2023 IEEE Symposium Series on Computational Intelligence (SSCI)
Mexico City, Mexico. December 5-8, 2023

Structural Analysis of the Mexico-Toluca Interurban Train with Data Science

Osmar David Arellano-Magdaleno*, Rosa María Valdovinos-Rosas*,
Angélica Guzmán-Ponce†, and David Joaquín Delgado-Hernández *
*Faculty of Engineering, Autonomous University of the Mexico State, Toluca, Mexico
{oarellanom001,rvaldovinosr,david.delgado}@uaemex.mx
† Institute of New Imaging Technologies, Department of Computer Languages and Systems
Universitat Jaume I, 12071 Castelló de la Plana, Spain. aguzman@uji.es

***Abstract*—As part of an initiative to mitigate transportation congestion of more than 230,000 daily travellers between Mexico City and Mexico State, the Mexican Federal Government, through the Secretariat of Communications and Transportation, began the construction of the Mexico-Toluca Interurban Train in January 2015. In this paper, an exhaustive analysis of the train infrastructure in an event of a high-magnitude earthquake is carried out. For that, a scenario of a large earthquake in which the train moves at high speed transporting passengers is considered. Specifically, we analyse the structure behaviour when it is exposed to an earthquake comparable in magnitude to those experienced on September 19th 1985, in Mexico, and May 22nd 1960, in Chile.**

Preliminary results confirm the useful of data science techniques for the study, and offer a comprehensive analysis of the train structural integrity under earthquake conditions via simulations conducted using the SAP 2000 software.

***Index Terms*—Interurban train, Earthquake, Data Science, Clustering, Machine Learning**

I. Introduction

The goal in this study is to subject the structure of the Mexico-Toluca Interurban Train to a critical load condition before its actual operation, to ensure proper performance once the infrastructure is officially opened.

In this regard, Thai [1] conducts a comprehensive review of Machine Learning applications in structural engineering, structural health monitoring, and concrete structure design. On the other hand, Mangalathu S. and Jeon [2] collected 311 experimental samples of circular columns, considering three types of failures: flexure, flexure-shear and shear; and 6 machine learning methods.

Nguyen et al. [3] introduce a method for damage detection in bridges considering changes in modal curvature using two machine learning models, the Bonghi bridge was used as an illustrative example.

In the same way Mahmoudi et al. [4] introduce a framework for detecting the extent of damages in concrete shear wall buildings during seismic movements using machine learning. The results demonstrate the efficiency of the proposed framework to accurately identify damages in the concrete shear walls of the building.

Abdelijaber et al. [5] introduced an efficient system for real-time detection and localization of structural damages using 1D Convolutional Neural Networks. This approach stands out for its ability to automatically extract optimal features from acceleration signals and its computational efficiency, demonstrated in experiments conducted on a simulator.

In addition, Chow and Ghaboussi [6] proposed a genetic algorithm-based approach for detecting structural damage. Unlike traditional methods, this approach avoids complex analyses and can accurately determine the approximate location of the damage using few measures.

In the Sarothi et al. [7] study, the failure modes in double shear bolted connections in structural steel were explored. For that, different algorithms on a database of 455 experimental results to identify the failure modes were used.

Given the empirical evidence from the studies mentioned above, to evaluate the potential response of the Mexico-Toluca Interurban Train under the circumstances that involve simultaneously a high-magnitude earthquakes and excessive load conditions is imperative. This analysis can offer assessment of the train reliability and the prospective users safety.

II. Preliminaries

The main contribution of this paper is presented from two perspectives:

1) Descriptive Phase: The aim is to analyze the data quality. The original data are distributed into four categories related to the seismic amplification factor. By applying clustering algorithms, its anticipate determining the purity of the data regarding its original assignment to each class. This consideration is particularly significant because the data were synthetically obtained. The k-means and Expectation-Maximization (EM) algorithms were used in this analysis.
2) Predictive Phase: Once the data clustering has been identified, the prediction achievable will be analyzed using three state-of-the-art Machine Learning models: Multilayer Perceptron, J48 decision tree, and Support Vector Machine.

A. Data Understanding

The structural analysis incorporates two specific earthquakes: one that took place on May 22nd 1960, in Chile and another occurred on September 19th 1985, in Mexico (magnitude of 8.1). The Chilean earthquake is significant in this analysis, being the highest earthquake recorded in global seismic history, with a moment magnitude of 9.5 Mw. The Mexican earthquake was considered due to its proximity to the construction site and status as one of the highest magnitude events in recent years (8.1 Mw).

Based on the reference values related to the physical structure of the Mexico-Toluca Interurban Train, a structural model was formulated [8]. From this model and employing the Structural Analysis Program (SAP 2000) several simulations were conducted: 996 for the Chilean earthquake and 10,000 for the Mexican one (Figure 1).

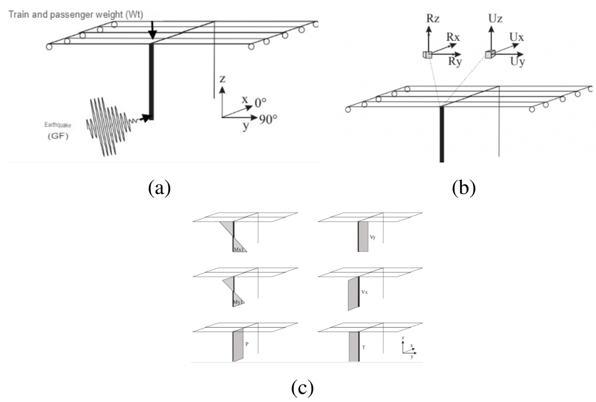

Fig. 1. Overview of key variables applied in designing and assessing a ”typical” support pile for the elevated viaduct of the Mexico-Toluca train. (a) Earthquake considerations: the total weight of the train and the weight of the passengers, (b) Variables concerning displacements at the top node, and (c) Variables associated with force exertions.

III. REVELATIONS IN RESEARCH

For this study, we use the 10-fold cross-validation, being 80% of the instances for training and the remaining 20% for testing. The performance of the models was evaluated using two standard measures obtained from a confusion matrix: accuracy and geometric mean. Table I shows the average values for overall accuracy and the geometric mean from the three Machine Learning models used. The best results are highlighted in bold.

TABLE I
COMPARATIVE RESULTS OF EARTHQUAKE DATASET CLASSIFICATION.

CD	**Accuracy**		**Geometric Mean**	
	Chile	Mexico	Chile	Mexico
Multilayer Perceptron				
Original	80.8	40.68	85.1	57.2
EM	97.8	**99.99**	**98.3**	**99.9**
K-Means	**99.3**	99.50	99.1	99.5
Support Vector Machine				
Original	76.9	40.74	82.0	57.7
EM	97.6	**99.99**	98.1	**99.9**
K-Means	**99.0**	99.62	**98.8**	99.6
Decision Tree J48				
Original	80.2	37.33	86.6	54.6
EM	98.3	**99.99**	**98.7**	**99.9**
K-Means	**98.7**	99.12	98.5	99.1

Based on results shown in the Table I, we can observe that: the use of any dataset generated by clustering methods significantly improves classification performance. When analyzing the accuracy and the geometric mean, both measures display minimal discrepancies between them. This last suggests the class imbalance problem does not affect significantly the performance measures.

IV. CONCLUDING REMARKS

This study examines the behaviour that the structure of interurban train might have in earthquakes with magnitudes of either 9.5 or 8.1.

Preliminary results suggest a class overlap problem in the original data. As a result, the performance derived from the clustering algorithms challenges the distribution obtained with the SAP 2000 software. Because it overlap influences the performance of models dificulting the correct recognition between classes.

On the other hand, in terms of the classification, considering both the initially defined classes and the determined through clustering algorithms, the results of the classifier were better when they are trained with datasets clustered. In fact, the clustering obtained by EM gives a superior performance.

Open lines include to cleaning the decision boundering to handle the class overlap issue, and to use a synthetic data generation conserving the statistical distributions as approach to augment data volume and, in this way, increase the performance of the classifiers. From the civil engineering viewpoint, further research will determine the reliability of the structure.

ACKNOWLEDGMENT

This work was partially supported by the UAEMex 6523/2022 CIA project, by the postdoctoral contract MGS/2021/23 (UP2021-021), funded by the European Union-Next Generation EU.

REFERENCES

[1] H.-T. Thai, “Machine learning for structural engineering: A state-of-the-art review,” *Structures*, vol. 38, pp. 448–491, 2022.

[2] S. Mangalathu and J.-S. Jeon, “Machine learning–based failure mode recognition of circular reinforced concrete bridge columns: Comparative study,” *Journal of Structural Engineering*, vol. 145, no. 10, p. 04019104, 2019.

[3] D. H. Nguyen, Q. B. Nguyen, T. Bui-Tien, G. De Roeck, and M. Abdel Wahab, “Damage detection in girder bridges using modal curvatures gapped smoothing method and convolutional neural network: Application to bo nghi bridge,” *Theoretical and Applied Fracture Mechanics*, vol. 109, p. 102728, 2020.

[4] H. Mahmoudi, M. Bitaraf, M. Salkhordeh, and S. Soroushian, “A rapid machine learning-based damage detection algorithm for identifying the extent of damage in concrete shear-wall buildings,” *Structures*, vol. 47, pp. 482–499, 2023.

[5] O. Abdeljaber, O. Avci, S. Kiranyaz, M. Gabbouj, and D. J. Inman, “Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks,” *Journal of Sound and Vibration*, vol. 388, pp. 154–170, 2017.

[6] J.-H. Chou and J. Ghaboussi, “Genetic algorithm in structural damage detection,” *Computers & Structures*, vol. 79, no. 14, pp. 1335–1353, 2001.

[7] S. Zakir Sarothi, K. Sakil Ahmed, N. Imtiaz Khan, A. Ahmed, and M. L. Nehdi, “Machine learning-based failure mode identification of double shear bolted connections in structural steel,” *Engineering Failure Analysis*, vol. 139, p. 106471, 2022.

[8] J. A. Jímenez García, “Función del enlace técnico dentro de la obra del tren Interurbano México Toluca.” Memoria de Experiencia Laboral, Ingeniería Civil, Facultad de Ingeniería, Universidad Autónoma del Estado de México, 2017.