Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 7s (2025)

Data-Driven Seismic Analysis of Structures Using Machine Learning

Santha kumari Kambala [1], Dr. A. Suganthi [2], Dr. Damodar S. Hotkar [3], Siva Sankar Namani [4], Umasree Mariappan [5], Ghanta Rajeswari Mallika [6] Dr.R.Senthamil Selvan [7]

[1] Department of BS & H, Narasaraopet Engineering college (Autonomous), Narasaraopet. samsarasan123@gmail.com

[2] Department of Computer Applications, A.V.C. College of Engineering, Mannampandal, Mayiladuthurai. suganthikarthik09@gmail.com. ORCID: 0000-0002-6059-3050

[3] Department of Computer Science and Engineering, R.T.E. Society's Rural Engineering College, Hulkoti, Rural Engineering College, Hulkoti, Gadag, Karnataka- 582205. Email ID: hotkards22@gmail.com

[4] Department of CSE (AI & ML), G Narayanamma Institute of Technology and Science (GNITS), Shaikpet, Hyderabad. sivagnits@gmail.com

[5] Department of Information Technology, P.S.R. Engineering College, Sivakasi -626140, Tamilnadu, India. umasree24@gmail.com

[6] Department of Engineering Mathematics &Humanities, SRKR Engineering College, Bhimavaram rajeswarimallikaghanta@gmail.com

[7] Department of Electronics and Communication Engineering, Annamacharya Institute of Technology and Sciences, Tirupati, Andra Pradesh. selvasenthamil2614@gmail.com

Article History:

***Received:** 25-10-2024*

***Revised:** 10-12-2024*

***Accepted:** 18-12-2024*

Abstract:
Reinforced concrete shear walls are among the most important building structural components for supporting lateral loads. In spite of its significance, the inadequate safety margins of shear walls have been brought to light by post-earthquake reconnaissance and current experimental investigations. Current shear walls cannot have their failure modes quickly identified due to the absence of models based on mechanics and empirical data. To find out how shear walls fail depending on geometric configurations, material qualities, and reinforcing details, this research uses machine learning (ML), which has recently made some strides. Results from 395 experiments with shear walls of different geometric configurations make up the study's exhaustive database. In this research, the optimal prediction method was determined by evaluating eight machine learning methods, which included K Nearest Neighbours (KNN), Naive Bayes, Randomized Forest, XG Boost, Decision Tree, Ada Boost, Cat Boost and LightGBM. An exhaustive examination led to the proposal of a Random Forest based ML method in this research. When it comes to determining how shear walls break, the suggested approach is 87% accurate. According to the study, aspect ratios, bordering element reinforcement indices, and wall length to wall ratio of thickness are key factors in shear wall failure. Lastly, this research offers a data-driven categorization approach that is open-source and may be utilized by design firms worldwide. Additional experimental data that provide new insights may be easily included into the suggested method.

Keywords: Earthquake, Shear Walls, Machine Learning, Naive Bayes, Experimental Data.

1. Introduction

Reinforcing a building's shear walls is essential for keeping it upright in the event of an earthquake [1]. Existing shear walls aren't always up to snuff when it comes to ductile details. Experimental study shows that shear barriers collapse differently depending on geometry and material [2-4]. These approaches include flexion, shear forces. flexure-shear, sliding shear, and out-of-plane. Seismic vulnerability curves for infrastructure systems to be used in regional risk assessments have also been developed using data-driven methods [5]. Despite the progress made in using machine learning techniques to analyse damage and failure modes in infrastructure systems, data-driven methods to identifying failure modes of shear walls have not been studied [6-8]. Since shear structures are the most prevalent system that resists lateral loads, and since no model based on mechanics or empirical data exists to recognize the failure device of shear structures [9], this kind of research is crucial.

Furthermore, the majority of the current research on earthquake engineering sticks to more conventional ML techniques like Decision Tree, Discriminant analysis, Naïve Bayes, and Random Forest [10-12]. But nowadays, a lot of people use gradient boosting algorithms for machine learning since they are efficient, accurate, and easy to understand [13]. Theory supports these techniques by showing how to build strong predictors from weaker ones by greedily combining base predictors in an iterative fashion [14-16]. There have been no investigations on these approaches in earthquake engineering, despite their widespread use in other domains [17].

The purpose of this paper is to address this knowledge gap and make use of state-of-the-art ML techniques by investigating the potential uses of boosting models like AdaBoost, XGBoost, LightGBM, and CatBoost [18-20]. This is the first study using boosting methods for failure mode diagnosis, so be aware. Since the model's effectiveness is data-dependent, shear- wall mode of failure prediction requires cutting-edge machine learning.

The following were the specific aims of this study:

(1) To compile a data base of shear structures experimental results. Data about geometric features, reinforcing patterns, and materials is stored in the database [21]. The data-driven method to evaluating buildings' seismic performance benefits from an open-source data.

(2) Figure out which machine learning models are best for predicting which shear wall failure situations to use [22]. The classification model in this study is built utilizing a number of ML methods, such as Naive Bayes, K Nearest Neighbours, Decision Tree, Randomized Forest, AdaBoost, Light GBM, XG Boost and Cat Boost [23].

(3) Determine the input factors' significance in determining the shear wall failure mode. In order to comprehend the seismically activity of shearing wall structures [24], researchers might benefit from this identification by designing appropriate experimental studies.

(4) Develop a world-class, freely-available data-driven classification model for shear wall failure mode prediction that design firms may utilize [25]. Since the open-source model may easily absorb new experimental data, it can be used to continuously enhance the categorization model.

The following is an overview of failure modes and available methods for forecasting them. Next, the experimental database is analysed for relevant results. The next sections describe machine learning methods and methodologies used to create the most accurate prediction method. The primary results of the research are presented in the conclusion.

2. Shear Wall Failure Mechanism

The shear forces, axial loads, and bending moments that concrete- reinforced shear walls experience are typical in construction systems. The aspect ratio, which is the length of the shear span divided by the length of the wall, is a common way to classify shear walls as either skinny (tall, high-rise) or short (low-rise). When it comes to ductile failure mechanisms, flexural yielding toward the base is the most common in thin walls. As a result of their shape, squat walls are more prone to shear-controlled failure mechanisms, which may be defined as an unexpected loss of stiffness and strength in the face of seismic pressures.

Aspect ratios less than 1.5 are deemed short or squat, whereas those more than 3.0 are deemed slim according to ASCE 41-06. Walls behave in a way that is influenced by shearing and flexure if their ratio is between 1.5 and 3.0. According to FEMA 306, shear walls that are sufficiently tall and have been thoughtfully constructed are more likely to experience ductile flexural failure. Concrete crushing or longitudinal reinforcement fracture at the plastic hinge zone causes flexural failure, as seen in Figure 1(a). Particularly for aspect ratios below 1.0, flexural failure is uncommon in short shear walls. There is a possibility of shear failure and flexural failure in these walls due to the peculiarities of their reinforcing.

Figure 1(b)–(d) shows the three cases of shear failure that researchers identified for squat shear walls: sliding shear failure, diagonal compression failure and tension failure based on diagonal. So, to conclude, this is how they described various failure scenarios. A wall with inadequate horizontal shear reinforcement is more prone to diagonal tension failure, which manifests as one or more fractures running diagonally from corner to corner, as seen in Figure 1(b). In the presence of massive and sufficient horizontal shear reinforcement, walls may experience diagonal compression failure, as seen in Figure 1(c). This kind of failure causes the concrete to collapse under diagonal compression, resulting in extensive fracture patterns. Because they can endure larger flexural strengths, shear walls with boundary features increase the shear demand on the web, making them more likely to fail diagonally under compression than shear walls with rectangular cross sections. One way to prevent diagonal tension or compression failure is to reduce the nominal shear stress and ensure that there is enough horizontal shear reinforcement. Large fractures at the base of the wall or a narrow band of crushed concrete and buckling rebars at the base of the wall (Figure 1(d)) following substantial bending in the flexural reinforcement are the two main causes of sliding shear failure. The present article does not address the out-of-plane failure scenario, requiring more research to forecast it using data-driven methods.

(a)Flexural Failure (b)Diagonal Tension Failure

(c)Diagonal Compression Failure (d) Sliding Shear Failure

Figure 1 Three Cases of Shear Failure

3. Database For Experimental Purposes

3.1 Explanation of Reinforced Concrete Shearing Wall Experimental Database

The 395 reinforced shearing walls, some with rectangular and some without, that make up the experimental database in this research all have one storey and one bay. While the majority of the specimens included in this collection were culled from two preexisting databases, the writers also compiled some new specimens based on previous experimental test findings.

The objective of the research is to create a ML method that can identify failure modes in a broad range of scenarios; the database also includes specimens that did not undergo axial strain. One hundred thirty-nine examples have square cross sections, ninety-six have rectangular cross sections, and sixty-one have flanged cross sections (Figure 2). When compared to shear walls with rectangular cross sections, those containing boundary elements, such as barbell or flanged sections, often have a greater peak shear strength. The boundary elements' dual roles as web confinement and reinforcement contribute to this. The cross sections of all the shear walls are symmetrical.

Furthermore, all of the samples have straight and deformed reinforcement in addition to continuous longitudinal reinforcement free of lap splices.

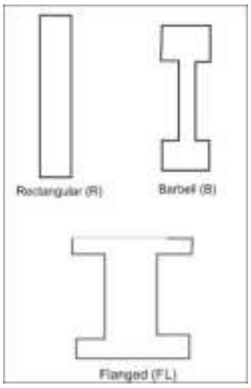

Figure 2 Cross Section Shape of Walls

A total of 153 specimens, 96 specimens with flexure-shear failure, 123 specimens with shear failure, and 23 specimens with sliding shear failure are included in the database. Visualized in the spatial distributions of the design elements and the structures of the wall cross-sections are shown in Figure 3 includes barbell, rectangle, and flanged.

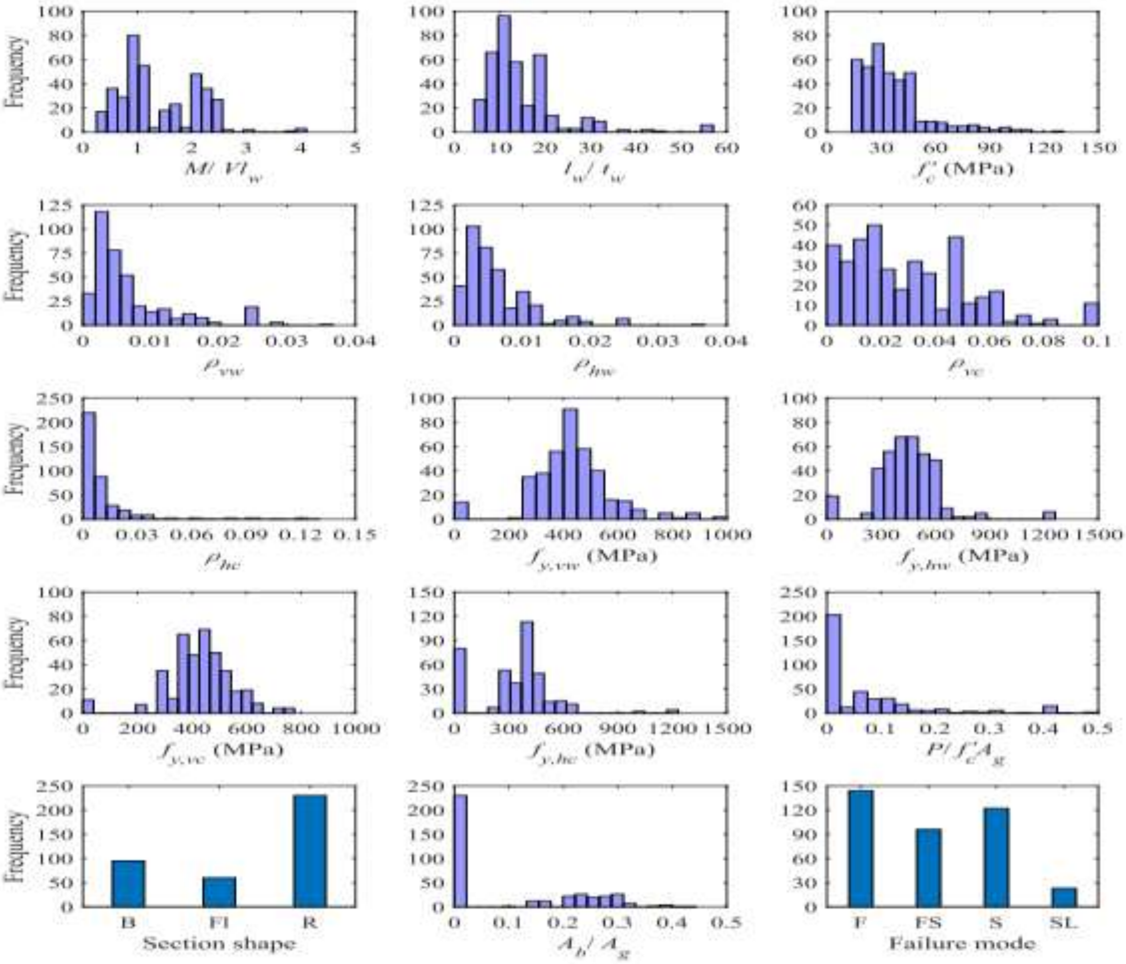

Figure 3 Distributions of Design Elements and Structures of the Wall Cross-Sections

3.2 Input Parameters for Failure Mode Detection

Drawing on earlier work on identifying failure modes in the remaining components, this study establishes input parameters using the design parameters mentioned earlier. The rationale for this is the dearth of research on the topic of probabilistic failure mode categorization in concrete- reinforced shear wall structures.

Input parameters that do not have dimensions provide fitted prediction models with dimensionless coefficients that do not rely on changes to the system of units. To make things easier for prospective users, this study's input parameters are all dimensionless according to the design specifications mentioned before. A few of the design parameters are considered input parameters. Shear wall configuration is accounted for by the first three factors.

The reinforcement index, formerly utilized to classify the failure approach of concrete members, is introduced to take into account design characteristics pertaining to the reinforcement, including the ratio of reinforcement and yield strength. Multiplying the reinforcement ratio by the yield strength that is normally expressed as a percentage of the concrete's compression strength yields the reinforcement index. This section's reinforcing index may be thought of as a strength ratio, which takes into account both the amount and the strength of the materials used.

This research utilizes four reinforcement indices: web vertical, web horizontal, boundary element vertical, and boundary element horizontal. Shear wall cross section shape is also considered as an input component. This article examines three cross section shapes: rectangle, barbell, as well as flanged. See Figure 4 for input parameter distribution based on design parameters. The input parameters were previously displayed in Figure 3; thus, they are not shown in Figure 4. Input parameter distributions clearly do not follow any kind of regular pattern.

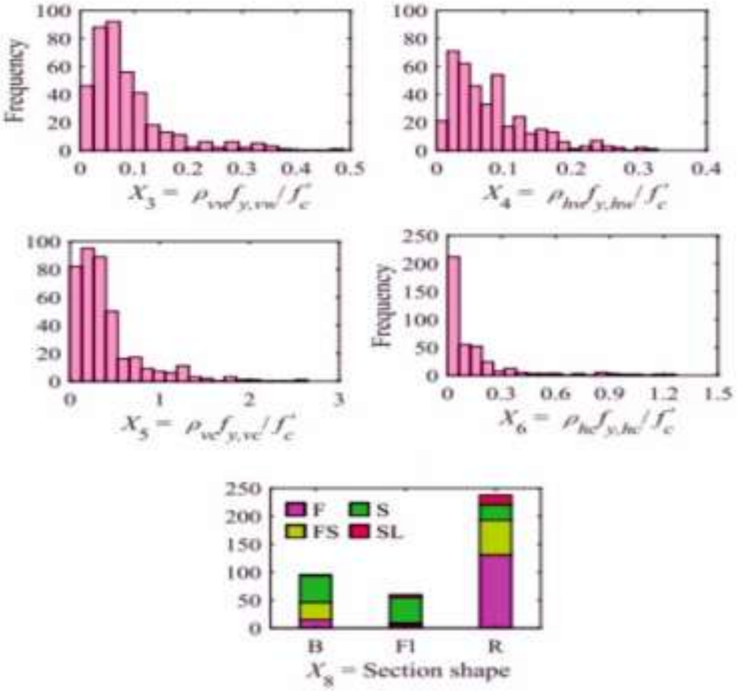

Figure 4 Input Parameter Distributions

3.3 Selection of Response Parameter Failure Modes

Researchers used the five classes of failure modes described in Section 2 to construct models for predicting the stability and deformation capacity of reinforced concrete walls: flexure, diagonal compression as well, sliding shear, and the unique scenario of squatting walls failing in shear. Most of the samples in their database that failed in shear after flexural yielding also failed in tension and compression shear. A number of the previously cited works, however, failed to identify a particular flexure shear failure mechanism.

This study categorizes the many ways in which reinforced concrete walls might fail according to the model code: flexural, flexure-shear, shear forces. and slide shear. Figure 3 shows that there were 153 flexural failure specimens, 96 flexure-shear failure specimens, 123 shear failure specimens, and 23 sliding shear failure specimens in the database.

4. Overview of ML methods

This research employed eight different ML models to determine which failure mode classification method was the most effective: First, Naive Bayes (NB) second, K-Nearest Neighbours (KNN) third, Decision Tree (DT) fourth, Random Forest (RF) fifth, Ada Boost (AB) sixth, XG Boost (XB) seventh, Light GBM (LGBM) and eighth, Cat Boost. While structural engineers have investigated algorithms like Naive Bayes, K-Nearest Neighbours, Decision Trees, and randomly generated forests, they have not investigated advanced algorithms like AdaBoost, XG Boost, Light GBM, and Cat Boost.

4.1 Nive Bayes

The assumption that the impact of the value of a characteristic on a specific period is unrelated to other attributes value is what makes learning easier for a Naive Bayes classifier. Based on Bayes' theorem, the classifier:

$$p_{r.}(x) = p_r(Y = f|X = x) = \frac{\pi_f \int f(x)}{\sum_{i=1}^{K} \pi_i f_i(x)} \qquad (1)$$

in wherever f is the previous possibility that an opinion from the f-th class is drawn at random, and $\int f(x)$ is the ratio of the density functional that an opinion from the class f is drawn. In this research, we assume that $\int f(x)$ follows a normal distribution and calculate f by finding the percentage of training observations that belong to the f th class.

4.2 K-Nearest Neighbours

The decision boundaries between failure modes are not assumed by the non-parametric machine learning algorithm known as KNN. This approach classifies samples as belonging to a failure mode if the K most comparable examples in the feature space are all members of the same mode. Then, in failure method f, the conditional possibility of x is approximated as:

$$p_r(x) = p_r(Y = f|X = x)\frac{1}{k}\sum_{i \in N_K} I(y_i = f) \qquad (2)$$

where N_K stands for the K training data points that are most closely related to the opinion, f.

4.3 Decision Tree

A non-parametric categorization approach, Decision Tree involves breaking the classification issue into a series of basic judgments, each of which is based on a single or many input characteristics. Building the tree and then pruning it are the two main components of Decision Tree modelling.

In order to construct a tree, the space of the training is divided into non-overlapping areas using the Gini Indices. The tree that was built up during the construction phase can have a lot of branches, therefore the next step is to prune it so it doesn't become overfit. Using the 10-fold validation approach, this work estimates the pruning parameter and adopts a cost complexity pruning.

4.4 Randomized Forest

In a Randomly Forest classifier, which relies on a collection of classifiers, to each classifier is created by a randomly sampled path that is separate from the initial vector. Two crucial approaches are included in Random Forest: out-of-bag estimates and random feature subspace. The first one makes tree building a lot quicker, while the second one opens the door to the idea of ranking the input features' significance. The sources provide a more in-depth explanation of Random Forest.

4.5 Methods for Boosting: AdaBoost, XG Boost, Light GBM, and CatBoost

The boosting approaches, which include AdaBoost, XGBoost, LightGBM, and CatBoost, enhance a model's performance by merging many weak classifiers into a single strong one. The idea is to choose the weak classifiers in a manner that makes them much stronger when used together. All the observations are given equal weight in the first phase of AdaBoost. Retraining the model is the next stage, after which the misclassified observations are given greater weight than the properly classified ones.

Consequently, the prior learners' weighted categorization accuracy serves as the basis for the learners' training. Scientists came up with a new way to boost that they term gradient boosting.

This method uses regression on an expression of the loss function's gradient vector from the previous iteration. A few examples of gradient methods include Light GBM, XG Boost and Cat Boost. All of them start with DT as their weak student and utilize gradient boosting to repeatedly fit a succession of these trees. By decreasing the prior model's misclassification error, XG Boost fits the updated method.

The sequential model training that XG Boost employs, however, makes it a sluggish implementation. Rather of building the model from the ground up, Light GBM generates it leaf-wise using decision tree methods. Accurate, more complicated trees are produced by generating them leaf by leaf. Cat Boost is a technique that effectively deals with input parameters that include categorical characteristics and makes use of them during training instead of during preprocessing.

5. Failure Mode Identification with the Use of Machine Learning Methods

Reinforced concrete shear wall failure modes were extracted from the collected information using the machine learning methods detailed in the preceding section. The writers used the machine learning models by transforming the shear wall's material, structural inequality, and geometric attributes into the nine input variables outlined in Section 3.2. The models discussed in Section 4 were trained using machine learning algorithms that are available in Section 7, which are built utilizing the open python program scikit learn.

The prediction model is constructed using a training set that contains 71% of the data from the supplementary material section, and its performance is evaluated using a test set that contains the other 29 percent of the data. By randomly splitting the entire collection into a set to be trained and a test set, we were able to gauge the model's results on the former, which in turn indicated how well it handled the unknown data. Machine learning methods are constructed using the methods outlined in Section 4 utilizing 70% of the collected data.

In figure 8 displays that (noted) one hundred percent is the total of all the numbers that are beyond the bars that are vertical. Using a grid search method, we first determined the best Random Forest parameters, and then we determined the comparative relevance of every constraint by observing the fluctuation in the out-of-bag (OOB) error.

The essential elements that influence the failure mode of shear walls are the aspect ratio, border component indices based on reinforcement, and the wall-to-wall thickness percentage (Figure 5). Compared to other characteristics, the cross-section shape is found to have a less impact on the failure mode.

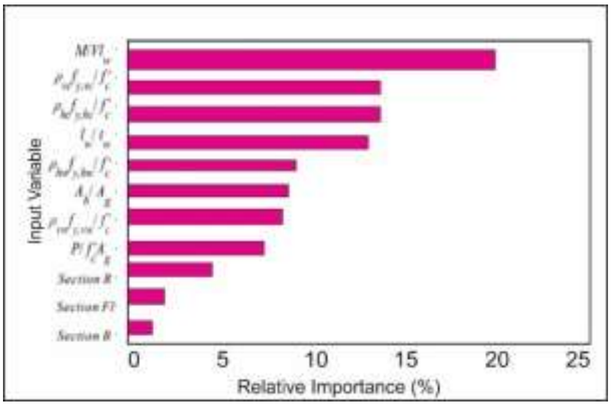

Figure 5 Total of all the Numbers are Beyond the Vertical Bars

6. Conclusion

As part of a system to resist lateral forces, shear structures are used to withstand wind and seismic forces. Shearing walls may fail in shear, flexion, flexure shearing or sliding, depending on the geometry, materials, and force. Despite a wealth of experimental and seismic reconnaissance data on wall performance under various configurations, no empirical or technique based on mechanics has been developed to forecast how shear walls would break. In order to quickly evaluate damage, seismic risk, and retrofitting decision methods, an intuitive failure mode forecast is required. The potential of AI and ML to detect failure modes in concrete shear walls is investigated in this research.

As a first step, the writers compile all of the available experiments into a comprehensive database. A total of 395 reinforced concrete shear walls, each having one bay and either rectangular or non-rectangular sections, are contained within the data. Out of all the samplings recorded in the database, 153 had flexural failure, 96 had flexure-shear failure, 123 had shear failure, and 23 had sliding shear failure. Drawing on previous research, the authors developed a set of 10 input parameters that accurately represent shear wall geometry, material characteristics, and reinforcing details. Two subsets, the training set and the test set, make up the whole dataset. The prediction model is built using the training set, and its performance is assessed using the test set.

In this research, eight computer vision models were examined: NB, K-Nearest Neighbours (KNN), Randomized Forest, DT, Ada Boost, XG Boost, Light GBM, and Cat Boost. International precision, accuracy, and recall were the three criteria used to assess the model's performance. According to the study's findings, Random Forest outperformed Cat Boost and XG Boost in terms of training set accuracy. Predicting the flexure-shear mode is really rather challenging. When it came to determining the mechanism of failure for the test set, the suggested Random Forest model achieved a recall of 71% and a precision of 85%. Research into Random Forest led to the conclusion that the aspect ratio, boundaries component reinforcement indexes, and wall length to wall thickness percentage are the most important input constraints controlling the shear wall failure method.

This research provides a quick, data-driven approach to design offices throughout the globe by demonstrating that machine learning methods can anticipate the failure method for shearing walls. The open-source model may be continuously improved by adding additional experimental data. Researchers may use it to prepare studies for the future as well. The method can detect the mechanism of failure, but it has not yet determined the root cause. Damage progression tracking and improved failure mode diagnosis need more investigation.

Reference:

[1] Zhang, Tianyang, et al. "Seismic response prediction of a damped structure based on data-driven machine learning methods." Engineering Structures 301 (2024): 117264.

[2] Luo, Huan, and Stephanie German Paal. "Data-driven seismic response prediction of structural components." Earthquake Spectra 38.2 (2022): 1382-1416.

[3] Xu, Ji-Gang, et al. "Seismic performance assessment of corroded RC columns based on data-driven machine-learning approach." Engineering Structures 255 (2022): 113936.

[4] Nguyen, Hoang D., et al. "Development of data-driven models to predict seismic drift response of RC wall structures: An application of deep neural networks." Soil Dynamics and Earthquake Engineering 186 (2024): 108952.

[5] Harati, Mojtaba, and John W. van de Lindt. "Data-driven machine learning for multi-hazard fragility surfaces in seismic resilience analysis." Computer-Aided Civil and Infrastructure Engineering (2024).

[6] Zhang, Ruiyang, Yang Liu, and Hao Sun. "Physics-guided convolutional neural network (PhyCNN) for data-driven seismic response modeling." Engineering Structures 215 (2020): 110704.

[7] Kumar, T. Rajasanthosh, G. Laxmaiah, and S. Solomon Raj. "A Framework of Intelligent Manufacturing Process by Integrating Various Function." AI-Driven IoT Systems for Industry 4.0. CRC Press 241-254.

[8] Gao, Zhaoqi, et al. "Building large-scale density model via a deep-learning-based data-driven method." Geophysics 86.1 (2021): M1-M15.

[9] Patro, Pramoda, et al. "A hybrid approach estimates the real-time health state of a bearing by accelerated degradation tests, Machine learning." 2021 Second International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE). IEEE, 2021.

[10] Kazemi, Farzin, Neda Asgarkhani, and Robert Jankowski. "Machine learning-based seismic response and performance assessment of reinforced concrete buildings." Archives of Civil and Mechanical Engineering 23.2 (2023): 94.

[11] Liao, Yangyang, et al. "Response prediction for linear and nonlinear structures based on data-driven deep learning." Applied Sciences 13.10 (2023): 5918.

[12] Shabbir, Khurram, Mohamed Noureldin, and Sung-Han Sim. "Data-driven model for seismic assessment, design, and retrofit of structures using explainable artificial intelligence." Computer-Aided Civil and Infrastructure Engineering (2024).

[13] Kim, Sungil, et al. "Data-driven signal–noise classification for micro seismic data using machine learning." Energies 14.5 (2021): 1499.

[14] Chen, Peng-Yu, and Xingquan Guan. "A multi-source data-driven approach for evaluating the seismic response of non-ductile reinforced concrete moment frames." Engineering Structures 278 (2023): 115452.

[15] Ghimire, Subash. Data-driven-based models for seismic damage assessment of buildings: using machine learning at a regional scale and in-situ structural monitoring at building scale. Diss. Université Grenoble Alpes [2020-....], 2023.

[16] Todorov, Borislav, and AHM Muntasir Billah. "Machine learning driven seismic performance limit state identification for performance-based seismic design of bridge piers." Engineering structures 255 (2022): 113919.

[17] R.Senthamil Selvan "INTEGRATING THE BIGDATA ANALYTICS AND DEEP LEARNING analysis human movement to improve the sports" by 2023 International Conference on New Frontiers in Communication, Automation, Management and Security (ICCAMS), ISSN: 0018-9219, E-ISSN: 1558-2256, December 2023. 10.1109/AECE59614.2023.10428236.

[18] R.Senthamil Selvan "MULTI OBJECTIVES EVALUATOR MODEL DEVELOPMENT FOR ANALYZE THE CUSTOMER BEHAVIOUS" by 2023 International Conference on New Frontiers in Communication, Automation, Management and Security (ICCAMS), ISSN:0018-9219,E-ISSN:1558-2256,December 2023. 10.1109/AECE59614.2023.10428189.

[19] Mokhtari, Fardad, and Ali Imanpour. "Hybrid data-driven and physics-based simulation technique for seismic analysis of steel structural systems." Computers & Structures 295 (2024): 107286.

[20] Yin, Huichao, et al. "A deep learning-based data-driven approach for predicting mining water inrush from coal seam floor using micro-seismic monitoring data." IEEE Transactions on Geoscience and Remote Sensing (2023).

[21] Zhang, Tianyang, et al. "Study on the evolution of dynamic characteristics and seismic damage of a self-centering concrete structure based on data-driven methods." Engineering Structures 316 (2024): 118529.

[22] Xu, Ji-Gang, et al. "Data-driven rapid damage evaluation for life-cycle seismic assessment of regional reinforced concrete bridges." Earthquake Engineering & Structural Dynamics 51.11 (2022): 2730-2751.

[23] Sadhukhan, Bikash, et al. "Climatic and seismic data-driven deep learning model for earthquake magnitude prediction." Frontiers in Earth Science 11 (2023): 1082832.

[24] Meng, Shiqiao, et al. "SeisGPT: A Physics-Informed Data-Driven Large Model for Real-Time Seismic Response Prediction." arXiv preprint arXiv:2410.20186 (2024).

[25] Dang-Vu, Hoang, et al. "Frequency-based data-driven surrogate model for efficient prediction of irregular structure's seismic responses." Journal of Earthquake Engineering 26.14 (2022): 7319-7336.

[26] Xiong, Qingsong, et al. "Machine learning-driven seismic failure mode identification of reinforced concrete shear walls based on PCA feature extraction." Structures. Vol. 44. Elsevier, 2022.

[27] Wang, Tianyu, et al. "From model-driven to data-driven: A review of hysteresis modeling in structural and mechanical systems." Mechanical Systems and Signal Processing 204 (2023): 110785.