Canadian Conference - Pacific Conference on Earthquake Engineering 2023

Vancouver, British Columbia

June 25th – June 30th, 2023

Numerical Validation of a Seismic Rapid Visual Screening Method

Niloofar Elyasi[1*], Wilson Carofilis[1], Eugene Kim[2] and Chul Min Yeum[2]

[1]PhD Student, Department of Civil Engineering, University of Waterloo, Waterloo, ON, Canada
[2]Assistant Professor, Department of Civil Engineering, University of Waterloo, Waterloo, ON, Canada
[*]n2elyasi@uwaterloo.ca (Corresponding Author)

ABSTRACT

Seismic vulnerability assessments of existing buildings are essential to proactively ensure occupant safety and minimize losses of earthquake events. Rapid visual screening (RVS) procedures have been introduced to quickly evaluate the vulnerability of existing buildings and prioritize actions. In a previous study, the authors proposed a new RVS method for low-rise RC buildings using machine learning (ML) techniques. A random forest (RF) classification algorithm was utilized to predict the earthquake damage based on seven basic geometric parameters (e.g., number of stories, floor area, column size, and wall cross-sectional areas) and two earthquake intensity measures, modified Mercalli intensity (MMI) and peak ground acceleration (PGA). Data from six past earthquake surveys were used to train and test the classification models. The ML-based RVS method showed high accuracy in predicting the damage state. However, the validation was conducted on a small set of survey data, which may lead to overfitting and reducing its reliability. In this study, a numerical simulation is conducted to validate the proposed RVS method. A numerical model of a low-rise RC building located in Ecuador is developed to represent the buildings in the databases used to train the ML models. Uncertainties and structural details such as poor concrete quality, insufficient reinforcement and confinement are also considered in the numerical model. Due to unavailability of the ground motion records for the six earthquake events in the database, similar earthquake records are selected to simulate them. The probable damage state of the numerical model hit by each earthquake scenario is determined by both the RF classification model and the nonlinear analysis. The results indicate high accuracy and applicability of the proposed ML-based RVS method as a preliminary risk reduction technique.

Keywords: Rapid visual screening, machine learning, damage state, low-rise RC building, numerical model

INTRODUCTION

There are numerous structures all over the world that may experience severe damage or collapse in future earthquake events. Rapid Visual Screening (RVS) is a quick and straightforward procedure to identify seismic vulnerable buildings based on simple visual observations recorded during sidewalk surveys [1]. Many RVS methods have been proposed in the past few decades. FEMA 154 was the first sidewalk RVS approach introduced in 1988 to facilitate rapid vulnerability assessments [2]. This method has been revised twice in 2002 and 2015. The most recent version of FEMA 154 uses information such as material type, number of stories, seismic force resisting system type, soil type, and regional seismicity to calculate a basic score [3]. Buildings with a basic score below a cut-off score (2 is recommended in the guideline) are considered as potential risk and prioritized for further evaluation or mitigation strategies. Several FEMA-based screening approaches have been developed in other countries according to their local seismic circumstances. EMS-98 and RISK-UE project introduced by The European union, OASP-0 organized by the earthquake planning and protection organization of Greece, and the structural scoring system proposed by National Research Council Canada (NRCC) are examples of the FEMA-based RVS methods [4-8]. Besides these scoring approaches, several empirical methods have been derived by researchers like Tesfamariam and Saatcioglu [9], Ningthoujam and Nada [10], and Coskun et.al [11]. However, one of the easiest and most practical empirical RVS so far, was proposed by Hassan and Sozen in 1997 [12]. In their study, a Priority Index (PI) was defined using the geometric features of low-rise reinforced concrete (RC) buildings including number of stories, floor area, column area, and wall area. The PI was then used to predict the severity of earthquake damage. The simplicity of the PI has made it a popular data-collecting approach in post-earthquake inspections all around the world. In fact, a large Hassan-Sozen PI building inventory exists on data sharing platforms such as Datacenterhub. The wide availability of the data and advances in machine learning (ML) motivated the

authors to improve the traditional PI method and introduce more widely applicable RVS. In a previous study, the PI data from six earthquake events (Duzce (1999) [13], Bingol (2003) [13], Nepal (2015) [14], Taiwan (2016) [15], Ecuador (2016) [16], and Pohang (2017) [17]) were first used to evaluate the Hassan-Sozen assessment method. Then, the method was improved using machine learning algorithms including logistic regression and Random Forest (RF). Finally, a new ML-based RVS model was proposed that incorporated two earthquake intensity features, modified Mercalli intensity (MMI) and peak ground acceleration (PGA), as extra input features in addition to the PI geometric features. The results demonstrated the capability of the proposed RVS method to accurately identify buildings with a high likelihood of being severely damaged. Although the method showed high accuracy in predicting the damage level, the validation was only conducted using the surveyed data.

In this study, first, a summary of the RVS method proposed in the previous study is described since the relevant paper is still not publicly accessible. A numerical model of a real low-rise RC building from Ecuador is then created based on a case study investigated after the 2016 Ecuador earthquake. Ten different earthquake records are considered to simulate all the six earthquake events whose data were used to develop the ML-based RVS model. The damage states of the building model hit by the earthquake records are predicted using the RF model. These results are then compared to the numerical results obtained from the nonlinear response history analysis (NRHA) of the building model. The outcomes from the comparison numerically validate the accuracy and ability of the proposed non-region-specific RVS model to rapidly identify high-risk low-rise RC buildings.

METHODLOGY

Hassan and Sozen PI

The PI developed by Hassan and Sozen (1997) [12] is computed as the sum of Wall Index (WI) and Column Index (CI). These two metrics are obtained from seven geometric parameters of low-rise RC buildings in total: number of stories, floor area, column cross-sectional areas, and areas of concrete walls and masonry infill walls in two horizontal directions, east-west (E-W) and north-south (N-S). WI and CI are computed as follows:

$$\mathbf{WI} = \left[(\boldsymbol{A_{cw}} + \boldsymbol{A_{mw}}/\mathbf{10})/\boldsymbol{A_{ft}}\right] \times \mathbf{100} \tag{1}$$

$$\mathbf{CI} = \left[(\boldsymbol{A_{col}}/\mathbf{2})/\boldsymbol{A_{ft}}\right] \times \mathbf{100} \tag{2}$$

Where A_{cw} and A_{mw} are the total cross-sectional areas of RC walls and unreinforced masonry infill walls in a given horizontal direction at the base, respectively. A_{ft} is the total building floor area above the base and A_{col} is the total cross-sectional area of columns at the base. The PI in calculated for both horizontal directions and the minimum value is taken as the index for the building. The lower the PI is, the more vulnerable the building would be in future seismic events. To define damage severity boundaries, Hassan and Sozen proposed a graphical procedure in which the values of WI (y-axis) are plotted against the CI (x-axis). The example figure produced by Hassan and Sozen (1997) for the 1992 Erzincan earthquake is shown in Figure 1.

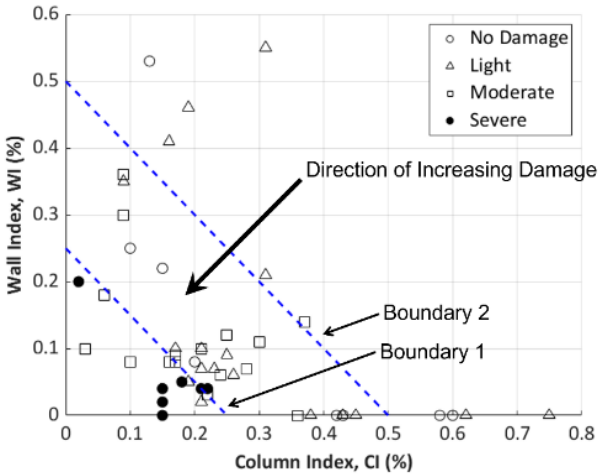

Figure 1. Graphical seismic vulnerability evaluation method proposed by Hassan and Sozen (1997).

As can be seen, buildings which were severely damaged in the Erzincan earthquake were closer to the origin (i.e., PI → 0) and fell under or close to Boundary 1. Despite the simplicity of the approach, the damage boundaries in Figure 1 are specific to the

data from Erzincan earthquake and for any other region and earthquake event, new boundaries must be manually defined based on the available data.

The post-earthquake PI inventories of 658 low-rise RC buildings collected after six earthquake events including Duzce (1999) [13], Bingol (2003) [13], Nepal (2015) [14], Taiwan (2016) [15], Ecuador (2016) [16], and Pohang (2017) [17] were used in the previous study to develop the ML models. The data is obtained from Datacenterhub, an open data sharing platform developed by Purdue University and University of Nebraska (www.datacenterhub.org). A summary of the six earthquake datasets is presented in Table 1.

Table 1. Characteristics of earthquake datasets.

Earthquake	Moment Magnitude (M_w)	Sample Size
Duzce, Turkey (1999)	7.2	116
Bingol, Turkey (2003)	6.4	55
Nepal (2015)	7.8	135
Taiwan (2016)	6.4	106
Ecuador (2016)	7.8	172
Pohang (2017)	5.4	74

The structural damage states of the buildings from the six datasets are classified into two damage classes: non-severe (including none and light damages) and severe (including moderate and severe damages). A description of each damage class is presented in Table 2.

Table 2. Description of structural damage classes [18-19].

Damage Class	Definition
Non-Severe	Hairline cracks (< 0.25 mm) and/or flexural cracks.
Severe	Inclined cracks, spalling, loss of concrete cover, loss of concrete core, reinforcement buckling or fracture, permanent drift in structure visually apparent, shear failure, or failure of any structural element.

Machine Learning Models

In the previous study, two ML algorithms, logistic regression and RF, were used to evaluate and improve the Hassan and Sozen evaluation method as well as propose a new RVS approach. In the following sections, a summary of logistic regression and RF models are presented.

Logistic Regression

Logistic regression works based on the sigmoid function $p(\mathbf{x}) = 1 / (1 + \exp(-f(\mathbf{x}))$ where $f(\mathbf{x})$ is a linear function of the form $f(\mathbf{x}) = b_0 + b_1x_1 + \cdots + b_nx_n$ with $x_1, x_2, \ldots, x_n$ as the input features, and $b_0, b_1, \ldots, b_n$ as the coefficients or the predicted weights. The logistic regression function predicts the probability that the output for a given set of input features, $\mathbf{x}$, is equal to 1. Therefore, $1 - p(\mathbf{x})$ is the probability that the output is equal to 0. Logistic regression determines the best predicted weights $b_0, b_1, \ldots, b_n$ such that the function $p(\mathbf{x})$ is as close as possible to all actual responses, y_i, where $i = 1, \ldots, n$ and n is the number of observations. For a binary classification problem (e.g., severe damage versus non-severe damage in our study), the classification boundary is defined by the set of all points that satisfy $p(\mathbf{x}) = 0.5$ which would be given by $b_0 + b_1x_1 + \cdots + b_nx_n = 0$ [20].

To evaluate the PI method, logistic regression was first used to make a model which predicted the damage state based on PI as the only input feature. As mentioned above, the PI is region-specific so, the model was trained and tested on each of the six earthquake datasets individually using 5-fold cross validation technique. The model accuracies of 54%, 56%, 60%, 65%, 64%, and 53% were obtained respectively for the Duzce, Bingol, Nepal, Taiwan, Ecuador, and Pohang datasets. In the next step, the relative importance of the WI and CI on the damage state was evaluated by applying the Hassan and Sozen graphical approach.

Logistic regression was employed to define the best fit classification boundaries. In none of the six datasets, the slope of the classification boundary was similar to the one proposed by Hassan and Sozen in Figure 1. In other words, the weights for WI and CI were not equal as it was assumed in Eq. (1). The authors concluded that logistic regression could more accurately define the classification boundaries by customizing the weighting on CI and WI.

Random Forest

RF is an ensemble learning algorithm which is made of several decision trees. Each tree is trained using a random subset of the dataset (bootstrapping). A random subset of the input features is then considered for the nodes of every tree to split on. These steps are repeated many times to build a forest with a wide variety of trees. To make a prediction from an RF model, the prediction from every individual tree is considered, then an overall aggregated prediction is obtained. In classification problems, the final prediction is determined by a vote in which the most frequent prediction is the outcome. RF has many advantageous features such as fast training, scalability, relatively high accuracy on small datasets, and its robustness against overfitting [21].

As an improvement to the PI method, RF classification models were developed to directly use the seven raw geometric parameters instead of considering the CI and WI. The accuracy of 72%, 78%, 70%, 71%, 71%, and 60% were reached from the RF models for the Duzce, Bingol, Nepal, Taiwan, Ecuador, and Pohang earthquakes, respectively. Comparing these results with the accuracy obtained from the PI models indicated that using the raw geometric parameters on their own could be more effective to classify buildings likely to be severely damaged without the need for defining a classification boundary. However, the developed RF models were still region-specific. To develop a generally applicable RVS approach that is not region-specific, the datasets were combined and two seismic intensity measures, MMI and PGA, were introduced as additional input features to the seven raw geometric parameters. An RF model was trained and tested on the mixed dataset containing all the buildings from the six datasets. The high accuracy of 71% demonstrated that the proposed ML-based model was a simple and widely applicable RVS approach.

NUMERICAL VALIDATION

To further validate the ML-based models proposed by the authors in the previous study, numerical results from a three-dimensional (3D) model of a representative building investigated after the Ecuador earthquake (2016) are evaluated and discussed [22]. Although this building was originally hit just by the Ecuador earthquake, its 3D model can be analyzed using the records of all the six earthquake events: Duzce (1999), Bingol (2003), Nepal (2015), Taiwan (2016), Ecuador (2016), and Pohang (2017). Since not all the ground motion records for the six events are available, ten similar ground motion records are selected for NRHA. In fact, the MMI and PGA values of the selected records must be similar to the ones of the six earthquakes ranged between 6 and 9, and 0.15g and 0.89g, respectively. The ground motions records are obtained from Peer Ground Motion Database (https://ngawest2.berkeley.edu). The characteristics of the records are presented in Table 3.

Table 3. Characteristics of the selected ground motion records.

Earthquake	Record Sequence Number (RSN)	PGA (g)	MMI
Superstition Hills-02 (1987)	719	0.15	6
Superstition Hills-02 (1987)	721	0.36	6
Chuetsu-oki, Japan (2007)	5259	0.16	7
Parkfield-02 (2004)	4140	0.25	7
Georgia, USSR (1991)	818	0.65	8
Chi-Chi, Taiwan (1999)	1205	0.64	8
Northridge-01 (1994)	949	0.35	8
Christchurch, New Zealand (2011)	8064	0.48	9
Chi-Chi, Taiwan (1999)	1504	0.5	9
Kobe, Japan (1995)	1119	0.69	9

Modeling and Analysis

The Ecuadorian case study is a three-storey, two- by three-bay RC building with masonry infills and flat slabs investigated by Kagermanov et al. [22]. Columns have 20×20 cm^2 square cross sections reinforced with eight uniformly distributed 16 mm bars and 8 mm stirrups spaced at 200 mm. 20 cm thick slabs are reinforced with 16 mm bars in the slab-column region and 8 mm bars in the outside region. For both columns and slabs, a concrete cover of 15 mm is assumed. The concrete compressive strength and the steel yield strength are 28 MPa and 462 MPa, respectively. For masonry infills, three different types, corresponding to windows, doors, and full panels are considered. A general configuration of the building as well as the location of the infills is reproduced and shown in Figure 2. All dimensions displayed in the figure are in meters.

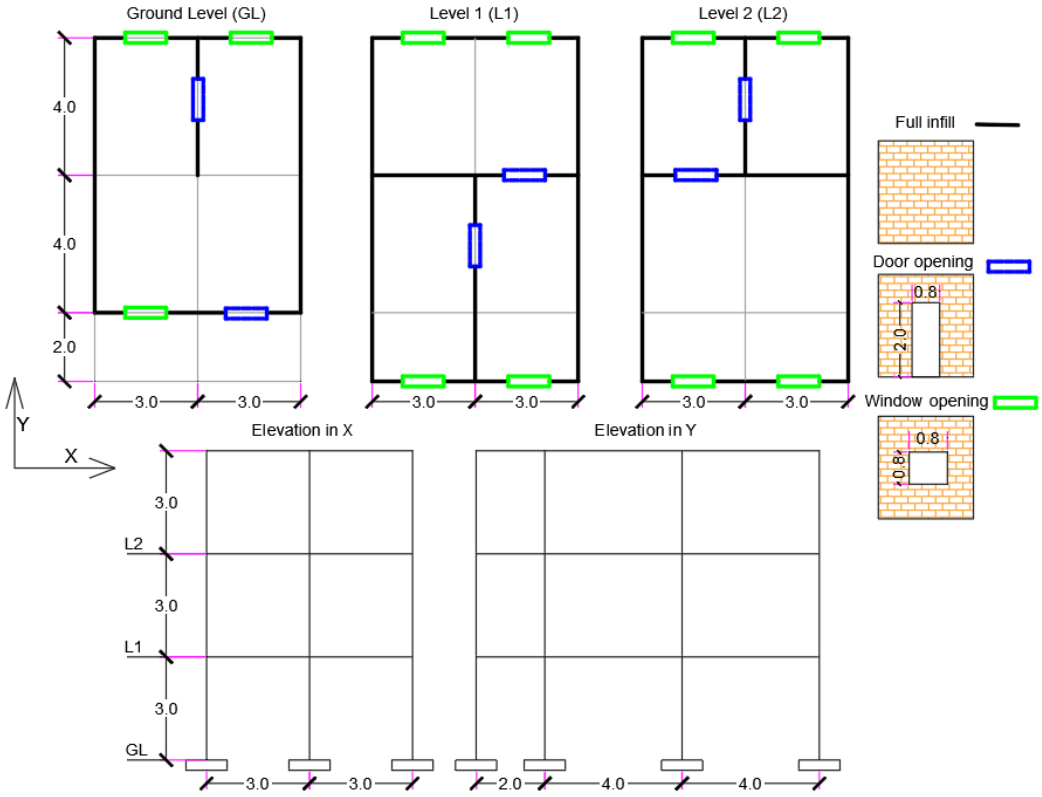

Figure 2. Configuration of the case study building.

A 3D finite element model of the case study building was created in OpenSees [23]. The slabs were modelled as equivalent beams and fiber-based beam-column elements were used for the beams and columns. The masonry infills were modelled as a single compressive strut connected to opposite corners in a frame (i.e., two single struts per frame). Also, the behavior of the infills under cyclic loading was implemented using the uniaxial material Pinching4 to account only for the compressive capacity. To avoid any convergence problem with this command, tension capacity is set to an extremely low value so that the strut capacity in tension has a null effect. All the details of the modeling can be found in [22].

After making the model, NRHA was performed for all the selected earthquake records in Table 3. In order to determine the damage state of the model hit by each of the records, the maximum inter-storey drift limit of 0.4 proposed by Ghobarah (2004) was adopted [24]. The definitions of the damage states described by Ghobarah are in accordance with the ones considered for the datasets (Table 2). 0.4 is the limit to separate the non-severe and severe classes for a RC Moment Resisting Frame Building with Infills. In other words, if the maximum inter-storey drift of the building model hit by an earthquake is less than 0.4, the damage state is classified as non-severe; otherwise, it is severe. It is worth mentioning that inter-storey drift is defined as the relative translational displacement between two sequential floors divided by the height of the storey.

Results and Discussion

The seven raw geometric parameters, WI, CI, and PI for the numerical model were computed and presented in Table 4. The 3D model was first evaluated using the logistic regression models made based on the PI as the only input feature. All the six logistic regression models trained on the six earthquake datasets predicted the damage class of the model as severe.

Table 4. The Hassan and Sozen characteristics of the numerical model.

No. of stories	Floor area (m^2)	Column area at base (m^2)	N-S concrete wall area at base (m^2)	E-W concrete wall area at base (m^2)	N-S masonry infill wall area at base (m^2)	E-W masonry infill wall area at base (m^2)	WI	CI	PI
3	60	0.48	0	0	2.208	1.012	0.06	0.13	0.19

The CI-WI models developed by logistic regression were then used to predict the damage state of the numerical model. Figure 3 shows the scatter plots in which CI and WI are on the horizontal and vertical axes respectively, the classification boundary defined by logistic regression is shown using red dashed lines, and the model accuracy is also displayed. The point representing the numerical model is displayed by a red star. As can be seen, for all the six datasets, the predicted damage state is severe as the red star is located under the decision boundary.

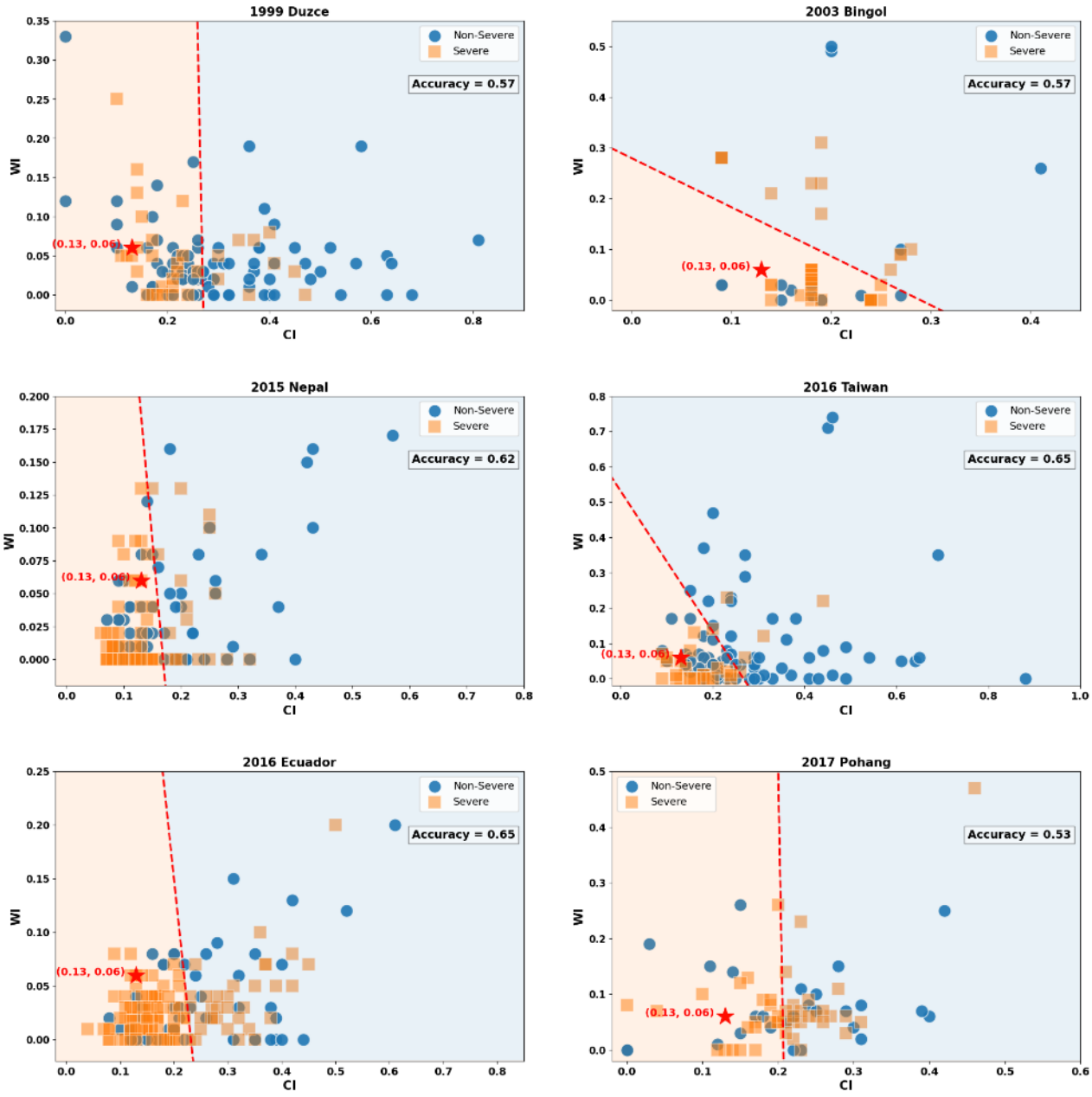

Figure 3. The building model displayed as a red star on the CI-WI scatter plots with classification boundaries defined by logistic regression.

In the next step, the region-specific RF models which were introduced as an improvement to the PI approach and trained directly using the seven raw geometric parameters, were used to predict the damage state of the model. All the six RF models also predicted severe damage state for the numerical model.

Eventually, the generalized non-region-specific RVS model proposed by the authors was validated. In this regard, the numerical results of the NRHA of the building model hit by the ten earthquake records were compared with the results predicted by the RF model trained on the combined datasets with considering MMI and PGA values as two additional input features. Table 5 presents the damage states obtained from NRHA as well as predicted by RF model.

Table 5. Damage states obtained from the NRHA and predicted by the generalized RF model.

Earthquake RSN	**719**	**721**	**5259**	**4140**	**818**	**1205**	**949**	**8064**	**1504**	**1119**
NRHA	Non-Severe	Non-Severe	Non-Severe	Severe	Severe	Severe	Severe	Severe	Severe	Severe
RF model	Non-Severe	Severe	Non-Severe	Severe	Severe	Severe	Severe	Severe	Severe	Severe

As can be seen, the RF model was able to correctly predict the damage state of the numerical model in nine out of ten earthquake scenarios. The results indicate that despite a region-specific model that is only able to predict the probable damage state of the low-rise RC buildings inside the region, the non-region-specific RF model can be used as a rapid screening method to identify high-risk buildings in the regions with similar seismic backgrounds. The numerical validation carried out in this study demonstrated the accuracy and applicability of the proposed RVS method.

CONCLUSIONS

This study utilized the numerical results of a low-rise RC building to validate an RVS method proposed by the authors in a previous study. In the proposed method, a RF model was trained and tested on the data from six earthquake events including Duzce (1999), Bingol (2003), Nepal (2015), Taiwan (2016), Ecuador (2016), and Pohang (2017). The seven geometric parameters from the Hassan and Sozen method (1997) in addition to two earthquake intensity measures, MMI and PGA, were employed as the input features of the RF model. The high accuracy of the model demonstrated the robustness of the proposed approach for vulnerability assessment of low-rise RC buildings. To further validate the applicability and accuracy of the proposed method, a 3D model of a real building from Ecuador was created in OpenSees. Ten earthquake records similar to the six earthquake events whose data were used to train the RF classification model, were also selected. The 3D model was first evaluated using the region-specific ML-based models developed based on the Hassan and Sozen method. The damage state of the building model was predicted for each region independently using the specific model for that region. The generalized not region-specific RF model was then utilized to predict the damage states of the building model for all the ten selected earthquake records. The predicted damage states were compared with the ones obtained from the NRHA of the numerical model. The outcomes of this numerical validation demonstrated the accuracy and effectiveness of the proposed ML-based RVS method.

REFERENCES

[1] Kassem, M.M., Nazri, F.M. and Farsangi, E.N., 2020. The seismic vulnerability assessment methodologies: A state-of-the-art review. *Ain Shams Engineering Journal*, *11*(4), pp.849-864.

[2] Rojahn, C., 1988. *Rapid visual screening of buildings for potential seismic hazards: A handbook* (Vol. 21). Federal Emergency Management Agency.

[3] Federal Emergency Management Agency (US) ed., 2017. *Rapid visual screening of buildings for potential seismic hazards: a handbook*. Government Printing Office.

[4] Bektaş, N. and Kegyes-Brassai, O., 2022. Conventional RVS Methods for Seismic Risk Assessment for Estimating the Current Situation of Existing Buildings: A State-of-the-Art Review. *Sustainability*, *14*(5), p.2583.

[5] Grünthal, G., 1998. *European macroseismic scale 1998*. European Seismological Commission (ESC).

[6] Milutinovic, Z.V. and Trendafiloski, G.S., 2003. Risk-UE An advanced approach to earthquake risk scenarios with applications to different european towns. *Contract: EVK4-CT-2000-00014, WP4: Vulnerability of Current Buildings*, pp.1-111.

[7] OASP (Greek Earthquake Planning and Protection Organization), 2000. Provisions for Pre-Earthquake Vulnerability Assessment of Public Buildings (Part A).

[8] Fathi-Fazl, R., Cai, Z., Jacques, E. and Cortés-Puentes, W.L., 2021. Methodology for seismic risk screening of existing buildings in Canada: Structural scoring system. *Canadian Journal of Civil Engineering*, *48*(3), pp.250-262.

[9] Tesfamariam, S. and Saatcioglu, M., 2008. Risk-based seismic evaluation of reinforced concrete buildings. *Earthquake Spectra*, *24*(3), pp.795-821.

[10] Ningthoujam, M.C. and Nanda, R.P., 2018. Rapid visual screening procedure of existing building based on statistical analysis. *International journal of disaster risk reduction*, *28*, pp.720-730.
[11] Coskun, O., Aldemir, A. and Sahmaran, M., 2020. Rapid screening method for the determination of seismic vulnerability assessment of RC building stocks. *Bulletin of Earthquake Engineering*, *18*, pp.1401-1416.
[12] Hassan, A.F. and Sozen, M.A., 1997. Seismic vulnerability assessment of low-rise buildings in regions with infrequent earthquakes. *ACI structural journal*, *94*(1), pp.31-39.
[13] Sim, C., Song, C., Skok, N., Irfanoglu, A., Pujol, S. and Sozen, M., 2016. Database of low-rise reinforced concrete buildings with earthquake damage. Retrieved from https://datacenterhub.org/deedsdv/publications/view/454. Accessed 2023
[14] Shah, P., Pujol, S., Puranam, A. and Laughery, L., 2016. 2015 Nepal Earthquake Building Performance Database. retrieved from https://datacenterhub.org/deedsdv/publications/view/537. Accessed 2023
[15] NCREE, Purdue University. 2016. 2016 Taiwan (Meinong) Earthquake. Retrieved from https://datacenterhub.org/deedsdv/publications/view/534. Accessed 2023
[16] Sim, C., Villalobos, E., Smith, JP., Rojas, P., Puranam, A.Y., Laughery, L. and Pujol. S., 2016. 2016 Ecuador Earthquake. Retrieved from https://datacenterhub.org/deedsdv/publications/view/535. Accessed 2023
[17] Sim, C., Laughery, L., Chiou, T.C. and Weng, P.W., 2018. 2017 Pohang Earthquake-Reinforced Concrete Building Damage Survey. Retrieved from https://datacenterhub.org/deedsdv/publications/view/296. Accessed 2023
[18] Johnson, M.G. and Fick, D.R., 2018. Generalized Trends of Severe Damage Observed from Building Surveys of Seven Different Earthquakes.
[19] Villalobos, E., Sim, C., Smith-Pardo, J.P., Rojas, P., Pujol, S. and Kreger, M.E., 2018. The 16 April 2016 Ecuador earthquake damage assessment survey. *Earthquake Spectra*, *34*(3), pp.1201-1217.
[20] Kleinbaum, D.G., Dietz, K., Gail, M., Klein, M. and Klein, M., 2002. *Logistic regression* (p. 536). New York: Springer-Verlag.
[21] Lundberg, S.M. and Lee, S.I., 2017. A unified approach to interpreting model predictions. *Advances in neural information processing systems*, *30*.
[22] Kagermanov, A., Ceresa, P., Morales, E., Poveda, J. and O'Connor, J., 2017. Seismic Performance of RC Buildings during the Mw 7.8 Muisne (Ecuador) Earthquake on April 2016: field observations and case study. *Bulletin of Earthquake Engineering*, *15*, pp.5167-5189.
[23] OpenSees version 3.3.0. http://opensees.berkeley.edu
[24] Ghobarah, A., 2004, June. On drift limits associated with different damage levels. In *International workshop on performance-based seismic design* (Vol. 28). Ontario, Canada: Department of Civil Engineering, McMaster University.