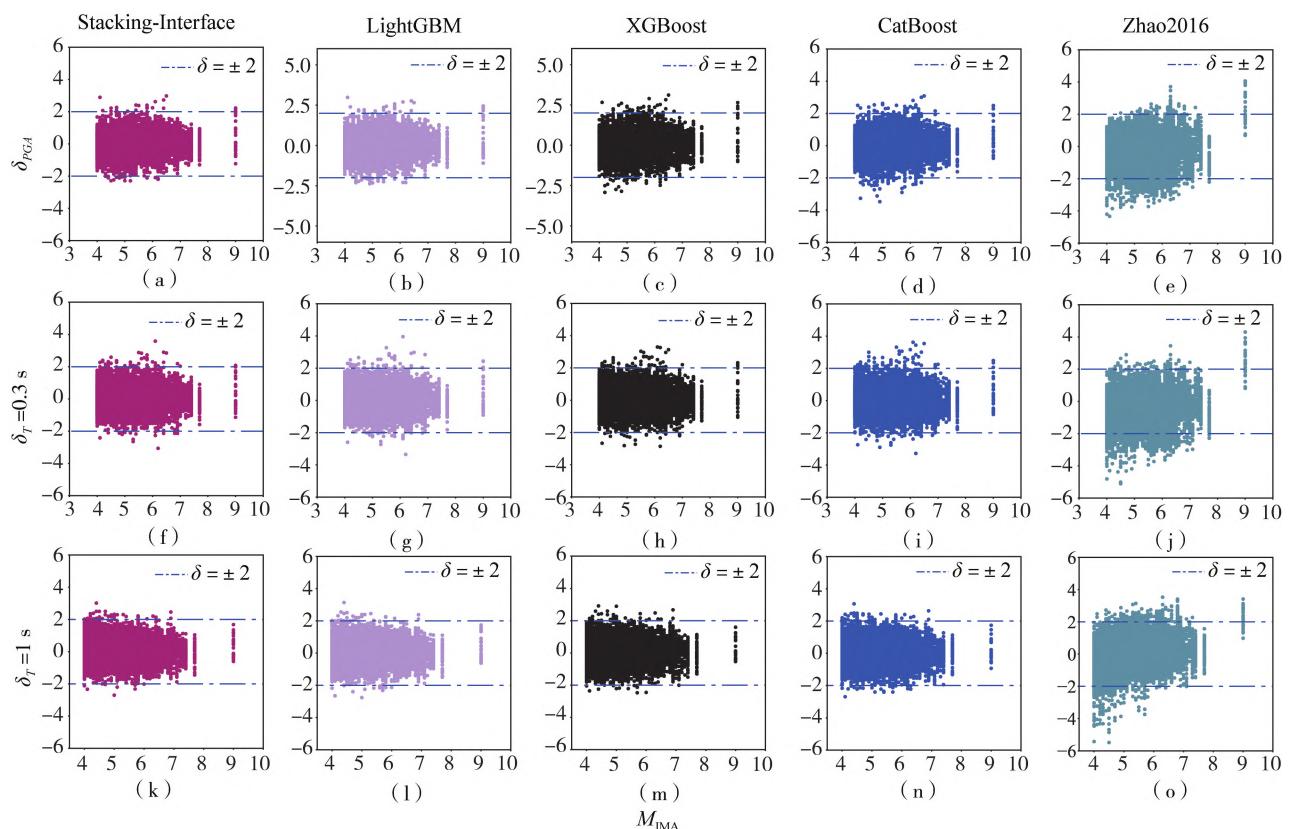


图 4 不同算法模型真实值与预测值残差随震级变化分布情况

Fig. 4 Distribution of the residuals between true and predicted values of different algorithms models with the variation of M_{JMA}

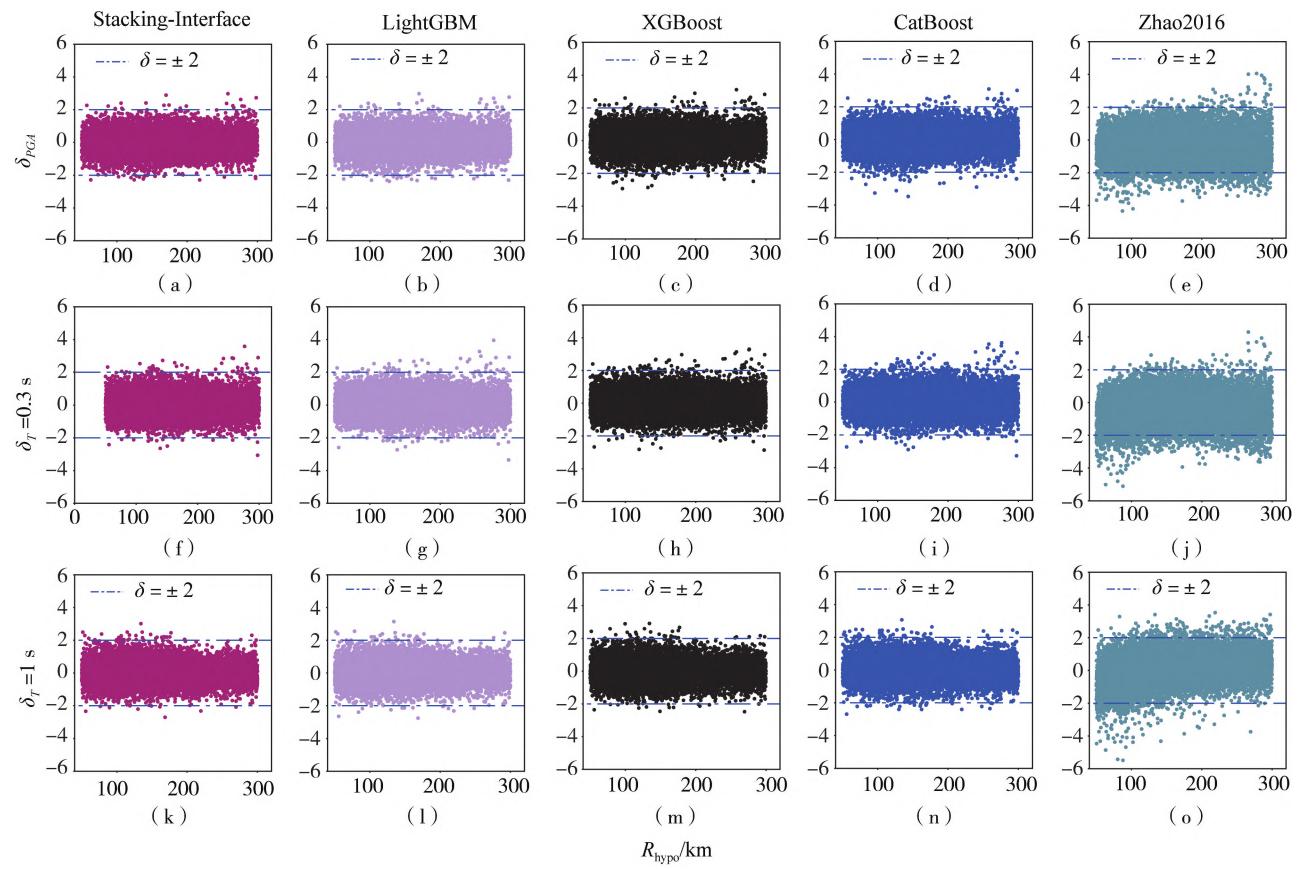


图5 不同算法模型真实值与预测值残差随震源距变化分布情况

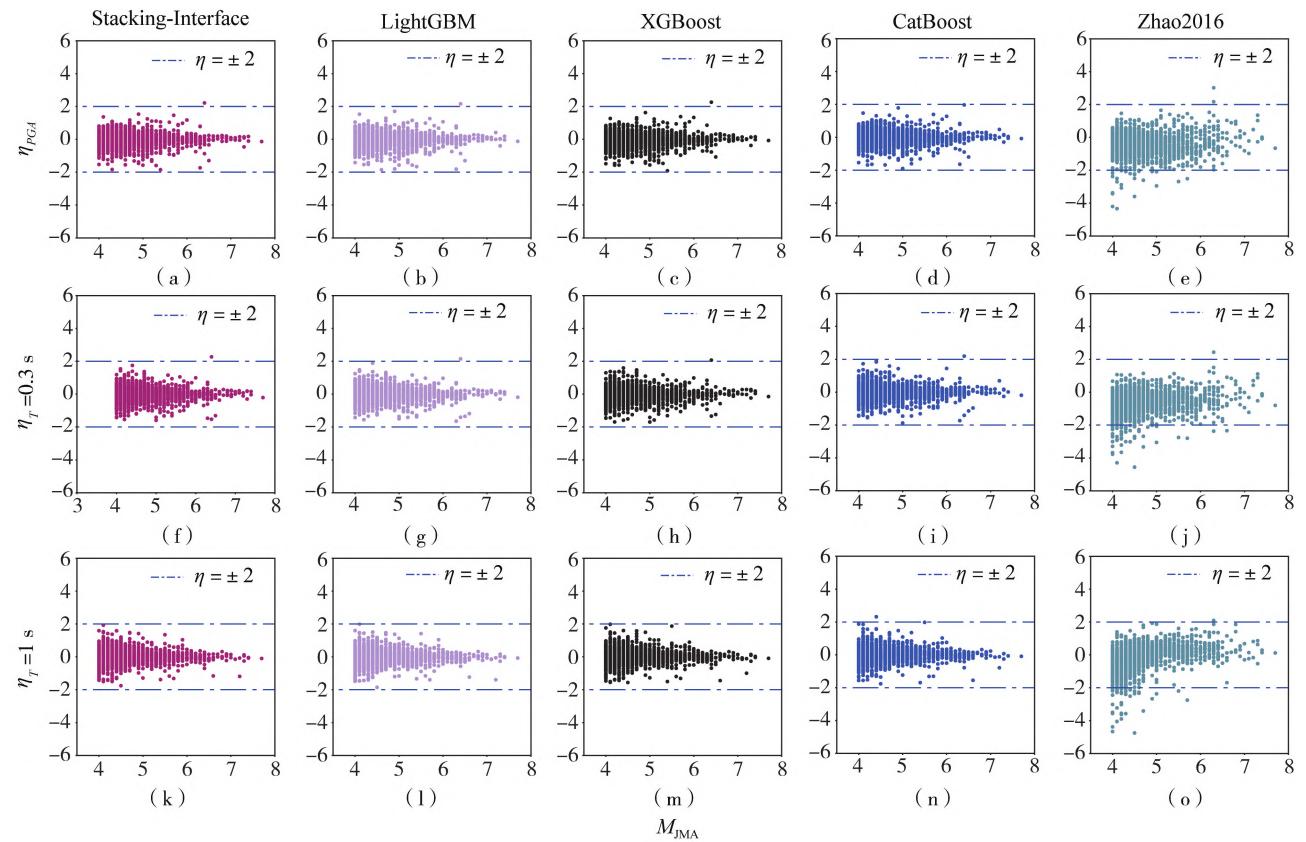
Fig. 5 Distribution of residuals between true and predicted values of different algorithms models with the variation of R_{hypo} 

图6 不同算法模型事件间残差随震级变化分布情况

Fig. 6 Distribution of inter-event residual of different algorithms models with the variation of M_{JMA}

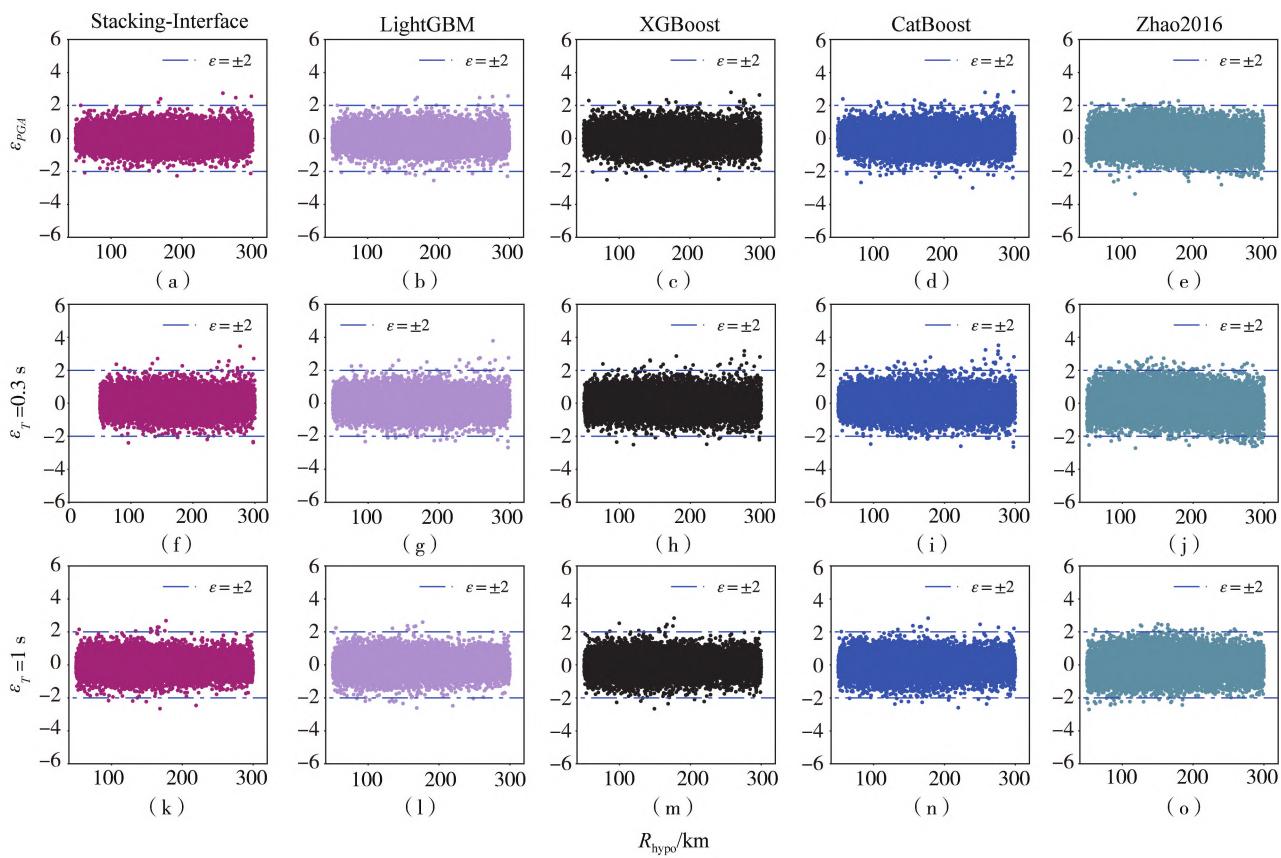


图7 不同算法模型事件内残差随震源距分布情况

Fig. 7 Distribution of intra-event residual of different algorithms models with the variation of R_{hypo}

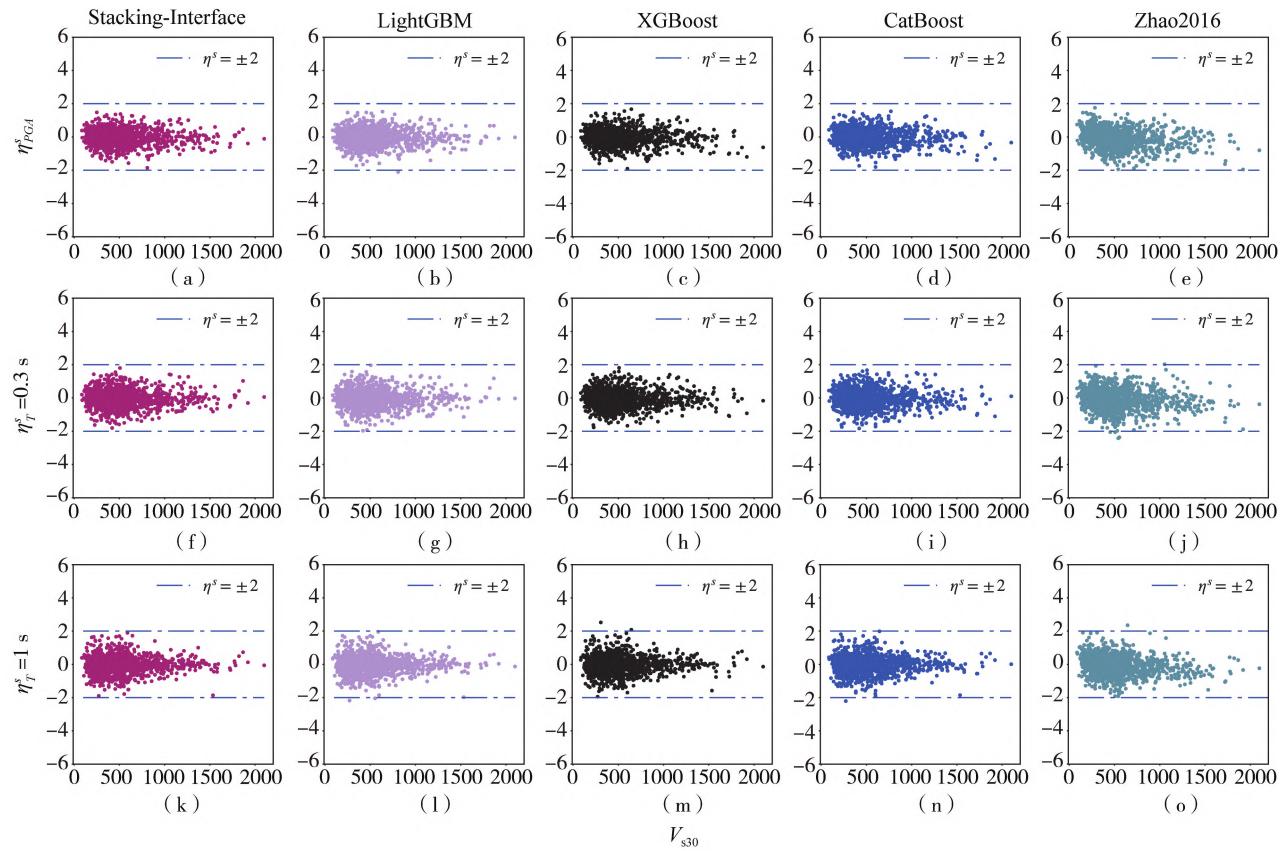
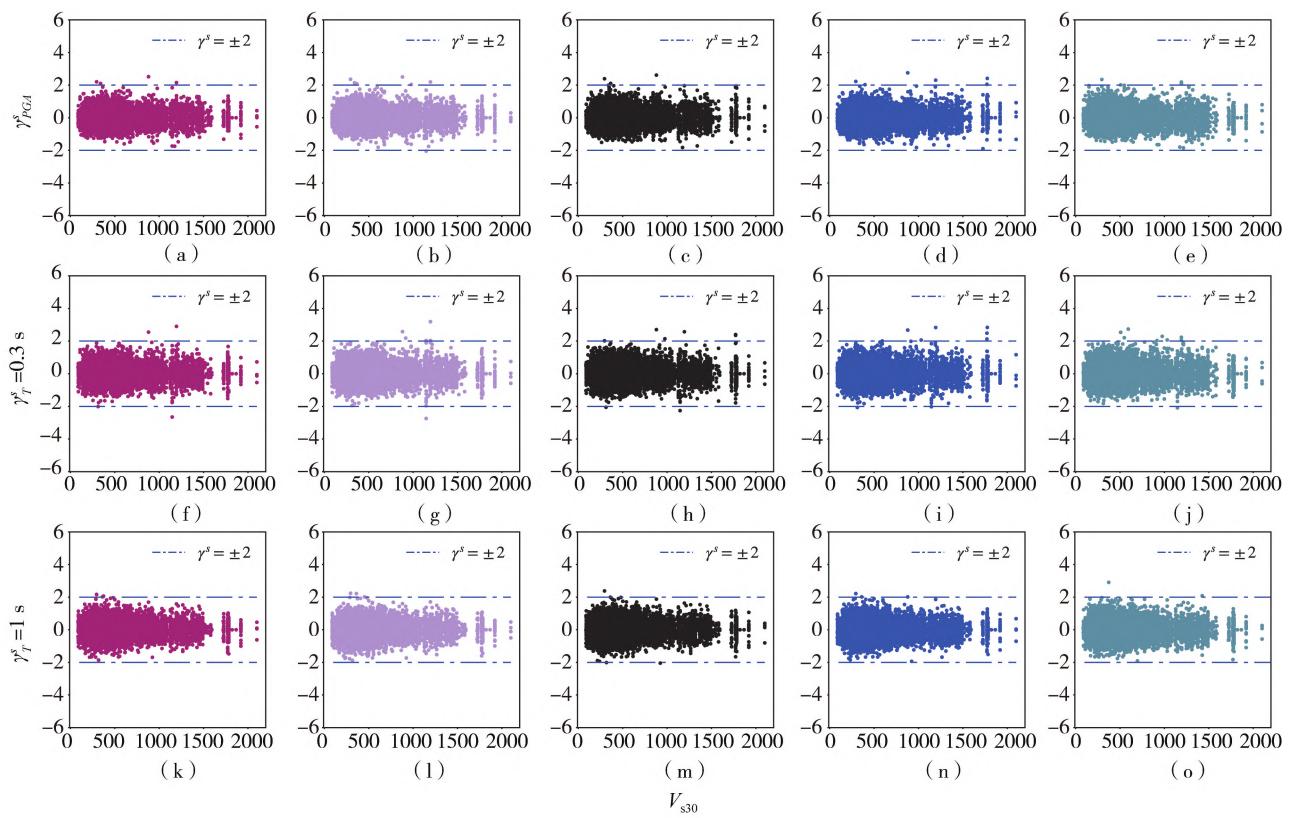


图8 不同算法模型场地间残差随 V_{s30} 变化分布情况

Fig. 8 Distribution of inter-event residual of different algorithms models with the variation of V_{s30}

图 9 不同算法模型场地内残差随 V_{s30} 变化分布情况Fig. 9 Distribution of intra-event residual of different algorithms models with the variation of V_{s30}

以上结果表明:机器学习方法预测结果相较传统方法所产生的残差更小,残差分布更稳定,不随震级或震源距的改变而产生系统性偏差,此外,相较于 LightGBM、XGBoost 和 CatBoost,Stacking-Interface 模型残差数值分布在 $[-1, 1]$ 内的数据占比分别增加了 0.28%、1.03% 和 2.16%,说明相较单一机器学习算法,Stacking 模型融合策略能够带来一定的性能进步。

3.2 评价指标对比

为更加直观的评价模型预测精度,本文引入三个指标,分别为:

- 1) 标准差 σ ,用于度量残差离散程度的统计量,当标准差较小时,说明残差的离散程度较小。
- 2) 均方误差 MSE,即模型预测值与真实值之间的差距。MSE 越小,说明模型的预测值与真实值之间的差距越小,预测精度越高。
- 3) 皮尔逊相关系数 R ,用于度量两个变量之间相关性强度的统计量,其值越大,表明真实值与预测值的相关联程度越高。

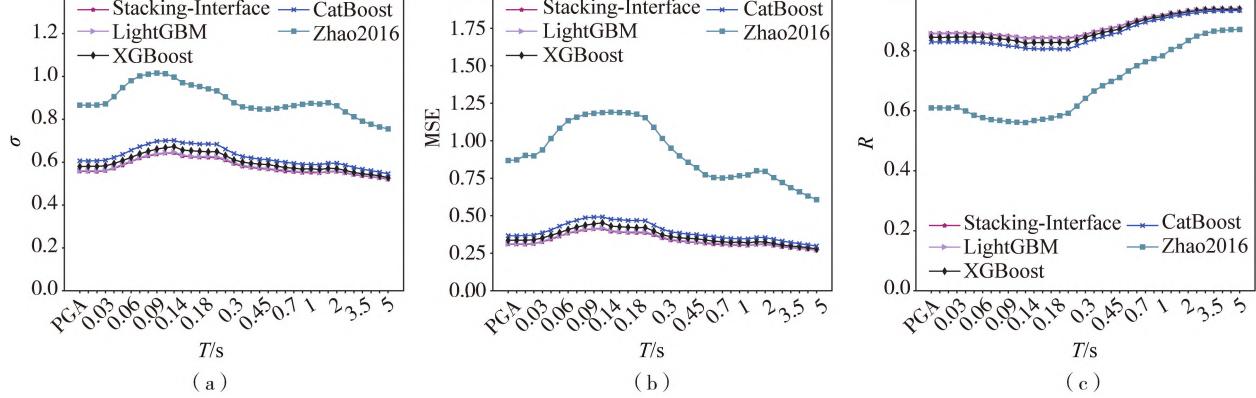


图 10 不同模型评价指标随周期变化对比情况

Fig. 10 Comparison of evaluation metrics for different models with period variation

不同算法模型在不同周期下的 σ 、MSE 和 R 如图 10 所示。从标准差来看:Stacking-Interface 模型在各周

期上均为最小,而 Zhao2016 模型较大,说明其模型预测结果的离散性较大。同样地,对于 MSE 与 R 这两个评价指标,Zhao2016 模型精度表现不如机器学习算法,且随周期的变化,出现了不稳定的变化:在较短周期内($T<0.03$ s)较好,随周期的增加,精度逐渐下降,而到了长周期($T>0.3$ s),精度逐步有所提升。而机器学习算法不仅在各周期上精度较高,且在各周期上表现相对稳定。

图 11 为各模型在测试集中各周期精度指标平均值的对比情况,可以看出:机器学习算法模型的各评价指标均优于 Zhao2016 模型,且 Stacking-Interface 模型相较其他单个算法模型在各评价指标上均为最优, σ 、MSE 和 R 分别为 0.5789、0.3364 和 0.8816。精度提升情况见表 3,说明使用 Stacking 模型融合策略,融合不同机器学习算法模型,通过二次学习数据分布规律,能够有效地提升预测精度。

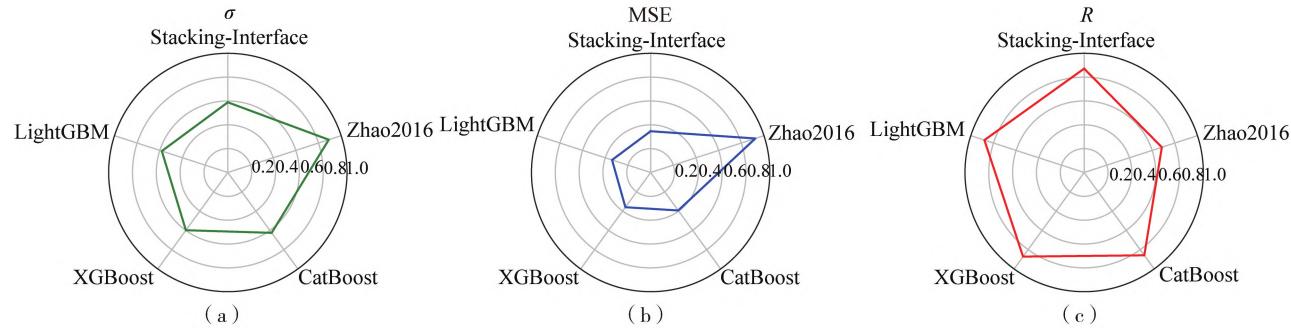


图 11 不同模型各周期平均精度指标

Fig. 11 Average accuracy metrics for different models at each period

表 3 Stacking-Interface 相较于单个机器学习算法模型与 Zhao2016 模型精度提升情况

Table 3 Accuracy improvement of Stacking-Interface compared to individual machine learning models and Zhao2016 model

对比模型	σ 降低百分比/%	MSE 降低百分比/%	R 提升百分比/%
LightGBM	0.59	1.18	0.17
XGBoost	3.30	6.57	1.08
CatBoost	7.49	14.54	2.65
Zhao2016	34.99	63.57	28.59

3.3 实例验证

通过以上分析,表明本文所提出的 Stacking-Interface 模型相较其他模型预测精度较好,且相对于单个机器学习算法模型,有一定的性能提升。为探究 Stacking-Interface 模型在地震动衰减曲线预测中的表现情况,本文将震级和震源距划分为不同的数值区间,观察模型在不同区间内的平均预测情况。此外,通过不同地震动特征的组合,观察单条地震动衰减曲线的预测情况。

将震级($4 \leq M_{JMA} < 7$)与震源距($50 \leq R_{hypo} < 200$ km)划分为不同的数值区间进行组合,所得到的平均衰减曲线如图 12 所示,从图中可以看出:多数情况下机器学习算法所预测的衰减曲线在数值与变化趋势都与真实衰减曲线都较为接近,机器学习模型在多数情况下都能达到较好的结果,而 Zhao2016 模型在不同震级与震源距的组合中,表现出不同的预测情况,如在震源距较近时,如图 12(a) 和图 12(d) 所示,出现了预测衰减曲线偏大的情况,而在震级较大时,如图 12(g)–图 12(h),长周期预测结果偏小。

为分析 Stacking-Interface 模型在具体地震动记录中的表现,图 13(a)–图 13(d) 为本文测试集中四次地震动记录,由于本文采用数据集为 2005—2019 年所记录到地震动记录,因此引入 2021 年 1 月 22 日 1 时 8 分(日本当地时间)发生在日本九州岛海域的地震记录观察本文所提出的模型在测试集之外的预测能力,如图 13(e) 和图 13(f) 所示。此外,为测试在日本之外地区 Stacking-Interface 模型的适用性,本文引入智利数据作为泛化测试,数据集来自地震动数据库 SIBER-RISK (simulation based earthquake risk and resilience of interdependent systems and networks),可以发现:尽管精度不如测试集理想,但所提模型的结果仍然可以较为准确地刻画出地震动实际的频谱特性。

结果表明:相较于传统模型与单一机器学习算法模型,本文提出的 Stacking-Interface 模型不仅能够更为准确的预测各周期地震动数值,衰减曲线变化趋势也与真实衰减曲线较为相近,且在本文所用测试集之外的日本地震动预测中,也表现出较为良好的预测能力。同时,在智利俯冲带板缘地震动的预测中也能够达到较好的效果,表明本文所提出的模型拥有一定的泛化能力。

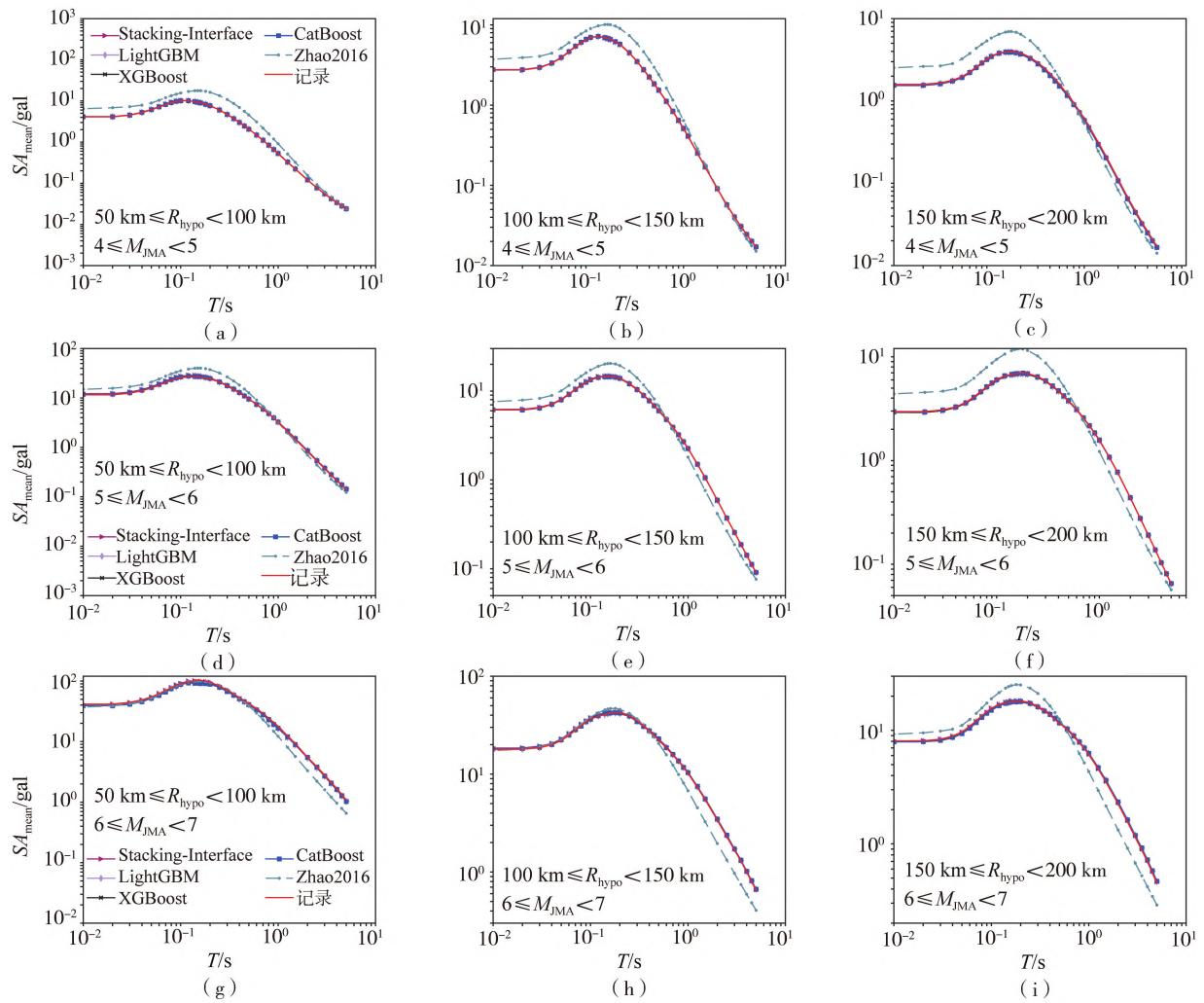
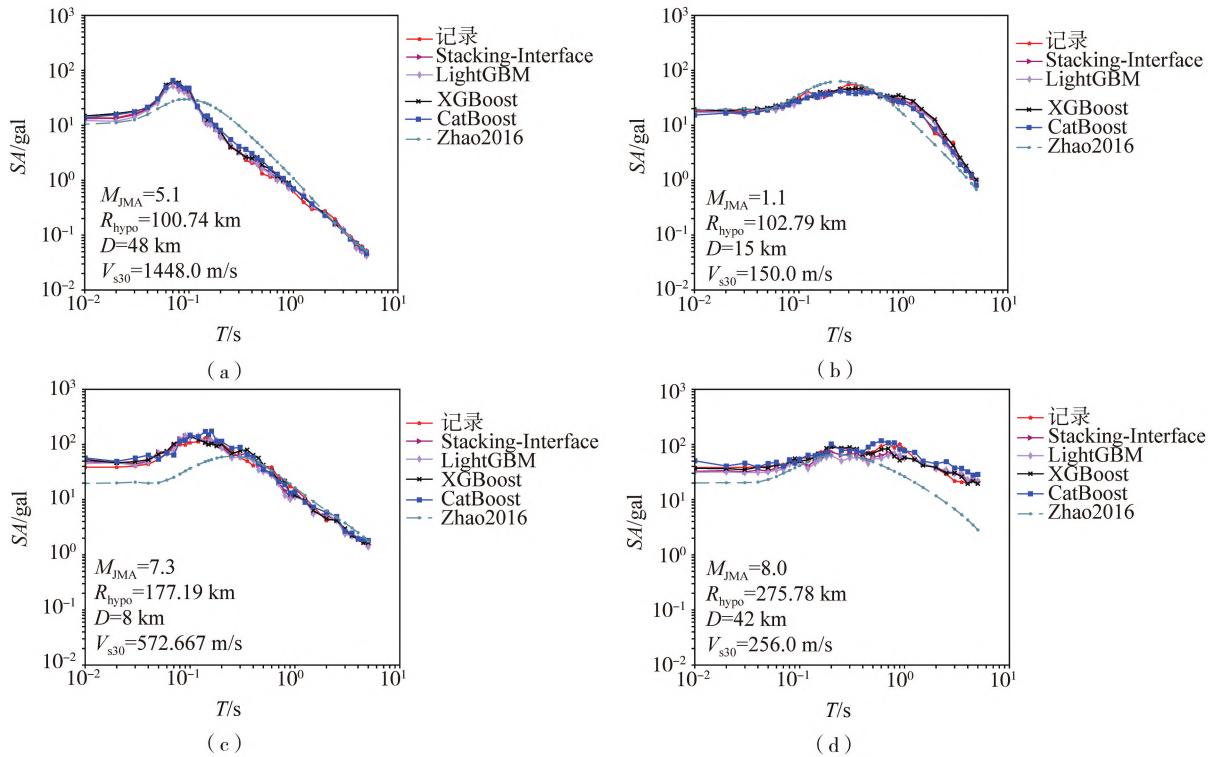


图 12 不同震级与震源距区间下,不同模型的预测衰减曲线与真实衰减曲线对比

Fig. 12 Comparison of predicted attenuation curves and true attenuation curves for different models across various magnitude and hypo-central intervals



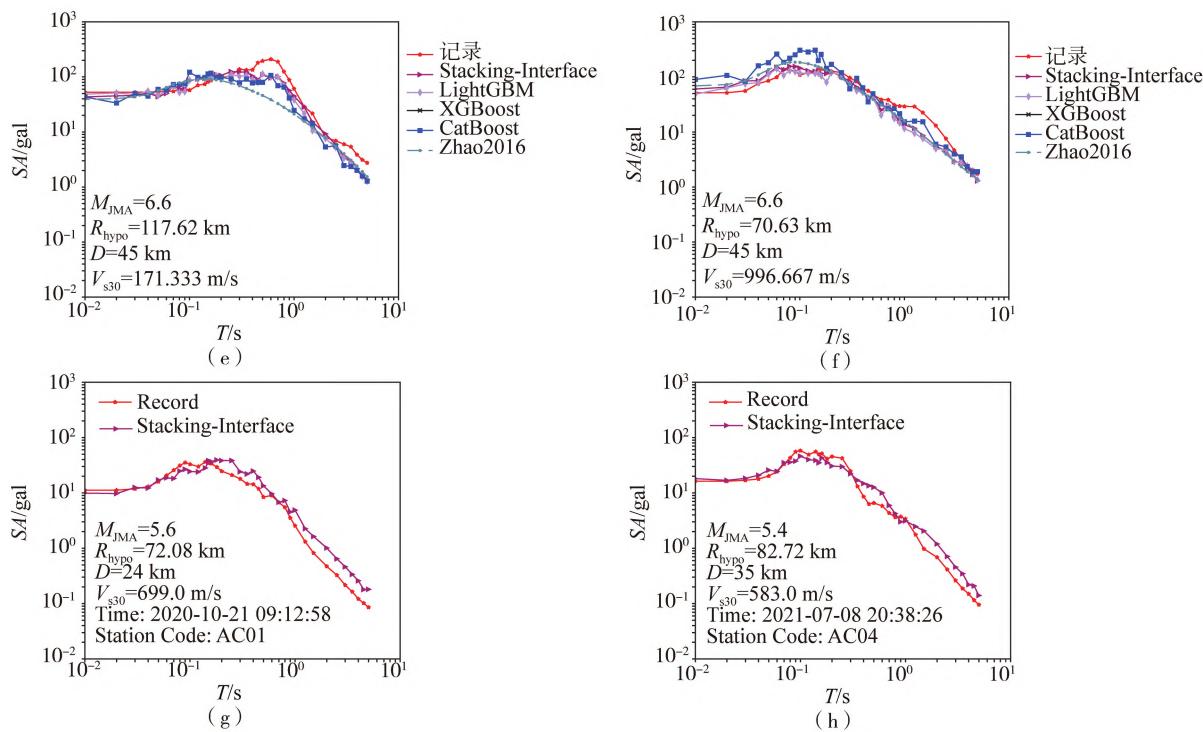


图 13 不同模型预算地震动衰减曲线实际地震动曲线对比

Fig. 13 Comparison of predicted ground motion attenuation curves and actual ground motion curves for different models

4 特征参数敏感性分析

机器学习模型是一种数据驱动的算法模型,虽然在许多研究领域被广泛利用并能够达到较好的精度,但是由于其本身作为一种无具体方程形式的非参数模型,在可解释性中存在一定局限。因此为分析本文所提出的 Stacking-Interface 模型是否符合物理意义,引入部分依赖图(partial dependence plot, PDP)对特征参数(震级、震源距、震源深度和 V_{s30})与 PGA、 $T=0.1$ s 和 $T=1$ s 时模型输出结果之间的交互关系进行敏感性分析^[38]。

PDP 期望值代表不同特征数值下模型输出结果的期望值大小,计算公式如式(9),可以通过保持其他特征参数数值不变,仅改变指定特征的数值,计算并观察 PDP 期望值的变化,实现对特征参数输入与模型输出关系的分析。

$$\hat{F}_{X_s}(X_s) = E_{X_e}[\hat{F}_{X_s}, X_e] \quad (9)$$

式中: X_s 为指定的特征参数, X_e 为其他特征参数, \hat{F} 为机器学习模型, $E_{X_e}[\hat{F}_{X_s}, X_e]$ 为 PDP 期望值。

一维关系 PDP 可以观察在单个特征变化下,模型输出结果的改变情况,图 14 展示了随震源深度、震级、震源距和场地条件数值变化情况下不同周期的 PDP 期望值。从图中可以看出:在不同周期下,虽然模型所输出的 PDP 期望值不同,但其随特征参数数值变化趋势相似,如震级越大,模型所输出结果越大,震源距越大,模型输出结果越小。本文发现:随着震源距与 V_{s30} 数值的改变,PDP 期望值变化较为剧烈。对于震源深度,当深度较浅时,随深度的增加,PDP 期望值出现下降的趋势,可能是由于发震深度增加导致的地震应力降的增加或品质因子 Q 的增大所致^[39-40]。 V_{s30} 的变化表明场地条件与模型输出有着强烈的非线性关系,但是随着周期的增加,剧烈程度逐渐变弱,此现象与姜治军等^[41]对场地建立放大系数模型时所发现变化规律一致。此外,相同周期时,随 V_{s30} 增大,产生的 PDP 期望值呈现非线性减小的趋势,表明松软土层场地上的场地放大效应通常更为明显^[42],因此传统方法构建经验方程时,会引入非线性场地项来表征特定场地上的地震动相对于基岩场地地震动的放大效应^[43]。

PDP 不仅可以观察单个特征参数与模型输出结果的一维关系,还可以观察在两个特征共同作用时,模型输出结果的变化情况,本文选择震级与震源距作为研究对象,不同周期下的部分依赖图如图 15 所示,部分依赖图的总体趋势表明:在震级越大,震源距越小的情况下,模型能够产生较大的输出结果,且能够发现,在相同震级下,震源距由近至远,模型输出结果呈现出非线性减小的情况。

通过使用部分依赖图对模型分别进行一维关系和二维关系分析,揭示了模型内在机制与特征参数之间的相互作用规律,结果表明:本文所提出的模型符合物理规律。

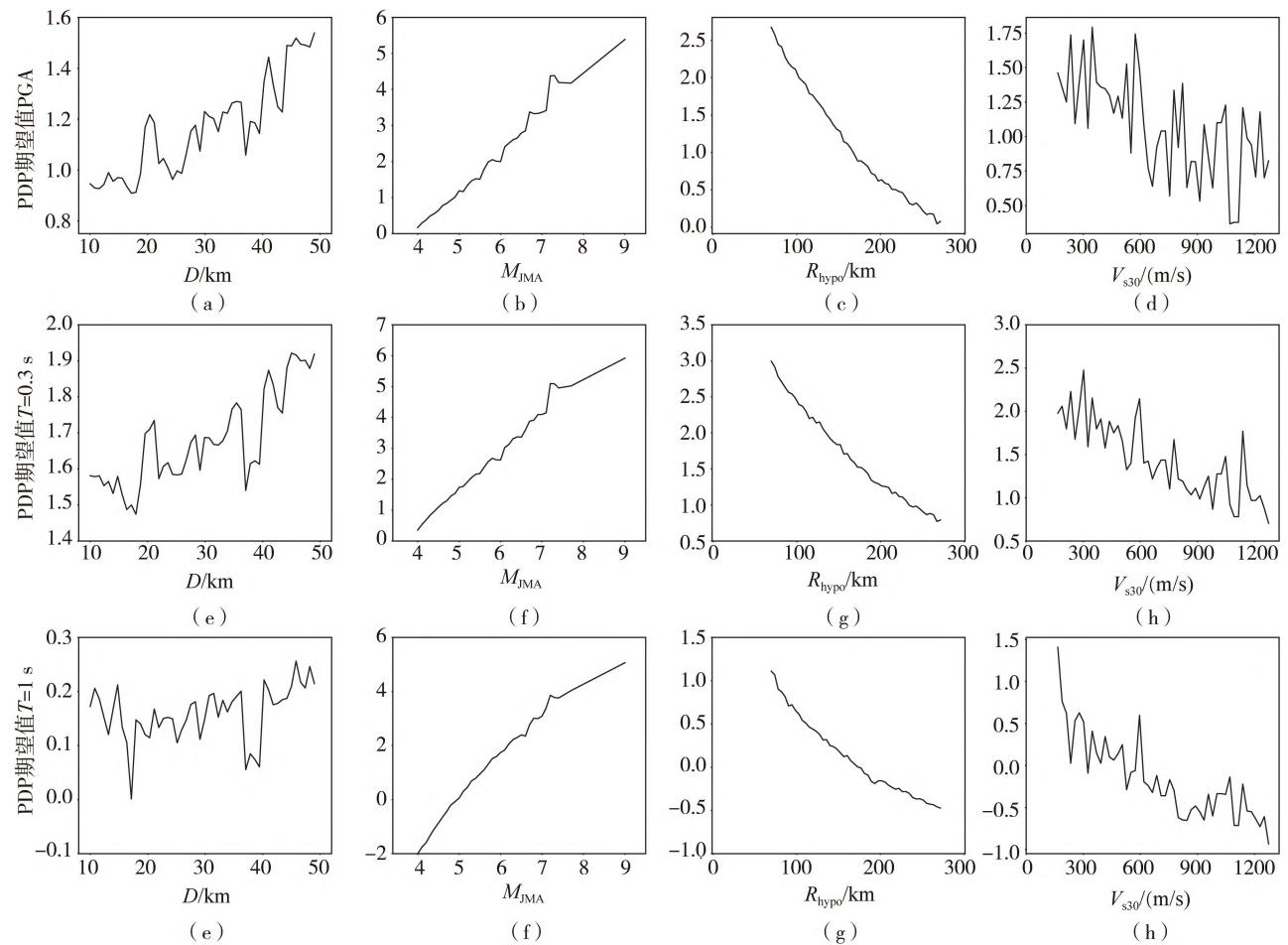


图 14 不同周期下,Stacking-Interface 模型 PDP 期望值随特征参数值变化趋势图

Fig. 14 Trend of Stacking-Interface model's partial PDP expected values with feature parameter values for different periods

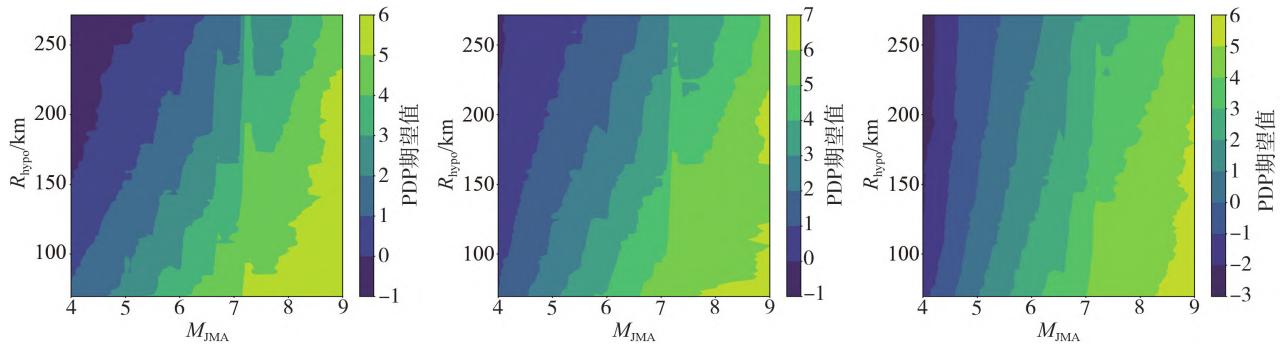


图 15 震级与震源距相互作用下,不同周期的部分依赖图

Fig. 15 PDP for different periods with interaction between magnitude and hypo-central

5 结论

本文通过从 KiK-net 和 K-Net 获取到的地震动记录,经筛选和处理后形成大数据库,通过不同算法的对比选择,采用 Stacking 模型融合策略集成 LightGBM、XGBoost 和 CatBoost 训练融合模型 Stacking-Interface,并引入贝叶斯优化算法对超参数进行调整以提高所提模型的精度,对比了所提模型、单一机器学习模型和传统模型的精度,测试了模型的泛化能力,并引入部分依赖图从一维特征和二维特征角度分析模型的内在机理,本文已将模型托管至 GitHub (<https://github.com/heroic98/Stacking-Interface>),供下载使用和验证本文结论,此外,本文所得结论如下:

1) 通过 Stacking 模型融合策略训练 Stacking-Interface 模型, 结果表明: 相较预测精度最好的单一机器学习算法模型 LightGBM, 测试集的标准差、均方误差分别降低 0.59%、降低 1.18%, 皮尔逊相关系数提升 0.17%, 相较传统方法, 标准差和均方误差分别降低 34.99% 和降低 63.57%, 皮尔逊相关系数提升 28.59%。并且在泛化测试中, Stacking-Interface 模型预测衰减曲线更贴近真实衰减曲线。

2) 通过对比遗传优化算法与贝叶斯优化算法, 发现在搜索相同次数时, 所消耗时间相近, 但贝叶斯优化算法能够搜索到精度更高的超参数配置。

3) 引入部分依赖图方法, 分别从单个特征角度与两个特征共同作用角度下对模型进行特征参数敏感性分析, 结果表明本文所提出的 Stacking-Interface 模型符合物理规律。

4) 限于计算机硬件与数据数量限制, 本文未尝试加入更多的模型进行融合, 未来将会加入更多数据, 融合更多模型, 以期能够带来更多的性能提升。同时由于机器学习模型为数据驱动, 暂无法做到直接输出平滑的反应谱, 因此在未来研究中, 将会考虑在模型中加入物理约束, 如修改损失函数, 以获得精度较高的平滑谱。

参考文献:

- [1] 梁基冠, 苏开滩, 杨青松. 不同类型俯冲带地震竖向地震动阻尼修正系数对比研究[J]. 世界地震工程, 2019, 35(4): 171–181.
LIANG Jiguan, SU Kaiwei, YANG Qingsong. Comparison of damping modification factors for the vertical ground motions of different earthquake types in subduction zones[J]. World Earthquake Engineering, 2019, 35(4): 171–181. (in Chinese)
- [2] 胡进军, 张辉, 张齐. 基于四川西昌中强震数据的区域地震动预测模型[J]. 振动与冲击, 2021, 40(2): 279–286.
HU Jinjun, ZHANG Hui, ZHANG Qi. Regional ground motion prediction model for moderate magnitude earthquakes in Xichang, Sichuan, China [J]. Journal of Vibration and Shock, 2021, 40(2): 279–286. (in Chinese)
- [3] 余聪, 宋晋东, 李山有. 基于支持向量机的现地地震预警地震动峰值预测[J]. 振动与冲击, 2021, 40(3): 63–72, 80.
YU Cong, SONG Jindong, LI Shanyou. Prediction of peak ground motion for on-site earthquake early warning based on SVM[J]. Journal of Vibration and Shock, 2021, 40(3): 63–72, 80. (in Chinese)
- [4] GUTENBERG B, RICHTER C F. Earthquake magnitude, intensity, energy, and acceleration[J]. Bulletin of the Seismological Society of America, 1942, 32(3): 163–191.
- [5] ZHAO J X, XU H. Magnitude-scaling rate in ground-motion prediction equations for response spectra from large subduction interface earthquakes in Japan[J]. Bulletin of the Seismological Society of America, 2012, 102(1): 222–235.
- [6] 陶正如, 陶夏新. 美国 2014 地震区划中采用的地震动衰减关系[J]. 世界地震工程, 2015, 31(3): 78–84.
TAO Zhengru, TAO Xiaxin. Ground motion attenuation relationships adopted in 2014 update of the US National Seismic Hazard Maps[J]. World Earthquake Engineering, 2015, 31(3): 78–84. (in Chinese)
- [7] ABRAHAMSON N A, SILVA W J, KAMAI R. Summary of the ASK14 ground motion relation for active crustal regions[J]. Earthquake Spectra, 2014, 30(3): 1025–1055.
- [8] BOORE D M, STEWART J P, SEYHAN E, et al. NGA-West2 equations for predicting PGA, PGV, and 5% damped PSA for shallow crustal earthquakes[J]. Earthquake Spectra, 2014, 30(3): 1057–1085.
- [9] CAMPBELL K W, BOZORGNAI Y. NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5% damped linear acceleration response spectra[J]. Earthquake Spectra, 2014, 30(3): 1087–1115.
- [10] CHIOU B S J, YOUNGS R R. Update of the chiou and youngs NGA model for the average horizontal component of peak ground motion and response spectra[J]. Earthquake Spectra, 2014, 30(3): 1117–1153.
- [11] IDRISI I M. An NGA-West2 empirical model for estimating the horizontal spectral values generated by shallow crustal earthquakes[J]. Earthquake Spectra, 2014, 30(3): 1155–1177.
- [12] PARKER G A, STEWART J P, BOORE D M, et al. NGA-subduction global ground motion models with regional adjustment factors[J]. Earthquake Spectra, 2022, 38(1): 456–493.
- [13] ZHAO J X, ZHOU S L, ZHOU J, et al. Ground-motion prediction equations for shallow crustal and upper-mantle earthquakes in Japan using site class and simple geometric attenuation functions[J]. Bulletin of the Seismological Society of America, 2016, 106(4): 1552–1569.
- [14] ZHAO J X, JIANG F, SHI P, et al. Ground-motion prediction equations for subduction slab earthquakes in Japan using site class and simple geometric attenuation functions[J]. Bulletin of the Seismological Society of America, 2016, 106(4): 1535–1551.
- [15] ZHAO J X, LIANG X, JIANG F, et al. Ground-motion prediction equations for subduction interface earthquakes in Japan using site class and simple geometric attenuation functions[J]. Bulletin of the Seismological Society of America, 2016, 106(4): 1518–1534.
- [16] 肖亮, 俞言祥. 俯冲带板缘与板内地震长周期地震动参数预测模型研究[J]. 震灾防御技术, 2021, 16(1): 71–80.
XIAO Liang, YU Yanxiang. Study on the long-period ground motion predicting model for interface and intraslab earthquakes in subduction zone [J]. Technology for Earthquake Disaster Prevention, 2021, 16(1): 71–80. (in Chinese)
- [17] 高孟潭.《中国地震动参数区划图(GB 18306—2015)》宣贯教材[M]. 北京: 中国质检出版社, 中国标准出版社, 2015.
GAO Mengtan. Chinese ground motion intensity measures zoning map publicity materials (GB18306—2015) [M]. Beijing: Standards Press of

- China Zhijian Publishing House, 2015. (in Chinese)
- [18] VEMULA S, KP S, RAGHUKANTH S T G. Neural network-based subduction ground motion model and its application to New Zealand and the andaman and Nicobar Islands [J]. Journal of Earthquake Engineering, 2023, 27(10): 2863–2886.
- [19] SEO H, KIM J, KIM B. Machine-learning-based surface ground-motion prediction models for South Korea with low-to-moderate seismicity [J]. Bulletin of the Seismological Society of America, 2022, 112(3): 1549–1564.
- [20] JI D F, LI C X, ZHAI C H, et al. Prediction of ground-motion parameters for the NGA-West2 database using refined second-order deep neural networks [J]. Bulletin of the Seismological Society of America, 2021, 111(6): 3278–3296.
- [21] KE G L, MENG Q, FINLEY T, et al. LightGBM: A highly efficient gradient boosting decision tree [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: ACM, 2017: 3149–3157.
- [22] CHEN T, GUESTRIN C. XGBoost: A scalable tree boosting system [C]//In proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2016: 785–794.
- [23] DOROGUSH A V, ERSHOV V, GULIN A. CatBoost: Gradient boosting with categorical features support [EB/OL]. 2018: arXiv: 1810.11363. <https://arxiv.org/abs/1810.11363.pdf>
- [24] PROKHOREKOVA L, GUSEV G, VOROBEV A, et al. CatBoost: Unbiased boosting with categorical features [C]//Proceedings of the 32nd International Conference on Neural Information Processing Systems. Montréal, Canada. New York: ACM, 2018: 6639–6649.
- [25] 胡进军, 郑鹏. 基于日本滨海强震数据的不同震源类型衰减关系比较 [J]. 建筑结构, 2017, 47(增刊1): 669–77.
- HU Jinjun, ZHENG Peng. Comparison of attenuation relationships of different seismic types based on strong earthquake data in coastal areas Japan [J]. Building Structure, 2017, 47(S1): 669–77. (in Chinese)
- [26] CONVERSE, A, BRADY, A. G. Basic strong-motion accelerogram processing software. USGS Open-file Report, 1992.
- [27] BOORE D M. Estimating s(30) (or NEHRP site classes) from shallow velocity models (depths < 30 m) [J]. Bulletin of the Seismological Society of America, 2004, 94(2): 591–597.
- [28] BOORE D M, THOMPSON E M, CADET H. Regional correlations of VS30 and velocities averaged over depths less than and greater than 30 m [J]. Bulletin of the Seismological Society of America, 2011, 101(6): 3046–3059.
- [29] TSAMPAS A D, SCORDILIS E M, PAPAZACHOS C B, et al. Global-magnitude scaling relations for intermediate-depth and deep-focus earthquakes [J]. Bulletin of the Seismological Society of America, 2016, 106(2): 418–434.
- [30] WOLPERT D H. Stacked generalization [J]. Neural Networks, 1992, 5(2): 241–259.
- [31] BREIMAN L. Random forests [J]. Machine Learning, 2001, 45(1): 5–32.
- [32] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. Nature, 2015, 521: 436–444.
- [33] SNOEK J, LAROCHELLE H, ADAMS R P. Practical Bayesian optimization of machine learning algorithms [C]//Proceedings of the 25th International Conference on Neural Information Processing Systems-Volume 2. Lake Tahoe, Nevada. New York: ACM, 2012: 2951–2959.
- [34] KATOCH S, CHAUHAN S S, KUMAR V. A review on genetic algorithm: Past, present, and future [J]. Multimedia Tools and Applications, 2021, 80(5): 8091–8126.
- [35] 崔佳旭, 杨博. 贝叶斯优化方法和应用综述 [J]. 软件学报, 2018, 29(10): 3068–3090.
- CUI Jiaxu, YANG Bo. Survey on Bayesian optimization methodology and applications [J]. Journal of Software, 2018, 29(10): 3068–3090. (in Chinese)
- [36] 葛继科, 邱玉辉, 吴春明, 等. 遗传算法研究综述 [J]. 计算机应用研究, 2008, 25(10): 2911–2916.
- GE Jike, QIU Yuhui, WU Chunming, et al. Summary of genetic algorithms research [J]. Application Research of Computers, 2008, 25(10): 2911–2916. (in Chinese)
- [37] ABRAHAMSON N A, YOUNGS R R. A stable algorithm for regression analyses using the random effects model [J]. Bulletin of the Seismological Society of America, 1992, 82(1): 505–510.
- [38] GREENWELL B M. Pdp: An R package for constructing partial dependence plots [J]. The R Journal, 2017, 9(1): 421.
- [39] ABERCROMBIE R E, TRUGMAN D T, SHEARER P M, et al. Does earthquake stress drop increase with depth in the crust? [J]. Journal of Geophysical Research: Solid Earth, 2021, 126(10): e2021JB022314.
- [40] 陈蒙, 王华. 地震动强度参数估计的可解释性与不确定度机器学习模型 [J]. 地球物理学报, 2022, 65(9): 3386–3404.
- CHEN Meng, WANG Hua. Explainable machine learning model for prediction of ground motion parameters with uncertainty quantification [J]. Chinese Journal of Geophysics, 2022, 65(9): 3386–3404. (in Chinese)
- [41] 姜治军, 胡进军, 张齐, 等. 考虑土层非线性效应的四川地区场地放大系数模型 [J]. 岩土工程学报, 2016, 38(9): 1650–1659.
- JIANG Zhijun, HU Jinjun, ZHANG Qi, et al. Site amplification factor model for Sichuan region considering nonlinear soil effects [J]. Chinese Journal of Geotechnical Engineering, 2016, 38(9): 1650–1659. (in Chinese)
- [42] 胡进军, 郑旭, 谢礼立. 基于混合方法的南海海域地震动衰减关系研究 [J]. 土木工程学报, 2018, 51(7): 36–49.
- HU Jinjun, ZHENG Xu, XIE Lili. Derivation of ground motion attenuation relation for earthquake in the South China Sea areas based on a hybrid method [J]. China Civil Engineering Journal, 2018, 51(7): 36–49. (in Chinese)
- [43] 张雨婷, 任叶飞, 米欣雪, 等. 区域场地分类方法研究综述 [J]. 世界地震工程, 2023, 39(2): 189–199.
- ZHANG Yuting, REN Yefei, MI Xinxue, et al. Research progress of the regional site classification method [J]. World Earthquake Engineering, 2023, 39(2): 189–199. (in Chinese)