Czech Journal of Civil Engineering

2025
Volume 11
Issue 01

Data-Driven Prediction of Seismic Drift Responses Using DNNs and XAI

Ing. Mayaz Uddin Gazi[1]; Ing. Md. Lahin Kabir[1*]; Ing. Md. Rayhan Chowdhury[1]; Asst. Prof. Sibbir Ahmed[1]; W.M Ing Md. Titumir Hsaan[1];

[1] *Sylhet Engineering College, Airport Road, Department of Civil Engineering, Sylhet, 3100, Bangladesh*

* *corresponding autor:2019333537@student.sec.ac.bd*

ORGINAL SCIENTIFIC PAPER

ABSTRACT

Structural drift is an important way to measure how well a building can withstand earthquakes. Too much drift can cause serious damage or even collapse, especially in reinforced concrete frames. For seismic risk assessment and resilient design in tectonically active, resource-limited areas, it is important to be able to accurately predict drift. This study offers a clear, data-driven framework that combines nonlinear time-history analyses in OpenSees with machine learning and explainable AI techniques. For supervised regression, a synthetic dataset of different types of buildings and ground motion scenarios for Bangladesh was used. SHAP and LIME were used to test Random Forest and deep neural network models. Random Forest was more accurate (R^2 = 0.897, MAE = 0.491). The framework offers a comprehensible and effective methodology for seismic risk assessment and urban resilience planning.

Key words: *Drift; Time Period; Ground Motion; OpenSees; Random Forest; SHAPE Analysis.*

1 INTRODUCTION

In Bangladesh, where major cities like Dhaka, Chattogram, and Sylhet are extremely vulnerable due to their proximity to active fault lines like the Dauki, Jamuna, Chittagong-Myanmar, Sylhet, and Tripura faults, seismic risk assessment is crucial for local and international safety. One of the deadliest natural disasters in the world, earthquakes cause a great deal of death, displacement, and economic disruption. Approximately 747,000 people died as a result of earthquakes and tsunamis between 1998 and 2017, which accounted for more than half of all disaster-related deaths worldwide. Particularly in tectonically active zones, the frequency and magnitude of seismic events have increased in recent decades, underscoring the necessity of efficient mitigation and preparation measures.

According to the United Nations Office for Disaster Risk Reduction (UNDRR), between 1998 and 2017, earthquakes and related tsunamis were responsible for 56 % of all disaster-related fatalities worldwide, resulting in approximately 747,234 deaths during this period [1]. The economic consequences are equally severe, with direct disaster-related losses—including those from earthquakes–reaching hundreds of billions of dollars annually [1][2]. In recent years, both the frequency and intensity of seismic events have exhibited a notable increase. Data from geological observatories indicate that between 2015 and 2025, the number of earthquakes with magnitudes of 6.0 or greater rose by more than 30 % compared to the previous decade [3][4]. This trend is especially evident in tectonically active regions such as the Pacific Ring of Fire, where densely populated urban centers remain at heightened risk of catastrophic seismic events [3][4]. Despite significant advancements in seismological research, the ability to predict the exact timing, location, and magnitude of earthquakes remains beyond current scientific capabilities.

The U.S. Geological Survey (USGS) and leading researchers emphasize that although long-term probabilistic models are available, the chaotic behavior of fault systems and the absence of reliable precursory signals render short-term earthquake prediction scientifically unfeasible [5]. As a result, seismic events often occur with little or no warning, underscoring the urgent need for comprehensive preparedness and risk-mitigation strategies. Preparedness efforts—including the rapid seismic assessment of building inventories, the implementation of resilient structural design practices, and the promotion of widespread public education—are essential to minimizing both human casualties and economic losses [6]. In addition, retrofitting vulnerable infrastructure and the strict enforcement of seismic building codes can substantially reduce the likelihood of structural collapse, which remains the principal cause of fatalities during seismic events [7]. Bangladesh faces an elevated risk of devastating earthquakes due to its unique geographic location at the convergence of the Indian, Eurasian, and Burma tectonic plates—a region commonly referred to as the Himalayan collision zone [8][9].

2 METHODOLOGY

This chapter details the methodology adopted to achieve the research objectives. It begins with an outline of the dataset generation process, including nonlinear dynamic analyses performed using OpenSees to simulate seismic responses of building models. Data preprocessing steps, feature selection strategies, and handling of missing values are described in detail. The supervised machine learning models employed for seismic drift prediction are introduced, followed by an explanation of model training, validation, and performance evaluation procedures. The chapter also covers the application of explainable AI (XAI) techniques to interpret model outputs.

2.1 Dataset Creation and Construction Workflow

The foundation of this research is the creation and meticulous preparation of a high-fidelity, simulation-based dataset tailored for advanced machine learning applications in earthquake engineering. Unlike datasets that are simply assembled from existing measurements, the dataset developed in this study was systematically engineered through a two-stage process designed to maximize both predictive power and physical interpretability. The process began with the development of a synthetic input dataset for OpenSees simulations. This dataset was programmatically structured to represent a broad spectrum of reinforced concrete (RC) frame buildings subjected to diverse seismic scenarios. Each record encoded critical structural parameters—including the number of stories, story height, bay width, and the cross-sectional properties of beams and columns—along with essential material characteristics such as concrete compressive strength fc′and steel yield strength fy. These parameters were systematically varied to ensure adequate representation across both low-rise and high-rise RC frames, thereby enabling robust generalization and meaningful analysis of structure–seismic interactions. The complete workflow of the study is presented in.

doi.org/10.51704/cjce.2025.vol11.iss1.pp32-38
ISSN (online) 2336-7148

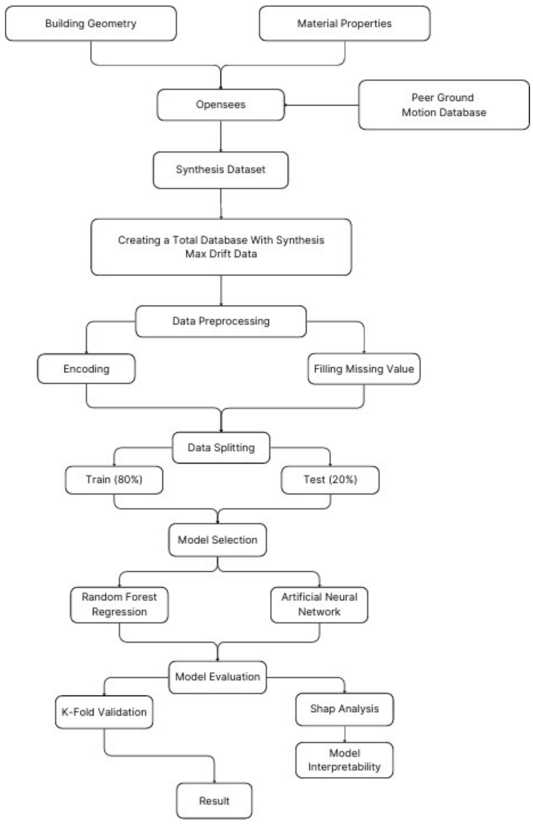

Fig 1 Total Study Workflow Diagram

2.2 Machine Learning Model Development and Explainability

For this study, we chose two models: one is a random forest tree-based model, and the other is a Neural network-based model.

2.2.1 Random Forest Regression Model

An ensemble-based regression model (Random Forest) was developed by aggregating the predictions of multiple decorrelated decision trees. Each tree was constructed using bootstrap sampling of the dataset and random feature selection at each split, with the ensemble output obtained by averaging the predictions from all trees. The model minimizes the following loss function for each tree:

follows:

$$L = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \tag{3.4}$$

where y_i is the observed drift and $\hat{y}_i$ is the predicted drift. The architecture of the model is showing as an example in Figure 2.

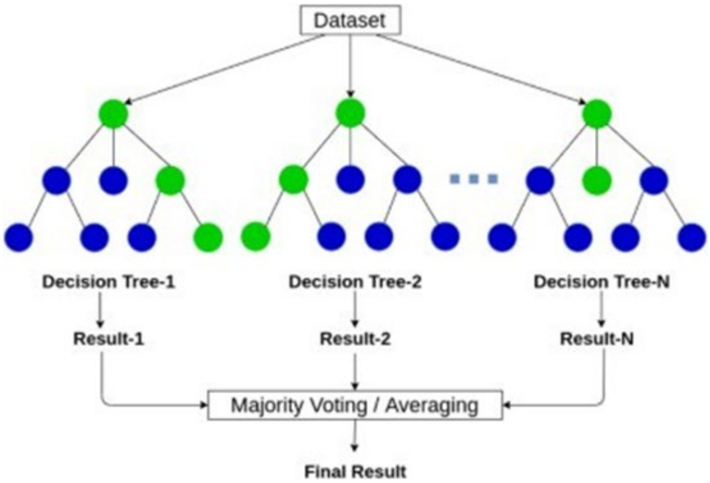

Fig 2 Random Forest Architecture

2.2.2 Deep Neural Networks (DNNs)

An artificial neural network (ANN) regression model was developed to capture the complex, nonlinear relationships present in the structural engineering dataset. The model was implemented using the Keras API within the TensorFlow framework, employing a deep, fully connected feedforward architecture. Integrated regularization techniques were applied to enhance generalization and mitigate overfitting. The architecture of the model is illustrated as Fig 3.

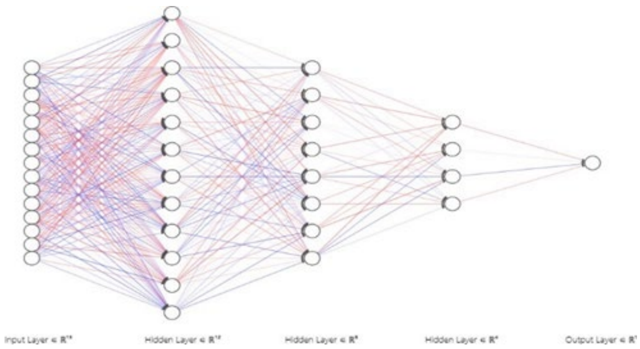

Fig 3 Deep neural Network Architecture

3 RESULTS

This chapter presents findings from applying machine learning models to the seismic drift dataset, focusing on both predictive accuracy and interpretability. The Random Forest (RF) regressor was selected for its robustness, flexibility, and ability to model complex, nonlinear relationships without extensive feature engineering. As an ensemble of decision trees built from bootstrap samples and random feature subsets, RF mitigates overfitting and provides stable predictions across diverse input scenarios. Its capacity to capture nonlinear interactions among structural, ground motion, and site features is particularly valuable in earthquake engineering, where responses are rarely linear. RF is also largely insensitive to feature scaling, simplifying preprocessing for multidisciplinary datasets. The model was trained on 80% of the data and evaluated on the remaining 20%, with hyperparameters optimized via grid search. Performance metrics – R^2 = 0.897, MAE = 0.491 – demonstrate high predictive accuracy. These results validate RF as a reliable baseline and a foundation for future, more advanced data-driven seismic drift modeling.

3.1 Model Comparision

Tab. 1 Test–set performance comparison for MaxDrift prediction. Best scores are in bold

Metric	RF	DNN
R^2	**0.897**	0.868
RMSE	**1.129**	1.275
MAE	**0.491**	0.6999

The Random Forest (RF) model outperforms the Deep Neural Network (DNN) across all evaluated metrics (Table 4.3). RF achieves a higher coefficient of determination (R^2 = 0.897) compared to the DNN (R^2 = 0.868), indicating that it explains a greater proportion of variance in the observed MaxDrift values. In addition, RF exhibits lower prediction errors, with RMSE reduced from 1.275 (DNN) to 1.129 and MAE reduced from 0.6999 to 0.491. These results demonstrate that RF not only provides more accurate predictions on average (lower MAE) but also ensures more stable and consistent performance (lower RMSE). The superior performance of RF is attributed to its ensemble-based structure, which effectively captures complex feature interactions while mitigating overfitting, particularly in small-to-moderate datasets. This makes RF a highly reliable choice for seismic drift prediction in data-limited scenarios.

4 CONCLUSION

This chapter summarizes the key outcomes of the study, emphasizing contributions to seismic drift prediction and interpretable machine learning in earthquake engineering. Research limitations are acknowledged, and directions for future studies are proposed. The chapter concludes with practical recommendations for policymakers, engineers, and stakeholders aiming to improve seismic safety and resilience in Bangladesh and other high-risk regions.

Seismic risk continues to pose a significant challenge in civil engineering, especially in rapidly urbanizing and tectonically vulnerable regions like Bangladesh. This thesis addresses a critical aspect of this challenge: the accurate, efficient, and interpretable prediction of seismic drift responses in reinforced concrete (RC) buildings using advanced data-driven approaches. By integrating physics-based simulations, machine learning, and explainable artificial intelligence (XAI), the research contributes to both the methodological foundations and the practical toolkit for performance-based earthquake engineering.

Acknowledgment

The author gratefully acknowledges the invaluables supervision, guidance, and encouragment provided by Md. Titumir Hasan and Sibbir Ahmed, whose expertise in Structural Engineering and commitment to research excellence were instrumental throughout every phase of this thesis. The constructive feedback and continuous support from the academic staff of the Department of Civil Engineering at Sylhet Engineerig College are sincerely appereciated.

Special thanks are extended to the developers and the user community of OpenSees for proivding the sumulation tools fundamental to this study, and to the PEER NGA database for access to ground motion records.

The author is indebted to friends and colleages for their encouragement, motivation, and ciritical discusions, which enriched both the research and writing process. Lastly heartfelt gratitude is offered to the authors family for their patience, understaning, and unwavering support throughout the entire journey.

Funding

The article is not financed from external sources.

Literature Review

This chapter provides a comprehensive review of existing literature on seismic risk assessment, vulnerability analysis, and recent advances in data-driven and machine learning approaches for earthquake engineering. It examines key developments in traditional methods, including Rapid Visual Screening (RVS) and the Seismic Vulnerability Index (SVI), alongside contemporary research utilizing artificial intelligence and explainable models. Special emphasis is placed on studies relevant to the Bangladeshi context, highlighting critical knowledge gaps and motivating the need for the present research.

[1] UNITED NATIONS OFFICE FOR DISASTER RISK REDUCTION. (2020). *UN 20-Year Review: Earthquakes and Tsunamis Kill More People While Climate Change is Driving Up Disaster Numbers*. [Online]. Available: https://www.undrr.org [Accessed: Jul. 6, 2025].

[2] UNESCO. (2024). *IPRED Knowledge – International Platform for Reducing Earthquake Disasters*. [Online]. Available: https://en.unesco.org/ipred [Accessed: Jul. 6, 2025].

[3] EVRIM AĞACI. (2024). *Surge in earthquake activity sparks global alarm*. [Online]. Available: https://evrimagaci.org [Accessed: Jul. 6, 2025].

[4] THOMSON REUTERS FOUNDATION. (2013). *Bangladesh not ready for major earthquake, experts say*. 2013. [Online]. Available: https://news.trust.org [Accessed: Jul. 6, 2025].

[5] U.S. GEOLOGICAL SURVEY. (2024) *Can you predict earthquakes?*. [Online]. Available: https://www.usgs.gov [Accessed: Jul. 6, 2025].

[6] APPLIED TECHNOLOGY COUNCIL. (2014). *Review of Past Performance and Further Development of Modeling Techniques for Collapse Assessment of Existing Reinforced Concrete Buildings*, Tech. Rep. ATC-95. Redwood City, CA, USA: Applied Technology Council. [Online]. Available: https://www.nist.gov [Accessed: Jul. 6, 2025].

[7] GLOBAL FACILITY FOR DISASTER REDUCTION AND RECOVERY (GFDRR). (2016). *Disaster Risk Profile: Bangladesh*. [Online]. Available:

https://www.gfdrr.org/sites/default/files/publication/Bangladesh%20v4.pdf [Accessed: Jul. 6, 2025].

[8] M. AHMED, D. BOSU, R. I. BABUL, A. A. MAHIN, AND M. A. SAYEED. (2024). Seismic vulnerability assessment of buildings in Sylhet city using rapid visual screening. SSRN. [Online]. Available: https://ssrn.com/abstract=4967136 [Accessed: Jul. 6, 2025].

[9] DHAKA TRIBUNE. (2023). Sylhet quake: Expert calls for extensive survey, third-party vetting to reduce losses, *Dhaka Tribune*. [Online]. Available: https://www.dhakatribune.com [Accessed: Jul. 6, 2025].

doi.org/10.51704/cjce.2025.vol11.iss1.pp32-38
ISSN (online) 2336-7148