Proceedings of the Sixteenth International Conference on
Computational Structures Technology
Edited by: P. Iványi, J. Kruis and B.H.V. Topping
Civil-Comp Conferences, Volume 14, Paper 3.1
Civil-Comp Press, Edinburgh, United Kingdom, 2026
ISSN: 2753-3239, doi: 10.4203/ccc.14.3.1

Predicting Earthquake-Induced Building Damage Grades Using Machine Learning and Explainable AI

E. K. Nyarko[1], S. Czarnecki[2], A. D. R. Troncoso Garcia[3], M. Hadzima-Nyarko[4], F. Martinez Alvarez[3], B. Bulajić[5], E. Isik[6] and D. Radu[7]

[1] Faculty of Electrical Engineering, Computer Science and Information Technology Osijek, Josip Juraj Strossmayer University of Osijek, Croatia
[2] Department of Materials Engineering and Construction Processes, Wroclaw University of Science and Technology, Wrocław, Polska
[3] Data Science & Big Data Lab, Pablo de Olavide University, Spain
[4] Faculty of Civil Engineering and Architecture, Josip Juraj Strossmayer, University of Osijek, Croatia
[5] Faculty of Technical Sciences, University of Novi Sad, Serbia
[6] Department of Civil Engineering, Bitlis Eren University, Turkey
[7] Faculty of Civil Engineering, Transilvania University of Braşov, Romania

Abstract

Machine learning techniques are increasingly being applied in civil engineering to improve the assessment of structural performance and post-earthquake damage evaluation. In this study, several machine learning approaches, including regression models, classification models, and Artificial Neural Networks, were investigated for the prediction of seismic building damage grades using a database composed of structural, geometrical, and seismic parameters collected after major earthquakes in Croatia and Türkiye. Different data stratification strategies were considered to evaluate the robustness and predictive capability of the developed models. The results demonstrated that machine learning methods can effectively capture complex nonlinear relationships associated with seismic damage assessment and provide accurate prediction of building damage grades. Among the investigated approaches, ensemble-based classification models and ANN architectures achieved the most favorable predictive performance depending on the adopted stratification strategy. To improve model transparency and engineering interpretability, SHapley Additive exPlanations (SHAP) analysis was additionally performed to investigate the influence of individual input variables on the predicted damage grades. The obtained results confirmed the potential of machine learning and explainable artificial intelligence techniques as efficient tools for rapid seismic damage assessment and decision support in post-earthquake engineering applications.

Keywords: building damage grade, machine learning, seismic performance, PGA, random forest, neural network, gradient boosting

1 Introduction

Civil engineering is increasingly benefiting from the rapid development of computational intelligence and data-driven methodologies [1]. Traditional approaches used for structural assessment and post-disaster evaluation are commonly based on empirical observations, simplified analytical assumptions, and deterministic procedures [2]. Although such methods remain fundamental in engineering practice, they often have limited capability in describing highly nonlinear interactions and complex dependencies associated with structural damage processes. Consequently, there is a growing demand for advanced analytical tools capable of improving the accuracy and efficiency of damage assessment in civil engineering applications [3]. Recent advances in machine learning (ML) have created new possibilities for analysing complex engineering problems using data-oriented approaches [4], [5]. In structural and earthquake engineering, ML methods have been successfully applied to tasks such as structural health monitoring, seismic vulnerability analysis, damage detection, and structural performance evaluation [6]. These techniques enable the identification of hidden relationships between structural characteristics and observed damage patterns without the need for explicitly defined physical equations [7]. Among the emerging applications of ML in civil engineering, building damage grade prediction has attracted particular attention due to its importance in post-earthquake assessment and disaster management [8], [9]. Accurate prediction of building damage levels is essential for rapid safety evaluation, emergency response planning, and allocation of reconstruction resources. Conventional field inspections are often labor-intensive, time-consuming, and dependent on expert interpretation, whereas ML models offer the possibility of fast and automated classification of structural damage states based on available building data [10]. The increasing availability of post-disaster databases and structural datasets provides favorable conditions for implementing ML-based predictive systems. Since building damage is influenced by multiple interacting parameters, including structural, geometrical, and environmental factors, the resulting relationships are often highly nonlinear and difficult to capture using traditional analytical methods. In such cases, ML algorithms are particularly effective due to their ability to learn complex patterns directly from data [11]. In recent years, various ML approaches, including Random Forests [12], Gradient Boosting techniques [13], Logistic Regression, and Artificial Neural Networks (ANNs), have demonstrated promising performance in classification and prediction problems related to seismic damage assessment [14]. However, despite their high predictive capabilities, ML models are frequently criticized for limited interpretability and the difficulty of understanding the decision-making mechanisms behind their predictions. Therefore, explainable artificial intelligence methods are becoming increasingly important in engineering applications. In this study, several ML techniques were investigated for the prediction of building damage grades using a dataset containing 22 building-related parameters and five categorical damage classes. Regression models, classification models, and ANN architectures were evaluated under different data stratification strategies. Furthermore, hyperparameter optimization and cross-validation procedures were applied to improve model robustness and predictive performance. To enhance interpretability, SHapley Additive exPlanations (SHAP) analysis was additionally employed to investigate the influence of individual input features on the predicted damage grades.

2 Methodology

2.1 Dataset Description

Earthquakes represent one of the most significant hazards affecting the built environment, particularly in regions characterized by vulnerable building stock and insufficient seismic resistance. Rapid and reliable assessment of structural damage following seismic events is essential for emergency response, safety evaluation, economic loss estimation, and long-term mitigation planning. In this study, a comprehensive database was developed to investigate the relationship between building characteristics and observed seismic damage.

The dataset was assembled using post-earthquake inspection data collected after the 2020 Petrinja earthquake in Croatia and the 2023 Kahramanmaraş earthquake in Türkiye. The final database consisted of 560 building samples described by 23 variables, including 22 input parameters and one target variable representing the building damage grade.

The input data included both categorical and numerical parameters describing the geometrical, structural, and functional properties of the buildings. The categorical variables included information related to floor structure, building position within a block, construction material, building purpose, roof type, and other structural characteristics. Numerical variables described parameters such as the year of construction, building dimensions, floor area, and peak ground acceleration (PGA). The target variable corresponded to the observed building damage grade classified into five categories ranging from minor damage to severe destruction.

Prior to model development, the dataset was preprocessed and divided into training and testing subsets. Three different stratification strategies were considered to evaluate the robustness of the developed models: stratification by City, stratification by Damage Grade, and combined stratification by both City and Damage Grade. In each case, 80% of the data were used for model training and validation, while the remaining 20% were reserved for testing purposes. The stratification process ensured that the distribution of selected class labels remained consistent between the training and testing datasets.

The primary objective of the study was to identify the most influential building-related parameters affecting seismic damage grades and to evaluate the capability of different ML methods for accurate damage prediction. In addition to predictive performance assessment, explainability techniques were later applied to improve understanding of the relationship between input variables and predicted damage classes.

2.2 Machine Learning Models

Several groups of machine learning models were investigated in this study, including regression models, classification models, and Artificial Neural Networks (ANNs). The objective was to compare the predictive performance of different ML approaches for the classification of building damage grades.

2.2.1 Regression Models

Several regression-based machine learning models were implemented in this study, including Linear Regression, Ridge Regression, Lasso Regression, Random Forest Regressor, and Gradient Boosting Regressor. Since the target variable represented categorical building damage grades ranging from 1 to 5, the continuous outputs generated by the regression models were rounded to the nearest integer value to obtain the final predicted damage class.

Linear Regression was employed as the baseline regression model due to its simplicity and interpretability. Ridge Regression and Lasso Regression were additionally implemented as regularized extensions of linear regression, incorporating L2 and L1 regularization techniques, respectively. Ensemble-based approaches were also investigated using Random Forest Regressor and Gradient Boosting Regressor models. Hyperparameter optimization for all regression models was performed using grid search combined with 5-fold cross-validation, where the average Mean Squared Error (MSE) was used as the optimization metric.

Model	Hyperparameter	Investigated Values
Ridge Regression	α	$(0.01, 0.1, 1.0, 10.0, 100.0)$
Lasso Regression	α	$(0.0001, 0.001, 0.01, 0.1, 1.0, 10)$
Random Forest Regressor	$n\_estimators$	$(50, 100, 200, 500)$
	max_depth	$(None, 5, 10, 20)$
	$min\_samples\_split$	$(2, 5, 10)$
Gradient Boosting Regressor	$n\_estimators$	$(100, 200, 300)$
	max_depth	$(3, 5, 7)$
	$learning\_rate$	$(0.01, 0.1, 0.2)$

Table 1: Investigated hyperparameters for regression-based machine learning models.

2.2.2 Classification Models

Three classification-based machine learning approaches were investigated in this study, namely Logistic Regression, Random Forest Classifier, and Gradient Boosting Classifier. These models were directly trained to predict categorical building damage grades.

Logistic Regression was employed as the baseline classification model due to its computational efficiency and interpretability. Ensemble-based approaches were additionally investigated using Random Forest Classifier and Gradient Boosting Classifier models. Hyperparameter optimization for all classification models was carried out using grid search combined with 5-fold cross-validation, where the weighted F1-score was used as the optimization metric.

2.2.3 Artificial Neural Network Model

A feedforward Artificial Neural Network (ANN) model was additionally developed for the prediction of building damage grades. The ANN architecture consisted of an input layer containing 22 neurons corresponding to the input parameters, two hidden layers, and an output layer consisting of five neurons representing the five damage grade classes. The softmax activation function was applied in the output layer to perform multiclass classification.

The optimal ANN structure was determined using hyperparameter optimization based on grid search and 5-fold cross-validation. The average classification accuracy obtained during cross-validation was used as the optimization metric.

Model	Hyperparameter	Investigated Values
Logistic Regression	$penalty$	$(l1, l2, elasticnet, None)$
	C	$np.logspace(-4, 4, 20)$
	$solver$	$(lbfgs, newton\text{-}cg, liblinear, sag, saga)$
	max_iter	$(100, 1000, 2500, 5000)$
Random Forest Classifier	$n\_estimators$	$(50, 100, 200, 500)$
	max_depth	$(None, 5, 10, 20)$
	$min\_samples\_split$	$(2, 5, 10)$
	$min\_samples\_leaf$	$(1, 2, 4)$
	$max\_features$	$(sqrt, log2, None)$
	$bootstrap$	$(True, False)$
Gradient Boosting Classifier	$n\_estimators$	$(50, 100, 200, 300)$
	max_depth	$(3, 5, 7)$
	$learning\_rate$	$(0.01, 0.1, 0.2)$
	$min\_samples\_split$	$(2, 5, 10)$
	$min\_samples\_leaf$	$(1, 2, 4)$

Table 2: Investigated hyperparameters for classification-based machine learning models.

Hyperparameter	Investigated Values
$batch\_size$	$(10, 20)$
$epochs$	200
$n\_neurons1$	$(2, 5, 10, 15)$
$n\_neurons2$	$(2, 5, 10, 15)$
$optimizer$	$(adam, sgd)$
$activation$	$(relu, tanh)$

Table 3: Investigated hyperparameters for the ANN model.

3 Model Performance

This section presents the performance comparison of the investigated machine learning models under three different data stratification strategies. The predictive capabilities of regression models, classification models, and Artificial Neural Networks

(ANNs) were evaluated using the testing datasets corresponding to each stratification regime.

3.1 Models Trained Using Data Stratified by "City"

The first training regime considered stratification based on the "City" class label. In this approach, both the training and testing datasets preserved the same proportion of samples associated with each city. The objective of this strategy was to evaluate the predictive capability of the developed models under geographically balanced data conditions.

The obtained results indicated that the classification models achieved the highest predictive performance. Among all investigated approaches, the RandomForestClassifier produced the best overall results, achieving a test accuracy of 0.9286 and a weighted F1-score of 0.9294. The model demonstrated particularly high recall values for damage grades 3 and 4, indicating strong classification capability for these categories.

The regression-based approaches also provided satisfactory performance, with the RandomForestRegressor producing the most accurate predictions among the regression models. However, the ANN model achieved lower predictive accuracy compared to the best-performing classification approach.

Model Type	Best Estimator	Accuracy	Weighted F1
Regression	RandomForestRegressor	0.8482	0.8459
Classification	RandomForestClassifier	0.9286	0.9294
ANN	ANN	0.8393	0.8425

Table 4: Comparison of best-performing models on the test dataset for the "City" stratification strategy.

Damage Grade	1	2	3	4	5
1	31	1	1	0	1
2	1	26	0	0	0
3	0	0	27	0	0
4	0	0	0	8	0
5	0	0	0	4	12

Table 5: Confusion matrix of the RandomForestClassifier model for the dataset stratified by "City".

3.2 Models Trained Using Data Stratified by "Damage Grade"

The second training regime involved stratification based exclusively on the "Damage Grade" class label. This approach ensured that the proportions of damage categories remained consistent between the training and testing datasets.

The obtained results demonstrated that the RandomForestClassifier again achieved the highest predictive performance among all investigated models, reaching a test accuracy of 0.8571 and a weighted F1-score of 0.8580. Compared to the "City" stratification strategy, the predictive performance slightly decreased, suggesting that damage grade prediction represented a more challenging classification task.

Among the regression-based approaches, the GradientBoostingRegressor demonstrated competitive predictive capability after rounding the regression outputs to integer damage classes. The ANN model achieved relatively high cross-validation accuracy during training; however, its generalization capability on the testing dataset was lower compared to the classification-based approaches.

Model Type	Best Estimator	Accuracy	Weighted F1
Regression	GradientBoostingRegressor	0.8482	0.8502
Classification	RandomForestClassifier	0.8571	0.8580
ANN	ANN	0.8036	0.8063

Table 6: Comparison of best-performing models on the test dataset for the "Damage Grade" stratification strategy.

Damage Grade	1	2	3	4	5
1	27	3	0	0	1
2	2	24	1	0	0
3	3	0	21	1	0
4	0	0	11	3	0
5	1	0	0	1	13

Table 7: Confusion matrix of the RandomForestClassifier model for the dataset stratified by "Damage Grade".

3.3 Models Trained Using Data Stratified by "City" and "Damage Grade"

The final training regime considered simultaneous stratification using both "City" and "Damage Grade" class labels. This approach ensured that the distributions of both variables remained balanced within the training and testing datasets.

The obtained results indicated that the ANN model achieved the highest predictive performance under this stratification strategy. The optimized ANN architecture achieved a test accuracy of 0.8929 and a weighted F1-score of 0.8947, outperforming both regression and classification-based approaches.

The results additionally demonstrated that the GradientBoostingClassifier achieved highly competitive predictive capability, while the RandomForestRegressor remained the most accurate regression-based approach. Overall, the joint stratification strategy produced robust predictive performance across all investigated model categories.

Model Type	Best Estimator	Accuracy	Weighted F1
Regression	RandomForestRegressor	0.8571	0.8565
Classification	GradientBoostingClassifier	0.8839	0.8838
ANN	ANN	0.8929	0.8947

Table 8: Comparison of best-performing models on the test dataset for the combined "City" and "Damage Grade" stratification strategy.

Damage Grade	1	2	3	4	5
1	29	0	0	0	2
2	0	25	0	1	1
3	2	1	22	0	0
4	0	1	12	1	0
5	0	0	3	12	0

Table 9: Confusion matrix of the ANN model for the dataset stratified by "City" and "Damage Grade".

3.4 Comparative Performance Analysis and SHAP Interpretation

A comparative summary of the obtained model performances indicated that the adopted data stratification strategy significantly influenced the predictive capability of the investigated machine learning models. The highest predictive performance was achieved

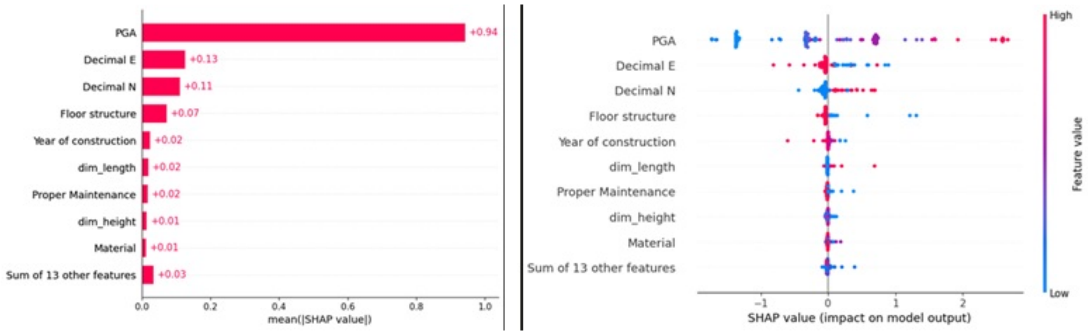

Figure 1: SHAP plots for the RandomForestClassifier trained on the dataset stratified according to the "City" class label.

for the dataset stratified according to the "City" class label, where the RandomForest-Classifier achieved a test accuracy of 0.9286 and a weighted F1-score of 0.9294. In contrast, the dataset stratified solely according to the "Damage Grade" label represented the most challenging prediction scenario, with the best-performing models achieving test accuracies in the range of approximately 85–86%. The joint stratification strategy based on both "City" and "Damage Grade" labels provided favorable conditions for the ANN model, which achieved a test accuracy of 0.8929 and a weighted F1-score of 0.8947. Overall, all investigated machine learning approaches demonstrated satisfactory predictive performance for seismic building damage assessment.

To improve the interpretability and transparency of the developed machine learning models, SHapley Additive exPlanations (SHAP) analysis was additionally performed for the best-performing models obtained under each stratification strategy. The SHAP method enabled quantitative evaluation of the influence of individual input variables on the predicted building damage grades. Global feature importance was assessed using SHAP feature importance plots, while local feature contributions and the direction of feature influence were analyzed using SHAP beeswarm plots. These analyses provided additional insight into the relationships between structural, geometrical, and seismic parameters and the predicted seismic damage grades.

Figure 1 presents the SHAP plots obtained for the RandomForestClassifier trained on the dataset stratified according to the "City" class label. The figure illustrates both the global feature importance ranking and the local contribution of individual variables to the prediction results.

4 Conclusions

This study investigated the applicability of various machine learning techniques for the prediction of seismic building damage grades using a dataset composed of structural, geometrical, and seismic parameters collected after the earthquakes in Croatia

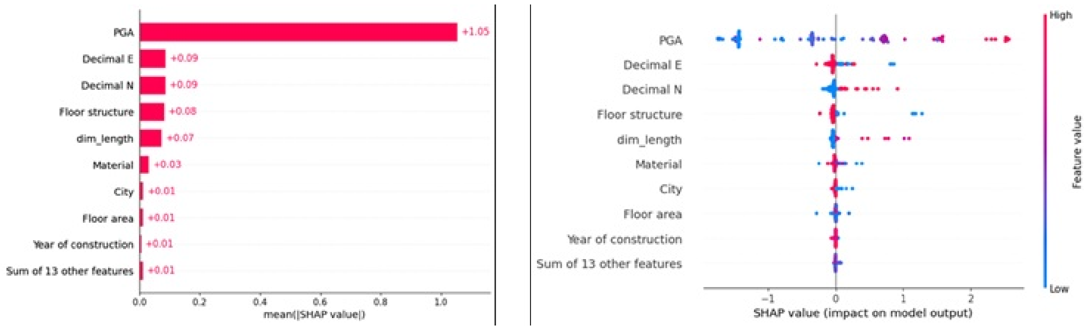

Figure 2: SHAP plots for the RandomForestClassifier trained on the dataset stratified according to the “Damage Grade” class label.

and Türkiye. Several regression-based, classification-based, and Artificial Neural Network (ANN) models were developed and evaluated under three different data stratification strategies, namely stratification by “City”, stratification by “Damage Grade”, and combined stratification by both “City” and “Damage Grade”.

- The obtained results demonstrated that machine learning methods were capable of accurately predicting building damage grades and effectively capturing complex nonlinear relationships between the investigated input parameters and observed seismic damage. Among the considered approaches, classification models generally achieved the highest predictive performance. In particular, the RandomForestClassifier demonstrated superior performance for datasets stratified according to “City” and “Damage Grade”, while the ANN model achieved the best results for the combined stratification strategy.

- The highest predictive capability was achieved for the dataset stratified according to the “City” class label, where the RandomForestClassifier reached a test accuracy of 0.9286 and a weighted F1-score of 0.9294. The results additionally indicated that prediction based solely on the “Damage Grade” stratification represented the most challenging classification scenario due to the increased complexity and imbalance of the target variable. Nevertheless, all investigated approaches achieved satisfactory predictive performance, with test accuracies generally ranging between approximately 0.84 and 0.93.

- The performed hyperparameter optimization demonstrated the importance of proper model configuration for improving predictive capability and model robustness. In the case of ANN models, the Adam optimizer and lower batch sizes generally produced more favorable results compared to alternative configurations.

- To improve the interpretability of the developed machine learning models, SHAP-based explainable artificial intelligence analysis was additionally performed. The SHAP methodology enabled both global and local interpretation of model predictions by quantifying the contribution of individual input parameters to the predicted building damage grades. The obtained SHAP results demonstrated that explainable AI methods can significantly enhance the transparency and engineering interpretability of machine learning approaches applied to seismic damage assessment problems.

The conducted analyses confirmed that machine learning techniques constitute a promising and efficient tool for rapid seismic damage prediction and post-earthquake assessment. The proposed methodology may support decision-making processes related to emergency response, structural safety evaluation, and long-term seismic risk mitigation. Future research should focus on extending the database, incorporating additional structural and geospatial parameters, and investigating more advanced deep learning and explainable artificial intelligence techniques to further improve predictive capability and model generalization.

Acknowledgements

This work presents part of the research activities within the project 2023-1-HR01-KA220-HED-000165929 ”Intelligent Methods for Structures, Elements and Materials” [https://im4stem.eu/en/home/] co-funded by the European Union under the program Erasmus+ KA220-HED - Cooperation partnerships in higher education.

References

[1] A.R. Byrd, M. Geheran, “Engineering AI Trustworthiness: A Framework of Civil Engineering Ethics and Explainable AI to Build AI Bridges from Research to Practice”, Journal of Computing in Civil Engineering, 40(3), 2026. https://doi.org/10.1061/JCCEE5.CPENG-6649.

[2] S.Q. Li, Y.S. Chen, H.B. Liu et al., “Empirical seismic vulnerability assessment model of typical urban buildings”, Bulletin of Earthquake Engineering, 21, 2217–2257, 2023. https://doi.org/10.1007/s10518-022-01585-8.

[3] J.C. Reyes-Suárez, M. Buitrago, B. Barros, S. Mammeri, N. Makoond, C. Lázaro, B. Riveiro, J.M. Adam, “Latent resistance mechanisms of steel truss bridges after critical failures”, Nature, 645, 101–107, 2025.

[4] N. Bonyadi, R. Talami, B. Lee, “Breaking-down building design problems with decomposition approaches: A review”, Energy and Built Environment, 2025. https://doi.org/10.1016/j.enbenv.2025.05.005.

[5] V. Saback, C. Popescu, T. Blanksvärd, B. Täljsten, “Analysis of Digital Twins in the Construction Industry: Practical Applications, Purpose, and Parallel with other Industries”, Buildings, 14(5), 1361, 2024. https://doi.org/10.3390/buildings14051361.

[6] S. Hu, T. Guo, M.S. Alam, Y. Koetaka, E. Ghafoori, T.L. Karavasilis, “Machine learning in earthquake engineering: A review on recent progress and future

trends in seismic performance evaluation and design", Engineering Structures, 340, 120721, 2025. https://doi.org/10.1016/j.engstruct.2025.120721.

[7] M. Azimi, A.D. Eslamlou, G. Pekcan, "Data-Driven Structural Health Monitoring and Damage Detection through Deep Learning: State-of-the-Art Review", Sensors, 20(10), 2778, 2020. https://doi.org/10.3390/s20102778.

[8] Z. Li, H. Lei, E. Ma, J. Lai, J. Qiu, "Ensemble technique to predict post-earthquake damage of buildings integrating tree-based models and tabular neural networks", Computers & Structures, 2023.

[9] M. Al-Rawashdeh, M. Al Nawaiseh, I. Yousef, M. Bisharah, S. Alkhadrawi, "Predicting building damage grade by earthquake: a Bayesian Optimization-based comparative study of machine learning algorithms", Asian Journal of Civil Engineering, 2024. https://doi.org/10.1007/s42107-023-00771-6.

[10] S.K.C. Sajan, A. Bhusal, D. Gautam, R. Rupakhety, "Earthquake damage and rehabilitation intervention prediction using machine learning", Engineering Failure Analysis, 2023.

[11] W. Chen, L. Zhang, "Building vulnerability assessment in seismic areas using ensemble learning: A Nepal case study", Journal of Cleaner Production, 2022. https://doi.org/10.1016/j.jclepro.2022.131418.

[12] M.M. Rosso, A. Aloisio, N. Di Battista, T. d'Alfonso, C. Demartino, G. Quaranta, "Machine learning model for predicting post-earthquake reconstruction grants of masonry buildings from seismic vulnerability features", International Journal of Disaster Risk Reduction, 140, 106123, 2026. https://doi.org/10.1016/j.ijdrr.2026.106123.

[13] S. Öztürk, "Beyond PGA: Role of tank geometry and filling level in seismic fragility of atmospheric storage tanks-a machine learning approach", Journal of Loss Prevention in the Process Industries, 102, 106033, 2026. https://doi.org/10.1016/j.jlp.2026.106033.

[14] P. Si, L. Tang, S. Tian, Y. Liu, X. Ling, "Seismic risk assessment of railway embankments on spatially variable loose deposit slopes: A stacking ensemble machine learning-based approach", Computers and Geotechnics, 192, 107841, 2026. https://doi.org/10.1016/j.compgeo.2025.107841.