

基于机器学习和东日本大地震破坏数据的 地震动参数筛选

李 遂¹, 王自法^{2,3,1}, 赵登科^{2,3}, 王昭栋^{2,3}, 李兆焱^{2,3}

(1. 河南大学 建筑工程学院, 河南 开封 475004; 2. 中国地震局工程力学研究所 地震工程与工程振动重点实验室, 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)

摘要: 地震动参数量化了地面运动的强度及其对建筑结构的影响程度, 选取合理的参数对震前抗震设计与震后震害评估具有重要意义。地震动参数和结构地震反应间的统计关系常通过传统的相关性分析和回归方法建立, 且数据多来源于数值模拟, 难以捕捉两者间真实的非线性映射关系。为此, 该文收集整理了2011年3月11日东日本大地震后近128万条房屋实际震害数据, 基于XGBoost (eXtreme gradient boosting)、随机森林(random forest, RF)、LightGBM (light gradient boosting machine) 和 CatBoost (categorical boosting) 共4种机器学习分类模型建立地震动参数与建筑破坏等级之间的复杂映射关系。引入SMOTE过采样和贝叶斯超参数优化法对模型进行优化, 结合2种重要性评价方法选出包含7个参数的最优地震动参数组合。结果表明: XGBoost 算法表现最佳, 在测试集上的整体准确率可以达到71.39%; 最优地震动参数组合为峰值地面加速度 (peak ground acceleration, PGA)、 T_d 、 V_{sl} 、峰值地面位移 (peak ground displacement, PGD)、峰值地面速度 (peak ground velocity, PGV)、PGV/PGA、谱加速度 S_a ; 地震动幅值、频谱及持时参数与震后建筑损失的相关性较强, 而累积能量参数与其相关性较弱。最后, 通过新西兰3次地震的实际震害数据建立基于XGBoost算法的地震损失预测模型, 验证了所选参数组合的完备性、可靠性和地区泛化性能。研究成果能够为建筑抗震设计及地震风险评估等工作提供一定的理论依据和工程参考。

关键词: 地震损失数据; 机器学习; 建筑破坏等级; 最优地震动参数组合; 特征重要性

中图分类号: P315.9

文献标志码: A

Selection of ground motion parameters based on machine learning and damage data from the Great East Japan Earthquake

LI Sui¹, WANG Zifa^{2,3,1}, ZHAO Dengke^{2,3}, WANG Zhaodong^{2,3}, LI Zhaoyan^{2,3}

(1. School of Civil Engineering and Architecture, Henan University, Kaifeng 475004, China; 2. Key Laboratory of Earthquake Engineering and Engineering Vibration, Institute of Engineering Mechanics, China Earthquake Administration, Harbin 150080, China;
3. Key Laboratory of Earthquake Disaster Mitigation, Ministry of Emergency Management, Harbin 150080, China)

Abstract: Ground motion parameters quantify the intensity of ground motion and their impact on building structures, making the selection of appropriate parameters crucial for pre-earthquake seismic design and post-earthquake damage assessment. Ground motion parameters and structural seismic responses are often statistically related through traditional correlation analysis and regression methods. However, data are mostly sourced from numerical simulations, which makes it difficult to capture the true nonlinear mapping relationship between the two. Therefore, this paper collected and organized nearly 1.28 million actual damage records from the Great East Japan

收稿日期: 2024-10-09; 修回日期: 2024-11-06

基金项目: 国家自然科学基金面上项目 (52378544, 52378543); 国家重点研发计划项目 (2023YFC3805203); 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, 2023A01)

作者简介: 李 遂 (2001—), 女, 硕士研究生, 主要从事地震动参数和损失间相关性研究。E-mail: ls1295950841@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

Earthquake on March 11, 2011, and the complex mapping relationship between ground motion parameters and building damage levels was established based on four machine learning classification models, namely, XGBoost (eXtreme gradient boosting), RF (random forest), LightGBM (light gradient boosting machine), and CatBoost (categorical boosting). The SMOTE oversampling and Bayesian hyperparameter optimization algorithms were introduced to optimize the model, and the optimal combination of seven ground motion parameters was selected using two methods for evaluating feature importance. The results indicate that the XGBoost algorithm performs the best, with an overall accuracy of 71.39% on the test set. The optimal combination of ground motion parameters includes PGA , T_d , V_{SI} , PGD , PGV/PGA , PGV , and S_a . The amplitude, spectrum, and duration parameters of the ground motion show a strong correlation with post-earthquake building damage, while the cumulative energy parameters exhibit a weaker correlation. Finally, an earthquake loss prediction model based on the XGBoost algorithm was established using actual damage data from three earthquakes in New Zealand, validating the completeness, reliability, and regional generalization capability of the selected parameter combination. The research results can provide a theoretical basis and engineering reference for the seismic design of buildings and earthquake risk assessment.

Key words: earthquake loss data; machine learning; building damage level; optimal ground motion parameter combination; feature importance

0 引言

强震地面运动十分复杂,地震动参数(intensity measure, IM)量化了地面运动的强度,是连接地震灾害和建筑破坏程度之间的关键参数^[1]。因此,在地震工程中,寻找关键的地震动参数及其组合来描述地面运动对结构的破坏潜力是一个重要问题^[2-3]。

传统方法往往基于相关性评价或回归方法确定地震动参数与结构响应指标(damage measure, DM)之间的统计关系,并进一步筛选出最具相关特征的单一地震动参数^[4]。近年来,许多专家学者围绕该方面做了大量研究^[5-6]。现有研究已提出数十种地震动参数,能表征地震动破坏潜力的参数通常只有有限几个。单一强度指标如峰值地面加速度(peak ground acceleration, PGA)、峰值地面速度(peak ground velocity, PGV)和考虑结构基本周期的谱加速度(S_a),因其形式简单且物理意义明确而被广泛使用。叶列平等^[7]基于皮尔逊相关性系数分析了33个地震动参数和结构峰值反应的线性相关性,指出已有的单一指标无法在各个周期内达到最佳效果;陈健云等^[8]研究了13种常用地震动参数与框架结构地震反应的相关性,提出结构在线性和非线性反应阶段的峰值反应与地震动参数的相关性不同,建议选用多个参数来提高抗震性能评估的可靠性。综上,单一地震动参数无法准确描述地震对结构响应的影响,因此,有必要使用多元地震动参数或者找到最优的地震动参数组合。

为了综合考虑多个地震动参数的影响,已有学者通过对单一参数进行线性或指数形式的组合得到多元地震动参数^[9-12]。LUCO等^[13]考虑到结构非线性和高阶振型的影响,提出了与位移谱相关的组合型地震动参数;刘亭亭等^[14-16]分别采用主成分分析、偏最小二乘法和典型相关性分析建立了多元地震动参数。与单个参数相比,多元参数与结构响应虽然在不同周期内整体相关性更强,但现有研究在组合方法上存在较大的主观性,并且对于如何构建最佳的组合型参数尚且缺乏系统讨论。此外,传统方法在拟合地震动参数和建筑破坏程度间的非线性关系时存在一定局限性,具有较大的改进空间。

机器学习作为一种有效的数据分析工具,在拟合非线性关系时更加灵活和准确,并且在参数选择上具有显著的优势,近年来其也被应用于地震动参数比选的研究中。XU等^[17]基于机器学习方法建立了地震动参数和结构损伤状态之间的非线性映射关系,并迭代比选出几个关键指标;YAN等^[18]提出了一种顺序浮动特征选择算法,可以从地震动参数候选集中找到最佳的参数组合,并验证了其有效性。相比于传统的相关性分析和多元回归方法,虽然基于机器学习方法的地震动参数筛选研究已经取得了一些成果,但仍存在以下不足:

1) 由于缺乏实际震害资料,目前大多数研究的数据都来自数值模拟,并非实际的房屋震害数据,且采用的样本数据量有限,难以满足大数据样本的需求。

2) 现有研究往往采用单一和经验性的结构响应指标(如最大层间位移、最大楼面速度和最大楼面加速度等)来判断结构的破坏状态,无法有效反映建筑物实际的破坏程度,进一步增加了地震动参数和结构损伤状态之间的不确定性。

3) 目前的研究通常仅针对特定地区的单一结构类型进行建模和分析,缺乏区域和建筑结构类型泛化能力,无法验证所选取的最优指标是否具有普适性。

为了解决上述问题,本文整理了2011年3月11日东日本大地震中近128万栋建筑的实际震害资料。首先基于台站实际观测信息,利用地震动空间插值方法计算出建筑物处的14个地震动参数;随后基于机器学习模型建立地震动参数、建筑物信息和场地条件与建筑破坏等级之间的复杂非线性映射关系,并结合2种重要性评价方法迭代筛选出最有效的地震动参数组合;最后采用新西兰3次地震事件中共115739条建筑详细破坏数据进行验证,证明所选参数组合在新西兰地区的适用性和有效性。该研究能够为地震损失估计中地震动参数的选择提供参考,同时能够有效地帮助提高地震巨灾风险分析精度。

1 数据

1.1 数据集介绍

日本位于环太平洋地震带,是世界上发生地震频数最多的国家之一。2011年3月11日,日本东北海岸 $38^{\circ}02'N, 143^{\circ}09'E$ 的地区发生里氏9.0级特大地震,造成约2.8万人遇难,数百万栋建筑受损,给社会和经济带来了巨大冲击。美国再保险公司Validus收集了震后建筑物的详细破坏资料,包括建筑物的建筑年代、层数、结构类型、柱子类型和损失比等详细建筑特征与破坏信息。本文收集整理了此次地震中1278861栋建筑的详细破坏资料作为本研究的数据集。

1.2 地震动参数计算

地震发生后,各空间位置地震动强度一般用地震动预测方程(ground motion prediction equation, GMPE)获取。GMPE通过传统回归方法来拟合地震动的复杂衰减关系,因此预测结果与实际结果往往存在偏差^[19]。本研究采用ZHAO等^[20]提出的地震动参数的空间插值计算方法,该方法基于实际的台站观测记录,利用GMPE作为权重分别对空间距离和场地条件进行修正,最终空间插值得到台站附近每栋建筑物处的地震动参数。本文采用日本国家防灾科学技术研究所(National Research Institute for Earth Science and Disaster Resilience, NIED)提供的实际台站观测以及适用于日本地区的GMPE^[21-23],获得了共14个广泛应用的地震动参数作为原始地震动参数集见第1.5节。需要说明的是,由于频谱参数中谱速度(S_v)和谱位移(S_d)也较为关键^[24],且目前没有适用于日本地区 S_v 和 S_d 的GMPE,因此上述2种地震动参数的GMPE分别采用MORASCA等^[22]和CAUZZI等^[23]提出的全球模型。建筑物的场地条件通过地下30m深度内的等效剪切波速(V_{s30})判别,计算方法参考WALD等^[25]。

1.3 数据预处理

获得的数据集中一般包含缺失值、异常值和重复数据等问题,影响后续的模式训练,因此模型训练前需要进行数据预处理。本文基于Python语言中的Pandas、Numpy等数据处理工具对数据集进行以下预处理:

1) 数据清理。剔除数据集中建筑年代、结构类型和屋顶类型等特征不明类数据;删除由地震外的其他灾害(如海啸、滑坡和火灾等)造成的建筑破坏数据。

2) 模型算法的本质是数值运算,无法识别非数字信息,因此需要将建筑信息中的建筑年代、结构类型等文本信息利用Scikit-learn库提供的Label Encoder编码工具转换为整数,实现文本数值化。

3) 数据集划分。采用机器学习领域最典型的“八二原则”对样本数据进行划分,即从震后建筑资料数据库中随机抽取80%数据(1023088条)作为训练集,剩余20%(255773条)作为测试集。

1.4 地震损伤标记

本文中建筑破坏程度信息由美国再保险公司Validus统计计算,它是由建筑物损失价值与建筑物总价值的比值计算得到,记为损失比(damage ratio, Dr)。参考美国应用技术协会所发布的地震震害评估标准ATC-13^[26],根据 r_0 将建筑物划分为4种破坏等级,(0,10%]为轻微破坏,(10%,30%]为中等破坏,(30%,60%]为严重破坏,(60%,100%]为毁坏。将未破坏建筑的破坏等级标记为0,其他四类按照破坏

等级由小到大依次标记为 1、2、3、4,标记的样本用于机器学习分类模型的训练。5 种破坏等级的数据分布见表 1。

表 1 依据 ATC-13 划分的建筑破坏等级数据分布

Table 1 Distribution of building damage levels according to ATC-13

建筑破坏等级标准	未破坏 $r_D = 0$	轻微破坏 $r_D \in (0, 10\%]$	中等破坏 $r_D \in (10\%, 30\%]$	严重破坏 $r_D \in (30\%, 60\%]$	毁坏 $r_D \in (60\%, 100\%]$	总数量/条
样本数量	888 008	157 490	172 100	31 937	29 326	1 278 861

1.5 模型特征参数选择

特征是机器学习模型的基本组成部分,选择和处理好特征可以显著提高模型的预测性能。地震发生后,建筑物的破坏是由强震地面运动、建筑结构和场地条件等因素之间相互复杂作用而造成的,因此特征选择要综合考虑以上影响因素。为了准确反映强震地面运动,研究采用在 1.2 节中计算得到的 14 个地震动参数,包括 PGA、PGV 和 S_a 等,以及震中距来描述地震动特性,这些选定的地震动参数涵盖了多种地震动信息,如振幅、频谱和累积能量等;结合数据库及现有研究中常用的建筑特征,最终确定建筑年代、屋顶类型等 7 个特征来代表建筑物信息;为了研究不同的场地类型对于建筑结构破坏的影响,引入 V_{S30} 来代表不同的场地条件。综上所述,最终确定 23 个特征作为模型输入,各个特征的具体信息见表 2。

表 2 模型的特征解释

Table 2 Feature explanation of the model

编号	类别	特征	特征解释	计算方式和来源	
1	地震信息	峰值地面加速度	$PGA = \max a(t) $	地震动空间插值方法	
2		峰值地面速度	$PGV = \max v(t) $		
3		峰值地面位移	$PGD = \max d(t) $		
4		谱加速度	$S_a = S_a(T)$		
5		谱速度	$S_v = S_v(T)$		
6		谱位移	$S_d = S_d(T)$		
7		Arias 强度	$I_a = \frac{\pi}{2g} \int_0^T a^2(t) dt$		
8	建筑物信息	特征强度	$I_c = A_{rms}^{1.5} T_d^{0.5}$	数据库	
9		重要持时	$T_d = t_2 - t_1$		
10		均方根加速度	$A_{rms} = \sqrt{\frac{1}{t_2 - t_1} \int_{t_1}^{t_2} a^2(t) dt}$		
11		累积绝对速度	$CAV = \int_0^T a(t) dt$		
12		加速度谱强度	$A_{SI} = \int_{0.1}^{0.5} S_a(T) dt$		
13		速度谱强度	$V_{SI} = \int_{0.1}^{2.5} S_v(T) dt$		
14		峰值比	PGV/PGA		
15		震中距	目标建筑到震中的地面距离		Haversine 公式
16		建筑年代	建筑物建造年代		数据库
17		外墙类型	外墙材料类型		数据库
18	结构类型	建筑物结构类型	数据库		
19	柱子类型	柱子材料类型	数据库		
20	屋顶类型	屋顶材料类型	数据库		
21	用途分类	建筑物使用用途	数据库		
22	层数	建筑物层数	数据库		
23	场地信息	V_{S30}	地下 30 m 深度内的等效剪切波速	USGS	

注: T 为目标建筑物的基本自振周期; t_1 和 t_2 分别为 5% 和 95% Arias 强度对应时间; t_f 为地震动总时长; 反应谱阻尼比取 5%。建筑物信息编码为建筑年代: ①1867-1910 年; ②1911-1924 年; ③1925-1987 年; ④1988-2011 年。外墙类型: ①混凝土; ②蒸压轻质混凝土; ③砂浆; ④土藏造; ⑤金属陶瓷; ⑥金属板; ⑦砌块。结构类型: ①木结构; ②钢结构; ③混凝土结构。柱子类型: ①木框架; ②钢结构; ③混凝土柱; ④防火涂层钢结构; ⑤其他。屋顶类型: ①瓷砖瓦; ②金属板; ③石板; ④混凝土; ⑤合成树脂; ⑥茅草。用途分类: ①住宅; ②店铺。

2 研究方法

2.1 研究技术路线

本研究的目的是得出与地震后建筑物损失相关性最强的地震动参数组合,从而提高损失预测的准确性、优化抗震设计和制定防灾减灾措施提供依据和参考。本文研究路线见图1,具体包括:

- 1) 收集 2011 年 3 月 11 日东日本大地震后实际震害资料,利用地震动空间插值方法计算每栋建筑处的 14 个地震动参数并为每栋建筑匹配 V_{S30} ,形成较为完善的实际震害数据库。
- 2) 对数据进行切片处理,指定模型输入特征及输出标签的类别。
- 3) 按照 8:2 的比例将数据集划分为训练集和测试集。
- 4) 选用 XGBoost (eXtreme gradient boosting)、随机森林 (random forest, RF)、LightGBM (light gradient boosting machine) 和 CatBoost (categorical boosting)^[27-30] 共 4 种机器学习分类算法,调用函数并设置模型初始超参数,建立 4 种初始预测模型。
- 5) 引入 SMOTE 过采样、贝叶斯超参数优化法后获得最优模型,对测试集进行预测并打分。
- 6) 通过 OA、P、R 及 F1 指标对 4 种模型进行综合评估并获得最优预测模型。
- 7) 利用最优预测模型对各输入特征进行重要性分析,迭代选出最优地震动参数组合,并进行结果分析。
- 8) 通过新西兰震害数据对筛选得到的最优参数组合进行结果验证。



图1 研究技术路线图

Fig. 1 Research technical roadmap

2.2 机器学习分类模型建立

数据集收集整理完毕后,将 1.5 节中确定的 23 个特征及标签输入机器学习模型进行训练。虽然地震动参数之间可能存在较高的相关性,但本研究并未利用相关性对特征进行剔除,原因有:①相关性较高的参数之间仍包含固有信息的差异^[17];②在地震动参数组合中,参数之间存在着复杂的相互作用,简单依据相关性剔除某些参数可能会忽略这些复杂关系和潜在影响。此外,由于不同的地震动参数及 V_{S30} 的量纲和数据分布不同,为了消除这些影响,在建立模型之前,利用式(1)对数据集中的连续型特征进行归一化处理,以确保数据的等效性。

$$X'_i = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

式中: X_i 为某一特征中第 i 个样本数据; X_{\max} 为该特征的最大值; X_{\min} 为该特征的最小值。

2.3 样本不均衡处理

由地震动衰减的客观规律可知,地震发生后影响场内未破坏和发生轻微破坏的建筑往往多于发生严重破坏的建筑,见表 1,导致实际震害数据分布不均衡。NGUYEN 等^[31]的研究表明,训练数据的不均衡分布会降低分类器的预测性能。因此,有必要采用相关方法降低数据不均衡分布带来的不利影响,提升模型分类性

能。现有研究中,非平衡数据的处理方法主要有过采样方法和欠采样方法。欠采样方法可能会导致多数数据丢失,从而破坏数据的原有特性,因此本研究不予采用。合成少数过采样技术 (synthetic minority oversampling technique, SMOTE)^[32] 是一种著名的过采样方法,通过在邻近的少数类样本间进行线性插值来合成新的样本,从而达到数据平衡,该方法可以在一定程度上减少过拟合风险,同时保持类别分布的自然性。本研究采用 SMOTE 过采样方法处理数据不均衡问题。

2.4 模型超参数优化

超参数是影响机器学习模型性能的关键参数,通过优化获得适当的超参数可以防止过拟合和欠拟合,以此提升模型的泛化能力和预测精度。超参数优化方法有网格搜索、随机搜索和贝叶斯优化三大类。网格搜索法 (grid search)^[33] 是通过穷举搜索的方式在给定的参数范围内遍历所有的参数组合,在测试集上精度最高的参数组合将作为最优超参数组合。但该方法在处理高维和大规模数据集时效率较低,且训练成本高。随机搜索 (random search)^[34] 是在搜索范围内随机选择超参数组合来训练和评估模型,以找到最优的超参数组合。尽管计算效率高,但由于搜索过程没有针对性地探索参数空间,可能会错过某些最优解。贝叶斯优化 (Bayesian optimization)^[35] 是一种基于贝叶斯概率论超参数调整方法,通过构建参数的后验概率分布来推断最优的超参数组合,该方法具有较高的运算效率,并且能够有效探索全局最优解,尤其适合于复杂的参数空间优化。因此,本研究采用贝叶斯优化方法选择最优超参数组合,4 种模型筛选得到的最优关键超参数见表 3,非关键参数采用默认数值。

表 3 基于贝叶斯优化后的 4 种模型最优超参数

Table 3 Optimal hyperparameters of four models based on Bayesian optimization

模型	超参数名称	最优取值	模型	超参数名称	最优取值
XGBoost	n_estimators	624.000	LightGBM	n_estimators	564.000
	max_depth	20.000		max_depth	11.000
	gamma	0.040		reg_lambda	5.210
	reg_lambda	27.536		reg_alpha	9.344
	reg_alpha	42.131		learning_rate	0.858
	learning_rate	0.546		num_leaves	50.000
Random Forest	n_estimators	898.000	CatBoost	depth	14.000
	max_depth	64.000		iterations	152.000
	min_samples_split	20.000		learning_rate	0.966
	min_samples_leaf	16.000		12_leaf_reg	47.502

2.5 性能评估

本研究选取整体准确率 OA、召回率 R、精确率 P、F1 值作为模型的性能评估指标,具体计算公式见表 4。准确率 OA 为模型正确预测目标的精度,即模型预测正确的样本个数占总样本数的比率。召回率 R 为模型正确分配实际标签的百分比,精确率 P 为模型正确分配预测标签的百分比。召回率 R 与精确率 P 在一定程度上存在权衡关系,通常使用 F1 值来平衡这 2 个指标并协助评估模型性能。

表 4 模型评价指标

Table 4 Evaluation index of model

评价指标	计算公式	评价标准
整体准确率	$OA = \frac{TP+TN}{TP+TN+FP+FN}$	OA、R、P、F1 ∈ [0, 1] 指标数值越大,模型性能越好
召回率	$R = \frac{TP}{TP+FN}$	
精确率	$P = \frac{TP}{TP+FP}$	
F1 值	$F1 = \frac{2PR}{P+R}$	

注:TP 为真阳性; FP 为假阳性; TN 为真阴性; FN 为假阴性。

3 结果分析

3.1 模型预测结果

采用 SMOTE 过采样和贝叶斯优化对 4 种机器学习模型优化后各模型的评价指标见表 5。对比 4 种模型得出:基于 XGBoost 算法模型的整体准确率和 F1 值最高,分别达到了 0.714、0.535。因此,本研究选择训练完毕的 XGBoost 模型为最终的损失预测模型。

表 5 模型性能对比

Table 5 Comparison of model performance

模型	召回率	精确率	F1 值	准确率	模型	召回率	精确率	F1 值	准确率
XGBoost	0.601	0.498	0.535	0.714	LightGBM	0.607	0.485	0.523	0.697
RF	0.609	0.493	0.531	0.703	CatBoost	0.599	0.480	0.517	0.695

为了直观展示 XGBoost 模型在 5 个破坏标签上的分类表现,采用混淆矩阵(confusion matrix)来评估其分类性能,最终模型在测试集上的分类结果见图 2。主对角线上的值分别表示模型对 5 种破坏等级的预测准确率,分别为 80.05%、48.59%、51.54%、49.60%、70.57%。模型的整体准确率为 71.39%。值得注意的是,模型可以较为准确地预测未破坏和发生毁坏的建筑物,这对灾后紧急救援具有重要意义。类别 1、类别 2、类别 3 的准确率较低,均在 50.00%左右,主要是因为建筑物发生破坏之后,不同破坏程度之间的建筑物属性区别度较低,相邻破坏等级的建筑物容易出现匹配失误,但这在建筑破坏等级评估中是可以接受的,如果机器学习错误将实际标签分配到跨越 2 个及以上的标签就需要引起较大的关注^[36]。

3.2 特征重要性分析

XGBoost 模型在训练的过程中可以检测各个特征对分类结果的影响。本研究采用置换特征重要性(permutation feature importance, PFI)^[37]和沙普利可加性模型解释方法(SHapley Additive exPlanation, SHAP)^[38]对模型进行解释。PFI 算法基于已经训练好的模型,通过随机打乱测试集中某一变量的数据顺序,统计打乱前后模型精度下降的程度,精度下降越多,表明模型对该特征的依赖程度越大;SHAP 是基于 Shapley 值的模型解释工具,该方法将传统方法与博弈论和局部解释联系起来用于解释各种黑箱模型,通过计算样本中每个特征的贡献,对模型进行解释,Shapley 值表示具有特征归因可加性的线性模型,将模型的预测值转换为每个特征的归因值之和,明确了各个特征对模型的影响程度。为了更好地对比衡量 2 种方法下的特征重要度,对每种方法得到的重要性进行归一化处理,评价结果见表 6。

表 6 置换重要性和 SHAP 重要性结果对比

Table 6 Comparison of the results of permutation importance and SHAP importance

特征	PFI	排名	SHAP	排名	特征	PFI	排名	SHAP	排名
震中距	1.000	1	1.000	1	CAV	0.083	14	0.113	13
PGA	0.480	2	0.690	2	A_{SI}	0.087	13	0.100	15
T_d	0.478	3	0.408	4	S_d	0.066	18	0.115	12
V_{SI}	0.328	5	0.498	3	A_{rms}	0.067	17	0.096	16
PGD	0.310	6	0.359	5	S_v	0.081	15	0.072	18
V_{S30}	0.380	4	0.282	6	I_a	0.038	20	0.086	17
屋顶类型	0.286	7	0.227	8	用途分类	0.071	16	0.021	20
建筑年代	0.264	8	0.237	7	外墙类型	0.048	19	0.006	21
PGV/PGA	0.166	9	0.170	9	I_c	0.018	21	0.030	19
PGV	0.126	11	0.127	11	柱子类型	0.004	22	0.000	23
层数	0.098	12	0.145	10	结构类型	0.000	23	0.002	22
S_a	0.130	10	0.104	14					

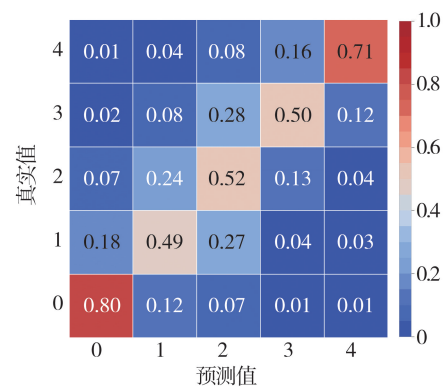


图 2 XGBoost 模型混淆矩阵

Fig. 2 Confusion matrix of XGBoost model

由于 2 种重要性评价方法的原理和评价指标不同,因此所得到的排序结果存在一定差异,但 2 种方法得到的特征排序基本一致。为了更加准确全面地评估各个特征对模型的贡献度,最终的特征重要性排序通过结合 2 种评价方法归一化后的加权特征重要性来确定,见图 3。

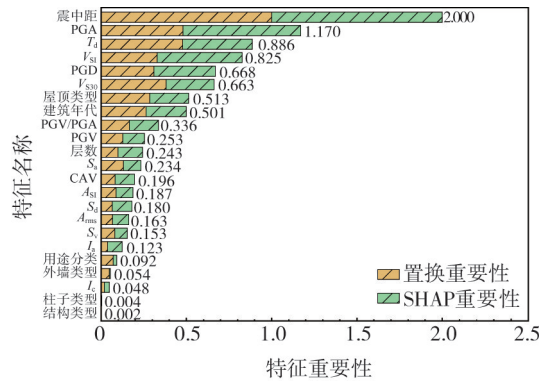


图 3 加权特征重要性排序

Fig. 3 Weighted feature importance ranking

3.3 关键地震动参数筛选

为了筛选出与震后建筑损失相关性最强的关键地震动参数组合,基于 3.2 节中加权后的特征重要性排序,采用前向特征搜索法逐步扩充特征集合 H ,将数据集按照 8:2 的比例划分为训练集和测试集,分别用于训练 XGBoost 模型和评估集合 H 对建筑损失等级的预测精度。关键参数组合应尽可能以较少的数量准确反映震后建筑损失。本文认为,当通过前向特征搜索法获得参数组合对应的模型精度与全部特征对应的最终模型精度相差 1.00%时,该参数组合即为可充分预测建筑破坏等级且包含全面地震动信息的最优地震动参数组合。

由图 3 可知,震中距是影响建筑损失最主要的因素,因此以震中距为基础创建特征集合 H_0 ,根据特征重要性排序从前到后逐步在集合 H_0 中加入特征,利用训练集样本构建模型,通过模型的整体准确率 OA 评估集合 H_0 对测试集建筑损失的预测精度,结果见图 4(a),其中不同颜色的图形代表不同类型的特征。其中,红色为地震动峰值参数;绿色为地震动频谱参数;黄色为地震动累积能量参数;蓝色为建筑物基本信息;黑色为震中距;紫色为地震动重要持时;浅蓝色为场地条件。

图 4(a) 中模型精度整体呈上升趋势,当集合 H_0 具有 2 个及以上特征时,模型精度总体趋于稳定。当模型中加入 12 个参数时,整体准确率可以达到 70.00%;相比之下,包含所有特征的模型准确率为 71.39%,相差不到 1.00%。表明选择前 12 个特征能够较好地保留关键信息,同时保持较高的模型准确率。所选 12 个特征分别是震中距、PGA、 T_d 、 V_{st} 、PGD、 V_{ss0} 、屋顶类型、建筑年代、PGV/PGA、PGV、层数、 S_a 。其中包含建筑物基本信息,场地条件,震中距以及 7 个地震动参数构成的组合。地震动参数数量居多,表明他们在联系地震动和建筑损失上起到了重要的桥梁作用。所选 7 个地震动参数的组合涵盖了振幅、频谱和持时 3 个地震动的关键要素。其中包含 PGA、PGD、PGV 这 3 个幅值参数; V_{st} 、PGV/PGA、 S_a 这 3 个频谱参数;1 个持时参数 T_d 。由此可见,幅值和频谱参数对模型性能有着重要影响,而累积能量参数对模型的贡献有限。幅值参数物理含义明确且可以直接反应地震动的强弱,在抗震设计中广泛应用,3 个幅值参数中 PGA 表现出最强的有效性,其次是 PGD 和 PGV;频谱参数中包含积分型谱强度参数 V_{st} ,这一现象可以解释为:在地震中结构发生非线性变形会导致周期延长,而 V_{st} 考虑了较长周期范围内(0.1~2.5 s)的反应谱值,因此表现出较好的效果;地震动持时的排序仅次于震中距和 PGA,其影响主要体现在非线性体系的最大反应和能量累积两个方面^[39],实际地震发生之后,地震动持续时间越长,对建筑物造成潜在破坏的可能性越大,因此地震动持时需要引起重视。

模型精度随特征集合 H_1 中地震动参数逐次递增的变化趋势见图 4(b),其能更加准确地评估地震动参数对模型精度的影响。模型精度随着特征集合 H_1 地震动参数数目的递增逐渐增加,说明地震动参数的加入能有效提升模型精度。当 H_1 中仅包含建筑信息、场地信息及震中距时,模型精度仅为 57.45%。加入 PGA 后,模型精度显著提升至 68.33%,进一步增加至 7 个参数后,模型精度达到 70.80%。当加入全部参数时,模型精度为 71.39%,与包含 7 个参数的模型精度相差不到 1.00%。以上结果说明 PGA、 T_d 、 V_{st} 、PGD、PGV/PGA、PGV、 S_a 这 7 个参数为最优参数组合,即可以用较少的参数且较为准确地反映地震损伤。

为了进一步与单一地震动参数作对比,本文在初始特征集合 H_1 的基础上分别增加 PGA 和 S_a 这 2 个广泛研究的地震动参数,训练后模型精度分别达到 68.33% (PGA) 和 67.80% (S_a)。表明与单一地震动参数 (PGA/ S_a) 相比,所选地震动参数组合可以弥补单一参数包含信息不足的问题,且能够更为全面地反映地震动特征。此外发现,上述过程中,PGA 的结果优于 S_a ,可能是因为本研究在计算 S_a 时参考了 AMANAT 等^[40] 的研究,经验性地使用 $N/10$ 估计房屋地自振周期(其中 N 为房屋层数),而建筑物进入非线性阶段后刚度会降低,使得获取可靠的自振周期变得困难,最终导致 S_a 的计算结果存在一定偏差,相反地,PGA 反映的是地面运动地强度,能够更加直接地反映建筑物遭受的地震作用。

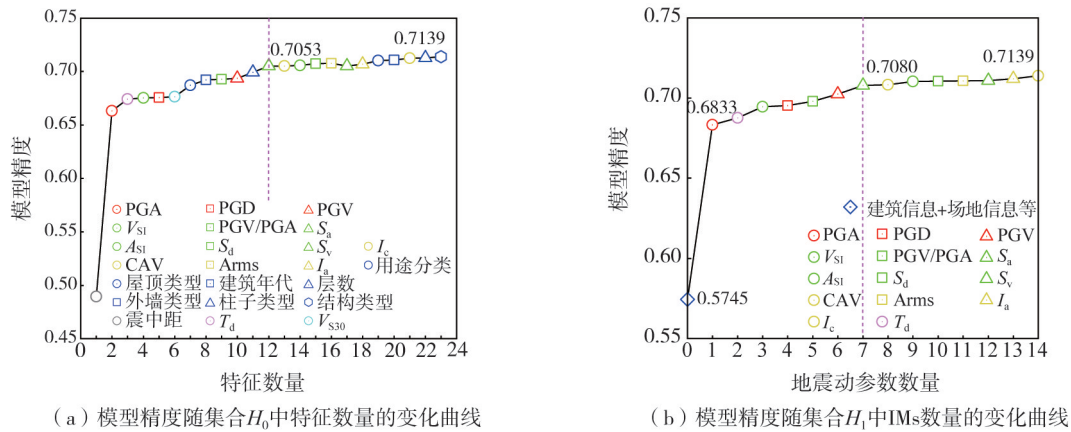


图 4 模型精度随特征集合 H 中特征数量的变化曲线

Fig. 4 Curves of model accuracy changing with the number of features in feature set H

4 基于新西兰震害资料的结果对比

新西兰位于太平洋板块和澳大利亚板块交界处,2 个板块的相互作用产生了显著的地质活动,导致地震频发^[41]。新西兰地震委员会 (Earthquake Commission, EQC) 成立至今积累了大量详实的房屋震后损失资料。为了验证所选最优参数组合的有效性,基于新西兰造成损失建筑较多的 3 次地震事件共 115739 条详细震害数据,建立震后建筑损失预测模型。由于新西兰和日本数据库中记录的基本信息不同,因此所选特征也存在一定差异,考虑到 3 次地震事件基本信息不同,因此将震级和震源深度作为 2 个新增的输入特征。

目前尚未针对新西兰地区提出专门的 GMPE,但新西兰与日本同处于太平洋板块的边缘,且均为狭长的岛国,因此采用基于日本观测数据得到的 GMPE 来计算新西兰地区的地震动参数是较为合理的选择,赵登科等^[42] 已经证明了日本地区的 GMPE 应用于新西兰地区的有效性。在计算 14 个地震动参数时,仍采用 1.2 节中适用于日本及其他地区的 GMPE,结合震害详细资料最终确定 29 个影响因素作为 XGBoost 分类模型的输入特征,同样以 ATC-13 划分的破坏等级作为标签。3 次地震事件的基本信息及破坏建筑分布见表 7。

表 7 新西兰 3 次地震基本信息及破坏建筑分布

Table 7 Basic information and distribution of damaged buildings of three earthquakes in New Zealand

事件时间	震级/ M_w	震源深度/km	构造分类	震源位置	破坏建筑分布
2010-09-04	7.2	11	浅地壳地震	43°32'S 172°10'E	建筑数量/条 57 506
2010-02-22	6.2	5	浅地壳地震	43°35'S 172°41'E	建筑数量/条 42 767
2011-06-13	6.0	7	浅地壳地震	43°35'S 172°44'E	建筑数量/条 15 466

在对 XGBoost 模型训练和优化后其整体精度约达到 71.85%。结合 FPI 和 SHAP 这 2 种方法对模型进行解释并获取加权后的特征重要性评分,排序结果见图 5。由图可知,排名靠前的 7 个地震动参数分别是 PGA、PGD、 S_a 、 V_{SI} 、 T_d 、PGV/PGA、CAV。其中包含 2 个地震动幅值参数为 PGA、PGD;3 个频谱参数为 S_a 、 V_{SI} 、PGV/PGA;1 个地震动持时参数为 T_d ;1 个累积能量参数为 CAV。这一结果与日本的研究发现相符,表明幅值、频谱和持时参数与震后建筑损失间存在关键联系,而累积能量参数只有 1 个且排序靠后,与损失间

相关性较弱。此外,新西兰排序前 7 位的地震动参数和日本筛选出的最优参数组合整体上基本一致,7 个参数中有 6 个相同,差异在于日本最优地震动参数组合中的 PGV 在新西兰的前 7 位参数中被 CAV 替代,而 PGV 在新西兰地震动参数中的排序为 8,仅次于 CAV。

为了验证所选地震动参数组合的有效性,基于建筑信息、场地信息等创建特征集合 H_2 ,初始模型精度 56.53%。将 3.3 节中选出的地震动参数组合 PGA、 T_d 、 V_{SI} 、PGD、PGV/PGA、PGV、 S_a 依次加入模型并记录精度变化情况,见图 6。当模型中包含最优参数组合时,精度为 70.97%,与包含所有特征的最终模型精度仅相差 0.88%,证明了最优参数组合的有效性。当在初始特征集合 H_2 中分别增加单一参数 PGA 和 S_a 后,模型精度分别为 68.07%、63.63%。由初始建筑物自振周期计算得到的 S_a 无法准确表征建筑物在非线性阶段的响应,而 PGA 直接反映了建筑物承受的地震强度,这造成了模型精度的差异。以上结果表明,最优参数组合比单一参数更具优势,同时涵盖了大部分关键地震动信息,显著降低了地震动维度。

综上所述,以新西兰 3 次地震震后建筑损失资料作为验证,建立建筑损失等级预测模型,并对 29 个输入特征进行重要性分析。将关键地震动参数与最优参数组合进行比较,结果显示 2 个地区筛选得到的关键参数基本一致,此外,包含最优参数组合的模型精度与最终模型精度相差不到 1.00%。研究表明,所选参数组合具有较高的信息覆盖率,并验证了其信息完备性且具有良好的地区泛化性能。

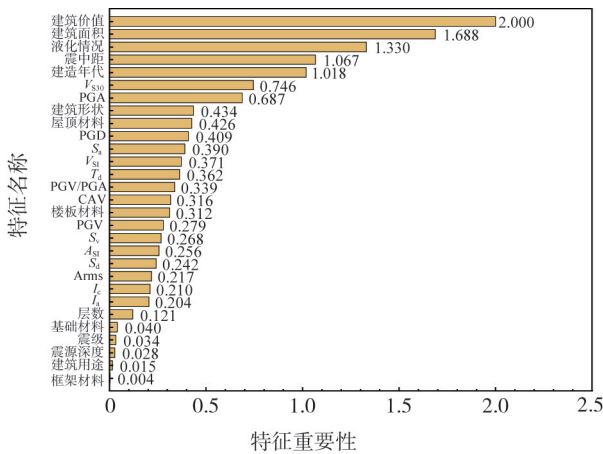


图 5 新西兰加权特征重要性排序
Fig. 5 Weighted feature importance ranking in New Zealand

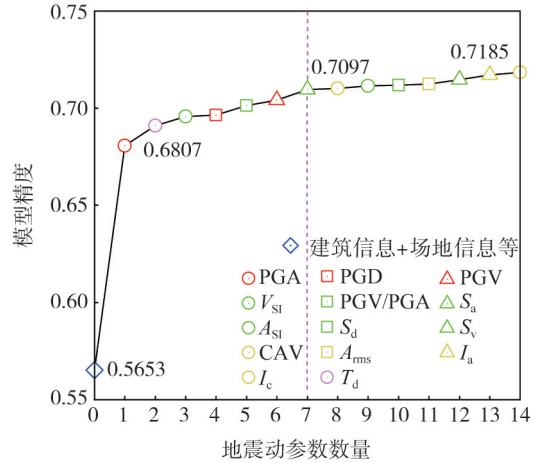


图 6 模型精度随集合 H_2 中 IMs 数量的变化曲线
Fig. 6 Curves of model accuracy changing with the number of IMs in set H_2

5 结论

本文基于 2011 年 3 月 11 日东日本大地震后大量详实的建筑损失数据,利用机器学习算法建立结构、场地及地震动特性与建筑破坏等级之间的非线性映射关系,并比选出关键地震动参数组合。以新西兰 3 次地震实际损失资料作为验证,证明了所选参数组合用于描述地震动信息的完备性和有效性,得出如下主要结论:

1) 与传统利用数值模拟开展的研究不同,本研究利用大量详实的地震损失数据,采用机器学习方法从 14 个地震动参数中比选出影响地震动破坏势的关键参数组合。解决了单一参数包含地震信息不足和多元参数构建复杂的问题;极大地改善了现有同类研究中数据受限、地震动参数间的内在关联被忽视、采用的结构响应指标无法反映建筑真实破坏程度、所选参数缺乏足够区域及结构类型泛化能力的问题。

2) 基于 4 种机器学习算法(XGBoost、RF、LightGBM、Catboost)建立震后建筑损失等级预测模型。采用 SMOTE 过采样和贝叶斯超参数优化法后,对比 4 种模型在测试集上的表现,结果表明 XGBoost 模型精度最高,整体准确率达到 71.39%,其次是 RF、LightGBM、CatBoost。

3) 结合置换重要性和 SHAP 重要性 2 种评价方法对训练完毕的 XGBoost 模型进行特征重要性分析,利用前向特征搜索法选出 7 个和损失相关性最强的地震动参数组合,依次为 PGA、 T_d 、 V_{SI} 、PGD、PGV/PGA、PGV、 S_a 。结果表明幅值参数和频谱参数与损失的相关性最强,而累积能量参数与损失的相关性较弱;另外,地震动持时的重要性需要引起关注。

4)以新西兰3次地震实际损失资料作为验证,证明所选参数组合涵盖了较为全面的地震动信息且具有一定的地区和未知事件泛化能力。研究成果有助于准确评估地震动的破坏潜力,同时为震前优化建筑抗震设计、震后应急评估与恢复提供有力依据。

5)本文在进行地震动参数筛选时,主要关注地震动参数与震后建筑损失间的联系,但未考虑到地震动参数与震后建筑损失分布的随机性和离散性之间的统计关系。因此,未来的研究应综合考虑这些复杂多元因素,以提升对建筑物在地震中的风险识别和损伤评估。

参考文献:

- [1] BURATTI N. A comparison of the performances of various ground-motion intensity measures[C]//Proceedings of the 15th World Conference on Earthquake Engineering, Lisbon, Portugal; Sociedade Portuguesa de Engenharia Sismica, 2012: 24-28.
- [2] 刘巴黎, 胡进军, 谢礼立. 基于弹性网络回归的地震动参数排序与比选[J]. 哈尔滨工业大学学报, 2024, 56(1): 54-62.
LIU Bali, HU Jinjun, XIE Lili. Ranking and comparison of ground motion parameters based on elastic net regression[J]. Journal of Harbin Institute of Technology, 2024, 56(1): 54-62. (in Chinese)
- [3] 吴梓楠, 韩小雷, 马建峰, 等. 基于机器学习的震动强度指标敏感性分析与破坏势评估[J]. 建筑结构学报, 2023, 44(11): 216-225, 235.
WU Zinan, HAN Xiaolei, MA Jianfeng, et al. Sensitivity analysis of ground motion intensity measures and evaluation of potential damage based on machine learning[J]. Journal of Building Structures, 2023, 44(11): 216-225, 235. (in Chinese)
- [4] 李爽, 谢礼立, 郝敏. 地震动参数及结构整体破坏相关性研究[J]. 哈尔滨工业大学学报, 2007, 39(4): 505-509, 560.
LI Shuang, XIE Lili, HAO Min. Correlation between seismic ground motion parameters and their relationship with overall damage to structure[J]. Journal of Harbin Institute of Technology, 2007, 39(4): 505-509, 560. (in Chinese)
- [5] PINZON L A, VARGAS-ALZATE Y F, PUJADES L G, et al. A drift-correlated ground motion intensity measure; Application to steel frame buildings[J]. Soil Dynamics and Earthquake Engineering, 2020, 132: 106096.
- [6] KOSTINAKIS K, ATHANATOPOULOU A, MORFIDIS K. Correlation between ground motion intensity measures and seismic damage of 3D R/C buildings[J]. Engineering Structures, 2015, 82: 151-167.
- [7] 叶列平, 马千里, 缪志伟. 结构抗震分析用震动强度指标的研究[J]. 地震工程与工程振动, 2009, 29(4): 9-22.
YE Lieping, MA Qianli, MIAO Zhiwei. Study on earthquake intensities for seismic analysis of structures[J]. Journal of Earthquake Engineering and Engineering Dynamics, 2009, 29(4): 9-22. (in Chinese)
- [8] 陈健云, 李静, 韩进财, 等. 震动强度指标与框架结构响应的相关性研究[J]. 振动与冲击, 2017, 36(3): 105-112, 144.
CHEN Jianyun, LI Jing, HAN Jincan, et al. Correlation between ground motion intensity indexes and seismic responses of frame structures[J]. Journal of Vibration and Shock, 2017, 36(3): 105-112, 144. (in Chinese)
- [9] LIU T T, YU X H, LU D G. An approach to develop compound intensity measures for prediction of damage potential of earthquake records using canonical correlation analysis[J]. Journal of Earthquake Engineering, 2020, 24(11): 1747-1770.
- [10] OZMEN H B. Developing hybrid parameters for measuring damage potential of earthquake records; Case for RC building stock[J]. Bulletin of Earthquake Engineering, 2017, 15(7): 3083-3101.
- [11] GEHL P, SEYEDI D M, DOUGLAS J. Vector-valued fragility functions for seismic risk evaluation[J]. Bulletin of Earthquake Engineering, 2013, 11(2): 365-384.
- [12] 胡进军, 刘巴黎, 谢礼立. 基于因子分析的地震动特征提取及潜在破坏势评估[J]. 工程力学, 2022, 39(10): 140-151, 172.
HU Jinjun, LIU Bali, XIE Lili. Factor analysis-based ground motion feature extraction and the measurement of the potential structural damage[J]. Engineering Mechanics, 2022, 39(10): 140-151, 172. (in Chinese)
- [13] LUCO N, CORNELL C A. Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions[J]. Earthquake Spectra, 2007, 23(2): 357-392.
- [14] 刘亭亭, 于晓辉, 吕大刚. 震动多元强度参数主成分与结构损伤的相关性分析[J]. 工程力学, 2018, 35(8): 122-129, 137.
LIU Tingting, YU Xiaohui, LV Dagang. Analysis of correlation between principal components of multivariate earthquake intensity measures and structural damage[J]. Engineering Mechanics, 2018, 35(8): 122-129, 137. (in Chinese)
- [15] LIU T T, LU D G, YU X H. Development of a compound intensity measure using partial least-squares regression and its statistical evaluation based on probabilistic seismic demand analysis[J]. Soil Dynamics and Earthquake Engineering, 2019, 125: 105725.
- [16] 刘亭亭, 于晓辉, 吕大刚. RC 框架基于典型相关分析的地震动多元破坏势评估[J]. 土木工程学报, 2019, 52(1): 27-36, 107.
LIU Tingting, YU Xiaohui, LU Dagang. Multivariate damage potential estimation of earthquake ground motions for RC frames based on canonical correlation analysis[J]. China Civil Engineering Journal, 2019, 52(1): 27-36, 107. (in Chinese)
- [17] XU Y J, LU X Z, TIAN Y, et al. Real-time seismic damage prediction and comparison of various ground motion intensity measures based on machine learning[J]. Journal of Earthquake Engineering, 2022, 26(8): 4259-4279.
- [18] YAN Y X, XIA Y, YANG J P, et al. Optimal selection of scalar and vector-valued seismic intensity measures based on Gaussian Process Regression[J]. Soil Dynamics and Earthquake Engineering, 2022, 152: 106961.
- [19] DANG H T, WANG Z F, ZHAO D K, et al. Ground motion prediction model for shallow crustal earthquakes in Japan based on XGBoost with

- Bayesian optimization[J]. *Soil Dynamics and Earthquake Engineering*, 2024, 177: 108391.
- [20] ZHAO D K, WANG Z F, WANG J M, et al. A rapid estimation method for post-earthquake building losses[J]. *International Journal of Disaster Risk Science*, 2023, 14(3): 428-439.
- [21] PODILI B, RAGHUKANTH S T G. Ground motion prediction equations for higher order parameters[J]. *Soil Dynamics and Earthquake Engineering*, 2019, 118: 98-110.
- [22] MORASCA P, ZOLEZZI F, SPALLAROSSA D, et al. Ground motion models for the Molise Region (southern Italy)[J]. *Soil Dynamics and Earthquake Engineering*, 2008, 28(3): 198-211.
- [23] CAUZZI C, FACCIOLO E, VANINI M, et al. Updated predictive equations for broadband (0.01~10 s) horizontal response spectra and peak ground motions, based on a global dataset of digital acceleration records[J]. *Bulletin of Earthquake Engineering*, 2015, 13: 1587-1612.
- [24] GUPTA A K. *Response spectrum method in seismic analysis and design of structures*[M]. New York; Routledge, 2017.
- [25] WALD D J, ALLEN T I. Topographic slope as a proxy for seismic site conditions and amplification[J]. *Bulletin of the Seismological Society of America*, 2007, 97(5): 1379-1395.
- [26] MCCORMACK T C, RAD F N. An earthquake loss estimation methodology for buildings based on ATC-13 and ATC-21[J]. *Earthquake Spectra*, 1997, 13(4): 605-621.
- [27] CHEN T Q, GUESTRIN C. XGBoost: A scalable tree boosting system[C]//*Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco California USA. New York: ACM, 2016: 785-794.
- [28] BREIMAN L. Random forests[J]. *Machine Learning*, 2001, 45: 5-32.
- [29] KE G L, MENG Q, FINLEY T, et al. LightGBM: A highly efficient gradient boosting decision tree[C]//*Proceedings of Advances in Neural Information Processing Systems*. Long Beach, CA, USA; Neural Information Processing Systems Foundation, 2017: 3149-3157.
- [30] PROKHORENKOVA 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*. Cambridge, USA; MIT Press, 2018: 6639-6649.
- [31] NGUYEN H M, COOPER E W, KAMEI K. Borderline over-sampling for imbalanced data classification[J]. *International Journal of Knowledge Engineering and Soft Data Paradigms*, 2011, 3(1): 4-21.
- [32] CHAWLA N V, BOWYER K W, HALL L O, et al. SMOTE: Synthetic minority over-sampling technique[J]. *Journal of Artificial Intelligence Research*, 2002, 16: 321-357.
- [33] SUN Y T, DING S F, ZHANG Z C, et al. An improved grid search algorithm to optimize SVR for prediction[J]. *Soft Computing*, 2021, 25(7): 5633-5644.
- [34] BERGSTRA J, BENGIO Y. Random search for hyper-parameter optimization[J]. *Journal of Machine Learning Research*, 2012, 13(2): 281-305.
- [35] SNOEK J, LAROCHELLE H, ADAMS R P. Practical bayesian optimization of machine learning algorithms[C]//*Proceedings of Advances in Neural Information Processing Systems*. New York, USA; Curran Associates Inc, 2012, 25: 2951-2959.
- [36] MANGALATHU S, SUN H, NWEKE C C, et al. Classifying earthquake damage to buildings using machine learning[J]. *Earthquake Spectra*, 2020, 36(1): 183-208.
- [37] ALTMANN A, TOLO I L, SANDER O, et al. Permutation importance: A corrected feature importance measure[J]. *Bioinformatics*, 2010, 26(10): 1340-1347.
- [38] LUNDBERG S M, LEE S-I. A unified approach to interpreting model predictions[C]//*Proceedings of Advances in Neural Information Processing Systems*, New York, USA; Curran Associates Inc, 2017, 30: 4765-4774.
- [39] 周锡元, 王广军, 苏经宇. 场地·地基·设计地震[M]. 北京: 地震出版社, 1990.
ZHOU Xiyuan, WANG Guangjun, SU Jingyu. *Site·Foundation·Design earthquake*[M]. Beijing: Seismological Press, 1990. (in Chinese)
- [40] AMANAT K M, HOQUE E. A rationale for determining the natural period of RC building frames having infill[J]. *Engineering Structures*, 2006, 28(4): 495-502.
- [41] 刘渊, 王自法, 赵登科. 基于新西兰详细地震破坏数据的建筑物构件损失分析[J]. *地震工程与工程振动*, 2021, 41(5): 186-195.
LIU Yuan, WANG Zifa, ZHAO Dengke. Building component level loss analysis based on detailed earthquake loss data in New Zealand[J]. *Earthquake Engineering and Engineering Dynamics*, 2021, 41(5): 186-195. (in Chinese)
- [42] 赵登科, 王自法, 刘渊, 等. 基于新西兰实际震害资料的地震损失不确定性分析[J]. *地震工程与工程振动*, 2021, 41(2): 84-95.
ZHAO Dengke, WANG Zifa, LIU Yuan, et al. Earthquake loss uncertainty based on detailed loss data in New Zealand[J]. *Earthquake Engineering and Engineering Dynamics*, 2021, 41(2): 84-95. (in Chinese)