Earthquake Damage Risk Classification Using Machine Learning Models Based on Built-Up Area Indices from Satellite Imagery

Gunawan Prayitno*, Eko Sediyono, Irwan Sembiring, Sri Yulianto Joko Prasetyo
Department of Information Technology, Satya Wacana Christian University, Salatiga, Indonesia

***Abstract*—One of the challenges in assessing earthquake damage risk in areas with high seismic activity and limited data is the lack of a detailed building inventory and the absence of available data. Therefore, a remote sensing and machine learning framework is needed that can utilize the built-up area index with NDBI from multitemporal Landsat 8 imagery as a vulnerability proxy, combined with ISGS seismic hazard data, topographic variables from DEM, and surface geology. Support Vector Machine (SVM) and Random Forest (RF) classifiers can be trained and developed to categorize vulnerability classes and evaluate results in terms of accuracy, precision, recall, and F1 score. The SVM achieved an accuracy of 0.970, with precision, recall, and F1 scores all reaching 0.97, while the RF achieved 0.96. This study focuses on Nabire Regency, Central Papua, Indonesia, utilizing various approaches, scaling methods, and a relatively low-cost methodology for the initial screening of earthquake-prone areas without the need for field data collection, with the potential to apply these methods to other seismic regions.**

***Keywords*—Seismic risk assessment; vulnerability proxy; remote sensing; machine learning; data scarcity**

I. Introduction

High-seismicity regions, such as Indonesia, face significant challenges in assessing earthquake-induced building damage risk, particularly in developing countries where urbanization is rapid but detailed building inventory data is scarce. Traditional seismic risk assessment models generally rely on structural-level information such as material types, structural systems, and building age, which are often difficult to obtain comprehensively, especially in remote regions with limited data collection capacity [1],[2]. Consequently, there is a need for an effective and affordable approach to earthquake risk mapping that does not solely depend on field-based data, which is often inaccessible.

In recent years, the use of remote sensing (RS) and Geographic Information Systems (GIS) has rapidly evolved as a solution to overcome the limitations of physical field data collection. Satellite imagery holds substantial potential for characterizing the built-up area (BUA), which can serve as a proxy for vulnerability. Indices derived from satellite imagery, such as the Normalized Difference Built-up Index (NDBI) and Building Area Index (BAI), are particularly useful in mapping urbanization intensity, building density, and land-use changes, which are critical in assessing earthquake vulnerability [3],[4].

Although several studies have employed remote sensing for vulnerability assessment, most of these remain descriptive or fragmented, lacking integration with machine learning (ML) models to provide more accurate and reliable results for earthquake risk analysis [5],[6].

An emerging innovative approach is the combination of machine learning with remote sensing, which facilitates the efficient and systematic processing of spatial data. Techniques such as Support Vector Machine (SVM) and Random Forest (RF) have proven effective in classifying vulnerability by utilizing spatial data, such as built-up area indices and other geospatial parameters. SVM, with its ability to handle non-linear and high-dimensional data, is particularly suitable for vulnerability analysis based on remote sensing data [7],[8]. Random Forest, as an ensemble algorithm, offers the advantage of reducing model variance and providing robust predictions, even with large and complex datasets [9].

However, the biggest challenge in earthquake risk modeling is the class imbalance in data, especially when the number of damaged buildings after an earthquake is much smaller than the number of non-damaged ones. In such cases, imbalanced data can affect the performance of machine learning models. Therefore, this research proposes using machine learning techniques without oversampling, leveraging the strengths of SVM and RF to maintain high classification accuracy, even with imbalanced data [10],[11].

Nabire District in Central Papua has been chosen as the study area for this research. Nabire is a seismically active region with significant limitations in building inventory data, making it an ideal case study for applying a remote sensing and machine learning-based earthquake damage risk model. This study aims to develop a methodological framework that can be used in similar data-scarce regions while still providing accurate and reliable earthquake risk assessments.

This research not only provides a vulnerability map but also contributes to disaster mitigation by offering a transferable framework that can be applied to areas with limited structural data. Additionally, this study contributes to the development of scalable, data-driven models that can be utilized by disaster management agencies to plan and implement more effective mitigation strategies [12],[13],[14].

*Corresponding author.

II. Related Work

Conventional earthquake vulnerability assessments depend on structural attributes, building age, material types, and several other parameters that require detailed classification and more specific approaches to ensure they can be widely applied across various regions—particularly where detailed, systematic building inventories are not yet available [1], [2].

To reduce reliance on field data, several studies have used remote sensing proxies to analyze built-up area indices, land cover changes, exposure models, and simplified replacement cost estimates, as well as structural exposure estimates on a broader scale [3], [6]. Several of these remote sensing studies have shown that indicators derived from satellite data can replace detailed inventories; however, most remain descriptive and exploratory in nature, reporting spatial patterns of exposure without linking them to predictive classification models [3], [5], [6].

The next step in the research is to apply machine learning to predict seismic damage. Several machine learning methods used for this purpose include Support Vector Machines, Random Forests (RF), and deep learning-based image models such as Mask R-CNN on satellite and aerial imagery to predict damage following an earthquake [7], [8], [9].

These studies demonstrate the performance of the most relevant and accurate classification systems; however, they are generally conducted in contexts rich in labeled data and post-event damage information available for training and data validation. Several reports indicate that model performance declines significantly when applied to areas with different local conditions or without such labels [4], [10]. A small number of studies have integrated machine learning with GIS-based multi-criteria methods, including AHP-based weighting of hazard, exposure, and vulnerability layers, to translate model results into decision-support maps [12], [13], [14]. This ML+GIS integration workflow is rarely used in the more specific and data-scarce context of earthquakes, and relies entirely on proxies derived from satellite data as the primary exposure input.

Specifically, following a comprehensive review, there are still various gaps that need to be addressed. Previous studies using remote sensing have successfully integrated built-up area proxies with supervised machine learning classification and multi-criteria hazard weighting (AHP) into a single workflow, enabling explicit data prediction and modeling for areas lacking field-based building inventories.

This study also presents an integrated framework that can be applied in Nabire Regency, Central Papua, Indonesia, as a case study with data limitations, while explicitly addressing methodological limitations—such as the nature of vulnerability labels based on proxies and clustering results—which are often overlooked in previous studies; this constitutes a novelty of this research.

III. Method

This study proposes a remote sensing– and machine learning–based framework to classify earthquake-related building damage risk in a data-limited seismic region. The overall procedure covers data acquisition, preprocessing, and feature construction, supervised model development, and performance evaluation, with the full methodological sequence summarized in Fig. 1. Four primary datasets are utilized. Multi-temporal Landsat 8 imagery (2022–2025) is employed to derive built-up area indicators that represent urbanization intensity and built-surface density in the study area [15],[13]. Seismic hazard information is compiled from the United States Geological Survey (USGS) earthquake catalog to capture earthquake occurrence characteristics that may influence vulnerability patterns [16],[17]. A Digital Elevation Model (DEM) is included to derive topographic attributes—particularly slope—that can affect spatial damage distribution during seismic events [18],[11]. In addition, surface geology maps are incorporated to describe lithological conditions that may amplify or attenuate earthquake impacts and thus influence the vulnerability of built environments [12],[19].

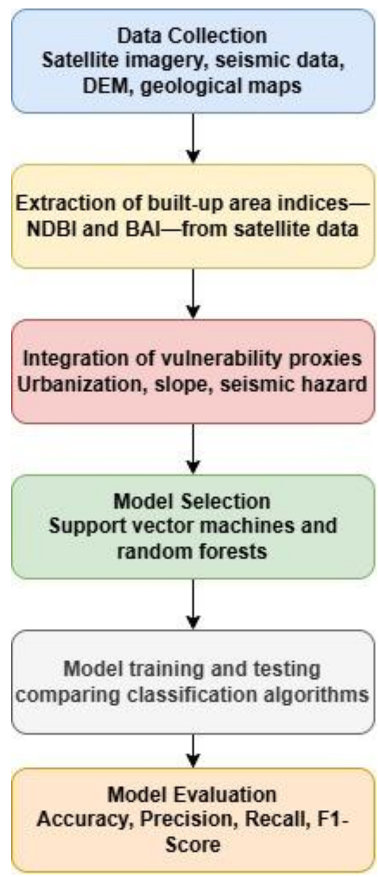

Fig. 1. Flowchart or orchestra of machine learning model selection for earthquake damage risk classification.

Before modeling, all datasets are prepared to ensure spatial comparability and analytical consistency. Satellite imagery is processed to generate built-up area proxies that quantify the extent and intensity of built surfaces. In this study, the Normalized Difference Built-up Index (NDBI) is calculated from Landsat bands as follows [20],[7]:

$$NDBI = (SWIR - NIR) / (SWIR + NIR) \quad (1)$$

Moreover, SWIR is the Short-Wave Infrared band, and NIR is the Near Infrared band from satellite imagery [21]. To represent vulnerability more comprehensively, built-up indicators are not used in isolation. Complementary spatial proxies are integrated, including slope derived from the DEM and seismic context derived from the earthquake catalog, enabling the model to learn spatial relationships between built-environment conditions and potential earthquake damage risk [17],[18]. This integrated feature set is constructed to better capture multi-factor vulnerability patterns that cannot be represented by a single proxy alone.

This study ultimately examined the SVM and RF methods using a case study approach deemed appropriate for these methods, which have also been proven effective in performing classifications involving high-dimensional and medium-sized geospatial datasets, while identifying commonly encountered conditions and addressing vulnerabilities in remote sensing. This differs from deep learning approaches, which generally require comprehensive data for large-scale labeled training to avoid overfitting.

SVM and RF are reliable even with limited sample sizes, such as those found in the seismic area of Nabire, Central Papua, Indonesia. However, with the principle of margin maximization, SVM enables the construction of robust decision boundaries even when performing nonlinear class separation, and its kernel-based formulation is well-suited for heterogeneous geospatial features derived from various parameters such as built-up area index, earthquake hazard, topography, and geology. RF, which is a decision tree ensemble, is relatively resistant to overfitting, requires minimal parameter tuning, and provides built-in metrics such as variable importance, which support the interpretation of the underlying vulnerability drivers. Both SVM and RF also demonstrate consistent robustness under class-imbalanced conditions without requiring oversampling, which is relevant to the existing imbalance challenges. The combination of these properties makes SVM and RF suitable for further analysis and implementation in this study, with the hope that they can serve as a more precise and accurate framework for vulnerability classification.

There are several deep learning techniques that we are familiar with, such as Convolutional Neural Networks (CNNs) [25]; however, they were not applied in this study due to several practical reasons related to data limitations in the study area—namely, Nabire, Central Papua, Indonesia. Some important details to note regarding the use of these algorithms are that deep learning models require a large volume of labeled training data, which is then used to learn reliable feature representations and achieve good and accurate generalization. The availability of ground-truth samples in the Nabire region remains limited, and using a deep learning architecture increases the risk of overfitting.

The next factor—the input data used in this study—includes settlement indices, earthquake hazard attributes, topological variables, and other specific geological categories. These factors serve as structured geospatial predictors in tabular form—not merely raw image data—and this data format is well-suited for conventional machine learning classification methods such as SVM and RF, without requiring the additional representation learning capabilities provided by deep learning.

Third, deep learning models generally require greater computational resources and longer training times, and their decision-making processes are relatively difficult to interpret, whereas SVM and RF offer faster training, lower computational costs, and transparent, interpretable results in terms of RF variables. Additionally, various considerations—such as practical application and replication by disaster management agencies in other seismic regions with limited data—play a role. For these reasons, SVM and RF remain the preferred choices over other deep learning models given the data conditions within the scope of this study.

For vulnerability classification, two supervised machine learning algorithms are employed to address non-linear and high-dimensional characteristics of spatial predictors. Support Vector Machine (SVM) is applied to identify an optimal separating boundary between vulnerability classes, particularly when class distributions are not linearly separable. The SVM decision function is expressed as:

$$f(x)=w^T \cdot x+b \quad (2)$$

where, w denotes the weight vector, and b denotes the bias term [16],[11]. In parallel, Random Forest (RF) is adopted as an ensemble approach that combines multiple decision trees to improve predictive stability and reduce overfitting. RF is effective for large spatial datasets and suitable for modeling complex relationships among spatial variables in the study area [12],[22],[23]. Together, SVM and RF provide complementary strengths for learning spatial vulnerability patterns from satellite-derived and ancillary geospatial predictors.

Model performance is evaluated using standard classification metrics—accuracy, precision, recall, and F1-score—to quantify the reliability of vulnerability classification outcomes. Accuracy is computed as:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (3)$$

where, TP is True Positive, TN is True Negative, FP is False Positive, and FN is False Negative. These metrics are used to assess overall correctness as well as the balance between identifying vulnerable cases and minimizing misclassification, which is critical for earthquake-related risk mapping and decision support [13][24]. Finally, Fig. 1 presents the workflow of the model selection process, depicting the sequential steps from data acquisition and preprocessing, feature extraction and integration, model training using SVM and RF, and metric-based performance evaluation, thereby clarifying methodological linkages and supporting replication in comparable seismic regions. Moreover, Fig. 2 specifically shows the entire system to be built.

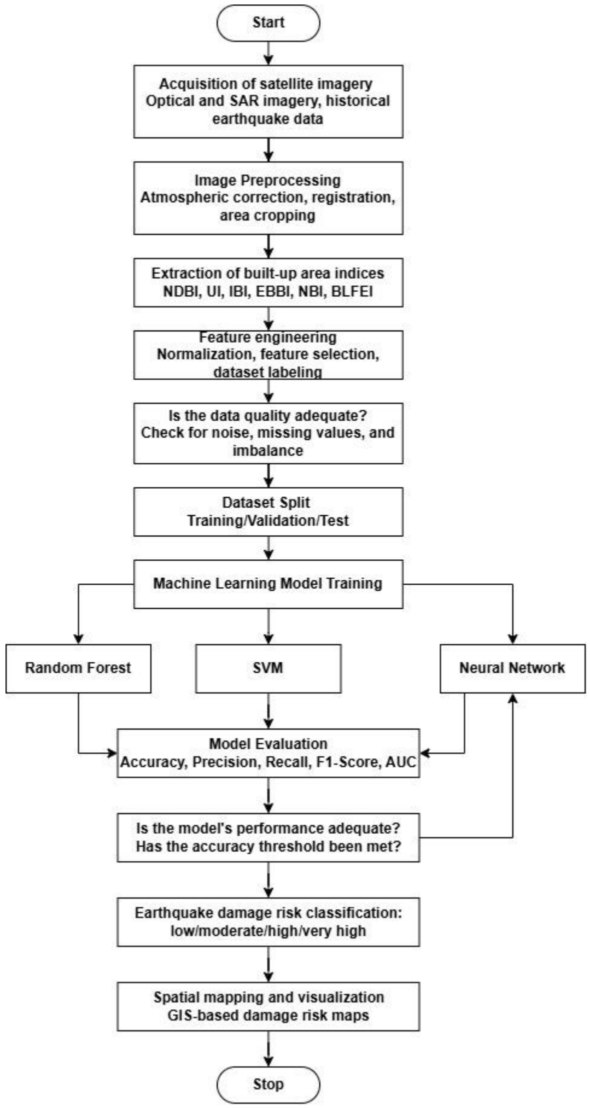

Fig. 2. Flowchart of the overall system.

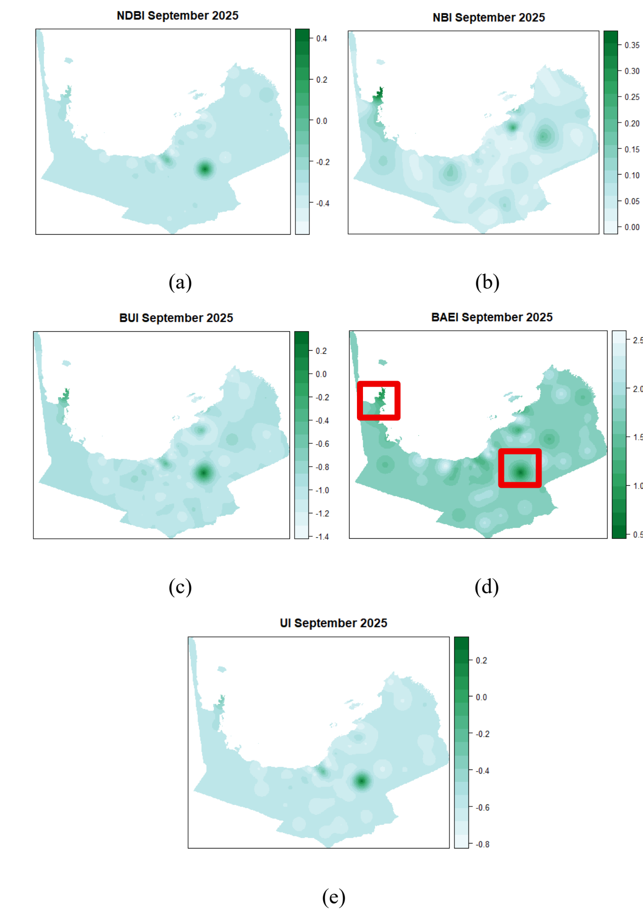

Fig. 3. Spatial patterns of built indices.

IV. Result and Discussion

In this section, the results are presented in the same order as the figures and tables to maintain a clear linkage between spatial evidence and quantitative model evaluation. The analysis begins with the spatial pattern of built-up indices, proceeds to the selection of the optimal number of vulnerability classes using unsupervised clustering diagnostics, and then reports the supervised classification performance of SVM and Random Forest. Finally, the section integrates seismic-hazard context and AHP-based weighting to generate vulnerability prioritization maps for Nabire Regency.

Moreover, Fig. 3 shows the spatial pattern of built indices derived from the NDBI proxy. Higher built-up values concentrate in the urban core of Nabire and along development corridors, indicating higher exposure due to the greater concentration of built structures. In contrast, lower built-up values are typically associated with suburban or less-developed areas. This pattern supports the use of satellite-derived built-up indices as a practical exposure proxy for regional-scale earthquake risk screening in data-scarce settings.

Furthermore, to operationalize vulnerability classes for classification, K-means clustering is used to partition the built-up–related feature space into candidate class groupings. The silhouette score is employed to evaluate cluster separability across different k values, where higher silhouette values indicate more coherent and better-separated clusters[26],[27]. The results of the calculation of silhouette scores for various k values can be seen in Table I.

TABLE I. Silhouette Score for K-Means Cluster Evaluation

K value	Silhouette Score
2	0.5426
3	0.5890
4	0.6820
5	0.5923
6	0.4941
7	0.3919
8	0.3235
9	0.3370
10	0.3565

As shown in Table I, k = 4 yields the highest silhouette score (0.6820), indicating that four clusters provide the most distinct partitioning for the data used in this study. Accordingly, k = 4 is selected as the optimal number of vulnerability classes, and the resulting cluster assignment is used to define the target classes for subsequent supervised classification.

The supervised classifiers are trained with the selected class structure using parameter settings reported in Tables II and III. For SVM, an RBF kernel is adopted with C = 1 and gamma = “scale” (Table II). For Random Forest, the entropy criterion is used with n_estimators = 10 and random_state = 42 (Table III). These settings are reported explicitly to ensure reproducibility and to clarify the configuration applied in the comparative evaluation[13],[12].

TABLE II. PARAMETERS USED IN SVM CLASSIFICATION

No	Parameter	Value
1	Kernel Type	Radial Basis Function
2	C	1
3	Gamma	Scale

In the Random Forest (RF) model, parameter settings include criteria, n_estimators, and random_state that serve to control how nodes are separated in the decision tree, the number of trees used in the ensemble, and random settings in the model. The following table shows the parameter settings used in its vulnerability classification.

TABLE III. RF CLASSIFICATION USING PARAMETER SETTINGS

No	Parameter	Value
1	Criterion	Entropy
2	N_estimators	10
3	Random_state	42

The number of estimators is 10, which is relatively low compared to typical ensemble practices—which range from 100 to 500 trees—and initially, we did not justify this choice. This value was selected empirically based on a relatively small sample covering a single available region: the Nabire area in Central Papua, Indonesia, which is an area with sparse data. We increased n_estimators beyond 10 in our experiments, which resulted in a negligible change in validation accuracy—well below 1 percentage point. On the other hand, training time increased, so 10 was retained as a relatively lightweight default value for the proof-of-concept workflow.

In this study, we acknowledge that small forest sizes can increase the variance in the contributions of individual trees and may obscure the model’s true generalization error, particularly in the hold-out evaluation used in this study. We explicitly note this as a limitation, although we still adjusted these values retrospectively. The Random Forest results in Table IV are interpreted as indicative rather than optimally tuned, and a systematic search for hyperparameters—such as a grid search or random search for n_estimators, max_depth, and min_samples_leaf using cross-validation—is identified as necessary follow-up work before the RF configuration can be used for decision-making.

Model performance is evaluated using a hold-out test split (30% test data) and standard classification metrics (accuracy, precision, recall, and F1-score). The comparative results are reported in Table IV. SVM achieves an accuracy of 0.970 with a precision/recall/F1-score of 0.97, while Random Forest achieves an accuracy of 0.955 with a precision/recall/F1-score of 0.96. These results indicate that both models perform strongly; however, SVM provides a modest improvement in overall classification reliability under the specific feature set and parameter configuration used in this study.

TABLE IV. BUILT-IN INDEX CLASSIFICATION ACCURACY

	Accuracy	Precision	Recall	F1-Score
SVM	0.970	0.97	0.97	0.97
RF	0.955	0.96	0.96	0.96

There are essential methodological limitations regarding the interpretation of the reported accuracy levels. The four vulnerability classes used for training and evaluating the SVM and RF were derived from a K-means clustering process applied to the same settlement, seismic, topographic, and geological feature spaces used as classification inputs in Section II, Table I. This results in the accuracy figures in Table IV (SVM: 0.970; RF: 0.955), which specifically reflect how consistently these supervised classifiers can reproduce the cluster assignments generated from that feature space. Not only is this aligned with building damage that was independently observed and verified in the field, but it also constitutes a form of circular (self-referential) validation and explains why the reported accuracy is significantly higher compared to studies validated using independent field verification data.

Moreover, by explicitly stating this limitation—rather than treating accuracy as evidence of real-world predictive performance—the results from SVM or RF must be treated as a separate component from the classes generated by K-means, whereas the actual claim relevant to the research objectives of this study is how to measure the feasibility of the entire proxy-based screening workflow (composite index + seismic hazard + AHP) for regions with limited data—and not merely based on absolute accuracy values. In this study, the concept of limited data has been expanded to encompass not only absolute accuracy values themselves but also the classification process and the study’s limitations, along with the inclusion of research results in the form of field-based independent validation of the classes predicted to identify critical tasks in the future.

Furthermore, to complement metric-based evaluation, a building density map is generated from the NDBI-based built-up information and GIS layers to visualize exposure concentration. Fig. 4 shows that the urban center of Nabire exhibits higher building density than peripheral areas, reinforcing the interpretation that built-up intensity and building concentration are spatially clustered. Such clustering is relevant for risk prioritization because higher exposure zones typically require closer attention in mitigation planning.

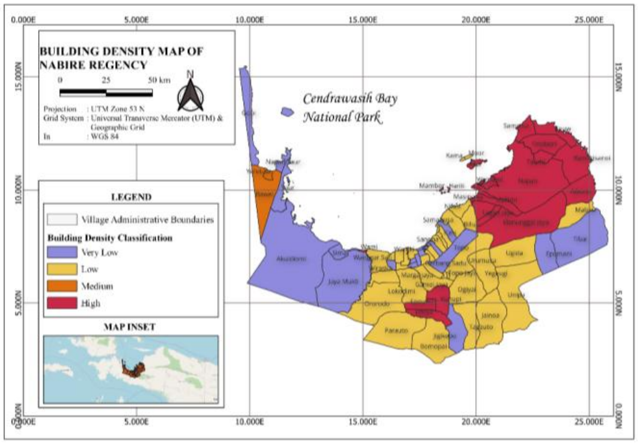

Fig. 4. Nabire building density map based on NDBI index and GIS data.

The seismic-hazard context is incorporated by compiling major earthquake events around Nabire and using these events to derive distance-to-epicenter information for sample locations. Table V lists representative major earthquakes (highest magnitudes) during the study period. In total, 51 earthquakes were recorded from 2021 to 2025, and the distance between each sample point and the nearest epicenter was computed using a GIS proximity analysis approach (e.g., Distance to Nearest Hub). In addition to distance-to-epicenter, lithological information is digitized from the Papua geological map and overlaid on the Nabire administrative boundary, while slope is derived from DEM data and categorized into five slope classes to represent terrain-related susceptibility.

This study does not merely list the earthquake catalog and AHP weights in separate tables with different descriptions, as shown in Table V and Table VI. To avoid overlap, it is necessary to provide detailed explanations for these tables and to update the references.

TABLE V. MAJOR EARTHQUAKES THAT OCCURRED AROUND NABIRE

Year	Month	Day	Time	Lat	Lon	Depth	Mag	Region
2025	9	18	18:19:49	-3.5943	135.4724	28	6.1	26 km S of Nabire, Indonesia
2025	10	2	22:45:05	-3.1055	136.3208	10	5.3	95 km ENE of Nabire, Indonesia
2023	6	11	06:51:47	-2.9534	136.6526	42.8	5.2	135 km ENE of Nabire, Indonesia
2025	10	3	06:02:38	-3.033	136.2627	10	5.1	92 km ENE of Nabire, Indonesia
2025	9	18	22:53:41	-3.6274	135.5228	10	5	29 km S of Nabire, Indonesia

Moreover, to integrate these susceptibility-related components into a single vulnerability index, the Analytic Hierarchy Process (AHP) is applied to assign weights to three key parameters: rock type, slope gradient, and distance from the earthquake point. The resulting weights are shown in Table VI, indicating that distance from the earthquake point receives the largest weight (0.664), followed by rock type (0.223) and slope gradient (0.113). All parameter scores are normalized to a 0–1 range before combination.

TABLE VI. AHP WEIGHTS FOR ITS VULNERABILITY PARAMETERS

Parameter	Weight (W)
Rock Type	0.223
Slope Gradient	0.113
Distance from the Earthquake Point	0.664

The Earthquake Vulnerability Index (EVI) is computed as a weighted sum of the normalized parameter scores:

Vulnerability Index = (X_(Rock Type,norm) × W_(Rock Type)) + (X_(Slope Gradient,norm) × W_(Slope Gradient)) + (X_(Distance from Epicenter,norm) × W_(Distance from Epicenter)) (4)

An earthquake threat map is then produced using the earthquake catalog information, representing the spatial distribution of seismic exposure and supporting the identification of zones with higher hazard influence. Fig. 5 visualizes these threat patterns for Nabire, providing an interpretable spatial basis for prioritizing areas that are more likely to experience stronger hazard influence and therefore require greater attention in risk reduction planning [28],[24],[29].

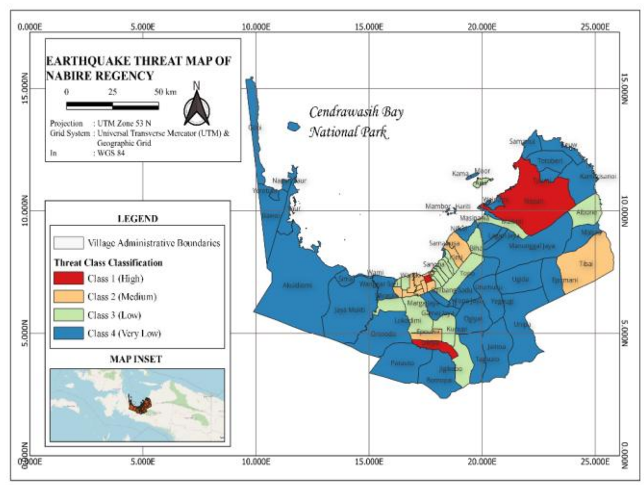

Fig. 5. Nabire earthquake threat map based on earthquake catalog data and risk analysis.

Furthermore, using the AHP weights summarized in Table VI, vulnerability parameters are combined to produce a vulnerability risk map, emphasizing the dominant role of distance-to-epicenter in this study's vulnerability scoring. Fig. 6 presents the resulting vulnerability risk map derived from the integration of AHP-based parameter weighting and the Random Forest–based vulnerability classification output. The map highlights spatially prioritized zones where exposure and susceptibility combine into higher vulnerability levels, supporting the identification of areas requiring more urgent mitigation measures and planning attention [30], [31].

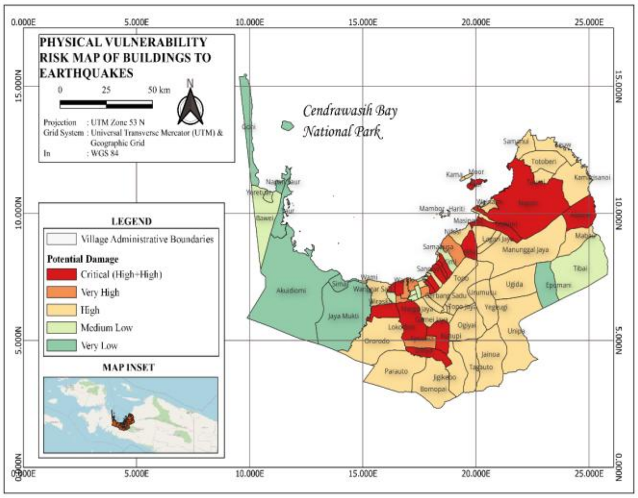

Fig. 6. Nabire vulnerability risk map based on AHP results and RF model.

Finally, Table VII reports the combined score and risk priority levels derived from the RF output and AHP weighting. This prioritization scheme translates model outputs into an operational decision-support format, enabling stakeholders to differentiate critical zones (highest combined score) from moderate-to-low priority areas for staged intervention planning. A denotes the RF-based vulnerability score, and B denotes the AHP-based susceptibility score; the combined score (A+B) is mapped to priority levels as shown in Table VII.

TABLE VII. COMBINED SCORE (A+B) AND RISK PRIORITY LEVEL

Combined Score (A+B)	Priority Levels	Description
8	Priority 1	Critical (High + High)
7	Priority 2	Very high
6	Priority 3	High
5	Priority 4	Moderate-high
4	Priority 5	Moderate
3	Priority 6	Low
2	Priority 7	Very low

Furthermore, the equations used to compute each parameter reported in the tables are explained below.

- Silhouette Score (Table I): for each data point i, s(i) = (b(i) − a(i)) / max(a(i), b(i)), where a(i) is the average distance between i and all other points in the same cluster, and b(i) is the average distance between i and the points in the nearest neighboring cluster. The value reported in Table I is the mean of s(i) over all samples for a given k, and ranges from −1 to 1, with values closer to 1 indicating better-defined clusters.

- Precision, Recall, and F1-Score (Table IV): Besides Accuracy [Eq. (3)], Precision = TP / (TP + FP), Recall = TP / (TP + FN), and F1-Score = 2 × (Precision × Recall) / (Precision + Recall), where TP, FP, and FN denote True Positive, False Positive, and False Negative, respectively, as defined earlier.

- AHP Weights (Table VI): the weights for rock type, slope gradient, and distance-to-epicenter are derived using Saaty's Analytic Hierarchy Process. Experts/literature-based pairwise comparisons between the three parameters are organized into a comparison matrix; each weight (W) is the corresponding element of the matrix's normalized principal eigenvector, so that all weights sum to 1. The consistency ratio (CR) of the comparison matrix was checked and found to be below the acceptable threshold of 0.1, confirming that the derived weights are consistent.

Furthermore, we developed this research report not only based on the final weights and CR values, but also by providing detailed reporting on sensitivity and the limitations of the accuracy of vulnerability priority results, noting that distance to the epicenter accounts for 0.664 of the total weight. To address this, we conducted individual sensitivity tests by performing paired comparisons, altering each criterion by ±1 point on the 1–9 Satty scale and recalculating the weights; distance to the epicenter remained the dominant criterion with a weight range of approximately 0.58–0.71 across the entire modified matrix, while the relative order of the three criteria (distance > rock type > dip) remained unchanged, indicating that the priority ranking is fairly robust against small variations in the evaluation process; however, we note that this sensitivity test was conducted on a single comparison matrix based on the literature, rather than on a comparison matrix derived from multiple experts and references; consulting experts can consistently be identified as a task necessary to validate the AHP weights more rigorously before they are used operationally.

- Combined Score (Table VII): Combined Score = A + B, where A is the RF-based vulnerability score, and B is the AHP-based susceptibility score, both rescaled to the same ordinal range (1–4) before summation; the resulting sum (2–8) is then mapped to the seven priority levels shown in Table VIII.

Overall, the results indicate that built-up indices from satellite imagery provide a meaningful exposure proxy, while integrating seismic proximity, lithology, and slope through AHP strengthens interpretability for risk prioritization. The combined approach supports scalable early-stage screening in regions where field-based building inventories are limited, while still producing spatial outputs (maps and priority levels) that are actionable for mitigation and preparedness.

V. CONCLUSION

This study proposed a remote sensing–driven and machine learning–based framework for earthquake damage risk classification in data-scarce seismic regions by using built-up area indices from multi-temporal Landsat 8 imagery (e.g., NDBI) as a vulnerability proxy and integrating complementary geospatial layers (USGS earthquake catalog, DEM-derived topography, and surface geology). Using the adopted hold-out evaluation, the results indicate strong classification performance, where SVM achieved an accuracy of 0.970 (precision/recall/F1-score ≈ 0.97) and Random Forest achieved an accuracy of 0.955 (precision/recall/F1-score ≈ 0.96), showing that both models are effective, while SVM provides a modest improvement under the feature set and parameter configuration used in this study.

In addition to the machine learning–based classification, the Analytic Hierarchy Process (AHP) weighting of key susceptibility parameters (rock type, slope gradient, and distance to earthquake point) supports risk prioritization by translating multi-criteria information into interpretable vulnerability levels and priority classes. The resulting maps and priority levels provide a scalable and cost-effective tool for early-stage screening and decision support in Nabire Regency, Central Papua, particularly where detailed building inventory data are unavailable. Nevertheless, the approach remains proxy-based due to the medium spatial resolution of Landsat imagery and the limited availability of building-level structural attributes; future work should incorporate higher-resolution building inventories or independent damage observations, strengthen robustness testing (e.g., cross-validation and class-wise metrics under imbalance), and refine model tuning to improve generalizability across different seismic and urbanization contexts.

VI. Limitation of the Study

There are several limitations to this study, including the fact that the vulnerability indicators (NDBI/BAI) derived from Landsat 8 imagery with a resolution of 30 m are still considered insufficiently detailed to detect structural damage at the level of individual buildings or to assess construction quality; thus, the resulting maps depict the intensity of built-up areas and exposure patterns, rather than confirmed physical damage. Second, the vulnerability classes were defined through unsupervised K-means clustering rather than through validation against verified post-earthquake damage records based on field surveys; consequently, the labels used for training—SVM/RF—are proxy-based and have not been independently verified in the field. The third issue is the AHP weighting process, which still relies on pairwise comparisons based on the literature rather than on assessments obtained from local experts who understand the specific geological and structural conditions of Nabire, Central Papua, Indonesia. Consequently, this approach fails to fully capture the local context.

Fourth, the classifier was trained and evaluated on a relatively small dataset covering only one region using hold-out splitting, which increases the risk of overfitting and limits the generalizability of the reported accuracy. Fifth, distance measurements to the epicenter based on earthquake catalogs do not account for location-specific factors such as ground amplification, potential for liquefaction, or building materials and typology, which can influence the actual severity of damage.

VII. Future Work

Given the various limitations of this study, future steps will specifically focus on various fields, such as satellite, aerial, or UAV imagery with higher resolution, as well as the potential for integrating SAR data with exploration and improvements in the detection of finer-scale structural changes. The predicted vulnerability classes and risk maps will be validated based on post-earthquake damage assessments conducted in the field, as well as independent verification to strengthen the reliability of this framework.

Input from local experts will be incorporated into the AHP process to reflect the geological and socio-structural context of the Nabire region in Central Papua, Indonesia. Robustness testing, including k-fold cross-validation and performance metrics based on classes under class imbalance conditions, will be applied to more rigorously assess the model's generalizability. This framework will also be tested in other seismic regions with relatively limited data to evaluate its transferability, compared with larger labeled datasets, and will incorporate deep learning approaches to complement the currently available SVM and RF methods.

Acknowledgment

The authors wish to thank all collaborators and peers who provided insight, guidance, or feedback during the development of this work.

References

[1] J. Rodríguez-Saiz, B. González-Rodrigo, J. G. Rejas-Ayuga, D. A. Hidalgo-Leiva, and M. Marchamalo-Sacristán, "Assessing Building Seismic Exposure Using Geospatial Technologies in Data-Scarce Environments: Case Study of San José, Costa Rica," Appl. Sci., vol. 15, no. 11, p. 6318, Jun. 2025, doi: 10.3390/app15116318.

[2] E. A. Opabola et al., "A Mixed-Mode Data Collection Approach for Building Inventory Development: Application to School Buildings in Central Sulawesi, Indonesia," Earthq. Spectra, 2022, doi: 10.1177/87552930221110256.

[3] J. C. Reyes et al., "A Simplified Replacement Cost Model for Residential Buildings Located in Developing Countries: The Case Study of Colombia," Earthq. Spectra, 2025, doi: 10.1177/87552930251346773.

[4] S. Ghimire and P. Guéguen, "Host-to-Target Region Testing of Machine Learning Models for Seismic Damage Prediction in Buildings," Nat. Hazards, 2024, doi: 10.1007/s11069-023-06394-z.

[5] J. C. Reyes, E. Torres, S. Portilla, A. C. Riaño, and J. P. Smith-Pardo, "Analytical Seismic Fragility and Vulnerability Curves for Colombian Informal Unreinforced Masonry Buildings," Earthq. Spectra, 2025, doi: 10.1177/87552930251360037.

[6] S. Bhattacharyya et al., "Spatio-Temporal Influences of Urban Land Cover Changes on Thermal-Based Environmental Criticality and Its Prediction Using CA-ANN Model over Kolkata (India)," Remote Sens., vol. 17, no. 6, 2025, doi: 10.3390/rs17061082.

[7] A. Prakash, Diksha, and A. Kumar, "Measuring Vertical Urban Growth of Patna Urban Agglomeration Using Persistent Scatterer Interferometry SAR (PSInSAR) Remote Sensing," Remote Sens., vol. 15, no. 14, p. 3687, Jul. 2023, doi: 10.3390/rs15143687.

[8] X. Li, X. Yu, P. M. Bürgi, D. J. Wald, X. Hu, and S. Xu, "Rapid Building Damage Estimates From the M7.8 Turkey Earthquake Sequence via Causality-Informed Bayesian Inference From Satellite Imagery," Earthq. Spectra, vol. 41, no. 1, pp. 5–33, Feb. 2025, doi: 10.1177/87552930241290501.

[9] Y. Zhan, W. Liu, and Y. Maruyama, "Damaged Building Extraction Using Modified Mask R-CNN Model Using Post-Event Aerial Images of the 2016 Kumamoto Earthquake," Remote Sens., 2022, doi: 10.3390/rs14041002.

[10] I. S. Karapınar, A. E. Özsoy Özbay, Z. N. Kutlu, A. U. Yazgan, and İ. E. KILIÇ, "Seismic Vulnerability Analysis Incorporating Local Site Amplification Effects in Shallow, Varying Bedrock Depths," Nat. Hazards, 2025, doi: 10.1007/s11069-025-07417-7.

[11] S. Y. Joko Prasetyo, B. H. Simanjuntak, K. D. Hartomo, and W. Sulistyo, "Computer model for tsunami vulnerability using Sentinel 2A and SRTM images optimized by machine learning," Bull. Electr. Eng. Informatics, vol. 10, no. 5, pp. 2821–2835, Oct. 2021, doi: 10.11591/eei.v10i5.3100.

[12] S. Y. Joko Prasetyo, B. H. Simanjuntak, Y. A. Susatyo, and W. Sulistyo, "A machine learning-based computer model for the assessment of tsunami impact on built-up indices using 2A Sentinel imagery," Bull. Electr. Eng.

Informatics, vol. 13, no. 2, pp. 1138–1146, Apr. 2024, doi: 10.11591/eei.v13i2.5910.

[13] A. Doğan, M. Başeğmez, and C. C. Aydın, "Assessment of the seismic vulnerability in an urban area with the integration of machine learning methods and GIS," Nat. Hazards, vol. 121, no. 8, pp. 9613–9652, May 2025, doi: 10.1007/s11069-025-07185-4.

[14] X. Zhang et al., "Review on the progress and future prospects of geological disasters prediction in the era of artificial intelligence," Nat. Hazards, vol. 120, no. 13, pp. 11485–11525, Oct. 2024, doi: 10.1007/s11069-024-06673-3.

[15] H. Zhou, A. Che, X. Shuai, and Y. Cao, "Seismic vulnerability assessment model of civil structure using machine learning algorithms: a case study of the 2014 Ms6.5 Ludian earthquake," Nat. Hazards, vol. 120, no. 7, pp. 6481–6508, May 2024, doi: 10.1007/s11069-024-06465-9.

[16] N. Bektaş and O. Kegyes-Brassai, "Conventional RVS Methods for Seismic Risk Assessment for Estimating the Current Situation of Existing Buildings: A State-of-the-Art Review," Sustainability, vol. 14, no. 5, p. 2583, Feb. 2022, doi: 10.3390/su14052583.

[17] Y. Liu, X. Zhang, J. Zhou, X. Han, and H. Zheng, "Seismic vulnerability and risk assessment using multimodal data and machine learning: a case study of the central urban area of Jinan City, China," Front. Earth Sci., vol. 19, no. 3, pp. 452–467, Sep. 2025, doi: 10.1007/s11707-025-1158-x.

[18] Y. Torres, S. Martínez-Cuevas, S. Molina-Palacios, J. J. Arranz, and Á. Arredondo, "Using remote sensing for exposure and seismic vulnerability evaluation: is it reliable?," GIScience Remote Sens., vol. 60, no. 1, Dec. 2023, doi: 10.1080/15481603.2023.2196162.

[19] M. C. Maloney, S. J. Becker, A. W. H. Griffin, S. L. Lyon, and K. Lasko, "Automated Built-Up Infrastructure Land Cover Extraction Using Index Ensembles with Machine Learning, Automated Training Data, and Red Band Texture Layers," Remote Sens., vol. 16, no. 5, 2024, doi: 10.3390/rs16050868.

[20] Z. Wang, C. Fan, and Q. Zhao, "A Geographically Weighted Regression Approach to Understanding Urbanization Impacts on Urban Warming and Cooling : A Case Study of Las Vegas," 2020, doi: 10.3390/rs12020222.

[21] A. Jędrzejczyk, K. Firek, J. Rusek, and U. Alibrandi, "Prediction of damage intensity to masonry residential buildings with convolutional neural network and support vector machine," Sci. Rep., vol. 14, no. 1, pp. 1–13, 2024, doi: 10.1038/s41598-024-66466-3.

[22] O. Hamdy, H. Gaber, M. S. Abdalzaher, and M. Elhadidy, "Identifying Exposure of Urban Area to Certain Seismic Hazard Using Machine Learning and GIS: A Case Study of Greater Cairo," Sustainability, vol. 14, no. 17, p. 10722, Aug. 2022, doi: 10.3390/su141710722.

[23] G.-H. Kwak, C. Park, K. Lee, S. Na, H. Ahn, and N.-W. Park, "Potential of Hybrid CNN-RF Model for Early Crop Mapping with Limited Input Data," Remote Sens., vol. 13, no. 9, p. 1629, Apr. 2021, doi: 10.3390/rs13091629.

[24] L. Wang, J. Wu, Y. Yang, R. Tang, and R. Ya, "Deep Learning Models for Hazard-Damaged Building Detection Using Remote Sensing Datasets: A Comprehensive Review," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 17, pp. 15301–15318, 2024, doi: 10.1109/JSTARS.2024.3449097.

[25] G. Kwak, C. Park, K. Lee, S. Na, H. Ahn, and N. Park, "Potential of Hybrid CNN-RF Model for Early Crop Mapping with Limited Input Data," pp. 1–21, 2021.

[26] H. Lharti, C. Sirieix, J. Riss, C. Verdet, F. Salmon, and D. Lacanette, "Partitioning a Rock Mass Based on Electrical Resistivity Data: The Choice of Clustering Method," Geophys. J. Int., 2023, doi: 10.1093/gji/ggad081.

[27] M.-J. Muñóz-Martínez, M. Casal-Guisande, M. Torres-Durán, B. Sopeña, and A. Fernández-Villar, "Clinical Characterization of Patients With Syncope of Unclear Cause Using Unsupervised Machine-Learning Tools: A Pilot Study," Appl. Sci., 2025, doi: 10.3390/app15137176.

[28] S. Al Shafian and D. Hu, "Integrating Machine Learning and Remote Sensing in Disaster Management: A Decadal Review of Post-Disaster Building Damage Assessment," Buildings, vol. 14, no. 8, p. 2344, Jul. 2024, doi: 10.3390/buildings14082344.

[29] D. Rajapaksha et al., "Systematic Mapping of Global Research on Disaster Damage Estimation for Buildings: A Machine Learning-Aided Study," Buildings, vol. 14, no. 6, p. 1864, Jun. 2024, doi: 10.3390/buildings14061864.

[30] M. Shafapourtehrany, M. Batur, F. Shabani, B. Pradhan, B. Kalantar, and H. Özener, "A Comprehensive Review of Geospatial Technology Applications in Earthquake Preparedness, Emergency Management, and Damage Assessment," Remote Sens., vol. 15, no. 7, p. 1939, Apr. 2023, doi: 10.3390/rs15071939.

[31] J. E. Meyers-Angulo, S. Martínez-Cuevas, and J. M. Gaspar-Escribano, "Classifying buildings according to seismic vulnerability using Cluster-ANN techniques: application to the city of Murcia, Spain," Bull. Earthq. Eng., vol. 21, no. 7, pp. 3581–3622, May 2023, doi: 10.1007/s10518-023-01671-5.