

基于随机森林算法的建筑物构件对地震损失的重要度分析

刘良巧^{1,2}, 王自法^{1,2,3}, 梁梓豪^{1,2}, 刘渊^{1,2}

(1. 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;
2. 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:建筑物地震损失是从构件破坏开始的, 确定建筑物构件对地震损失的重要度有助于寻找结构的抗震薄弱构件。传统的抗震研究多数关注建筑物的结构构件破坏, 对非结构构件的重要性研究不足, 这也造成非结构构件的地震损失远超过结构构件的地震损失。基于新西兰地震保险数据集中大量详实的建筑物地震构件损失数据, 采用随机森林的特征重要性(feature importance, FI)、排列重要性(permutation importance, PI)以及SHAP值(shap_values)这3种重要性分析方法对建筑物各损失构件在总体损失中的重要性大小进行排序, 并采用赋值排序相加法和归一化相加的方法统筹表示重要性排序。归一化相加法不仅能判别构件的相对重要性, 还可以定量地表达不同构件的重要性差异。结果表明, 在所分析的12种构件类型中, 非结构构件的重要度明显高于结构构件的重要度, 其中墙面装修最重要, 板式基础次之, 屋顶框架的重要性最弱, 未来需要更加注重墙面装修和建筑基础的抗震性能。

关键词:地震损失; 结构构件损失; 非结构构件损失; 随机森林算法; 重要性排序
中图分类号:P315.9 **文献标识码:**A

Importance analysis of building components for earthquake losses based on the random forest algorithm

LIU Liangqiao^{1,2}, WANG Zifa^{1,2,3}, LIANG Zihao^{1,2}, LIU Yuan^{1,2}

(1. Key Laboratory of Earthquake Engineering and Engineering Vibration, Institute of Engineering Mechanics, China Earthquake Administration, Harbin 150080, China; 2. Key Laboratory of Earthquake Disaster Mitigation, Ministry of Emergency Management, Harbin 150080, China)

Abstract: Earthquake damage to buildings always starts from a damaged component, so the importance analysis of building components helps identify the weakest component in the structure. Traditional researches often focused on structural components, and the importance of non-structural components was not well studied, resulting in the loss of non-structural components being much larger than the loss by structural components. In this paper, based on a large amount of detailed component earthquake loss data from the New Zealand earthquake, three importance analysis methods from random forest, namely the feature importance(FI), the permutation importance(PI) and the SHAP value(shap_values), were used to rank the importance of each component for earthquake losses. Assignment sort summation and normalized summation were used to assemble the results from the three methods representing the combined importance of each component. The normalized summation method can not only identify the relative importance of the components, but also quantify the significance of the differences among different components. The results demonstrate that of the 12 types of components analyzed, non-structural components are more important than

收稿日期:2022-11-19; 修回日期:2023-03-15

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:刘良巧(1997—),女,硕士研究生,主要从事地震易损性的研究。E-mail:2504582606@qq.com

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

structural components in earthquake loss contribution, and wall ornaments are the most important component, followed by slab foundations, but roof framing is the least influential. The results indicate that future earthquake resistance efforts should put more focus on wall ornaments and slab foundations.

Key words: earthquake loss; structural component loss; non-structural component loss; random forest algorithm; importance ranking

0 引言

建筑物是人类生存和发展的不可或缺的生活保障之一。以往的震害统计表明,建筑物破坏是造成社会经济损失和人员伤亡的主要原因^[1]。建筑物地震损失总是从某个构件的破坏开始的,如果能科学地分析出各构件对地震损失重要性的大小,并改善重要性比较大的构件的抗震性能,对减轻建筑物地震损失有着至关重要的作用。

现有的从建筑物构件层次进行地震损失重要性分析的研究依然较少,且多关注于在地震中发生破坏的单一构件对建筑物整体的影响,如墙板^[2]、墙体结构^[3]、窗户^[4]、电器设备^[5]、吊顶^[6]、管线系统^[7]和电梯^[8]等。少有的多构件重要性分析中,GHARAIBEH等^[9]首先提出了一个用于评估结构中各元素的结构重要性框架,并确定了单个结构构件对建筑物整体系统性能的贡献,但其只分析了结构构件这一个类型。PERRONE等^[10]详述了在2016年意大利地震中观察到的填充墙和天花板系统等多种非结构构件的破坏情况,明确了填充墙是受影响最大的非结构构件,然而该文只是对最常见的非结构性抗震性能进行了概述,并没有与结构构件进行对比分析,所以并没有得出全面的结论。GAUTAM等^[11]建立了受2015年尼泊尔地震影响的钢筋混凝土建筑的地震脆弱性函数,认为梁柱连接是建筑中最容易受到破坏的结构构件,并提出了非结构构件比结构构件更脆弱的想法,但这仅仅是通过观察单个建筑的破坏情况得到的猜想,缺乏大量详细震害资料的验证。刘渊等^[12]基于新西兰的详细破坏资料分别对结构构件和非结构构件的破坏进行了地震易损性分析,得出了在非结构构件中墙面饰品更易发生破坏的结论,但该方法没有将2种类型的多个构件放在一起进行分析,同样也没有得出各类损失构件在地震损失中的整体重要性排序。传统的地震损失分析方法主要依赖于经验法、统计法^[13]和组合法^[14]等,但由于受到样本数量、成本等的限制,更无法体现各损失构件在地震损失中的重要程度。

为了进一步探究建筑物构件对地震损失贡献的重要性,本研究基于新西兰地震保险数据集中的16次地震事件中超过12万栋建筑物的近60万条构件详细破坏记录,利用随机森林算法中的3种重要性分析方法对建筑物各损失构件进行重要性排序。基于大量真实建筑物构件地震损失数据进行分析不仅能厘清各构件在地震损失中的重要度,也能为我国今后抗震设防和震后救援乃至韧性城乡的深入发展等提供重要的参考资料。

1 数据来源

1.1 建筑物数据来源及损失构件分类

新西兰处在环太平洋火山地震带上,是一个地震频发的国家。据不完全统计,新西兰每年记录在案大大小小的地震就多达15000次。虽然每年大约只有3000次的地震有感,但是其中造成经济损失和人员伤亡的破坏性地震数量不多,所以详细的地震破坏资料极其珍贵。经过新西兰地震保险制度多年来的不断完善以及地震灾害经验积累,其地震保险制度成为了全世界最成熟的地震保险制度之一,同时也形成了新西兰特有的地震保险资料数据库。这些数据是开展建筑物地震损失详细研究的最难得资源之一。为了能够得到更加可靠的建筑物损失构件重要性排序,本文通过检索这个数据库中包含的新西兰地震相关的基本信息数据集、保险索赔数据集和建筑物详细破坏信息数据集,共整理出了新西兰16次地震事件中受到破坏的124427栋建筑物的52种构件加以重要性分析。

建筑物构件在地震中受到的破坏可以归类为结构构件破坏和非结构构件破坏。结构构件(structural components)主要指结构(施工图上)的承重构件,支承荷载起骨架作用的构件或由其组成的整体,如房屋中

的梁、柱、屋架和基础等构件^[15]。非结构构件(non-structural components)由水平非结构构件、垂直非结构构件和建筑物内部物品三部分组成^[16]。根据上述构件的定义,本文整理出的52种构件中有5种结构构件和47种非结构构件。结构构件分别是墙体框架、环形基础、板式基础、桩和屋顶框架,它们在所有损失构件数量中所占的比例分别是3.1%、6.0%、3.9%、3.9%、0.5%。考虑到本文数据集中大部分的非结构损失构件数量不多,损失不高,进行构件层次的单独分析意义不大,故将剩下的47种非结构构件重新整合为7种。其中墙面装修是一种应用于垂直表面以装饰和保护墙体的材料,本文主要指的是墙布和涂料等;墙面涂层是用材料覆盖建筑物的外墙或内墙以改善其外观、耐用性和隔热性的过程,本文的墙面覆层包括了填充墙和蜂窝板等。各类别的构件数量分布如图1所示,其中屋内财物由门、窗户、供暖设备、工作台、车库门、淋浴、洗手池和马桶等22种屋内使用设备组成,烟囱等余下的20种非结构构件则用其他非结构来表示。

1.2 建筑物构件损失等级划分

基于新西兰详细的地震破坏资料,也为了能够从构件层次更加详尽地分析建筑物地震损失,本文选择借鉴破坏比^[17]的概念,用损失比来描述建筑物构件的破坏情况。破坏比是建筑物损失与建筑物总价值的比值,那么本文的总损失比(damage ratio, Dr)可由公式计算为

$$\text{总损失比}(Dr) = \frac{\text{建筑构件总损失}}{\text{建筑总价值}} \quad (1)$$

总损失比(Dr)是区间(0,1]上的连续性指标,相比其他结构损伤指数或状态参数,其能够更直观地展现建筑物各构件的破坏程度。

不同的国家由于规范和国情的不同,总损失比和破坏状态之间的对应关系无法完全一致。如美国ATC-13《应用技术委员会》^[18]针对78种地震工程设施类别给出了对应于完好、基本完好、轻微破坏、中等破坏、严重破坏、极重破坏和毁坏等7个损伤状态的损伤系数;我国国家标准GB/T 18208.4—2011《地震现场工作 第4部分:灾害直接损失评估》^[19]给出了我国常见结构体系对应于基本完好、轻微破坏、中等破坏、严重破坏和毁坏等5个损伤状态的损失比的取值范围。考虑到新西兰国家标准将建筑物地震破坏分成了4个等级,却并未给出损失比的具体数值。故本文采用国际上通用的美国ATC-13的损伤等级划分标准,如表1所示。

表1 ATC-13 建筑物损失比划分标准

Table 1 ATC-13 standard for the division of building loss ratio

| 划分标准 | (0,1%) | [1%,10%) | [10%,30%) | [30%,60%) | [60%,100%) | 100% |
|------|--------|----------|-----------|-----------|------------|---------|
| 损伤等级 | 基本无损坏 | 轻度损坏 | 中度损坏 | 严重损坏 | 重度损坏 | 全部损坏及倒塌 |

由于本文仅研究在地震中受到破坏的构件,故根据美国ATC-13损伤等级划分标准,将数据中总损失比为(0,1%)的建筑归类为“轻微破坏”,将总损失比[10%,30%)的建筑归类为“中等破坏”类,将总损失比为[30%,60%)的建筑归类为“严重破坏”,将总损失比为[60%,100%)以及极重破坏和毁坏划分为“毁坏”一类,得到了如图2所示的0、1、2、3这4个标签类别数据集。随机森林具备的样本可容量大的优势,不仅可以减少异常值对结果的影响,还能解决由于样本量过小导致的总体缺乏足够代表性的问题^[20]。同时在随机森林模型中可以引入class_weight = “balanced”来平衡样本,从而保障了本文在标签0和标签3的数据量存在较大差异的情况下仍然可以得到比较可靠的结果。

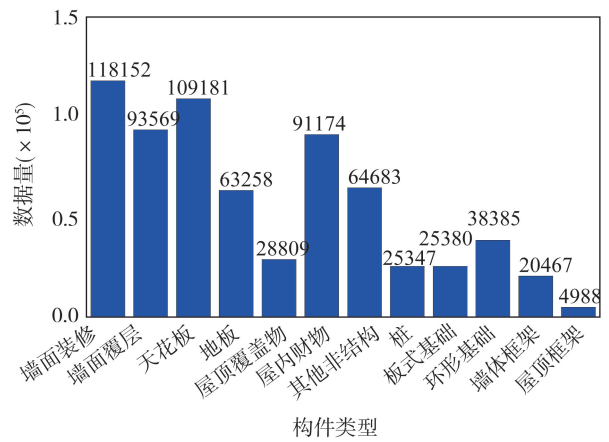


图1 建筑物各损失构件数量分布

Fig. 1 Distribution of the number of damaged components

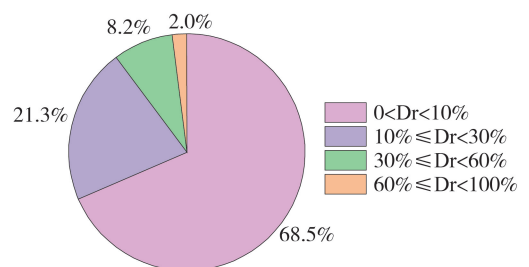


图2 不同标签类别数量分布

Fig. 2 Distribution of the number of different label categories

2 研究方法

2.1 随机森林重要性分析原理

随机森林算法已经被成功地应用于各种实际问题中,其准确性、足够灵活的机制以及处理小样本和高维特征空间的能力也得到了普遍认可^[21]。作为一种基于决策树的集成学习算法,随机森林不仅能够很好地处理连续型数据的回归任务,而且能快速解决离散型数据的分类问题。最主要的是,随机森林为每个变量提供了可解释的全局变量重要性度量^[22],故本文可以通过随机森林算法非常方便地获取建筑物各构件对地震损失贡献的重要性排序。

特征的重要程度可以通过随机改变特征值,计算特征值改变前后测试数据集的误差得到,这个误差又被称为该特征的袋外误差(out-of-bag errors, errOOB)^[23]。假设随机森林中有 N 棵树,那么特征的重要度 I_{RF} 计算公式为

$$I_{\text{RF}} = N_i \sum_{x=1}^{N_i} |\text{errOOB}_2 - \text{errOOB}_1| \quad (2)$$

式中: N_i 为特征在随机森林中出现的次数; errOOB_2 为 N 棵决策树中的特征值加入噪声干扰后再次计算得到的袋外误差; errOOB_1 为 N 棵决策树的特征值所对应的袋外数据计算出的袋外误差。这就意味着在给某个特征随机加入噪声之后,袋外数据的准确率下降得越多,这个特征对于样本的分类结果影响就越大,它的重要程度也越高。袋外误差与交叉验证的结果比较接近,所以在随机森林中就无需再进行交叉验证或误差无偏估计计算^[24]。

本文通过袋外误差方法获取到决策树生长的最优参数后,再用随机森林中的特征重要性(feature importance, FI)、排列重要性(permutation importance, PI)以及 shap_values 这 3 种重要性分析方法来进行损失构件的重要性排序。FI 方法可以通过信息熵增益(entropy)计算出每个特征对于模型的重要度,但不足之处在于其无法判断特征与最终预测结果之间的关系;shap_values 方法则可以很好地弥补 FI 方法的不足,既可以解释模型,还能反映出每个样本中特征的影响力,同时还能准确地体现出影响的正负性;PI 方法是将测试集数据中某一特征列数据打乱,更改数据后根据模型得分变化来判断该特征对于模型结果的影响程度,可以作为 FI 方法的结果的检验。

2.2 各损失构件相关性分析

模型的准确率与各个损失构件之间的相关性有关,相关性越强,表明模型的准确率就越低,所以进行相关性分析是很有必要的。容差和方差膨胀因子都是相关性强弱的表征,其中容差和方差膨胀因子互为倒数。当容差的值大于 0.1 或方差膨胀因子小于 10 时,表示各个因子共线性程度低^[25]。多重共线性越低则说明各个损失构件之间的相关性越弱,反之则认为相关性较强。一旦出现了构件强相关性,就需要对数据进行降维从而保证模型的准确率。

本文采用 SPSS 对 12 种类型的损失构件进行多重共线性检验,计算出的容差以及方差膨胀因子如表 2 所示。由表 2 可知,所有构件类型的容差均大于 0.1,各方差膨胀因子的大小也均在 10 以下,说明各类型的损失构件之间的共线性程度低,相关性较弱。综上所述,本研究不需要进行降维处理,各类别损失数据可以应用到随机森林模型中。

表 2 多重共线性诊断表

Table 2 Multi-collinearity diagnostic table

| 损失构件 | 墙面 装修 | 墙面 覆层 | 天花板 | 地板 | 屋顶 覆盖物 | 屋内 财物 | 其他 非结构 | 桩 | 板式 基础 | 环形 基础 | 墙体 框架 | 屋顶 框架 |
|--------|----------|----------|-------|-------|-----------|----------|-----------|-------|----------|----------|----------|----------|
| 容差 | 0.538 | 0.672 | 0.513 | 0.762 | 0.690 | 0.616 | 0.674 | 0.684 | 0.923 | 0.566 | 0.759 | 0.693 |
| 方差膨胀因子 | 1.858 | 1.488 | 1.948 | 1.313 | 1.449 | 1.624 | 1.483 | 1.463 | 1.084 | 1.766 | 1.317 | 1.443 |

2.3 随机森林模型参数的调整

2.3.1 随机森林测试集占比选择

本文通过对比实验来获得随机森林模型最优测试集占比。为了排除其他参数的干扰,首先将各模型参数设置为默认值,再按照 0.1 的步长将测试集占比划分为 9 组,得到的 9 个测试集占比下模型准确率如表 3 所示。

表3 原始数据不同测试集占比与准确率对照表

Table 3 Comparison of the percentage of different test sets and accuracy rates for all data

| 测试集占比 | 测试集准确率/% | 测试集占比 | 测试集准确率/% | 测试集占比 | 测试集准确率/% |
|-------|----------|-------|----------|-------|----------|
| 0.1 | 96.30 | 0.4 | 96.18 | 0.7 | 95.52 |
| 0.2 | 96.33 | 0.5 | 95.97 | 0.8 | 95.05 |
| 0.3 | 96.22 | 0.6 | 95.77 | 0.9 | 94.53 |

由表3可知,准确率刚开始处于上升的趋势,在测试集占比为0.2时达到峰值,之后便一直下降。由于准确率越高,模型的学习精度也越高。所以本文选择的测试集占比大小为0.2,训练集占比选择0.8。

2.3.2 随机森林模型参数的优化

最大特征数、决策树数目和决策树深度是影响决策树生长的关键因素,也是随机森林模型精度的最基础保障。只有当模型准确率越高时,重要性排序才会越可靠。本文通过改变以上3个参数的大小来实现随机森林模型的参数优化。袋外误差通常是参数调整的第一选择^[26]。这3个参数的最优解均可以通过袋外误差的方法来获得。但由于本文需要分析每类构件的重要性排序,所以将最大特征数设置为None,其余的参数均设置为默认值。在设置最大特征数为None的前提下,检索袋外误差曲线上的稳定值就可以找到最优决策树数目。首先将决策树的数目定义到[1,1000]区间内,然后分别计算当最大决策树深度为1、10、15、20、25、30时随机森林模型的袋外误差值,并绘制了袋外误差曲线,如图3所示。由图可知,袋外误差大小随着决策树数目和决策树深度的增加首先呈现出急剧下降的趋势,当决策树数目约等于100时下降速度明显放缓,在400后曲线逐渐趋于平稳;当决策树深度为30时,袋外误差率曲线在最底端,此时的袋外误差值是最低的,大小为0.0366。故综合考虑模型误差、模型效率和计算成本等影响因素,将随机森林模型中的决策树深度取为30,决策树数目大小定为400。

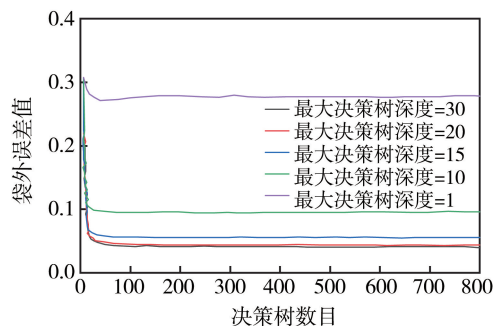


图3 各参数遍历后袋外误差率曲线

Fig. 3 Out-of-bag error rate curves after traversal of each parameter

3 结果与分析

3.1 损失构件重要性排序结果

基于参数优化后的随机森林模型,再分别用FI、PI和shap_values这3种方法对12种损失构件进行重要性排序,得到的结果如图4(a)~(c)所示。由图可知,墙面装修和板式基础在3种方法中的排序结果均位列第一和第二位,且墙面装修的重要性明显大于其他类型的构件;屋顶框架的重要性处于最后一位;其余的构件重要性排序分别为桩、墙面覆层、环形基础、天花板、其他非结构、屋内财物、墙体框架、地板和屋顶覆盖物,3种方法得到的结果略有差异。

随机森林中的3种重要性分析方法之间存在量级的差异,直接综合比较不具有足够的说服力。所以为了能够得出更加客观的结论,本文接下来会分别采用赋值排序相加法和归一化相加法来综合3种重要性分析方法的排序结果。赋值排序相加法是将FI、PI和shap_values这3种方法得到的重要性排序结果分别进行赋值。赋值方法是将 k 个变量通过重要性排序的高低依次赋予 $k, k-1, k-2, \dots, 2, 1$,并将三者所产生的新赋值结果进行相加后再进行重新排序^[27],从而确定不同损失构件的重要性大小。由于本文中一共有12种损失构件,故令赋值排序相加法的 k 值为12,最终得到的赋值排序相加法重要性排序结果以数字的形式呈现在表4中。归一化相加法将3种方法得到的结果分别归一化到0~1之间如图5所示。归一化方法可以用式(3)表示:

$$X' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (3)$$

式中: x 为某一损失构件在3种排序方法中的原始结果; x_{\min} 和 x_{\max} 分别对应3种排序结果之中的最小值和最大值; X' 为3种方法各自的归一化处理结果。

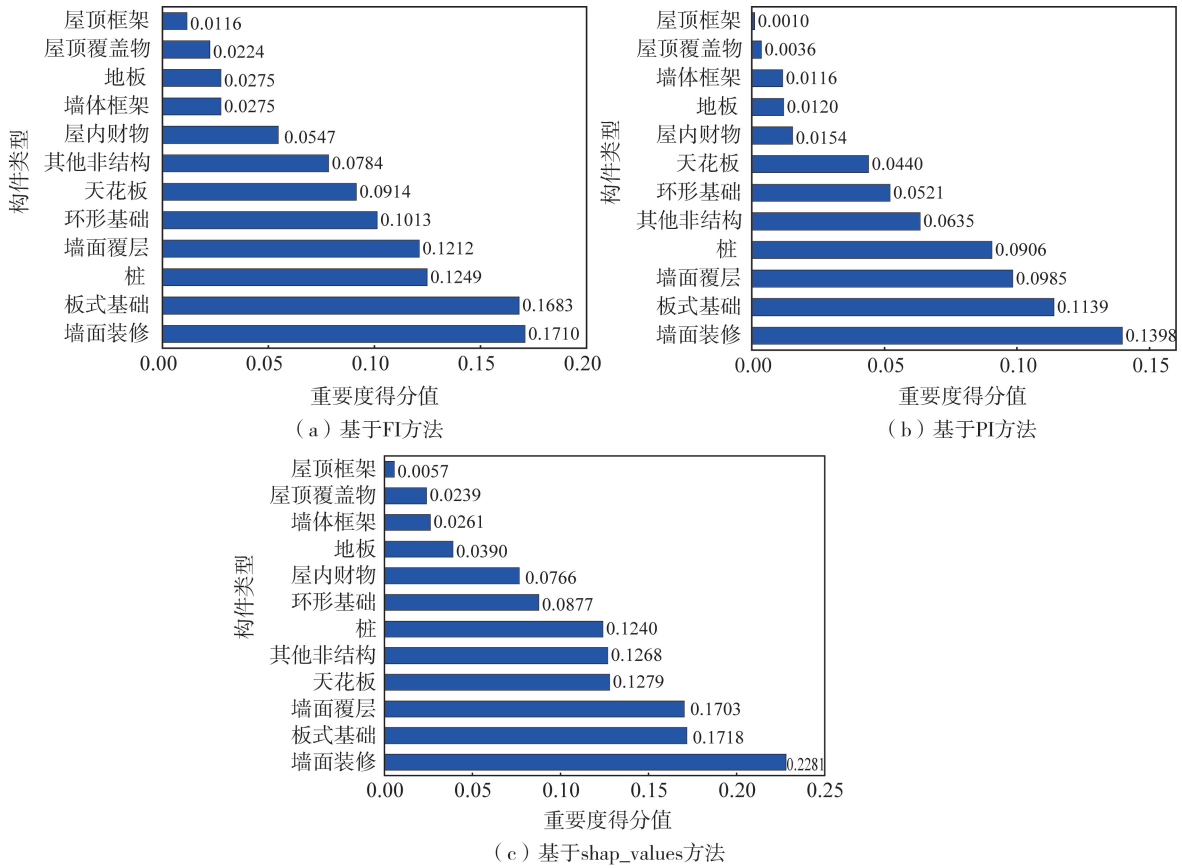


图 4 建筑物各损失构件重要性排序

Fig. 4 Ranking of the importance of the various loss components of the building

表 4 赋值排序相加法的重要性排序列表

Table 4 List of importance rankings of the sorted assignment summation method

| 构件类别 | FI 排名 | PI 排名 | shap_values 排名 | 赋值等权重和 | 综合排名 | 构件类别 | FI 排名 | PI 排名 | shap_values 排名 | 赋值等权重和 | 综合排名 |
|-------|-------|-------|----------------|--------|------|-------|-------|-------|----------------|--------|------|
| 墙面装修 | 1 | 1 | 1 | 36 | 1 | 环形基础 | 5 | 6 | 7 | 21 | 7 |
| 板式基础 | 2 | 2 | 2 | 33 | 2 | 屋内财物 | 8 | 8 | 7 | 16 | 8 |
| 墙面覆层 | 4 | 3 | 3 | 29 | 3 | 地板 | 10 | 9 | 8 | 12 | 9 |
| 桩 | 3 | 4 | 6 | 26 | 4 | 墙体框架 | 9 | 10 | 10 | 10 | 10 |
| 天花板 | 6 | 7 | 4 | 22 | 5 | 屋顶覆盖物 | 11 | 11 | 11 | 6 | 11 |
| 其他非结构 | 7 | 5 | 5 | 22 | 5 | 屋顶框架 | 12 | 12 | 12 | 3 | 12 |

由表 4 和图 5 可知,当损失构件的重要性存在明显的差异时,赋值排序相加法和归一化相加法均可以很好地反映出建筑物各损失构件的重要程度。在这 2 种相加法中,重要性排序的前 4 位均是墙面装修(墙布和涂料等)(赋值等权重和为 36)、板式基础(赋值等权重和为 33)、墙面覆层(蜂窝板等)(赋值等权重和为 29)、桩(赋值等权重和为 26),但归一化相加法不仅能够同步的体现重要性排序大小,还可以通过数据来展现各损失构件重要性排序的差异值,其中墙面装修重要度得分为 2.36,板式基础重要度得分为 1.98,墙面覆层重要度得分为 1.71,桩重要度得分为 1.49;在重要性大小差异不太显著时,由于在赋值排序相加法中天花板和其他非结构的排序均是 5,那么它们的重要性就很难被准确区分。而归一化相加法则可以通过比较数值的大小对 2 种损失构件的细微差距进行区分(其他非结构重要度得分为 1.17,天花板重要度得分为 1.15)。

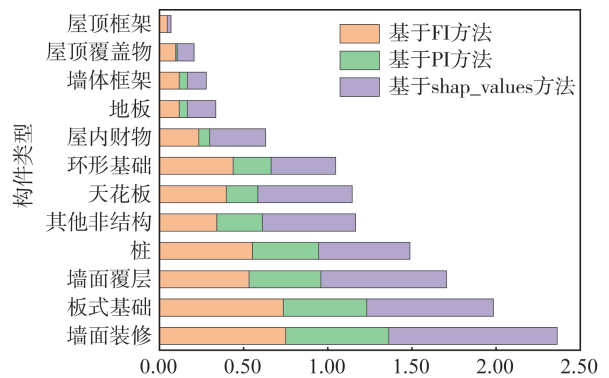


图 5 Importance ranking results of normalized summation

Fig. 5 Importance ranking results of normalized summation

可以通过比较数值的大小对 2 种损失构件的细微差距进行区分(其他非结构重要度得分为 1.17,天花板重要度得分为 1.15)。

3.2 损失构件重要性排序分析

通过赋值排序相加法和归一化相加法的重要性排序分析,研究结果表明:

1) 在本文选取的 12 种建筑物地震损失构件的重要性大小排序中,墙面装修的重要性远远领先于其他损失构件,在 3 种随机森林重要性分析结果中都位列第一。在归一化相加法中,其重要性约是排在第四位的桩的 1.5 倍。墙面不仅是家居家装中占地面积最大的部分,而且据调查发现,墙面装修在地震发生时极易发生大面积脱落,在一些中等地震下,当主体结构保持完好或轻微破坏时,墙面发生的破坏尤为严重^[28],且由于在以往的研究中墙面装修的重要性最容易被忽略,所以造成了大地震损失的局面,故墙面装修的重要性在第一位是符合实际情况的。因此对于建筑物地震损失而言,墙面装修应当得到比其他损失构件更高的关注度。

2) 板式基础的重要性仅次于墙面装修,排在第二,是结构构件中重要性最高的。在本文的地震事件中,存在地震液化影响规模最大的 2011 年基督城地震。此次地震导致了基督城地区大量建筑的倒塌或震陷^[29],一定程度地加重了建筑物的损失。由本文所选取的 124 427 栋建筑物的液化情况可知,存在液化现象的建筑物约占了其中的 2/5。同时我们将数据集按照无液化和液化 2 种状态进行划分后带入到随机森林模型中,得到了如图 6 所示的重要性排序图。由图可知,在无液化的状态下,结构构件的重要性排序相较于上文所得到的结果有了一定程度的下降;在仅存在液化的情况下,板式基础的重要性是第一位的,桩重要性排序也超过墙面覆层成为了第三。刘渊等^[12]通过大量的新西兰灾害数据分析,证实了该地区液化对于建筑物结构构件的损失加重,尤其对板式基础和桩的影响更大。根据上述结论,可以说明板式基础在整体重要性大小排序中排在第二、桩处于第四位具有一定的合理性。

3) 结构构件中,屋顶框架的重要性处于最后一位。在地震荷载的作用下,楼层越高,侧向力越小,层间位移角也越小^[30]。屋顶框架受到的地震作用越小,破坏程度越轻,对地震损失的影响自然也就越小。而非结构构件中,屋顶覆盖物的重要性最低,但其重要性仍高于屋顶框架,这也从侧面强调了开展对非结构构件的研究是很有必要的。

4) 在归一化相加法中,对结构构件和非结构构件的重要性数值分别进行相加,得到的结果为结构构件的总体重要性大小约为 4.8,而非结构构件的总体重要性值为 7.5。非结构构件的整体重要度约为结构构件重要度的 1.5 倍,所以减少非结构构件的损失是减少建筑物整体损失的关键点。

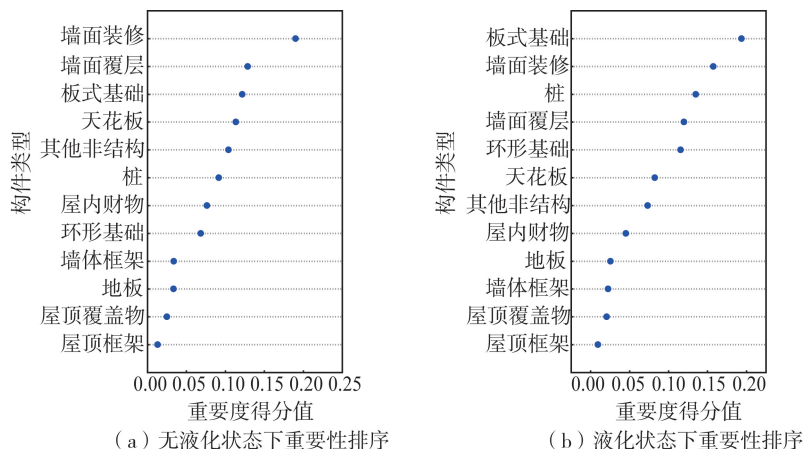


图 6 液化与非液化状态下的损失构件重要性排序

Fig. 6 Importance ranking of loss components in liquefied and non-liquefied sites

进一步通过分别删除本文所用的损失构件类型来计算新的随机森林的袋外误差,可以非常直观地观察到缺乏某一类型构件之后的袋外误差准确率的变化,其结果如表 5 所示,其中原始数据集的袋外误差准确率为 0.963 4。

表 5 分别删除各构件后袋外误差准确值表

Table 5 Accurate values of out-of-bag errors values after removing each component separately

| 删除构件 | OBB 准确率 | 删除构件 | OBB 准确率 | 删除构件 | OBB 准确率 |
|-------|---------|-------|---------|------|---------|
| 墙面装修 | 0.9184 | 墙面覆层 | 0.9241 | 天花板 | 0.9511 |
| 地板 | 0.9576 | 屋顶覆盖物 | 0.9623 | 屋内财物 | 0.9601 |
| 其他非结构 | 0.9383 | 桩 | 0.9478 | 板式基础 | 0.9160 |
| 环形基础 | 0.9590 | 墙体框架 | 0.9606 | 屋顶框架 | 0.9630 |

由表5可知,原始的袋外误差准确率与删除了屋顶框架与屋顶覆盖物之后的并无太大差异,再一次说明了这2种构件对建筑物地震损失的影响较小;而删除板式基础和墙面装修差异相当,删除环形基础和屋内财物 OBB 准确率相差也不大。为了进一步说明环形基础和屋内财物对于建筑物损失的重要性大小,将二者同时删除后进行袋外误差计算,得到的准确率为 0.956 5。分别删除环形基础和屋内财物以及同时删除两者后,其准确率分别下降了 0.004 4、0.003 3、0.006 9,准确率下降得越多,则说明该构件的重要性越高。由此可知两者对于建筑物地震损失的贡献大小相当且处于中等水平。

4 结论和讨论

本文采用随机森林模型对整理好的新西兰地震损失数据库中的 12 种建筑物构件进行了重要性排序。通过袋外误差方法来选择决策树数目和决策树最大深度的参数后,再将随机森林 3 种不同的方法得到的重要性排序结果进行排序赋值法和归一化相加法处理,得出的结论如下:

1) 随机森林算法可以用于研究建筑物地震损失构件的重要性分析。FI、PI 和 shap_values 这 3 种方法得到的重要性排序结果差异不大。归一化相加法不仅可以定量地表达不同构件重要性的差异,而且所得到的结果与当前研究现状较为吻合,具有一定的合理性。

2) 无论是从单独的重要性分析方法还是从 3 种重要性分析方法结果的综合比较结果来看,墙面装修、墙面覆层的损失和填充墙等墙面的大面积破坏有着直接联系,因此这两者的重要性排序比较靠前;板式基础的重要性大小次于墙面装修,这也从侧面印证了新西兰地区存在的液化现象会加重结构构件损失的说法;环形基础和屋内财物的重要性是相当的,均处于中等水平;屋顶框架在地震中受到的破坏是最轻的,所以它的重要性在整体排序中最低。

3) 本研究首次将结构构件和非结构构件放在一起进行整体构件的重要性排序。结果证明,在地震灾害中,非结构构件破坏造成了建筑物的主要损失。从建筑物地震损失构件层面来分析,非结构构件的整体重要度约为结构构件的 1.5 倍。因此加强对非结构构件的研究,提高非结构构件特别是墙体装饰等的抗震能力有益于未来的防震减灾工作。

参考文献:

- [1] KIRCHER C A, REITHERMAN R K, WHITMAN R V, et al. Estimation of earthquake losses to buildings[J]. *Earthquake Spectra*, 1997, 13(4): 703-720.
- [2] 于敬海,丁永君,李久鹏,等. 设置耗能外挂墙板结构的抗震性能[J]. *天津大学学报(自然科学与工程技术版)*, 2015, 48(B07): 122-126. YU Jinghai, DING Yongjun, LI Jiupeng, et al. Seismic behavior of structure with energy dissipation external wall panel[J]. *Journal of Tianjin University (Science and Technology)*, 2015, 48(B07): 122-126. (in Chinese)
- [3] LIN Y J, ZHANG X Z, XU W Q, et al. Importance assessment of structural members based on elastic-plastic strain energy[J]. *Advances in Materials Science and Engineering*, 2019, 2019: 1-17.
- [4] CARRADINE D, KUMAR A, FAIRCLOUGH R, et al. Serviceability fragility functions for New Zealand residential windows[J]. *Bulletin of the New Zealand Society for Earthquake Engineering*, 2020, 53(3): 137-143.
- [5] WANG S J, YANG Y H, LIN F R, et al. Experimental study on seismic performance of mechanical/electrical equipment with vibration isolation systems[J]. *Journal of Earthquake Engineering*, 2017, 21(3): 439-460.
- [6] 李威齐,曲哲,解全才,等. 我国公共建筑中吊顶的震害特征及其易损性分析[J]. *工程力学*, 2019, 36(7): 207-215. LI Qiqi, QU Zhe, XIE Quancai, et al. Seismic damage characteristics and fragility of suspended ceilings in Chinese public buildings[J]. *Engineering Mechanics*, 2019, 36(7): 207-215. (in Chinese)
- [7] 尚庆学,李吉超,王涛. 管线系统抗震性能研究方法综述[J]. *世界地震工程*, 2018, 34(4): 174-181. SHANG Qingxue, LI Jichao, WANG Tao. State of the art of study on seismic performance of piping system[J]. *World Earthquake Engineering*, 2018, 34(4): 174-181. (in Chinese)
- [8] 韩庆华,张晓龙,芦燕,等. 基于增量动力法的电梯对重系统地震易损性分析[J]. *地震工程与工程振动*, 2022, 42(5): 53-62. HAN Qinghua, ZHANG Xiaolong, LU Yan, et al. Seismic fragility analysis of elevator counterweight system based on incremental dynamic analysis[J]. *Earthquake Engineering and Engineering Dynamics*, 2022, 42(5): 53-62. (in Chinese)
- [9] GHARAIBEH E S, FRANGOPOL D M, ONOUFRIOU T. Reliability-based importance assessment of structural members with applications to complex structures[J]. *Computers & Structures*, 2002, 80(12): 1113-1131.
- [10] PERRONE D, CALVI P M, NASCIMBENE R, et al. Seismic performance of non-structural elements during the 2016 Central Italy earthquake[J]. *Bulletin of Earthquake Engineering*, 2019, 17(10): 5655-5677.

- [11] GAUTAM D, ADHIKARI R, RUPAKHETY R. Seismic fragility of structural and non-structural elements of Nepali RC buildings[J]. *Engineering Structures*, 2021, 232: 111879.
- [12] 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)
- [13] WANG Z F. A preliminary report on the Great Wenchuan Earthquake[J]. *Earthquake Engineering and Engineering Vibration*, 2008, 7(2): 225-234.
- [14] 王自法, Sangioon Park, Selina Lee, 等. 提高地震灾害损失估计精度的几点研究[J]. *地震工程与工程振动*, 2014, 34(4): 110-114.
WANG Zifa, PARK S, LEE S, et al. Quantification improvement of earthquake loss estimation[J]. *Earthquake Engineering and Engineering Dynamics*, 2014, 34(4): 110-114. (in Chinese)
- [15] GB 50011—2010 建筑抗震设计规范[S]. 北京: 中国建筑工业出版社, 2010.
GB 50011—2010 Code for seismic design of buildings[S]. Beijing: China Architecture & Building Press, 2010. (in Chinese)
- [16] DHAKAL R P, POURALI A, TASLIGEDIK A S, et al. Seismic performance of non-structural components and contents in buildings: An overview of NZ research[J]. *Earthquake Engineering and Engineering Vibration*, 2016, 15(1): 1-17.
- [17] 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)
- [18] COUNCIL A T. Earthquake damage evaluation data for California[M]. Redwood: Applied Technology Council, 1985.
- [19] GB/T 18208.4—2011 地震现场工作 第4部分: 灾害直接损失评估[S]. 北京: 中国标准出版社, 2011.
GB/T 18208.4—2011 Post-earthquake field works—Part 4: Assessment of direct loss [S]. Beijing: Standards Press of China, 2011. (in Chinese)
- [20] 陈妙金, 汪小钦, 吴思颖. 基于随机森林算法的水土流失影响因子重要性分析[J]. *自然灾害学报*, 2019, 28(4): 209-219.
CHEN Miaojin, WANG Xiaochin, WU Siying. Importance analysis of soil erosion influencing factors based on random forest[J]. *Journal of Natural Disasters*, 2019, 28(4): 209-219. (in Chinese)
- [21] NICODEMUS K K. Letter to the Editor: On the stability and ranking of predictors from random forest variable importance measures[J]. *Briefings in Bioinformatics*, 2011, (4): 369-373.
- [22] BIAU G, SCORNET E. A random forest guided tour[J]. *Test: An Official Journal of the Spanish Society of Statistics and Operations Research*, 2016, 25(2): 197-227.
- [23] 范晓, 邹文, 刘璞, 等. 随机森林储层预测及关键参数探讨——以 SC 某工区储层预测为例[J]. *地球科学前沿(汉斯)*, 2021(5): 701-709.
FAN Xiao, ZOU Wen, LIU Pu, et al. Random forest reservoir prediction and key parameters discussion—Taking reservoir prediction of a work area in SC as an example[J]. *Advances in Geosciences (Hans)*, 2021(5): 701-709. (in Chinese)
- [24] BREIMAN L. Out-of-bag estimation[R]. Berkeley: Statistics Department, University of California, 1996.
- [25] 郑迎凯, 陈建国, 王成彬, 等. 确定性系数与随机森林模型在云南芒市滑坡易发性评价中的应用[J]. *地质科技通报*, 2020, 39(6): 131-144.
ZHENG Yingkai, CHEN Jianguo, WANG Chengbin, et al. Application of certainty factor and random forests model in landslide susceptibility evaluation in Mangshi City, Yunnan Province[J]. *Bulletin of Geological Science and Technology*, 2020, 39(6): 131-144. (in Chinese)
- [26] JANITZA S, HORNING R. On the overestimation of random forest's out-of-bag error[J]. *PLoS One*, 2018, 13(8): e0201904.
- [27] HAN H, GUO X L, YU H. Variable selection using mean decrease accuracy and mean decrease gini based on random forest[C]//2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS). Beijing: Piscataway, NJ, USA: IEEE, 2017: 219-224.
- [28] 刘小娟, 蒋欢军. 非结构构件基于性能的抗震研究进展[J]. *地震工程与工程振动*, 2013, 33(6): 53-62.
LIU Xiaojuan, JIANG Huanjun. State-of-the-art of performance-based seismic research on nonstructural components[J]. *Earthquake Engineering and Engineering Dynamics*, 2013, 33(6): 53-62. (in Chinese)
- [29] 陈龙伟, 袁晓铭, 孙锐. 2011 年新西兰 M_w 6.3 地震液化及岩土震害述评[J]. *世界地震工程*, 2013, 29(3): 1-9.
CHEN Longwei, YUAN Xiaoming, SUN Rui. Review of liquefaction phenomena and geotechnical damage in the 2011 New Zealand M_w 6.3 earthquake[J]. *World Earthquake Engineering*, 2013, 29(3): 1-9. (in Chinese)
- [30] 郭仕群, 吴传文, 褚云朋. 高层建筑结构设计[M]. 重庆: 重庆大学出版社, 2022.
GUO Shiqun, WU Chuanwen, CHU Yunpeng. Structural design of high-rise buildings[M]. Chongqing: Chongqing University Press, 2022. (in Chinese)