18TH WORLD CONFERENCE ON EARTHQUAKE ENGINEERING
WCEE2024
MILAN

MILAN, ITALY
30TH JUNE - 5TH JULY 2024

www.wcee2024.it

AMPLIFICATION FACTORS OF TORSIONALLY COUPLED RC FRAME BUILDINGS USING ARTIFICIAL INTELLIGENCE

J.C. de la Llera[1], M. Arré[2], F. Bertin[2], J.P. Muñoz[2] & J.C. Reyes[3]

[1] Pontificia Universidad Católica de Chile, Research Center for Integrated Disaster Risk Management (CIGIDEN), ANID/FONDAP/1522A0005, Santiago, Chile, jcllera@ing.puc.cl

[2] Pontificia Universidad Católica de Chile, Research Center for Integrated Disaster Risk Management (CIGIDEN), ANID/FONDAP/1522A0005, Santiago, Chile

[3] Universidad de Los Andes, Faculty of Engineering, Department of Civil and Environmental Engineering, Bogotá, Colombia

Abstract: *Inelastic lateral-torsional coupling is a problem of estimating displacement demands that are unevenly distributed across the building plan. Several natural and accidental factors influence this behavior, such as ground motion characteristics, dynamic properties of the system, and the plan and heightwise distribution of mass, stiffness and strength. The problem still remains elusive to simple design equations and is a relevant topic in earthquake-resistant building design. This study works with the expected response amplification factors (RAFs) of seismic demands to simplify the problem, and uses a data science perspective by considering over 100,000 structures and machine learning (ML) techniques to predict RAFs in frame structures. This is only a first step of a larger effort to use ML to predict RAFs for different structural configurations, and describes the simplest case of inelastic single-story monosymmetric reinforced concrete frame buildings, which were parameterized using properties such as beam/column dimensions, plan distribution of resisting planes, and geometric properties of the plan. The systems were automatically designed according to the Chilean code and modeled with force-based fiber elements in OpenSees. Since this initial study only aims to capture the effect of structural properties on RAFs, a single seismic record was considered and scaled so its spectral acceleration in the average fundamental period of the dataset corresponds to a hazard level with a return period of about 2500 years. Different structural responses were examined, including drifts at the center of mass and building edges, base-shear and torque, and curvature ductility demand in the elements. Two companion models were considered for each structure—an asymmetric elastic, and a nominally symmetric inelastic system. Responses were normalized relative to these models to compute RAFs, and a ML model was trained (and validated) to predict them. It was concluded that the predictions are accurate enough for design purposes with estimation errors less than 20%*.

1 Introduction

Structures with lateral-torsional coupling have been the subject of study in structural dynamics for several decades (Kuang, Jiang & Jiang, 2018; Yiu et al., 2013; Tso & Ying, 1992; Bozorgnia & Tso, 1986), due to their larger sensitivity to failure as a result of the uneven lateral inelastic deformations of the resisting planes. Because of the rotation of the plan, the resisting planes located at the edges of the building, or more distant to the rotation point of the plan, may undergo larger deformations and stresses, and if not properly designed, may have an increased probability of severe damage, or even collapse.

Because many buildings in practice present irregularities and uneven distributions of stiffness, mass, and strength, several studies have analyzed this phenomenon in search of elastic and inelastic system parameters that can help predict these torsional amplifications. Parameters such as stiffness, strength, and mass eccentricities, normalized yield strength, uncoupled lateral period ratios, and the uncoupled torsional-to-lateral

frequency ratio, have been proposed to characterize the effects of torsion, considering the elastic and inelastic response of different groups of structures (Kuang, Jiang & Jiang, 2018; Yiu et al., 2013; Tso & Ying, 1992; Bozorgnia & Tso, 1986). Despite the significant efforts carried in this direction, no general conclusions have been achieved for different types of structures and ground motions. There are many reasons for this, but as an example please think of the inelastic excursions of a structure, which imply instantaneous changes in the stiffness of resisting planes, which depend, in turn, on the dynamic characteristics of the structure and frequency content of the seismic record. Thus, structural responses may vary between structures with similar parameters.

Given the discrepancies observed between the inelastic torsion results of previous studies, which are mostly attributed to different assumptions in the models, there is a need to establish more robust trends for a wider range of structures. Hence, this study aims to analyze a large number of lateral-torsional coupled structures in order to bound and predict such amplifications. As a starting point, more than 130,000 inelastic models with lateral-torsional coupling were created and their earthquake response histories computed. Two benchmark model results were used to compare: one representing the elastic model of the structure, and the other, a symmetric inelastic counterpart. Response amplification factors (RAFs) are normalized relative to these two base cases. To process all this information, a Machine Learning (ML) algorithm approach is used.

Machine learning (ML) algorithms use statistical techniques to learn from data and make predictions or decisions based exclusively on the data and have gained importance in some structural dynamics applications (Yu, Yao & Liu, 2020; Cunha et al., 2023; Birky et al., 2021). ML models also have the capabilities to give information of the relevant variables for each input feature (Thai, 2022), helping investigators in understanding the effects of the parameters considered in different analyses. In this research, ML algorithms are used to predict RAFs from a large database of response history inelastic analysis of single-story systems, to better understand the impact of the torsional irregularities in the responses of different structures. This study represents an initial effort to develop an adequate ML model for RAF prediction, and hence, to reduce the complexity in estimating this critical parameter in more complex structures. This is only an initial effort, and the variability of the ground motion is not yet considered in the analyses presented. A following study will also consider the effect of varying ground motion input and other building configurations.

2 Structural analysis methodology

2.1 Structural models

This study considers the responses of single-story RC, torsionally coupled structures. The structures were modelled using OpenSees for Python (University of California, 2000), considering the inelastic behavior of beams and columns. For the models shown in Figure 1, 9 parameters were defined in a physical range of values to build a variety of different structures. First, the lengths of the spans in the X direction, L_1, L_2 and L_3, were defined. The values of the lengths L_1 and L_3 can be 0, or from 5 to 12 m, while the values for L_2 can be 0, or ranging from 2.5 to 4.5 m. Then, the total length of the slab in the X-direction, a, is the sum of L_1, L_2 and L_3. Different rules were defined for the cases in which one, two or the three lengths of the spans were 0, so as to force that the structure has at least two resistant planes in the Y-direction. Considering these definitions, parameter a took values between 6.5 m and 39.0 m. The total length of the slab in the Y direction, b, was obtained through the parameter defined as the aspect ratio of the slab, $r_s = a/b$, which was set to ⅓, 0.5, 1, 2 or 3. Therefore, parameter b took values between 5 and 72 m. The number of resistant planes in the X-direction was defined so as to keep the lengths of the spans between 5 and 12 m, and keep the symmetry of the structure in that direction. Thus, all the buildings created were monosymmetric in the X-axis and asymmetric in the Y-axis given by the location of the resistant planes. The height of the structure, H_s, was set between 2.5 and 5 m. With these first parameters, the geometry of the structure looks like to one in Figure 1.

Two other parameters were defined to set the dimensions of beams and columns. First, the height of the beams was set to be either 1/10 or 1/15 of the length of the span, and then the width of the beam was set by using the parameter r_b, defined as the ratio between the width and height of the beam, which values can be ⅓ or 0.75. The columns were all built as squared sections, whose sizes were set by the parameter r_{bc}, defined as the ratio between the height of the beam and the width of the column, and took values between 0.6 and 1. Finally, parameters *L* and *D* define the live and dead loads for the structure, which were considered to be

uniformly distributed on the slab. The live load took values between 200 and 600 kg/m^2, and the dead load between 400 and 1200 kg/m^2. All these parameter values are very common in Chilean practice.

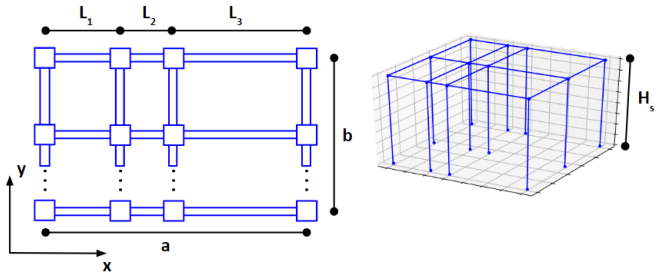

Figure 1. Geometric definition of the models

Considering different combinations of the parameters, a total of 136,545 structures were created. We also assumed an elastic modulus of *E* = 23.42 GPa and a shear modulus of *G* = 9 GPa. These linear structures were automatically designed according to the Chilean code NCh433 (INN, 1996) and the ACI 318 (American Concrete Institute, 2019), assuming that the characteristic resistance of concrete was f'_c = 250 kg/cm^2 and the yield tension of the steel f_y= 4200 kg/cm^2. Given its proof-of-concept nature, and understanding that this only a first attempt to the topic, all the columns and beams in the same direction in each structure were designed identical, considering the required reinforcement of the most solicited element. The distribution of the parameters and properties of these structures and their designs are shown in Figure 2, in which T_x, T_y and T_θ are the uncoupled periods of the linear structure, e_x/a is the normalized eccentricity of the structure, r is the radius of gyration of the slab, Ω_x= T_x/T_θ and Ω_y= T_y/T_θ are the uncoupled lateral-to-torsional frequency ratios, ρ_{col} is the average reinforcement ratio of the columns, $T_{y,asym}$ is the linear coupled period in the Y-direction and *m* = *D* + 0.25*L* is the seismic distributed mass.

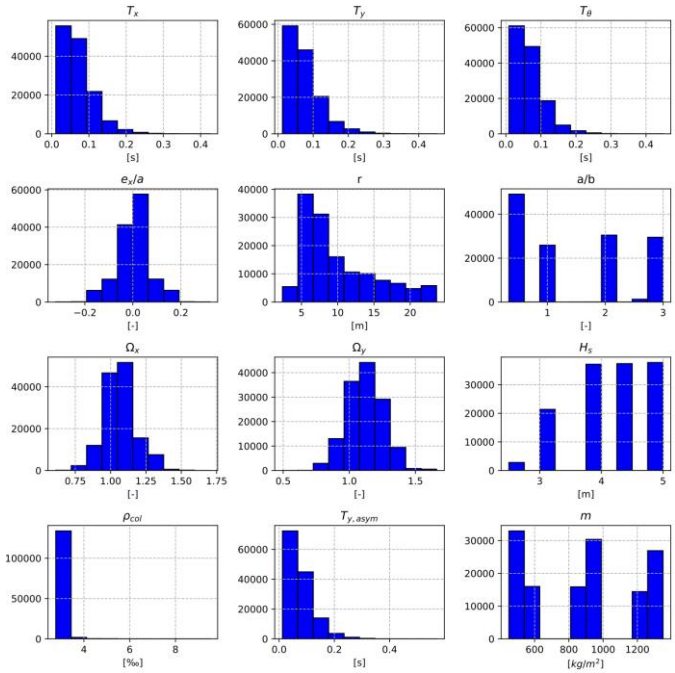

Figure 2. Parameter distribution for the linear structures.

The designs for the structural elements were obtained taking into account the axial-bending interaction curves of beam and column elements, checking the resistance to shear force, and assuring that the criteria of bending moment strength between beams and columns was fulfilled. Inelastic structures were modeled considering these designs, using the *forceBeamColumn* element available in OpenSees. The sections were created as a fiber section, using materials *Concrete02* and *Steel02* (University of California, 2000). The concrete strain-

stress relationship was regularized to maintain the concrete crushing energy G_f/L_p constant, where L_p is the length of the plastic hinge. As stated by multiple authors (Jansen & Shah, 1997; Nakamura & Higai, 2001), this property is rather invariant in the material.

The structures were then modeled with frames formed by nonlinear columns and beams, and an in-plane rigid diaphragm was defined at the floor slab level, so that the X- and Y-displacements and the rotation in the Z direction of the Center of Mass (CM) can describe all of the displacements of the structure. The total seismic mass was computed as $M = abm$, and then divided proportionally to each of the nodes defined at the slab level. The corresponding percentage was assigned as translational mass in the three directions, and each of the structures has a 5% modal damping of Rayleigh's type in the first and third mode.

2.2 Seismic record

As a first step in this research, one seismic record with two components was used to compute the response history analyses. A record from the 2010 8.8Mw Maule earthquake in Chile was obtained from the Hualañé seismic station. This accelerogram was selected due to its high intensity (e.g. PGA = 0.42 g) and because the station is located on firm soil (Vs30 = 547 m/s), which is consistent with the test building design assumptions (soil type B, 500 m/s ≤Vs30 ≤900 m/s). A Uniform Hazard Spectrum (UHS) with a return period of 2500 years was computed for the building assumed location in Santiago, using the SeismicHazard toolbox (Candia et al., 2019), and considering the Poulos 2018 recurrence model, five Ground Motions Models with equal weights (Zhao 2006, Montalva-Bastías 2017, Kuehn 2020, Parker 2020, and Abrahamson-Gulerce 2020), and the Strasser 2010 scaling model. The selected record was scaled to match the hazard level of the UHS at the fundamental vibration period of the structure (T_n = 0.1 s), in a geometric mean sense, resulting in a scaling factor of 2.68. For computational efficiency, the record was also trimmed to consider only the accelerations between 1% and 95% of the peak Arias intensity. A value of 1% was considered instead of 5% for the lower bound to avoid using a seismic record with considerable initial acceleration, thus an iterative code in MATLAB was used to obtain a scaling factor consistent with the record trimming. Figure 3 shows the pseudo-acceleration response spectrum of the record.

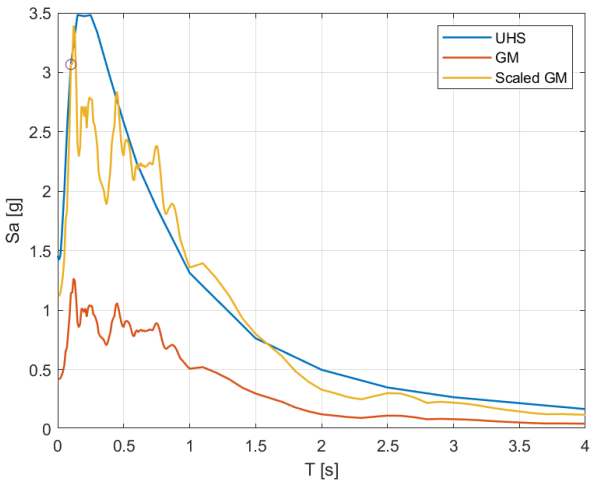

Figure 3: Target Uniform Hazard Spectrum and Hualañé unmodified and scaled response spectra.

2.3 Procedure

For each model in the database, two other related structures were built to compute RAFs, thus three analyses were made for each of the more than 130,000 structures. First, a nonlinear response history analysis (NRHA) was performed on the nonlinear base structure with lateral-torsional coupling. For these monosymmetric structures, a rigid diaphragm was considered, so the equation of motion for each nonlinear model corresponds to Equation 1, where $r(t) = [x(t), y(t), \theta(t)]^{\mathrm{T}}$ is the state vector in which *x* and *y* are the displacements of the center of mass, and θ is the rotation of the slab, $[\ddot{x}_g(t), \ddot{y}_g(t)]^{\mathrm{T}}$ is the ground motion vector where the record to consider was explained in section 2.2, and $L^T f(t)$ are the nonlinear forces of the elements.

Responses such as base reactions, forces and deformations of the structural elements and CM, and rotation were obtained from these NLTHA. The maximum interstory drift was computed as shown in Equation 2, which is used as one of the output parameters for the ML analysis.

$$M\ddot{r}(t) + C\dot{r}(t) + L^T f(t) = -M\begin{bmatrix}1 & 0\\ 0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}\ddot{x}_g(t)\\ \ddot{y}_g(t)\end{bmatrix} \quad (1)$$

$$\delta_{max} = \max\{x(t)b/2\,; y(t)a/2\}/H_s \quad (2)$$

The second model built for this study corresponds to a linear version of the base structure, in which the beams and columns are created with the *elasticBeamColumn* element. As in practice, this is the model used to design the structure. The maximum interstory drift was obtained for this model considering the responses of a linear time history analysis, and with that parameter the first RAF was computed. The normalized nonlinear to linear amplification, A_{NL-L}, corresponds to Equation 3, which is defined (pairwise) as the ratio between the maximum drift of the nonlinear monosymmetric structure and the maximum drift of the linear monosymmetric structure.

$$A_{NL-L} = [\max\left\{\{x_{NL}(t)b/2\}\,; \max\left\{\frac{y_{NL}(t)a}{2}\right\}\right\}]/[\max\left\{\frac{x_L(t)b}{2}\right\}; \max\left\{\frac{y_L(t)a}{2}\right\}] \quad (3)$$

The third and last model built for all structures in the database corresponds to a nonlinear symmetric equivalent structure, considered to evaluate the amplification of the responses due to torsional effects. This model was created by moving the resisting planes in the Y-direction in such a way that the resulting eccentricity was zero, this is, the CM and stiffness were located in the same place. Since the resisting planes were moved, the fundamental period of the structure would change, therefore the total mass of the structure was also modified so as to have a fundamental period that was equal to the one of the base structure. Considering this model, the asymmetric to symmetric amplification factor, A_{As-S}, was computed, as shown in Equation 4.

$$A_{As-S} = [\max\left\{x_{As}(t)b/2\right\}\,; \max\left\{\frac{y_{As}(t)a}{2}\right\}]/[\max\left\{\frac{x_S(t)b}{2}\right\}\,; max\{y_S(t)a/2\}] \quad (4)$$

Figure 4 shows a representation of the three models considered. The base structure which is a monosymmetric nonlinear model, the linear asymmetric reference structure and the nonlinear symmetric structure.

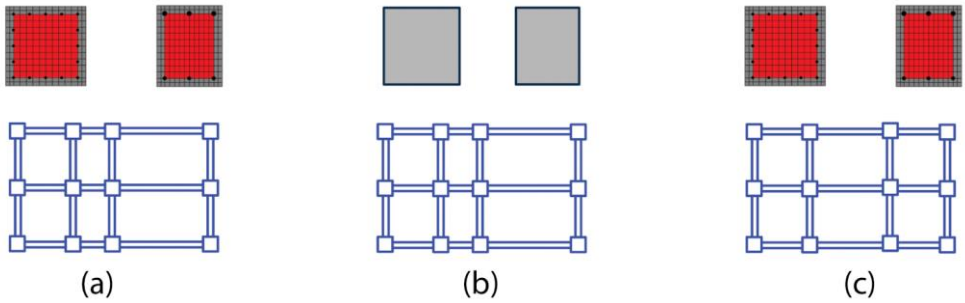

Figure 4. Models considered for the analysis. (a) Monosymmetric nonlinear structure, base case; (b) monosymmetric linear equivalent structure; and (c) nonlinear symmetric equivalent structure.

3 Overview of machine learning algorithms

3.1 Machine learning algorithms considered

ML algorithms work as an input/output system, in which the input variables are processed to obtain the output results. The testing inputs used for the algorithm are the ones presented in Figure 2, which include uncoupled periods, dimensions of the structure and the reinforcement ratio of the columns. For each of the three parameters mentioned previously, A_{NL-L}, A_{As-S} and δ_{max}, a different model was trained so as to predict its output value. Table 1 shows the range of values of these three output parameters.

Table 1: ML model output parameters.

Output parameter	Description	Range of values
A_{NL-L}	Nonlinear to linear amplification	0.394 – 5.596
A_{As-S}	Asymmetric to symmetric amplification	0.155 – 16.549
δ_{max}	Maximum interstory drift	0.00001 – 0.044

After the inputs and outputs of the ML model are defined, the next step is to fit the ML algorithm to the data. In order to do this, the database is divided into two groups. The first group is used to train the algorithm, which adjusts the internal parameters so as to predict the corresponding output by using the input values. This group is called training dataset/set in computer science. The second group, which is not considered in the training step, is used to evaluate the precision of the model by comparing its predictions to the actual output results. This group is called testing dataset/set in computer science, since it tests the accuracy of the model.

In this study, 3 different ML algorithms were evaluated to work with the problem described.

1. Random Forest (RF) is a ML algorithm notable for its ability to reduce overfitting and improve prediction accuracy. Instead of relying on a single decision tree, RF creates a collection of trees during training and calculates the final prediction by taking the average of the predictions from these individual trees (Thai, 2022).
2. Gradient Boosting (GB) is another ML algorithm that consists of building multiple models called weak learners (generally they are decision trees) sequentially, in which each model reduces the error of the previous one. The process of building models stops when it is not possible to further reduce the error. The final prediction is a combination of the individual model predictions (Thai, 2022).
3. Support Vector Regression (SVR) is a supervised learning algorithm that is used to do regressions. Unlike traditional regression methods, SVR is based on the idea of finding an optimal hyperplane that minimizes the difference between model predictions and actual values, allowing some flexibility in accepting errors within a specific margin (Thai, 2022).

3.2 Metrics considered

To evaluate the performance of the ML models, three metrics of accuracy of the predictions were considered for each output parameter.

1. Root Mean Square Error (RMSE): this metric computes the square root of the average of the squared errors between the model predictions and the actual values. The expression of this metric is shown in Equation 5, where A_i is the real output of the structure and F_i is the value predicted by the ML model.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(A_i - F_i)^2} \tag{5}$$

2. Coefficient of Determination (R^2): this metric indicates how well the model fits the observed data. The expression is presented in Equation 6, where $\bar{A}$ is the average value of the specimens.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(F_i - \bar{A}_i)^2}{\sum_{i=1}^{n}(A_i - \bar{A}_i)^2} \tag{6}$$

3. Mean Absolute Percentage Error (MAPE): this metric measures the average percentage absolute error between model predictions and actual values, as it is presented on Equation 7.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|A_i - F_i|}{|A_i|} \tag{7}$$

4 Results and discussion

4.1 Time history analyses

To have a better understanding of the effects of lateral-torsional coupling in these structures, we examined in detail some individual responses of the structures in the database. Figure 5(a) shows the structure number 85439 and Figure 5(b) the displacement responses of the center of mass (CM) of the structure, whose parameters are those in Table 2. The CM displacements in the Y-direction shows that the structure experiences inelastic behavior, having a residual displacement of approximately 5 cm, which means a residual drift of 1.1%. For the same structure, the drifts at the edges are shown in Figure 5(c), and it is apparent that the responses at the left and right edges are different. Since the structure is asymmetric, both edges experience different displacements because of the rotational effects. This difference increases after the first 25 s of the record, time for which the structure starts its inelastic behavior. Given the existing eccentricity, one of the edges has bigger displacements, and, hence, more inelastic behavior in its columns and beams, resulting in a reduction of stiffness that makes the eccentricity even bigger. Thus, the right edge plane ends the analysis with a residual drift of 1.5%, while the left edge has a residual drift of 0.9%.

Table 2: Example of parameters for base structure number 85439.

Input parameters	T_y (s)	T_x (s)	T_θ (s)	e_x/a	$T_{y,asym}$(s)	r (m)	Ω_x	Ω_y	H_s (m)	ρ_{col}	a/b	a (m)	b (m)	m (kg/m^2)
Value	0.312	0.262	0.283	0.03	0.294	7.26	0.927	1.101	4.375	0.0053	2	22.5	11.25	1250

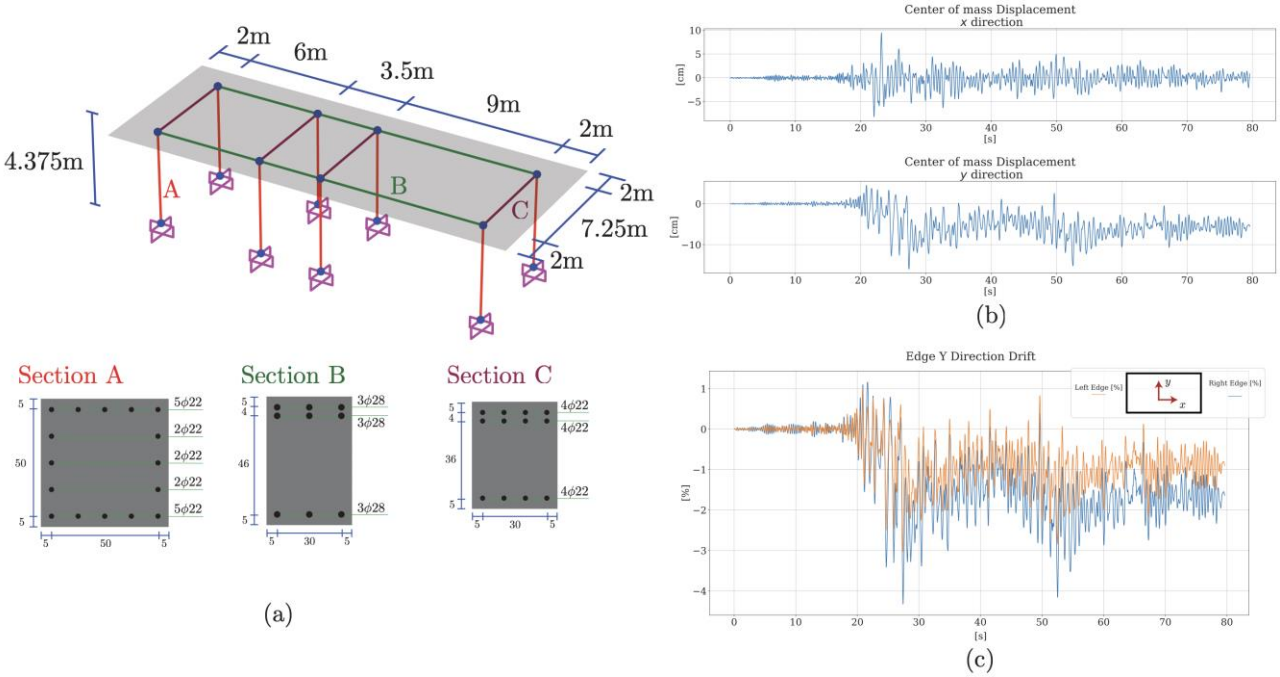

Figure 5. Example of structure. (a) Structure 85439; (b) center of mass displacements of structure 85439; and (c) Edge drift response of the structure 85439.

Figure 6 presents the edge drift responses for the equivalent symmetric structure. Since this structure has symmetry with respect to the CM, the responses at both edges must be equal. In this case the residual drift is of 1.1% for the left and right edges, which is the same value as the residual drift at the CM.

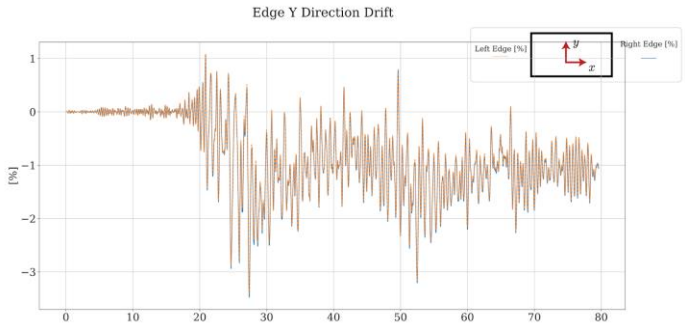

Figure 6. Edge drift response of the counterpart equivalent symmetric structure of the selected sample structure in the database.

4.2 Training and testing datasets for ML model

The three ML algorithms were trained with multiple datasets of different sizes, to compare their relative accuracy. The sizes of the training datasets are described as percentages of the total database, and in this case there were 9 training datasets considered with a different quantity of structures: (i) 0.1% (100 samples); (ii) 0.7% (1000 samples); (iii) 3.7% (5000 samples); (iv) 7.3% (10000 samples); (v) 14.6% (20000 samples); (vi) 36.6% (50000 samples); (vii) 73.2% (100000 samples); (viii) 84.2% (115000 samples); and (ix) 95.2% (130000 samples). For each case, the size of the testing dataset corresponds to the complementary percentage. For example, if the size of the training dataset is 36.6%, then the size of the testing dataset is 63.4% (86545 samples).

Figure 7 presents the results of the performance metrics for the ML models considering the 9 different testing datasets. The first important aspect to highlight is that the metrics improve when the size of the training datasets increase, but some metrics reach the best value for training datasets smaller than 95.2%. For example, the RMSE for the A_{As-S} has the minimum value for the 84.2% dataset for SVR and GB algorithms. Also, for the amplification parameter A_{NL-L} predicted with the GB algorithm, it is possible to see that the MAPE is reduced only by 0.5% from the 3.7% to the 95.2% datasets. Given that the results of the performance metrics are similar for the three biggest datasets, and that there is a recommendation to consider a testing dataset of at least a 20% of the database (Thai, 2022), the size chosen to do the analyses next is the 73.2%, which has 100000 samples for training.

Considering this training dataset, it is easy to see that the algorithm that predicts the best output is the RF model. The coefficient of determination for the parameter A_{NL-L} has values of 0.967 for the RF while the R^2 values are 0.863 for GB and 0.76 for the SVR algorithm. For the drift, the value for the RF reaches 0.991, while for the GB is 0.98. The RMSE values also show that the best algorithm is RF. For example, for the asymmetric to symmetric amplification, A_{As-S}, it is easy to see that the blue line is the lower one, having RMSE(A_{As-S}) = 0.0243, and also for the nonlinear to linear amplification, where RMSE(A_{NL-L}) = 0.0108, while the same values for GB are RMSE(A_{As-S}) = 0.1014 and RMSE(A_{NL-L}) = 0.0516. The smallest error in this case corresponds to the one of the δ_{max} models, since these values range from 10^{-2} to 10^{-5}, proving again that the RF has the smallest one with RMSE(δ_{max}) = 1.17×10^{-7}. The MAPE metric can be used to compare the performance of the models with different outputs, since this metric evaluates the percentage error of the prediction. The ML algorithm with the lowest MAPE is also the RF, with values of less than 5% for all the output parameters, in contrast to the other two ML algorithms that have values over 10% or 30% in some cases.

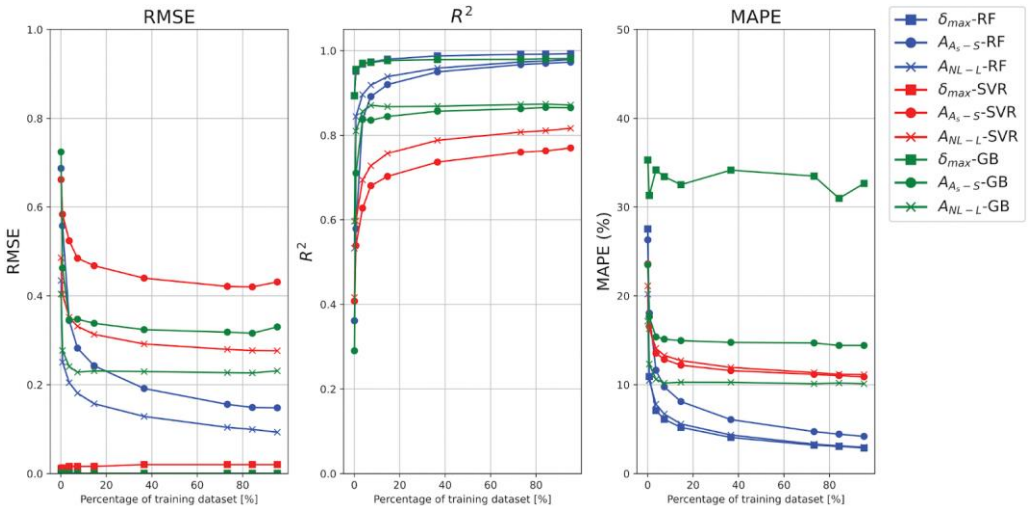

Figure 7. Performance metrics of ML models for different sizes of testing dataset.

After the training process for the algorithms, a SHAP (SHapely Additive exPlanations) values analysis was performed to look at the relevance of each parameter in the model using the Python library shap (Lundberg et al., 2020). SHAP is a powerful and widely used technique in the field of machine learning explainability. It provides a unified measure of the importance of each feature or input parameter, helping to interpret complex models and understanding the impact of different features on model predictions. Derived from cooperative game theory, SHAP analysis assigns a value to each feature representing its contribution to the prediction. The SHAP values are the contribution of each feature for the final prediction of the responses of the structure (Lundberg et al., 2020).

Figure 8 shows SHAP value for each feature and for each structure, for the three outputs obtained from the analyses performed considering the testing dataset. For the maximum story drift, δ_{max}, the two most important features of the ML model are the uncoupled period in Y-direction, T_y, and the story height, H_s, which are the parameters that control the response of the structure and the normalization for the drift. For the uncoupled period it is possible to conclude that the bigger its value, the bigger the effect it has on the drift, making it increase. This can be explained by the typical displacement response spectra, in which for small periods the displacement increases with the period. On the other hand, for the story height, we can see that when it has a small value the drift response increases.

For the A_{As-S}, the features with the biggest SHAP values are the torsional to lateral frequency ratio for the X and Y directions, and the coupled and uncoupled period in Y direction. The lateral periods control the responses of the system, while the torsional-to-lateral frequencies are the parameters that help explain the difference between the displacements of the asymmetric and symmetric equivalent structures. Figure 8 shows that the frequency ratios have antagonistic effects over the ML model for the A_{As-S}, but this effect changes with the values of the parameters. Figure 9 (a) and (b) shows that when $\Omega_x < 1$ this variable has great importance for the ML predictions, since the structure is more rigid in the X-direction. On the other hand, when $\Omega_y < 1.2$, this feature increases its importance for the predictions, given that the structure is more flexible in the Y direction. Figure 10 also shows that the periods T_y and $T_{y,asym}$ present antagonistic trends. Figure 9 (c) and (d) shows that for periods bigger than 0.11 s, the contribution of both features is constant, having a constant SHAP value. For smaller periods, the coupled period increases its importance, having bigger SHAP values.

Finally, for the inelastic to elastic amplification, the features with the biggest SHAP values are both uncoupled periods. The lower the period in the Y-direction, the smaller the amplification since displacements are going to be smaller and hence, a smaller excursion in the inelastic range will happen.

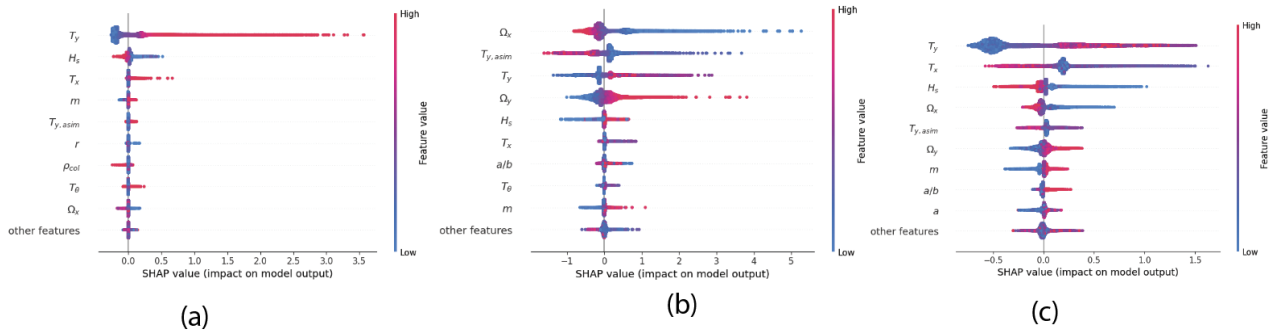

Figure 8. SHAP values for the testing dataset for (a) maximum story drift δ_{max} ; (b) asymmetric to symmetric amplification A_{As-S} ; (c) nonlinear to linear amplification A_{NL-L}.

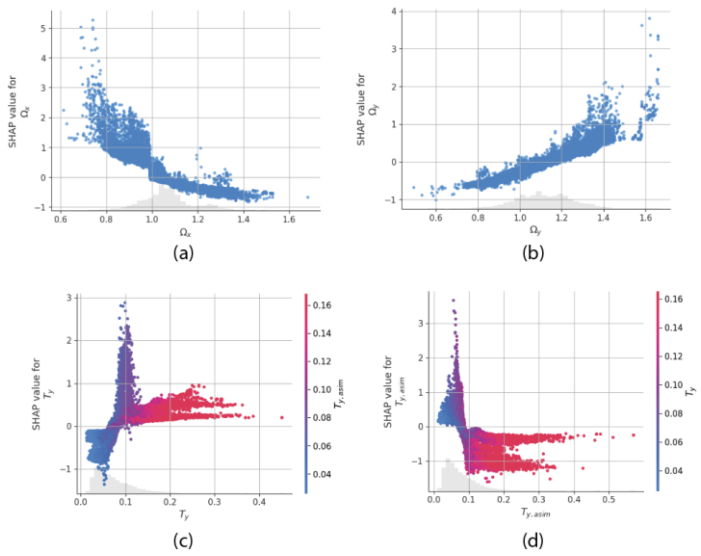

Figure 9. Analyses for A_{As-S} ML model: (a) SHAP values for feature Ω_x ; (b) SHAP values for feature Ω_y ; (c) SHAP values for feature T_y ; and (d) SHAP values for feature $T_{y,asym}$.

Figures 10 and 11 present the results for both amplification factors considered in this study. For the A_{As-S}, Figure 10(a) shows a good correlation between the real values and the predicted ones by the ML model, having most of the points near the identity line. Figure 10(b) shows the amplification factor with respect to the frequency ratio in the X-direction. Results show that the larger the value of the frequency ratio, the lower the asymmetric to symmetric amplification. This is due to the fact that structures with $\Omega_x > 1$ are torsionally stiff,

therefore the amplification produced by torsional effects is not as significant as it is when $\Omega_x < 1$. Although this is a tendency, a frequency ratio larger lower than 1 does not necessarily means a large value of A_{As-S}, since this factor depends on other features as well. Figure 10(c) shows that there is a difference in the structures with periods larger than 0.1 s and the ones with periods lower than 0.1 s. The more rigid structures have a larger asymmetric to symmetric amplification factor, while the more flexible ones have smaller amplifications.

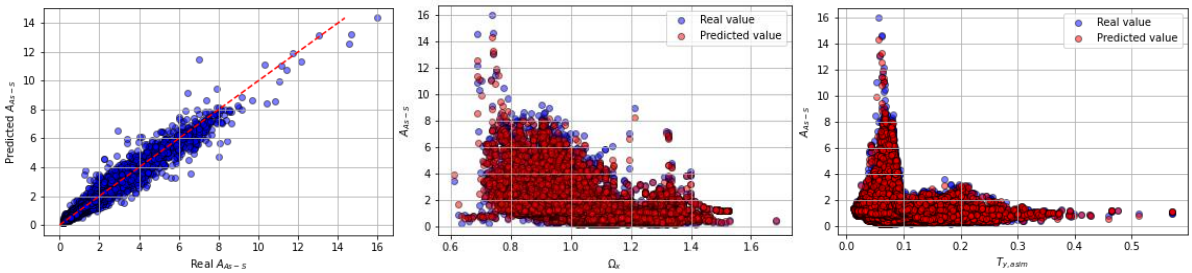

Figure 10. Results of A_{As-S} ML model for testing dataset: (a) relation between the real amplification factors and the ones predicted by the ML model; (b) amplification factors and frequency ratio Ω_x; and (c) amplification factors and coupled period in Y-direction.

Figure 11(a) shows that the ML model for the nonlinear to linear amplification also predicts well the values. Figures 11(b) and (c) show that there is a peak of amplifications for a 0.1 s period in the Y-direction and a 0.06 s period in the X-direction. These peaks may be explained by the response spectrum of the seismic record shown in Figure 3, where the higher value of the pseudo-acceleration is around a period of 0.1 s. Besides those peaks, results show that when the period increases, the A_{NL-L} also increases, which is explained by the fact that a bigger period means a larger value of pseudo-displacement.

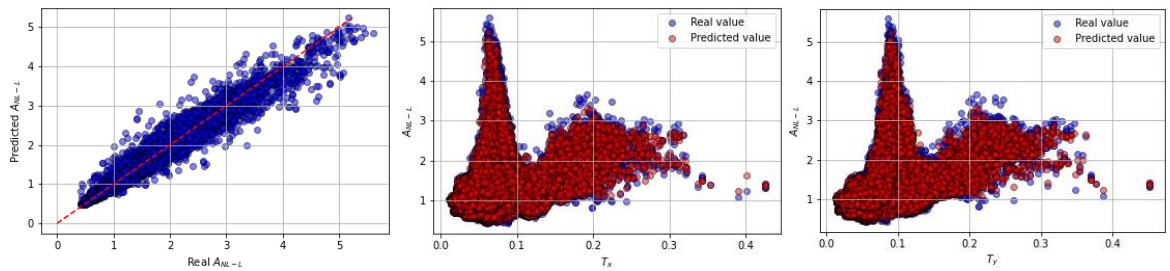

Figure 11. Results of A_{NL-L} ML model for testing dataset: (a) relation between the real amplification factors and the ones predicted by the ML model; (b) amplification factors and uncoupled period in X-direction; and (c) amplification factors and uncoupled period in Y-direction.

4.3 Additional structure

In order to further test the proof-of-concept ML algorithm developed, an additional structure was created and designed manually, which was not included in the database for the training of the model. The input properties for the ML model of this structure are shown in Table 3. The structure is presented in Figure 12, and it incorporates some differences from the structures of the database. This system has two types of columns, shown in yellow and green, and three different types of beams.

Table 3: Input parameters for ML model of additional structure.

Input parameters	T_y (s)	T_x (s)	T_θ (s)	e_x/a	$T_{y,asym}$(s)	r (m)	Ω_x	Ω_y	H_s (m)	ρ_{col}	a/b	a (m)	b (m)	m (kg/m^2)
Value	0.145	0.145	0.358	0.09	0.148	9.02	0.404	0.405	3.60	0.0075	1.2	24	20	625

The RF model results are shown in Table 4, and they seem to predict well the response. The errors of the predictions made by the ML algorithm are lower than 20% for the three outputs, which is a reasonable approximation of the results.

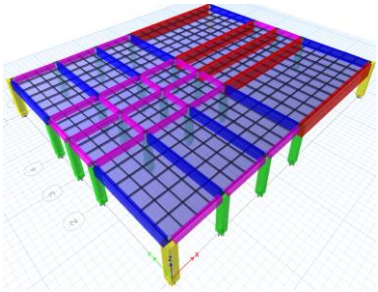

Figure 12. Additional structure.

Table 4: Output predictions and error of ML model for additional structure

Outputs	Real value	Predicted value	Percentage error
A_{As-S}	0.786	0.906	15.3%
A_{NL-L}	2.10	1.73	17.5%
δ_{max}	1.30%	1.13%	12.9%

5 Conclusions

This study is a first attempt to investigate the benefits of using artificial intelligence in the prediction of structural responses in lateral-torsional coupled structures. Over 130000 single-story reinforced concrete structures were analyzed, comparing their responses with two reference models. The results of the time history analyses show that structures that have torsional irregularities present bigger displacements at the building edges. The implementation of the ML models showed that these algorithms are capable of predicting with sufficient accuracy responses such as story drift, asymmetric to symmetric amplification and nonlinear to linear amplification of the drift. The algorithm that worked best with the data presented was the random forest, having predictions with errors under 10% for the structures in the testing datasets and under 20% for the additional structure presented. Uncoupled lateral periods and coupled lateral periods appeared as the most relevant features for the ML models, followed by the frequency ratios and story height. The real and predicted values show that when the frequency ratio is lower than 1, the structure will have larger asymmetric to symmetric amplifications, since it is torsionally flexible. Also, results showed that for larger periods the nonlinear to linear amplifications are bigger.

The next steps for this investigation are to create a more diverse database of structures, considering more typologies and stories, and to consider more seismic records. Once this is achieved, the evaluation of the accuracy of the ML models needs to consider more performance metrics, and another set of structures that are not included in the database for testing. The goal is to be able to predict RAFs without having to perform a full inelastic response history analysis and using only system parameters.

Powered@NLHPC: This research was partially supported by the supercomputing infrastructure of the NLHPC (CCSS210001).

6 References

Abrahamson, N.A. & Gulerce, Z. (2022). Summary of the Abrahamson and Gulerce NGA-SUB ground-motion model for subduction earthquakes, *Earthquake Spectra*, 38(4). DOI: 10.1177/87552930221114374

American Concrete Institute, (2019). Building Code Requirements for Structural Concrete, ACI 318 - 19. Texas: American Concrete Institute.

Birky, D., Ladd, J., Guardiola, I. & Young, A. (2021). Predicting the dynamic response of a structure using an artificial neural network, *Journal of Low Frequency Noise, Vibration and Active Control*, 41(1). DOI: 10.1177/14613484211038408

Bozorgnia, Y. & Tso, W.K. (1986). Inelastic Earthquake Response of Asymmetric Structures. *Journal of Structural Engineering*, 112(2). DOI: 10.1061/(ASCE)0733-9445(1986)112:2(383)

Candia, G., Macedo, J., Jaimes, M.A. & Magna-Verdugo, C. (2019). A New State-of-the-Art Platform for Probabilistic and Deterministic Seismic Hazard Assessment, *Seismological Research Letters*, 90(6), 2262-2275. DOI: 10.1785/0220190025

Cunha, B., Droz, C., Zine, A., Foulard, S. & Ichchou, M. (2023). A Review of Machine Learning Methods Applied to Structural Dynamics and Vibroacoustic, *Mechanical Systems and Signal Processing*, 200. DOI: 10.1016/j.ymssp.2023.110535

Instituto Nacional de Normalización, (1996). Diseño sísmico de edificios, NCh 433. Santiago: Instituto Nacional de Normalización.

Jansen, D. & Shah, S. (1997). Effect of Length on Compressive Strain Softening of Concrete, *Journal of Engineering Mechanics*, 123(1), 25-35. DOI: 10.1061/(ASCE)0733-9399(1997)123:1(25)

Kuang, Y., Jiang, X. & Jiang, N. (2018). Inelastic Parametric Analysis of Seismic Responses of Multistorey Bidirectional Eccentric Structure, *Shock and Vibration*. DOI: 10.1155/2018/7023205

Kuehn, N., Bozorgnia,Y., Campbell, K.W. & Gregor, N. (2020). Partially Non-Ergodic Ground-Motion Model for Subduction Regions using the NGA-Subduction Database, *Pacific Earthquake Engineering Research Center (PEER)*. DOI: 10.55461/NZZW1930

Lundberg, S.M., Erion, G., Chen, H. DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N. & Lee, S.-I. (2020). From local explanations to global understanding with explainable AI for trees, *Nat Mach Intell* 2, 56–67. DOI: 10.1038/s42256-019-0138-9

Montalva, G.A., Bastías, N. & Rodríguez-Marek, A. (2017). Ground-Motion Prediction Equation for the Chilean Subduction Zone, *Bulletin of the Seismological Society of America,* 107(2), 901-911. DOI: 10.1785/0120160221

Nakamura, H. & Higai, T. (2001). Compressive Fracture Energy and Fracture Zone Length of Concrete, Reston, VA: American Society of Civil Engineers.

Parker, G.A., Stewart, J.P., Boore, D.M., Atkinson, G.M. & Hassani, B. (2021). NGA-subduction global ground motion models with regional adjustment factors, *Earthquake Spectra*, 38(1). DOI: 10.1177/87552930211034889

Poulos, A., Monsalve, M., Zamora, N. & de la Llera, J.C. (2018). An Updated Recurrence Model for Chilean Subduction Seismicity and Statistical Validation of Its Poisson Nature, *Bulletin of the Seismological Society of America*, 109(1), 66-74. DOI: 10.1785/0120170160

Pozo, J., Hube, M. & Kurama, Y. (2022). Effective Nonlinear Simulations of RC Columns with Force-Based Elements, Journal of Earthquake Engineering, 27(2), 340-361. DOI: 10.1080/13632469.2021.2001395

Strasser, F.O., Arango, M.C. & Bommer, J.J. (2010). Scaling of the Source Dimensions of Interface and Intraslab Subduction-zone Earthquakes with Moment Magnitude, *Seismological Research Letters,* 81(6), 941-950. DOI: 10.1785/gssrl.81.6.941

Thai, H.-T. (2022). Machine learning for structural engineering: A state-of-the-art review, *Structures*, 38, 448-491. DOI: 10.1016/j.istruc.2022.02.003

Tso, W.K. & Ying, H. (1992). Lateral strength distribution specification to limit additional inelastic deformation of torsionally unbalanced structures, *Engineering Structures,* 14(4), 263-277. DOI: 10.1016/0141-0296(92)90014-H

University of California (2000) *OpenSees* (version 3.5) [software]. Available at: https://opensees.berkeley.edu/index.php

Yiu, C.-F., Chan, C.-M., Huang, M. & Li, G. (2013). Evaluation of lateral-torsional coupling in earthquake response of asymmetric multistory buildings, *The Structural Design of Tall and Special Buildings.* DOI: 10.1002/tal.1102

Yu, Y., Yao, H. & Liu, Y. (2020). Structural dynamics simulation using a novel physics-guided machine learning method, *Engineering Applications of Artificial Intelligence*, 96. DOI: 10.1016/j.engappai.2020.103947

Zhao, J.X., Zhang, J., Asano, A., Ohno, Y., Oouchi, T., Takahashi, T., Ogawa. H., Irikura, K., Thio, H.K., Somerville, P.G., Fukushima, Y. & Fukushima, Y. (2006). Attenuation Relations of Strong Ground Motion in Japan Using Site Classification Based on Predominant Period, *Bulletin of the Seismological Society of America*, 96(3), 898-913. DOI: 10.1785/0120050122