

<https://doi.org/10.3799/dqkx.2022.481>



基于机器学习的砖砌体房屋震害快速预测方法

刘 丽^{1,2}, 沈俊凯^{1,2}, 张令心^{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: 由于现有的震害预测方法不能对砖砌体结构做出高效的预测, 基于机器学习模型, 提出了一种综合考虑地震动特性与结构特性的砖砌体结构震害快速预测方法. 该方法利用机器学习模型, 从时域、频域、反应谱和持时 4 个方面初步选取了能够代表地震动特性的 12 个参数, 从承载力、刚度等方面初步选取了与砖砌体结构破坏相关性较强的 7 个结构参数; 将地震动参数与结构参数相结合作为输入变量, 分别给出了基于支持向量机(SVM)、随机森林(RF)、人工神经网络(ANN)三种机器学习模型的砖砌体结构的震害快速预测方法, 并进行了性能比较; 采用相关性分析对输入参数进行进一步优化, 给出了优化输入参数后的最优预测模型. 结果表明, 当采用 19 个输入参数时, ANN 模型的预测准确率最高, 达到 91.56%. 当采用优化后的 12 个参数作为输入时, 基于 RF 模型的预测性能更加稳定, 预测的准确率也更高, 可达到 90.01%. 优化输入参数后的基于 RF 模型的预测方法可以实现对砖砌体结构震害的快速预测; 与只考虑结构参数或只考虑地震动参数作为输入的方法相比, 同时考虑结构和地震动参数作为输入的方法极大地提高了预测的准确性.

关键词: 机器学习; 震害快速预测; 砖砌体结构; 地震动特性; 工程地质.

中图分类号: P694

文章编号: 1000-2383(2023)05-1769-11

收稿日期: 2022-09-23

A Machine Learning-Based Method for Rapid Prediction of Earthquake Damage in Brick Masonry Houses

Liu Li^{1,2}, Shen Junkai^{1,2}, Zhang Lingxin^{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: Since the existing earthquake damage prediction methods cannot make rapid predictions for brick masonry structures. A rapid prediction method for earthquake damage of brick masonry structures is proposed. The method uses a machine learning model, considering the ground motion characteristics and structural characteristics. 12 ground motion parameters that represent the ground motion characteristics and 7 structural parameters that have a strong correlation with the damage of brick masonry structures are selected. The ground motion parameters are considered in four aspects: time domain, frequency domain, response spectrum and holding time, and the structural parameters are considered in terms of bearing capacity and stiffness. Three machine learning models based on support vector machine, random forest and artificial neural network are given for fast prediction of

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重点项目(No.2019A01); 国家自然科学基金项目(No.U2139209); 黑龙江省头雁行动计划项目.

作者简介: 刘丽(1998—), 女, 硕士研究生, 主要从事砌体结构震害研究. ORCID:0000-0002-6533-0268. E-mail: Kelly_ll1998@163.com

* **通讯作者:** 张令心, E-mail: lingxin_zh@126.com

引用格式: 刘丽, 沈俊凯, 张令心, 2023. 基于机器学习的砖砌体房屋震害快速预测方法. 地球科学, 48(5):1769-1779.

Citation: Liu Li, Shen Junkai, Zhang Lingxin, 2023. A Machine Learning-Based Method for Rapid Prediction of Earthquake Damage in Brick Masonry Houses. *Earth Science*, 48(5):1769-1779.

seismic damage of brick masonry structures. The input parameters were further optimized using correlation analysis, and the optimal model after optimizing the input parameters was given. The results show that the ANN model has the highest prediction accuracy of 91.56% when 19 input parameters were used. The prediction accuracy of the RF model-based earthquake damage prediction method was higher when 12 optimized parameters were used as inputs, reaching 90.01%. The prediction performance of the RF-based model was more stable when the input parameters were gradually reduced. The optimized input parameters of the RF model-based prediction method can achieve rapid prediction of seismic damage to brick masonry structures. The method that considers both structural and ground vibration parameters as input greatly improves the accuracy of prediction compared to the method that considers only structural parameters or only ground vibration parameters as input.

Key words: machine learning; rapid prediction of earthquake damage; brick masonry structure; ground motion characteristics; engineering geology.

0 引言

地震造成的人员伤亡和经济损失主要是由建筑物发生严重破坏或毁坏引起的。一方面,在震前对建筑物进行震害预测是防震减灾基础工作的重要组成部分,通过判断建筑物在遭受不同地震动可能发生的破坏状态、分析导致破坏的薄弱环节并采取防震减灾措施可以减轻地震造成的人员伤亡和经济损失,这对城乡建筑群的防灾减灾工作和韧性城乡的建设具有重要意义。另一方面,在震后对建筑物的破坏情况做出快速、准确的评估能够为应急救援提供依据。传统的建筑物震害预测的方法包括经验统计法(张桂欣和孙柏涛,2010)、理论计算法等。经验统计法需要耗费大量的人力统计震害历史数据;理论计算法则需要在震后建立相应结构的有限元模型并进行分析,流程复杂且耗时。这些方法对一个区域内的大量房屋进行震害预测时,前期收集各类图纸资料需要耗费大量的时间、人力和物力,对各个房屋进行逐个建模分析更是难以实现。因此,有必要研究一种仅需准备少量且容易获取数据依然可以对其震害进行准确、快速的预测方法,以在保证精度的情况下提高预测效率的目标。

近年来,随着人工智能技术的发展,机器学习方法已逐步被引入各类灾害风险评估(Shaheen and Hakan, 2022; 吴波等, 2022; 张志强等, 2023)、建筑结构的震害预测当中。使用机器学习对各类建筑物进行震害预测时,输入向量的选取至关重要。目前,大部分研究者仅单独选取结构参数或地震动参数作为输入变量开展相关研究。Harirchian *et al.* (2020)将层数、建筑面积等 8 个结构参数作为输入变量,采用支持向量机(Support Vector Machine, SVM)对韩国等 4 个不同国家钢筋混凝土(Rein-

forced Concrete, RC)建筑物的损伤状态进行预测,评估准确率为 68%。该研究仅考虑了结构参数对结构震害的影响,没有考虑地震动对建筑结构破坏状态的影响。然而,有专家研究表明,针对同一栋建筑,不同地震动造成的破坏差异很大(潘志宏和洪博, 2014)。Morfidis and Kostinakis(2017)选取 14 个地震动参数作为输入,使用多层前馈感知器神经网络对地震作用下 RC 建筑物的损伤进行预测。Xu *et al.* (2020)选用了 48 种地震动参数,采用 SVM、决策树(Decision Tree, DT)与逻辑回归 3 种机器学习算法对框架结构及砌体结构的建筑物进行破坏状态预测,评估区域内建筑物的平均预测精度达到 81.12%。上述这类方法充分考虑了地震动对结构破坏的影响,但却未考虑结构本身的影响。另一部分研究者将结构参数和地震动参数相结合作为输入变量,例如 Mangalathu *et al.* (2020)采用 DT 和随机森林(Random Forest, RF)等 4 种不同的机器学习算法对建筑物破坏状态进行分类,输入变量包括断层距离和剪切波速等 3 个地震动参数和结构层数,建筑年代等 5 个结构参数,但 4 种模型的分类准确率较低,准确率最高的 RF 模型仅有 66%。Lautour and Omenzetter (2009)和 Morfidis and Kostinakis (2018)采用人工神经网络(Artificial Neural Network, ANN)预测 RC 框架结构在地震中的破坏情况时,输入变量选取了结构参数与地震动参数并取得了良好的预测效果。但目前更多的研究是仅采用一项地震动参数,其他均为结构参数作为输入变量,例如,张令心等(2019)将层数、柱面积率等 11 个结构参数与地震烈度参数作为输入变量,采用 ANN 对 RC 框架结构进行震害预测。刘本玉等(2002)采用 ANN 对多层砖房进行震害预测,根据与多层砖房震害程度有关的震害因子选取了房屋层数及地震烈度等 9 个震害影响因子。汤皓等(2006)针对传统的基于地震烈度的建筑物震害预

测方法的不足,将地震动峰值加速度(PGA)作为建筑物震害预测的地震动指标,将PGA和砖砌体的结构参数作为输入变量,建立了基于BP神经网络的震害预测模型.虽然考虑了地震动对结构破坏的影响,但所采用的地震动参数较少,无法有效、全面地描述地震动对结构破坏状态的影响程度.

另一方面,现有的基于机器学习算法对建筑物进行震害预测的研究大部分都是针对RC结构开展的,针对砌体结构开展的研究相对较少.而砌体结构是我国城乡广泛采用的一种结构形式,也是历次地震中震害较重的结构.在汶川大地震中,砌体结构是损坏数量最多、震害最为严重的一类结构.由此可见,对砌体结构震害进行快速准确的预测极为重要.本文以砖砌体结构为研究对象,以机器学习模型为核心,提出一种综合考虑结构特性和地震动特性的砖砌体结构震害快速预测方法.该方法初步选取了与地震动频谱、持时、峰值特性相关的12个地震动参数和与结构震害相关性较强的7个结构参数,共19个参数作为输入向量,分别建立了基于SVM、RF和ANN三种模型的砖砌体结构震害预测快速方法.以精确率、召回率和F1分数为评价指标对3个模型进行评估.然后通过相关性分析对相关性较强的参数进行优化,研究了输入参数的数量对3种模型预测精度的影响,最终给出经输入参数优化后的最优砖砌体震害快速预测方法.

1 研究方法

目前,基于机器学习的结构震害快速预测方法主要包括数据集的建立、模型的建立与评估和模型的验证3个部分.实现的具体过程如图1所示.

本文方法最主要的特点是选取了结构和地震动两类参数作为输入.首先初步选取12个地震动参数与7个结构参数,其中,地震动参数主要考虑了地震动三要素以及参考文献(Morfidis and Kostinakis, 2017; Xu *et al.*, 2020)的建议,选取了12个地震动参数,具体参数及计算方法如表1所示.这些地震动参数大致可分为时域参数、反应谱参数、频域参数及持时参数4类,较全面地考虑了地震动主要特性对结构震害的影响.影响砌体结构抗震能力的因素有多种,在选取影响结构抗震性能的参数时,既要考虑影响砖砌体结构震害相关的刚度、强度,也要考虑这些参数是否易于获取.综合考虑以上2种因素,并参考相关的文献(刘本玉等,2002)以及震害

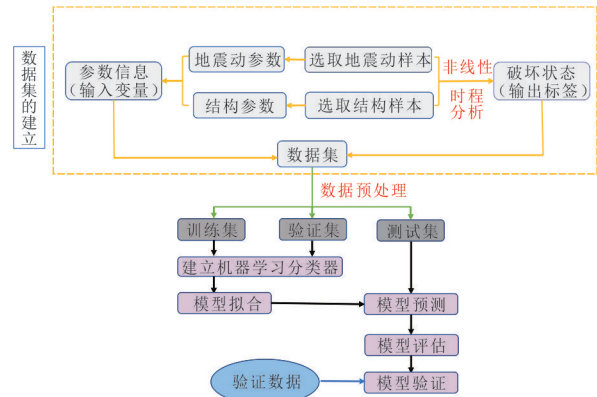


图1 基于机器学习的砖砌体结构震害快速预测流程
Fig. 1 Flow chart of rapid earthquake damage prediction of brick masonry structures based on machine learning

统计资料,本文初步选取了以下7个结构参数:(1)建筑物层数:历次震害资料表明,结构的层数对结构的震害影响很大;(2)建筑物总高:一般情况下,随着结构总高的增加建筑的震害会更加严重;(3)底层砂浆强度:多层砖砌体的震害现象主要表现为砖墙出现裂缝,裂缝的程度取决于砂浆强度;(4)砌体强度:砖砌体强度越高,抗压强度亦高,从而提高结构的抗震性能;(5)砖墙面积率:该参数决定了砌体房屋的刚度和强度等;(6)设防烈度:设防烈度对结构的抗震性能影响极大;(7)地基基础:采用不同地基基础形式的砖砌体抗震性能差异较大.

由于本次选取的结构样本均为20世纪90年代建造且均按照《建筑抗震设计规范》GBJ11-89设计,因此,本文没有将建筑年代列为输入参数.采用机器学习进行训练和测试时,除地基基础参数外,其他参数按实际取值.地基基础为条形基础时取1,砖基础取2.

在现有的机器学习模型中,SVM、RF、ANN是3种采用不同理念的代表性模型.这3种模型在分类任务中应用非常广泛,分类表现也比较突出.SVM的基本思想是在多维空间中构造超平面将对象进行分类(Gunn, 1998).SVM实现简单且具有良好的鲁棒性,在高维数据和低维数据上均表现良好.RF结合了装袋(Breiman, 1996)和随机特征选择(Ho, 1998)的概念,是对决策树模型的一种改进,可以提高预测效果且简化了超参数调整的步骤.ANN凭借可累加的隐含层和非线性神经元激活方法在许多实际应用领域中得到了广泛的应用.因此,本文基于SVM、RF和ANN建立了砖砌体结构震害快速预测方法,并对比不同机器学习模型在砖砌体震害快

表 1 地震动参数及对应的计算方法
Table 1 Ground motion parameters and the corresponding calculation methods

参数名称	简称	计算方法	分类
地面峰值加速度	PGA	$\max a(t) $	时域参数
地面峰值速度	PGV	$\max v(t) $	
地面峰值位移	PGD	$\max d(t) $	
Arias 强度	Ia	$\frac{\pi}{2g} \int_{t_1}^{t_2} [a(t)]^2 dt$	
累计绝对加速度	CAA	$\int_{t_1}^{t_2} a(t) dt$	
有效峰值加速度	EPA	$\frac{1}{2.5} \{ \bar{S}_a(\xi=0.05, T) \}_{0.1}^{0.5}$	反应谱参数
加速度谱强度	ASI	$\int_{0.1}^{0.5} S_a(\xi=0.05, T) dT$	
Housner 强度	HI	$\int_{0.1}^{2.5} PSV(\xi=0.05, T) dT$	
括号持时	TBD	第一次和最后一次超过 0.025 g 的时间	持时参数
显著持时/能量持时	TSD	$5\% \leq \int_0^t a(t)^2 dt / \int_0^T a(t)^2 dt \leq 95\%$	
V_{\max}/A_{\max}	PGV/PGA	$\max v(t) / \max a(t) $	频域参数
卓越周期	PP	$T[\max S_a(\xi=0.05, T)]$	

注: $a(t)$ 、 $v(t)$ 、 $d(t)$ 分别表示加速度、速度和位移时程; S_a 表示加速度反应谱; ξ 表示阻尼比; t_1 、 t_2 分别对应于 Arias 强度的 5% 和 95%。

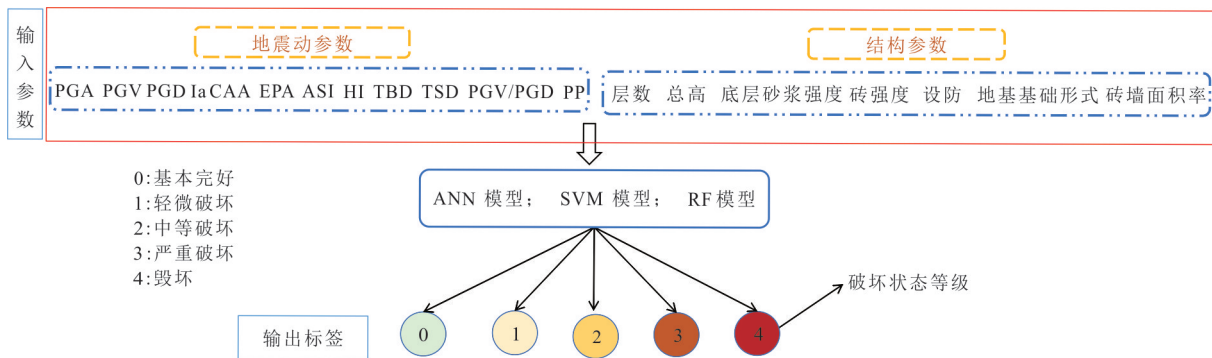


图 2 输入参数与输出标签

Fig.2 Input parameters and output labels

速预测中的性能。

2 数据集的建立

数据集的建立过程如图 1 数据库建立部分所示,主要是准备输入参数与对应的输出标签。首先选取砖砌体结构样本并提取第 1 节所提到的 7 个结构参数,同时选取地震动样本并按照表 1 计算相应的 12 个地震动参数,将得到的结构参数与地震动参数作为训练机器学习模型的输入参数,如图 2 所示。然后,使用所选取的地震动样本对结构样本进行非线性时程分析,将得到的破坏状态进行编码,最终得到用于训练机器学习模型

的输出标签,如图 2 所示。具体过程如下。

2.1 选取砖砌体结构样本

为了保证砌体结构样本的多样性,本文在选取砖砌体结构样本时,在考虑第 1 节的结构参数的基础上,尽可能多地扩大采样范围,最终选取了 55 栋使用用途为住宅的砖砌体结构样本。由于搜集的样本库的砌体结构多数为 6 度和 7 度设防,因此本文仅选取了设防烈度为 6、7 度的建筑物,其中 6 度设防有 29 栋,7 度设防有 26 栋,选取的建筑物中结构参数的信息如表 2 所示。

2.2 选取地震动样本

本文从太平洋地震工程研究中心(Pacific earthquake Engineering Research Center, PEER)的数据

表 2 所选砖砌体结构样本的信息

Table 2 The information of selected brick masonry structure samples

结构特性	取值范围
总高	6~16.4 m
砖强度	MU7.5、MU10
底层砂浆强度	M2.5、M5、M7.5、M10
砖墙面积率	8%~15%
层数	2~6层
地基基础	条形基础、砖基础
设防烈度	6~7度

库中选取了 50 条地震动记录,选取地震动记录的原则如下:(1)所选取的地震动记录所属的地震事件的里氏震级大于 6.5;(2)断层距小于 200 km(20 组小于 20 km,30 组大于 20 km)。

为了扩充数据集,本文对地震动的幅值进行了调整,对每条地震动按照调幅系数 2、3、4 进行调幅。若调幅后 PGA 大于 1.0 g,则不再对该地震动进行后续的调幅。使用调幅后的地震动对 55 栋结构按 2.3 节所述的非线性分析方法和破坏指标进行分析和计算得到对应的破坏状态,共得到 8 580 个样本。

2.3 砖砌体非线性时程分析计算方法及破坏指标的选取

本文采用由张令心等(2002, 2021)提出的一种基于自平衡力的弹塑性动力反应分析方法来完成对砌体结构的非线性动力反应分析。该方法的合理性和有效性在张令心等(2021)得到验证。本文砌体结构的破坏状态指标采用文献(张令心等,2021)中给出的延伸系数指标,破坏状态划分的具体方法见文献(张令心等,2021)。在使用机器学习进行训练和测试的过程中,采用从 0 到 4 的 5 个整数表示这 5 种破坏状态,其中,0 表示基本完好,1 表示轻微破坏,2 表示中等破坏,3 表示严重破坏,4 表示毁坏。

2.4 数据集预处理

将数据输入模型进行训练前,需要对数据进行归一化、样本平衡抽样等预处理。由于输入数据幅值变化过大会导致神经网络这类算法在训练过程中难以收敛,因此需要对输入数据进行归一化处理。具体方法是按参数类别对每一类参数进行归一化,将某一类输入参数归一化到 [0, 1] 区间内,具体公式如下:

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}}, \tag{1}$$

式中: X_{scaled} 为归一化后的数值; X 为具体的参

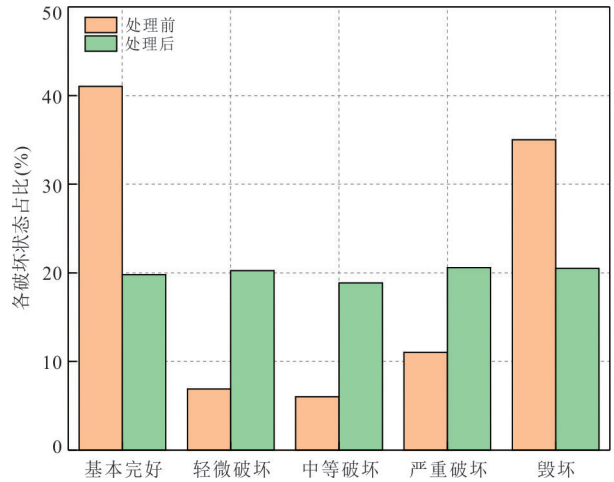


图 3 数据预处理前后各破坏状态所占比例

Fig.3 Proportion of each failure state before and after data preprocessing

数值; X_{min} 为该类参数中的最小值; X_{max} 为该类参数中的最大值。

所得到的数据集集中的 8 580 个样本中基本完好状态占比为 41.04%,而中等破坏的占比仅为 6.03%(如图 3 所示)。为了避免样本不均导致机器学习模型性能下降,本文采用 SMOTE 算法来实现各类损伤样本的抽样平衡问题。SMOTE 算法的基本思想为:对每个少数类样本 x_i ,从它的最临近样本中随机选择一个样本 \hat{x}_i (\hat{x}_i 也为少数类中的一个样本),然后在 x_i 和 \hat{x}_i 之间的连线上随机选择一点作为新合成的少数类样本。对于每一个随机选出的近邻 \hat{x}_i ,分别与 x_i 按照如下公式构建新样本:

$$x_{new} = x_i + (\hat{x}_i - x_i) \times \delta, \tag{2}$$

式中: $\delta \in [0, 1]$ 是一个随机数。

经过该方法抽样后的数据与原始数据的对比如图 3 所示。最后,将数据集随机打乱并按照 8:1:1 划分为训练集、验证集和测试集,其中训练集用于优化机器学习模型,而验证集则是用来观测超参数对训练的影响,测试集用来测试训练模型的性能。

3 模型建立及其验证

数据集建立后,本文基于 SVM、RF 和 ANN 模型分别建立基于地震动和结构两类参数的砖砌体结构震害快速预测方法。将初步确定的 19 个参数作为输入对砖砌体结构的破坏状态进行预测,以精确率、召回率和 F1 分数为评价指标对 3 个机器学习模型的性能进行了评价。然后,以各输入参数之间的相关系数为依据,通过逐步减少输入参数数量研究

表 3 最优机器学习模型的主要超参数数值

Table 3 Main hyperparameter values of the optimal machine learning model

机器学习算法	主要超参数的数值
SVM	$C=2, \text{gamma}=0.1, \text{kernel}=\text{rbf}$
RF	$n_estimators = 16$
ANN	$\text{hidden_layer_sizes}=(100,100,80,80,60,50,30,20); \text{activation}=\text{tanh}; \text{slover}=\text{adam}; \text{learning_rate}=\text{constant}; \text{max_iter}=200$

了输入参数数量对 3 种机器学习模型的影响, 最终筛选出一个最优方法。

3.1 超参数的选取

机器学习模型中的超参数无法训练, 往往只能通过观测模型在验证集上的表现进行手动调整。在 SVM 模型中, 影响 SVM 模型性能的超参数为正则化参数 (C) 和 gamma 。其中正则化参数 C 对模型性能的影响较大。因此, 本文对 C 从 0.1 变化到 100 进行训练, 而 gamma 取默认值 0.1。RF 模型中的树木的个数 ($n_estimators$) 对模型的过拟合现象影响较大, 因此, 本文研究了该参数从 1 变化到 100 对模型性能的影响。影响 ANN 的超参数较多且影响机制复杂, 包括隐含层数及神经元数量、激活函数等。对此, 本文进行了大量的训练, 在经过对比分析后, 本文使用 19 个参数进行预测时, 3 个模型所采用的最终超参数如表 3 所示。

3.2 模型评价指标

本文选取精确率 (Precision)、召回率 (Recall) 和 $F1$ 分数作为各个模型的评价指标。计算每个类别的精确率和召回率分别为公式 (3) 和 (4)。

$$P = \frac{TP}{TP + FP}, \quad (3)$$

$$R = \frac{TP}{TP + FN}, \quad (4)$$

式中, TP 表示正类样本被预测为正类; FP 表示负类样本被预测为正类; FN 表示正类样本被预测为负类。以计算基本完好预测结果的精确率和召回率为例, TP 代表基本完好标签的样本被预测为基本完好状态, FP 代表非基本完好标签的样本被预测为基本完好状态, FN 代表基本完好标签的样本被预测为其他破坏状态。最后以 5 个类别的精确率取平均值作为最后的精确率值, 5 个类别的召回率取平均作为最后的召回率值, 即:

$$p = \frac{1}{5} \sum_{i=1}^5 P_i, \quad (5)$$

$$r = \frac{1}{5} \sum_{i=1}^5 R_i, \quad (6)$$

式中, P_i 表示第 i 个类别的精确率; p 表示最后的精

表 4 基于 SVM 模型的精确率和召回率

Table 4 Precision and recall of SVM-based method

	基本完好	轻微破坏	中等破坏	严重破坏	毁坏	召回率 (%)
基本完好	292	17	6	2	0	92.11
轻微破坏	17	316	15	1	0	90.54
中等破坏	13	10	346	9	3	90.81
严重破坏	7	4	14	320	6	91.17
毁坏	0	2	4	19	319	92.70
精确率 (%)	88.75	90.54	89.87	91.17	97.26	91.47
						91.52

确率值; R_i 表示第 i 个类别的召回率; r 表示最后的召回率值。 $F1$ 分数计算公式为:

$$F1 = 2 \cdot \frac{p \cdot r}{p + r}. \quad (7)$$

3.3 预测结果

根据上面给出的评价指标, 使用初步确定的 19 个输入参数进行砖砌体结构破坏状态的预测, 得出 3 个模型在测试集上的预测结果, 并得到测试集中每个样本预测需要的时间。基于混淆矩阵计算得到各个模型的精确率和召回率如表 4~表 6 所示, $F1$ 分数和预测每个样本需要的时间如表 7 所示。

从表 4~表 6 可以看出, 基于 SVM 模型的精确率和召回率分别为 91.52% 和 91.47%; 基于 RF 模型的精确率和召回率分别为 90.63% 和 90.40%; 基于 ANN 模型的精确率和召回率分别为 91.73% 和 91.59%, ANN 模型的精确率和召回率略高于其他 2 个模型。从表 7 可以看出, 3 个模型的 $F1$ 分数均为 90% 左右。综合 3 项评价指标, 3 个模型均能达到良好的预测效果。从表 7 中 3 个模型对每个样本做出预测的时间可以看出, 其中 SVM 模型和 RF 模型可以达到每个样本预测时间为 0.2 ms (计算机的处理器为: Intel Core i7-12700), ANN 模型对测试集每个

表 5 基于 RF 模型的精确率和召回率

Table 5 Precision and recall of RF-based method

	基本完好	轻微破坏	中等破坏	严重破坏	毁坏	召回率 (%)
基本完好	299	18	4	1	0	92.86
轻微破坏	6	333	21	9	2	89.76
中等破坏	5	18	338	18	2	88.71
严重破坏	2	6	21	305	11	88.40
毁坏	0	2	4	19	298	92.26
精确率 (%)	95.83	88.33	87.11	86.65	95.24	90.40
						90.63

表 6 基于 ANN 模型的精确率和召回率

Table 6 Precision and recall of ANN-based method

	基本完好	轻微破坏	中等破坏	严重破坏	毁坏	召回率 (%)
基本完好	317	13	1	0	0	95.77
轻微破坏	12	330	15	5	0	91.16
中等破坏	3	15	349	19	2	89.55
严重破坏	0	0	12	325	18	91.55
毁坏	0	1	3	28	274	89.54
精确率 (%)	95.48	91.92	91.84	86.21	93.20	91.59
						91.73

表 7 三种模型的 F1 分数与预测时间

Table 7 F1 scores and prediction speed of the three methods

机器学习算法	F1 分数 (%)	预测时间 (s)
SVM	91.49	0.000 20
RF	90.51	0.000 20
ANN	91.64	0.000 36

样本的预测时间也可达到 0.36 ms,就预测效率来看,3 个模型均能够到达快速预测的目的。

3.4 选取不同输入变量的预测结果对比分析

为了验证本文选用的结构参数与地震动参数相结合作为输入参数的合理性和有效性,本文还将只选用结构参数或地震动参数作为输入参数与本文提出的方法预测结果进行了对比分析,不同输入

表 8 具有不同输入变量方法的准确率

Table 8 Accuracy of methods with different input variables

	结构参数 (%)	地震动参数 (%)	结构+地震动 (%)
SVM	53.53	63.13	91.45
RF	55.28	62.89	90.30
ANN	61.34	60.32	91.56

参数得到的准确率如表 8 所示,准确率表示为预测正确的样本个数与全部样本的比值。

从表 8 可以看出,仅选用结构参数或地震动参数作为输入参数时,基于 3 种机器学习模型建立的方法的准确率均为 60% 左右。而将结构参数和地震动参数共同作为输入参数时,3 种方法均取得了较高的准确率,远远高于仅选用结构参数或地震动参数作为输入参数时的准确率。由此可见,本文将结构参数和地震动参数共同作为输入参数极大地提高了基于机器学习模型的砖砌体震害快速预测的效果。

3.5 输入参数优化

如上所述,由初步选定的 19 个输入参数训练得到的基于 SVM、RF 及 ANN 三个模型的砖砌体震害快速预测方法虽然均能取得良好的预测效果,但表 1 中输入参数间仍存在较高的相关性,需要进一步优化。为此,本节将通过计算各参数之间的相关系数和逐步去除相关程度较高的参数的方法对输入参数的数量进行优化。各参数之间的 Pearson 相关系数 ρ_{xy} 的计算公式如下所示:

$$\rho_{xy} = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y}, \tag{8}$$

$$\text{cov}(x, y) = E(xy) - E(x)E(y), \tag{9}$$

式中, $E(x)$ 与 $E(y)$ 分别是 x 与 y 的数学期望; $E(xy)$ 表示 xy 的数学期望; σ_x 和 σ_y 分别为变量 x 和 y 的标准差; $|\rho_{xy}|$ 值越大则说明参数间相关程度越高。

图 4 给出了各参数间的线性相关系数热力图,颜色越深说明相关性程度越高。参考文献(陈健云等,2017)后,认为相关系数高于 0.8 时,两者之间存在着较强相关性,即可去除其中一个参数。从图 4 中可以看出,结构参数中楼层数与总高具有较强的相关性,具有较强相关性的地震动参数有 PGA、Ia、CAA、ASI、EPA、HI、PGV。考虑到实际中更容易获取建筑的楼层数,因此可去除总高这一参数。由于 Morfidis and Kostinakis(2017)已经研究了机器学习模型中地震动参数对于结构损伤状态的影响程度,在此基础上,综合考虑各系数之间的相关性系数,

表 9 输入参数数量与每个模型对应的准确率

Table 9 The number of input parameters and the corresponding accuracy of each model

	输入参数数量							
	19	18	17	16	15	14	13	12
SVM(%)	91.45	90.99	90.59	90.01	89.78	89.38	89.04	88.64
RF(%)	90.30	90.47	90.11	90.24	90.36	90.59	90.41	90.01
ANN(%)	91.56	90.70	90.37	90.30	90.07	89.84	89.27	87.09

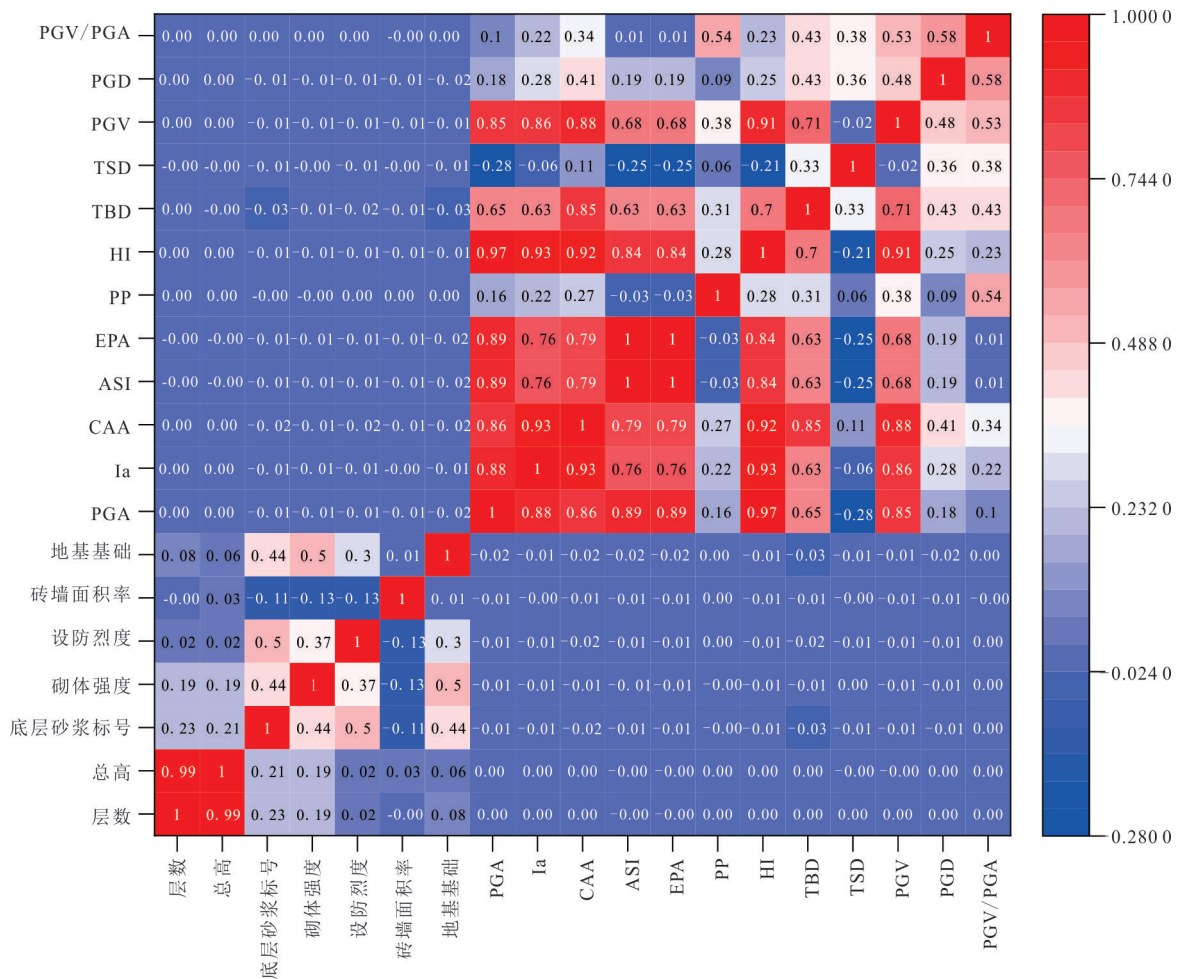


图 4 相关系数热力图

Fig.4 Correlation coefficient heat map

依次去除地震动参数 EPA、PGA、Ia、CAA、ASI、PGV。表 9 列出了按照上述方法依次去除总高、EPA、PGA、Ia、CAA、ASI、PGV 等参数后的 3 个机器学习模型在测试集上的准确率。从表 9 中可以看出, ANN 模型与 SVM 模型受参数数量的变化影响较大。而 RF 模型随着参数的减少, 模型的准确率并无明显下降, 受参数数量变化的影响较小, 且当采用 12 个参数进行预测时亦能达到 90.01% 的准确率。此外, 相比于 SVM 与 ANN 模型, RF 模型中的超参数调参过程更加简单。因此, 优化输入参数

后, RF 模型更适合用于建立砖砌体结构的震害快速预测方法, 最终优化后的输入参数为层数、底层砂浆强度、砌体强度、砖墙面积率、设防烈度、地基基础、PGD、HI、PP、PGV/PGA、TBD、TSD。

3.6 模型验证

如上所述, 优化输入参数后, 本文最终筛选出的最优方法为基于 RF 模型的砖砌体结构震害快速预测方法。为了进一步验证优化输入参数后 RF 模型的性能, 在已有的数据集之外重新选取了有代表性的 15 栋砖砌体房屋和 8 条地震动, 在进行非线性

表 10 新测试集的混淆矩阵

Table 10 Confusion matrix for the new test set

	基本完好	轻微破坏	中等破坏	严重破坏	毁坏
基本完好	46	0	0	0	0
轻微破坏	2	7	4	0	0
中等破坏	0	1	8	1	0
严重破坏	0	0	6	5	5
毁坏	0	0	4	1	30

表 11 试验与数值模拟样本的混淆矩阵

Table 11 Confusion matrix of experimental and numerical simulation samples

	基本完好	轻微破坏	中等破坏	严重破坏	毁坏
基本完好	10	0	0	0	0
轻微破坏	0	1	0	0	0
中等破坏	0	0	1	1	0
严重破坏	0	0	0	1	4
毁坏	0	0	0	0	3

时程分析后得到 120 组样本作为新测试集,采用新测试集对优化输入参数后的 RF 模型进行了验证,发现优化输入参数后的 RF 模型的预测准确率可达到 80%,其中,有 108 个样本预测的破坏状态与非线性时程分析得到的破坏状态一致.不一致的 24 个样本中 20 个样本破坏等级相差一级,4 个差两个破坏等级,具体预测结果见表 10.由此表明,在面对全新的未知数据时,优化输入参数后的 RF 模型依然展现了优秀的准确预测能力.

除此之外,本文从现有的试验结果与数值模拟(罗梓桐,2018;杨腾飞,2018;王晓等,2019;张令心和孔建辉,2021)中选出 21 个样本,对优化输入参数后的 RF 模型进行验证,验证结果如表 11 所示.从表中可以看出,其中 16 个样本与试验结果或数值模拟的结果一致,不一致的 5 个样本中破坏等级与试验或数值模拟的破坏等级仅相差一级.通过与试验、数值模拟结果的对比,进一步验

证了优化输入参数后的 RF 模型具有良好的预测能力,也进一步证明了基于优化输入参数后 RF 模型可以实现砖砌体结构的震害快速预测.

4 结论

本文基于 3 种机器学习模型(SVM、RF、ANN)建立了一种基于地震动参数与结构参数的砖砌体结构震害快速预测方法.初步选取与房屋震害程度相关性较强的 7 个结构参数和与地震动的频谱、持时、峰值特性相关的 12 个地震动参数作为输入变量,利用精确率、召回率和 F1 分数作为评价指标对各模型的性能进行分析.然后对输入参数进行优化,研究了输入参数数量对预测模型的影响,从中筛选出优化输入参数后预测性能最优的模型,并对最优模型进行验证.得到结论如下:

(1)采用 19 个参数基于 3 种机器学习模型的砖砌体结构震害快速预测方法的准确率均达到了 90% 以上,单栋建筑物的预测时间可达到毫秒级,3 种方法均能够准确、快速对砖砌体结构震害进行预测.

(2)与只选用结构参数或地震动参数作为输入参数相比,本文将结构参数和地震动参数共同作为输入参数能极大地提高对砖砌体结构震害预测的效果,预测准确率高达 90% 以上,远远高于仅使用一类参数作为输入的准确率.

(3)经参数优化后,随着输入参数数量的减少,SVM 模型与 ANN 模型的预测能力有所下降,而 RF 模型受影响程度较小,仍具有良好的预测效果,且当采用 12 个参数进行预测时,预测准确率同样可达到 90.01%.

(4)采用新测试集对所筛选出的基于优化输入参数后的 RF 模型的预测方法进行验证并发现,预测的准确率可达到 80%.且从现有的试验结果与数值模拟中选出 21 个样本,对优化输入参数后的 RF 模型进一步验证发现,RF 模型仍具有良好的预测能力.因此,基于优化输入参数后的 RF 模型的预测方法可作为砖砌体结构震害快速预测方法.

值得一提的是,在本文方法研究中,所建立的数据库中仅考虑结构用途为住宅的房屋,且房屋均按照《建筑抗震设计规范》GBJ11-89 设计,因此,并没有将用途、建造年代、场地类型等信息列入输入参数,且设防烈度不完整.为了方法具有普适性,需要考虑用途、建造年代等参数的影响.如果扩充多

种用途、不同建造年代的房屋样本,可以进一步完善结构参数的选取,提高预测准确性。

References

- Breiman, L., 1996. Bagging Predictors. *Mach Learn*, 24(2): 123–140.
- Chen, J.Y., Li, J., Han, J.C., et al., 2017. Correlation between Ground Motion Intensity Indexes and Seismic Responses of Frame Structures. *Journal of Vibration and Shock*, 36(3): 105–112, 144 (in Chinese with English abstract).
- Gunn, S. R., 1998. Support Vector Machines for Classification and Regression. *ISIS Technical Report*, 14(1): 5–16.
- Harirchian, E., Kumari, V., Jadhav, K., et al., 2020. A Machine Learning Framework for Assessing Seismic Hazard Safety of Reinforced Concrete Buildings. *Applied Sciences*, 10(20): 7153. <https://doi.org/10.3390/app10207153>
- Ho, T. K., 1998. The Random Subspace Method for Constructing Decision Forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8): 832–844. <https://doi.org/10.1109/340709601>
- Lautour, O. R., Omenzetter, P., 2009. Prediction of Seismic-Induced Structural Damage Using Artificial Neural Networks. *Engineering Structures*, 31(2): 600–606. <https://doi.org/10.1016/j.engstruct.2008.11.010>
- Liu, B.Y., Ye, L.Y., Jiang, J.J., 2002. Forecasting Seismic Damage in Multistory Masonry Buildings with a Neuro-Fuzzy Approach. *Journal of Tsinghua University (Science and Technology)*, 42(6): 843–846 (in Chinese with English abstract).
- Luo, Z.T., 2018. Study on Seismic Performance of Masonry Structure with Falling Floors Based on Dynamic Elastic-Plastic Analysis (Dissertation). Southwest Jiaotong University, Chengdu (in Chinese with English abstract).
- Mangalathu, S., Sun, H., Nweke, C. C., et al., 2020. Classifying Earthquake Damage to Buildings Using Machine Learning. *Earthquake Spectra*, 36(1): 183–208. <https://doi.org/10.1177/8755293019878137>
- Morfidis, K., Kostinakis, K., 2017. Seismic Parameters' Combinations for the Optimum Prediction of the Damage State of R/C Buildings Using Neural Networks. *Advances in Engineering Software*, 106: 1–16. <https://doi.org/10.1016/j.advengsoft.2017.01.001>
- Morfidis, K., Kostinakis, K., 2018. Approaches to the Rapid Seismic Damage Prediction of R/C Buildings Using Artificial Neural Networks. *Engineering Structures*, 165: 120–141. <https://doi.org/10.1016/j.engstruct.2018.03.028>
- Pan, Z. H., Hong, B., 2014. Influence of Spectral Characteristics and Duration of Ground Motions on Results of IDA. *Journal of Vibration and Shock*, 33(5): 155–159, 199 (in Chinese with English abstract).
- Shaheen, M. S. A., Hakan, G., 2022. Robust Multi-Output Machine Learning Regression for Seismic Hazard Model Using Peak Crust Acceleration Case Study, Turkey, Iraq and Iran. *Journal of Earth Science*, 1–54. <https://doi.org/10.1007/s12583-022-1616-2>
- Tang, H., Chen, G. X., Li, F. M., 2006. Seismic Damage Prediction of Multistory Masonry Buildings Based on BP Neural Network Model. *Earthquake Engineering and Engineering Vibration*, 26(4): 141–146 (in Chinese with English abstract).
- Wang, X., Sun, B.T., Yan, P.L., et al., 2019. Influence Analysis of Constructional Column and Masonry Mortar on Seismic Resistance of Masonry Structure. *Earthquake Prevention Technology*, 14(3): 501–512 (in Chinese with English abstract).
- Wu, B., Qiu, W. X., Xu, S. X., et al., 2022. A Method for Assessing the Probability of Tunnel Collapse Based on Artificial Intelligence Deformation Prediction. *Earth Science*, 1–16 (in Chinese with English abstract).
- Xu, Y. J., Lu, X. Z., Tian, Y., et al., 2020. Real-Time Seismic Damage Prediction and Comparison of Various Ground Motion Intensity Measures Based on Machine Learning. *Journal of Earthquake Engineering*, 1–21. <https://doi.org/10.1080/13632469.2020.1826371>
- Yang, T.F., 2018. Research on the Influence Factors of Seismic Performance of Masonry Structures Based on Numerical Simulation (Dissertation). Xi'an University of Architecture and Technology, Xi'an (in Chinese with English abstract).
- Zhang, G. X., Sun, B. T., 2010. A Method for Earthquake Damage Prediction of Building Groups Based on Multiple Factors. *World Earthquake Engineering*, 26(1): 26–30 (in Chinese with English abstract).
- Zhang, L.X., Dai, J.H., Shen, J.K., et al., 2019. Rapid Prediction Model of Earthquake Damage to Frame Structure Based on LM-BP Neural Network. *Journal of Natural Disasters*, 28(2): 1–9 (in Chinese with English abstract).
- Zhang, L.X., Jiang, J.R., Liu, J.P., 2002. Seismic Vulnerability Analysis of Multistory Dwelling Brick Buildings. *Earthquake Engineering and Engineering Vibration*, 22(1): 49–55 (in Chinese with English abstract)
- Zhang, L. X., Kong, J. H., 2021. Comparative Analysis for

Seismic Performance of Masonry Structure with Bottom Frame Designed with New and Old Codes. *Journal of Shenyang University of Technology*, 43(2): 220–227 (in Chinese with English abstract).

Zhang, L.X., Lu, R.F., Zhu, B.J., 2021. Determination and Verification for the Nonlinear Seismic Response Analysis Method and the Damage State Index of Brick Masonry Buildings. *Earthquake Engineering and Engineering Vibration*, 41(3): 1–10 (in Chinese with English abstract).

Zhang, Z. Q., Fan, J. Q., Zeng, P., et al., 2023. Probabilistic Classification Prediction of Tunnel Squeezing Based on Bayesian Network and Its Application during the Investigation and Design Stage. *Earth Science*, 48(5): 1923–1934 (in Chinese with English abstract).

附中文参考文献

陈健云, 李静, 韩进财, 等, 2017. 地震动强度指标与框架结构响应的相关性研究. *振动与冲击*, 36(3): 105–112, 144.

刘本玉, 叶燎原, 江见鲸, 2002. 用模糊人工神经网络方法预测多层砖房震害. *清华大学学报(自然科学版)*, 42(6): 843–846.

罗梓桐, 2018. 基于动力弹塑性分析的掉层砌体结构抗震性能研究(硕士学位论文). 成都: 西南交通大学.

潘志宏, 洪博, 2014. 地震动频谱特性和持时对 IDA 结果影

响的研究. *振动与冲击*, 33(5): 155–159, 199.

汤皓, 陈国兴, 李方明, 2006. 基于 BP 神经网络模型的多层砖房震害预测方法. *地震工程与工程振动*, 26(4): 141–146.

王晓, 孙柏涛, 闫培雷, 等, 2019. 构造柱和砌筑砂浆对砌体结构抗震能力影响分析. *震灾防御技术*, 14(3): 501–512.

吴波, 邱伟兴, 徐世祥, 等, 2022. 基于人工智能变形预测隧道坍塌失效概率评估方法. *地球科学*, 1–16.

杨腾飞, 2018. 基于数值模拟的砌体结构抗震性能影响因素研究(硕士学位论文). 西安: 西安建筑科技大学.

张桂欣, 孙柏涛, 2010. 多因素影响的建筑物群体震害预测方法研究. *世界地震工程*, 26(1): 26–30.

张令心, 戴静涵, 沈俊凯, 等, 2019. 基于 LM-BP 神经网络的钢筋混凝土框架结构震害快速预测模型. *自然灾害学报*, 28(2): 1–9.

张令心, 江近仁, 刘洁平, 2002. 多层住宅砖房的地震易损性分析. *地震工程与工程振动*, 22(1): 49–55.

张令心, 孔建辉, 2021. 新旧规范设计的底框砌体结构抗震性能对比分析. *沈阳工业大学学报*, 43(2): 220–227.

张令心, 鲁若帆, 朱柏洁, 2021. 砖砌体房屋非线性地震反应分析方法和破坏状态指标的确定与验证. *地震工程与工程振动*, 41(3): 1–10.

张志强, 范俊奇, 曾鹏, 等, 2023. 基于贝叶斯网络的隧道勘察设计期大变形灾害概率分级预测与应用研究. *地球科学*, 48(5): 1923–1934.