frontiers | Research Topics

Machine learning applications in civil engineering

Edited by
Yasmin Murad, Husam Abu Hajar and Iftikhar Azim

Published in
Frontiers in Built Environment

ISSN 1664-8714
ISBN 978-2-83250-917-3
DOI 10.3389/978-2-83250-917-3

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Machine learning applications in civil engineering

Topic editors
Yasmin Murad — The University of Jordan, Jordan
Husam Abu Hajar — The University of Jordan, Jordan
Iftikhar Azim — Shanghai Jiao Tong University, China

Topic coordinator
Ahmad N. Tarawneh — Hashemite University, Jordan

Citation
Murad, Y., Hajar, H. A., Azim, I., Tarawneh, A. N., eds. (2022). *Machine learning applications in civil engineering*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-83250-917-3

Table of contents

frontiers | Frontiers in Built Environment

ORIGINAL RESEARCH
published: 01 April 2022
doi: 10.3389/fbuil.2022.858020

Evaluation Circular Failure of Soil Slopes Using Classification and Predictive Gene Expression Programming Schemes

Shadi Hanandeh *

Assistant Professor Civil Engineering Department, Al-Balqa Applied University, Al-Salt, Jordan

This article presents an analysis of a dataset related to slope stability circular failure cases using the GeneXproTools program. Two models were developed: The first model was utilized to find a classification model that shows the stability status (S), and the second model was utilized to estimate the factor of safety and predict the mathematical model to serve as an indicator of slope stability. The proposed models include cohesion, unit weight, slope angle, coefficient of pore water pressure, internal friction angle, slope height, internal friction angle, and safety factor. The results showed that the values of accuracy and error were 93.1 and 6.9%, respectively, for classification model 1. The relationship between the actual values of the factor of safety and the output of model 2 was an R2 value of 0.96. The models developed were subjected to sensitivity analysis to determine the input variables' impact on output for the two models. The selected slope locations with different material characterizations were introduced into geox5 software, and the results were compared to those of the two developed models. Furthermore, when the proposed models were compared to the available models in the literature, the results showed that the proposed models in this study performed better than the available models.

Keywords: slope stability, genetic algorithm, machine learning, soil properties, hazards slope

OPEN ACCESS

Edited by:
Yasmin Murad,
The University of Jordan, Jordan

Reviewed by:
Bara' Al-Mistarehi,
Jordan University of Science and Technology, Jordan
Ahmad N. Tarawneh,
Hashemite University, Jordan

***Correspondence:**
Shadi Hanandeh
hanandeh@bau.edu.jo

Specialty section:
This article was submitted to Computational Methods in Structural Engineering, a section of the journal Frontiers in Built Environment

Received: *19 January 2022*
Accepted: *15 February 2022*
Published: *01 April 2022*

Citation:
Hanandeh S (2022) Evaluation Circular Failure of Soil Slopes Using Classification and Predictive Gene Expression Programming Schemes. Front. Built Environ. 8:858020. doi: 10.3389/fbuil.2022.858020

INTRODUCTION

Slope stability assessment is utilized to evaluate earth stability, including soil or rock slope. Analyses are commonly targeted to understand the influences that possibly activate the slope movement or the reasons for a recent slope failure. Slope failure happens when the downward material movements of shear stresses and gravity become greater than the soil shear strength. Thus, slope failure occurs when there is a decrease in soil shear strength or an increase in applied shear stresses. Cracks, swelling, pore pressure, decomposition fills of creep loads, clayey rock, strain softening, leaching, weathering, and cyclic loading can all cause a decrease in slope shear strength. In addition, shear stress might increase because of extra loads at the top of the soil slope and an increase in soil weight because of an increase in the water content, an increase in water pressure in cracks at the top of the slope, and seismic effects at the bottom of the slope. Moreover, these causes are related to slope failure, such as stress state, slope geometry, erosion, and temperature. The water impact on the slope can be measured in two locations: One is the aquifer under the groundwater or the surface that produces pore water pressure, and the second is the infiltration of rainwater that leaks through flows and the surface. It is associated with the nearby levels of rainfall, near water masses, and the geo-hydrological and topographic soil features.

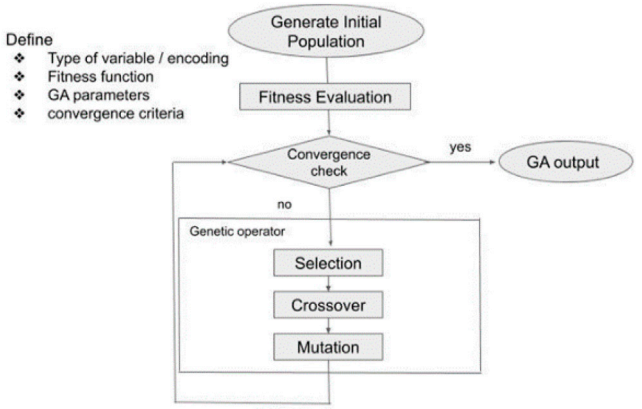


FIGURE 1 | Flowchart of the genetic algorithm.

It is noticed that groundwater is almost present everywhere under the Earth's surface. Groundwater fills the whole space between the grains in the soil, and it can leak into the gap existing in the rock mass, changing the air in the whole space, which leads to an increase in soil weight.

The process of avoiding the initiation movement, decelerating it through countermeasures, or it down, starting with Terzaghi's 1950 work, landslide mechanism. Natekin and Knoll (2013) investigated the slope stability in terms of analytical, theoretical, numerical, statistical, and experimental methods that have been used throughout the world. Xu and Low (2006) used the finite element method and the limit equilibrium method. Even with these methods, certain slope failure mechanisms were too complicated for them to be effective, including internal deformation and brittle cracking, as well as creeping and liquefaction of weaker soil layers (Zhou et al., 2019). However, assessment of slope stability depends on historical performances and geotechnical characteristics of the collected slope cases by using an artificial neural network (Lu and Rosenbaum, 2003). Wang et al. (2005) also used the ANN to forecast the slope stability and precisely estimate the factor of safety for the Yodonghe landslide. The support vector machine was used to discover the non-linear relationship between its influence factors and slope stability (Zhao et al., 2012; Zhao and Li, 2021). Lollino et al. (2016) investigated two slopes in Italy: on the first slope with stiff clay, the dissipation of negative surplus pore water pressures created by earlier mine excavation at the slope toe caused a first-time collapse, and on the second slope, the stress–strain development of the old landslide reactivates mostly in rainy seasons. The benefits of using artificial intelligence and genetic algorithms in geotechnical applications have been studied in different studies (Al Bodour et al., 2022; Hanandeh et al., 2020a; Hanandeh et al., 2020b).

Additionally, for circular failure, the factors used were cohesion, unit weight, slope height, slope angle, ratio of pore water pressure, and internal friction angle. These six factors were consistent with those normally accepted for circular failure evaluation in a previous research (Sah et al., 1994). If the slope height increased, the slope stability decreased sharply, and when the angle developed like that between the normal force and the limit equilibrium state, the resultant force extended and failure happened because of shearing. Zhao (2008) worked on spherical mode failure using the incline boosting machine method and accuracy rate classification, and the coefficient of Cohen's Kappa was used for determining the proposed model performance. The slope stability state is expressed as some classification difficulties to estimate the output either failed or stable. Ahour et al. (2020) used a numerical model that is dependent on artificial neural networks (ANN) to calculate the rock slope stability by using the comprehensive Hoek–Brown failure principle. In this research, the FOS was calculated using the finite element approach using the SSR method. Sakellariou and Ferentinou (2005) studied slope stability estimation using neural networks. This model offers forecasting abilities; therefore,

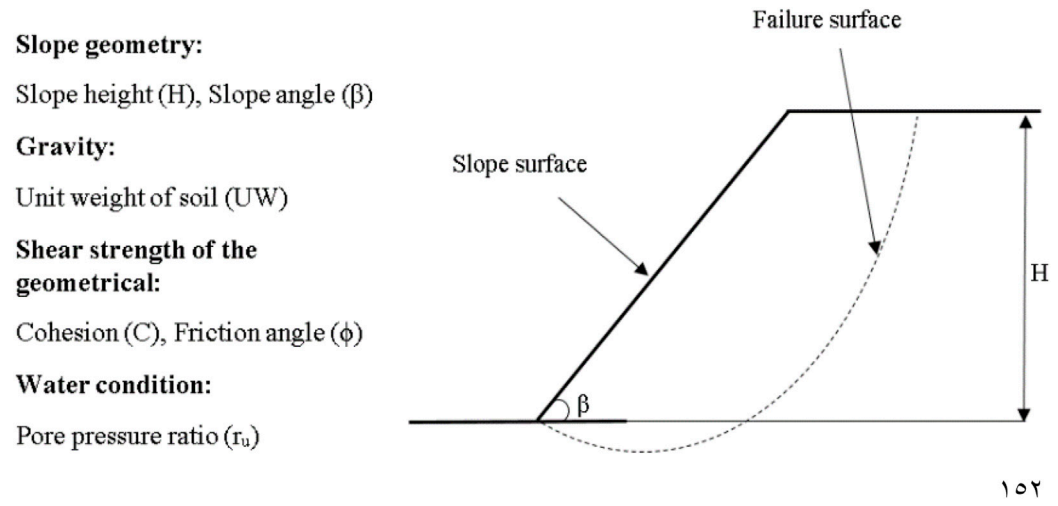


FIGURE 2 | Effective factors affecting the stability of slopes undergoing circular collapse.

TABLE 1 | The descriptive statistical analysis of the classifcation data for model 1.

Parameter	Min	Max	Mean
H	3.66	565	97
β	10	59	34
γ	12	31	22
C	0	300	30
Φ	0	50	29
ru	0	1	0

TABLE 2 | The descriptive statistical analysis for model 2.

Parameter	Mean	Max	Min
γ	19.49	28	12
C	23.46	150	0
Φ	25.01	45	0
β	32.03	53	16
H	46.76	214	3.66
ru	0.09	05	0
FS	1.28	2	0.78

it has been used for soil and rock parameter prediction and identification. In addition, substantial progress has been made in the study of slope stability difficulties. The prepared equation may be used to forecast the supplement for a given set of input variables without having to do any further work. This is regarded as among the most significant benefits of the standard soft computing methods used in today's world. In the literature on slope stability, successful implementations of these soft computing approaches may be identified (Wang et al., 2005; Choobbasti et al., 2009; Li et al., 2009; Chakraborty and Goswami, 2017; Kumar and Basudhar, 2018; Qian et al., 2019; Ahour et al., 2020; Li et al., 2020; Ray et al., 2020; Zheng et al., 2020; Che Mamat et al., 2021; Palazzolo et al., 2021); (Das et al., 2011; Erzin and Cetin, 2012; Erzin and Cetin, 2014; Abdalla et al., 2015; Ai and Zsaki, 2017; Chakraborty and Goswami, 2018; Rukhaiyar et al., 2018; Moayedi et al., 2019; Bui et al., 2020; Chen et al., 2020; He et al., 2020; Liao and Liao, 2020; Che Mamat et al., 2020; Markovic Brankovic et al., 2021; Meng et al., 2021).

This study used GeneXproTools to propose two models: The first model was classified by including different input factors such as the cohesion (c), slope angle (), unit weight, internal friction angle, slope height, and pore water pressure coefficient, and the output was the occurrence of slope failure. The second model was a regression-based function approximation model that included various input parameters such as cohesion (c), unit weight, internal friction angle, angle of slope, coefficient of pore water pressure, slope height, and factor of safety. The output was the existence of slope failure in order to predict the failure state and function approximation to obtain a formula and a factor of safety of 2.

METHODOLOGY

The study utilized GeneXproTools, which is an enormously flexible modeling tool considered for logistic regression, linear regression, time series estimation, classification, and logic mixture. For the objective of developing an analytical or classified model for the prediction of FS on slopes with the potential for failure, a GA was used in this research. The primary goal was to establish whether variables should indeed be quantitatively converted using equations such as power, exponential, and logarithm. The quality of the regression models was improved by first converting the variables and then computing the coefficients of the regression model. In addition, it may be used to

TABLE 3 | Confusion matrix for data.

Confusion matrix		
Yes (predicted)		**No (predicted)**
Yes (actual)	110 (TP)	5 (FN)
No (actual)	10 (FP)	95 (TN)

demonstrate non-linear interactions between dependent factors and independent variables. The planned GA comprised four major phases as follows: First, a collection of alternative functions was determined; next, each independent variable was assigned a transforming function; then, a combination of multiple solutions was used to identify the best one; and last, the predictive estimated values were calculated. A regression model was created to test our hypothesis. Before any more steps were taken, a group of mathematical functions was created that seem to capture the essence of the issue and have no chance of being incorrect mathematically. Each independent variable was randomly modified using one of these mathematical functions. This function has a general form independent variable and is called the conversion function. Decision was made as to what kind of modification should be applied to the variable. As a result, each of these forms was determined by a digit, and each variable was allocated to it via the use of a random technique. It was necessary to utilize this approach to generate arbitrary parameters for each of the independent variables in the system. This means that the encoded values of the numeric variables make up each possible solution. It is known as the power of the number of all feasible solutions for the problem with various transformation forms and l variable. It demonstrates that this GA was capable of creating and testing a huge number of different regression models. Following the formation of the solutions, the prediction parameters of the model for each solution were defined in order to generate a fitness value, and the regression model (a mathematical criterion such as the root-mean-square error, standard deviation, and mean square error) for every solution was calculated. Using mutation and crossover operators, the GA progressed toward the optimal solution as quickly as possible. By evaluating the best fitness function, this GA created a vast number of alternative solutions and then picks the best one from among them. Many scientific issues need the development of models to clarify the statistical connection between distinct variables, but finding the most

FIGURE 3 | Tree Expression to classified slope stability.

TABLE 4 | Estimating function historical data.

No.	γ	C	Φ	β	H	ru	F.O.S	GA model one	GA model two
1	19	26	15	35	8.23	0	1.09	1	1.03
2	17	11.49	0	30	3.66	0	1	0	0.99
3	19	24	16	28	12	0	2	1	1.59
3	16	14	22	20	18	0	2	1	1.68
4	25	7	28	33	7	0	1.78	1	1.28
5	21	33	10	20	40	0	1.99	0	1.42
6	23	16	31	27	14	0	1.25	1	1.86
7	24	2	22	21	8	0	1.13	1	1.70
8	22	0	25	24	70	0	1.02	0	1.58
9	20	13	29	32	32	0	1.3	1	1.03
10	19	26	15	35	8.23	0	1.2	0	0.99

appropriate regressing model is not always straightforward. A diagram illustrating how a genetic algorithm framework is organized is shown in **Figure 1**. Cramer NL was the first to create genetic programming.

DATA PROCESS

The database for model 1 contains 253 cases that were collected from real data for slope stability cases by (Sah et al., 1994), (Wang et al., 2005), (Chen and Chameau, 1983), and (Zhou et al., 2019) is shown in **Table 1**. The data for model 2, which includes of 220 historical cases of slope conditions and collected by Sah et al. (1994), Sakellariou and Ferentinou (2005), Samui (2008), Zhou and Chen (2009), and Li and Dong (2012), are shown in **Tables 2**, **3**.

Figure 2 shows that the slope's geometric form includes the slope angle, slope height, cohesiveness, unit weight, pore pressure ratio, and internal friction angle, which is the pore pressure-to-overburden pressure ratio.

RESULTS AND ANALYSIS

Genetic Algorithm Classification Model One

It should be noted that the binary system was used to classify the slope state, with 0 representing "failure" and 1 representing "stable." Slope height (H), overall angle of slope (*ß*), unit weight (γ), cohesion (C), and the internal friction angle (φ) were used as input parameters in the program, which are in line with whatever is frequently used for the analysis of circular failure. The geometrical simple slope design factors (H), (*ß*), and () typically determine soil slope failure conditions and are mechanical parameters related to slope stability in geomaterials that are dependent on the Mohr–Coulomb

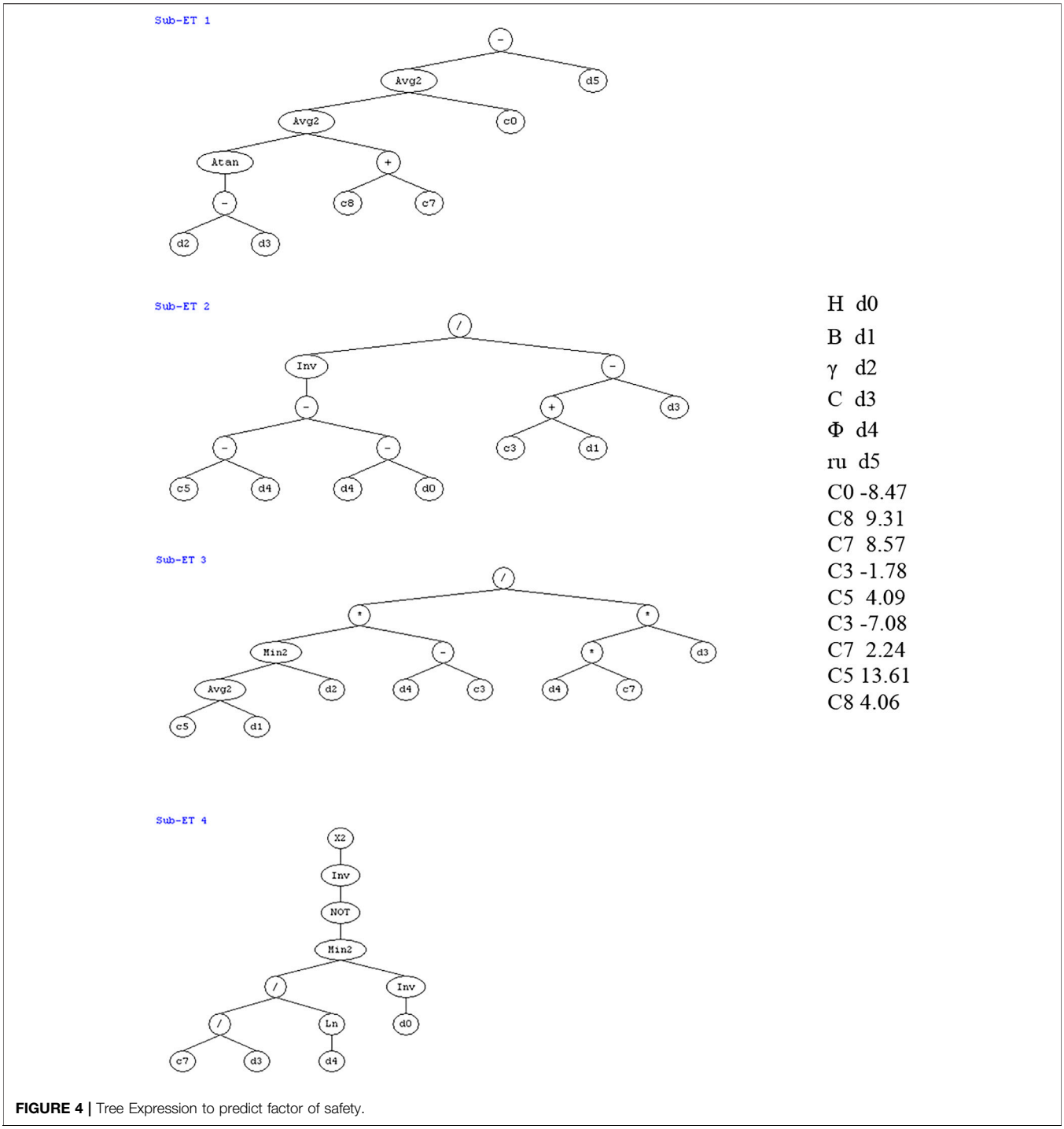


FIGURE 4 | Tree Expression to predict factor of safety.

failure standard (ru), which is the pore water pressure ratio to the overload pressure. The output was the slope state (actual) and can take two options: OK (correct match) or wrong (false match). An expression tree, as the term recommends, was terms arranged in a tree-like information configuration. It helps understand the program algorithms since it adapts to various programming languages to express the results. Another advantage of expression trees is that they can provide a formula where the same kind of inputs can be used to get an output, which in our case was stable (1) or failure (0). The number of genes in the classification stage was three genes. It means that we will have three sub-trees in **Figure 3** that were produced and transformed into the numerical scheme as displayed (classification formula). The interesting part about this program is that it can deliver results with an algorithm based on the data entered. The algorithm in

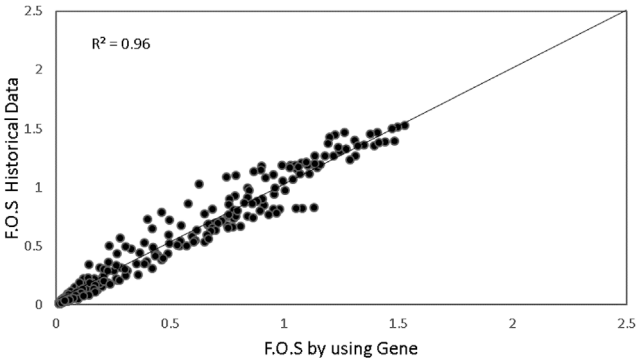


FIGURE 5 | Predicted results using gene software plot against factor of safety for measured data.

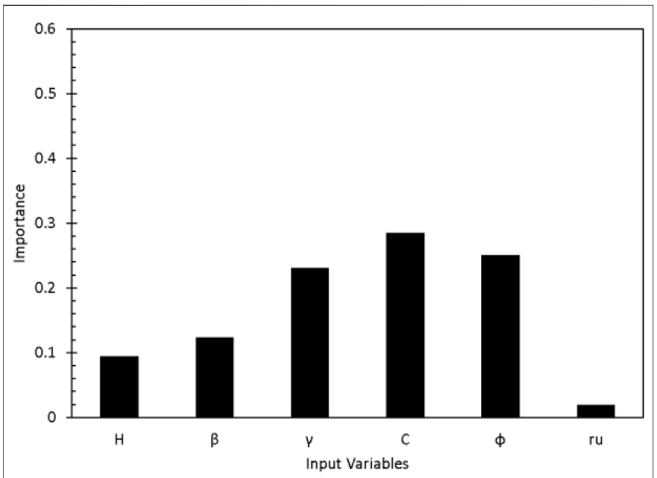


FIGURE 6 | Importance of a Feature for Classification Variable importance.

this program is compatible with expression trees to obtain more realistic results.

The data were organized into txt files before being loaded into the program Gene X, which contains an algorithm for generating a non-linear classification model. Then, deployment of models and ensembles to Excel was carried out. In machine learning and specifically in statistical classification difficulties, a confusion matrix is an exact table layout that permits the visualization performance of an algorithm and a classification model on a group of test data for which the true standards are identified. Every row of the matrix links to an expected class. Every column of the matrix agrees with an actual class. The program compares the results using a confusion matrix and shows them as part of the results via the match column that we discussed earlier. In **Table 4**, as for the 104 values that are the same in the yes predicted and the yes actual columns, take a TP symbol; as for the 19 values that are in the yes predicted and the no actual, take the FP symbol. The genetic algorithm was first evaluated on the same training data (all 220 slope instances in the database). The confusion matrix is shown in **Table 4**, and the variables TP (stable) and TN (failure)

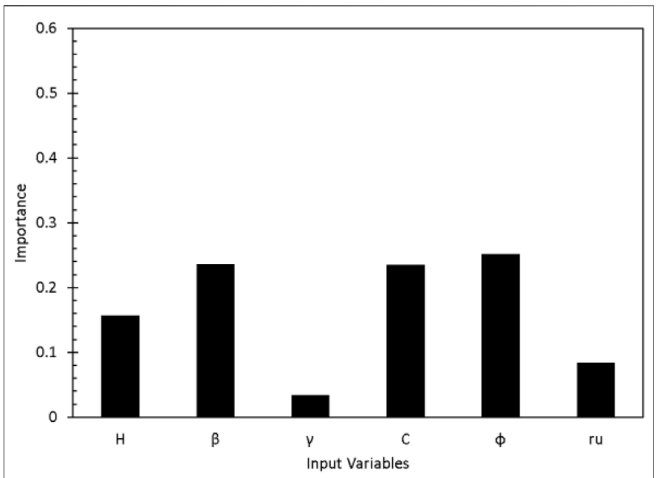


FIGURE 7 | Importance of a Feature for Function estimation.

denote the stable and failing slopes numbers that are accurately sorted into the relevant classes. False positive (FP) stands for the number of failure slopes that are incorrectly identified as stable, whereas false negative (FN) denotes the number of stable slopes that are incorrectly classified as failed. The precision of 115 stable slopes is 95 percent (referred to as sensitivity, which is defined as TP/(TP + FN)). The accuracy (referred to as "specificity," which equals TN divided by (TN + FP)) of the 28 failed slopes is 90.4 percent. Total accuracy (TP + TN)/(TP + TN + FP + FN) is approximately 93.1 percent, which is very good for real-world engineering.

Genetic Algorithm Prediction Model 2

To achieve an accurate estimate of the safety factor, we need to run the data through a function approximation issue, in addition to a prediction of whether the slope state is a failure or is stable. Slope height (H), overall angle slope (β), unit weight (γ), internal friction angle (φ), cohesion (C), and ratio of pore water (ru) were used as input parameters in the program which are in accordance with what is frequently used for circular failure analysis. The number of genes in regression stage was four genes, which means we will have four sub-trees, as shown in **Figure 4**.

This software's working idea is based on chromosomes, which include expression trees and genes that specify the genetic information contained in them. The Gene Expression Program (GeneXpro 5.0) saves time and effort in calculating the factor of safety. Furthermore, by establishing a link between soil types, features, and bearing capability, when the generation is 101,934, the computer tries numerous paths to get the ideal model. The trial was paused, and the best model was chosen based on certain criteria, like $R^2 > 0.90$ for all data. **Figure 5** demonstrates the expected factor of safety alongside the observed safety factor. The evolutionary algorithm model also accurately predicts the destination values. However, the genetic algorithm model delivers an excellent estimate, with an R^2 value of 0.96.

FIGURE 8 | Landslide area at Amman Jerash highway.

The number of genes in the regression stage was four genes. That means we will have four equations used in this generation. The estimated equations are as follows:

$$\text{F.O.S} = \left[\left[\frac{\tan^{-1}(\Phi-\beta)}{4} + 0.411 - \text{ru}\right]\right] + \frac{c_3 + C - \beta}{(c5 - H)\cdot(H-\gamma)} + \frac{\min[(\text{Av}13.62.C).\Phi*(H+7.87)]}{2.24H*\beta} + \left(\frac{1}{1-\min\left(\frac{4.06\text{Ln}\,H}{\beta}, \frac{1}{\gamma}\right)}\right)^2.$$

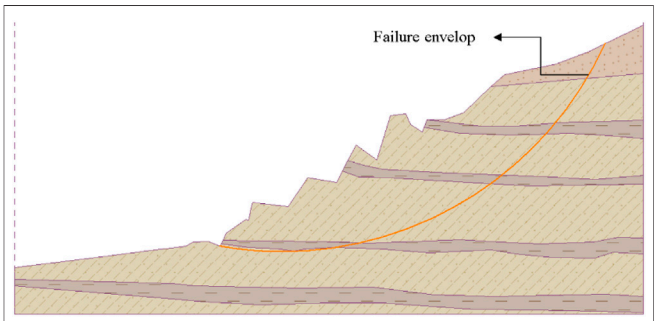


FIGURE 9 | Geox5 for selected site at Irbid Jerash highway.

SENSITIVITY ANALYSIS

Sensitivity analysis involves adjusting an entry parameter's amount and observing how the genetic algorithm model responds. This may reveal which variables have the most influence on genetic algorithm models and which variables affect simulation results. This assessment demonstrates how input factors affect the safety factor of slope failure. The significance of a feature is assessed by the change in the estimation of designated prediction models, and the length of the bars represents the final assignment's feature relevance. **Figure 6** shows the influence of input variables on model 1 as follows: cohesion (C), angle of internal friction (φ), unit weight (γ), overall angle slope (), slope height (H), and pore water ratio (ru). **Figure 7** summarizes the features relevant to model 2. The angle of internal friction (φ), overall angle slope (ß), cohesion (C), slope height (H), pore water ratio (ru), and unit weight (γ) were the main input factors for factor safety prediction modeling.

VALIDATION BY USE SLOPE STABILITY CASES IN JORDAN

The common landslides that happened along the routes of major roads in Jordan, and mainly along the Dead Sea-Na'our–Amman road and the Amman–Jerash–Irbid road, have considerably increased. Over the last decades, excessive hillslope disturbances

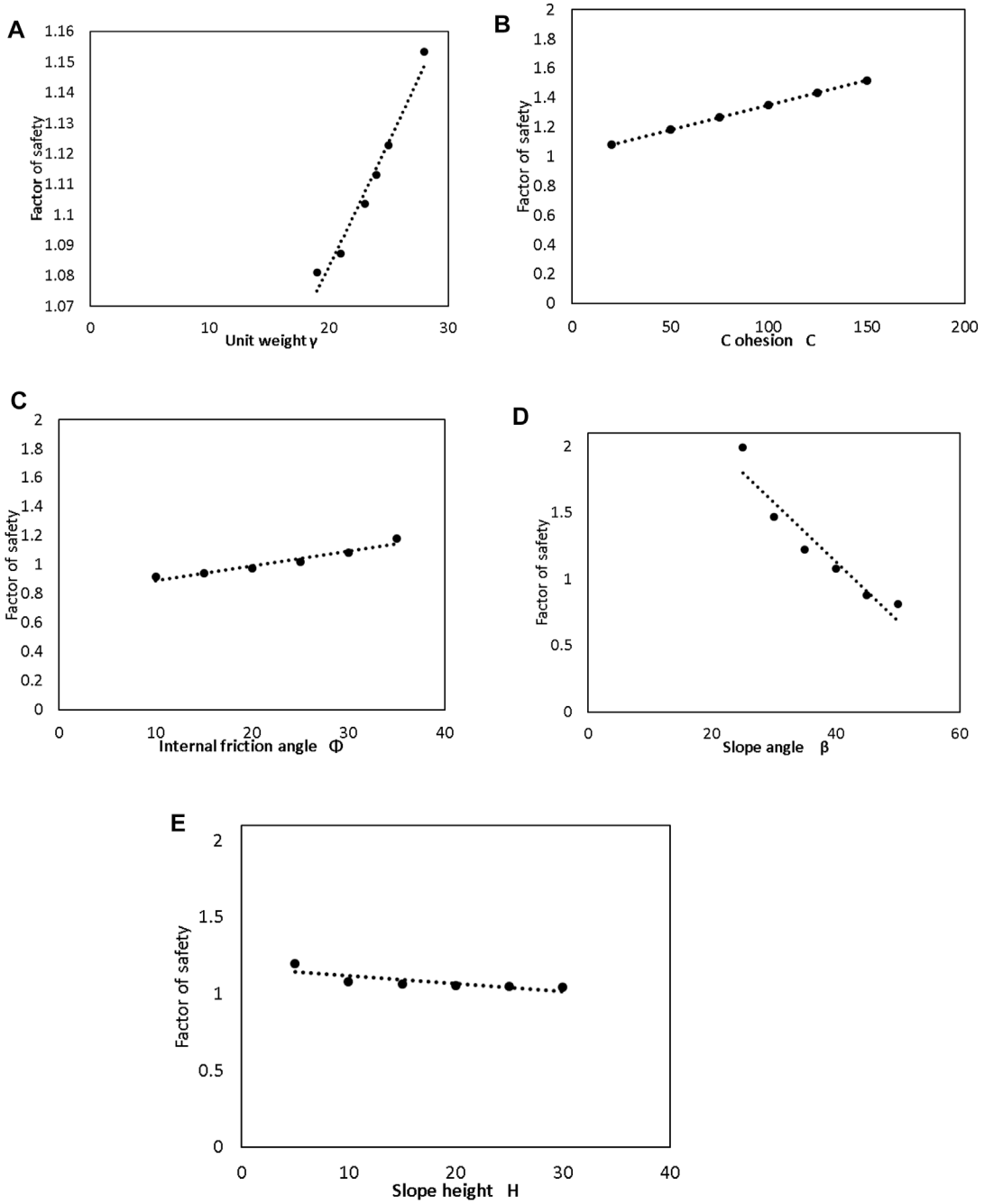


FIGURE 10 | Parametric study for input variables. **(A)**. Unit weight vs. factor of safety. **(B)**. Cohesion vs. Factor of Safety. **(C)**. Internal Friction Angle vs. Factor of Safety. **(D)**. Slope Angle vs. Factor of Safety. **(E)**. Slope height vs. Factor of safety.

have been identified during highway construction and urban development in central and northern Jordan. Most of the major engineering projects were implemented on natural, steep slopes, which were related to old landslide complexes, weak geological structures, and low shearing resistance. Thus, a wide range of landslides were reactivated, causing serious damage to highways and urban areas. Engineering activities often take place without previous appreciation of the geological and hydrological settings. The human impact of the increased probability of landslides was also ignored. The final triggering factor responsible for these slope failures was the buildup of high pore-water pressures in the materials, caused by recurrent, high-intensity rainstorms. A geotechnical stability evaluation was carried out along the Amman-Irbid highways as shown in **Figure 8**. It is concluded that hillslope instability problems are likely to increase with continued highway and urban development. In order to

TABLE 5 | The statistical assessments of the responses for model 2 with other models.

Model	R^2 (%)	SSE	RMSE
This study	92	0.12	0.018
Manouchehrian et al. (2014)	79	0.26	0.12
Yang et al. (2004)	61	0.54	0.18

investigate the slope failure at the Amman–Irbid Highway, the site investigation and soil physical and mechanical properties were performed in the laboratory. Six boreholes were drilled in slope failure sites and the soil classification was sandstone with an unconfined compressive strength of between 3 and 45 MPa. The direct shear test was performed on the collected samples, the cohesion (C kPa) ranged between 32 and 51%, and the internal friction angle was between 7 and 49°.

geox5 software was used to calculate the factor of safety for the slope failure in **Figure 9**. The soil properties with all properties were assigned to geox5. **Table 4** depicts the slope stability situation at 10 slope locations in the Irbid–Jerash Highway. geox5 was used to find the factor of safety based on several material properties. Furthermore, **Table 4** shows the comparison between geox5 and the 2 moles that were proposed in this study.

In order to verify the factor of safety models developed in this study, 10 slopes at Irbid Jerash were utilized by the Ministry of Housing. These slope cases were used to test the factor of safety models developed in this study. The factor of safety was predicted using the models created in this study first, and then geox5 was performed on the variables shown in the table.

PARAMETRIC STUDY BY USING GENE EXPRESSION PROGRAMMING

A parametric study was carried out for the established genetic expression programming model in order to investigate the impact of the input values on the anticipated factor of safety. As a result, particular input data are altered, while the other input data are left constant in order to examine the impact of each input on the anticipated factor of safety. **Figures 10A, B,C** show that as unit weight, cohesion (c), and internal friction angle increase, the factor of safety also increases in a direct proportion. The impact of cohesion (c) on the factor of safety is increased in a direct proportion. **Figures 10D,E** show that slope angle (*ß*) and slope height are inversely proportional to the factor of safety.

DISCUSSION AND ANALYSIS

In this study, an evolutionary algorithm was used to construct a regression model for the estimation of FS for soil slopes that have the possibility of circular failure. The suggested genetic algorithm models provide a higher coefficient of determination and minimize the root-mean-square error. The sensitivity analysis shows that increasing unit weight, soil cohesion, and angle of friction increase FS, whereas increasing slope height, slope angle, and pore pressure reduce FS. In addition, the findings revealed that the most efficient FS parameters were slope angle, unit weight, and friction angle, but as unit weight increased, the model's efficiency increased as well. The proposed genetic algorithm models were compared to the models from the literature. Compared to the model being used for validation, the offered model predicted more accurately. As a result of its simple formulation, the regression model proposed by the researchers is capable of accurately forecasting FS with a high degree of confidence, making it an effective alternative to LEM and numerical approaches in the early phases of the design process. To summarize, the suggested GA's notable features include its easy formulation and high accuracy, which may lead to the construction of extremely helpful regression models for a wide range of issues. **Table 5** summarizes the statistical assessments of the responses for model 2 compared with the other models. The results demonstrate that the genetic algorithm model proposed was more effective in assessing FS.

CONCLUSION

The genetic algorithm technique has been effectively used to examine the slope stability situation using (252) slope conditions of verified historical cases (regression and classification). Using the GeneXproTools model, the state slope stability is expressed as a classification difficulty in which estimation outputs are either failed or stable. The study shows that the solutions from the genetic algorithms method are accurate for most cases. The results showed values of accuracy and error of 93.1 and 6.9%, respectively, where the parameter cohesion (C) was the most important aspect in the classification case. For the function estimating case, independent variables in this research involved material variables (height of slope and slope angle). Using the regression feature, an equation was established to forecast FOS of slopes. The genetic algorithm models were observed and compared to show that all approaches have a respectable agreement in forecasting the safety factor. The results of the parametric study show that increasing unit weight, cohesion, and internal friction angle increase the factor of safety, whereas increasing slope height, slope angle, and pore pressure reduce the factor of safety.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

The author has participated in conception and design; analysis and interpretation of the data; drafting the manuscript or revising it critically for important intellectual content; and approval of the final version.

REFERENCES

Abdalla, J. A., Attom, M. F., and Hawileh, R. (2015). Prediction of Minimum Factor of Safety against Slope Failure in Clayey Soils Using Artificial Neural Network. *Environ. Earth Sci.* 73 (9), 5463–5477. doi:10.1007/s12665-014-3800-x

Ahour, M., Hataf, N., and Azar, E. (2020). A Mathematical Model Based on Artificial Neural Networks to Predict the Stability of Rock Slopes Using the Generalized Hoek–Brown Failure Criterion. *Geotech. Geol. Eng.* 38, 587–604. doi:10.1007/s10706-019-01049-y

Ai, X., and Zsaki, A. M. (2017). Stability Assessment of Homogeneous Slopes Loaded with mobile Tracked Cranes—An Artificial Neural Network Approach. *Cogent Eng.* 4, 1. doi:10.1080/23311916.2017.1360236

Al Bodour, W., Hanandeh, S., Hajij, M., and Murad, Y. (2022). Development of Evaluation Framework for the Unconfined Compressive Strength of Soils Based on the Fundamental Soil Parameters Using Gene Expression Programming and Deep Learning Methods. *J. Mater. Civ. Eng.* 32 (2). doi:10.1061/(asce)mt.1943-5533.0004087

Bui, X. N., Nguyen, H., Choi, Y., Nguyen-Thoi, T., Zhou, J., and Dou, J. (2020). Prediction of Slope Failure in Open-Pit Mines Using a Novel Hybrid Artificial Intelligence Model Based on Decision Tree and Evolution Algorithm. *Sci. Rep.* 10, 9939. doi:10.1038/s41598-020-66904-y

Chakraborty, A., and Goswami, D. (2018). Prediction of Critical Safety Factor of Slopes Using Multiple Regression and Neural Network. *J. Geo-engineering Sci.*. [Preprint], 1–10. doi:10.3233/jgs-170047

Chakraborty, A., and Goswami, D. (2017). Slope Stability Prediction Using Artificial Neural Network (ANN). *Int. J. Eng. Comput. Sci* 6 (6), 21845–21848. doi:10.18535/ijecs/v6i6.49

Che Mamat, R., Manan Samad, A., Kasa, A., Razali, S. F. M., Ramli, A., and Omar, M. B. H. C. (2020). Slope Stability Prediction of Road Embankment on Soft Ground Treated with Prefabricated Vertical Drains Using Artificial Neural Network. *Prediction embankment Stab.* 9, 239–243. doi:10.11591/ijai.v9.i2.pp236-243

Che Mamat, R., Ramli, A., Samad, A. M., Kasa, A., Razali, S. F., and Omar, M. B. (2021). Artificial Neural Networks in Slope of Road Embankment Stability Applications: A Review and Future Perspectives. *Int. J. Adv. Technol. Eng. Explor.* 8, 304–319. doi:10.19101/ijatee.2020.762127

Chen, L., Zhang, W., Gao, X., Wang, L., Li, Z., Böhlke, T., et al. (2020). Design Charts for Reliability Assessment of Rock Bedding Slopes Stability against Bi-planar Sliding: SRLEM and BPNN Approaches. *Georisk: Assess. Manag. Risk Engineered Syst. Geohazards*, 1–16. doi:10.1080/17499518.2020.1815215

Chen, R. H., and Chameau, J.-L. (1983). Three-dimensional Limit Equilibrium Analysis of Slopes. *Géotechnique* 33, 31–40. doi:10.1680/geot.1983.33.1.31

Choobbasti, A. J., Farrokhzad, F., and Barari, A. (2009). Prediction of Slope Stability Using Artificial Neural Network (Case Study: Noabad, Mazandaran, iran). *Arab. J. Geosci.* 2 (4), 311–319. doi:10.1007/s12517-009-0035-3

Das, S. K., Biswal, R. K., Sivakugan, N., and Das, B. (2011). Classification of Slopes and Prediction of Factor of Safety Using Differential Evolution Neural Networks. *Environ. Earth Sci.* 64, 201–210. doi:10.1007/s12665-010-0839-1

Erzin, Y., and Cetin, T. (2014). The Prediction of the Critical Factor of Safety of Homogeneous Finite Slopes Subjected to Earthquake Forces Using Neural Networks and Multiple Regressions. *Geomech. Eng.* 6, 1–15. doi:10.12989/gae.2014.6.1.001

Hanandeh, S., Alabdullah, S. F., Aldahwi, S., Obaidat, A., and Alqaseer, H. (2020a). Development of a Constitutive Model for Evaluation of Bearing Capacity From CPT and Theoretical Analysis Using ann Techniques. *Int. J. Geomate.* 19 (74), 229–235. doi:10.21660/2020.74.36965

Hanandeh, S., Ardah, A., and Abu-Farsakh, M. (2020b). Using Artificial Neural Network and Genetics Algorithm to Estimate the Resilient Modulus for Stabilized Subgrade and Propose New Empirical Formula. *Transp. Geotech.*. doi:10.1016/j.trgeo.2020.100358

Erzin, Y., and Cetin, T. (2012). The Use of Neural Networks for the Prediction of the Critical Factor of Safety of an Artificial Slope Subjected to Earthquake Forces. *Sci. Iran.* 19, 188–194. doi:10.1016/j.scient.2012.02.008

He, X., Xu, H., Sabetamal, H., and Sheng, D. (2020). Machine Learning Aided Stochastic Reliability Analysis of Spatially Variable Slopes. *Comput. Geotech* 126, 103711. doi:10.1016/j.compgeo.2020.103711

Kumar, S., and Basudhar, P. K. (2018). A Neural Network Model for Slope Stability Computations. *Geotech. Lett.* 8, 149–154. doi:10.1680/jgele.18.00022

Li, A.-J., Lim, K., and Fatty, A. (2020). Stability Evaluations of Three-Layered Soil Slopes Based on Extreme Learning Neural Network. *J. Chin. Inst. Eng.* 43, 628–637. doi:10.1080/02533839.2020.1719899

Li, J., and Dong, M. (2012). "Method to Predict Slope Safety Factor Using SVM," in Earth and Space 2012-Proceedings of the 13th ASCE Aerospace Division Conference and the 5th NASA/ASCE Workshop on Granular Materials in Space Exploration (Pasadena, CA: American Society of Civil Engineers). doi:10.1061/9780784412190.095

Li, S., Shangguan, Z., Duan, H., Liu, Y., and Luan, M. (2009). Searching for Critical Failure Surface in Slope Stability Analysis by Using Hybrid Genetic Algorithm. *Geomech. Eng.* 1 (1), 85–96. doi:10.12989/gae.2009.1.1.085

Liao, Z., and Liao, Z. (2020). Slope Stability Evaluation Using Backpropagation Neural Networks and Multivariate Adaptive Regression Splines. *Open Geosci.* 12, 1263–1273. doi:10.1515/geo-2020-0198

Lollino, P., Cotecchia, F., Elia, G., Mitaritonna, G., and Santaloia, F. (2016). Interpretation of Landslide Mechanisms Based on Numerical Modelling: Two Case-Histories. *Eur. J. Environ. Civ. Eng.* 20, 1032–1053. doi:10.1080/19648189.2014.985851

Lu, P., and Rosenbaum, M. S. (2003). Artificial Neural Networks and Grey Systems for the Prediction of Slope Stability. *Nat. Hazards* 30, 383–398. doi:10.1023/b:nhaz.0000007168.00673.27

Manouchehrian, A., Gholamnejad, J., and Sharifzadeh, M. (2014). Development of a Model for Analysis of Slope Stability for Circular Mode Failure Using Genetic Algorithm. *Environ. Earth Sci.* 71, 1267–1277. doi:10.1007/s12665-013-2531-8

Markovic Brankovic, J., Markovic, M., Andrejevic Stosovic, M., Zivkovic, S., and Brankovic, B. (2021). Ann Model for Prediction of Rockfill Dam Slope Stability. *Teh. Vjesn.* 28, 1488–1494. doi:10.17559/TV-20200707150903

Meng, J., Mattsson, H., and Laue, J. (2021). Three-dimensional Slope Stability Predictions Using Artificial Neural Networks. *Int. J. Numer. Anal. Methods Geomech* 45, 1988–2000. doi:10.1002/nag.3252

Moayedi, H., Tien Bui, D., Gör, M., Pradhan, B., and Jaafari, A. (2019). The Feasibility of Three Prediction Techniques of the Artificial Neural Network, Adaptive Neuro-Fuzzy Inference System, and Hybrid Particle Swarm Optimization for Assessing the Safety Factor of Cohesive Slopes. *ISPRS Int. J. Geo-inf.* 8, 391. doi:10.3390/ijgi8090391

Natekin, A., and Knoll, A. (2013). Gradient Boosting Machines, a Tutorial. *Front. Neurorobot.* 7, 21. doi:10.3389/fnbot.2013.00021

Palazzolo, N., Peres, D., Bordoni, M., Meisina, C., Creaco, E., and Cancelliere, A. (2021). Improving Spatial Landslide Prediction with 3D Slope Stability Analysis and Genetic Algorithm Optimization: Application to the Oltrepò Pavese. *Water* 13, 801. doi:10.3390/w13060801

Qian, Z. G., Li, A. J., Chen, W. C., Lyamin, A. V., and Jiang, J. C. (2019). An Artificial Neural Network Approach to Inhomogeneous Soil Slope Stability Predictions Based on Limit Analysis Methods. *Soils Found.* 59, 556–569. doi:10.1016/j.sandf.2018.10.008

Ray, A., Kumar, V., Kumar, A., Rai, R., Khandelwal, M., and Singh, T. N. (2020). Stability Prediction of Himalayan Residual Soil Slope Using Artificial Neural Network. *Nat. Hazards* 103, 3523–3540. doi:10.1007/s11069-020-04141-2

Rukhaiyar, S., Alam, M. N., and Samadhiya, N. K. (2018). A PSO-ANN Hybrid Model for Predicting Factor of Safety of Slope. *Int. J. Geotech. Eng.* 12, 556–566. doi:10.1080/19386362.2017.1305652

Sah, N. K., Sheorey, P. R., and Upadhyaya, L. N. (1994). Maximum Likelihood Estimation of Slope Stability. *Int. J. Rock Mech. Min. Sci.* 31, 47–53. doi:10.1016/0148-9062(94)92314-0

Sakellariou, M. G., and Ferentinou, M. D. (2005). A Study of Slope Stability Prediction Using Neural Networks. *Geotech. Geol. Eng.* 23, 419. doi:10.1007/s10706-004-8680-5

Samui, P. (2008). Slope Stability Analysis: A Support Vector Machine Approach. *Environ. Geol.* 56, 255. doi:10.1007/s00254-007-1161-4

Wang, H. B., Xu, W. Y., and Xu, R. C. (2005). Slope Stability Evaluation Using Back Propagation Neural Networks. *Eng. Geol.* 80 (3), 302–315. doi:10.1016/j.enggeo.2005.06.005

Xu, B., and Low, B. K. (2006). Probabilistic Stability Analyses of Embankments Based on Finite-Element Method. *J. Geotech. Geoenvironmental Eng.* 132, 1444–1454. doi:10.1061/(asce)1090-0241(2006)132:11(1444)

Yang, C. X., Tham, L. G., Feng, X. T., Wang, Y. J., and Lee, P. K. K. (2004). Two-Stepped Evolutionary Algorithm and its Application to Stability Analysis of Slopes. *J. Comput. Civ. Eng.* 18, 145–153. doi:10.1061/(asce)0887-3801(2004)18:2(145)

Zhao, H., and Li, S. (2021). Determining Geomechanical Parameters and a Deformation Uncertainty Analysis of the Longtan Hydropower Station Slope, China. *Bull. Eng. Geol. Environ.* 80 (8), 6429–6443. doi:10.1007/s10064-021-02339-7

Zhao, H. (2008). Slope Reliability Analysis Using a Support Vector Machine. *Comput. Geotech.* 35, 459–467. doi:10.1016/j.compgeo.2007.08.002

Zhao, H., Yin, S., and Ru, Z. (2012). Relevance Vector Machine Applied to Slope Stability Analysis. *Int. J. Numer. Anal. Methods Geomech.* 36, 643–652. doi:10.1002/nag.1037

Zheng, Y., Chen, C., Meng, F., Liu, T., and Xia, K. (2020). Assessing the Stability of Rock Slopes with Respect to Flexural Toppling Failure Using a Limit Equilibrium Model and Genetic Algorithm. *Comput. Geotech.* 124, 103619. doi:10.1016/j.compgeo.2020.103619

Zhou, J., Li, E., Yang, S., Wang, M., Shi, X., Yao, S., et al. (2019). Slope Stability Prediction for Circular Mode Failure Using Gradient Boosting Machine Approach Based on an Updated Database of Case Histories. *Saf. Sci.* 118, 505–518. doi:10.1016/j.ssci.2019.05.046

Zhou, K. P., and Chen, Z. Q. (2009). "Stability Prediction of Tailing Dam Slope Based on Neural Network Pattern Recognition," in 2nd International Conference on Environmental and Computer Science, ICECS 2009 (Dubai, United Arab Emirates: IEEE), 380–383. doi:10.1109/ICECS.2009.55

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Built Environment

ORIGINAL RESEARCH
published: 06 April 2022
doi: 10.3389/fbuil.2022.873546

Applicability of Convolutional Neural Networks for Calibration of Nonlinear Dynamic Models of Structures

Angela Lanning, Arash E. Zaghi and Tao Zhang*

Department of Civil and Environmental Engineering, University of Connecticut, Storrs, CT, United States

The objective of this study is to examine a machine learning (ML) framework for calibrating the parameters of analytical models of complex nonlinear structural systems where experimental data is significantly limited. Because of the high cost of large-scale structural tests, analytical models are widely used to enhance the understanding of structural performance under complex loading environments. In this study, an ML framework is proposed and evaluated for the calibration of an analytical model representing a shake table test performed on a composite column developed in OpenSees software. A large number of parameters for modeling the constitutive behavior of the concrete core, steel reinforcement, exterior composite tube, as well as the interactions between the concrete core and the tube, base fixity, and nonlinear shear deformations are included. A convolutional neural network (CNN) architecture was used to calibrate these parameters by using the lateral load, displacement, and axial load time histories as input variables. First, a synthetic dataset is generated for permutations of different model parameters. Next, four CNNs were trained to evaluate the presentation of input data in time-domain and time-frequency domain. Finally, the trained model was prompted with real experimental data and the values of peak lateral force, residual displacement, and hysteresis energy dissipation from the analytical model were compared with those from the experiment. The results show that the proposed framework is appropriate for calibration of complex nonlinear structural models when experimental data is limited.

Keywords: convolutional neural network, model calibration, image representation, shake table testing, nonlinear structural analysis

OPEN ACCESS

Edited by:
Husam Abu Hajar,
The University of Jordan, Jordan

Reviewed by:
Shehata E. Abdel Raheem,
Assiut University, Egypt
Yasmin Murad,
The University of Jordan, Jordan

****Correspondence:***
Angela Lanning
angela.lanning@uconn.edu

Specialty section:
This article was submitted to Computational Methods in Structural Engineering, a section of the journal Frontiers in Built Environment

Received: *10 February 2022*
Accepted: *15 March 2022*
Published: *06 April 2022*

Citation:
Lanning A, E. Zaghi A and Zhang T (2022) Applicability of Convolutional Neural Networks for Calibration of Nonlinear Dynamic Models of Structures. Front. Built Environ. 8:873546. doi: 10.3389/fbuil.2022.873546

1 INTRODUCTION

This study examines the applicability of a convolutional neural network (CNN)-based machine learning (ML) framework for calibrating the parameters of analytical models of nonlinear mechanical/structural systems with limited experimental data. Large-scale experiments on structural elements, such as bridge columns, are cost prohibitive. Traditionally, analytical models are used to supplement experimental data and investigate effects of a variety of design variables. However, complex nonlinear models contain a large number of parameters that require approximation. This is conventionally done through basic statistical or data-fitting methods to achieve a desired performance, such as an accurate prediction of load-displacement relationships (Willard et al., 2020). The conventional approaches, however, may fail to capture complex interactions of modeling parameters or lead to subjective outcomes when analytical models

involve a large number of parameters requiring calibration. In this work, the seismic response of a composite bridge column is used as the case study for the evaluation of a transfer learning framework, where an ML model is trained with a large synthetic dataset and then used to obtain calibrated parameters for one or a few sets of experimental data. The composite column consisted in this study is a fiber-reinforced polymer (FRP) tube with a reinforced concrete (RC) core. A large-scale model of this column was previously tested at the University of Nevada, Reno on a shake table (Zaghi et al., 2012). An analytical model of the column capturing material plasticity, non-linear composite action of the tube and the concrete core, base rotations, and inelastic shear deformation was developed in the OpenSees structural modeling platform. The large number of modeling variables and their complex nonlinear interactions in a such model may lead traditional calibration methods to have subjective outcomes. Thus, ML, in particular a CNN, is used to calibrate the parameters of the analytical model. The network is trained using the lateral load, axial load, and displacement time histories as the input variables and the analytical modeling parameters as the output variables. The trained network is then used to obtain the calibrated parameters for the experimental data.

ML techniques have been used successfully to identify mechanical characteristics of different types of materials under complex loading conditions, including variable temperatures, high-rate loads, and triaxial stresses (Tang, 2010; Gandomi et al., 2012; Rafiei et al., 2016; Huang and Burton, 2019; Kang et al., 2021). ML techniques, such as artificial neural networks and gene expression programming, have also been used to identify strength of reinforced concrete columns and walls using experimental data (Ilkhani et al., 2019; Murad et al., 2020; Murad, 2021; Naderpour et al., 2022). However, ML techniques require large datasets for the successful training of the model, which are expensive to obtain experimentally. Thus, these studies often rely on previously conducted experiments, which can limit their applicability to well-established materials with ample experimental data, such as reinforced concrete. However, the recent enhancements in computation power have enabled development of large datasets using analytical simulations. These large datasets have promoted the application of deep learning methods for complex nonlinear materials (Liu et al., 2019; Chen et al., 2021), such as the damage progression of composites (Yan et al., 2020). However, the applicability of ML to the calibration of nonlinear mechanical/structural systems under complex loading has not been fully explored, which the framework proposed in this study aims to address. Complex structural systems often contain different types of material and various sources of nonlinearity in addition to the material plasticity. Traditional ML techniques, such as shallow artificial neural networks, often require pre-processing steps to select optimal features from raw data. Deep neural networks, such as CNNs, offer the advantage of extracting inherent features in high-dimensional spaces directly from the data (Sadoughi and Hu, 2019; Kiranyaz et al., 2021). This will enable "matching" the entire shape of the hysteresis curve rather than a few engineering parameters or the backbone curve. A significant portion of the information about the nonlinear response of a complex system is encoded in the hysteresis curves. Thus, a deep learning model that is capable of learning complex visual features will offer promise for the successful calibration of difficult parameters of nonlinear analytical models, and ultimately will enhance the accuracy of simulations.

CNNs are traditionally used for two-dimensional (2D) signals, such as images or videos. The convolutional layers apply kernels over regions of the input to extract complex visual features of an image at different scales, which are presented as feature maps. CNNs capture the spatial relationship between features on an image. Alternatively, one-dimensional (1D) CNNs have been used for signal processing applications, such as structural health monitoring, high-power engine fault monitoring, and structural damage detection by extracting complex features of time-series data (i.e., signals) (Kiranyaz et al., 2015; Avci et al., 2017; Chen and Wang, 2019; Yu et al., 2019; Kiranyaz et al., 2021). In 1D CNNs, the kernel is applied along the time-series in one direction, extracting features from the adjacent timesteps. Researchers have also proposed reshaping the signals into 2D arrays to take advantage of 2D kernels, which use the adjacent rows and columns for feature extraction (Hoang and Kang, 2017; Zhang et al., 2017). In addition to reshaped time-domain signals, researchers have proposed using time-frequency representations to take advantage of the 2D feature extraction of CNNs. Time-frequency transformations let us illustrate the temporal variations of frequency features of a signal, resulting in a 2D matrix that can be processed as an image data. In engineering applications, this is an efficient and conventional approach for analyzing and extracting temporal variations of the features from nonstationary time-series data (Spanos et al., 2007; Nagarajaiah and Basu, 2009; Balli and Palaniappan, 2010; Kong et al., 2017; Kim et al., 2017; Silik et al., 2021). The most common representations are spectrograms and scalograms, which have recently been used with CNNs to extract deep features of signals that typically require manual feature extraction performed by an expert (Xu et al., 2018; Khan et al., 2019; Khare and Bajaj, 2020; Pham et al., 2020; Zhang et al., 2020). While previous studies have shown image representations of engineering signals are suitable for training CNNs, the applications are typically limited to classification tasks or damage identification. Thus, the ability to model calibration of a system with different materials and many sources of nonlinearity has not been fully explored. The framework. To address this, the applicability of the framework to multiple image-based representations of the signals was investigated to establish the suitability of different representation methods to parameter calibration tasks.

This work proposes a novel deep learning framework for calibrating the parameters of analytical models representing complex nonlinear mechanical/structural systems when experimental data is limited (e.g., results from one or two

experiments.) The framework was evaluated by calibrating the parameters of an analytical model representing the shake table response of a composite structural column. The developed analytical model included new modeling approaches for aspects of the composite column that were neglected in previous studies, such as FRP-to-concrete core bond, nonlinear shear deformations, and bond-slip, in an effort to improve the accuracy capturing the lateral force and dissipated energy. The analytical model representing the system was used to generate a large-enough synthetic dataset given different permutations of the parameters requiring calibration. A CNN was trained using the lateral load, axial load, and displacement time histories as the input variables and the model parameters as the output variables. Using a synthetic dataset allowed deep learning techniques to be used without sufficient experimental data and allowed the learning to be centered on the underlying relationships and interactions of these parameters. The synthetic dataset (i.e., analytical results) complied with the physical phenomena, such as constitutive relationships, equilibrium, and energy and momentum conservation, is captured by the analytical model. Thus, the deep learning model learns these concepts from the analytical results rather than a large dataset of physical experiments. Additionally, as the framework uses a dataset of only model inputs and outputs, it can be adapted to analytical models of various complexities with little to no change to its structure. The applicability of the framework to multiple image-based representations of the signals was investigated to establish the suitability of different representation methods to parameter calibration tasks. This included presenting the time histories in the time-domain and time-frequency domain when training a CNN. The time-domain input was used to train two CNNs to compare 1D and 2D convolution kernels. Two time-frequency domain inputs, spectrograms and scalograms, were used. The performance of the trained networks in capturing peak lateral force, residual displacement, and hysteresis energy dissipation in addition to the individual model parameters were used for evaluation. Finally, the trained networks were prompted with the experimental data from the shake table test to obtain the analytical model parameters for the composite column system.

The predicted parameters from the proposed framework captured the experimental response with higher accuracy across the time-series than the previously proposed analytical models for the composite column system. This accuracy demonstrates for the first time that image representations of engineering signals are suitable for training CNNs for model calibration of nonlinear structural systems. The various sources nonlinearity and damping as well as the large number of parameters, allowed benefits of both the time-series and time-frequency presentations to be realized. For this application, the time-series networks had superior performance. The results of the proposed framework also demonstrated that a synthetic dataset can be sufficient for training a CNN that will subsequently be used with experimental data without requiring additional training. This indicates further potential for transfer learning applications using the framework, where it can be retrained with small datasets to performance a related task, such as calibration of a different structural system.

2 METHODS AND APPROACH

An analytical model was developed to simulate the shake table response of a composite bridge column with an FRP tube and RC core that was previously tested at the University of Nevada, Reno (Zaghi et al., 2012). These composite columns were developed as a durable alternative to conventional RC columns. While they have shown superior structural performance to RC columns, the exterior composite tube and its interaction with the concrete core introduce additional modeling complexities that lack established modeling strategies. Additionally, even well-established materials, such as steel and concrete, have properties that vary due to strain-rate effects or parameters characterizing their constitutive relationships that are not easily identifiable from mechanical tests (Kulkarni and Shah, 1998; Zadeh and Saiidi, 2007; Carreño et al., 2020). As such, parameter calibration for the shake table testing of the composite column provided both a relevant and challenging case study for the evaluation of the proposed framework. A total of 25 model parameters requiring calibration were selected from the analytical model. These parameters included the constitutive behavior of the concrete core, steel reinforcement, and exterior composite tube, as well as interactions between the concrete core and the composite tube, global and local damping coefficients, base fixity, and nonlinear shear deformations. The measured displacement and axial load time histories from the shake table test were used as the input signals in the OpenSees model, which then generated the lateral load time history. A synthetic dataset was generated by varying the 25 model parameters. A supervised learning framework was used to learn the relationships between the model parameters as the target vector and displacement, axial load, and lateral load time histories as the input data. The experimental test, composite column system, and analytical model are detailed in the following sections.

2.1 Experimental Test and Composite Column System

The shake table test was performed on a 370-mm diameter by 1511-mm long concrete-filled FRP tube (CFFT) column system (Zaghi et al., 2012). The experimental test setup is shown in **Figure 1A**. The CFFT column had an FRP tube with glass fibers oriented at ±55°, an inner diameter of 355.6 mm, and a thickness of 7 mm. The unconfined concrete had a compressive strength of 47.4 MPa on the test day. The steel reinforcement included 8-#13 longitudinal bars (**Figure 1B**). The FRP tube provided confinement for the concrete core and functioned as transverse reinforcement for shear resistance.

The input motion for the shake table test was a ground acceleration recorded during the 1994 Northridge, CA earthquake (Hauksson et al., 1995). The input motion was applied in seven progressive runs, with acceleration scale factors increasing from 0.1 to 1.9. An axial load of 222 kN was

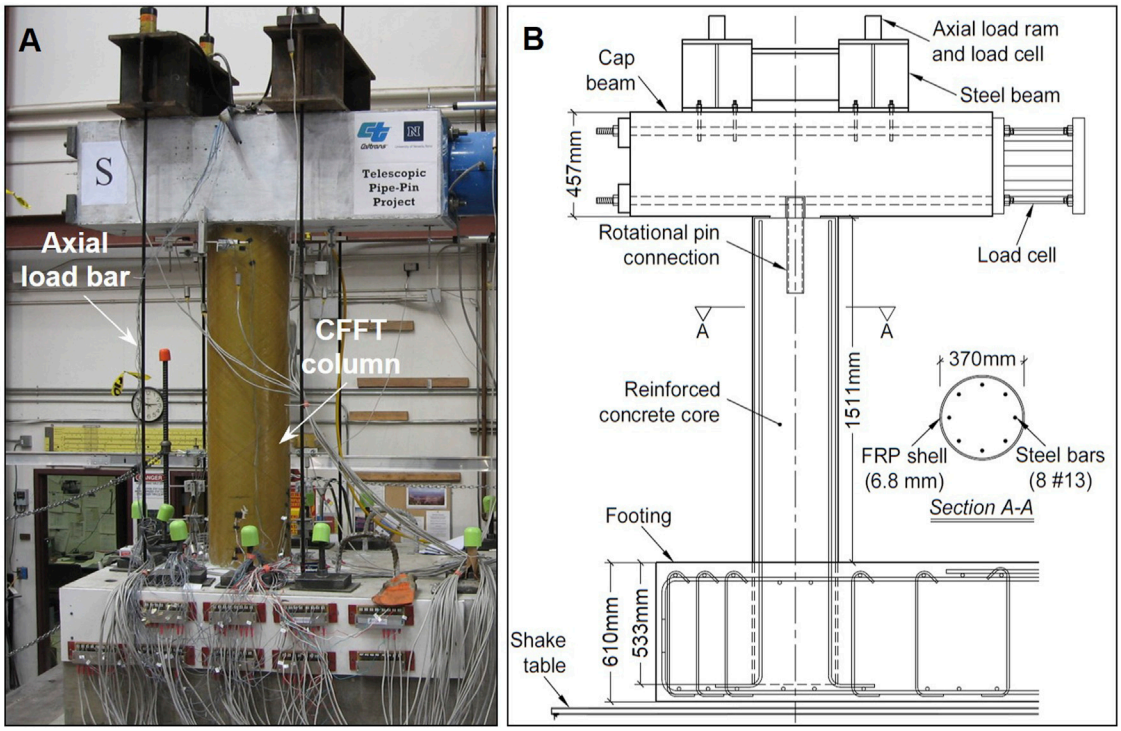


FIGURE 1 | (A) Experimental test setup showing CFFT column on the shake table prior to testing and **(B)** structural detail including components of the test setup and internal steel reinforcement.

applied on the column using four high-strength rods and measured by two load cells during the tests. The axial load fluctuated between 85 and 390 kN during the tests due to the small vertical displacement of the cap beam causing the axial load rods to stretch. Additional details regarding the CFFT column system, the construction of the specimen, and the experimental setup can be found in Zaghi et al. (Zaghi et al., 2012).

2.2 Analytical Model

The analytical modeling was performed in OpenSees, which is an open-source, object-oriented structural analysis framework (Mazzoni et al., 2009). The model was developed to capture the global and local responses of the CFFT column. The schematic view of the nonlinear analytical model is shown in **Figure 2** along with the element types, material commands, and constitutive material models. The CFFT column was modeled using displacement-based elements with fiber cross-sections representing the longitudinal steel bars, FRP-confined concrete core, and longitudinal behavior of the FRP tube. The lateral displacement, $\delta(t)$, and axial load, $P(t)$, time histories recorded during the experimental shake table test were used as input signals for the analytical model. **Figure 2** shows these input functions during Runs 4 through 7. The general analytical model details are discussed below, with a focus on the parameters requiring calibration.

The FRP-confined concrete core behavior was defined using the model proposed by Teng et al. (Teng et al., 2009) and assigned to the uniaxial material "Concrete02". The steel reinforcement was modeled using the uniaxial material "Steel02", defined by the Guiffre-Menegotto-Pinto model with isotropic strain hardening (Filippou et al., 1983). The longitudinal behavior of the FRP tube was modeled using the uniaxial material "Hysteretic" in combination with the uniaxial material "Maxwell" to represent the viscoelastic response of the composite tube. The initial envelope curve for the FRP tube was based on previous experimental and analytical models from Shao et al. (Shao, 2003) and Zaghi et al. (Zaghi et al., 2012) for filament wound FRP tubes with fibers at ±55°. The constitutive material models are provided in **Figure 2**. The reinforced concrete core and the FRP tube were modeled using separate elements, as shown in **Figure 2**, to capture any bond slippage between the FRP tube and concrete core elements (i.e., partial composite action), which was previously observed during the shake table test (Zaghi et al., 2012). To capture this behavior, the cantilever part of each element was discretized into four sections and connected with nonlinear shear and rotational springs using "ZeroLength" elements, with the constitutive behaviors defined in **Figure 2**. The preliminary break point was estimated using the strains measured on the FRP tube and steel bars during the experiment (Zaghi, 2009). Multi-point constraints were defined between the nodes for the FRP tube and the concrete core in the transverse direction (**Figure 2**). The lateral displacement and axial load were applied to the top node of the concrete core element. The initial shear stiffness of the CFFT column was defined using the uncracked concrete section properties. The post-cracking shear stiffness was defined by the cracked concrete section and FRP

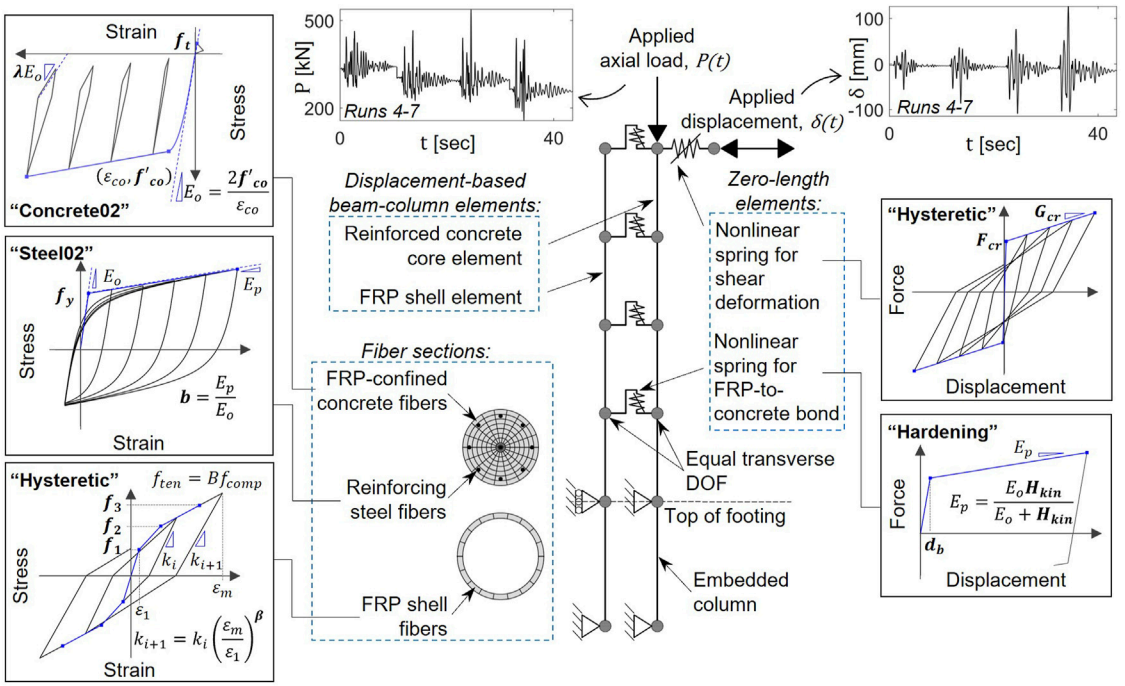


FIGURE 2 | Schematic view of the nonlinear analytical model of a CFFT column modeled using OpenSees. Plots of the applied axial load and displacement time histories and constitutive material behavior are presented. The material commands are identified along with the general shape of the hysteresis curves. The parameters included for calibration are shown in boldface.

TABLE 1 | Details of model parameters to be calibrated by the CNN, including the modeling component, interval of input values, and OpenSees commands.

Analytical Component	Model Parameter	Range	Unit	OpenSees Command
Longitudinal behavior of FRP tube	Backbone stress 1, f_1	20.5–26.5	MPa	uniaxialMaterial Hysteretic
	Backbone stress 2, f_2	48.0–66.5	MPa	
	Backbone stress 3, f_3	75.5–93.0	MPa	
	Tensile backbone factor, B	1.0–1.4	--	
	Unloading stiffness degradation, β_{FRP}	0.1–0.5	--	
	Damping coefficient, C_d	50–2000	kN-s/mm	uniaxialMaterial Maxwell
FRP-confined concrete	Compressive strength, f'_{co}	40.0–70.0	MPa	uniaxialMaterial Concrete02
	Ratio of unloading slope to initial slope, λ	0.2–0.8	--	
	Tensile strength, f_t	1.4–6.2	MPa	
	Hoop stiffness of FRP tube, E_{frp}	7.0–20.0	GPa	
	Hoop rupture strain of FRP tube, ε_{frp}	0.04–0.09	mm/mm	
Steel reinforcement	Yield strength, f_y	480–560	MPa	uniaxialMaterial Steel02
	Strain-hardening ratio, b	0.005–0.04	--	
	Transition from elastic to plastic, R	17.0–19.0	--	
Shear deformation	Shear cracking strength, F_{cr}	0.70–1.50	kN/mm	uniaxialMaterial Hysteretic
	Post-cracking stiffness, G_{cr}	3.50–12.3	kN/mm	
	Unloading stiffness degradation, β_s	0.3–0.7	--	
FRP-concrete composite action	Bond breaking point, d_b	0.001–0.006	--	uniaxialMaterial Hardening
	Kinematic hardening, H_{kin}	0.01–350	--	
	Rotational stiffness, k_θ	2.8e7 - 6.2e8	kN-mm/rad	uniaxialMaterial ElasticPP
Construction misalignment	Compressive yielding force, F_{yc}	275–620	MPa	uniaxialMaterial ElasticPPGap
	Initial gap length, L_{gap}	0.00–7.60	mm	
	Hardening ratio, η	0.01–0.30	--	
Mass- and stiffness-proportional damping	Damping ratio, ξ	2.0–8.0	%	rayleigh
	Frequency 2 at defined damping ratio, ω_j	9.0–18.0	Hz	

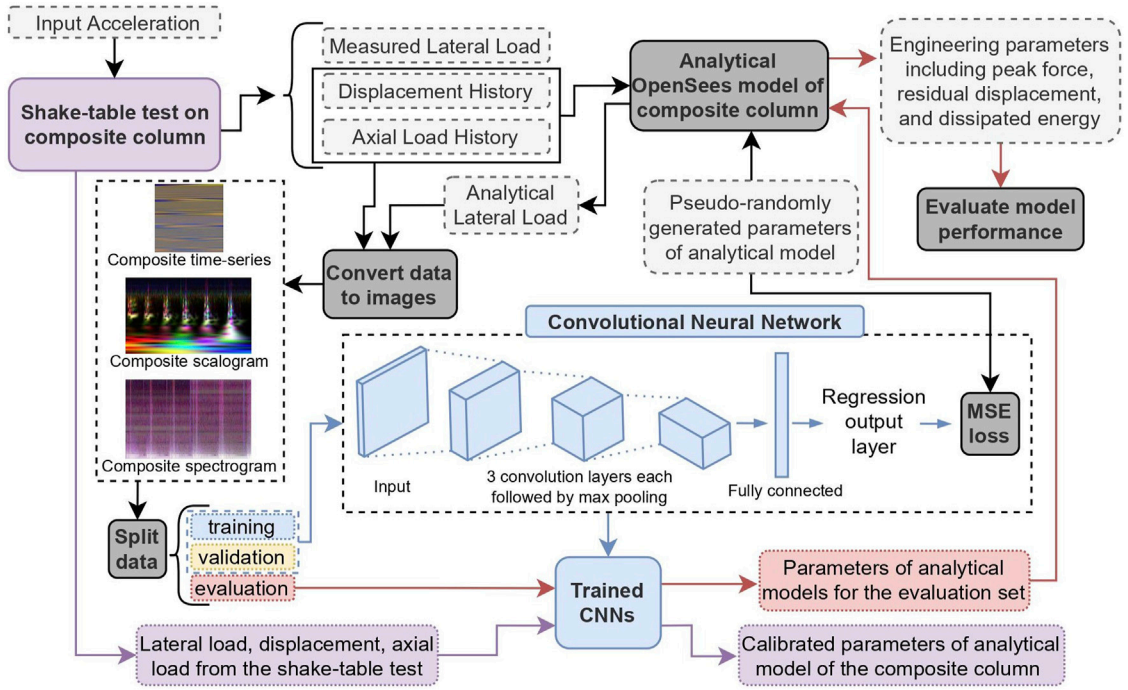


FIGURE 3 | Complete workflow of the proposed framework, including generating the training data, evaluating the trained networks using the engineering metrics, and obtaining calibrated parameters for the composite column system.

tube properties. The embedded portion of the CFFT column was modeled using an embedded element to explicitly capture the effects of base rotation. The effect of potential construction misalignment was captured by two zero-length, compression-only springs at the top of the column with a gap for initial surface asymmetry. Mass- and stiffness-proportional damping were defined using the "rayleigh" command.

2.2.1 Target Variables for the Convolutional Neural Network

Twenty-five analytical model parameters, shown in **Table 1**, that required calibration were defined as target variables when training the CNN. These include material properties for the individual components of the CFFT column, such as the FRP tube, reinforced-concrete core, and steel reinforcement. Parameters defining system-level interactions were also included, such as the nonlinear shear behavior, FRP-to-concrete bond behavior, and construction misalignment. Additionally, the critical-damping ratio and frequency were included, which define the mass- and stiffness-proportional damping coefficients (Wilson, 2004). The parameter details are included in **Table 1**. The schematic constitutive material behaviors are presented in **Figure 2** and include definitions of the calibration parameters, which are shown in bold. The ranges of the parameters were selected to encompass the expected values estimated through material testing, specified material properties, and engineering judgment (Wilson, 2004; Burgueño and Bhide, 2006; Zhu et al., 2006; Teng et al., 2009; Systems, 2012; Zaghi et al., 2012).

3 FRAMEWORK FOR ANALYTICAL MODEL PARAMETER CALIBRATION

The following sections discuss the training data generation, data preparation, and network architectures for the different input representation methods. Next, the evaluation metrics and process for obtaining the calibrated parameters for the composite column system are detailed. The complete workflow of the proposed framework is shown in **Figure 3**. In the proposed workflow, the analytical model representing the composite column is used to generate a synthetic dataset for training a CNN. The input to the analytical models includes the model parameters (**Table 1**) as well as the displacement and axial load time histories measured during the shake table test. The lateral load responses from the analytical model are combined with the displacement and axial load and used as the input data when training the CNN. The corresponding model parameters are the target outputs. After training and evaluating with the synthetic dataset, the CNN is prompted with the experimental data for the composite column and the calibrated model parameters are obtained.

3.1 Training Data

The analytical model previously introduced was used to generate the synthetic dataset for training the CNNs. The model parameters to be calibrated were defined using a quasi-random number generator (QRNG) within the bounds shown in **Table 1**. QRNGs minimize discrepancy between generated points to fill the multi-dimensional space in a uniform manner (Skublska-Rafajłowicz and Rafajłowicz, 2012). This results in a

TABLE 2 | Summary of inputs and trained networks.

Id	Description	Input size (Width x Height x Channels)
TS-1D	Time-series, 1D kernel	123 × 123 × 3
TS-2D	Time-series, 2D kernel	
TF-SC	Time-frequency, Spectrogram	152 × 243 × 3
TF-SP	Time-frequency, Scalogram	150 × 250 × 3

comprehensive representation of the 25 parameters and their effect on the analytical lateral load. The quasi-random number dataset was generated using the Sobol sequence (Sobol', 1967). The first 1,000 points were omitted, and the sequence was scrambled using Owen's scrambling algorithm to reduce correlations and improve uniformity (Matoušek, 1998).

In addition to the target parameters, i.e., calibration parameters, the input displacement and axial load were varied within the synthetic dataset. This was done by adding Gaussian white noise (GWN) to the displacement and axial load time histories prior to running the analytical simulations, as a method of data augmentation and regularization to mitigate overfitting (Rochac et al., 2019). This approach also provided supplementary information on the instantaneous stiffness of the system and time-dependent behavior, such as Rayleigh damping and viscoelasticity. The magnitude of the noise added to the displacement signals varied between 0.25 and 2.5 mm. Random weighting of the two load cells measuring the axial load was applied in addition to GWN with a signal-to-noise ratio (SNR) of 30 dB. Varying the axial load was necessary to provide a comprehensive representation of the asymmetric behavior of the FRP tube and concrete core in tension and compression.

A total of 50,000 samples were generated for the synthetic dataset, where the analytical lateral load varied based on the input displacement, axial load, and model parameters. As one objective of this work is to evaluate and compare the input methods, a large dataset was used to ensure that the performance was not being meaningfully limited by the amount of training data.

3.2 Data Preparation

CNNs were developed primarily for images, which are simply 2D arrays where the cell value defines the pixel intensity. An RGB image stores the red, green, and blue pixel values in separate color channels, which can also be viewed as three different 2D arrays. As such, after converting to 2D arrays, the lateral force, displacement, and axial load signals can be stored in separate channels and used as a single input image for training a CNN. A variety of approaches for converting the signals to images have been implemented in previous studies (Kiranyaz et al., 2021). The main approaches typically fall into two categories 1) using the raw signal in the time domain or 2) transforming the signal into the time-frequency domain. In the first approach, the signal is rescaled so that the amplitude represents the pixel intensity (typically to an integer between 0 and 255). The signal can be reshaped to create an $m \times n$ array, i.e., an image. The engineering applications of this approach are for vibration and electrocardiogram signals (Kolar et al., 2020; Wang et al., 2020; Wang et al., 2021; Zare and Ayati, 2021). In the second approach, a time-frequency transformation is used to represent the signal in the time and frequency domains simultaneously, resulting in a 2D signal that is plotted as an image to visualize the signal activity. The most popular representations are spectrograms and scalograms, which are commonly used in engineering applications to extract features from signals (Spanos et al., 2007; Nagarajaiah and Basu, 2009; Balli and Palaniappan, 2010; Kong et al., 2017; Kim et al., 2017; Silik et al., 2021). Recently, these representations have been used with CNNs to automatically extract features that typically require manual methods informed by expert knowledge (Xu et al., 2018; Khan et al., 2019; Khare and Bajaj, 2020; Pham et al., 2020; Zhang et al., 2020). This work investigated training CNNs for parameter calibration using both the raw time-series signals and the transformed time-frequency representations. The time-series data was used to train two CNNs to evaluate 1D and 2D convolution kernels; these networks are referred to as TS-1D and TS-2D, respectively. Additionally, the time-frequency representations, spectrograms and scalograms, were used to train two CNNs, referred to as TF-SP and TF-SC, respectively. A summary of the input for each of the four networks is provided in **Table 2** and representative inputs are shown in **Figure 4**. The

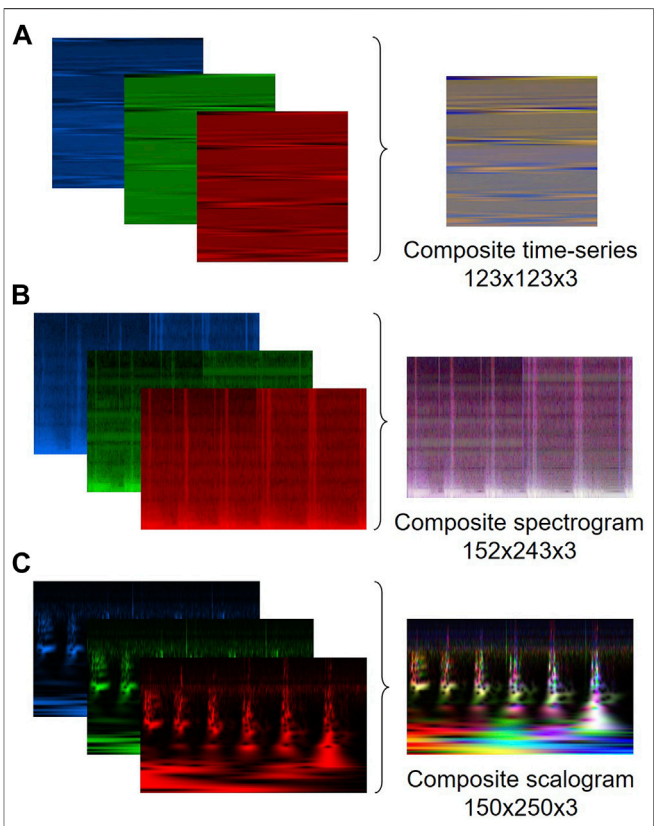


FIGURE 4 | Sample composite input and corresponding lateral load, displacement, and axial load components for the **(A)** time-series **(B)** spectrogram, and **(C)** scalogram inputs.

following sections discuss generating the training dataset as well as background on the approaches.

3.2.1 Generating Time-Domain Training Dataset

Previous applications of this approach have typically been in 1D CNNs, where the convolution kernels extract features from the adjacent points along the time-series (Kiranyaz et al., 2021). These applications are commonly using CNNs for real-time monitoring or damage detection/localization and have limited available training data, making the computationally efficient 1D CNN advantageous. As the proposed framework uses synthetic data, sufficiently large datasets can be generated to train a 2D CNN. A 2D convolution kernel uses both the adjacent rows and columns for feature extraction, which provides information across different cycles in addition to the adjacent timesteps. This may provide a more robust feature extraction as the model parameters typically influence the response across multiple cycles.

This work reshaped the signals by wrapping the 1D arrays into 2D arrays, which can be used to train a CNN with either a 1D or 2D kernel. The lateral load, displacement, and axial load signals reshaped into a 123-by-123 matrix. The signals values were normalized to be integers between 0 and 255. Next, the normalized lateral load, displacement, and axial load signals were assigned to separate color channels within an RGB image to create a single input with dimensions of 123-by-123-by-3. A sample RGB image and the lateral load, displacement, and axial load components are shown in **Figure 4A**. Additional information on 1D and 2D convolution kernels is provided when presenting the CNN architectures.

3.2.2 Generating Time-Frequency Domain Training Datasets

Representing the signals in the time-frequency domain was investigated using spectrograms and scalograms, which are generated by a short-time Fourier transform (STFT) and continuous wavelet transform (CWT), respectively. Time-frequency representations provide a visualization of the signal activity and have been used to extract modal properties and detect damage in structures (Nagarajaiah and Basu, 2009; Kong et al., 2017; Silik et al., 2021). As such, these inputs may facilitate the extraction of frequency-related parameters, such as damping and stiffness. However, this is at the cost of temporal resolution as there is a limit to the degree a signal can be localized both in time and frequency, as shown by Heisenberg's uncertainty principle (Cohen, 1995). The wavelet transform aims to mitigate the limitations due to the trade-off between time and frequency resolutions by applying a variable-size wavelet. This approach allows the time-frequency resolution to change within the scalogram, where the STFT applies a window function with a fixed length, resulting in a constant time-frequency resolution. As a result, the spectrogram is less accurate in capturing time-localized information than the scalogram, especially when the duration of events in a signal differs significantly. CWTs are computationally more complex than STFT (Bouchikhi et al., 2011); thus, slight improvements in accuracy may not warrant the application of scalograms when large datasets are required.

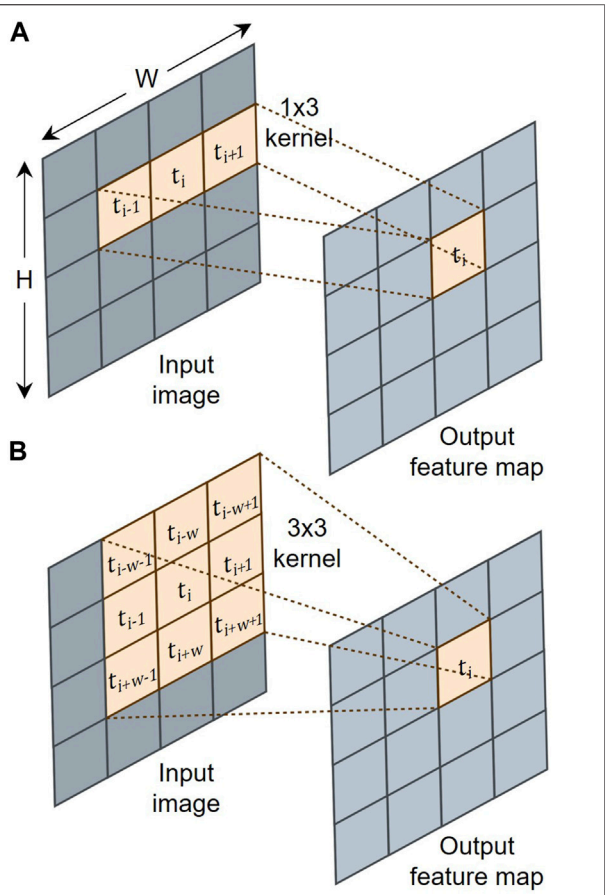


FIGURE 5 | Effect of kernel size and the image width on the output feature map for **(A)** 1D and **(B)** 2D kernels.

The spectrograms were generated using a Hanning window with a width of 120 data points and an overlap length of 50% to prevent information loss due to the Picket Fence effect (Cerna and Harvey, 2000). The magnitude of the STFT was plotted as a function of both time and frequency, shown on the y- and *x*-axis, respectively. The spectrograms were generated for the lateral load, displacement, and axial load signals and assigned to the red, green, and blue channels of a 2D RGB image to create a single composite image as shown in **Figure 4B**.

The second TF representation, scalograms, were obtained by applying a CWT, which dilates and translates the signal along the time axis using a defined wavelet. Analytic Morse wavelets were used with a shape parameter of 3 and time-bandwidth product of 60. These wavelets are symmetric in the time domain and nearly symmetric in the frequency domain while maintaining a high time/frequency concentration (Lilly and Olhede, 2008; Lilly and Olhede, 2012; Lilly, 2017). The modulus of the complex-valued wavelet coefficients obtained from the CWT were scaled by the maximum values at each frequency. The scaled wavelet coefficients were plotted as a function of frequency and time, shown on the y- and *x*-axis, respectively. The frequency was plotted in log-scale and the magnitude of the wavelet coefficients was shown by the color

TABLE 3 | Architecture details for the four trained networks.

Layer type	Description	TS-1d	TS-2d	TF-SP	TF-SC
Image input	*Image size*	123 × 123 × 3	123 × 123 × 3	150 × 242 × 3	150 × 250 × 3
1st Convolution	*Kernel size*	1x3	3x3	3x3	3x3
	No. of feature maps	8	8	8	8
	Output size	123 × 123 × 8	123 × 123 × 8	150 × 242 × 8	150 × 250 × 8
Max pooling	*Pooling size*	1x2	1x2	2x2	2x2
	Output size	123 × 62 × 8	123 × 62 × 8	75 × 121 × 8	75 × 125 × 8
2nd Convolution	*Kernel size*	1x3	3x3	3x3	3x3
	No. of feature maps	16	32	32	32
	Output size	123 × 62 × 16	123 × 62 × 32	75 × 121 × 32	75 × 125 × 32
Max pooling	*Pooling size*	1x2	1x2	2x2	2x2
	Output size	123 × 31 × 16	123 × 31 × 32	38 × 61 × 32	38 × 63 × 32
3rd Convolution	*Kernel size*	1x3	3x3	3x3	3x3
	No. of feature maps	32	128	128	128
	Output size	123 × 62 × 32	123 × 62 × 128	75 × 121 × 128	75 × 125 × 128
Max pooling	*Pooling size*	1x2	1x2	2x2	2x2
	Output size	123 × 16 × 32	123 × 16 × 128	19 × 61 × 128	19 × 32 × 128
Dropout	*Output size*	123 × 16 × 32	123 × 16 × 128	19 × 61 × 128	19 × 32 × 128
Fully connected	*Output size*	1 × 1 × 25	1 × 1 × 25	1 × 1 × 25	1 × 1 × 25

scale. **Figure 4C** shows an example of the composite image obtained using this approach.

3.3 Network Architecture

General CNN architectures contain an image input layer, hidden layers, and an output layer. Hidden layers consist of convolutional layers, pooling layers, and a fully connected layer. Each convolution layer uses a kernel to extract features from local regions of the input. The kernel is a matrix of weights that slides across the input and is convolved with the pixels from a small area of the input to form the output feature maps. Typically, kernels are small and are symmetric around the center pixel, i.e., $k \times k$ where $k > 1$ and is odd. However, applications with time-domain signals commonly use 1D kernels, i.e., $1 \times k$, which require substantially less computational complexity than 2D kernels. For example, using a $k \times k$ kernel to apply convolutions to an $n \times m$ image has a computational complexity of $O(nmk^2)$ while a $1 \times k$ kernel applied under the same conditions has a computational complexity of $O(nmk)$. Despite higher computation costs, a 2D kernel allows the extracted features to be informed by points from different cycles in the time-series depending on the defined width of the reshaped array, in addition to the adjacent points in time. This is demonstrated in **Figure 5**, which compares the receptive field of 1D and 2D kernels. The 2D kernels may be beneficial for parameter calibration as the parameter effects are typically present across cycles in the time-series. As such, both 1D and 2D kernels were investigated for the time-domain input. Additionally, the time-frequency inputs were each used to train a CNN, resulting in a total of four CNNs used to evaluate the proposed framework.

The network architectures are discussed below with the key components provided in **Table 3**. Each network had three convolution layers. Kernel sizes of 3x3 were used, excluding TS-1D, which used 1x3 kernels as discussed. Zero-padding was used so that the convolution layers did not alter the spatial dimensions of the input. Each convolution layer included rectified linear unit (ReLU) as the activation function, which introduced nonlinearity to system. A pooling layer succeeded each convolution layer, which down-sampled values in the feature maps to reduce the number of parameters and dimensionality of the feature maps. A pooling size and stride of 1x2 was used for the time-series networks, which down-sampled the adjacent points in the signals without distorting the order. A pooling size and stride of 2x2 was used for the time-frequency networks. Dropout was introduced prior to the fully connected layer as a regularization method to prevent overfitting (Srivastava et al., 2014). A dropout rate of 20% was used. Finally, a fully connected layer was used to convert the feature map into a feature vector the size of the number of output variables. A sample network architecture is depicted in **Figure 6** for the TS-2D network.

3.4 Training and Validation

The synthetic dataset was divided into training, validation, and evaluation subsets using an 80/15/5 split. Adaptive moment estimation (Adam) optimization algorithm (Kingma and Ba, 2014) was used with an initial learning rate of 0.001. The learning rate decreased by a factor of 0.1 every 10 epochs. During training, accuracy and loss were calculated using the validation dataset to monitor the progress. Mean-squared-error (MSE) was defined as the loss function. Early stopping was used to automatically terminate training when the current validation loss had exceeded the lowest validation loss 10 times.

3.5 Evaluation Metrics

The four trained networks were compared using the evaluation dataset, which included 5% of the overall synthetic data and was unseen during training. As the individual model parameters do not contribute equally to the global response and different combinations can produce similar force-displacement responses, a set of engineering metrics was defined to provide a more meaningful performance criterion. The defined metrics include the peak lateral force (F_p), residual displacement (δ_r), and

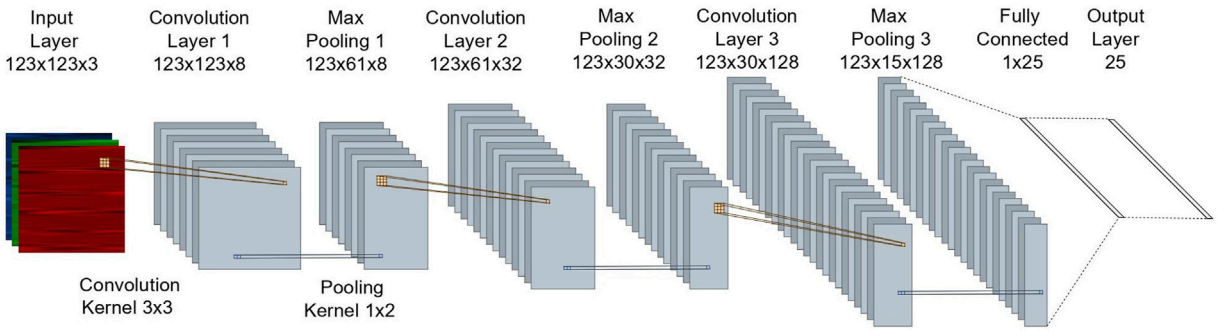


FIGURE 6 | Sample architecture for the TS-2D network using the time-series input and 2D kernel.

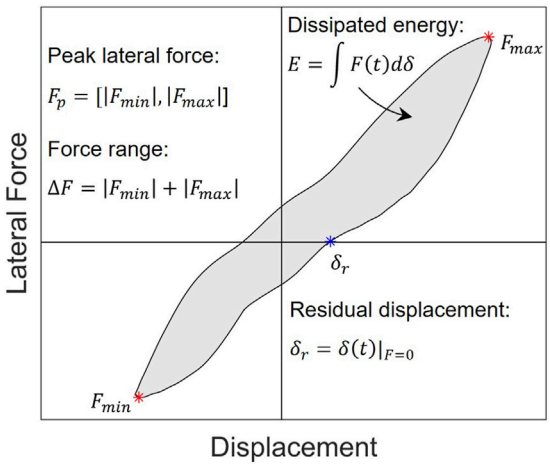


FIGURE 7 | Definition of engineering metrics for a sample force-displacement hysteresis loop.

cumulative dissipated energy (E). These were selected as they are commonly used to evaluate the seismic design or performance of a bridge column. Additionally, the engineering performance metrics can be used to evaluate the predicted parameters for the experimental data. The engineering metrics are defined in **Figure 7** for a representative force-displacement hysteresis curve. For future applications, relevant performance metrics should be selected for evaluation.

The trained networks were prompted with input data from the evaluate set and the predicted parameters were fed back to the analytical model to generate new simulations. The engineering metrics were calculated from the new analytical responses and compared to the actual values for the original dataset. The mean absolute percentage error (MAPE) was used to calculate the error for the engineering metrics as it is a scale-independent accuracy measure, allowing comparison between the metrics with different units. The MAPE is defined as:

$$\text{MAPE}(\%) = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_a - y_p}{y_a}\right| \quad (1)$$

where y_p is the predicted value, y_a is the actual value, and n is the number of samples. The residual displacement was calculated after the largest displacement in the positive and negative loading directions for each run. The average error of the two loading directions was reported for the residual displacement and peak lateral force. As the magnitude of the applied excitations increased throughout the shake table test, the average errors across Runs 1 – 4 and Runs 5 – 7 are reported to compare performance after mid- and high-amplitude excitations.

3.6 Testing Trained Networks With Experimental Data

The model parameters for the composite column system were obtained by prompting the trained network with the lateral load, displacement, and axial load measured during the experimental shake table test. To evaluate the objectivity and sensitivity of the predicted model parameters, the experimental input data was varied by adding GWN with a SNR of 30 dB to the measured time-histories. A total of 10 variations of the experimental data were generated for prompting the trained network. The averages of the predicted values were used in an analytical model to predict the lateral force response of the column system. Finally, the engineering metrics were calculated from the analytical responses and compared to the experimental results to evaluate the accuracy. These steps were completed for each trained network.

4 RESULTS AND DISCUSSION

4.1 Performance of the Networks on Synthetic Data

The performance of the four trained networks was assessed using the accuracy capturing the individual model parameters and the defined engineering metrics. The accuracy in predicting the individual model parameters revealed five parameters were not captured by all four of the trained networks. These parameters include the FRP hoop stiffness, E_{FRP}, hoop rupture strain, ε_{FRP}, and unloading coefficient, λ, defined for the FRP-confined concrete model, as well as the damping coefficient, C_d, defined in the FRP longitudinal model and the rotational stiffness, k_θ,

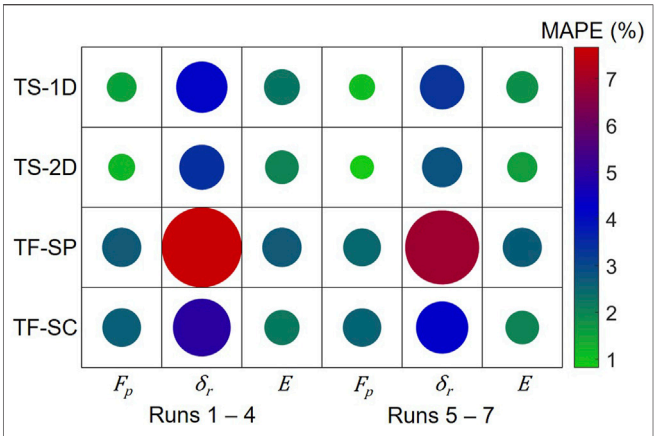


FIGURE 8 | Comparison of each trained network in capturing the engineering metrics when prompted with the evaluation dataset. The size and color of the circles are proportional to the error values.

defined in the FRP-to-concrete bond model. The R^2 values for these parameters were less than 0.10 across the four networks. The uniform low prediction accuracy for these parameters indicates there was insufficient information in the training data, likely due to the small effect the parameters had on the global response. If calibration of parameters controlling localized behavior is required, additional measurements capturing these effects can be included when training, such as strain gauge measurements. These parameters were not removed from the training data as the small effect on the generated responses served as an additional source of noise, mitigating overfitting. The accuracy capturing the individual model parameters is discussed further in relation to the engineering metrics. The cumulative MSE for the 25 model parameters were 1.11, 1.02, 1.55, and 1.16 for TS-1D, TS-1D, TF-SP, and TF-SC, respectively. The engineering metrics were calculated from the simulated responses given the network predicted parameters, as previously discussed. The MAPE values for the trained networks are compared in **Figure 8** for Runs 1—4 and Runs 5—7. The TS-2D network had the lowest average error for the engineering metrics as well as the individual model parameters, followed by TS-1D, TF-SC, and TF-SP, respectively.

The time-series networks performed notably better on the peak lateral force and residual displacement than the time-frequency networks. Residual displacement is the displacement corresponding to zero force, which requires maintaining information on the relationship between the force and displacement signals, in addition to adequate time resolution that is also required to capture the peak lateral force. This can be accomplished by preserving the order or phase information of the signals. The time-series inputs maintain temporal information as the order and increment of the signals are preserved. While the time-frequency inputs maintain some temporal information, the resolution is limited due to the time-frequency resolution trade-off. As previously discussed, scalograms mitigate this by applying a variable-sized wavelet, allowing TF-SC to outperform TF-SP for residual displacement. The overlapping window length for spectrograms lessens information loss at peaks but does not improve temporal resolution. As such, the accuracy of the time-frequency networks was comparable for the peak lateral force.

The energy dissipation mechanisms in the analytical model were the material hysteretic relationships and Rayleigh damping, which is a form of viscous damping commonly used to approximate the non-hysteretic sources of energy dissipation in structures (Priestley and Grant, 2005; Petrini et al., 2008). Energy dissipation can be measured by the lag between when the displacement reaches its maximum and when the force does, which can be identified in the inputs through temporal shifts between the force and displacement signals. As the Rayleigh damping is also frequency dependent, the time-frequency representations contain additional information that can be used for identification. In the analytical model, the level of Rayleigh damping was primarily controlled by the critical damping ratio, ξ. The hysteretic sources of energy dissipation were controlled by the inelastic material properties, such as the post-cracking shear stiffness, G_{cr}, and unloading coefficient, β_s, for the shear deformation model. The accuracy in capturing these three model parameters as well as the cumulative energy dissipation, E, is provided in **Figure 9**, which includes the actual versus predicted values and the corresponding R^2 value.

As shown in **Figure 9**, the four networks had comparable accuracy for the parameters impacting the hysteretic sources of energy dissipation (G_{cr} and β_s), which are frequency independent and can be clearly characterized by the force-displacement relationship. However, the time-frequency networks performed considerably better than the time-series networks for the critical damping ratio, as expected due to the frequency-dependent behavior. Despite this, the time-series networks captured the cumulative energy dissipation with comparable accuracy to the time-frequency networks due to the large contribution of material nonlinearity. As such, the time-series networks had higher accuracy for the energy dissipation after the high-amplitude excitation than after the mid-amplitude excitation when there was less material plasticity. This allowed them to outperform the time-frequency networks after the high-level excitation, as shown in **Figure 8**. The sources of energy dissipation should be considered when selecting the input method.

4.2 Performance in Capturing the Experimentally Measured Response

The model parameters for the composite column system were obtained by prompting the trained network with the experimental data with added random noise as previously discussed. The predicted parameters for the four trained networks were within the range of estimated values from material testing, constitutive equations, or manufacturer specifications. This indicates that the networks did not compensate with unrealistically large or small values when prompted with experimental data. **Table 4** includes the predicted values for 11 parameters. These parameters strongly influence the global response and demonstrate key sources of

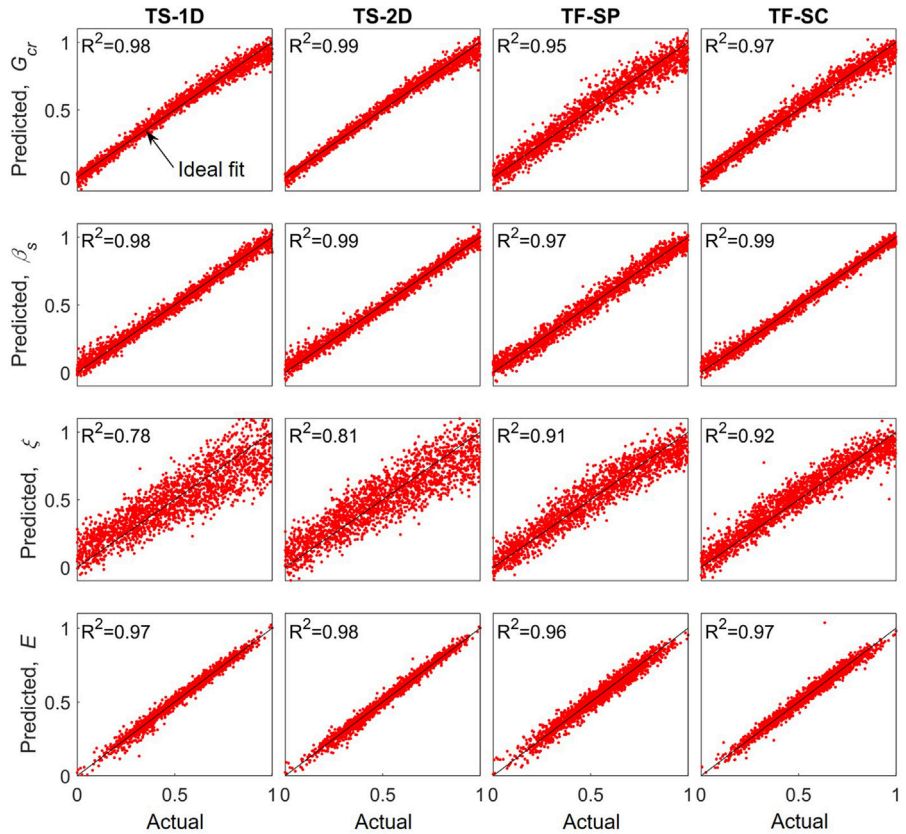


FIGURE 9 | Actual versus predicted post-cracking shear stiffness, G_{cr}, shear deformation unloading coefficient, β_s, damping ratio, ξ, and cumulative energy dissipation, E, for each representation method.

TABLE 4 | Predicted values of key model parameters for experimental data with added noise.

Analytical Component	Model Parameter	Predicted Parameter Values (Mean ± SD)				Units
		TS-1d	TS-2d	TF-SP	TF-SC	
Longitudinal behavior of FRP tube	Backbone stress 1, f_1	25.5 ± 0.1	22.1 ± 0.1	42.1 ± 0.5	25.1 ± 0.6	MPa
	Backbone stress 2, f_2	50.0 ± 0.2	50.3 ± 0.3	54.4 ± 0.4	37.2 ± 0.4	MPa
	Backbone stress 3, f_3	85.3 ± 0.1	88.4 ± 0.3	87.5 ± 0.6	92.6 ± 0.5	MPa
	Unloading stiffness degradation, β_{FRP}	0.34 ± 0.01	0.30 ± 0.01	0.01 ± 0.02	0.54 ± 0.02	--
FRP-confined concrete	Compressive strength, f'_{co}	65.5 ± 0.4	62.7 ± 0.8	55.2 ± 1.7	55.4 ± 0.7	MPa
Steel rebar	Yield strength, f_y	475 ± 1	465 ± 2	437 ± 6	503 ± 3	MPa
Shear deformation	Shear cracking strength, F_{cr}	1.01 ± 0.01	1.16 ± 0.01	1.50 ± 0.04	0.64 ± 0.02	kN/mm
	Post-cracking stiffness, G_{cr}	10.2 ± 0.1	10.3 ± 0.1	2.8 ± 0.1	9.8 ± 0.2	kN/mm
	Unloading stiffness degradation, β_s	0.68 ± 0.002	0.66 ± 0.005	0.74 ± 0.005	0.59 ± 0.007	--
Construction misalignment	Initial gap length, β_s	1.7 ± 0.1	1.5 ± 0.1	-0.2 ± 0.3	2.5 ± 0.2	mm
Rayleigh damping	Damping ratio, ξ	2.1 ± 0.1	1.8 ± 0.1	3.4 ± 0.2	4.6 ± 0.2	%

discrepancy between the trained networks. This discrepancy is discussed below in relation to the global response and engineering metrics. The accuracy of the calibrated models in capturing the engineering metrics was also compared to an analytical model previously developed by (Zaghi et al., 2012), referred to herein as the "Reference Model". The model was developed in OpenSees using conventional calibration approaches, with the full details found in (Zaghi, 2009). The lateral displacement was applied directly to the Reference Model instead of the ground acceleration to facilitate comparison to the current study.

The average predicted parameters were fed to the OpenSees analytical model to simulate the response of the composite column. The resulting force-displacement response and corresponding energy dissipation are shown in **Figure 10A**

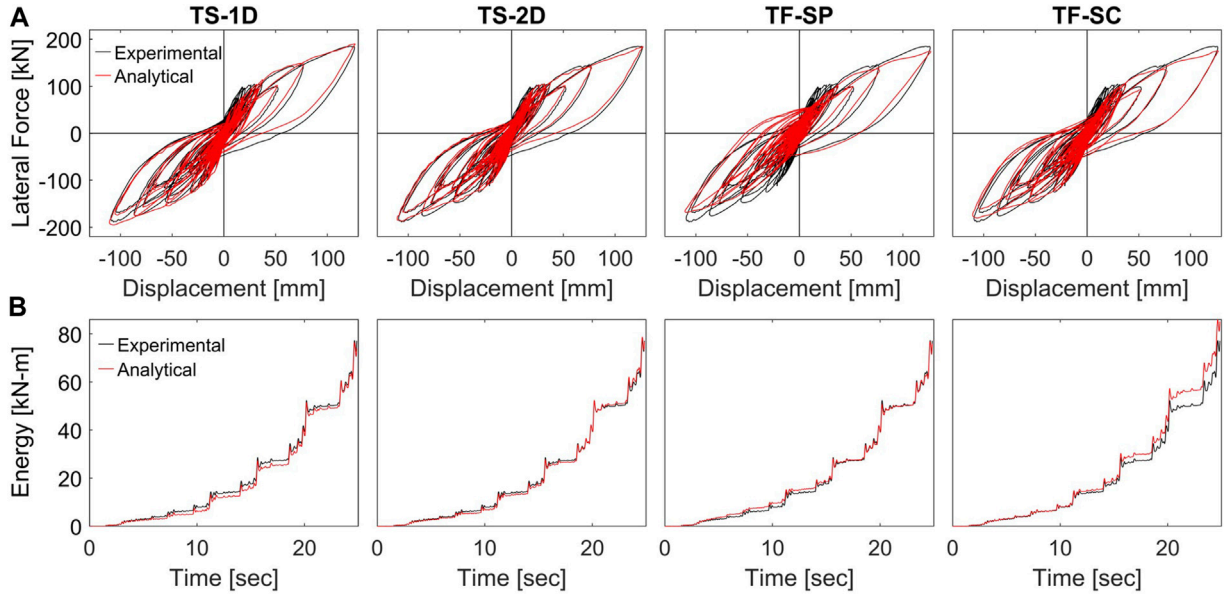


FIGURE 10 | Analytical model response given network predicted parameters including **(A)** global lateral force-displacement and **(B)** energy dissipation.

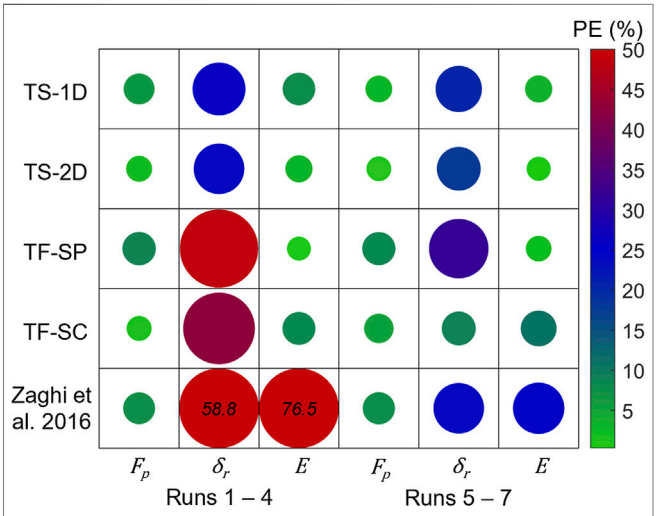


FIGURE 11 | Performance of the CNN-predicted model parameters in capturing the engineering metrics when prompted with the experimental data. The size and color of the circles are proportional to the error values.

and **Figure 10B**, respectively. The energy dissipation was calculated by integrating the measured force-displacement hysteresis curves. The time between cycles was removed for presentation purposes. The R^2 values were calculated using the predicted force and energy and the measured lateral force and dissipated energy to demonstrate the accuracy across the time histories. The R^2 values for the lateral force were 0.947, 0.976, 0.852, 0.926, and 0.936 for TS-1D, TS-1D, TF-SP, TF-SC, and the Reference Model, respectively. The R^2 values for the dissipated energy were 0.993, 0.997, 0.996, 0.924, and 0.831 for TS-1D, TS-1D, TF-SP, TF-SC, and the Reference Model, respectively. Additionally, the engineering metrics were calculated from the analytical responses and compared to the experimental values. The percent error for each metric is compared in **Figure 11**, including the accuracy of the Reference Model developed by (Zaghi et al., 2012). The color scale axis was set to 50 to facilitate comparison of the smaller errors. Errors over this are reported on the figure.

The accuracy capturing the engineering metrics was generally lower for the experimental data than for the synthetic dataset, which was anticipated as no experimental data was included in training and the analytical model is limited by the approximations in the experimental setup and constitutive relationships. Overall, the trends in performance across the engineering metrics and between the four networks were largely comparable to what was observed for the synthetic dataset. The predicted responses were reasonable for TS-1D, TS-2D, and TF-SC, demonstrating the ability of the networks to generalize to experimental data after training on a synthetic dataset. Additionally, the strong performance of the time-series networks and TF-SC indicate that the framework is applicable to both time-domain and time-frequency domain inputs. The poor performance of TF-SP is likely due to differences between the high frequency artifacts present in the experimental data and synthetic data. This may be mitigated by additional filtering of the predicted responses; however, this may remove information useful during training and is not a necessary step for the other inputs methods. While the strong performance of TF-SC proved the framework is applicable to time-frequency inputs, scalograms require substantially more computation time than the time-series inputs and did not provide substantial improvement in accuracy to justify the use for

this application. However, if the objective is calibration of frequency-dependent parameters, using scalograms may be warranted.

The TS-2D network had the lowest errors across the engineering metrics, as shown in **Figure 11**. Additionally, the difference in performance from the synthetic dataset was smaller for TS-2D than TS-1D. This suggests the 2D kernel provided additional benefit for generalizing to experimental data that was not observed or necessary for the synthetic dataset. The 2D kernel provides information at multiple cycles, such as the loading and unloading slope, which allows TS-2D to balance the contribution of parameters to multiple points of the signal and likely provide more robust feature extraction.

Despite the stronger performance of TS-2D, the time-series networks had comparable predicted parameters with low standard deviations, as shown in **Table 4**. The time-frequency networks, in contrast, had less consistency between predicted values. A key difference in the predicted parameters is the source of energy dissipation. The time-series networks relied on a similar balance between the hysteretic and non-hysteretic sources of energy dissipation. This is demonstrated by the predicted values for ξ, which controls Rayleigh damping, in comparison to $G_{cr,}$ β_s, and β_{FRP}, which control material nonlinearity, shown in **Table 4**. As the energy dissipation is not provided during training, the time-series networks rely on the relationship between force and displacement to accurately capture the energy dissipation. The time-frequency inputs use frequency-dependent behavior of Rayleigh damping, which allowed for higher accuracy on the synthetic dataset. However, Rayleigh damping is a mathematically convenient method of approximating non-hysteretic sources of energy dissipation and is not directly related to one physical process. To ensure the contribution of hysteretic sources of nonlinearity is accurate, local measurements, such as strain gauges, can be used. Future application of the trained network could be feasible through the use of transfer-learning, which repurposes the trained network to perform a second related task. This could be done to calibrate parameters for a different experimental setup or column geometry. As this requires less training data, this may also be done with experimental data, if available.

The errors for the Reference Model are larger in the early cycles. This was attributed to an underestimation of hysteretic damping at low amplitudes in the analytical model, as noted in (Zaghi et al., 2012). This hysteretic damping is largely due to concrete cracking under small deformations, which was accounted for in the current study through the nonlinear shear model. As a result, the energy dissipation as well as the residual displacements in the low-amplitude cycles were predicted with higher accuracy. Additionally, a high damping ratio of 15% was used in the Reference Model to account for the high damping of the FRP tube and other sources of damping; however, the energy dissipation was still under-predicted. The current model incorporated individual components to represent this (i.e., viscoelastic behavior of the FRP shell and FRP-concrete composite action) and was able to calibrate the additional parameters with the framework. As a result, the energy dissipation was captured across the entire time history with higher accuracy than the Reference Model. Overall, the performance of the Reference Model validates the strong performance of the network predicted parameters.

5 CONCLUSION

This work proposed a novel supervised learning framework to calibrate the model parameters of complex structural systems with limited experimental data. The framework was evaluated by calibrating the parameters of an analytical model representing the shake table response of a composite column. A CNN was trained using a synthetic dataset to learn the relationships between the analytical model parameters and the displacement, axial load, and analytical lateral load time histories. The applicability of the framework to presenting the signals in both the time-domain and time-frequency domain was investigated. The time-domain input was used to train two CNNs to compare 1D and 2D convolutional kernels. The time-frequency inputs were used to train two CNNs to compare scalograms and spectrograms. Each trained network was prompted with the experimental data from the shake table test to obtain analytical model parameters for the composite column system.

The results of the study demonstrated a CNN can perform parameter calibration for nonlinear dynamic models of structures. When prompted with the experimental data, the trained CNN predicted reasonable model parameters and captured the measured seismic response of the composite column. The accuracy capturing the measured lateral force-displacement response and dissipated energy was higher than the previously developed analytical model for the composite column system. The results demonstrate that a synthetic dataset can be sufficient for training a CNN and the train network can be subsequently used with experimental data without requiring additional training. This indicates additional potential for transfer learning applications using the framework, where it can be retrained with small datasets to performance a related task, such as calibration of a different structural system. Additionally, the results demonstrated image representations of engineering signals are suitable for training CNNs for model calibration of nonlinear structural systems. Presenting the input signals in both the time-domain and time-frequency domain both proved to be a feasible and effective approach for parameter calibration. The networks trained with time-series inputs had lower errors for both the synthetic dataset and the experimental data than for the time-frequency inputs. The time-series inputs also required fewer preprocessing steps and less computation time. Additionally, the time-series inputs benefitted from 2D convolution kernels in place of 1D kernels, which are typically used for signal-processing applications of CNNs. In conclusion, CNN-based parameter calibration presents an effective way of calibrating many parameters of complex structural models with limited experimental data.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

AL designed the study, implemented the program, and wrote the initial draft of the manuscript. AE contributed to conception of the study and interpretation of data. TZ contributed to implementation of the program. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This project was supported by the National Science Foundation Partnership for Innovation Award #1500293. The student support for this project was funded by the National Science Foundation Graduate Research Fellowship under Grant No. 1747453.

ACKNOWLEDGMENTS

The authors gratefully acknowledged the guidance of Dr. Hadi Meidani of the University of Illinois at Urbana-Champaign. The support of Dr. Alexandra Hain of the University of Connecticut and Michael Vaccaro Jr is also appreciated.

REFERENCES

Avci, O., Abdeljaber, O., Kiranyaz, S., and Inman, D. (2017). "Structural Damage Detection in Real Time: Implementation of 1D Convolutional Neural Networks for SHM Applications," in *Structural Health Monitoring & Damage Detection* (Berlin, Germany: Springer), Vol. 7. doi:10.1007/978-3-319-54109-9_6

Balli, T., and Palaniappan, R. (2010). Classification of Biological Signals Using Linear and Nonlinear Features. *Physiol. Meas.* 31, 903–920. doi:10.1088/0967-3334/31/7/003

Bouchikhi, E. H., Choqueuse, V., Benbouzid, M., Charpentier, J.-F., and Barakat, G. (2011). "A Comparative Study of Time-Frequency Representations for Fault Detection in Wind Turbine," in IECON 2011-37th Annual Conference of the IEEE Industrial Electronics Society, Melbourne, Australia, November 7–10, 2011 (IEEE), 3584–3589. doi:10.1109/iecon.2011.6119891

Burgueño, R., and Bhide, K. M. (2006). Shear Response of concrete-filled FRP Composite Cylindrical Shells. *J. Struct. Eng.* 132, 949–960. doi:10.1061/(ASCE)0733-9445(2006)132:6(949)

Carreño, R., Lotfizadeh, K., Conte, J., and Restrepo, J. (2020). Material Model Parameters for the Giuffrè-Menegotto-Pinto Uniaxial Steel Stress-Strain Model. *J. Struct. Eng.* 146, 04019205. doi:10.1061/(ASCE)ST.1943-541X.0002505

Cerna, M., and Harvey, A. F. (2000). *The Fundamentals of FFT-Based Signal Analysis and Measurement. Application Note 041*. Austin, TX: National Instruments Corporation.

Chen, J., Wan, L., Ismail, Y., Ye, J., and Yang, D. (2021). A Micromechanics and Machine Learning Coupled Approach for Failure Prediction of Unidirectional CFRP Composites under Triaxial Loading: A Preliminary Study. *Compos. Structures* 267, 113876. doi:10.1016/j.compstruct.2021.113876

Chen, Y.-Y., and Wang, Z.-B. (2019). Feature Selection Based Convolutional Neural Network Pruning and its Application in Calibration Modeling for NIR Spectroscopy. *Chemometrics Intell. Lab. Syst.* 191, 103–108. doi:10.1016/j.chemolab.2019.06.004

Cohen, L. (1995). *Time-frequency Analysis*. Englewood Cliffs, NJ: Prentice-Hall.

Filippou, F. C., Bertero, V. V., and Popov, E. P. (1983). *Effects of Bond Deterioration on Hysteretic Behavior of Reinforced concrete Joints*. Berkeley, California: Earthquake Engineering Research Center.

Gandomi, A. H., Babanajad, S. K., Alavi, A. H., and Farnam, Y. (2012). Novel Approach to Strength Modeling of concrete under Triaxial Compression. *J. Mater. Civ. Eng.* 24, 1132–1143. doi:10.1061/(asce)mt.1943-5533.0000494

Hauksson, E., Jones, L. M., and Hutton, K. (1995). The 1994 Northridge Earthquake Sequence in California: Seismological and Tectonic Aspects. *J. Geophys. Res.* 100, 12335–12355. doi:10.1029/95jb00865

Hoang, D.-T., and Kang, H.-J. (2017). "Convolutional Neural Network Based Bearing Fault Diagnosis," in International conference on intelligent computing, Liverpool, UK, August 7–10, 2017 (Springer), 105–111. doi:10.1007/978-3-319-63312-1_9

Huang, H., and Burton, H. V. (2019). Classification of In-Plane Failure Modes for Reinforced concrete Frames with Infills Using Machine Learning. *J. Building Eng.* 25, 100767. doi:10.1016/j.jobe.2019.100767

Ilkhani, M. H., Naderpour, H., and Kheyroddin, A. (2019). A Proposed Novel Approach for Torsional Strength Prediction of RC Beams. *J. Building Eng.* 25, 100810. doi:10.1016/j.jobe.2019.100810

Kang, M.-C., Yoo, D.-Y., and Gupta, R. (2021). Machine Learning-Based Prediction for Compressive and Flexural Strengths of Steel Fiber-Reinforced concrete. *Construction Building Mater.* 266, 121117. doi:10.1016/j.conbuildmat.2020.121117

Khan, A., Ko, D.-K., Lim, S. C., and Kim, H. S. (2019). Structural Vibration-Based Classification and Prediction of Delamination in Smart Composite Laminates Using Deep Learning Neural Network. *Composites B: Eng.* 161, 586–594. doi:10.1016/j.compositesb.2018.12.118

Khare, S. K., and Bajaj, V. (2020). Time-frequency Representation and Convolutional Neural Network-Based Emotion Recognition. *IEEE Trans. Neural networks Learn. Syst.* 32 (7), 2901–2909.

Kim, B., Jeong, H., Kim, H., and Han, B. (2017). Exploring Wavelet Applications in Civil Engineering. *KSCE J. Civ Eng.* 21, 1076–1086. doi:10.1007/s12205-016-0933-3

Kingma, D. P., and Ba, J. (2014). "Adam: A Method for Stochastic Optimization" in 3rd International Conference for Learning Representations (ICLR), San Diego, US, May 7–9, 2015. *arXiv preprint arXiv:1412.6980*.

Kiranyaz, S., Ince, T., Hamila, R., and Gabbouj, M. (2015). "Convolutional Neural Networks for Patient-specific ECG Classification," in 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2015, Milano, Italy, August 25–29, 2015 (IEEE), 2608–2611. doi:10.1109/EMBC.2015.7318926

Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., and Inman, D. J. (2021). 1D Convolutional Neural Networks and Applications: A Survey. *Mech. Syst. Signal Process.* 151, 107398. doi:10.1016/j.ymssp.2020.107398

Kolar, D., Lisjak, D., Pająk, M., and Pavković, D. (2020). Fault Diagnosis of Rotary Machines Using Deep Convolutional Neural Network with Wide Three axis Vibration Signal Input. *Sensors* 20, 4017. doi:10.3390/s20144017

Kong, X., Cai, C. S., Deng, L., and Zhang, W. (2017). Using Dynamic Responses of Moving Vehicles to Extract Bridge Modal Properties of a Field Bridge. *J. Bridge Eng.* 22, 04017018. doi:10.1061/(asce)be.1943-5592.0001038

Kulkarni, S. M., and Shah, S. P. (1998). Response of Reinforced concrete Beams at High Strain Rates. *Struct. J.* 95, 705–715. doi:10.14359/584

Lilly, J. M. (2017). Element Analysis: a Wavelet-Based Method for Analysing Time-Localized Events in Noisy Time Series. *Proc. R. Soc. A.* 473, 20160776. doi:10.1098/rspa.2016.0776

Lilly, J. M., and Olhede, S. C. (2012). Generalized Morse Wavelets as a Superfamily of Analytic Wavelets. *IEEE Trans. Signal. Process.* 60, 6036–6041. doi:10.1109/tsp.2012.2210890

Lilly, J. M., and Olhede, S. C. (2008). Higher-order Properties of Analytic Wavelets. *IEEE Trans. Signal Process.* 57, 146–160. doi:10.1109/TSP.2008.2007607

Liu, X., Gasco, F., Goodsell, J., and Yu, W. (2019). Initial Failure Strength Prediction of Woven Composites Using a New Yarn Failure Criterion Constructed by Deep Learning. *Compos. Structures* 230, 111505. doi:10.1016/j.compstruct.2019.111505

Matoušek, J. (1998). On Thel2-Discrepancy for Anchored Boxes. *J. Complexity* 14, 527–556. doi:10.1006/jcom.1998.0489

Mazzoni, S., Mckenna, F., Scott, M. H., and Fenves, G. L. (2009). *OpenSees Command Language Manual*. Berkeley: Pacific Earthquake Engineering Research Center, University of California.

Murad, Y., Hunifat, R., and Wassel, A.-B. 2020. "Interior Reinforced Concrete Beam-to-Column Joints Subjected to Cyclic Loading: Shear Strength Prediction Using Gene Expression Programming." *Case Stud. Constr. Mater.* 13, e00432. doi:10.1016/j.cscm.2020.e00432

Naderpour, H., Sharei, M., Fakharian, P., and Heravi, M. A. (2022). Shear Strength Prediction of Reinforced Concrete Shear Wall Using ANN, GMDH-NN and GEP. *Soft Comput. Civil Eng.* 6, 66–87. doi:10.22115/SCCE.2022.283486.1308

Nagarajaiah, S., and Basu, B. (2009). Output Only Modal Identification and Structural Damage Detection Using Time Frequency & Wavelet Techniques. *Earthq. Eng. Eng. Vib.* 8, 583–605. doi:10.1007/s11803-009-9120-6

Petrini, L., Maggi, C., Priestley, M. J. N., and Calvi, G. M. (2008). Experimental Verification of Viscous Damping Modeling for Inelastic Time History Analyzes. *J. Earthquake Eng.* 12, 125–145. doi:10.1080/13632460801925822

Pham, M. T., Kim, J.-M., and Kim, C. H. (2020). Accurate Bearing Fault Diagnosis under Variable Shaft Speed Using Convolutional Neural Networks and Vibration Spectrogram. *Appl. Sci.* 10, 6385. doi:10.3390/app10186385

Priestley, M. J. N., and Grant, D. N. (2005). Viscous Damping in Seismic Design and Analysis. *J. earthquake Eng.* 9, 229–255. doi:10.1142/s1363246905002365

Rafiei, M. H., Khushefati, W. H., Demirboga, R., and Adeli, H. (2016). Neural Network, Machine Learning, and Evolutionary Approaches for Concrete Material Characterization. *ACI Mater. J.* 113, 781–789. doi:10.14359/51689360

Rochac, J. F. R., Zhang, N., Xiong, J., Zhong, J., and Olandunni, T. (2019). "Data Augmentation for Mixed Spectral Signatures Coupled With Convolutional Neural Networks" in 9th International Conference on Information Science and Technology (ICIST) (Hulunbuir, China: IEEE), 402–407. doi:10.1109/ICIST.2019.8836868

Sadoughi, M., and Hu, C. (2019). Physics-based Convolutional Neural Network for Fault Diagnosis of Rolling Element Bearings. *IEEE Sensors J.* 19, 4181–4192. doi:10.1109/jsen.2019.2898634

Shao, Y. (2003). *Behavior of FRP-Concrete Beam-Columns Under Cyclic Loading*. Raleigh, NC: North Carolina State University.

Silik, A., Noori, M., A. Altabey, W., Ghiasi, R., and Wu, Z. (2021). Comparative Analysis of Wavelet Transform for Time-Frequency Analysis and Transient Localization in Structural Health Monitoring. *Struct. Durability Health Monit.* 15, 1–22. doi:10.32604/sdhm.2021.012751

Skublska-Rafajłowicz, E., and Rafajłowicz, E. (2012). Sampling Multidimensional Signals by a New Class of Quasi-Random Sequences. *Multidimensional Syst. Signal Process.* 23, 237–253. doi:10.1007/s11045-010-0120-5

Sobol', I. Y. M. (1967). On the Distribution of Points in a Cube and the Approximate Evaluation of Integrals. *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki* 7, 784–802. doi:10.1016/0041-5553(67)90144-9

Spanos, P. D., Giaralis, A., and Politis, N. P. (2007). Time-frequency Representation of Earthquake Accelerograms and Inelastic Structural Response Records Using the Adaptive Chirplet Decomposition and Empirical Mode Decomposition. *Soil Dyn. Earthquake Eng.* 27, 675–689. doi:10.1016/j.soildyn.2006.11.007

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: a Simple Way to Prevent Neural Networks from Overfitting. *J. machine Learn. Res.* 15, 1929–1958.

Systems, N. F. G. (2012). *Red Thread II Piping Systems*. San Antonio, TX: National Oilwell Varco.

Tang, C.-W. (2010). Radial Basis Function Neural Network Models for Peak Stress and Strain in plain concrete under Triaxial Stress. *J. Mater. Civ. Eng.* 22, 923–934. doi:10.1061/(asce)mt.1943-5533.0000077

Teng, J. G., Jiang, T., Lam, L., and Luo, Y. Z. (2009). Refinement of a Design-Oriented Stress-Strain Model for FRP-Confined Concrete. *J. Compos. Constr.* 13, 269–278. doi:10.1061/(asce)cc.1943-5614.0000012

Wang, Y., Zhou, J., Zheng, L., and Gogu, C. (2020). An End-To-End Fault Diagnostics Method Based on Convolutional Neural Network for Rotating Machinery with Multiple Case Studies. *J. Intell. Manufacturing* 33, 1–22. doi:10.1007/s10845-020-01671-1

Wang, Z., Zhao, W., Du, W., Li, N., and Wang, J. (2021). Data-driven Fault Diagnosis Method Based on the Conversion of Erosion Operation Signals into Images and Convolutional Neural Network. *Process Saf. Environ. Prot.* 149, 591–601. doi:10.1016/j.psep.2021.03.016

Willard, J., Jia, X., Xu, S., Steinbach, M., and Kumar, V. (2020). Integrating Physics-Based Modeling with Machine Learning: A Survey. *ACM Comput. Surv.* 1 (1), 1–34. *arXiv preprint arXiv:2003.04919*. doi:10.48550/arXiv.2003.04919

Wilson, E. L. (2004). Static and Dynamic Analysis of Structures, in *A Physical Approach with Emphasis on Earthquake Enginering*. 4th Edn. Berkeley, CA: Computers and Structures Inc..

Xu, C., Guan, J., Bao, M., Lu, J., and Ye, W. (2018). Pattern Recognition Based on Time-Frequency Analysis and Convolutional Neural Networks for Vibrational Events in φ-OTDR. *Opt. Eng.* 57, 016103. doi:10.1117/1.oe.57.1.016103

Yan, S., Zou, X., Ilkhani, M., and Jones, A. (2020). An Efficient Multiscale Surrogate Modelling Framework for Composite Materials Considering Progressive Damage Based on Artificial Neural Networks. *Composites Part B: Eng.* 194, 108014. doi:10.1016/j.compositesb.2020.108014

Yu, Y., Wang, C., Gu, X., and Li, J. (2019). A Novel Deep Learning-Based Method for Damage Identification of Smart Building Structures. *Struct. Health Monit.* 18, 143–163. doi:10.1177/1475921718804132

Zadeh, M., and Saiidi, M. (2007). *Effect of Constant and Variable Strain Rates on Stress-Strain Properties and Yield Propagation in Steel Reinforcing Bars*. Report No. CCEER-07-2 (Reno: Center for Civil Engineering Earthquake Research, University of Nevada).

Zaghi, A. E., Saiidi, M. S., and Mirmiran, A. (2012). Shake Table Response and Analysis of a concrete-filled FRP Tube Bridge Column. *Compos. Structures* 94, 1564–1574. doi:10.1016/j.compstruct.2011.12.018

Zaghi, A. E. (2009). *Seismic Design of Pipe-Pin Connections in concrete Bridges*. Reno: University of Nevada.

Zare, S., and Ayati, M. (2021). Simultaneous Fault Diagnosis of Wind Turbine Using Multichannel Convolutional Neural Networks. *ISA Trans.* 108, 230–239. doi:10.1016/j.isatra.2020.08.021

Zhang, W., Peng, G., and Li, C. (2017). "Bearings Fault Diagnosis Based on Convolutional Neural Networks with 2-D Representation of Vibration Signals as Input," in 2016 the 3rd International Conference on Mechatronics and Mechanical Engineering, Shanghai, China, October 21–23, 2016 (Shanghai, China: EDP Sciences), 13001. doi:10.1051/matecconf/20179513001

Zhang, Y., Xing, K., Bai, R., Sun, D., and Meng, Z. (2020). An Enhanced Convolutional Neural Network for Bearing Fault Diagnosis Based on Time-Frequency Image. *Measurement* 157, 107667. doi:10.1016/j.measurement.2020.107667

Zhu, Z., Ahmad, I., and Mirmiran, A. (2006). Fiber Element Modeling for Seismic Performance of Bridge Columns Made of concrete-filled FRP Tubes. *Eng. structures* 28, 2023–2035. doi:10.1016/j.engstruct.2006.03.031

Zuhair Murad, Y. (2021). Predictive Model for Bidirectional Shear Strength of Reinforced concrete Columns Subjected to Biaxial Cyclic Loading. *Eng. Structures* 244, 112781. doi:10.1016/j.engstruct.2021.112781

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Built Environment

ORIGINAL RESEARCH
published: 20 April 2022
doi: 10.3389/fbuil.2022.860805

Using Machine Learning Models to Forecast Severity Level of Traffic Crashes by R Studio and ArcGIS

Bara' W. Al-Mistarehi[1], Ahmad H. Alomari[2], Rana Imam[3] and Mohammad Mashaqba[3]*

[1]Department of Civil Engineering, Jordan University of Science & Technology (JUST), Irbid, Jordan, [2]Department of Civil Engineering, Yarmouk University (YU), Irbid, Jordan, [3]Department of Civil Engineering, The University of Jordan, Aljubeiha, Jordan

This study describes crash causes, conditions, and distribution of accident hot spots along with an analysis of the risk factors that significantly affect severity levels of crashes and their effects on pedestrian safety using machine learning (ML) techniques. Supervised ML algorithm–random forest and decision tree–based algorithm-AdaBoost algorithms are applied and compared to predict the severity level and future crashes based on road crash elements. Association rule, an unsupervised learning algorithm, is utilized to understand the association between driver characteristics, geometric elements of the highway, the environment, time, weather, and speed. Slight, medium, and severe injuries and fatalities in crashes are also considered to understand the behavior of road drivers, who are most likely to cause crashes. Fatalities and injuries are studied with spatial statistics analysis. The variables most affecting the severity of the crash are determined and discussed in detail. The results are checked for accuracy, sensitivity, specificity, recall, precision, and F1 score performance. The impact of drivers, vehicles, and road characteristics is investigated in traffic crashes. The random forest model was found to be the most suitable algorithm to predict crash severity levels.

Keywords: road safety, crashes, hot spots, machine learning, GIS

OPEN ACCESS

Edited by:
Iftikhar Azim,
Shanghai Jiao Tong University, China

Reviewed by:
Muhammad Javed,
COMSATS University, Pakistan
Muhammad Farjad Iqbal,
Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Pakistan

****Correspondence:***
Bara' W. Al-Mistarehi
bwmistarehi@just.edu.jo

Specialty section:
This article was submitted to Computational Methods in Structural Engineering, a section of the journal Frontiers in Built Environment

Received: *23 January 2022*
Accepted: *14 March 2022*
Published: *20 April 2022*

Citation:
Al-Mistarehi BW, Alomari AH, Imam R and Mashaqba M (2022) Using Machine Learning Models to Forecast Severity Level of Traffic Crashes by R Studio and ArcGIS. Front. Built Environ. 8:860805. doi: 10.3389/fbuil.2022.860805

INTRODUCTION

Road safety and reducing traffic accidents have been a major concern of transportation planners, engineers, and policymakers. Road accidents place a heavy burden on the world economy due to significant loss of life and property damage. The number of traffic accidents in Jordan is rising, and the consequences have a significant impact on people's lives and on their social and emotional well-being (Al-Masaeid, 2009; Alomari and Taamneh, 2020). It is a complex behavioral problem due to the multiplicity of variables and factors affecting it such as driver, pedestrian, passenger behavior, vehicle, highway, and infrastructure characteristics. To develop effective countermeasures, a spatial and temporal analysis should be conducted using geographic information systems (GIS) and artificial intelligence (AI) in countries that have witnessed a decline in the numbers of injuries and deaths to identify and distinguish accidents and then classify them according to the time of their occurrence, type, location, severity, and causes.

Understanding the factors under which people are primarily harmed in car crashes could improve the general safety standards. Variables that influence the level of extended injury to passengers in a vehicle crash incorporate attributes of the individual, natural elements, highway conditions at the time of the incident, and intrinsic qualities of the vehicle itself. In this research, considerable accident

information collection was utilized alongside learning techniques to reveal the connections between the number of accidents compared to various injury severity levels and accident-related factors.

The literature includes many studies carried out on traffic accidents and their seriousness. Labib et al. (2019) stated in their study that traffic accidents had become a significant problem in Bangladesh and the whole world. Their study used the following machine learning methods to analyze accident data: classified decision tree, k-nearest neighbors (KNN), naive Bayes, and AdaBoost. Radzuan et al. (2019) explained that the number of traffic accidents in Malaysia grows by 5% every year. It is essential to understand their occurrence to propose counter plans to reduce the number of severe injuries. The research used an artificial neural network method to predict the number of serious severe injuries by applying a time-series algorithm. The study by Al-Mamlook et al. (2019) compared different machine learning algorithm performance on predicting crash severity and found that the random forest algorithm outperformed logistic regression, naive Bayes, and AdaBoost algorithms.

Traffic accident data provided by the public authorities are essential in building a classification of these accidents based on their type and severity, which could later lead to developing a predictive model. These models help address and detect the severity of the injury in accidents and guide general road safety policies (Cuenca et al., 2018). In a study by Kumeda et al. (2019), machine learning algorithms were used to analyze the data to unlock invisible patterns and predict severity level. Their study applied six different machine learning techniques using the data from the United Kingdom: hierarchical LVQ, Fuzzy–FARCHD, random forest, multilayer perceptron, naive Bayes, and radial basis function network. The results revealed that the Fuzzy–FARCHD algorithm was the most effective method, with an accuracy of 85.94%. In another study by Cai et al. (2020), the locations of traffic accidents were analyzed and their relationships with the resulting behaviors based on vehicle OBD data. This research relied on the RTSE method implemented by the entropy weight method. Moreover, the method includes three factors: piecewise calculation of each index's weight, optimization of the logarithm base, and processing of zero-value secondary indices. In addition, Hu et al. (2020) investigated clusters and injuries in pedestrian crashes using GIS in Changsha, China. This study did not just include standard traffic accident reports but also integrated these reports with the GIS to analyze 791 pedestrian accidents. The study concluded that casualties have a strong relationship with road isolation facilities, light conditions, time of the accident, and age of pedestrians. For example, casualty numbers were higher at night than during the day, and pedestrians of school age and the elderly made up the most significant share of the casualties. Moreover, Al-Omari et al. (2020) studied the prediction of traffic accident hot spots using fuzzy logic and the GIS in Irbid city, north of Jordan. The study applied the fuzzy overlay method (FOM) and the weighted overlay method (WOM) since these two methods are generally used in decision-making and alternative analysis. Eight hotspots were identified; three are road sections and five are major intersections, which were investigated to obtain accident-contributing factors.

Lee et al. (2017) built an accident prediction model using two ML methods: artificial neural networks (ANN) and k-nearest neighbor. Due to the complexity and variety of factors involved in traffic incidents, various data such as traffic data, weather conditions, light, and road geometry were included. The accident characteristics are important as well. Based on the comparison analysis, they concluded that the ANN outperformed the k-nearest neighbor and provided a helpful prediction with less than 30% error values. Achu et al. (2019) used geospatial technology to examine traffic accident temporal and spatial behaviors based on the collected data between 2013 and 2015. The study used multiple methods (kernel density functions, Moran's-I, and Getis-Ord Gi hotspot analysis) for spatiotemporal behaviors of traffic accidents. Moreover, this research revealed that the results are helpful in terms of developing better safety policies for roads and highways in the accident hot spots. Al-Aamri et al. (2020) conducted their study in Muscat, Oman, based on 9,357 registered traffic accidents from 2010 to 2014 collected by the Royal Oman Police. The study used a combined analysis method based on the GIS and RTC data. In addition, this study used estimation techniques using kernel density estimation (KDE) for 1D and 2D space dimensions; network-based, forest algorithm, and K-function were used in the data analysis. In addition, the results of this study provided statistical evidence that proved the hypothesis of the study, which was that road junctions have a higher rate of causing RTCs than the rest of the road features. Moreover, the findings of the study identified the importance and influence of road- and traffic-related features in road crash spatial analysis.

Ando et al. (2018) investigated the links between traffic crashes and urban crime in Toyota City, Japan. They defined urban safety in terms of urban crime and traffic accidents using the GIS and statistical analysis. They found that there were some overlap and common factors, including narrow streets and medium-to-high rise residential buildings. Longer street segments tended to discourage both phenomena. A graphical modeling method for individual driving behavior and its application in driving safety analysis using GPS data was proposed by Chen et al. (2019). The study found that the different aspects of driving safety concerns originate from the skills and behaviors of drivers. The study applied a quantitative analysis method to evaluate the graphs extracted from the method mentioned previously. The study concluded that the graphical method could describe the individual features of a driver's longitudinal acceleration behavior and distinguish differences among drivers. Another study by Özcan and Küçükönder (2020) identified the spatial density of traffic accidents and targeted the modifications of spatiotemporal elements in high-density areas in the city of Kahramanmaras, Turkey. The researchers developed their methodology based on kernel and hotspot estimation (KDE) methods to identify the spatial densities in the GIS, which successfully detect the high density of traffic accidents in the selected areas.

According to the World Health Organization (WHO), about 1.35 million people die annually due to road traffic accidents, making them the eighth leading cause of death for people of all ages. The cost of road traffic accidents in most countries is estimated at 3% of their GDP (WHO, 2018). Road traffic

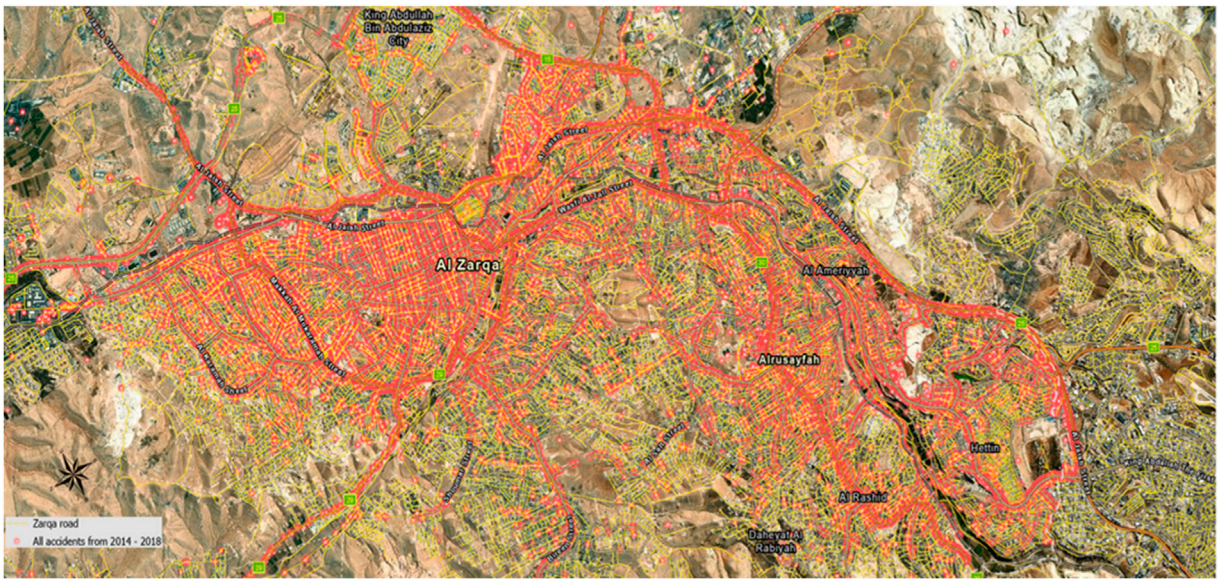


FIGURE 1 | Traffic accidents from 2014 to –2018 occurred in Zarqa city.

injuries remain an important public health problem at global, regional, and national levels; hence, improving traffic safety is a real challenge worldwide. As a result of the political and economic conditions surrounding it, Jordan has witnessed an increase in residents, vehicles, and accidents. In 2019, 161,511 traffic accidents occurred in Jordan, resulting in 17,013 accidents of human injuries; among them were 643 deaths, 792 seriously injured, 6062 medium injuries, and 10,159 slightly injured people; the financial cost is estimated to be $457 million. Jordan's population reached 10,554 million, and the number of registered vehicles was 1,637,981, in addition to 767,614 foreign vehicles entering Jordan annually (PSD, 2018).

The main objective of this research is to develop a framework to identify and traffic accident hot spots in urban regions, with Zarqa City in Jordan as a case study. The methodology is based on combining the GIS, indicators of the severity of traffic accidents, artificial intelligence, and statistical methods. This framework defines crash causes, conditions, crash risk factors, and their impact on severity levels. In addition, it involves analyzing the factors that significantly affect each type of crash and forecast the risk factors affecting the level of traffic crashes using ML techniques. In addition, this study identifies and maps high-density black spots, fatality crashes, and pedestrian crash locations using the Optimized Hot Spot Analysis method. Finally, it recommends appropriate remedial measures to decrease traffic crashes and their severity levels and improve crash reduction factors.

METHODOLOGY

The study area comprised the minor arterials and collectors of Zarqa City, Jordan. According to the number of residents, Zarqa is the center of the Zarqa governorate and the third major city in Jordan after the capital (Amman) and Irbid (Alomari and Taamneh, 2020). Its population constitutes 14.3% of all Jordanians, reaching 1,534,577 inhabitants. It is also the third-highest city in accident frequency. It is located to the northeast of the country and has an area of over 4761 km^2 with a road network of around 440 km^2. The data set for pedestrian, run-off-road, and collision accidents that occurred in urban areas in Zarqa City was obtained from the Jordan Police Traffic Department (PTD). It includes nearly 97,900 accidents in Zarqa City during 2014–2018, as shown in **Figure 1**. The data consist of the numbers of vehicles and category; road properties, type, speed limit, surface, and lane design; driver's gender, age, and mistakes; weather; light conditions; coordinates; accident type, time, day and date; severe, medium, and slight injuries; and fatalities.

Accident locations were defined using the projected coordinate system Universal Transverse Mercator (UTM), World Geodetic System (WGS 1984 Zone 36 North). X and Y coordinates were automatically calculated by the traffic cop using the GPS at each accident location. The time of crashes is also essential for accident analysis and finding the leading cause before giving suitable remedies. The time is required in three configurations: the time of day (morning, noon, and evening), the day of the accident (workday, holiday), and the time of the year (winter, spring, summer, or autumn). Several causes invariably characterize an accident. The alignment of the road is a significant impact factor: flat straight, upright straight, straight slope, flat curve, elevated curve, and slope curve. Other data groups that can represent independent variables and help in accident analysis study are as follows: weather conditions (clear, rainy, dusty, hazy), pavement surface conditions (dry, wet, glacier, sandy), light conditions (noon, sunrise, night, road lights are sufficient; night time and road

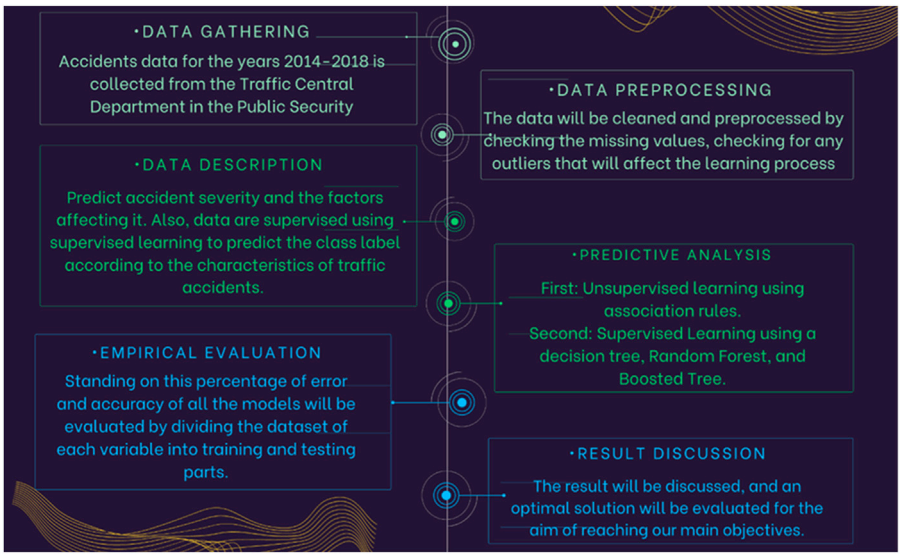


FIGURE 2 | Research methodology.

lights are not enough; sunset; and dark), road type (flexible pavement, rigid pavement, soil road, gravel road), road lane (two-way with central island, two-way without central island, one way, parking, and public square), vehicle category (small passenger cars, shipping vehicles, joint transport, trailer, mid-size passenger cars, bus, construction vehicle, and particular purpose vehicle), road speed limit, driver gender, the number of cars involved in the accident, the type of accident, and driver mistakes).

The data were cleaned and preprocessed by looking for missing values and inspecting any outliers that could interfere with the learning process. Feature selection using ML was applied to find the most relevant and correlated attributes affecting the learning process. These datasets were studied using supervised learning to predict the class label according to the characteristics of drivers and road properties. **Figure 2** shows the research methodology used in this study.

Association Rules

Association rule learning is an unsupervised learning approach that detects the reliance of one database object on another and maps it appropriately to make it more lucrative. It applies various techniques to identify interesting relationships between variables in a database. They are used to find the relations between variables in the dataset.

The association rule can be defined as follows: if we let D = {t1, t2, . . ., tm} be a set of transactions that causes the accidents and let I = {i1, i2, . . ., in} be the set of all considered items in the possible factors, each transaction in D has a unique transaction ID and contains a subset of the items. A rule is defined as $X \Rightarrow Y$ where $X, Y \subseteq I$ and $X \cap Y = \varnothing$. The sets of items (for short item sets) X and Y are called antecedent (left-hand-side or LHS) and consequent (right-hand side or RHS) of the rule. The rules are often restricted to only a single-item inconsequent. The association rules surpass the user-specified minimum support and confidence. The support, supp(X), of an itemset X is a measure of importance, defined as the proportion of transactions in the data set containing the itemset. The confidence of the rule is defined as conf $(X \Rightarrow Y)$ = supp $(X \cup Y)$/supp(X), measuring how likely it is to see Y in a transaction containing X. The association rule $X \Rightarrow Y$ needs to satisfy supp $(X \cup Y) \geq \sigma$ and conf $(X \Rightarrow Y) \geq \delta$, where σ and δ are the minimum support and minimum confidence thresholds, respectively.

This research will use association rules to find the relationship between variables. It will also check which variables affect slight injury, medium injury, severe injury, or fatalities. The association algorithm can be summarized as converting all variables to factors and fitting the model with an *a priori* algorithm. It also specifies which variable has the most significant impact on each type of injury (Galárraga et al., 2013).

Decision Tree

It is a supervised learning approach that may solve classification and regression issues. However, it is most often used to solve classification tasks. It is a tree-structured classifier in which internal nodes contain dataset characteristics, branches represent decision rules, and each leaf node represents the result. A decision tree has two nodes: the decision node and leaf node. Decision nodes judge and have numerous branches, whereas leaf nodes result from those selections and have no more branches. The judgments or assessments are based on the characteristics of the provided dataset. It is a graphical depiction of all potential answers to a problem/decision, given specific criteria. It is termed a decision tree because, like a tree, it begins with the root node and extends on subsequent branches to form a tree-like structure. The CART algorithm, which stands for classification and regression tree, creates a tree. A decision tree asks a question and divides the tree into subtrees based on the answer (yes/no).

The decision tree coding algorithm may be defined as follows. Two approaches are used: one with the class label (type of integer) and another with a class label (0 and 1), where zero indicates no injury and one indicates injuries. The data set is then divided into training and testing sets using the 70/30 technique and the (rpart) function for estimating the decision tree. The model's effectiveness is then evaluated. The decision tree classification model uses the dataset to create a tree structure. A decision tree is created by breaking down a large dataset into smaller components. A decision tree node is split into two or more branches at each step-in algorithm until it reaches leaf nodes. The leaf nodes show the class labels or outcomes. Di Noia et al. (2020) stated that the decision tree picks a feature that best divides the data at each stage.

Ensemble Method

Bagging and boosting are two techniques that use the ensemble method. Both models can predict the result of the type of injury in a traffic accident. Bagging (random forest) is a well-known ML algorithm from the supervised learning approach. It may be applied to both classification and regression issues in machine learning. It is built on the idea of ensemble learning, which integrates several classifiers to resolve a complicated issue and enhance the model's performance. It is a classifier that uses several decision trees on different subsets of a given dataset and averages them to enhance the prediction accuracy of that dataset. On the other hand, the boosting (AdaBoost) algorithm, short for adaptive boosting, is used in machine learning, adjusted to each instance, with increased weights applied to erroneously categorized instances. Boosting is used in supervised learning to minimize bias and variation. It is based on the progressive development of learners. The boosted methods are applied when there is a large amount of data to produce a forecast. It is used to transform weak optimization algorithms into strong learning algorithms. The boosted algorithm's basic principle combines several weak classifiers to create a strong classification.

This study uses adaptive boosting (AdaBoost). It fits a series of weak classifiers to various weighted training data. It begins by predicting the original data set and then distributing it evenly to each observation. If the first learner's prediction is incorrect, it rewards a greater score to the incorrectly predicted observation. As an iterative process, it continues to add trainees until the variety of models or accuracy is reached. AdaBoost combines multiple inadequate learners to create a more powerful model. The result is the weighted sum of the weak algorithms. Because it is a sequential model, a poor learner gets replaced at each phase in favor of misclassified data points in the preceding classifier. Weighting is used to train the first basic classifier, y1 (x). Every transaction is the same. The weighting will be increased in the following phases of reinforcement. For data, misclassified data point attributes increase and decrease. The points were correctly categorized. The epsilon value is the weighted error rate for each basic classifier. As a result, the weights provide alpha. A higher weight is assigned to a more successful classification (Mayr et al., 2014).

	Actually Positive (1)	Actually Negative (0)
Predicted Positive (1)	True Positives (TPs)	False Positives (FPs)
Predicted Negative (0)	False Negatives (FNs)	True Negatives (TNs)

FIGURE 3 | Confusion matrix (Draelos, 2019).

The Empirical Evaluation

This section explains the empirical evaluation. Because of the enormous data, the data have divided the model into 70% training data and 30% testing data. The following performance measurements used the confusion matrix, accuracy, recall, precision, sensitivity, F1 score, and errors. The confusion matrix is a matrix used to analyze the quality of classification models for a given set of test data. It can be determined only if the real values of the test data are given. The matrix itself is simple to grasp, but the associated terminology can be perplexing because it displays the model performance faults in a matrix. **Figure 3** shows the confusion matrix table, where:

1) True positive: It is the number of correctly classified instances.
2) True negative: It is the number of correctly classified instances.
3) False negative: It is the number of incorrectly classified instances.
4) False positive: It is the number of incorrectly classified instances.

Accuracy is an essential metric in determining the accuracy of classification issues. It specifies how frequently the model predicts the proper outcome. Furthermore, it may be computed as the ratio of the classifier's correct predictions to the total number of predictions produced by the classifiers. The error rate defines how frequently the model makes incorrect predictions. By dividing the number of wrong predictions by the total number of predictions made by the classifier, the error rate can be figured.

Recall is an essential metric for dealing with imbalanced data (where the number of observations per class is not equally distributed). It equals the true positive over the summation of true positive and false negative. The research team does not want to miss recording any injuries. Therefore, it makes the "false negative" as low as possible, so the recall should be high.

Precision isessential to using the essential metric for imbalanced data and very high accuracy. However, it is not a good model as the accuracy is misleading. It equals the true positive over the summation of true positives and false positives. In this research, the term "false positive" (which means we do not have an injury, but we predict having one).

F1 score is the harmonic mean of precision and recall. It takes both false positives and false negatives into account. This formula

can be used on the imbalanced data to decide which one is better in terms of high precision and low recall.

Identification of Hazardous Locations Using Geographic Information Systems

Optimized Hot Spot Analysis uses parameters obtained from input data characteristics to run the Hot Spot Analysis (Getis-Ord Gi*) application. The Optimized Hot Spot Evaluation method interrogates data to obtain the settings that will produce optimum hot spot performance. The application can aggregate incident point data into weighted features if the dataset of input features includes incident point data. The tool can figure out how big an analysis can be based on the distribution of weighted features. Spatial autocorrelation hot spot analysis is one of the spatial statistics toolset of inferential spatial pattern analysis tools. Probability theory underpins inferential analysis. Probability is a measure of variance, and all mathematical analyses (directly or indirectly) are based on probability equations that evaluate the impact of chance on the study results. Nonspatial statistics often include working with a random sample and attempting to calculate the probability that the data collected is a successful approximation (is reflective) of the population. Like the spatial autocorrelation statistics mentioned above, many spatial statistics frequently deal with all accessible data for the research area (such as all crimes, disease cases, and attributes for every census block).

The null hypothesis is the starting point for most statistical experiments. Complete spatial randomness (CSR), regardless of the features directly or the attributes associated with those features, is the null hypothesis for the pattern analysis methods (analyzing patterns framework and mapping clusters toolset). The pattern analysis technique's' z-scores and *p*-values indicate whether to reject the null hypothesis or not. According to the pattern analysis framework, it is likely that a random sampling technique generated the detected spatial pattern. When the *p*-value is very low, it means that the observed spatial trend is very unlikely (low probability) to be the product of random processes, and the null hypothesis can be rejected. Z-scores are standard deviations.

RESULTS AND DISCUSSION

Machine Learning Model Evaluation

The evaluation techniques used in this research are based on the results of the confusion matrix of each ML model. The performance of the model can be evaluated by checking accuracy, recall, precision, F1 score, error, and sensitivity. This research evaluates the rules to find the relations between variables. After applying ML algorithms, **Tables 1–3** describe the results for accuracy for every type of injury. The comparison shows that the random forest model has the highest accuracy rate measurement in each database compared to the other ML models. The slight injury rate was 97.93%, the medium injury rate was 99.78%, the severe injury rate was 99.81%, and the fatality rate was 99.96%. The results are imbalanced, so the researchers will also check the recall, precision, and F1 score to check the injuries.

TABLE 1 | Percentages of parameters in the result of the decision tree.

Parameter	Slight injury	Medium injury	Severe injury	Fatalities
Precision	33.08%	32.97%	32.96%	32.97%
F1 score	49.41%	49.59%	49.57%	49.59%
Accuracy	34.08%	32.97%	32.96%	32.97%
Error	0.5592	0.5703	0.5704	0.5703
Recall	97.60%	100%	100%	100%
Sensitivity	33.09%	33.00%	33.00%	32.97%
Specificity	68.50%	37.50%	0.00%	37.00%

TABLE 2 | Percentages of parameters in the result of random forest.

Parameter	Slight injury	Medium injury	Severe injury	Fatalities
Precision	99.96%	100%	100%	100%
F1 score	98.86%	99.88%	99.90%	99.97%
Accuracy	97.93%	99.78%	99.81%	99.96%
Error	0.0203	0.0012	0.0019	0.0004
Recall	97.80%	99.77%	99.81%	99.95%
Sensitivity	99.96%	99.78%	99.81%	100.00%
Specificity	72.30%	71.90%	83.30%	83.30%

TABLE 3 | Percentages of parameters in the result of AdaBoost.

Parameter	Slight injury	Medium injury	Severe injury	Fatalities
Precision	99.40%	99.60%	99.90%	99.90%
F1 Score	97.40%	98.70%	99.60%	99.70%
Accuracy	95.10%	97.60%	99.30%	99.90%
Error	0.049	0.0234	0.0062	0.0027
Recall	95.60%	97.90%	99.40%	99.70%
Sensitivity	95.60%	97.90%	99.40%	99.00%
Specificity	61.00%	57.10%	40.00%	50.00%

In the recall part, the focus is on the false negative, trying to make it as low as possible to obtain the highest recall. Then, ML models were checked to see the highest recall. The comparison showed that the decision tree model has the highest recall for slight injuries at 99.2%, the random forest has the highest recall for medium injuries at 100%, the decision tree has the highest recall for severe injuries at 100%, and the random forest has the highest recall for fatalities at 100%. The precision matrix has been checked because of the imbalanced data. The focus was on the false positives. The comparison showed that the random forest model has the highest precision rate measurement in each database compared to the other ML models. The slight injury rate was 99.96%, the medium injury rate was 100%, the severe injury rate was 100%, and the fatality rate was 100%.

The precision and recall were checked. The emphasis was on both false-positive and false-negative results. The abovementioned comparison showed that the random forest model had the highest accuracy rate measurement in each database compared to the other ML models. The slight injury rate was 98.86%, the medium injury rate was 99.88%, the severe injury rate was 99.9%, and the fatality rate was 99.97%. Based on the confusion matrix F1-score, the test results showed that random forest seemed to perform better than other models.

TABLE 4 | Slight injury results.

lhs	rhs	Support	Confidence	Coverage	Lift	Count
{}	=> {Slight.Injuries = 1}	0.07455724	0.07455724	1	1	7304
{vehicle. category = 0}	=> {Slight.Injuries = 1}	0.04820089	0.07486682	0.6438218	1.0041522	4722
{Driver.gender = 1}	=> {Slight.Injuries = 1}	0.05152861	0.07702872	0.6689532	1.0331487	5048
{Light = 0}	=> {Slight.Injuries = 1}	0.04818047	0.06900383	0.698229	0.9255148	4720
{Road.properties = 0}	=> {Slight.Injuries = 1}	0.06672791	0.07368789	0.9055479	0.9883398	6537
{Road.surface = 0}	=> {Slight.Injuries = 1}	0.07197468	0.07496757	0.9600776	1.0055036	7051
{The.type.of.accident = 0}	=> {Slight.Injuries = 1}	0.05023223	0.05213918	0.9634257	0.6993175	4921
{weather = 0}	=> {Slight.Injuries = 1}	0.07188282	0.07459351	0.9636605	1.0004864	7042
{Road.type = 0}	=> {Slight.Injuries = 1}	0.07280151	0.07524556	0.967519	1.0092321	7132
{Severe.Injuries = 0}	=> {Slight.Injuries = 1}	0.07088246	0.07174073	0.9880365	0.9622235	6944
{Medium.Injuries = 0}	=> {Slight.Injuries = 1}	0.07353647	0.07393571	0.9946001	0.9916637	7204
{Fatalities = 0}	=> {Slight.Injuries = 1}	0.07329148	0.07360328	0.9957638	0.987205	7180
{Road.surface = 0, vehicle. category = 0}	=> {Slight.Injuries = 1}	0.04672077	0.0753787	0.619813	1.011019	4577
{weather = 0, Vehicle. category = 0}	=> {Slight.Injuries = 1}	0.04678201	0.0752125	0.621997	1.008789	4583
{Road.type = 0, vehicle. category = 0}	=> {Slight.Injuries = 1}	0.0471597	0.0753609	0.625784	1.0107791	4620
{Medium.Injuries = 0, vehicle. category = 0}	=> {Slight.Injuries = 1}	0.04754759	0.0742369	0.640483	0.9957047	4658
{Fatalities = 0, vehicle. category = 0}	=> {Slight.Injuries = 1}	0.04754759	0.0740729	0.641902	0.9935037	4658
{Road.surface = 0, Driver.gender = 1}	=> {Slight.Injuries = 1}	0.04956872	0.0776699	0.6381973	1.0417486	4856
{weather = 0, Driver.gender = 1}	=> {Slight.Injuries = 1}	0.04941561	0.07708599	0.6410453	1.0339169	4841
{Road.type = 0, Driver.gender = 1}	=> {Slight.Injuries = 1}	0.05009953	0.07771971	0.644618	1.0424167	4908
{Severe.Injuries = 0, Driver.gender = 1}	=> {Slight.Injuries = 1}	0.04886439	0.07396706	0.6606237	0.9920842	4787
{Medium.Injuries = 0, Driver.gender = 1}	=> {Slight.Injuries = 1}	0.05102843	0.07656492	0.6664727	1.026928	4999
{Fatalities = 0, Driver.gender = 1}	=> {Slight.Injuries = 1}	0.05068137	0.07609545	0.6660236	1.0206313	4965
{Road.surface = 0, Light = 0}	=> {Slight.Injuries = 1}	0.04694534	0.0693864	0.6765784	0.930646	4599

TABLE 5 | Medium injury results.

lhs	rhs	Support	Confidence	Coverage	lift	Count
{}	=> {Medium.Injuries = 1}	0.0053998	0.0053998	1	1	529
{Road.properties = 0}	=> {Medium.Injuries = 1}	0.0051344	0.0056700	0.90554	1.05002	503
{Road.surface = 0}	=> {Medium.Injuries = 1}	0.0052978	0.0055181	0.96007	1.02189	519
{weather = 0}	=> {Medium.Injuries = 1}	0.0053182	0.005518775	0.9636605	1.0220167	521
{Road.type = 0}	=> {Medium.Injuries = 1}	0.0053692	0.005549518	0.967519	1.02771	526
{Severe.Injuries = 0}	=> {Medium.Injuries = 1}	0.0051957	0.0052586	0.98803	0.97384	509
{Fatalities = 0}	=> {Medium.Injuries = 1}	0.0052773	0.0052998	0.99576	0.98147	517
{Road.surface = 0,Road.properties = 0}	=> {Medium.Injuries = 1}	0.0050324	0.0057607	0.87356	1.06682	493
{weather = 0,Road.properties = 0}	=> {Medium.Injuries = 1}	0.0050528	0.0057719	0.87541	1.06889	495
{Road.type = 0,Road.properties = 0}	=> {Medium.Injuries = 1}	0.0051038	0.0058010	0.87981	1.07429	500
{Fatalities = 0,Road.properties = 0}	=> {Medium.Injuries = 1}	0.0050119	0.0055571	0.90189	1.02913	491
{Road.surface = 0,weather = 0}	=> {Medium.Injuries = 1}	0.0052774	0.0055583	0.94946	1.02933	517
{Road.surface = 0,Road.type = 0}	=> {Medium.Injuries = 1}	0.0052671	0.0056495	0.93232	1.04623	516
{Severe.Injuries = 0,Road.surface = 0}	=> {Medium.Injuries = 1}	0.0050936	0.0053699	0.94855	0.99445	499
{Fatalities = 0,Road.surface = 0}	=> {Medium.Injuries = 1}	0.0051753	0.0054133	0.95603	1.00248	507
{Road.type = 0,weather = 0}	=> {Medium.Injuries = 1}	0.0052978	0.0056539	0.93701	1.04704	519
{Severe.Injuries = 0,weather = 0}	=> {Medium.Injuries = 1}	0.0051242	0.0053807	0.95232	0.99646	502

After applying the ML model, we get the following results for sensitivity and specificity. Sensitivity was used to evaluate the model's ability to predict the true positive of each available class. Specificity was used to evaluate the model's ability to predict the true negatives of each available class. In general, the results reported higher sensitivity and specificity for the random forest model. The lower the error rate measurement is, the better the performance of the ML model will be. On comparing the ML models used in this research, the random forest gives the lowest error rate when predicting every type of injury. The slight injury rate was 0.0203, the medium injury rate was 0.0012, the severe injury rate was 0.0019, and the fatality rate was 0.0004.

The results demonstrate that the best ML technique is the random forest model compared with other models. The random forest has excellent ability to resist the noise by randomly selecting variables and data to generate plenty of classification trees. It can also process discrete data and continuous data.

Evaluating Result for Association Rules

In this part, the research checks the relationship between variables and every type of injury (slight injury, medium

TABLE 6 | Severe injury results.

lhs	rhs	Support	Confidence	Coverage	lift	Count
{}	=> {Severe.Injuries = 1}	0.011963456	0.011963456	1	1	1173
{Driver.gender = 1}	=> {Severe.Injuries = 1}	0.008329505	0.012451552	0.6689532	0.9276155	1027
{Road.properties = 0}	=> {Severe.Injuries = 1}	0.010473128	0.011565515	0.9055479	0.9202381	1130
{Slight.Injuries = 0}	=> {Severe.Injuries = 1}	0.008288675	0.008956442	0.9254428	0.6622587	1111
{Road.surface = 0}	=> {Severe.Injuries = 1}	0.011524524	0.012003743	0.9600776	0.9259468	1127
{weather = 0}	=> {Severe.Injuries = 1}	0.011330577	0.011757852	0.9636605	0.7515682	1153
{Road.type = 0}	=> {Severe.Injuries = 1}	0.011493901	0.011879767	0.967519	0.740365	1047
{Medium.Injuries = 0}	=> {Severe.Injuries = 1}	0.011759302	0.011823145	0.9946001	0.7271596	998
{Fatalities = 0}	=> {Severe.Injuries = 1}	0.010677283	0.010722706	0.9957638	0.7351375	978
{Road.surface = 0,Driver.gender = 1}	=> {Severe.Injuries = 1}	0.007962027	0.012475808	0.6381973	0.9669399	985
{Road.type = 0,Driver.gender = 1}	=> {Severe.Injuries = 1}	0.007962027	0.012351544	0.644618	0.9717923	1007
{Medium.Injuries = 0,Driver.gender = 1}	=> {Severe.Injuries = 1}	0.008247844	0.012375366	0.6664727	0.7087523	1104
{Road.surface = 0,Road.properties = 0}	=> {Severe.Injuries = 1}	0.010177104	0.011650054	0.8735671	1.0084241	1091
{weather = 0,Road.properties = 0}	=> {Severe.Injuries = 1}	0.00997295	0.011392257	0.8754147	0.953732	1110
{Road.type = 0,Road.properties = 0}	=> {Severe.Injuries = 1}	0.010044404	0.011416505	0.8798142	1.0732998	1013
{Medium.Injuries = 0,Road.properties = 0}	=> {Severe.Injuries = 1}	0.010268974	0.011404732	0.9004134	1.0891904	1073
{Fatalities = 0,Road.properties = 0}	=> {Severe.Injuries = 1}	0.009329863	0.01034475	0.9018935	1.0915744	1092
{Slight.Injuries = 0,Road.surface = 0}	=> {Severe.Injuries = 1}	0.008053897	0.009068653	0.8881029	1.0514448	998
{Slight.Injuries = 0,weather = 0}	=> {Severe.Injuries = 1}	0.007890573	0.008848139	0.8917777	1.003407	974

TABLE 7 | Fatality results.

lhs	rhs	Support	Confidence	Coverage	lift	Count
{}	=> {Fatalities = 1}	0.004236207	0.004236207	1	1	415
{Driver.gender = 1}	=> {Fatalities = 1}	0.002929618	0.004379406	0.6689532	1.0338037	287
{Light = 0}	=> {Fatalities = 1}	0.002980656	0.004268881	0.698229	1.0077131	292
{Road.properties = 0}	=> {Fatalities = 1}	0.003654366	0.004035531	0.9055479	0.9526284	358
{Slight.Injuries = 0}	=> {Fatalities = 1}	0.002970449	0.003209759	0.9254428	0.7576966	291
{Road.surface = 0}	=> {Fatalities = 1}	0.00404226	0.004210347	0.9600776	0.9938956	396
{The.type.of.accident = 0}	=> {Fatalities = 1}	0.002317154	0.00240512	0.9634257	0.5677531	227
{weather = 0}	=> {Fatalities = 1}	0.003970806	0.004120544	0.9636605	0.9726967	389
{Road.type = 0}	=> {Fatalities = 1}	0.003991221	0.004125212	0.967519	0.9737986	391
{Severe.Injuries = 0}	=> {Fatalities = 1}	0.002950033	0.002985753	0.9880365	0.7048176	289
{Medium.Injuries = 0}	=> {Fatalities = 1}	0.004113714	0.004136048	0.9946001	0.9763566	403
{Road.properties = 0,Driver.gender = 1}	=> {Fatalities = 1}	0.002541724	0.004215837	0.602899	0.9951916	249
{Road.surface = 0,Driver.gender = 1}	=> {Fatalities = 1}	0.002745879	0.004302554	0.6381973	1.015662	269
{weather = 0,Driver.gender = 1}	=> {Fatalities = 1}	0.002684632	0.004187898	0.6410453	0.9885962	263
{Road.type = 0,Driver.gender = 1}	=> {Fatalities = 1}	0.002705048	0.004196358	0.644618	0.9905933	265
{Medium.Injuries = 0,Driver.gender = 1}	=> {Fatalities = 1}	0.002847956	0.004273177	0.6664727	1.0087272	279
{Light = 0,Road.properties = 0}	=> {Fatalities = 1}	0.002613178	0.004107962	0.6361251	0.9697266	256
{Road.surface = 0, Light = 0}	=> {Fatalities = 1}	0.002868371	0.004239526	0.6765784	1.0007834	281

injury, severe injury, and fatalities). It also specifies which variable has the most significant influence on each type of injury. After applying the inspect code, the results found several rules that strongly influence the slight injury; according to the following rules: road-properties = flat–straight; road-surface = dry; weather = clear; road-type = flexible-pavement; and crash-types = collision, as shown in **Table 4**.

After applying the inspect code, the results found several rules that strongly influence the medium injury; according to what they get, the road-properties=flat–straight; road-type=flexible-pavement; crash-types collision; and weather = clear, as shown in **Table 5**.

After applying the inspect code, the results found several rules that strongly influence the severe injury; according to what they get, the road properties = flat-straight; road surface = dry; weather = clear; road type = flexible-pavement; and road surface = dry and clear weather, as shown in **Table 6**.

After applying the inspect code, the results found several rules that strongly influence the occurrence of fatalities; according to what they get: the crash-types = collision; driver-mistakes = tailgating; driver-age = [25,32]; road-lanes = two-way (with central island); vehicle-category = small passenger-cars, as shown in **Table 7**.

Generally, collision accidents on flat–straight roads and flexible pavement roads were more effective variables in the data set for every type of injury, with assigned support and confidence. According to the results, which confirm the support and confidence, it is confirmed that the road

TABLE 8 | Effect of driver characteristic on traffic accidents.

lhs	rhs	Support	Confidence	Coverage	lift	Count
	=> {A11 = 1}	0.050467464	0.05046746	1	1	1085
Driver_gender = mALE1	=> {A11 = 1}	0.027954789	0.04968995	0.56258431	0.9845939	601
{Driver.Mistakes = other}	=> {A11 = 1}	0.02367552	0.05030638	0.47062654	0.9968083	509
{speed--(40.60]}	=> {A11 = 1}	0.023489465	0.05123782	0.45843993	1.0152645	505
{Driver._gender = FEMALE}	=> {A11 = 1}	0.022512675	0.05146746	0.43741569	1.0198147	484
{Speed=(20.40]}	=> {A11 = 1}	0.022001023	0.05174489	0.42518257	1.0253118	473
{Driver._age=(18.36]}	=> {A11 = 1}	0.015442579	0.05144894	0.3001535	1.0194477	332
Driver_gender-mALE,driver.mistakes-otherl	=> {A11 = 1}	0.013488999	0.05064618	0.26633797	1.0035411	290
iSpeed=(20.40],Driver._gender = mALEl	=> {A11 = 1}	0.012977348	0.05366417	0.2418252	1.0633419	279
iSpeed-(40.60],Drivers_gender-MALEl	=> {A11 = 1}	0.012419182	0.04860732	0.25550026	0.9631417	267
ispeed--(40.60],driver._gender = FEmALE1	=> {A11 = 1}	0.011070282	0.05454962	0.20293967	1.0808869	238
{Speed-(40.60],Driver.mistakes-Other}	=> {A11 = 1}	0.010884227	0.04929429	0.22080097	0.9767539	234
driver._age=(18.36],driver._gender = mALEl	=> {A11 = 1}	0.010744686	0.05261959	0.20419554	1.0426438	231
iSpeed=(20,401, Driver.mistakes = 0therl	=> {A11 = 1}	0.010605144	0.05358402	0.19791618	1.0617538	228
fDriver._gender-FEMALE, Driver.Mistakes-Otherl	=> {A11 = 1}	0.01018652	0.04986339	0.20428857	0.9880304	219
fspeed=(20.40],driver._gender = FEmALE1	=> {A11 = 1}	0.009023676	0.0492136	0.18335737	0.975155	194
iSpeed=(20.40],Driver._age-(18,361}	=> {A11 = 1}	0.007628262	0.05608755	0.13600633	1.1113606	164
fdriver's_age=(18.36],driver.mistakes = otherl	=> {A11 = 1}	0.007209638	0.05122274	0.14075073	1.0149655	155
{Driver._age=(36.54]}	=> {A11 = 1}	0.00711661	0.04750078	0.14982092	0.9412158	153

TABLE 9 | Descriptive statistics of 5-year accidents in Zarqa City.

Year	Overall no. of accidents	Collision	Pedestrian	Run-off-road	No. of fatalities	No. of injuries		
						Severe	Medium	Slight
2014	14745	14024	563	158	91	334	456	1997
2015	18663	17895	548	220	112	267	549	2493
2016	20969	20181	547	241	98	370	745	2634
2017	22089	21441	486	162	58	154	813	1597
2018	21499	20841	510	148	56	118	511	1156
Sum	97965	94382	2654	929	415	1243	3074	9877

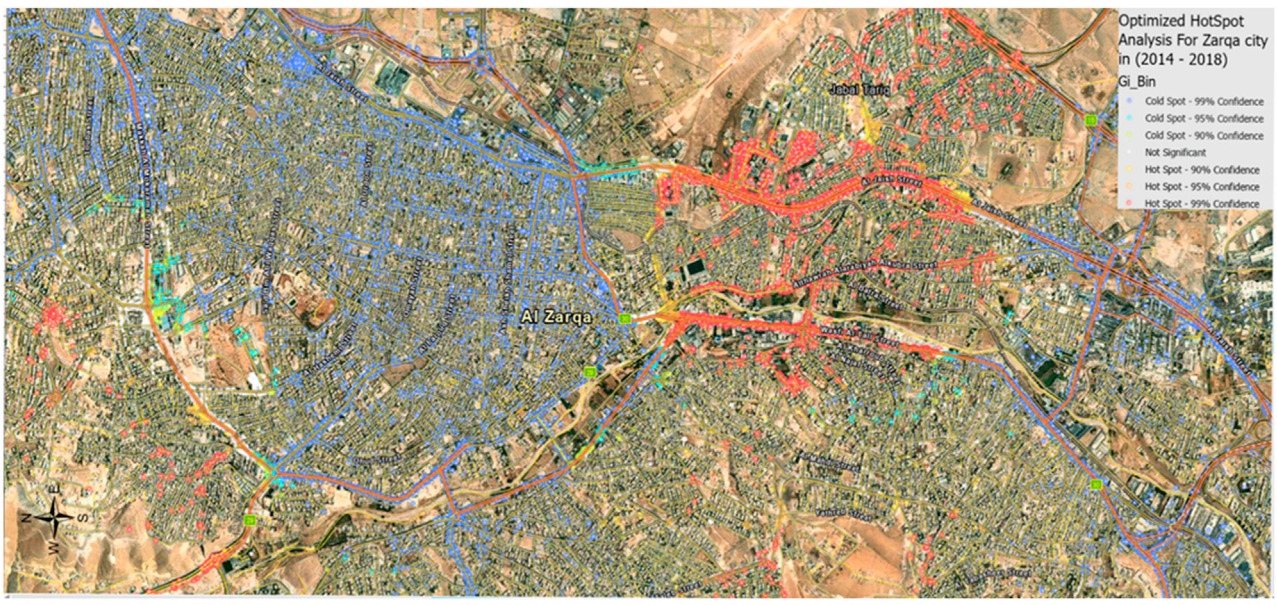


FIGURE 4 | Optimized hot spot analysis for Zarqa City in 2014–2018.

FIGURE 5 | Detailed sample of Optimized Hot spot Analysis on King Al Hussein Bin Talal Street.

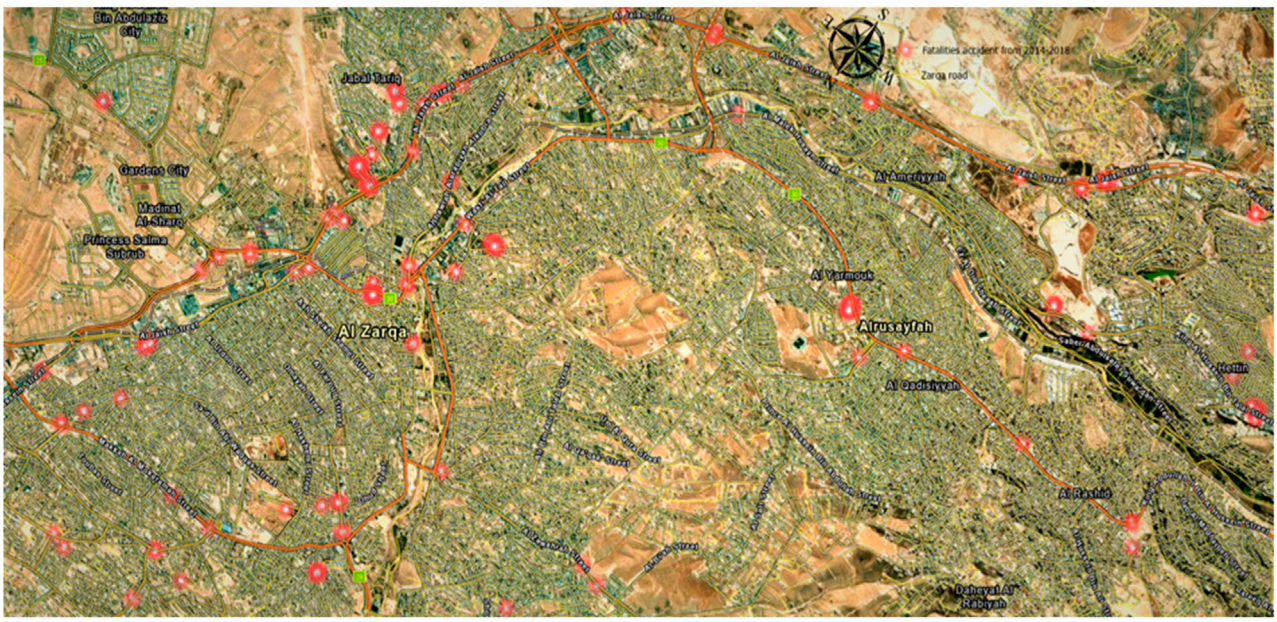


FIGURE 6 | Fatality crashes for Zarqa City in 2014—2018.

characteristics have more effect than the driver characteristics. However, the impact of driver characteristics on traffic accidents was investigated, as shown in **Table 8**. The data are split into two classes: 1) making an accident without injury and 2) making an accident with injury. Then, the researchers will study the association between driver characteristics and traffic accident variables.

Hazardous Locations Based on Optimized Hot Spot Analysis

As explained in the *Identification of Hazardous Locations Using Geographic Information Systems* section, the optimized hot spot analysis application is run using parameters derived from the characteristics of the input data. The optimized hot spot evaluation method

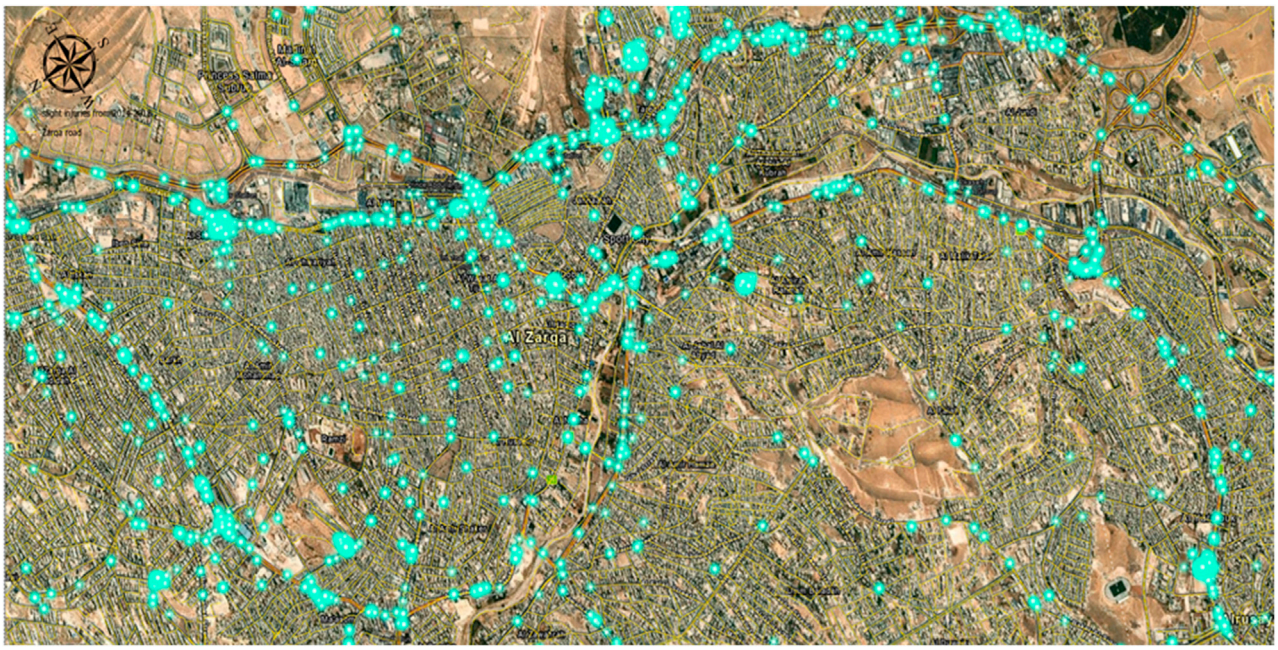


FIGURE 7 | Slight injuries crashes for Zarqa City in 2014—2018.

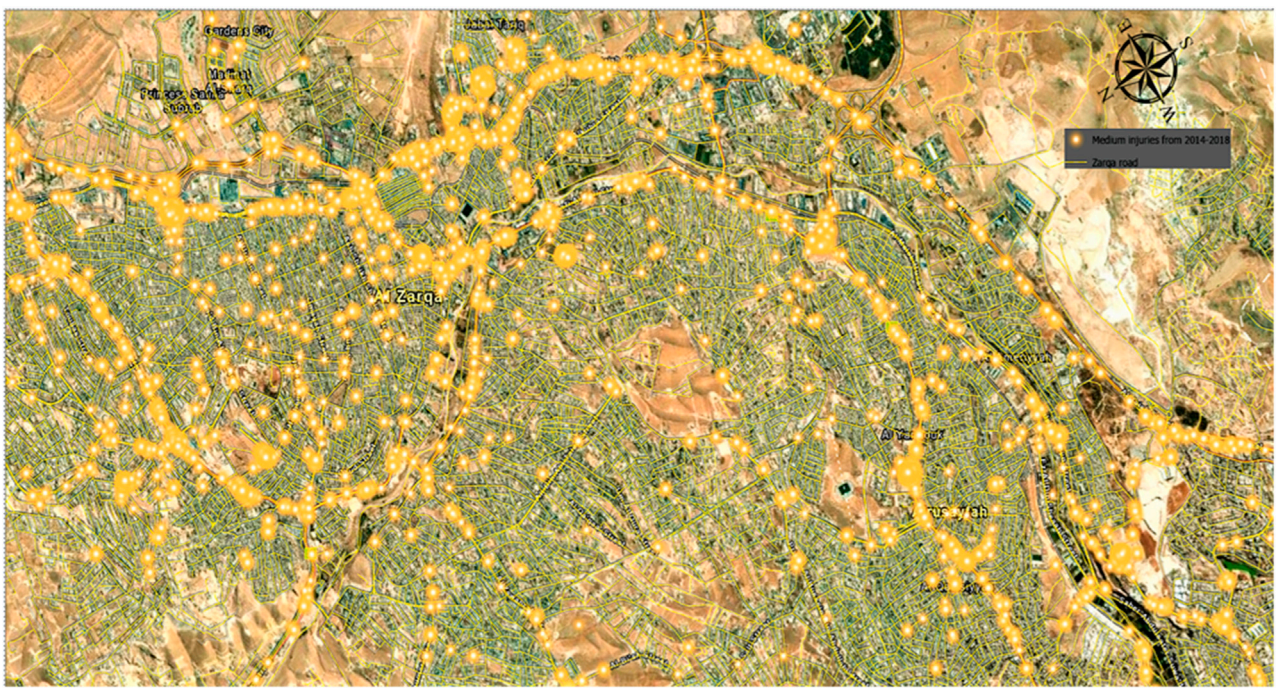


FIGURE 8 | Medium injury crashes for Zarqa City in 2014–2018.

interrogates data to obtain the settings that will produce optimum hot spot performance. The tool can determine an effective scale of analysis based on the distribution of weighted features. Accident locations were defined according to the projected coordinate system, with X and Y coordinating automatically by the traffic policeman using the GPS at

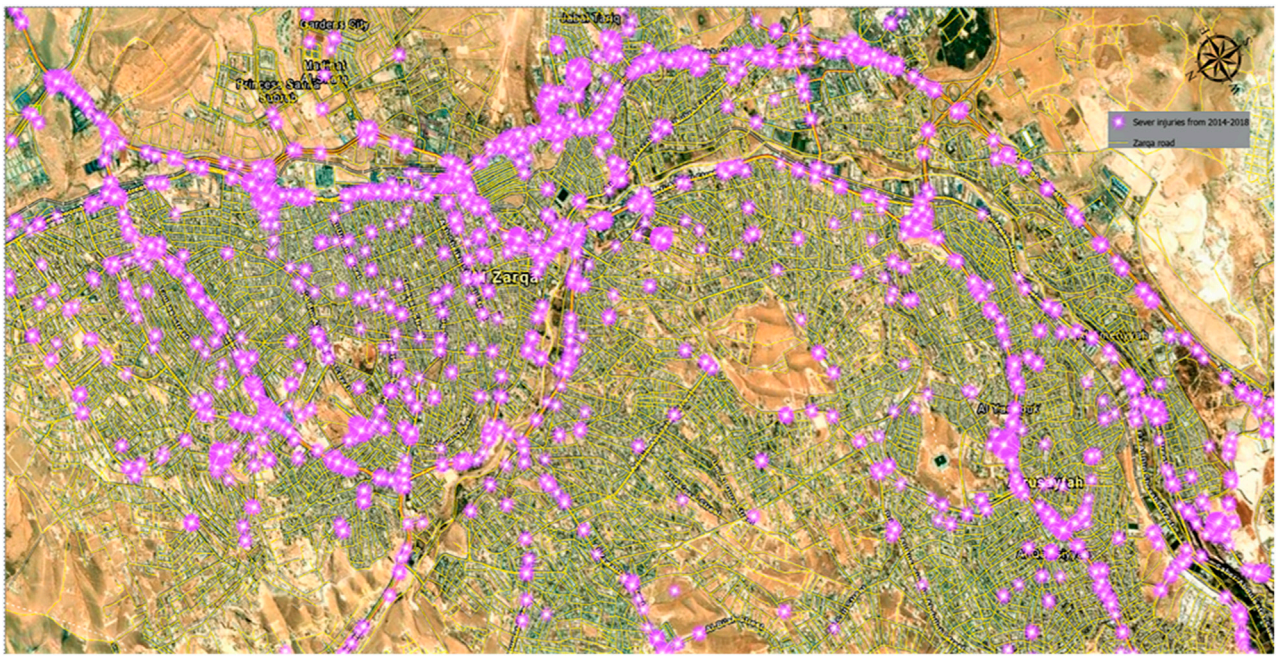


FIGURE 9 | Severe injury crashes for Zarqa City in 2014–2018.

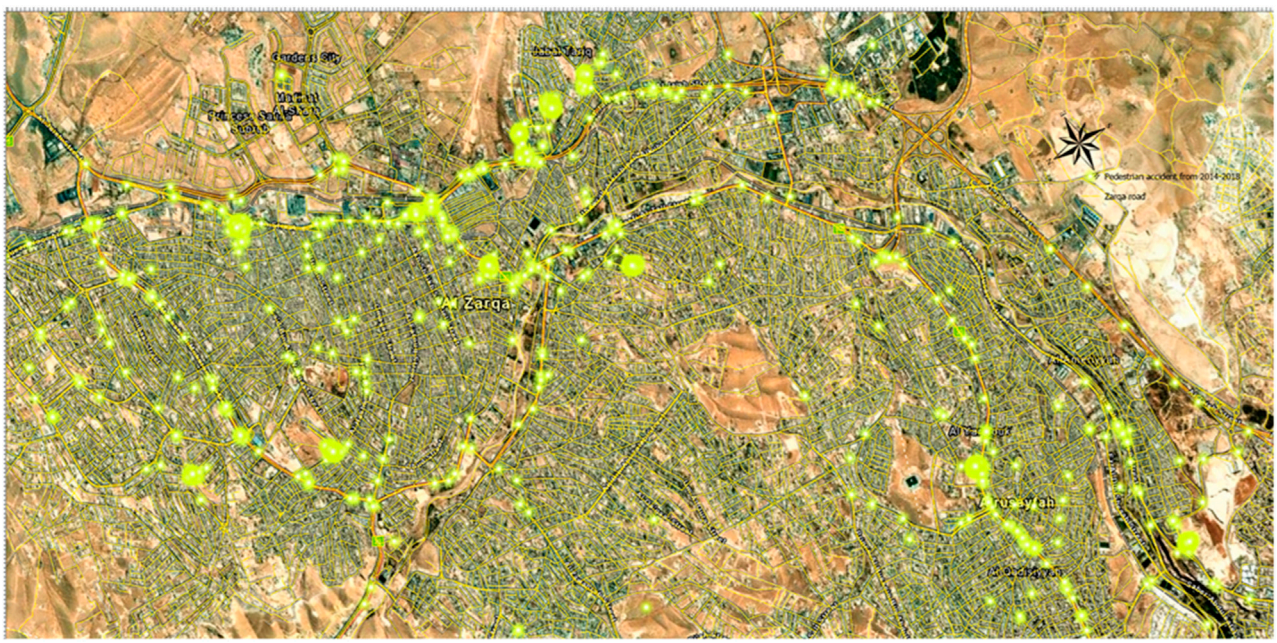


FIGURE 10 | Pedestrian crashes for Zarqa City in 2014–2018.

each crash location. **Table 9** shows the descriptive statistics of 5 year accidents in Zarqa city.

In 5 years, there were 9,877 slight injuries, 3,074 medium injuries, and 1,243 severe injuries accidents in Zarqa City. Around 2,654 (2.7%) of them were pedestrian accidents, 929 (0.9%) were run-off-road accidents, and 94,382 (96.4%) were collision accidents. Since the study area is an urban area, pedestrian accidents were expected due to the high pedestrian movements in such areas. **Figure 4** shows the optimized hot spot analysis on each street from 2014 to 2018. Hot spots were distributed on the

following streets: Al Jaish Street, Wasfi Al-Tall Street, Makkah Al-Mokaramah Street, King Al Hussein Bin Talal Street, King Abdullah II Bin al Hussein Street, the intersection of Bab Al Wad Street with King Talal Street, the intersection of Al Jaish Street with Al Shurtah Street, Abdullah Ghosheh Street, and King Talal Street. **Figure 5** shows a sample of a detailed optimized hot spot analysis on King Al Hussein Bin Talal Street.

Fatality accidents have a lower frequency than injury accidents in general. The study area recorded 415 fatality accidents over 5 years. **Figure 6** shows that fatality accidents are mainly distributed on the main streets with high traffic volumes and the outer border arterial roads of the city (outside the CBD). Fatalities were distributed on the following streets: Al Jaish Street, Wasfi Al-Tall Street, Makkah Al- Mokaramah Street, King Al Hussein Bin Talal Street, King Abdullah II Bin al Hussein Street, the intersection of Bab Al Wad Street with King Talal Street, and the intersection of Al Jaish Street with Al Shurtah Street. Traffic accidents resulting in injuries in the study area were more frequent than fatality accidents. As shown in **Figures 7–9**, there were slight, medium, and severe injury accidents in Zarqa City from 2014 to 2018. The most noticeable was the nature of locations, where most of these accidents occurred on high-speed corridors and pedestrian high-density areas.

The study area recorded 2,654 pedestrian crashes over 5 years. There are many reasons for them such as excessive speed, inadequate light at nighttime, slippery pavement, high traffic volume, inadequate signing, dual parking, inadequate lane width, inadequate maintenance, a high percentage of heavy vehicles, and poor signing. **Figure 10** shows the pedestrian crashes in each street from 2014 to 2018. Pedestrian crashes were distributed on the following streets: King Al Hussein Bin Talal Street, King Abdullah II Bin Al Hussein Street, King Talal Street, Al Hashemi Street, and the intersection of Al Jaish Street with Abdullah Ghosheh Street.

Considering the studied crashes and their causes, several factors can be assessed to enhance traffic safety and reduce accidents in the studied urban areas, such as offering parking in the urban area and near the commercial centers, improving street light, prohibiting on-street parking on one side of the narrow-old roadway due to lack of typical street widths, and finally, introducing a public transport system along with high standards and specifications, including trip schedules and clear trip routes.

CONCLUSION

This research aims to identify the causes and conditions of crashes; distribution of hot spots; analyze the factors that significantly affect each crash severity level; forecast the risk factors that affect these levels; and the effect of all these factors on pedestrian safety using ML techniques. The results showed that the random forest model was the most suitable method to predict slight, medium, and severe injuries, considering the factors specific to the highway, vehicle, and environment. There were significant factors that caused different types of injuries and fatalities, including the type of crash (collision), road properties (flat straight), road type (flexible pavement), road surface (dry), road lanes (two ways with median), weather (clear), vehicle category (small passenger car), driver mistake (not taking necessary safety precautions while driving), light conditions (night with sufficient road light), day of the week (Thursday), range of speed limit (20 km/h to 60 km/h), and driver age (18–36 years). However, the time of the day was the most significant variable in traffic crashes.

Future work based on this research could involve analyzing data from all regions in Jordan to investigate crashes in other parts of the country. In addition, it is recommended for future studies to better understand the relationship between accident occurrence (hot spots), enforcement, and land use in urban areas such as residential, commercial, educational, industrial, etc. Furthermore, field visits to hazardous locations (hot spots) are recommended to understand the causes of crashes specific to each location to help in formulating strategic plans to reduce or prevent future crashes at these sites.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Materials, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

REFERENCES

Achu, A. L., Aju, C. D., Suresh, V., Manoharan, T. P., and Reghunath, R. (2019). Spatio-temporal Analysis of Road Accident Incidents and Delineation of Hotspots Using Geospatial Tools in Thrissur District, Kerala, India. *KN J. Cartogr. Geogr. Inf.* 69 (4), 255–265. doi:10.1007/s42489-019-00031-1

Al Mamlook, R. E., Ali, A., Hasan, R. A., and Kazim, H. A. M. (2019). "Machine Learning to Predict the Freeway Traffic Accidents-Based Driving Simulation," in 2019 IEEE National Aerospace and Electronics Conference (NAECON) (Dayton, OH: IEEE), 630–634. doi:10.1109/naecon46414.2019.9058268

Al-Aamri, A. K., Hornby, G., Zhang, L. C., Al-Maniri, A. A., and &Padmadas, S. S. (2020). Mapping Road Traffic Crash Hotspots Using GIS-Based Methods: A Case Study of Muscat Governorate in the Sultanate of Oman. *Spat. Stat.* 42, 100458. doi:10.1016/j.spasta.2020.100458

Al-Masaeid, H. (2009). Traffic Accidents in Jordan. *Jordan J. Civil Eng.* 3 (4), 331–343.

Al-Omari, A., Shatnawi, N., Khedaywi, T., and Miqdady, T. (2020). Prediction of Traffic Accidents Hot Spots Using Fuzzy Logic and GIS. *Appl. Geomat* 12 (2), 149–161. doi:10.1007/s12518-019-00290-7

Alomari, A. H., and Taamneh, M. M. (2020). Front-seat Seatbelt Compliance in Jordan: an Observational Study. *Adv. transportation Stud.* 2020 (52), 101–116. doi:10.4399/97888255370317

Ando, R., Higuchi, K., and Mimura, Y. (2018). Data Analysis on Traffic Accident and Urban Crime: A Case Study in Toyota City. *Int. J. Transportation Sci. Techn.* 7 (2), 103–113. doi:10.1016/j.ijtst.2018.01.002

Cai, X., Lei, C., Peng, B., Tang, X., and Gao, Z. (2020). Road Traffic Safety Risk Estimation Method Based on Vehicle Onboard Diagnostic Data. *J. Adv. transportation* 2020, 13. doi:10.1155/2020/3024101

Chen, C., Zhao, X., Zhang, Y., Rong, J., and Liu, X. (2019). A Graphical Modeling Method for Individual Driving Behavior and its Application in Driving Safety Analysis Using GPS Data. *Transportation Res. F: traffic Psychol. Behav.* 63, 118–134. doi:10.1016/j.trf.2019.03.017

Cuenca, L. G., Puertas, E., Aliane, N., and Andres, J. F. (2018). "Traffic Accidents Classification and Injury Severity Prediction," in 2018 3rd IEEE International Conference on Intelligent Transportation Engineering (ICITE) (Singapore: IEEE), 52–57. doi:10.1109/icite.2018.8492545

Di Noia, A., Martino, A., Montanari, P., and Rizzi, A. (2020). Supervised Machine Learning Techniques and Genetic Optimization for Occupational Diseases Risk Prediction. *Soft Comput.* 24 (6), 4393–4406. doi:10.1007/s00500-019-04200-2

Draelos, R. (2019). Measuring Performance: The Confusion Matrix. Available at: https://towardsdatascience.com/measuring-performance-the-confusion-matrix-25c17b78e516 Accessed January 2022.

Galárraga, L. A., Teflioudi, C., Hose, K., and Suchanek, F. (2013). "AMIE: Association Rule Mining under Incomplete Evidence in Ontological Knowledge Bases," in Proceedings of the 22nd international conference on World Wide Web, Rio de Janeiro, Brazil (New York, NY: Association for Computing Machinery), 413–422.

Hu, L., Wu, X., Huang, J., Peng, Y., and Liu, W. (2020). Investigation of Clusters and Injuries in Pedestrian Crashes Using GIS in Changsha, China. *Saf. Sci.* 127, 104710. doi:10.1016/j.ssci.2020.104710

Kumeda, B., Zhang, F., Zhou, F., Hussain, S., Almasri, A., and Assefa, M. (2019). "Classification of Road Traffic Accident Data Using Machine Learning Algorithms," in 2019 IEEE 11th International Conference on Communication Software and Networks (ICCSN) (Chongqing, China: IEEE), 682–687. doi:10.1109/iccsn.2019.8905362

Labib, M. F., Rifat, A. S., Hossain, M. M., Das, A. K., and &Nawrine, F. (2019). "Road Accident Analysis and Prediction of Accident Severity by Using Machine Learning in Bangladesh," in 2019 7th International Conference on Smart Computing & Communications (ICSCC) (Sarawak, Malaysia: IEEE), 1–5. doi:10.1109/icscc.2019.8843640

Lee, Y., Wei, C. H., and Chao, K. C. (2017). Non-parametric Machine Learning Methods for Evaluating the Effects of Traffic Accident Duration on Freeways. *Arch. Transport* 43, 91–104. doi:10.5604/01.3001.0010.4228

Mayr, A., Binder, H., Gefeller, O., and Schmid, M. (2014). The Evolution of Boosting Algorithms. From Machine Learning to Statistical Modelling. *Methods Inf. Med.* 53 (06), 419–427. doi:10.3414/ME13-01-0122

Özcan, M., and Küçükönder, M. (2020). Investigation of Spatiotemporal Changes in the Incidence of Traffic Accidents in Kahramanmaraş, Turkey, Using GIS-Based Density Analysis. *J. Indian Soc. Remote Sensing* 48 (7), 1045–1056. doi:10.1007/s12524-020-01137-0

PSD, Public Security Directorate of Jordan. (2018). Jordan Traffic Institute. Available at: https://www.psd.gov.jo/images/traffic/traffic2017.pdf Annual Reports (Accessed 29 November 2020).

Radzuan, N. Q., Hassan, M. H. A., Abdul Majeed, A. P. P., Musa, R. M., Mohd Razman, M. A., and Abu Kassim, K. A. (2019). "Predicting Serious Injuries Due to Road Traffic Accidents in Malaysia by Means of Artificial Neural Network," in *Symposium on Intelligent Manufacturing and Mechatronics* (Singapore: Springer), 75–80. doi:10.1007/978-981-13-9539-0_8

WHO, World Health Organization. (2018). Global Status Report on Road Safety 2018: Summary.No.WHO/NMH/NVI/18.20. Available at: https://www.who.int/publications/i/item/9789241565684 (Accessed 29 November 2020).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Built Environment

ORIGINAL RESEARCH
published: 15 June 2022
doi: 10.3389/fbuil.2022.917017

Investigating the Factors Affecting Speeding Violations in Jordan Using Phone Camera, Radar, and Machine Learning

Bara' W. Al-Mistarehi[1], Ahmad H. Alomari[2], Rana Imam[3] and Tasneem K. Alnaasan[1]*

[1]Department of Civil Engineering, Jordan University of Science and Technology (JUST), Irbid, Jordan, [2]Department of Civil Engineering, Yarmouk University (YU), Irbid, Jordan, [3]Department of Civil Engineering, The University of Jordan, Amman, Jordan

Traffic accidents are considered the leading cause of death for the age group (5–29). They cause approximately 1.3 million deaths every year. Different factors cause traffic accidents; one of them is speeding violations. This work studied speeding violations and the factors that affect them in Irbid, Jordan. Several factors related to the driver, environment, vehicle, road, and weather conditions were studied. The analysis results showed that age, vehicle type, speed limit, day of the week, season, accident year, accident time, license category, and light condition affect speeding violations. To be more specific, young drivers, morning rush hours, clear visibility during daylight, the winter season, slippery roads (due to rain, snow, or ice at low temperatures), and a speed limit of 40 km/h in residential areas; were all associated with a high percentage of speeding violations. Seven hot spots for speeding violations in Irbid were specified, and the free flow speed was measured using the radar and video at these hot spots. The 85th percentile speed was estimated, and the results for all locations were less than the speed limit. In addition to that, the radar and video measurements were compared to specify the most accurate method. The results did not show any difference between the measurements.

Keywords: speeding violations, speed limit, radar, machine learning, classification and regression tree, traffic accidents

OPEN ACCESS

Edited by:
Yasmin Murad,
The University of Jordan, Jordan

Reviewed by:
Maxim A. Dulebenets,
Florida Agricultural and Mechanical University, United States
Wassim Raphael,
Saint Joseph University, Lebanon
Shadi Hanandeh,
Al-Balqa Applied University, Jordan
Ibrahim Khliefat,
Al-Balqa Applied University, Jordan

****Correspondence:***
Bara' W. Al-Mistarehi
bwmistarehi@just.edu.jo

Specialty section:
This article was submitted to Computational Methods in Structural Engineering, a section of the journal Frontiers in Built Environment

Received: *10 April 2022*
Accepted: *25 May 2022*
Published: *15 June 2022*

Citation:
Al-Mistarehi BW, Alomari AH, Imam R and Alnaasan TK (2022) Investigating the Factors Affecting Speeding Violations in Jordan Using Phone Camera, Radar, and Machine Learning. Front. Built Environ. 8:917017. doi: 10.3389/fbuil.2022.917017

INTRODUCTION

With the increasing demand for commercial shipments, people, and mobility by air, marine, and land, there is an urgent need not just to raise the capacity of the roads connecting the various service providers and infrastructure between the different types of transport modes (Belokurov et al., 2020; Dulebenets, 2020; Elmi et al., 2022). There is also a necessity to discover new policies to manage traffic flow and passenger movements, evolve public transport, and shift the dependency on passenger cars toward a shared mobility future with better use of intelligent transportation systems (Enoch et al., 2020; Albayrak et al., 2020; Rahimi et al., 2021). On the other hand, the increasing traffic volumes (passenger and freight) on roadways result in higher exposure for road users and an increased risk of accidents associated with the driver, environment, vehicle, road, and weather conditions. Traffic accidents are considered the leading cause of death for the age group (5–29). They cause approximately 1.3 million deaths every year (WHO and World Health Organization, 2021). In Jordan, traffic accidents are among the top ten causes of death (CDC, 2014). According to the Annual Report for Traffic Accidents in Jordan for the year 2019, there were

161,511 traffic accidents that resulted in 643 deaths, 792 severe injuries, 6062 moderate injuries, and 10,159 slight injuries. The cost of these accidents was estimated to be 324 million Jordanian Dinars (PSD and Jordanian Public Security Directorate, 2019; Alomari et al., 2021a; Bara'W et al., 2021; Al-Mistarehi et al., 2021; Hazaymeh et al., 2022).

Speeding is one of the causes of traffic accidents. According to a study of 345 Ghanaian drivers, speeding behavior dominated the drivers' behaviors compared with overtaking. Speeding behavior increases as the influence of the motivational factors increases, such as the presence of friends, feeling comfortable while speeding, and previous experience in driving (Atombo et al., 2016). In Riyadh, Saudi Arabia, speeding is the first and main cause of severe injury accidents and Property Damage Only (PDO) accidents (Al-Ghamdi, 2003). Also, high speed is responsible for the high percentage of pedestrian fatalities in Kuwait (Koushki et al., 2001). In Jordan, it was revealed that when the difference between the design speed and the speed limit is greater than 10 km/h, there would be an increase in the speed variance and its safety severity (Alomari et al., 2021b; Alomari et al., 2021c). Obaidat et al. (2008) studied speeding in Jordan as one of the most dangerous driver behaviors. The most important reasons behind this behavior were an emergency, running late for work, traffic jams, road grade, terrain type, social culture, low-speed limit, and wrong overtaking. Mujalli (2018) studied the main factors that lead to traffic accidents. The researcher found that speeding is one of the violations that can lead to traffic accidents with high severe injuries or fatalities. Jadaan et al. (2021) studied driving behavior in Amman, Jordan. Their study revealed that speeding is the most frequent traffic violation. Results showed that 30% of drivers ignore speed limits late at night and early morning. Almost 36% of the respondents reported that speed cameras are the most effective tool for traffic enforcement. Atieh et al. (2020) examined Jordanian speeding behavior in Amman, Jordan. The results showed that speeding is significantly influenced by three factors: being late for work, being alone in the vehicle, and emergencies. Also, it was found that 23.1% of respondents mentioned that speeding is a habit for them. Another study in Jordan conducted by Bener et al. (2019) indicated that excessive speed, mobile phone usage, and driving skills increase accidents. Mobile phones and excessive speed were the most vital factors that cause accidents.

Al-Omari et al. (2013) investigated the influence of vehicle and driver on the vehicle's speed in Jordan. It was found that higher speeds are associated with male and young adult drivers, passenger cars, new vehicles, and seat belt usage. On the other hand, loading vehicles and increasing the number of occupants in a vehicle were associated with lower speeds. Zamanov (2012) studied the factors that increase the percentage of speeding violations. Regression and analysis of variance (ANOVA) were used to analyze the data. It was revealed that the driver's age and gender were the most influential factors. Also, street lighting and median width affected speeding violations. Shawky et al. (2017) studied the effects of different factors on speeding violations to specify the best positions of speed cameras on highway in Abu Dhabi, United Arab Emirates (UAE). A negative binomial analysis was conducted to predict speeding violations. The examined variables were related to roads, traffic information, and camera characteristics. The results indicated that speed limits influenced speeding violations at a 95% confidence level. Also, speeding violations increased during weekends on rural highways because of the road-based trips between different cities. Wu and Hsu (2019) studied speeding violations in northern Taiwan using the association rule algorithm. Different factors, such as gender, age, violation area, and day of the week were investigated. The outcomes indicated that female drivers and those in the age group (18–24) were more likely to violate speed limits. Also, midnight and dawn time had a significant influence on speeding violations. (Inada et al., 2021) used time series to examine collisions caused by speed-related traffic violations in Japan during the COVID-19 lockdown period. They found no proof that the number of fatal accidents caused by speeding went up during the lockdown. According to Bolsunovskaya et al. (2021), 90% of accidents result from drivers violating the speed limit and other travel rules. They designed an information system structure to photograph and videotape traffic speed violations. A performance evaluation conducted at the conclusion demonstrated that developing an information system is cost-effective and will pay for itself in the shortest possible time. The developed complex outperforms existing systems in terms of precision and data reliability.

Javid and Al-Roushdi (2019) examined drivers' speeding behaviors in Oman. A questionnaire was used to study the effect of drivers' attitudes on speeding behavior. Three hundred samples were analyzed using multivariate statistical analysis. The outcomes indicated that young and male drivers were more likely to violate speed. Driving culture, the driver's personality, and situational constraints were associated with drivers' speeding behavior. Cheng et al. (2019) employed a binary logistic regression method to study speeding violations in Wujiang, China. The findings revealed that license plate, rainfall, speeding area, and season all impacted speeding violations. A decision tree was used to specify the factors that influence speeding violations. It was found that heavy rain, autumn, and foreign licenses were associated with a high level of speeding violations. Sutela and Aaltonen (2020) investigated speeding sanction rates in Finland in terms of weather conditions and temporal characteristics. Generalized additive models were used for the analysis. The outcomes revealed that rainfall reduced the rate of speeding sanctions. Lunchtime (11 a.m.–3 p.m.) scored the highest speeding sanction rates. (Ambo et al., 2020) investigated the factors that led to different types of traffic violations, such as speeding and illegal parking, in Luzhou, China. A multinomial logistic regression model was developed. The marginal impact and odd ratio were estimated to determine the effect of independent variables on the dependent one. The outcomes showed that buses and vehicles used for passenger transport were associated with speeding violations. Also, the odd ratios for the morning and rainy weather were more than one. This indicates that these variables had a significant effect on speeding violations.

A very limited number of researchers studied speeding violations, and no previous study used traffic accident data,

cameras, radar, or machine learning to study these violations. Machine learning has been used extensively in various civil engineering fields and applications. In structural engineering, Gene Expression Programming (GEP) was used by Murad Zuhair (2021) to predict the bidirectional shear strength of columns subjected to biaxial cyclic loads based on a database collected from different experimental studies found in the literature. In pavement evaluation, (Imam et al., 2021) adopted GEP for the first time to predict the Pavement Condition Index (PCI) from the International Roughness Index (IRI) using data that was half compiled from the existing literature and the other half was measured and collected in the field by the authors. Using genetic algorithms and artificial neural networks, (Hanandeh, 2022a) developed a mathematical model to estimate the quality index of flexible pavements. According to the results, the genetic algorithm model outperforms the neural network in performance. As for geotechnical engineering, (Tarawneh and Imam, 2014) developed Artificial Neural Network (ANN) models to predict pile setup for three pile types (pipe, concrete, and H-pile) using 169 dynamic load tests obtained from the published literature and the authors' files. An evaluation framework for the unconfined compressive strength (UCS) of soils was developed by Al Bodour et al. (2022) using gene expression programming and deep learning methods. The study found the main soil parameters, gave applicable models for figuring out UCS, found essential links between UCS and the void ratio, and found the clay content breakpoint. Circular Failure of Soil Slopes was studied using Classification and Predictive Gene Expression Programming Schemes by Hanandeh (2022b). Higher unit weight, cohesion, and internal friction angle increased safety, whereas higher slope height, incline, and pore pressure decreased safety. In traffic safety, (Al-Mistarehi et al., 2022) used R Studio and ArcGIS to forecast the severity of traffic collisions using machine learning models: random forest, decision tree, and Boost tree algorithms. The random forest model was the optimal algorithm for forecasting crash severity levels. Kuşkapan et al. (2021) used spatial analysis and machine learning algorithms to investigate speed violations by heavy vehicles on highways. The naive bayes (NB) algorithm outperformed support vector machines (SVM) and k-nearest neighbors (KNN) algorithms in terms of both effectiveness and error scales. Decision Trees are used in this research to investigate speeding violations in Irbid and identify the factors that influence this type of violation.

METHODOLOGY

The data for this work was obtained from the Jordan Traffic Institute (JTI) from 2015to 2019 (PSD and Jordanian Public Security Directorate, 2019). A total of 17,237 records were analyzed by applying two types of decision trees: classification and regression tree (CART) and J48 algorithms, using the WEKA 3.8.4 program (Witten et al., 2016). Humans, vehicles, and roadways are the subjects of traffic safety investigations (Alomari and Taamneh, 2020). It is essential to look at these three elements to determine what causes speeding violations, accidents, and the best solutions for each high-risk area with a high accident rate. This study investigated different variables related to the vehicle, driver, roadway, and environment. The most critical studied characteristics were identified after referring to several previous studies discussed in the literature in the introduction section. **Table 1** shows these factors and their categories.

A decision tree is a machine learning algorithm consisting of branches and nodes that can handle regression and classification problems. It starts with the root node, which splits into child nodes and depends on splitting the data into small subsets to reach a decision. It is a tree-structured classifier, where internal nodes represent the features of a dataset, branches represent the decision rules, and each leaf node represents the outcome. The tree uses splitting criteria to minimize node impurities such as gain ratio and Gini index. The pruning technique is used to increase model precision. There are different types of decision trees, such as classification and regression trees (CART) and J48 algorithms. CART is one of the most common decision tree types. It uses the Gini index as splitting criteria and the cost-complexity model for tree pruning. J48 is the version of the C4.5 algorithms in the WEKA program. It uses the splitting criterion that depends on information gain and the single-pass algorithm for tree pruning (Saravana et al., 2018; Shah et al., 2020; Nathanail et al., 2019).

In order to determine the locations where the speed measurements will be taken in the field, the history of traffic accidents in Irbid city from 2015 to 2019 was scanned and explored for the most frequent areas of traffic accidents that were caused by excessive speeds. Based on the available accident data extracted from the Jordan Traffic Institute (JTI) at the Jordanian Public Security Directorate (PSD and Jordanian Public Security Directorate, 2019), seven streets with the most frequent traffic accidents due to excessive speed were specified: Fawzi Almulki, Prince Hasan, Al-Hashmi, King Hussein, King Abdallah II, Petra, and Irbid-Alhoson Streets. Vehicles' speeds were measured in September 2021 using a radar gun and a phone camera between 9:00 and 11:00 a.m. The measurements were taken in clear weather, daylight, on dry pavement, and away from interruptions. No traffic law enforcement campaigns were conducted during measurements. One hundred speed records were obtained for each site and each method. **Figure 1** shows the hot spot locations for speeding violations.

Radar measurements were taken at points that are safe and have a clear line of sight on the roadside. Speeds for going and on-coming traffic were measured by firing the laser towards the target vehicle and recording the value that appears on the radar's screen. To obtain speed measurements from the videos, the distance between two definite points on the site was measured using a measuring tape; **Table 2** shows these lengths. **Figure 2** shows samples for the studied streets.

The time that a vehicle needs to pass the two definite points was measured by using a stopwatch. By knowing the distance and the time, the speed of each vehicle was calculated by the following equation:

TABLE 1 | Variables description.

Variable	Category	Count	Percent	Variable	Category	Count	Percent
Speeding Violation	Yes	328	1.9	Gender	Female	4,758	27.6
	No	16,909	98.1		Male	12,479	72.4
Traffic Way Direction	One way	1,470	8.5	Road Alignment	Straight	17,021	98.7
	Two way	15,767	91.5		Curve	216	1.3
Holiday	Yes	827	4.8	Road Type	Asphalt	17,038	98.9
	No	16,410	95.2		Other	200	1.1
Day of the week	Weekend	4,697	27.2	Vehicle Nationality	Jordanian	17,084	99.1
	Weekday	12,540	72.8		Other	153	0.9
Accident Year	2015	4283 3474 4519 3395 1566	24.8	Weather	Clear	16164	93.8
	2016	3,474	20.2		Rain	968	5.6
	2017	4,519	26.2		Other	105	0.6
	2018	3,395	26.2		—	—	—
	2019	1,566	19.7		—	—	—
Road Surface Condition	Dry	15,757	91.4	Road Grade	Level	15,833	91.9
	Wet	1,431	8.3		Ascending	816	4.7
	Other	49	0.3		Descending	588	3.4
Vehicle Type	Car	12,954	75.2	Season	spring	4,572	26.5
	Bus	838	4.9		summer	3,417	19.8
	Truck	1,422	8.2		autumn	3,950	22.9
	Other	2,023	11.7		winter	5,298	30.7
Age	≤25	2,509	14.6	Accident Time	0:00–6:59	1,580	9.2
	26–35	6,252	36.3		7:00–8:59	1,085	6.3
	36–45	3,928	22.8		9:00–11:59	2,380	13.8
	46–55	2,363	13.7		12:00–16:59	5,380	31.2
	56–65	1,103	6.4		17:00–19:59	3,665	21.3
	≥ 66	1,082	6.3		20:00–23:59	3,147	18.3
Speed Limit (km/h)	≤30	27	1.6	Light Conditions	Daylight	10,590	61.4
	40	6,338	36.8		Sunrise	197	1.1
	50	4,140	24.0		Sunset	754	4.4
	60	5,302	30.8		Dark	321	1.9
	70	585	3.4		Night light insufficient	1,460	8.5
	80	540	3.1				
	90	40	0.2		Night light sufficient	3,915	22.7
	100	18	0.1				
License Category	First	5	0.03	License Category	Fifth	1,658	9.6
	Second	19	0.1		Sixth	1,158	6.7
	Third	10,140	58.8		Seventh	50	0.3
	Forth	4,082	23.7		Other	125	0.7

$$Speed = \frac{Distance}{Time} \times 3.6 \quad (1)$$

Where:

- Speed: vehicle speed (km/h)
- Distance: distance traveled by a vehicle between the two definite points (m)
- Time: time a vehicle needs to pass the two definite points (second)
- 3.6: speed conversion factor from (m/sec) to (km/h).

RESULTS AND DISCUSSION

The important influencing factors were specified using the CART algorithm in the SPSS software. Age, vehicle type, speed limit, day of the week, season, accident year, accident time, license category, and light conditions affect speeding violations. Age groups (26–35) and (36–45) were associated with high speeding violations levels; this result is matched with Al-Omari et al., Tseng, Liang and Xiao's (Al-Omari et al., 2013; Tseng, 2013; Liang and Xiao, 2020) studies. This is because young drivers underestimate the speeding risk and overestimate their driving abilities, so they tend to drive at higher speeds. Another influencing factor is the vehicle type. Passenger cars violate speed more than other types of vehicles, and this is similar to Al-Omari et al. (2013), Balasubramanian and Sivasankaran (2021) concluded that light motor and two-wheeler vehicles increase speeding violations because of the high number of registered vehicles. The Annual Report of Traffic Accidents for 2019 mentioned that Jordan's number of registered vehicles increased from 1,412,817 in 2015 to 1,677,061 in 2019 (PSD and Jordanian Public Security Directorate, 2019).

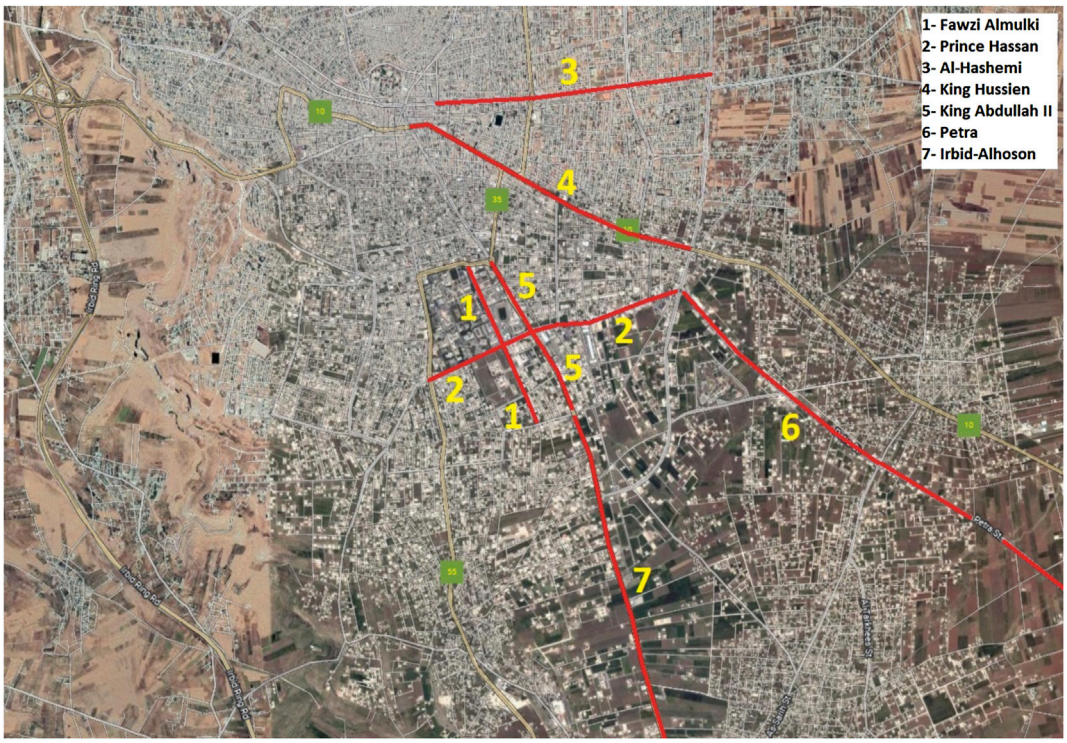


FIGURE 1 | Aerial photo for the hot spot locations (Google Map).

TABLE 2 | Segments lengths for speed measurements by video.

Street name	Segment Length (m)
Fawzi Almulki	27.8
Prince Hasan	49.0
Al-Hashmi	39.2
King Hussein	42.8
King Abdallah II	37.4
Petra	46.6
Irbid- Alhoson	41.75

The CART results showed a relationship between speeding violations and speed limits. A speed limit of 40 km/h affects speeding violations. Most speeding violations in Irbid are near residential areas and close to intersections where speed limits are relatively low. Regarding accident time, the speeding violations increase during the period (7:00–8:59), which is known as rush time, and that is similar to (Ambo et al., 2020) and (Zhang et al., 2014) results. In the morning, drivers are in a hurry to avoid being late to work, so they tend to increase their speed. Another influencing factor is the accident year. The years 2015 and 2016 were associated with high percentages of speeding violations. A decrease in speeding violations happened during the years from 2017 to 2019; this may be because of implementing some countermeasures that increase road safety, such as installing radar cameras and red-light cameras. Regarding the day of the week factor, weekdays experienced high-speed levels. This result is not compatible with

FIGURE 2 | Fawzi Almulki street (left) and Prince Hasan street (right).

TABLE 3 | Classifiers' evaluation metrics.

Classifier	Accuracy	Kappa statistic	F-measure	RMSE
CART	97.47	0.949	0.975	0.156
J48	97.22	0.944	0.972	0.162

(Shawky et al., 2017) studies. Balasubramanian and Sivasankaran (2021) concluded that there is no relation between weekdays and speeding, while Shawky et al., 2017) concluded that speeding violations increased during the weekends because of the road-based trips between different cities. Increasing speeding violations in Irbid on weekdays can be linked with the accident time factor, where most violations occur at rush hour on weekdays.

The season is another influencing factor. A high percentage of speeding violations was noticed in the winter; this result contradicts (Zhang et al., 2014) study, which indicated no relationship between season and speeding violations. Slippery roads in winter because of ice, snow, or rain with low temperatures increase braking distance, which increases the probability of speeding (Li et al., 2021). Some factors do not strongly correlate with speeding violations, such as light conditions and license category. Regarding light conditions, daylight affected speeding violations. This result is consistent with (Balasubramanian and Sivasankaran, 2021) results. The clear visibility at daylight encourages driving at high-speed levels. The third category was associated with speeding violations for the license category factor. This factor can be linked with the vehicle type factor since drivers with third category licenses drive passenger cars.

CART and J48 algorithms were applied to assess their efficiencies in predicting speeding violations. Accuracy, Kappa statistic, F-measure, and Root Mean Squared Error (RMSE) are

TABLE 4 | Descriptive statistics for radar and video measurements.

Street name	Fawzi Almulki	Prince Hasan	Al-Hashmi	King Hussein	King Abdallah II	Petra	Irbid-Alhoson
Radar Measurements							
Mean (km/h)	34.8	50.6	43.8	44	51.2	81.2	64.4
Mode (km/h)	34	45	42	43	56	74	71
Median (km/h)	35	50.5	43.5	43.5	52	80	63
50th percentile (km/h)	34.2	50	43	43	51.6	79.7	62.8
85th percentile (km/h)	43	59.7	48.8	51.3	58	93.7	73.7
Violators (%)	0	17	0	2	10	8	3
Speed Limit (km/h)	60	60	60	60	60	100	80
Video Measurements							
Mean (km/h)	42	46.2	50.8	45	48.6	72.7	58.5
Mode (km/h)	48.1	46.1	53.9	41.9	45.6	73.3	62.1
Median (km/h)	43	46.1	51	44.4	48.8	72.3	58.8
50th percentile (km/h)	42.95	46.03	50.3	44.3	48.7	72	58.7
85th percentile (km/h)	49.6	57.2	56.7	51.9	56.8	86.5	67.3
Violators (%)	0	3	7	5	10	2	1
Speed Limit (km/h)	60	60	60	60	60	100	80

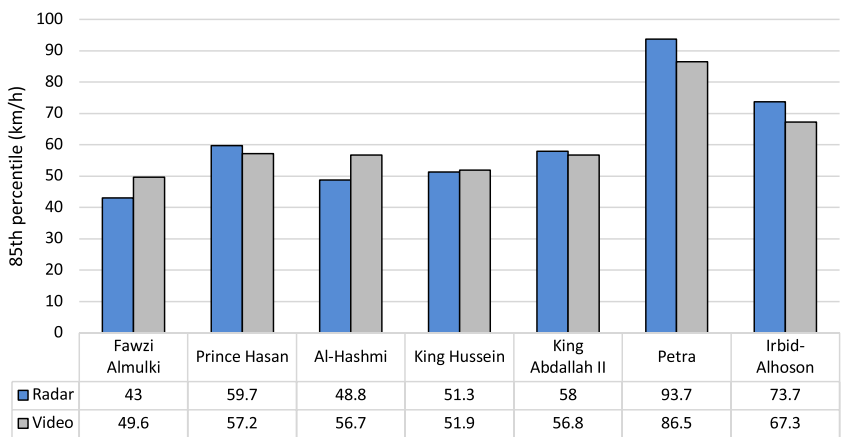


FIGURE 3 | 85th percentile speeds (km/h).

TABLE 5 | Kolmogorov-Smirnov test and independent samples *t*-Test results for each street.

Street name	Radar Kolmogorov-Smirnov (sig)	Video Kolmogorov-Smirnov (sig)	Leven's test	*t*-test (sig)	Mean (radar) km/h	Mean (video) km/h	More accurate method
Fawzi Almulki	0.20	0.20	0.03	0.00	34.8	42.0	Video
Prince Hasan	0.08	0.20	0.06	0.00	50.6	46.2	Radar
Al-Hashmi	0.20	0.20	0.31	0.00	43.8	50.8	Video
King Hussein	0.20	0.07	0.97	0.35	44	45	No Difference
King Abdallah II	0.20	0.20	0.30	0.02	51.2	48.6	Radar
Petra	0.20	0.20	0.75	0.00	81.17	72.68	Radar
Irbid- Alhoson	0.15	0.20	0.7	0.00	64.4	58.5	Radar

TABLE 6 | Mann-Whitney *U* test, results for all readings.

Method	Mean (km/h)	Kolmogorov-Smirnov (sig)	Mann-Whitney *U* test (sig)
Radar	52.8	0.00	0.76
Video	51.9	0.00	

the evaluation metrics for assessing the performance of the classifiers. **Table 3** shows the evaluation metrics for these classifiers.

According to **Table 3**, all evaluation metrics are nearly identical. The classifiers' accuracy scores of 97.47 for CART and 97.22 for J48 are both excellent. The Kappa statistic and F-measure results are greater than 0.5 for both algorithms, showing that the predicted values are very close to the true ones. The Kappa statistic values are 0.949 and 0.944, and the F-measure values are 0.975 and 0.972 for CART and J48, respectively. The RMSE results are relatively low, at 0.156 and 0.162 for CART and J48, respectively. The results for the two classifiers showed excellent performance, which means that these algorithms can predict speeding violations efficiently.

As shown in **Table 4**, the radar and video readings have 50th and 85th percentile speeds, respectively. Five streets have a speed limit of 60 km/h, which are Fawzi Almulki, Prince Hasan, Al-Hashmi, King Hussein, and King Abdallah II streets. The 50th and 85th percentile speeds for the radar measurements ranged between (34.2–51.6) km/h and (43–58) km/h, respectively. For the video, the percentiles ranged between 42.95 and 48.7 km/h and 49.6–56.8 km/h, respectively. As it was evident from these ranges, especially the 85th percentile speed, the vehicles' speeds do not exceed the speed limit of 60 km/h. This means that the speed limits are suitable for most drivers. Another observation to support this result is the low percentage of violators. The highest percentage (17%) was for Prince Hasan Street, which is not high. **Figure 3** shows the 85th percentile speeds for the radar and the video readings.

The speed limit on Petra Street is 100 km/h. The 50th percentiles for the two methods are 79.7 km/h and 72 km/h; however, 93.7 km/h and 86.52 km/h are the 85th percentile speeds for each approach. Speeds in the 50th and 85th percentiles are below the limit. The proportion of people who have broken the law (8%) is also low. These comparisons and observations indicate that the speed limit is considered suitable, safe, and acceptable for most vehicles.

Irbid-Alhoson Street has a speed limit of 80 km per hour. The radar speed limit for Irbid-Alhoson Street is 80 km per hour in the 50th and 85th percentiles. Radar readings show that the 50th and 85th percentiles are 62.8 and 73.7 km/h, respectively. The values for the video measurements are 58.7 km/h and 67.3 km/h, respectively. The values are considered acceptable by comparing these percentiles with the 80 km/h speed limit because they are less than the speed limit. From the values of the speed percentiles and the low violators' percentage of 3%, it can be

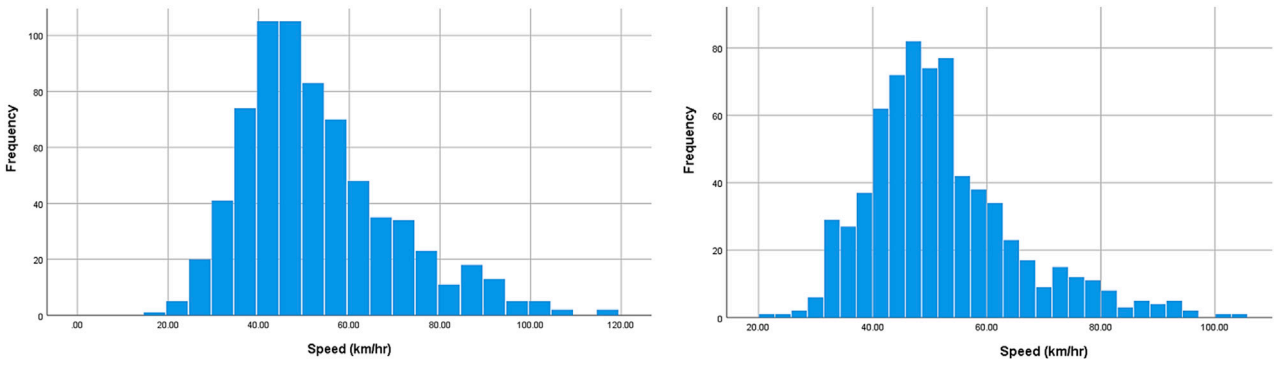


FIGURE 4 | Data distribution, Radar (left), Video (right).

concluded that the speed limit is suitable for the traffic stream. Therefore, speed limits at all locations are suitable and acceptable. This is further supported by the low number of traffic accidents at these sites.

A comparison between the radar and the video was conducted for each site and for all sites together to specify the most accurate method. The Independent Samples *t*-test was applied to examine the data for each site. The normal distribution of the data was checked using the Kolmogorov-Smirnov test for each site at a 95% confidence level. The test hypotheses are: H_o (the distribution is normal), and H_1 (the distribution is not normal).

As shown in **Table 5**, a comparison between the two approaches using the Kolmogorov-Smirnov test was performed. All of the numbers are greater than or equal to 0.05. (i.e., all the distributions are normal). Since all distributions are normal, the Independent Samples *t*-test could be applied at a 95% confidence level. The hypotheses are: H_o (the two distributions have equal means, and there is no difference between the radar and the video), and H_1 (the two distributions do not have equal means, and there is a difference between the radar and the video).

Also, Leven's test for equality of variances was conducted. The test hypotheses are: H_o (the variances are equal), and H_1 (the variances are not equal). **Table 5** also presents Leven's and *t*-test results. Leven's test results indicate that the variances are equal for all streets (p-value > 0.05), except Fawzi Almulki Street (p-value < 0.05). For Fawzi Almulki Street, the *t*-test value for unequal variances was taken. The p-values for the *t*-test are less than 0.05 for all streets except King Hussein Street. For values less than 0.05, the alternative hypothesis is accepted (i.e., there is a difference in performance between the radar and the video). The more accurate method is the one that achieved the highest mean. The radar is more accurate than the video at four locations: Prince Hasan, King Abdallah II, Petra, and Irbid-Alhoson Streets. On the other hand, the video outperforms the radar at Fawzi Almulki and Al-Hashmi streets. For King Hussein Street, the p-value for the *t*-test (0.35) is more than 0.05; we accept the null hypothesis (i.e., there is no difference in performance between the radar and the video at this location). This difference in performance can be attributed to the nature of each method. During the measurement process, drivers may have noticed the radar, which may have affected their speed. There may be many mistakes when measuring the distance and time, which makes the video readings less accurate.

The data for all locations were examined together. To determine the suitability of the Independent Samples *t*-test, the Kolmogorov-Smirnov test was used to check the normal distribution of the data. **Table 6** shows the Kolmogorov-Smirnov test results. The values for the two distributions are 0.00, which is less than 0.05 (i.e., the distributions are not normal); hence, the Independent Samples *t*-test cannot be applied here. The Mann-Whitney *U* test was applied to specify which method is more accurate at a 95% confidence level.

The similarity between sample distributions is a key assumption in Mann-Whitey U tests. The radar and video measurements are shown in **Figure 4**. The null and alternative hypotheses for the Mann-Whitney *U* test are: H_o (there is no difference between the radar and the video in measuring speed), and H_1 (there is a difference between the radar and the video in measuring speed).

As seen in **Table 6**, the Mann-Whitney *U* test result (0.76) is greater than 0.05; the null hypothesis is accepted (i.e., there is no difference between the radar and the video in measuring speed). Although there were differences in performance between the two methods when the data were analyzed separately, these differences disappeared when the data were examined together. According to this, the accuracy of the two approaches is reasonably similar.

CONCLUSION

This research aimed to study speeding violations and the factors that affect them in Irbid city, Jordan using phone camera, radar, and machine learning. In addition to that, seven hot spots (streets with the most frequent traffic accidents) for speeding violations were specified to assess speed limits at these locations. A total of 17,237 records were analyzed by applying two types of decision trees: CART and J48 algorithms, using the WEKA software. Different variables related to the vehicle, driver, roadway, and environment were investigated.

The significant influencing factors were specified using CART algorithm. Speeding violations were affected by age, vehicle type, speed limit, day of the week, season, accident year, accident time, license category, and lighting conditions. Young drivers tend to underestimate speeding risks and overestimate their driving abilities. Age groups (26–35) and (36–45) were associated with speeding violations, with violators' percentages equal to (36.3%) and (31.4%), respectively. Also, winter was associated with a high percentage of speeding violations. Slippery roads due to rain, snow, or ice at low temperatures increase braking distance, which increases speeding violations. Besides, speed limits affect speeding violations. The speed limit of 40 km/h at residential areas and near intersections was associated with speeding violations. Additionally, speeding violations experienced a rise during the morning rush hours (7:00–9:00 a.m.), where drivers hurry to avoid being late to work. Moreover, increasing the number of cars, morning rush hours during weekdays, and clear visibility during daylight increased speeding violations. Furthermore, results showed that CART and J48 algorithms predicted speeding violations efficiently. Finally, Mann-Whitney *U* test results did not show any significant difference between the radar and the video. These two methods almost have the same accuracy.

This research utilized CART and J48 algorithms to examine speeding violations. For future research, it is recommended to employ a variety of machine learning algorithms as well as other programs or programming languages to analyze speeding violations. Using established statistical and machine learning techniques, speed violations could also be investigated to compare how well each method predicts speeding violations. Moreover, advanced optimization algorithms can be used in

various fields as solution approaches to challenging decision problems, including online learning and learning automatons (LA) based on a decomposition-based multi-objective optimization framework (Zhao and Zhang, 2020), scheduling (Abd Elaziz et al., 2021), multi-objective optimization (Dulebenets et al., 2020; Liu et al., 2020), medicine (D'Angelo et al., 2019), and transportation (Chong and Osorio, 2018). Future recommendations could be based on studies that look at how well-advanced optimization algorithms work and how well they can predict factors that lead to speeding violations.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

REFERENCES

Abd Elaziz, M., Abualigah, L., and Attiya, I. (2021). Advanced Optimization Technique for Scheduling IoT Tasks in Cloud-Fog Computing Environments. *Future Gener. Comput. Syst.* 124, 142–154. doi:10.1016/j.future.2021.05.026

Al Bodour, W., Hanandeh, S., Hajij, M., and Murad, Y. (2022). Development of Evaluation Framework for the Unconfined Compressive Strength of Soils Based on the Fundamental Soil Parameters Using Gene Expression Programming and Deep Learning Methods. *J. Mater. Civ. Eng.* 34 (2), 04021452. doi:10.1061/(asce)mt.1943-5533.0004087

Albayrak, M. B., Özcan, İ. Ç., Can, R., and Dobruszkes, F. (2020). The Determinants of Air Passenger Traffic at Turkish Airports. *J. Air Transp. Manag.* 86, 101818. doi:10.1016/j.jairtraman.2020.101818

Al-Ghamdi, A. S. (2003). Analysis of Traffic Accidents at Urban Intersections in Riyadh. *Accid. Analysis Prev.* 35 (5), 717–724. doi:10.1016/s0001-4575(02)00050-7

Al-Mistarehi, B. W., Alomari, A. H., Obaidat, M. T., and Al-Jammal, A. A. (2021). Driver Performance through the Yellow Phase Using Video Cameras at Urban Signalized Intersections. *Transp. Probl.* 16 (1), 51–64. doi:10.21307/tp-2021-005

Al-Mistarehi, B. W., Alomari, A. H., Imam, R., and Mashaqba, M. (2022). Using Machine Learning Models to Forecast Severity Level of Traffic Crashes by R Studio and ArcGIS. *Front. Built Environ.* 8, 860805. doi:10.3389/fbuil.2022.860805

Alomari, A. H., and Taamneh, M. M. (2020). Front-Seat Seatbelt Compliance in Jordan: An Observational Study. *Adv. Transp. Stud.* 2020 (52), 101–116. http://www.atsinternationaljournal.com/index.php/2020-issues/lii-november-2020/1167-front-seat-seatbelt-compliance-in-jordan-an-observational-study

Al-Omari, B., Al-Ja'afari, A., and Abu-Lebdeh, G. (2013). "Effect of Driver & Vehicle Characteristics on Speeds at a Tangent Section of a Rural Highway in Jordan," in 2013 5th International Conference on Modeling, Simulation and Applied Optimization (ICMSAO), 1–4. doi:10.1109/ICMSAO.2013.6552655

Alomari, A., Aldalalah, M., Al-Dalaika, M. A.-D., Sawae, G., Sawae, S., and Al-Alawneh, D. (2021a). Evaluation of Urban Public Transport: A Case Study of Yarmouk University. *Saudi J. Civ. Eng.* 5 (1), 8–17. doi:10.36348/sjce.2021.v05i01.002

Alomari, A. H., Al-Omari, B. H., Al-Adwan, M. E., and Sandt, A. (2021b). Investigating and Modeling Speed Variability on Multilane Highways. *Adv. Transp. Stud.* 54, 5–16.

Alomari, A. H., Al-Omari, B. H., and Al-Adwan, M. E. (2021c). Analysis of Speed Variance on Multilane Highways in Jordan. in Proceedings of the 1st International Congress on Engineering Technologies. CRC Press, 206–216. doi:10.1201/9781003178255-27

Ambo, T. B., Ma, J., and Fu, C. (2020). Investigating Influence Factors of Traffic Violation Using Multinomial Logit Method. *Int. J. Inj. Control Saf. Promot.* 28 (1), 78–85. doi:10.1080/17457300.2020.1843499

Atieh, A. A., Naghawi, H., and Jadaan, K. (2020).Analysis of the Speeding Behaviour of Jordanian Drivers. *Resil. Sustain. Transp. Syst. Rest. VA Am. Soc. Civ. Eng.* 2020, 28–34. doi:10.1061/9780784482902.004

Atombo, C., Wu, C., Zhong, M., and Zhang, H. (2016). Investigating the Motivational Factors Influencing Drivers Intentions to Unsafe Driving Behaviours: Speeding and Overtaking Violations. *Transp. Res. Part F Traffic Psychol. Behav.* 43, 104–121. doi:10.1016/j.trf.2016.09.029

Balasubramanian, V., and Sivasankaran, S. K. (2021). Analysis of Factors Associated with Exceeding Lawful Speed Traffic Violations in Indian Metropolitan City. *J. Transp. Saf. Secur.* 13 (2), 206–222. doi:10.1080/19439962.2019.1626962

Bara'W, A. M., Alomari, A. H., and Al Zoubi, M. S. (2021). Investigation of Saturation Flow Rate Using Video Camera at Signalized Intersections in Jordan. *Open Eng.* 11 (1), 216–226. doi:10.1515/eng-2021-0021

Belokurov, V., Spodarev, R., and Belokurov, S. (2020). Determining Passenger Traffic as Important Factor in Urban Public Transport System. *Transp. Res. Procedia* 50, 52–58. doi:10.1016/j.trpro.2020.10.007

Bener, A., Jadaan, K., Crundall, D., and Calvi, A. (2019). The Effect of Aggressive Driver Behaviour, Violation and Error on Vehicle Crashes Involvement in Jordan. *Int. J. Crashworthiness* 25 (3), 276–283. doi:10.1080/13588265.2019.1583422

Bolsunovskaya, M., Leksashov, A., Shirokova, S., and Tsygan, V. (2021). Development of an Information System Structure for Photo-Video Recording of Traffic Violations, *E3S Web Conf*, 244, 07007. doi:10.1051/e3sconf/202124407007

CDC (2014). CDC Global Health – Jordan. [online] Available at: https://www.cdc.gov/globalhealth/countries/jordan/default.htm (Accessed February 15, 2022).

Cheng, Z., Lu, J., Zu, Z., and Li, Y. (2019). Speeding Violation Type Prediction Based on Decision Tree Method: A Case Study in Wujiang, China. *J. Adv. Transp.* 2019, 1–10. doi:10.1155/2019/8650845

Chong, L., and Osorio, C. (2018). A Simulation-Based Optimization Algorithm for Dynamic Large-Scale Urban Transportation Problems. *Transp. Sci.* 52 (3), 637–656. doi:10.1287/trsc.2016.0717

D'Angelo, G., Pilla, R., Tascini, C., and Rampone, S. (2019). A Proposal for Distinguishing between Bacterial and Viral Meningitis Using Genetic Programming and Decision Trees. *Soft Comput.* 23 (22), 11775–11791. doi:10.1007/s00500-018-03729-y

Dulebenets, M. A., Pasha, J., Kavoosi, M., Abioye, O. F., Ozguven, E. E., Moses, R., .Boot, W. R., and Sando, T. (2020). Multiobjective Optimization Model for Emergency Evacuation Planning in Geographical Locations with Vulnerable Population Groups. *J. Manage. Eng.* 36 (2), 04019043. doi:10.1061/(asce)me.1943-5479.0000730

Dulebenets, M. A. (2020). An Adaptive Island Evolutionary Algorithm for the Berth Scheduling Problem. *Memetic Comp.* 12 (1), 51–72. doi:10.1007/s12293-019-00292-3

Elmi, Z., Singh, P., Meriga, V. K., Goniewicz, K., Borowska-Stefańska, M., Wiśniewski, S., et al. (2022). Uncertainties in Liner Shipping and Ship Schedule Recovery: A State-Of-The-Art Review. *J. Mar.Sci. Eng.* 10 (5), 563. doi:10.3390/jmse10050563

Enoch, M. P., Cross, R., Potter, N., Davidson, C., Taylor, S., Brown, R., .Huang, H., Parsons, J., Tucker, S., Wynne, E., Grieg, D., Campbell, G., Jackson, A., and Potter, S. (2020). Future Local Passenger Transport System Scenarios and Implications for Policy and Practice. *Transp. policy* 90, 52–67. doi:10.1016/j.tranpol.2020.02.009

Hanandeh, S. (2022a). Introducing Mathematical Modeling to Estimate Pavement Quality Index of Flexible Pavements Based on Genetic Algorithm and Artificial Neural Networks. *Case Stud. Constr. Mater.* 16, e00991. doi:10.1016/j.cscm.2022.e00991

Hanandeh, S. (2022b). Evaluation Circular Failure of Soil Slopes Using Classification and Predictive GEP Schemes. *Front. Built Environ.* 8, 1–11. doi:10.3389/fbuil.2022.858020

Hazaymeh, K., Almagbile, A., and Alomari, A. H. (2022). Spatiotemporal Analysis of Traffic Accidents Hotspots Based on Geospatial Techniques. *ISPRS Int. J. Geo-Infor.* 11 (4), 260. doi:10.3390/ijgi11040260

Imam, R., Murad, Y., Asi, I., and Shatnawi, A. (2021). Predicting Pavement Condition Index from International Roughness Index Using Gene Expression Programming. *Innov. Infrastruct. Solutions* 6 (3), 139. doi:10.1007/s41062-021-00504-1

Inada, H., Ashraf, L., and Campbell, S. (2021). COVID-19 Lockdown and Fatal Motor Vehicle Collisions Due to Speed-Related Traffic Violations in Japan: a Time-Series Study. *Inj. Prev.* 27 (1), 98–100. doi:10.1136/injuryprev-2020-043947

Jadaan, K., Albeetar, N., Abuhalimeh, D., and Naji, Y. (2021). Analysis of Driver Behavior in Amman Using Manchester Driver Behavior Questionnaire. *Acta Tech. Jaurinensis* 14 (4), 440–454. doi:10.14513/actatechjaur.00599

Javid, M. A., and Al-Roushdi, A. F. A. (2019). Causal Factors of Driver's Speeding Behaviour, a Case Study in Oman: Role of Norms, Personality, and Exposure Aspects. *Int. J. Civ. Eng.* 17 (9), 1409–1419. doi:10.1007/s40999-019-00403-8

Koushki, P., Al-Saleh, O., Yaseen, S., and Ali, M. (2021). "On Fatal and Injurious Pedestrian Accidents," in Presented at the 80th Annual Meeting of the Transportation Research Board, Washington, D.C, January. 2001.

Kuşkapan, E., Çodur, M. Y., and Atalay, A. (2021). Speed Violation Analysis of Heavy Vehicles on Highways Using Spatial Analysis and Machine Learning Algorithms. *Accid. Analysis Prev.* 155, 106098. doi:10.1016/j.aap.2021.106098

Li, Y., Li, M., Yuan, J., Lu, J., and Abdel-Aty, M. (2021). Analysis and Prediction of Intersection Traffic Violations Using Automated Enforcement System Data. *Accid. Analysis Prev.* 162, 106422. doi:10.1016/j.aap.2021.106422

Liang, Z., and Xiao, Y. (2020). Analysis of Factors Influencing Expressway Speeding Behavior in China. *Plos One* 15 (9), e0238359. doi:10.1371/journal.pone.0238359

Liu, Z.-Z., Wang, Y., and Huang, P.-Q. (2020). AnD: A Many-Objective Evolutionary Algorithm with Angle-Based Selection and Shift-Based Density Estimation. *Inf. Sci.* 509, 400–419. doi:10.1016/j.ins.2018.06.063

Mujalli, R. O. (2018). Modeling Risk of Road Crashes Using Aggregated Data. *Jordan J. Civ. Eng.* 12(1), 45–60. https://jjce.just.edu.jo/issues/show_paper.php?pid=4150

Murad Zuhair, Y. (2021). Predictive Model for Bidirectional Shear Strength of Reinforced Concrete Columns Subjected to Biaxial Cyclic Loading. *Eng. Struct.* 244, 112781. doi:10.1016/j.engstruct.2021.112781

Nathanail, E. G., Prevedouros, P. D., Mintu Miah, M., and De Melo Barros, R. (2021). Predicting the Impact of Text-Reading Using Decision Trees. In EPIA Conference on Artificial Intelligence. doi:10.1007/978-3-030-30241-2_57

Obaidat, M. T., Al-Khateeb, G., Khedaywi, T., and Bannoura, B. (2008). "Evaluation of Drivers' Behavioral Mistakes in Jordan," in Fifth Traffic Safety Conference in Jordan, Jorden, 22–24.

PSD, Jordanian Public Security Directorate (2019). *Annual Report of Traffic Accidents for Jordan*. Amman: Public Security Directorate. Available at: https://www.psd.gov.jo/images/docs/TraddicRep2019.pdf (Accessed April 30, 2021).

Rahimi, A. M., Dulebenets, M. A., and Mazaheri, A. (2021). Evaluation of Microsimulation Models for Roadway Segments with Different Functional Classifications in Northern Iran. *Infrastructures* 6 (3), 46. doi:10.3390/infrastructures6030046

Saravana, N., Gayathri, V., and thri, V. G. (2018). Performance and Classification Evaluation of J48 Algorithm and Kendall's Based J48 Algorithm (KNJ48). *Int. J.Comput.Trends. Technol.* 59 (2), 73–80. doi:10.14445/22312803/ijctt-v59p112

Shah, T. N., Khan, M. Z., Ali, M., Khan, B., and Idress, N. C. A. R. T., J-48graft, J48, ID3 (2020). Decision Stump and Random Forest: A Comparative Study. *Univ. Swabi J. (USJ); Open access*. 2020, 1–6. http://www.uoswabi.edu.pk/data/downloads/1125.pdf

Shawky, M., Sahnoon, I., and Al-Zaidy, A. (2017). Predicting Speed-Related Traffic Violations on Rural Highways. In Proceedings of the 2nd World Congress on Civil, Structural, and Environmental Engineering (CSEE'17), Spain. þdoi:10.11159/icte17.117

Sutela, M., and Aaltonen, M. (2020). Effects of Temporal Characteristics and Weather Conditions on Speeding Sanction Rates in Automatic Traffic Enforcement. *Police J.* 94 (4), 590–615. doi:10.1177/0032258x20977024

Tarawneh, B., and Imam, R. (2014). Regression versus Artificial Neural Networks: Predicting Pile Setup from Empirical Data. *KSCE J. Civ. Eng.* 18 (4), 1018–1027. doi:10.1007/s12205-014-0072-7

Tseng, C.-M. (2013). Speeding Violations Related to a Driver's Social-Economic Demographics and the Most Frequent Driving Purpose in Taiwan's Male Population. *Saf. Sci.* 57, 236–242. doi:10.1016/j.ssci.2013.02.005

WHO, World Health Organization (2021). Road Traffic Injuries. [online] Available at: https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries (Accessed February 15, 2022).

Witten, I. H., Frank, E., Hall, M. A., and Pal, C. J. (2016). The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques". *Morgan Kaufmann*. doi:10.1016/C2015-0-02071-8

Wu, Y. W., and Hsu, T. P. (2019). Mining Characteristics of Speeding and Red Light Running Violations Using Association Rules. *J. East. Asia Soc. Transp. Stud.* 13, 2111–2125. þ.

Zamanov, M. S. (2012). Development of a Prediction Model for Speed Limit Violations on Tangent Road Sections. Master's Thesis. Gelderland: Delft University of Technology.

Zhang, G., Yau, K. K. W., and Gong, X. (2014). Traffic Violations in Guangdong Province of china: Speeding and Drunk Driving. *Accid. Analysis Prev.* 64, 30–40. doi:10.1016/j.aap.2013.11.002

Zhao, H., and Zhang, C. (2020). An Online-Learning-Based Evolutionary Many-Objective Algorithm. *Inf. Sci.* 509, 1–21. doi:10.1016/j.ins.2019.08.069

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor YM declared a shared affiliation with the author RI at the time of review.

frontiers | Frontiers in Built Environment

ORIGINAL RESEARCH
published: 13 July 2022
doi: 10.3389/fbuil.2022.949167

Prediction of Marshall Test Results for Dense Glasphalt Mixtures Using Artificial Neural Networks

Yazeed S. Jweihan[1], Roaa J. Alawadi[2], Yazan S. Momani[3] and Ahmad N. Tarawneh[4]*

[1]Civil and Environmental Engineering Department, Faculty of Engineering, Mutah University, Karak, Jordan, [2]Civil Engineering Department, Faculty of Engineering, Applied Science Private University, Amman, Jordan, [3]Civil Engineering Department, Faculty of Engineering, University of Petra, Amman, Jordan, [4]Civil Engineering Department, Faculty of Engineering, The Hashemite University, Zarqa, Jordan

Asphalt mixture comprising waste glass as an aggregate is referred to as "*glasphalt*". Limited studies have been oriented to investigate the Marshall test results of dense-graded glasphalt mixes considering a wide range of variables. This study aims to utilize the artificial neural networks (ANNs) to develop predictive models for Marshall stability and Marshall flow of dense glasphalt mixes based on a large experimental database collected from the literature. Eight independent variables covering the material and mix properties were utilized as inputs in the models. The proposed models resulted in an experimental-to-predicted ratio of 1.00 and 1.00, coefficient of variation of 8.6% and 8.7%, RMSE of 1.63 kN and 0.54 mm, and R-squared of 93.6% and 85.7% for the glasphalt stability and flow models, respectively. Comprehensive parametric analyses have been conducted to further validate the models by investigating the sensitivity of their parameters to the predicted stability and flow values. The analyses revealed some desirable design values that could be considered for a better performance of dense glasphalt mixes. The results indicate that 4% is the desired design air void content of glasphalt mixes. High stability value can be achieved for glasphalt mixes containing a crushed aggregate of 12.5 mm maximum size and 50% glass cullet of 4.75 mm maximum size. Lower viscosity asphalt binder would provide uniformly compacted mixes. Furthermore, glasphalt flow increases as the maximum size of ingredient particles, the penetration grade of asphalt cement, asphalt cement content, and VMA% increase.

Keywords: waste glass, prediction models, glasphalt, Marshall stability, flow

OPEN ACCESS

Edited by:
Yasmin Murad,
The University of Jordan, Jordan

Reviewed by:
Taqwa Alhadidi,
Al-Ahliyya Amman University, Jordan
Bara' Al-Mistarehi,
Jordan University of Science and Technology, Jordan
Taher Ahmed,
University of Anbar, Iraq

****Correspondence:***
Yazeed S. Jweihan
y.jweihan@mutah.edu.jo

Specialty section:
This article was submitted to Computational Methods in Structural Engineering, a section of the journal Frontiers in Built Environment

Received: *20 May 2022*
Accepted: *09 June 2022*
Published: *13 July 2022*

Citation:
Jweihan YS, Alawadi RJ, Momani YS and Tarawneh AN (2022) Prediction of Marshall Test Results for Dense Glasphalt Mixtures Using Artificial Neural Networks. Front. Built Environ. 8:949167. doi: 10.3389/fbuil.2022.949167

INTRODUCTION

Environmental waste materials increase every day as a result of domestic and industrial waste production, and landfill areas for those wastes are limited continuously (Ektas and Karacasu, 2012). Incorporating waste materials into asphalt mixtures has been investigated worldwide (Rahman et al., 2020). It has significant benefits in reducing the adverse impacts of waste materials on the environment, saving raw construction materials, using sustainable and cost-effective materials in asphalt mixes (Jony et al., 2011; Anochie-Boateng and George, 2016; Almuhmdi et al., 2021), and mitigating the high cost of landfill disposal (Disfani et al., 2011).

Many researchers have investigated the feasibility of utilizing waste crushed glass for road materials. Viswanathan (Viswanathan, 1996) investigated the applicability of utilizing glass cullet as a

highway material and found that the glass waste has properties similar to natural aggregates and could be used as a highway construction material. As the awareness on environmental and natural resource issues increases, waste materials such as glass powder can be beneficially used in road construction (Finkle and Ksaibati, 2007; Arabani et al., 2017), and the use of glass cullet, with a maximum size of 12.5 mm (1/2 inch) and up to 20% replacement rate, was recommended as a supplement to the road-base materials. The high angularity of the glass cullet can enhance asphalt mix stability. Other characteristics of the waste glass such as low absorption, low specific gravity, and low thermal conductivity could offer better heat retention for asphalt mixes (Salem et al., 2017). Furthermore, constructing asphalt pavements with recycled glass offers an appropriate light reflection that enhances night visibility (Su and Chen, 2002).

Asphalt mixture comprising waste crush glass as an aggregate is referred to as "glasphalt" (Kandahl, 1992; Flynn, 1993). Alhassan et al. (Alhassan et al., 2018) investigated the feasibility of incorporating fine cullet from waste glass bottles with a maximum size of 4.75 mm into asphalt mixes at different percentages of 5%–10%. The study proved the applicability of using waste cullet as a partial replacement of fine aggregate in asphalt mixes. The results of Marshall stability, flow, bulk density, and air void content achieved their best values at 8.0% glass content. Dalloul (Dalloul, 2013) has also studied the effect of adding a crushed glass with a maximum fraction size of 4.75 mm into asphalt mixes at different percentages ranging from 2.5% to 15% by aggregate weight. His study indicated that 7.5% is the optimum content for waste glass fractions in asphalt mixes, where the stability and bulk density of HMA satisfied the standard limits of acceptance.

Arabani (Arabani, 2011) assesses replacing partial aggregate of both wearing and binder course mixes with 5%–20% waste glass fractions with a maximum size of 4.75 mm. The Marshall test results revealed that 15% is the most suitable glass content for the tested mixes. Su and Chen (Su and Chen, 2002) studied the applicability of recycling waste glass into asphalt pavement. They concluded that the stability values of asphalt mixes utilizing 5%–15% recycled glass (4.75 mm maximum size) are lower than those of control mixes. However, the obtained values by glasphalt mixes satisfied the acceptance limit for Marshall stability.

Waste crushed glass can be utilized in asphalt mixtures with a maximum size of 2.36 mm and an optimal replacement ratio of 10% (Salem et al., 2017). However, the Marshall stability of glasphalt mixes increased when waste glass fractions with a maximum size of 2.36 mm were added to the mixes with a 20% replacement rate (Almuhmdi et al., 2021). An opposite trend was also noticed for glasphalt density, VMA, and air void content with a 10%–25% replacement rate (Almuhmdi et al., 2021). A study conducted by Aashish et al. (Aashish and Tamrakar, 2019) showed that 10% was the desired percentage to achieve high stability and low flow for HMA containing 1.18–0.075 mm glass fractions. Issa (Issa, 2016) concluded that a significant improvement in the Marshall properties was achieved in glasphalt mixes when compared with control asphalt mixes.

Moreover, the feasibility of incorporating glass powder as a mineral filler (passing 0.075 mm, sieve (No. 200)) into asphalt mixtures has been investigated in several studies. Jony et al. (Jony et al., 2011), Arabani et al. (Arabani et al., 2017), Kifile et al. (Kifile et al., 2020), and Saltan et al. (Saltan et al., 2015a) have added the glass powder into asphalt mixes at different percentages to assess the mixes' performance in terms of the Marshall test properties. Jony et al. (Jony et al., 2011) found that using 7% of glass powder as a mineral filler in asphalt mixes is sufficient to provide a high average value of Marshall stability and lower average values of flow and density when compared to asphalt mixes containing Portland cement or limestone powder as fillers. A study conducted by Arabani et al. (Arabani et al., 2017) also showed that utilizing 6% of glass powder as a mineral filler in asphalt mixes provides the best stability and lowest flow values among other mixes containing waste brick powder, rice-husk ash, and stone dust powder. Kifile et al. (Kifile et al., 2020) investigated the effect of replacing crushed stone-dust filler with waste glass powder in asphalt mixes at different replacement rates. They concluded that as the glass powder content increases in asphalt mixes, the Marshall stability, bulk density, and voids filled with asphalt values increased, whereas the values of Marshall flow, air voids, and void in mineral aggregates decreased. The results proved that the 75% replacement rate can provide optimum performance for heavy traffic conditions (Kifile et al., 2020).

Proper mix design of asphalt mixes with suitable ingredients and proportions enhances pavement performance (Brown et al., 2009; Asphalt Institute, 2014). Marshall's design method has been widely adopted for asphalt mix design. It is essentially used for designing dense-graded asphalt mixes that have well-graded aggregate particles distributed from course to fine throughout the entire range of sieves (Garcia and Hansen, 2001; Brown et al., 2009; Asphalt Institute, 2014). In general, the method involves several steps including aggregate and bitumen evaluation, test specimen preparation, density and void analysis, Marshall stability and flow tests, and data interpretation (Brown et al., 2009). The Marshall stability is defined as the maximum load carried by a test specimen of asphalt concrete before failure when tested at 60 °C and 51 mm/min loading rate. The flow is the amount of the specimen's deformation before failure. Both properties are determined physically from the Marshall test. Other test parameters such as bulk specific gravity of asphalt mixture (G_{mb}), voids in mineral aggregate (VMA), voids filled with asphalt (VFA), and air void content (Va) are determined for the design by carrying out extra calculations of volumetric analysis (Brown et al., 2009; Asphalt Institute, 2014).

Recently, the machine learning approaches have been utilized in several studies to predict models for different aspects of construction materials and civil engineering applications (Murad et al., 2021; Al Bodour et al., 2022; Iftikhar et al., 2022; Momani et al., 2022). The artificial neural networks (ANNs), which is a recent learning machine tool, was also utilized by researchers in asphalt mix design applications (Tapkın et al., 2010; Ozgan, 2011; Singh et al., 2013; Ozturk and Kutay, 2014; Shafabakhsh et al., 2015; Ozturk et al., 2016; Zavrtanik et al., 2016; Pasetto et al., 2019; Fadhil et al., 2022; Othman, 2022). An ANN–based model to predict the Marshall mix volumetric properties has been proposed by Ozturk et al. (Ozturk et al., 2016). Ozgan (Ozgan, 2011) modeled the Marshall

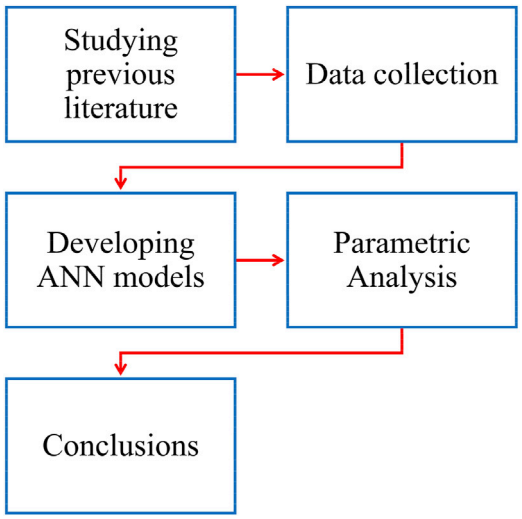


FIGURE 1 | Research program of this study.

stability of asphalt mixes under different testing temperatures and exposure time conditions using the ANN technique. Tapkın et al. (Tapkın et al., 2010) proposed a neural-network model to predict the Marshall stability, flow, and Marshall quotient for polypropylene-modified asphalt mixtures. Some physical properties of standard Marshall specimens such as asphalt content, specimen height, unit weight, VMA, VFA, Va, polypropylene types, and percentage were used as inputs in the proposed model (Tapkın et al., 2010). Fadhil et al. (Fadhil et al., 2022) have utilized the ANN approach as a tool for HMA design based on actual data of 252 mixes following the Marshall design method. They used five parameters including asphalt penetration, kinematic viscosity, aggregate surface area, abrasion, and binder content as inputs in the models to predict five parameters of HMA design including the bulk density, air voids, Marshall stability, Marshall flow, and Marshall stiffness. Othman (Othman, 2022) has predicted the optimum asphalt content (OAC) based on different aggregate gradations using ANN approaches of different activation functions, number of hidden layers, and number of neurons per layer. The volumetric properties of Superpave asphalt mixes were also modeled utilizing the ANN approach at different gyration levels (Ozturk and Kutay, 2014). In addition, the performance of asphalt pavements incorporating waste materials has been predicted using the ANN approach in several studies (Xiao et al., 2009; Ektas and Karacasu, 2012; Azarhoosh et al., 2018; Mansourian et al., 2018; Keskin and Karacasu, 2021). The fatigue behavior of glasphalt mixes has been modeled analytically by Arabani et al. (Arabani et al., 2012).

Previous studies in the literature indicate that incorporating glass waste in asphalt mixes, either as an aggregate or mineral filler substitute, provided some promising results, especially in terms of Marshall stability and Marshall flow. However, there is no mutual agreement on the effect of contributing different variables on glasphalt stability and flow. In addition, there is a lack of studies evaluating the behavior of glasphalt based on a large database with a wide range of variables including the maximum aggregate size (MAS), the maximum size of waste glass fractions (MSGF), average penetration grade of asphalt (AP), asphalt content (AC%), waste glass content (WGC%), the bulk specific gravity of the glasphalt mixture (G_{mb}), air void content (Va%), and voids in mineral aggregate (VMA%). This study utilizes the ANN technique to develop the robust predictive models for the Marshall stability and Marshall flow of the dense-graded glasphalt mixes based on a large experimental database collected from the literature. In addition, the study conducts parametric analyses to evaluate the effect of each variable on the predicted stability and flow values.

METHODOLOGY AND EXPERIMENTAL DATABASE

Figure 1 outlines the research program followed in this study. The study started with studying the previous literature on dense glasphalt mixes, their properties, and prediction models. A large worldwide database of 136 reliable experimental data points was then collected from thirteen different studies (Malisch et al., 1970; Arabani, 2011; Jony et al., 2011; Arabani and Kamboozia, 2013; Dalloul, 2013; Saltan et al., 2015b; Arabani et al., 2017; Salem et al., 2017; Al- Jameel and Al- Saeedi, 2018; Alhassan et al., 2018; Aashish and Tamrakar, 2019; Kifile et al., 2020; Almuhmdi et al., 2021) available in the literature and were carried out according to ASTM D1559 and ASTM D6927 testing procedures. Natural crushed aggregate, waste crushed glass, and asphalt cement were the main ingredients for the dense-graded glasphalt mixtures of the database test specimens. The specimens were prepared using a standard Marshall mold of 4 inches (101.6 mm) in diameter and 2.5 inches (63.5 mm) in height.

The experimental studies showed that the dense glasphalt mixes are mainly influenced by a wide variety of design and material properties. Therefore, eight independent variables including MAS, MSGF, AP, AC%, WGC%, G_{mb}, Va%, and VMA% were utilized as inputs in the models to predict the Marshall stability and Marshall flow of dense glasphalt mixes. A summary of the collected database with range values of variables is presented in **Table 1**. It should be mentioned that the database includes 24 data points of control mixes, that is, 0% WGC. The range and distribution of variables are shown in **Figure 2**. The collected data were then utilized to develop ANN models to predict Marshall test results of dense glasphalt mixes. Finally, comprehensive parametric analyses were performed to validate the models by investigating the sensitivity of each variable to the predicted stability and flow values.

ARTIFICIAL NEURAL NETWORK

ANN is a relatively new soft-computing method that has been utilized in several fields of civil engineering. The ANN is known for its ability as a powerful prediction tool that describes the relationship between input parameters and the output. The ANN is a multilayer perception-based procedure as shown in **Figure 3**.

TABLE 1 | Collected database used in the ANN models.

Reference	No. of data points	MAS (mm)	MSGF (mm)	AP (0.1 mm)	AC%	WGC %	Gmb (g/cm3)	Va%	VMA%
Dalloul Dalloul, (2013)	19	25	4.75	75	5.1	0–15	2.32–2.38	1.63–5.68	13.37–17.16
Almuhmdi, et al. Almuhmdi et al. (2021)	5	25	2.36	45	5	0–25	2.394–2.4	3.12–3.42	13.7–14.06
Aashish et al. Aashish and Tamrakar, (2019)	4	20	1.18	60	4.5–5.22	0–14	2.39–2.4	3–4	14.55–15.56
Arabani and Kamboozia Arabani and Kamboozia, (2013)	20	19	4.75	65	5–6.5	0–20	2.283–2.351	3.42–5.22	14.17–17.04
Arabani Arabani, (2011)	20	25	4.75	65	4–5.5	0–20	2.275–2.34	3.25–6.87	13.31–15.93
Alhassan et al. Alhassan et al. (2018)	7	19	4.75	65	5.9	0–10	2.31–2.35	3.17–4.36	15.94–16.99
Salem et al. Salem et al. (2017)	5	19	2.36	92.5	5.35–5.75	0–20	2.213–2.248	2.5–4.4	15.2–16.5
Jony et al. Jony et al. (2011)	4	19	0.075	45	5	0–10	2.09–2.35	3.29–5.2	12.79–21.67
Arabani et al. Arabani et al. (2017)	2	19	0.075	65	5.1–5.4	0–6	2.511–2.6	3.3–3.67	14.46–16.19
Kifile et al. Kifile et al. (2020)	5	25	0.075	92.5	5.1	0–6.5	2.35–2.36	4–4.3	14.4–14.8
Malisch et al. Malisch et al. (1970)	12	12.5	12.5	92.5	4–7	100	2.202–2.267	0.98–6.57	13.89–16.61
Saltan et al. Saltan et al. (2015b)	15	19	0.075	85	5.8–5.94	0–9	2.35–2.416	3.4–4.4	14.1–17.43
Al-Jamee and Al-Saeedi Al- Jameel and Al-Saeedi, (2018)	18	37.5	0.6–12.5	45	3.8	0–100	2.235–2.391	2.3–5.9	11.2–14.25

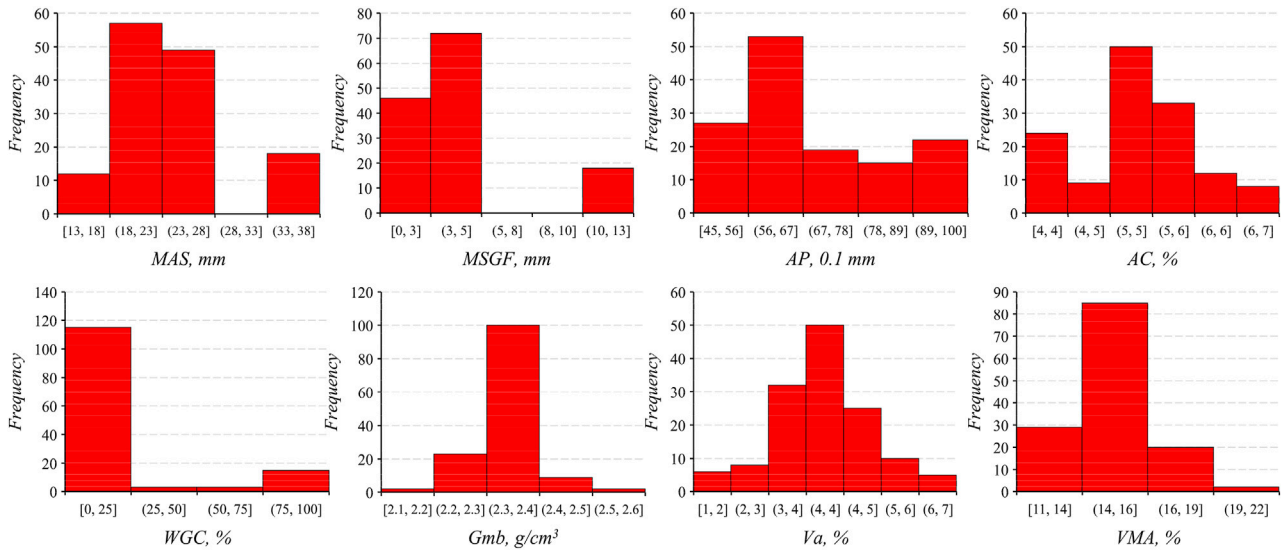


FIGURE 2 | Distribution of the selected variables used in the ANN models.

The first layer is called the *Input Layer*, which includes the inputs (independents) of the problem, a hidden layer, and an output layer that contains the output of the problem of interest. Each layer consists of processing units that are fully connected with the processing units in the subsequent layer.

Each signal or input xi that reaches the processing units is multiplied by a calibration weight w_{ji} that controls the signal's behavior and significance. The calibrated signals are added to a calibrated bias value B_{ji} at each processing unit, as shown in **Eq. (1)**. The combined input I_j is then sent through a nonlinear transfer function $f(I_j)$ to generate the processing output, which will be used as the input for the following layer's processing units (**Figure 3**). This research uses a hyperbolic tangent transfer sigmoid function.

$$I_j = \sum w_{ji}\, x_i + B_{ji}. \tag{1}$$

An algorithm adjusts the input weights and bias values using the Levenberg–Marquardt optimization approach to train the ANN model. To construct a network that generalizes well, the optimization technique minimizes a combination of squared errors and weights and then selects the optimum combination. The method is known as Bayesian regularization. The mean square error (MSE), also known as the performance function (**Eq. (2)**) [X], is used to express the error. The training procedure will continue until the MSE converges and no more improvement is possible. Eighty percent of the data was used for training, and the rest was used for testing and validation. MATLAB software was used to conduct the analysis.

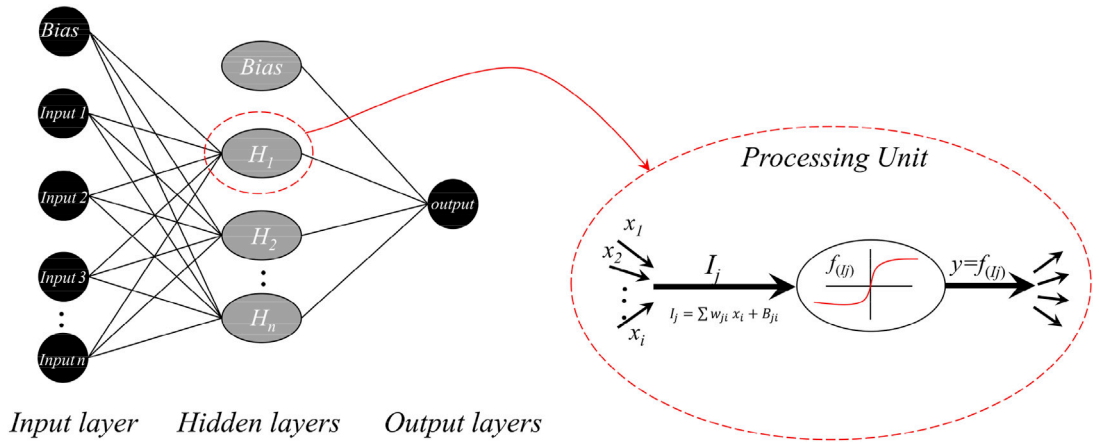


FIGURE 3 | General structure and processing unit of ANNs.

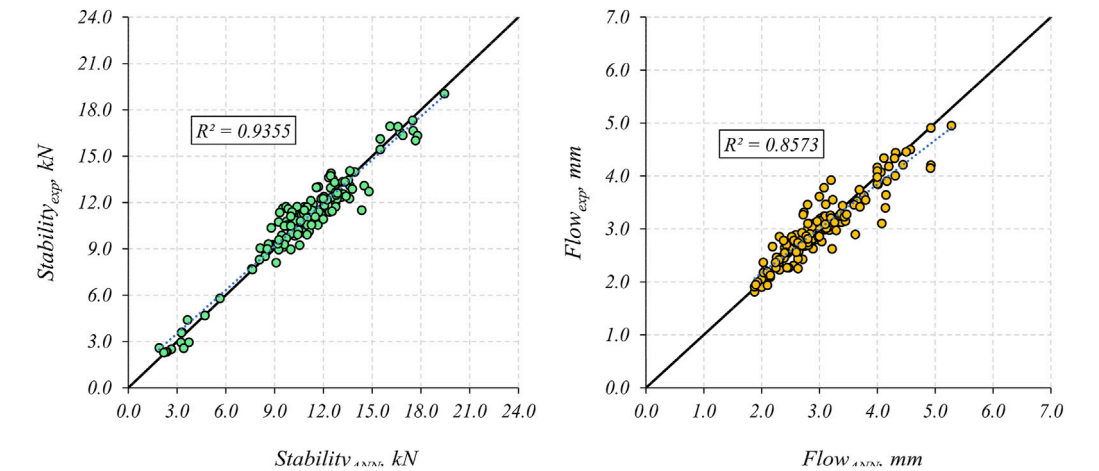


FIGURE 4 | General structure and processing unit of ANNs.

$$MSE = \frac{1}{N}\sum_{k=1}^{N}(Actual - Predicted)^2 . \quad (2)$$

Several iterations were used to choose the needed number of processing units in the hidden layer to represent and approximate the relationships between variables. The best number of layers was encountered by trying processing units from 1 to 10 while monitoring the MSE to minimize overfitting. Due to the use of Bayesian regularization and the amount of data in the current case, overfitting was not an issue.

In this study, two ANN models with two processing units in the hidden layers were constructed to predict the stability and flow of the dense glasphalt. The input layer for each model includes MAS, MSGF, AP, AC%, WGC%, G_{mb}, Va%, and VMA%. The proposed models resulted in an experimental-to-predicted ratio of 1.00 and 1.00, coefficient of variation (CoV) of 8.6% and 8.7%, and RMSE of 1.63 kN and 0.54 mm for the glasphalt stability and flow models, respectively. This indicates a good agreement with the test results. The prediction models are presented in **Figure 4**.

PARAMETRIC ANALYSES

The developed ANN models are used to perform parametric analyses to evaluate the effect of each independent variable on the Marshall stability and flow values of dense-graded glasphalt mixes. The parametric analyses were conducted by fixing the mean values of all variables and then varying the values of each concerned variable independently. **Figures 5** and **6** show the plots of each independent variable on the stability and flow values of the dense-graded glasphalt mixes, respectively. These plots are useful as they enable better and simple visualization of the influence of each studied variable on the output parameters.

It can be seen from **Figure 5A** that the stability of dense-graded glasphalt mixes of small maximum aggregate size (MAS)

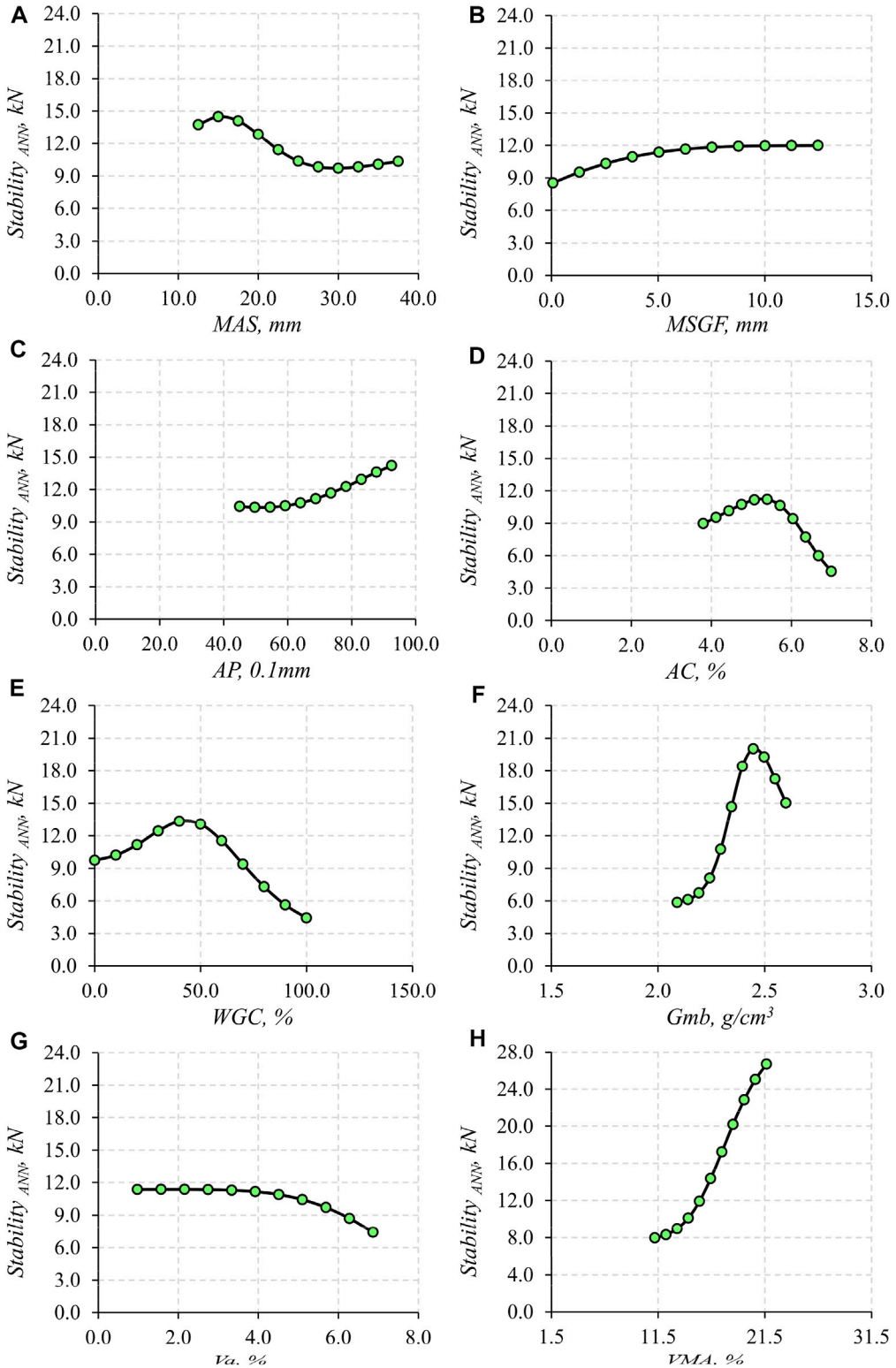


FIGURE 5 | Parametric study for glasphalt stability versus **(A)** Maximum aggregate size (MAS), **(B)** Maximum size of waste glass fractions (MSGF), **(C)** Average penetration grade of asphalt (AP), **(D)** Asphalt content (AC%), **(E)** Waste glass content (WGC%), **(F)** Bulk specific gravity of glasphalt mixture, **(G)** Air voids content (Va%), and **(H)** Voids in mineral aggregate (VMA%).

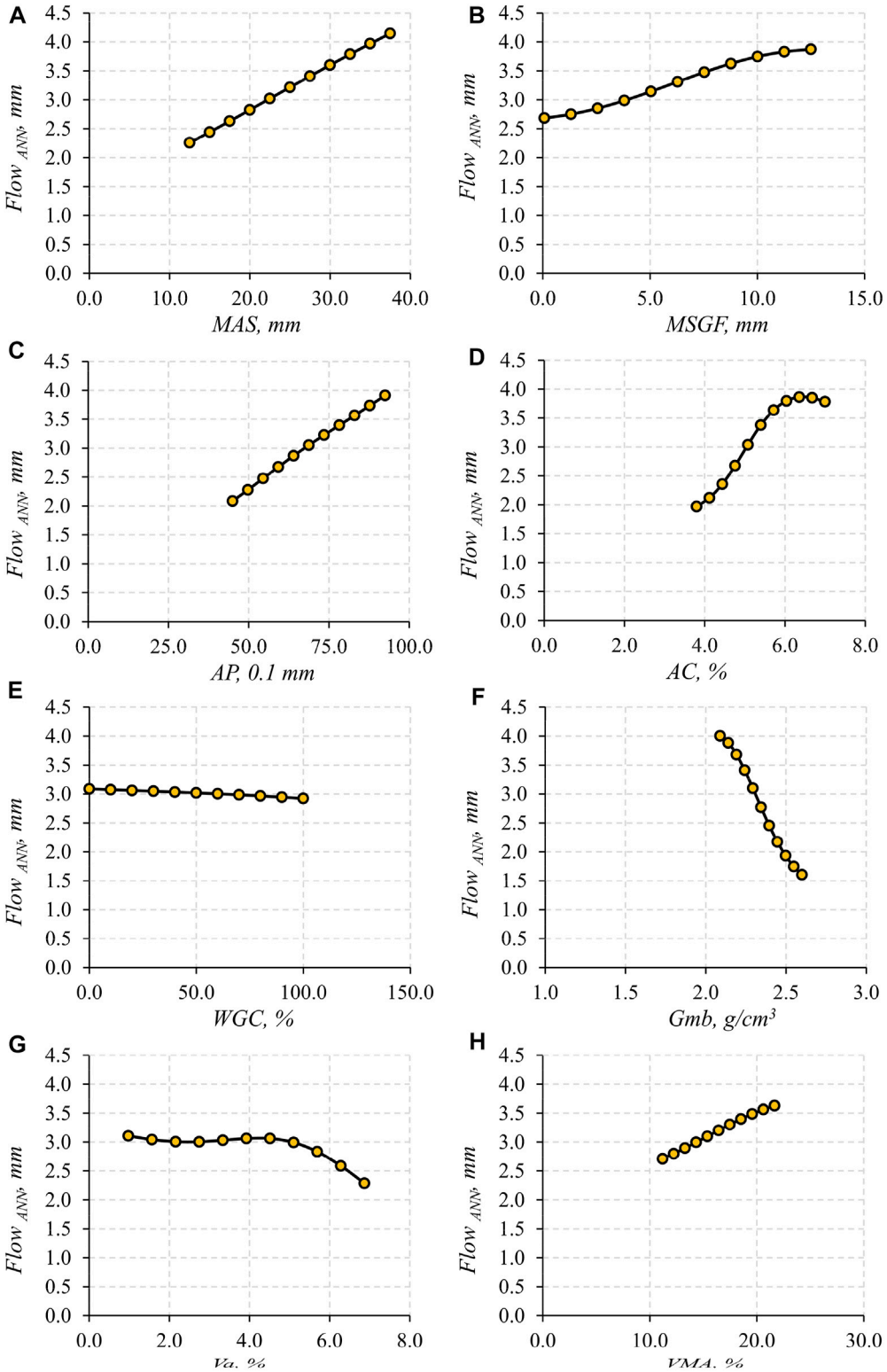


FIGURE 6 | Parametric study for glasphalt flow versus **(A)** Maximum aggregate size (MAS), **(B)** Maximum size of waste glass fractions (MSGF), **(C)** Average penetration grade of asphalt (AP), **(D)** Asphalt content (AC%), **(E)** Waste glass content (WGC%), **(F)** Bulk specific gravity of glasphalt mixture, **(G)** Air voids content (Va%), and **(H)** Voids in mineral aggregate (VMA%).

would have better stability than that with large MAS. The stability of mixes reaches its peak value at about 12.5 mm (1/2 inch) and then decreases gradually as the maximum size increases to about 25 mm (1.0 inch). No considerable change in the stability was observed when the MAS reaches 37.5 mm (1 ½ inch). **Figure 5B** indicates that glasphalt stability improved by increasing the maximum size of waste glass fractions (MSGF) to 4.75 mm (Sieve No. 4). Then, the stability maintains at a constant level for mixes incorporating glass fractions larger than 4.75 mm MSGF.

Figure 5C shows the relation between the average penetration grade of asphalts (AP) and glasphalt stability. The plot indicates that the stability increases as the AP increases, especially when the AP is higher than 60 (0.1 mm). This may be attributed to the low viscosity of the high penetration grade of asphalt cement, which provides proper coating and uniform dispersion of the angular glass cullet and aggregate particles that result in high stability. While the assumption that "soft" asphalt cement with a high penetration number is recommended to use in cold climate regions to resist thermal cracking (Brown et al., 2009), it can be argued that constructing glasphalt pavements with appropriate asphalt cement in cold regions would provide better stability than in warm regions.

As expected, in **Figures 5D and F**, the stability increases as the asphalt cement content (AC %) and bulk specific gravity of the glasphalt mixture (Gmb) increase to a certain limit and then decrease. This appreciable increase in the stability is attributed to the hot asphalt cement that initially lubricates the particles, enhances the interlock, and compacts them to a high density. The gradual decrease in the stability after that refers to the additional asphalt cement that produces thicker films around particles, pushing them apart and resulting in lower stability. It should be mentioned, however, that the considerable changes in the glasphalt density can be attributed to the waste glass content in the mixes, where the density of the glasphalt mixes reduces as the glass content increases (Al- Jameel and Al- Saeedi, 2018).

Figure 5E shows a gradual increase in the stability by increasing the waste glass content (WGC) up to 50%, and then the stability decreases as the WGC% achieves 100%. This behavior refers to the high angularity of the glass cullet compared to the crushed-stone aggregate, which increases the interlock between particles. After that the stability decreases gradually by increasing the WGC% because the crushed glass has less capability to bear the load (less abrasion) than the crushed aggregate in the mixes. It can be stated that a 50% replacement rate might be the ideal rate for incorporating waste crush glass as an aggregate substitute in glasphalt mixes.

Figure 5G shows the relation between the air-void content (Va %) and glasphalt stability. It can be seen that increasing the air voids up to 4% will keep the stability at a constant level. However, a gradual decrease is observed in the stability by increasing the air voids up to 7%. This agrees with the conventional design value of 4% air void content for a dense-graded asphalt mix design. Finally, the effect of VMA on glasphalt stability is shown in **Figure 5H**. The stability increases as the VMA increases. While the VMA represents the amount of space available for bitumen to properly coat each aggregate particle, sufficient asphalt film thickness around the glass and aggregate particles promotes good adhesion between them and increases the mix stability. It should be mentioned that the VMA refers to the total volume of air voids and effective (non-absorbed) asphalt cement between mineral particles of a compacted mix. Thus, the design should consider the minimum VMA values needed to achieve a durable asphalt film thickness around particles, which is determined based on the nominal maximum particle size and the design air voids of the mix (Brown et al., 2009; Asphalt Institute, 2014).

According to the flow results of the glasphalt mixes, **Figures 6A and B** indicate that the flow increases as the maximum size of used aggregate and glass cullet increases in the mix. This increase indicates that the larger the MAS and MSGF, the better the flexibility of the glasphalt mixes since the flow represents the total amount of vertical deformation of a test specimen. **Figures 6C** and 6H also show that the glasphalt flow increases linearly as the penetration grade of the used asphalt cement and the VMA% increases. This increase in the flow is expected since a higher penetration grade of asphalt cement indicates a lower viscosity of asphalt cement. Also, higher VMA% indicates a compacted mix with more spaces filled with non-absorbed asphalt cement and air voids, which provides better flexibility to the mixes. **Figure 6D** shows that the flow increases gradually as the AC% achieves 6%, and then no considerable change in the flow values was noticed.

Opposite to the stability trends, the flow values stood generally at a certain level as the waste glass content (WGC%) increases in the glasphalt mixes, as shown in **Figure 6E**. The flow values were also decreased gradually by increasing the bulk density of the glasphalt mixes (G_{mb}), as shown in **Figure 6F**. This decrease is attributed to the good compaction of glasphalt mixes, where the asphalt cement lubricates the mix ingredients and compacts them together to the dense state of high density. **Figure 6G** shows that the flow values maintain marginally at a constant level as the air-void percentage (Va%) increases to about 4%, which also agrees with the typical design values of air void content for the dense-graded asphaltic mixes.

CONCLUSION

As the awareness on environmental and natural resource issues increases globally, the waste glass cullet is used as an aggregate substitute in asphalt pavements. Asphalt mixture comprising waste crushed glass as an aggregate is referred to as "glasphalt". This study aims to investigate the feasibility of utilizing the ANN learning machine tool to predict new and effective models to predict the Marshall stability and Marshall flow of the dense-graded glasphalt mixes. The developed ANN models utilized a wide range of related Marshall test results collected from previous studies available in the literature. Eight independent variables were selected in this study as inputs for the models. They include the maximum aggregate size (MAS), the maximum size of waste glass fractions (MSGF), average penetration grade of the used asphalts cement (AP), asphalt cement content (AC%), waste glass content (WGC%), bulk specific gravity of the glasphalt mixture (G_{mb}), air void content (Va%), and voids in mineral aggregate (VMA%).

The proposed ANN models proved their high accuracy to predict the Marshall stability and Marshall flow of the dense glasphalt mixes. The results show an experimental-to-predicted ratio of 1.00 and 1.00, coefficient of variation (CoV) of 8.6% and 8.7%, RMSE of 1.63 kN and 0.54 mm, and R-squared of 93.6% and 85.7% for the glasphalt stability and flow models, respectively. The sensitivity of each variable on the Marshall stability and flow was investigated by conducting parametric analyses, and the following conclusions were obtained:

1) The desired maximum size of the used aggregate and waste crushed glass in glasphalt mixes is about 12.5 and 4.75 mm, respectively, where maximum stability values are achieved.
2) Using low viscosity asphalt cement would provide better coating and uniform dispersion of the used angular crushed glass and aggregate in the mix, resulting in high stability.
3) Glasphalt stability increases to a certain limit and then decreases gradually as the asphalt cement content and the bulk specific gravity increase for the mix.
4) The optimal content for incorporating waste crushed glass as an aggregate substitute in asphalt mixes is limited to 50%. This may refer to the high angularity and less bearing capability of the crushed glass compared to the natural aggregate, which limits mixing asphalt mixes with only waste crushed glass as an aggregate.
5) A 4.0% is the typical design content for the air voids in glasphalt mixes.
6) The stability increases as the VMA% increases. However, the minimum VMA% value that is required to assure a sufficient asphalt film thickness around the mix's ingredients should be considered while designing.
7) The flow increases as the maximum size of crushed glass and aggregate, the penetration grade of asphalt cement, asphalt cement content, and VMA% increase.
8) Varying the content of the waste glass in glasphalt mixes does not affect the flow values, whereas increasing the mix density reduces the flow values.

The developed ANN models are very important in the sense that they provide accurate, rational, general, and simple visualization for the performance of the dense glasphalt mixes in terms of Marshal stability and flow considering a wide range of different parameters. They are also helpful in saving time and effort to estimate the Marshall test properties of glasphalt mixes. It would be recommended to develop mathematical equations for estimating the performance of dense glasphalt mixes based on a wide range of design variables.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

Concept: YJ, AT, and YM. Data collection: YJ, RA, and YM. Analysis: AT, YM, YJ, and RA. Writing: YJ, RA, AT, and YM.

REFERENCES

Aashish, A., and Tamrakar, G. B. S. (2019). Effect of Adding Waste Crushed Glass to Asphalt Mix. *Int. J. Adv. Eng. Manag.* 4 (4), 10–18.

Al Bodour, W., Hanandeh, S., Hajij, M., and Murad, Y. (2022). Development of Evaluation Framework for the Unconfined Compressive Strength of Soils Based on the Fundamental Soil Parameters Using Gene Expression Programming and Deep Learning Methods. *J. Mater. Civ. Eng.* 34 (2), 04021452. doi:10.1061/(asce)mt.1943-5533.0004087

Al-Jameel, H. A. E., and Al-Saeedi, B. J. K. (2018). Sustainable Performance of Iraqi Asphalt Base Course Using Recycled Glass as Aggregate Replacement. *Jubes* 26 (5), 63–80. doi:10.29196/jub.v26i5.869

Alhassan, H. M., Yunusa, G. H., and Sanusi, D. (2018). Potential of Glass Cullet as Aggregate in Hot Mix Asphalt. *Nig. J. Tech.* 37 (2), 338–345. doi:10.4314/njt.v37i2.8

Almuhmdi, A., Muhmood, A., and Salih, A. (2021). Effects of Crushed Glass Waste as a Fine Aggregate on Properties of Hot Asphalt Mixture. *Tikrit J. Eng. Sci.* 28 (3), 129–145. doi:10.25130/tjes.28.3.10

Anochie-Boateng, J., and George, T. (2016). Use of Waste Crushed Glass for the Production of Hot-Mix Asphalt. *Sustain. Constr. Mater. Technol. (SCMT)*. doi:10.18552/2016/scmt4s198

Arabani, M. (2011). Effect of Glass Cullet on the Improvement of the Dynamic Behaviour of Asphalt Concrete. *Constr. Build. Mater.* 25 (3), 1181–1185. doi:10.1016/j.conbuildmat.2010.09.043

Arabani, M., and Kamboozia, N. (2013). The Linear Visco-Elastic Behaviour of Glasphalt Mixture under Dynamic Loading Conditions. *Constr. Build. Mater.* 41, 594–601. doi:10.1016/j.conbuildmat.2012.12.023

Arabani, M., Mirabdolazimi, S. M., and Ferdowsi, B. (2012). Modeling the Fatigue Behaviors of Glasphalt Mixtures. *Sci. Iran.* 19 (3), 341–345. doi:10.1016/j.scient.2012.02.021

Arabani, M., Tahami, S. A., and Taghipoor, M. (2017). Laboratory Investigation of Hot Mix Asphalt Containing Waste Materials. *Road Mater. Pavement Des.* 18 (3), 713–729. doi:10.1080/14680629.2016.1189349

Asphalt Institute (2014). *MS-2 Asphalt Mix Design Methods*. KY, United States: Asphalt Institute.

Azarhoosh, A. R., Hamedi, G. H., and Abandansari, H. F. (2018). Providing Laboratory Rutting Models for Modified Asphalt Mixes with Different Waste Materials. *Period. Polytech. Civ. Eng.* 62 (2), 308–317. doi:10.3311/ppci.10684

Brown, E. R., Kandhal, P. S., Roberts, F. L., Kim, Y. R., Lee, D. Y., and Kennedy, T. W. (2009). *Hot Mix Asphalt Materials, Mixture Design, and Construction*. MD, United States: NAPA research and education foundation.

Dalloul, K. N. (2013). *Study of the Effect of Crushed Waste Glass as Coarse Sand and Filler in the Asphalt Binder Course*. Master Thesis (Gaza: University of Gaza).

Disfani, M. M., Arulrajah, A., Bo, M. W., and Hankour, R. (2011). Recycled Crushed Glass in Road Work Applications. *Waste Manag.* 31 (11), 2341–2351. doi:10.1016/j.wasman.2011.07.003

Ektas, S., and Karacasu, M. (2012). Use of Recycled Concrete in Hot Mix Asphalt and an ANN Model for Prediction of Resilient Modulus. *Ekoloji Derg.* 21 (83). doi:10.5053/ekoloji.2012.836

Fadhil, T. H., Ahmed, T. M., and Al Mashhadany, Y. I. (2022). Application of Artificial Neural Networks as Design Tool for Hot Mix Asphalt. *Int. J. Pavement Res. Technol.* 15 (2), 269–283. doi:10.1007/s42947-021-00065-7

Finkle, I., and Ksaibati, K. (2007). "Recycled Glass Utilization in Highway Construction," in Proceedings of the Transportation Research Board 86th Annual Meeting (University of Wyoming, Laramie, WY, United States), MPC Report No. 07-192.

Flynn, L. (1993). 'Glasphalt'utilization Dependent on Availablity. *Roads Bridg.* 31 (2).

Garcia, J., and Hansen, K. (2001). *HMA Pavement Mix Type Selection Guide*. MD, United States: National Asphalt Pavement Association.

Iftikhar, B., Alih, S. C., Vafaei, M., Elkotb, M. A., Shutaywi, M., Javed, M. F., et al. (2022). Predictive Modeling of Compressive Strength of Sustainable Rice Husk Ash Concrete: Ensemble Learner Optimization and Comparison. *J. Clean. Prod.* 348, 131285. doi:10.1016/j.jclepro.2022.131285

Issa, Y. (2016). Effect of Adding Crushed Glass to Asphalt Mix. *Archives Civ. Eng.* 62 (2), 35–44. doi:10.1515/ace-2015-0063

Jony, H. H., Al-Rubaie, M., and Jahad, I. (2011). The Effect of Using Glass Powder Filler on Hot Asphalt Concrete Mixtures Properties. *Eng. Technol. J.* 29 (1), 44–57.

Kandahl, P. (1992). *Waste Materials in Hot Mix Asphalt*. Auburn University, United States: National Center for Asphalt Technology. NCAT Report 92–06.

Keskin, M., and Karacasu, M. (2021). Artificial Neural Network Modelling for Asphalt Concrete Samples with Boron Waste Modification. *J. Eng. Res.* doi:10.36909/jer.8124

Kifile, D., Quezon, E. T., and Tesfaye, A. (2020). Effect of Partial Replacement of Crushed Stone Dust Filler with Waste Glass Powder in Hot Mix Asphalt Concrete Production. *Xi'an Dianzi Keji Daxue Xuebao/Journal Xidian Univ.* 14 (9).

Malisch, W. R., Day, D. E., and Wixson, B. G. (1970). Use of Domestic Waste Glass as Aggregate in Bituminous Concrete. *Highw. Res. Rec.* 307, 1–10.

Mansourian, A., Ghanizadeh, A. R., and Golchin, B. (2018). Modeling of Resilient Modulus of Asphalt Concrete Containing Reclaimed Asphalt Pavement Using Feed-Forward and Generalized Regression Neural Networks. *J. Rehabilitation Civ. Eng.* 6 (1), 132–147. doi:10.22075/JRCE.2017.10876.1177

Momani, Y., Tarawneh, A., Alawadi, R., and Momani, Z. (2022). Shear Strength Prediction of Steel Fiber-Reinforced Concrete Beams without Stirrups. *Innov. Infrastruct. Solutions* 7 (1), 1–15. doi:10.1007/s41062-021-00703-w

Murad, Y., Tarawneh, A., Arar, F., Al-Zu'bi, A., Al-Ghwairi, A., Al-Jaafreh, A., et al. (2021). Flexural Strength Prediction for Concrete Beams Reinforced with FRP Bars Using Gene Expression Programming. *Structures* 33, 3163–3172. Elsevier. doi:10.1016/j.istruc.2021.06.045

Othman, K. (2022). Artificial Neural Network Models for the Estimation of the Optimum Asphalt Content of Asphalt Mixtures. *Int. J. Pavement Res. Technol.*, 1–13. doi:10.1007/s42947-022-00179-6

Ozgan, E. (2011). Artificial Neural Network Based Modelling of the Marshall Stability of Asphalt Concrete. *Expert Syst. Appl.* 38 (5), 6025–6030. doi:10.1016/j.eswa.2010.11.018

Ozturk, H. I., Saglik, A., Demir, B., and Gungor, A. G. (2016). "An Artificial Neural Network Base Prediction Model and Sensitivity Analysis for Marshall Mix Design," in Proceedings of the 6th Eurasphalt & Eurobitume Congress (Prague, Czech Republic. doi:10.14311/ee.2016.224

Ozturk, H. I., and Kutay, M. E. (2014). An Artificial Neural Network Model for Virtual Superpave Asphalt Mixture Design. *Int. J. Pavement Eng.* 15 (2), 151–162. doi:10.1080/10298436.2013.808341

Pasetto, M., Baldo, N., and Manthos, E. (2019). Asphalt Concrete Mechanical Behavior Prediction by Artificial Neural Networks*Bitum. Mix. Pavements VII*. CRC Press, 252–261. doi:10.1201/9781351063265-38

Rahman, M. T., Mohajerani, A., and Giustozzi, F. (2020). Recycling of Waste Materials for Asphalt Concrete and Bitumen: A Review. *Materials* 13 (7), 1495. doi:10.3390/ma13071495

Salem, Z. T. A., Khedawi, T. S., Baker, M. B., and Abendeh, R. (2017). Effect of Waste Glass on Properties of Asphalt Concrete Mixtures. *Jordan J. Civ. Eng.* 11 (1).

Saltan, M., Öksüz, B., and Uz, V. E. (2015). Use of Glass Waste as Mineral Filler in Hot Mix Asphalt. *Sci. Eng. Compos. Mater.* 22 (3), 271–277. doi:10.1515/secm-2013-0135

Saltan, M., Öksüz, B., and Uz, V. E. (2015). Use of Glass Waste as Mineral Filler in Hot Mix Asphalt. *Sci. Eng. Compos. Mater.* 22 (3), 271–277. doi:10.1515/secm-2013-0135

Shafabakhsh, G. H., Ani, O. J., and Talebsafa, M. (2015). Artificial Neural Network Modeling (ANN) for Predicting Rutting Performance of Nano-Modified Hot-Mix Asphalt Mixtures Containing Steel Slag Aggregates. *Constr. Build. Mater.* 85, 136–143. doi:10.1016/j.conbuildmat.2015.03.060

Singh, D., Zaman, M., and Commuri, S. (2013). Artificial Neural Network Modeling for Dynamic Modulus of Hot Mix Asphalt Using Aggregate Shape Properties. *J. Mat. Civ. Eng.* 25 (1), 54–62. doi:10.1061/(asce)mt.1943-5533.0000548

Su, N., and Chen, J. S. (2002). Engineering Properties of Asphalt Concrete Made with Recycled Glass. *Resour. conservation Recycl.* 35 (4), 259–274. doi:10.1016/s0921-3449(02)00007-1

Tapkın, S., Çevik, A., and Uşar, Ü. (2010). Prediction of Marshall Test Results for Polypropylene Modified Dense Bituminous Mixtures Using Neural Networks. *Expert Syst. Appl.* 37 (6), 4660–4670. doi:10.1016/j.eswa.2009.12.042

Viswanathan, K. (1996). *Characterization of Waste Recycled Glass as a Highway Material*. Doctoral dissertation. TX, United States: Texas Tech University.

Xiao, F., Amirkhanian, S., and Juang, C. H. (2009). Prediction of Fatigue Life of Rubberized Asphalt Concrete Mixtures Containing Reclaimed Asphalt Pavement Using Artificial Neural Networks. *J. Mat. Civ. Eng.* 21 (6), 253–261. doi:10.1061/(asce)0899-1561(2009)21:6(253)

Zavrtanik, N., Prosen, J., Tušar, M., and Turk, G. (2016). The Use of Artificial Neural Networks for Modeling Air Void Content in Aggregate Mixture. *Automation Constr.* 63, 155–161. doi:10.1016/j.autcon.2015.12.009

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor YM declared a past co-authorship with the author AT.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Built Environment

TYPE Original Research
PUBLISHED 23 August 2022
DOI 10.3389/fbuil.2022.991061

OPEN ACCESS

EDITED BY
Yasmin Murad,
The University of Jordan, Jordan

REVIEWED BY
Muhammad Rjoub,
Al-Balqa Applied University, Jordan
Eid Al-Sahawneh,
Al-Balqa Applied University, Jordan
Ahmad N. Tarawneh,
Hashemite University, Jordan

*CORRESPONDENCE
Mazen A. Musmar,
m.musmar@ju.edu.jo

SPECIALTY SECTION
This article was submitted to Computational Methods in Structural Engineering, a section of the journal Frontiers in Built Environment

RECEIVED 11 July 2022
ACCEPTED 25 July 2022
PUBLISHED 23 August 2022

CITATION
Musmar MA (2022), Structural performance of steel plates. *Front. Built Environ.* 8:991061. doi: 10.3389/fbuil.2022.991061

Structural performance of steel plates

Mazen A. Musmar*

Civil Engineering Department, The University of Jordan, Amman, Jordan

In general, plates are classified as thick plates when the minimum dimension to thickness ratio (b/h) is less than 10, thin plates when the b/h ratio ranges from 10 to about 100, provided that the plate maximum deflection to thickness ratio (w/h) is less than 0.2, and membranes when the b/h ratio approaches 100 and w/h $\geq$ 0.2. Thick plates develop internal stress resultants governed by three-dimensional elasticity similar to that of a solid body. Thin plates behave as plane stress members governed by two-dimensional elasticity. Membranes can only develop internal tensile stress within the plate's neutral plane. Few studies have adopted b/h ratios between 90 and 110 to investigate the feasibility of the utilization of such plates in their various available forms. The current study with the b/h ratio of 100 aims to fill the gap. Steel Plates are available in different forms such as intact plates, stiffened, perforated, and stiffened perforated plates. They are used in buildings, bridges, ships, as well as aerospace structures. In this study, the investigated steel plate has a square shape and is subjected to uniaxial loading. The plate edges are simply supported. The plate is 200 mm wide and 2 mm thick. In this case, the critical buckling strength is, in general, less than the plate maximum strength. With further loading, the plate would experience an undesirable sudden mode of failure owing to buckling instability. This study aims at investigating the performance of the different forms of square steel plates, such as intact, stiffened, perforated, and perforated stiffened, when the minimum dimension to thickness ratio is 100. A pushover finite element linear elastic buckling analysis as well as a nonlinear large deflection buckling analysis have been carried out. The study indicated that the increase in the plate maximum strength in single, double, and triple stiffener plates was 75.6%, 174%, and 196%, respectively, compared to that of the intact plate. Based on the obtained results, it is concluded that the utilization of plates having the b/h ratio of 100 is feasible provided that the appropriate plate form is adopted.

KEYWORDS

performance of plates, perforated plates, stiffened plates, steel plates, buckling

Introduction

Steel plates are commonly used in buildings, bridges, hydraulic structures, containers, ships, aerospace structures, and planes, as well as instruments and machines (Giovanni et al., 2014). They may be subjected to in plane loads or lateral loads or both.

Plates are generally classified as: 1) thick plates when the plate minimum dimension to thickness ratio (b/h) is less than 10. Thick plates develop internal load resultants, governed by three-dimensional elasticity to counterbalance the applied load. 2) Thin plates when b/h ranges from 10 to 100. Thin plates behave as plane stress members provided that the plate maximum deflection to plate thickness ratio (w/h) is less than 0.2. In this case, the plate develops internal load resultants, governed by two-dimensional elasticity to counterbalance the applied loading. 3) Membranes when the b/h ratio approaches 100 and w/h > 0.2. Membranes are only capable of developing internal tensile stress resultants, namely, membrane tensile force resultants acting within the plate middle plane (Yamaguchi and Wai-Fah, 1999; Ventsel and Krauthammer, 2001; Stephen et al., 2010).

Buckling instability is a mode of failure that the thin plate may experience under compression. This happens when the critical buckling load is less than the plate maximum strength. At low b/h values, strain hardening is generally attained without plate buckling. For medium b/h values, the plate imperfections as well as the plate residual stresses both give rise to inelastic buckling depicted as a transition curve. On the other hand, for large b/h values, the maximum plate strength exceeds the critical buckling load. In this case, the plate may experience elastic buckling, followed by nonlinear buckling when the incremental load increases. Buckling nonlinearity is caused by deformed thin plate geometry, geometric imperfections, residual steel stresses, and inelastic material behavior (Yamaguchi and Wai-Fah, 1999; Stephen et al., 2010). AISC (2017) provisions require preventing local plate buckling at any stress below the steel yield strength. On performing the pushover nonlinear buckling finite element (FE) analysis, subsequent to the critical buckling load, one of the following scenarios may take place: 1) the applied load remains constant whereas the lateral displacement upsurges, 2) the applied load declines however the lateral displacement upsurges, or 3) the applied load and the lateral displacement both upsurge, resulting in an additional cycle of buckling (Malhas et al., 2020). Post buckling response encompasses large displacement owing to the combined effect of both the geometric and the material nonlinearities due to the inelastic material behavior existing in the steel stress and strain constitutive relationship.

When plate buckling is experienced in a simply supported steel plate under a uniaxial compressive loading applied, membrane tensile stresses develop in the direction normal to the applied load. The tensile stresses are triggered by the stretching of the deformed neutral plane while the unloaded side edges are constrained against in plane translation. The developed tensile stresses hamper the lateral out of plane displacement.

Researches have been performed on buckling of intact plates as well as perforated plates in plane axial edge loading (Behzad et al., 2018). They indicated that in the case of large perforations, the critical buckling strength is generally higher than the plate maximum strength.

Aydin and Mustafa Sonmez (2015) stated that the perforations modify the plate buckling modal shape. Several researcher works (Ul-Nyeon et al., 2009; Aydin and Mustafa Sonmez, 2015; AISC, 2017; Behzad et al., 2018; Malhas et al., 2020) have indicated that cutouts reduce the elastic buckling load as well as the plate maximum strength. If large cutouts exist in the plate, the elastic buckling load becomes higher than the plate maximum strength. A plate with large openings may undergo either buckling or alternative yielding that may result in fractures (Saad-Eldeen et al., 2014). The local disturbance in the stress flow and the experienced peak stresses at the edges of the openings result in plate fracture when the steel's von Mises (σ_e) stresses at the edges of cutouts reach the steel ultimate stress (σ_u).

The plate maximum strength is influenced by its width to thickness ratio and the steel stress strain constitutive model, as well as the steel elastic properties such as the modulus of elasticity and yield strength. Several researcher works (Grondin et al., 1999; Gunay et al., 2013) have investigated the nonlinear buckling behavior of the stiffened plates, and they have concluded that the single sinusoidal half wave–shaped global buckling of an unstiffened intact plate converts to multiple local buckling in the form of half waves extending between every two nearby stiffeners. They also stated that the critical buckling stress is enhanced as the number of stiffeners is increased. Thus, in stiffened plate forms, instability might be experienced as local half waves existing between every two adjacent stiffeners, provided that the stiffeners possess relatively large flexural and torsional stiffnesses. Moreover, instability may be initiated as local buckling of the part of the plate spanning between the two adjacent stiffeners, followed by plate global buckling with further loading.

On the other hand, instability may be first experienced in stiffeners in the form of stiffener tripping, namely, stiffeners lateral torsional buckling owing to the relatively low values of torsional and flexural stiffnesses of the provided stiffeners.

A stiffened plate may be designed to experience local buckling for the parts of the plate spanning between each two adjacent stiffeners at a fraction of the failure load, followed by global buckling of the stiffened plate with further loading. Such design approach improves the critical buckling strength and is cost effective (Quin et al., 2009).

Plates are typified examples of plane stress members. In addition to the aforementioned factors, nonlinearity in the behavior of steel plates is also attributed to the low plate

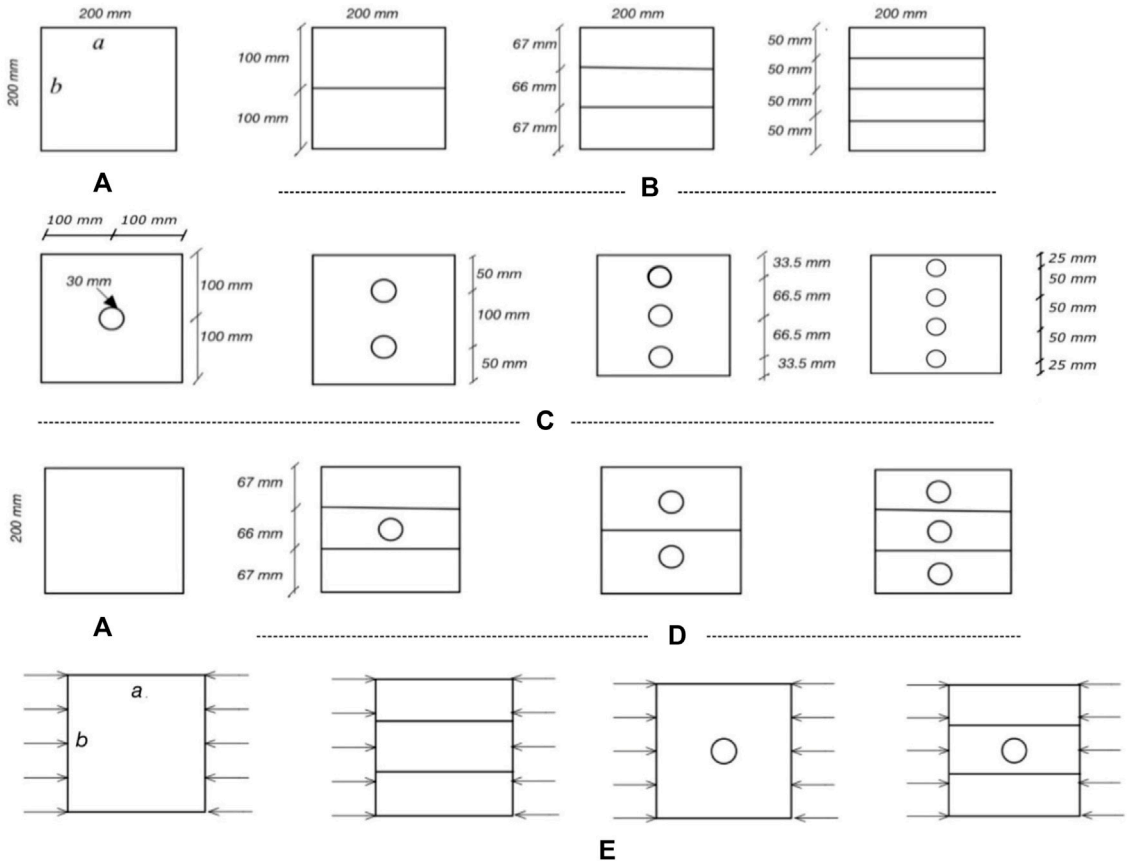


FIGURE 1
(A) Intact plates, **(B)** stiffened plates, **(C)** perforated plates, **(D)** stiffened perforated plates, **(E)** loading.

thickness, which is considerably low compared to the plate geometry. Such dimensions pave the way toward large deflection behavior. On the other hand, the steel behavior is linear elastic up to the steel yield stress. Subsequently, it is inelastic and involves both elastic and plastic strains. The plastic strains increase with further loading. Thus, nonlinearity is attributed to both geometry and material inelasticity.

In this study, the minimum plate dimension to thickness ratio is 100. Few studies have adopted b/h ratios between 90 and 110 to examine the performance and the feasibility of utilization of the plates in their various available forms such as intact, stiffened, perforated, and perforated stiffened under uniaxial uniform compression. The current study aims to fill the gap.

Problem statement

When the elastic buckling strength of the plate is less than the maximum plate strength, the plate in general experiences buckling failure. It is an undesirable mode of failure as it gives no warning. If the opposite is true, with further loading, nonlinear buckling may occur as the incrementally applied load approaches the plate maximum strength, which is defined as the highest load value in the load axial shortening curve. Nonlinearity is first caused by geometry as well as residual stresses and surface imperfections. When the developed internal steel's von Mises (σ_e) stresses reach the steel yield stress, the associated total strain becomes inelastic. They comprise both elastic and plastic strains. Nonlinearity is now attributed to the aforementioned factors as well as material plasticity. The plastic strains permit stress redistribution and result in the development of larger internal load resultants that counterbalance the incremental applied loading. Ultimately, failure occurs when the plate section is not capable anymore of developing the appropriate internal load resultants that would maintain equilibrium.

This study intends to investigate the performance, buckling behavior, and mode of failure of plates subjected to uniaxial edge loading at the two opposite edges. All plate forms have simply supported edges. The study has embraced the following cases, shown in Figure 1:

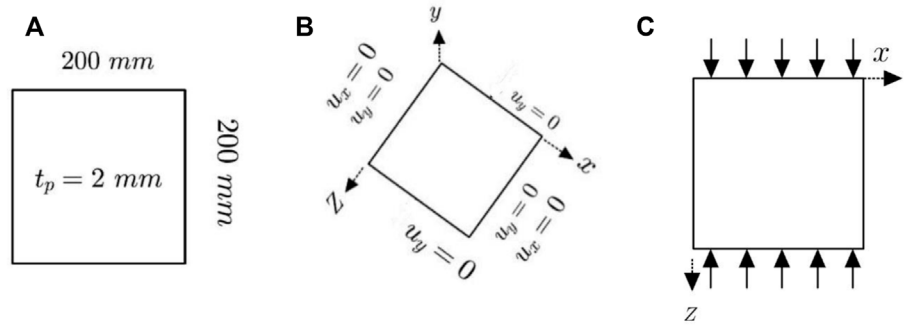


FIGURE 2
(A) Plate geometry, **(B)** plate boundary conditions, **(C)** plate loading.

- Square intact plates of 200 mm width (b) and thickness (tp) equals 2 mm.
- Stiffened plates; depth of stiffeners equals 20 mm and thickness of stiffeners (ts) is 4 mm. Thickness of stiffeners to thickness of plate (ts/tp) equals 2.
- Perforated plates; circular perforations of diameter (d) equals 30 mm.
- Perforated stiffened plates; circular perforations of diameter (d) = 30 mm, and stiffeners are identical in dimensions and shape to those in stiffened plates.

All the plate end edges are simply supported, shown in Figure 2

It is assumed that the plate and the stiffeners are steel and are rigidly attached together.

Analysis and numerical modeling:

The FE numerical analysis has been performed in two stages:

The initial stage was the pushover linear elastic buckling analysis. The second stage involved nonlinear post elastic buckling analysis due to nonlinearity in geometry, residual stresses, and surface imperfections, followed by material nonlinearity that has initiated when the steel's von Mises stresses (σ_e) exceeded the material yield stress (σ_y), as follows:

1) Initial stage: elastic buckling analysis.
2) In the second stage, the distorted shape from the initial stage was employed as the deformed geometry in the pushover nonlinear large deflection buckling analysis.

Modeling

Analysis has been performed using Ansys (2021). Shell 181 element has been used. It simulates both flexural and membrane behaviors. It is a four noded element. Each node possesses three translational as well as three rotational degrees of freedom. Modulus of elasticity of steel = 200 GPa, Poison ratio = 0.3, yield strength of steel = 250 MPa, and ultimate strength of steel = 460 Mpa. The maximum size of the shell element is b/20. In the regions close to perforations, the element max size is the least of b/20 or d/55, where d is the perforation size.

Validation

A validation analysis has been performed to confirm that the chosen element sizes are appropriate, and the adopted element type numerically simulates the behavior of the considered structural member. A sensitivity analysis for the mesh was performed to investigate the appropriateness of the size of mesh. Eventually the attained analysis results have been validated for the case of the intact plate in both elastic linear as well as nonlinear pushover buckling analyses. The recorded finite element results compared well with the results of the relevant mathematical expressions presented in literature. Subsequently, a buckling analysis has been performed for all considered plate forms subjected to uniaxial loading under simple span end conditions, shown in Figures 1, 2.

The obtained results for the elastic linear as well as the nonlinear buckling analyses for intact plates compare well with the closed form solutions as follows:

- Elastic buckling analysis

The recorded critical buckling load (P_{cr}) for the intact solid square plate using finite element simulation (P_{cr}) is 28.64 kN or N_{cr} = 143.2 N/mm.

It can be observed that it is relatively close to the critical buckling load N_{cr} of 144.61 N/mm for the solid square plate, obtained by Eq. 1 (Ventsel and Krauthammer, 2001).

$$D = \frac{Et^3}{12(1-\upsilon^2)} = \frac{2.0x10^5(2)^3}{12(1-0.3^2)} = 146520.146 \text{ mm}^4, \quad (1)$$

whereD: plate flexural rigidity;E: steel modulus of elasticity;υ: Poisson ratio.

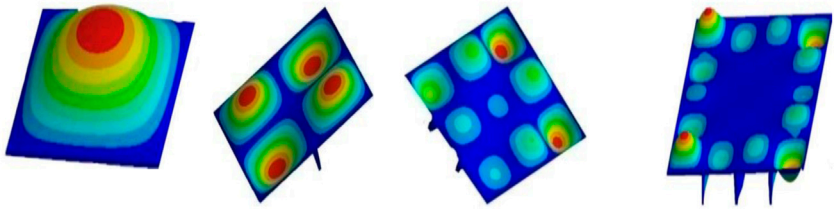


FIGURE 3
First modal vibration mode shapes for the intact and the stiffened square plates.

TABLE 1 Summary of analysis results for intact and stiffened plates.

Plate form	Critical buckling strength	Lateral displacement	Maximum strength P_{max}	% increase in maximum strength	End shortening at P_{max}	Out of plane deflection
	kN	mm	kN		mm	mm
Intact plate	28.64	1.50	47.43	0	0.3	3.78
One stiffener	-	-	83.3	75.6%	0.26	0.95
Two stiffeners	-	-	129.8	174%	0.22	0.53
Three stiffeners	-	-	140.2	196%	0.78	0.61

$$N_{cr} = 4\frac{\pi^2 D}{b^2} = 4\frac{\pi^2 * 146520.146}{200^2} = 144.61 N/mm, \quad (2)$$

where*Ncr*: critical buckling load;*b*: plate width.

Finite element analysis: P_{cr} = 28.64 *kN;*
N_{cr} = 28.64 *kN**1000.0/200.0 mm = 143.2 *N/mm*.

- Nonlinear buckling analysis

The intact square 200 mm plate maximum strength that has been determined by FE analysis is 47.43 kN. The calculated value according to Equation 4 suggested by Soares (1992) is 48.64 kN.

The difference is (47.43–48.64)/48.64 = −2.5%.

The results are acceptably close.

$$\lambda = \frac{b}{t}\sqrt{\frac{\sigma_y}{E}} = \frac{200}{2}\sqrt{\frac{250}{2.0E5}} = 3.53, \quad (3)$$

$$\frac{\sigma_u}{\sigma_y} = \frac{a_1}{\lambda} - \frac{a_2}{\lambda^2}, \qquad \lambda \geq 1, \quad (4)$$

$$a_1 = 2\,;\; a_2 = 1,$$

$$\frac{\sigma_u}{250} = \frac{2}{3.53} - \frac{1}{(3.53)^2} = 0.486,$$

$$\sigma_u = (250)(0.486) = 121.6\,\text{MPa},$$

$$P_{max} = (\sigma_u)(b)\left(t_p\right),$$

$$P_{max} = (121.6)\frac{(200)(2)}{1000} = 48.64\,kN.$$

Parametric study

This study investigates the performance of the following different available forms of steel plates under uniaxial compression. The intact plate smaller dimension to thickness ratio (b/h) is 100. The intact plate has been considered as the control specimen. The performance of the stiffened plates, perforated plates, and stiffened perforated plates had been compared with that of the control plate, and the results are shown as follows.

Stiffened plates

The following cases of stiffened plates have been investigated:

- Single stiffener stiffened plates.
- Two stiffeners stiffened plates.
- Three stiffeners stiffened plates.

A pushover elastic linear buckling analysis has been carried out for the aforementioned forms of square steel plates. The obtained mode shapes representing the first mode of vibration are illustrated in Figure3. It is obvious that the stiffeners converted the two global orthogonal identical half waves experienced by the intact plate into orthogonal multiple

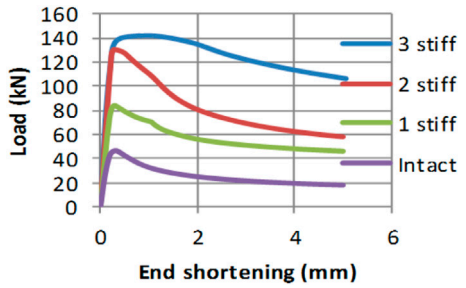


FIGURE 4
Load end shortening relationship of stiffened plates compared with the intact plate.

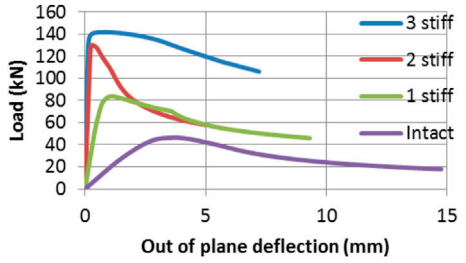


FIGURE 5
Load buckling relationship of stiffened plates compared with the intact plate.

identical half wave–shaped local buckling between nearby stiffeners. The space between the nearby stiffeners has controlled the length of the half waves in both directions as illustrated in Figure 3. If the plate is properly designed, the local buckling would result into global buckling for the whole stiffened plate. The analysis results indicate that the critical buckling load was 60.3% of the maximum plate strength in the intact plate. However, in all the stiffened plates, the critical buckling load was larger than the plate maximum strength as indicated in Table 1.

Figures 4, 5 show the end shortening as well as the lateral buckling versus load relationships, respectively, for all the aforementioned stiffened forms as well as the intact plate. Table 1 lists the recorded results. The maximum obtained load for the intact plate as well as the stiffened plates with single, double, and triple stiffeners having ratios of plate area to sum of stiffeners areas of 5.0, 2.50, and 1.67 were 47.43, 83.3, 129.8, and 140.2 kN, respectively. The increase in strength compared to that of the intact plate is 75.6%, 174%, and 196%.

The mere situation for stiffener tripping has been for single stiffener plate of Ap/As = 5.0, as shown in Figure 6. Stiffener tripping has been experienced when the summation of the values of the lateral torsional stiffness for all stiffeners was lower than the plate flexural stiffness value.

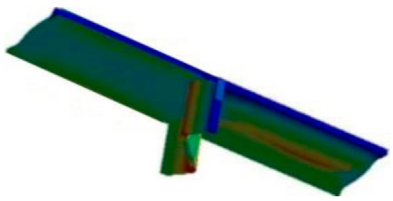


FIGURE 6
Tripping in a single stiffened plate.

Perforated plates

The following cases have been considered:

1) Single opening perforated plates.
2) Two openings perforated plates.
3) Three openings perforated plates.
4) Four openings perforated plates.

The mode shapes for the aforementioned forms of perforated plates are illustrated in Figure 7. The ratios of the sum of hole diameters to plate width for single, double, triple, and four openings are 0.150, 0.3, 0.45, and 0.60, respectively.

Table 2summarizes the critical buckling load for the intact solid plate as well as all other forms of perforated plates. The reduction in the critical buckling load for single, double, triple, and four perforations with respect to the intact plate is 5.50%, 11.1%, 14.%, and 43.1% respectively.

Figures 8, 9 indicate that the shapes of the load end shortening curves as well as the load buckling curves for all cases have kept an analogous form. Critical buckling strength was recorded as 0.58 of the plate maximum strength in single perforation plate and up to 0.61 of the plate maximum strength for triple perforations plate. In plates with four perforations, the critical buckling strength became 0.78 of the plate maximum strength.

The post elastic buckling stage has initiated, as nonlinear buckling due to nonlinearity in geometry, followed by material plasticity when steel's von Mises stresses (σ_e) have exceeded the material yield stress (σ_y) threshold as a direct result of excessive buckling. Kinematic instability was inevitable with further loading. However, the critical buckling strength was less than the plate maximum strength for all forms of perforated plates.

Stiffened and stiffened perforated plates

Figure 10 illustrates the modal shapes pertaining to the fundamental mode for the different forms of stiffened and stiffened perforated square plates. Figures 11, 12 illustrate the load end shortening as well as the load buckling curves for the

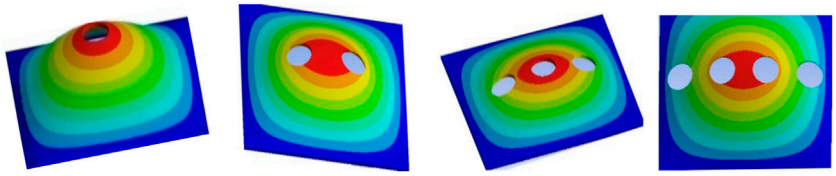


FIGURE 7
First modal vibration mode shapes for the perforated square plates.

TABLE 2 Summary of analysis results for intact and perforated plates.

Plate form	Critical buckling load	Out of plane deflection	Maximum strength Pmax	% decrease in maximum strength	End shortening at Pmax	Lateral deflection at Pmax
	kN	mm	kN		mm	mm
Intact plate	28.64	1.50	47.43	0	0.30	3.78
Single perforation	26.1	1.53	44.80	5.5%	0.30	3.79
Two perforations	25.8	1.54	42.18	11.1%	0.29	3.90
Three perforations	24.8	1.54	40.8	14%	0.25	3.38
Four perforations	20.9	1.56	27.0	43.1%	0.22	2.10

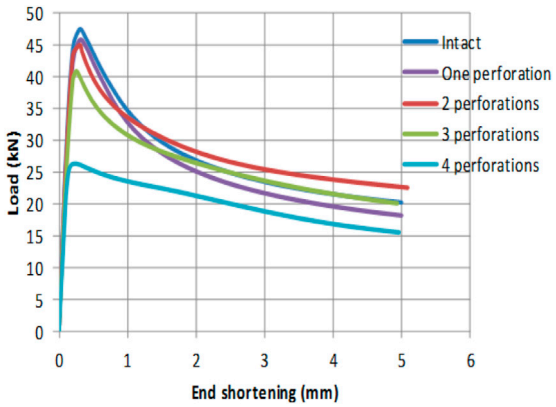


FIGURE 8
Load end shortening relationship of perforated plates compared with the intact plate.

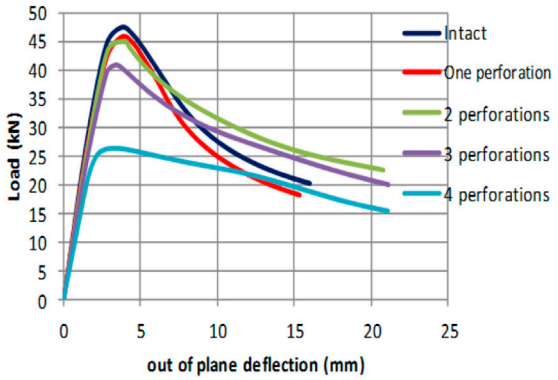


FIGURE 9
Load buckling relationship of perforated plates compared with the intact plate.

intact plate as well as single stiffener, single stiffener double perforations, double stiffeners, and double stiffeners with single perforation plate.

Table 3 summarizes the results for the stiffened and stiffened perforated plates. The analysis results indicate that the increase in plate maximum strength for the case of single stiffener, single stiffener with two perforations, double stiffeners, and double stiffeners with one perforation were 74.0%, 69.2%, 174%, and 133%, respectively, compared to that of the intact plate. Stiffener tripping has only been experienced in the single stiffener plate of Ap/As = 5. In all the remaining stiffened as well as stiffened perforated plate forms, the plate has not experienced any noticeable stiffener tripping. The effect of the two openings was dropping the amount of the plate flexural stiffness to values that were lower than the torsional stiffness of stiffeners. In such a case stiffener tripping can be avoided. The critical buckling loads for all the investigated forms of stiffened and perforated stiffened plates have been larger than the plate

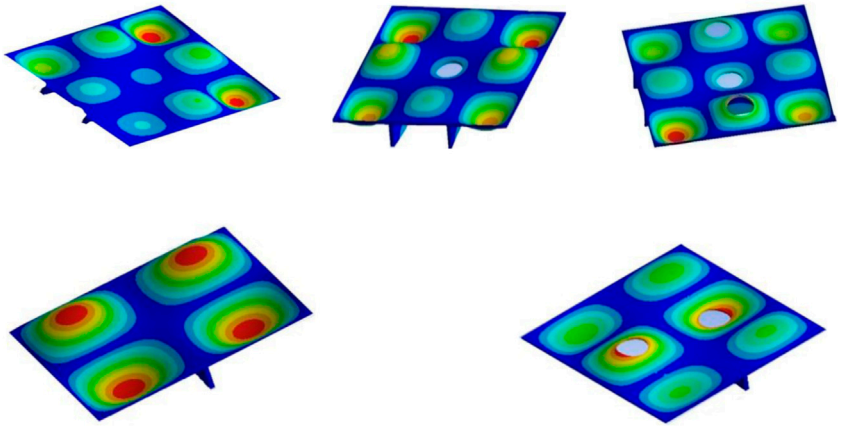


FIGURE 10
First modal vibration mode shape for perforated stiffened plates.

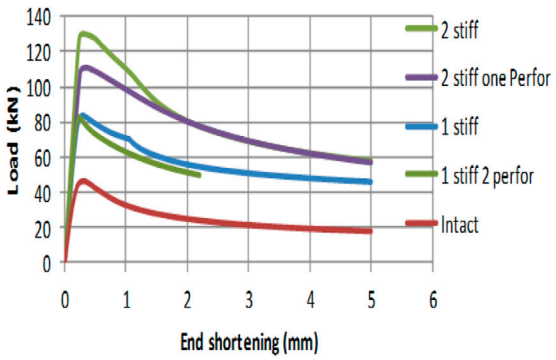


FIGURE 11
Load end shortening relationship for stiffened and stiffened perforated compared with the intact plates.

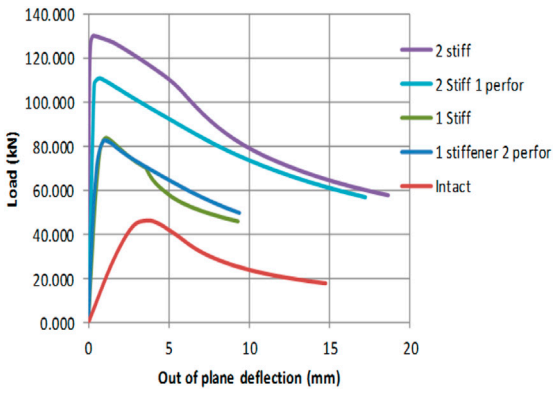


FIGURE 12
Load buckling relationship for stiffened and stiffened perforated compared with the intact plates.

maximum strength. All forms of the stiffened and stiffened perforated plates have not experienced buckling instability during the elastic linear stage as well as the post elastic nonlinear stage and up to failure. Initially, the plate response has been linear elastic. Upon surpassing the steel yield stress (σ_y) threshold, the plate has experienced inelastic deformation that resulted in stress redistribution, enabling the plate section to develop the appropriate internal loads that would counterbalance the incrementally applied load. Eventually, the developed internal load resultants were unable to counterbalance the incrementally applied load. A gradual failure has been experienced owing to geometric nonlinearity followed by material nonlinearilty that lead to the required ample warning. It is a desirable mode of failure compared to buckling instability that was experienced by unstiffened thin plate forms.

Results and discussion

Although in columns, the development of elastic buckling is an instability failure mode. This might not be the case in plate buckling because in general the plate has four supporting edges. Plate buckling develops in the two perpendicular directions. In general, the plate might experience kinematic based buckling instability failure. This may happen if the critical buckling strength is less than the maximum plate strength. However, if the opposite is true, the plate experiences strength-based failure when the developed internal force resultants are not capable anymore of counterbalancing the incrementally applied loading.

The current study has indicated that in the case of intact square plate having width to thickness ratio of 100, the critical

TABLE 3 Summary of analysis results for the stiffened and perforated stiffened plates.

Plate form	Critical buckling load	Out of plane deflection	Maximum strength	% increase in maximum strength	End shortening	Out of plane deflection
	kN	mm	kN		mm	mm
Intact plate	28.64	1.50	47.43	0	0.30	3.78
One stiffener	-	-	83.30	74%	0.27	0.96
Single stiffener double openings	-	-	80.3	69.2%	0.26	0.93
Double stiffeners	-	-	129.8	174%	0.21	0.53
Double stiffeners single opening	-	-	110.7	133%	0.48	0.84

buckling strength was 60.3% of the maximum strength of the plate. Elastic buckling has initiated while steel von Mises stresses (σ_{eq}) were less than σ_y. With further loading, more plate buckling was experienced. Boosted by geometric nonlinearity, residual stresses and imperfections in the geometry. Followed by material nonlinearity when steel von Mises stresses (σ_{eq}) exceeded σ_y. Both geometric and material nonlinearities raised the amount of plate buckling. Eventually, the resulting excessive buckling lead to kinematic instability. It is an undesirable mode of failure because it gives no warning.

In stiffened plates, the study has shown that stiffener tripping was merely experienced in single stiffener plate at (ts/tp) ratio of 2. In all the remaining cases, the large torsional stiffness values of stiffeners have obstructed stiffener tripping development.

The results also indicated that in all stiffened plates, the critical buckling strength was larger than the plate maximum strength. With further loading, the plate post buckling strength has initiated associated first with geometric nonlinearity, then with material nonlinearity when steel von Mises stresses (σ_{eq}) exceeded σ_y. Both geometric and material nonlinearities raised the amount of plate buckling. With further load, the amount of internally developed load resultant within the section was insufficient to balance the applied load. Steel plasticity lead to gradual failure with ample warning.

Perforations decreased the critical buckling strength as well as the plate maximum strength and lead to the development of larger stresses that surpassed σ_y owing to the plate reduced section. The steel plasticity enhanced the plate ductility, however the plate maximum strength was reduced.

The stiffened perforated plates lead to an increase in both plate maximum strength and ductility.

The adopted numerical analysis has been FE modeling using Ansys software (Ansys, 2021).

Conclusion

This study involves investigating the performance of different forms of simply supported square plates with a minimum dimension to thickness ratio of 100. At this slenderness ratio the intact plate flexural stiffness is relatively low. Based on the results and the pertaining discussion, it is concluded that:

1) When an intact plate with a relatively high slenderness value, is under uniaxial load, the critical buckling load is generally below the plate maximum strength. When the incremental load reaches the critical buckling load, buckling occurs in the form of one-half sinusoidal wave along the two orthogonal plate directions. With further loading and owing to the excessive lateral deformation, the plate experiences kinematically based buckling instability. It is sudden and gives no warning. It is an undesirable mode of failure and does not permit to fully exploit the full plate strength capacity.
2) Using another plate form, such as the stiffened plate modifies the buckling shape and increases the plate slenderness ratio to a level where the critical buckling load exceeds the plate maximum strength capacity. In this case, and depending upon the number of provided stiffeners, the failure mode changes from the undesirable sudden buckling instability to gradual failure that involves post buckling strength and gives the appropriate pre-failure warning owing to geometric nonlinearity and steel yielding that takes place prior to failure.
3) The perforations disturb the internal stress path and raises the peak values of the developed internal stresses within the reduced section. When the steel von Mises stresses surpass σ_y, steel plasticity enhances the ductility prior to failure. However Fracture may occur at the edge of the openings when steel von Mises stresses reach the ultimate stress (σ_u).
4) The perforated stiffened plate form if properly designed in terms of the number of stiffeners. The ratio of the sum of the diameters of perforations to the plate width, would alter the mode of failure from kinematic based buckling instability sudden failure to gradual failure involving geometric nonlinearity and material plasticity that pave the way for stress redistribution. Failure occurs when the developed

internal load resultants are not able anymore to counterbalance any further applied loading.

5) Stiffener tripping must be checked in stiffened and stiffened perforated plates. Both forms of square plates might experience stiffener tripping when the torsional stiffness of a stiffener becomes less than the plate flexural stiffness.
6) Using the available different forms of plates enables the effective utilization of thin stiff plates of plate minimum dimension to depth ratio of 100.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Acknowledgments

The author is pleased to acknowledge the support of The University of Jordan.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

AISC (2017). *Steel construction manual*. 15th edition. Illinois, United States: American Institute of Steel Construction.

Ansys (2021). *Ansys 2021 R1*. Canonsburg, US: Inc. Academic version.

Aydin, Komur M., and Mustafa Sonmez (2015). Elastic buckling behavior of rectangular plates with holes subjected to partial edge loading. *J. Constr. Steel Res.* 112, 54–60. doi:10.1016/j.jcsr.2015.04.020

Behzad, M., Eunsoo, C., and Woo, J. (2018). Compressive investigation of buckling behavior of plates considering the effect of holes. *Struct. Eng. Mech.* 68, 261–275. doi:10.12989/sem.2018.68.2.261

Giovanni, G., Antonio, M., Giuseppe, Z., and Casciaro, R. (2014). A Numerical Asymptotic Formulation for the post buckling analysis of structures in case of coupled instability. *Romanian J. Tech. Analysis* 59 (1-2), 38–55.

Grondin, G. Y., Elwi, A. E., and Cheng, J. J. R. (1999). Buckling of stiffened steel plates/a parametric study. *J. Constr. Steel Res.* 50 (2), 151–175. doi:10.1016/s0143-974x(98)00242-9

Gunay, E., Aygun, C., and Yildiz, Y. (2013). Nonlinear buckling finite element analysis of stiffened steel plates. *Adv. Mat. Res.* 669, 450–456. doi:10.4028/www.scientific.net/amr.699.450

Malhas, A., Johnson, E., and Salmon, G. (2020). *Steel structures design and behavior*. London, United Kingdom: Pearson.

Quin, D., Murphy, A., McEwan, W., and Lemaitre, F. (2009). Stiffened panel stability behavior and performance gains with plate prismatic sub-stiffening. *Thin-walled Struct.* 47 (12), 1457–1468. doi:10.1016/j.tws.2009.07.004

Saad-Eldeen, S., Garbatov, Y., and Soares, G. (2014). *Ultimate strength assessment of steel plates with a large opening*. London: Taylor and Francis Group. ISBN 978-1-138-00124-4.

Soares, G. C. (1992). Design equation for ship plate elements under uniaxial compression. *J. Constr. Steel Res.* 22, 99–114. doi:10.1016/0143-974x(92)90010-c

Stephen, P., Timoshenko, S., and Woinowsky, k. (2010). *Theory of plates and shells*. New York, United States: McGraw-Hill.

Ul-Nyeon, K., Ick-Heung, C., and Jeom-Kee, P. (2009). Buckling and ultimate strength of perforated plate panels subject to axial compression. *Taylor Francis* 4 (14), 337–361.doi:10.1080/17445300902990606

Ventsel, E., and Krauthammer, T. (2001). *Thin plates and shells theory analysis and applications*. New York, United States: Marcel Dekker.

Yamaguchi, E. (1999). in *Basic theory of plates and elastic stability" structural engineering handbook*. Editor Chen Wai-Fah (Boca Raton: CRC Press LLC).

frontiers | Frontiers in Built Environment

TYPE Original Research
PUBLISHED 23 August 2022
DOI 10.3389/fbuil.2022.899072

OPEN ACCESS

EDITED BY
Iftikhar Azim,
Shanghai Jiao Tong University, China

REVIEWED BY
Gennaro Senatore,
Swiss Federal Institute of Technology Lausanne, Switzerland
Ge Ou,
The University of Utah, United States

*CORRESPONDENCE
Chi-tathon Kupwiwat,
kupwiwat.chitathon.73c@st.kyoto-u.ac.jp

SPECIALTY SECTION
This article was submitted to Computational Methods in Structural Engineering, a section of the journal Frontiers in Built Environment

RECEIVED 18 March 2022
ACCEPTED 11 July 2022
PUBLISHED 23 August 2022

CITATION
Kupwiwat C-t, Hayashi K and Ohsaki M (2022), Deep deterministic policy gradient and graph convolutional network for bracing direction optimization of grid shells. *Front. Built Environ.* 8:899072. doi: 10.3389/fbuil.2022.899072

Deep deterministic policy gradient and graph convolutional network for bracing direction optimization of grid shells

Chi-tathon Kupwiwat*, Kazuki Hayashi and Makoto Ohsaki

Department of Architecture and Architectural Engineering, Graduate School of Engineering, Kyoto University, Kyoto, Japan

In this paper, we propose a method for bracing direction optimization of grid shells using a Deep Deterministic Policy Gradient (DDPG) and Graph Convolutional Network (GCN). DDPG allows simultaneous adjustment of variables during the optimization process, and GCN allows the DDPG agent to receive data representing the whole structure to determine its actions. The structure is interpreted as a graph where nodes, element properties, and internal forces are represented by the node feature matrix, adjacency matrices, and weighted adjacency matrices. DDPG agent is trained to optimize the bracing directions. The trained agent can find sub-optimal solutions with moderately small computational cost compared to the genetic algorithm. The trained agent can also be applied to structures with different sizes and boundary conditions without retraining. Therefore, when various types of braced grid shells have to be considered in the design process, the proposed method can significantly reduce computational cost for structural analysis.

KEYWORDS

bracing direction optimization, reinforcement learning, deep deterministic policy gradient, graph convolutional network, grid shell structures

1 Introduction

Structural optimization aims to obtain the best design variables that minimize/maximize an objective function under specified constraints (Christensen and Klarbring, 2009). For discrete structures, such as trusses and frames, typically, the design variables are cross-sectional properties, nodal locations and/or nodal connectivity (Ohsaki and Swan, 2002). Finding the best nodal locations is generally called geometry optimization, and the determination of nodal connectivity is called topology optimization. Structural optimization is important in early-stage design of large-span grid shells because their structural performance depends significantly on the shape and topology (Ohsaki, 2010). An optimization problem for grid shells can be formulated to maximize the stiffness against static loads through minimization of the compliance (i.e., elastic strain energy). Examples of such

formulation can be found in Refs. (Topping, 1983; Wang et al., 2002; Kociecki and Adeli, 2015).

In topology optimization of grid shells where bracing directions are to be optimized, the optimization problem can be formulated as a combinatorial problem and solved using heuristic approaches such as genetic algorithm (GA) and simulated annealing without utilizing gradient information (Dhingra and Bennage, 1995; Ohsaki, 1995; Kawamura et al., 2002). While this approach allows for simple implementation, it requires many evaluations of the structural response and therefore has a high computational cost especially for structures made of many elements (i.e., many design variables). In addition, the topology optimization problem to minimize compliance can be formulated as mixed-integer programming (MIP) which is practical for small- to medium-size optimization problems due to computational cost (Kanno and Fujita, 2018). Recent advances in MINLP enable to solve very large mixed-integer problems with quadratic and/or bilinear objective function and constraints. However, both heuristic and mathematical programming approaches do not allow the use of knowledge acquired from previously obtained solutions for similar structural configurations.

In recent years, machine learning (ML) approaches have been applied to structural optimization problems. ML can be classified into supervised learning, unsupervised learning, and reinforcement learning (RL). A supervised learning model learns to map (predict or classify) given input instances to specific output domains using sample data for training. Examples of this method for structural optimization can be found in (Berke et al., 1993; Hung et al., 2019; Mai et al., 2021). An unsupervised learning model learns to capture relationships between instances (data). Examples of unsupervised learning are t-distributed stochastic neighbor embedding (t-SNE) (van der Maaten and Hinton, 2008) and k-means clustering (MacQueen, 1967). Jeong and Yoshimura (Jeong and Yoshimura, 2002) also applied an improved unsupervised learning method to multi-objective optimization of plane trusses. Applications of unsupervised learning methods for structural design and structural damage detection can be found in (Eltouny and Liang, 2021; Puentes et al., 2021).

RL is a type of ML that has been developed from optimal control and dynamic programming (Sutton and Andrew, 2018). In RL, a model, or agent, is allowed to interact with an environment. The agent adjusts its policy to take actions according to given reward signals, which are designed to encourage the agent to do actions that change the environment into a desirable state such as winning a game or obtaining solutions to problems. RL has been successfully applied to various problems such as playing arcade games (Mnih et al., 2013) and controlling vehicles (Yu et al., 2019).

Deep Deterministic Policy Gradient (DDPG) (Lillicrap et al., 2016) is a type of RL algorithm that uses two neural networks (NN) (Rosenblatt, 1958; Ivakhnenko, 1968; Goodfellow et al., 2016) as an agent. The DDPG can be used in an environment where multiple agent actions are needed. Kupwiwat and Yamamoto (Kupwiwat and Yamamoto, 2020) studied various RL algorithms, using NNs as agents, and found that DDPG can be effectively applied to the geometry optimization problem of grid shell structures. However, the agents can only observe a constant number of inputs. Therefore, when structural dimensions are changed, it is difficult for the agents to detect the change of structural characteristics.

Hayashi and Ohsaki (Hayashi and Ohsaki, 2021) proposed a combined method of RL and graph representation for binary topology optimization of planar trusses. Graph representation allows an RL agent to observe the whole structure by transforming the structure into graph data consisting of nodes (vertices) and elements (edges) and implementing repetitive graph embedding operations to transmit signals of adjacent nodes and elements for estimating accumulated rewards associated with each action. Zhu et al. (2021) studied the applicability of RL and graph representation for stochastic topology generation of stable trusses which can be further used as initial structures for other topology optimization algorithms.

Graph neural networks are types of NNs, specifically designed for working with graph data. Graph Convolutional Network (GCN) (Kipf and Welling, 2017) is a class of graph neural networks that uses a convolution operator to process graph signals to the output domain. GCN has been successfully applied to problems such as node classification (Kipf and Welling, 2017) and link prediction (Kipf and Welling, 2016).

This paper proposes methods for the bracing directions optimization of grid shell structures for minimizing the strain energy using DDPG and GCN. The RL agent is trained to optimize the bracing directions from initial randomly generated directions. The proposed method is considered a part of the early-stage design of grid shells. The method takes an input the shape of the grid-shell which must be pre-determined. Bracing direction optimization can reduce the structure's strain energy without affecting the appearance of the shape because the braces are typically covered by finishing or ceiling. This paper is organized as follows: Section 2 gives the optimization formulations. Sections 3, 4 introduce existing approaches of GCN and a type of RL named DDPG, respectively. Section 5 explains the novelty of this research consisting of the vectors and matrices utilized in RL and the formulation of the Markov decision process for training the RL. Numerical examples are presented in Section 6 to benchmark the proposed method against the enumeration method and the genetic algorithm in terms of structural performance and computational cost.

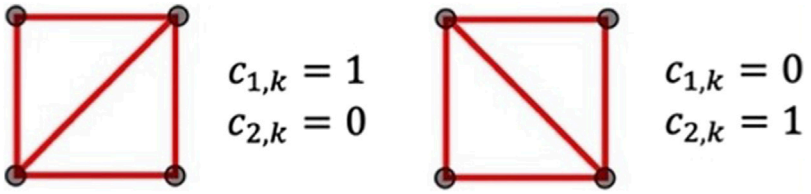


FIGURE 1
Bracing directions in the grid cell k and corresponding values of $c_{1,k}$ and $c_{2,k}$

2 Optimization problems of braced grid shells

2.1 Objective function: Total strain energy

The total strain energy of the structure subjected to static loads is chosen as the objective function to be minimized. The main grid elements are modeled using 3-dimensional beam elements with 12 degrees of freedom (DoFs), whereas the bracing elements are modeled using 3-dimensional truss elements with 6-DoFs. In the local coordinate system, the stiffness matrices of the frame element and the bracing element are denoted as $\mathbf{k}_f \in \mathbb{R}^{12\times12}$ and $\mathbf{k}_e \in \mathbb{R}^{6\times6}$, respectively. These matrices are converted into those with respect to the global coordinate system, and assembled into the global stiffness matrix $\mathbf{K} \in \mathbb{R}^{n_D\times n_D}$, where n_D is the number of DoFs of the structure after assigning the boundary conditions. Every node in the structure is subjected to a point load and every element is subjected to self-weight per unit length. These loads and weights are collected into the load vector $\mathbf{p} \in \mathbb{R}^{n_D}$ with respect to the global coordinates, and the nodal displacement vector $\mathbf{d} \in \mathbb{R}^{n_D}$ is obtained by solving the equilibrium and geometric compatibility equations:

$$\mathbf{Kd} = \mathbf{p} \tag{1}$$

Then, the total strain energy E is computed from

$$E = (1/2)\cdot \mathbf{d}^T\mathbf{Kd} \tag{2}$$

where the superscript T indicates the transpose of a vector or a matrix.

2.2 Bracing direction optimization problem

The combinatorial problem of bracing directions optimization is used to investigate the performance of the proposed method in terms of solution quality and the computational cost. Nodal coordinates and element cross-section size are not included in the variables. Given a grid shell with n_x by n_y square grids and diagonal bracing in each grid cell $k \in \{1, \ldots, n_x n_y\}$, there can be two possible directions for bracing indicated by $c_{1,k}, c_{2,k} \in \{0, 1\}$ which correspond to absence and presence of the brace in each direction as illustrated in Figure 1.

Since only one brace should exist in the grid cell k, the summation $c_{1,k} + c_{2,k}$ is always 1. Let $\mathbf{c}_1 \in \mathbb{R}^{n_x n_y}$ and $\mathbf{c}_2 \in \mathbb{R}^{n_x n_y}$, respectively, be vectors consisting of $c_{1,k}$ and $c_{2,k}$ for all bracing elements. The global stiffness matrix of the structure is a function of $\mathbf{c}_1$ and $\mathbf{c}_2$ denoted as $\mathbf{K}(\mathbf{c}_1, \mathbf{c}_2)$. The bracing direction optimization problem to minimize the strain energy is formulated as follows:

$$\text{minimize} \quad E(\mathbf{c}_1, \mathbf{c}_2) = (1/2)\cdot \mathbf{d}^T\mathbf{K}(\mathbf{c}_1, \mathbf{c}_2)\mathbf{d} \tag{3a}$$

$$\text{subject to} \quad c_{1,k}, c_{2,k} \in \{0, 1\}, \left(k = 1, 2, \ldots, n_x n_y\right) \tag{3b}$$

$$c_{1,k} + c_{2,k} = 1, \quad \left(k = 1, 2, \ldots, n_x n_y\right) \tag{3c}$$

where $\mathbf{d}$ is an implicit function of $\mathbf{c}_1$ and $\mathbf{c}_2$ obtained by solving Eq. 1.

3 Graph Convolutional Network

Consider a graph consisting of n nodes with g features in each node. The graph data can be represented using a node feature matrix $\mathbf{N} \in \mathbb{R}^{n\times g}$ representing features of each node in the graph, an adjacency matrix $\mathbf{M} \in \mathbb{R}^{n\times n}$ representing the connectivity of structural elements, a weighted adjacency matrix $\mathbf{P} \in \mathbb{R}^{n\times n}$ representing the connectivity of structural elements weighted by element forces, and a degree matrix $\mathbf{D} \in \mathbb{R}^{n\times n}$ representing the number of connections of each node in the graph. Kipf and Welling (Kipf and Welling, 2017) proposed a GCN that can process graph input data which are chosen from $\{\mathbf{N}, \mathbf{M}, \mathbf{P}, \mathbf{D}\}$, and map the inputs to the target domain of the graph for tasks such as node classification or link (connection) prediction (i.e., perform supervise learning by comparing mapped target domain from GCN to the node classification training data). A single GCN computation can be considered as a *layer*. Multiple GCN *layers* can be connected together to create a computational model for an RL agent. A GCN *layer* consists of a normalized form of adjacency matrix $\tilde{\mathbf{M}}$, and converts the input instances $\mathbf{N}$ to the output $\mathbf{O} \in \mathbb{R}^{n\times h}$ that has h embedding spaces (i.e., output for the GCN

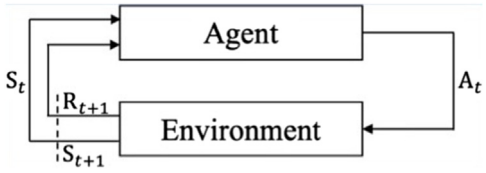


FIGURE 2
Diagram of the MDP

or output for a GCN layer that will be treated as **N** for the next GCN layer). The GCN layer can be formulated as follows:

$$\mathbf{O} = \sigma\left(\tilde{\mathbf{M}}\mathbf{N}\mathbf{w}\right) \tag{4}$$

where σ is a non-linear activation function, and $\mathbf{w} \in \mathbb{R}^{g\times h}$ is the weight matrix (i.e., the convolution filter parameters) (Kipf and Welling, 2017) in the GCN layer which is adjusted during the training. The convolutional filter parameters **w** weight $\tilde{\mathbf{M}}\mathbf{N}$, an aggregated signal of a node and its neighboring nodes. It should be noted that **w** can handle any graph with any number of nodes as long as the nodes have the same number of features.

The normalized adjacency matrix $\tilde{\mathbf{M}}$ can be computed using the following equation:

$$\tilde{\mathbf{M}} = \mathbf{D}^{-1/2}[\mathbf{M} + \mathbf{I}]\mathbf{D}^{-1/2} \tag{5}$$

where $\mathbf{I} \in \mathbb{R}^{n\times n}$ is the identity matrix, and $\mathbf{D}^{-1/2}$ is the inverse of the matrix $\mathbf{D}^{1/2}$ satisfying. $\mathbf{D}^{1/2}\mathbf{D}^{1/2} = \mathbf{D}$

In this paper, we utilize GCN to build an RL agent. The original GCN does not utilize the weighted adjacency matrix. However, for the structural optimization problem, the internal forces (i.e., the forces taken by the structural elements) should be utilized to guide the actions of the RL agent. Therefore, we propose a novel GCN-DDPG architecture that employs weighted adjacency matrices constructed from the internal forces for solving the bracing direction optimization problem. Details of the formulation are given in Section 5.

4 Reinforcement learning and Deep Deterministic Policy Gradient

RL is a type of ML that trains an agent to perform actions in an environment using reward signals. An RL algorithm consists of three main elements: a *policy* that determines the agent behavior, a *reward signal* that defines how good/bad the agent behavior is, according to the policy, and a *value function* that predicts how the agent performs based on the policy (Sutton and Andrew, 2018). The interaction of an agent and the environment is formulated using a Markov Decision Process (MDP) (Bellman, 1954; Bellman, 1957) as follows:

In a discrete step t:
The agent receives a representation of the environment as state S_t.
The agent performs actions A_t.
The agent receives quantitative reward R_{t+1} and next state S_{t+1} from the environment.

The diagram of the MDP can be represented as shown in Figure 2.

DDPG (Lillicrap et al., 2016) is a type of RL policy gradient algorithm that utilizes a parameterized policy function (*Actor*) π_{θ_1} to determine the probability of taking action A_t^i in a state S_t, denoted by $P(A_t^i|S_t)$, and another parameterized value function Q_{θ_2} (*Critic*) to predict the accumulated reward (Q-Value) from the actions of the agent follows:

$$\pi_{\theta_1}(S_t) = P(A_t^i|S_t) \tag{6}$$

$$Q_{\theta_2}(S_t, \pi_{\theta_1}(S_t)) = \sum_{v=1}^{\infty}\gamma^{v-1}R_{t+v} \tag{7}$$

where θ_1 and θ_2 are parameters of policy and value functions to be adjusted during the training, respectively. $\gamma \in [0, 1)$ is a discount factor for the reward.

During training, the value function adjusts its parameter θ_2 to increase the accuracy of its prediction of accumulated reward using a *replay buffer* that stores data of $\{S_t, A_t, R_{t+1}, S_{t+1}\}$, whereas the policy function adjusts its parameter θ_1 to increase the value predicted by the value function, which is equivalent to the obtained rewards, using the gradient ∇J_{θ_1} as described in the following DDPG algorithm. Since using the online policy and value functions will make the learning unstable, Haarnoja et al. (Haarnoja et al., 2018) proposed a *tau update* method that trains a surrogate policy function $\pi_{\theta_1'}$ and a surrogate value function $\mathbf{Q}_{\theta_2'}'$, and then gradually updates the parameters of these functions into the online functions using a small value of τ ($\tau \ll 1$) at every *tau update* interval. Note that the agent interacts with the environment to collect data for the replay buffer using the online policy function.

Let $\mathcal{L}(y, \hat{y})$ be a loss function between y which represents the correct value of training data and a predicted value $\hat{y}$. The training algorithm of DDPG is as follows:

DDPG algorithm:
1. Sample u training data $\{S_t, A_t, R_{t+1}, S_{t+1}\}$ from the replay buffer and convert them into a set of vectors $\{\mathbf{S}_t, \mathbf{A}_t, \mathbf{R}_{t+1}, \mathbf{S}_{t+1}\}$.
2. Update the parameters as follows:
$\pi_{\theta_1'}'(\mathbf{S}_t) = \hat{\mathbf{A}}_t$ *# Surrogate policy function* $\pi_{\theta_1'}'$ *decides actions from* $\mathbf{S}_t$
$\pi_{\theta_1}(\mathbf{S}_{t+1}) = \hat{\mathbf{A}}_{t+1}$ *# Online policy function* π_{θ_1} *decides actions from* $\mathbf{S}_{t+1}$
$Q_{\theta_2'}'(\mathbf{S}_t, \mathbf{A}_t) = \hat{\mathbf{Q}}_t$ *# Surrogate value function* $Q_{\theta_2'}'$ *predicts reward from* $\mathbf{S}_t$, $\mathbf{A}_t$
$Q_{\theta_2}(\mathbf{S}_{t+1}, \hat{\mathbf{A}}_{t+1}) = \mathbf{Q}_{t+1}$ *# Online value function* Q_{θ_2} *predicts reward from* $\mathbf{S}_{t+1}$, $\hat{\mathbf{A}}_{t+1}$

$\nabla Q'_{\theta'_2} = \nabla_{\theta'_2} Q'_{\theta'_2}(\mathbf{S}_t, \mathbf{A}_t) \nabla_{\hat{\mathbf{Q}}_t} \mathcal{L}(\mathbf{R}_{t+1} + \mathbf{Q}_{t+1}, \hat{\mathbf{Q}}_t)$ # *Gradient of* $Q'_{\theta'_2}$ *to the loss function*

$\nabla J'_{\theta'_1} = -\mathbb{E}[\nabla_{\theta'_1} \boldsymbol{\pi}'_{\theta'_1}(\mathbf{S}_t) \nabla_{\hat{\mathbf{A}}_t} Q'_{\theta'_2}(\mathbf{S}_t, \hat{\mathbf{A}}_t)|_{\hat{\mathbf{A}}_t = \boldsymbol{\pi}'_{\theta'_1}(\mathbf{S}_t)}]$ # *Gradient of* $\boldsymbol{\pi}'_{\theta'_1}$ *to the Q-Value.*

Update $\boldsymbol{\theta}'_2$ in $Q'_{\theta'_2}$ using $\nabla Q'_{\theta'_2}$

Update $\boldsymbol{\theta}'_1$ in $\boldsymbol{\pi}'_{\theta'_1}$ using $\nabla J'_{\theta'_1}$

If *tau update* interval is reached:

$\boldsymbol{\theta}_1 = (1-\tau)\boldsymbol{\theta}_1 + \tau\boldsymbol{\theta}'_1$ # *Update parameters of* $\boldsymbol{\pi}_{\theta_1}$

$\boldsymbol{\theta}_2 = (1-\tau)\boldsymbol{\theta}_2 + \tau\boldsymbol{\theta}'_2$ # *Update parameters of* Q_{θ_2}

In order to reduce training time of adjusting the weights in each layer of GCN using the gradients, optimizers such as stochastic gradient descent (SGD) (Robbins and Monro, 1951; Kiefer and Wolfowitz, 1952; Ruder, 2016) or Adam (Kingma and Ba, 2015), is used for updating $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$. The exploration of DDPG's policy is activated by adding small Ornstein-Uhlenbeck noise (Uhlenbeck and Ornstein, 1930) into the output value of the policy function.

5 Reinforcement learning for structural optimization

5.1 State

This research utilizes the graph representation to express structural data, such as nodal coordinates, boundary conditions, and internal forces, at each optimization step which is equivalent to a step t in the MDP formulation.

Suppose we have five node features, and the ith row of the node feature matrix $\mathbf{N} \in \mathbb{R}^{n\times 5}$ is represented as $\boldsymbol{n}_i = \{x_i/\max_p x_p \quad y_i/\max_p y_p \quad z_i/z_{\max} \quad k^i_{\text{free}} \quad k^i_{\text{fix}}\}$, in which x_i, y_i, and z_i are the coordinates of node i, $z_{\max}$ is the pre-determined upper-bound value of z_i, $\max_p x_p$ and $\max_p y_p$ are the maximum coordinate values in each axis. Note that the minimum coordinate values are assumed to be 0 for all coordinates. k^i_{free} and k^i_{fix} are determined depending on the boundary condition. $(k^i_{\text{free}}, k^i_{\text{fix}}) = (0, 1)$ if node i is fixed support, and $(k^i_{\text{free}}, k^i_{\text{fix}}) = (1, 0)$ if node i is not supported.

In the adjacency matrix for frame elements $\mathbf{M}_1 \in \mathbb{R}^{n\times n}$, the existence of a frame element e connecting nodes i and j is denoted as $m_{1ij} = m_{1ji} = k^e_{\text{frame}}$ where $m_{1_{ij}}$ indicates (i, j) component of matrix $\mathbf{M}_1$. k^e_{frame} indicates the existence and non-existence of a 12-DoFs frame element e that connects nodes i and j by $k^e_{\text{frame}} = 1$ and 0, respectively. In the adjacency matrix for truss elements $\mathbf{M}_2 \in \mathbb{R}^{n\times n}$, the existence of a truss element e connecting nodes i and j is represented as $m_{2ij} = m_{2ji} = k^e_{\text{truss}}$. k^e_{truss} indicates the existence and non-existence of a 6-DoFs truss element e that connects nodes i and j by $k^e_{\text{truss}} = 1$ and 0, respectively. The combined adjacency matrix for frame and truss elements $\mathbf{M}_3 \in \mathbb{R}^{n\times n}$ is obtained by $\mathbf{M}_3 = \mathbf{M}_1 + \mathbf{M}_2$.

In order to evaluate the efficiency of the structural configuration, this paper proposes weighted adjacency matrices to represent the element internal forces. For the frame element, only a single weighted adjacency matrix $\mathbf{P}_1 \in \mathbb{R}^{n\times n}$ is determined using the ratio between the bending moment and the axial force, which for this type of structure is a useful index that helps minimizing the strain energy. Suppose frame element e connects nodes i and j, the entry p_{1ij} in $\mathbf{P}_1$ is determined as follows:

$$p_{1_{ij}} = k^e_{\text{frame}} b'_{ei} \Big/ \left(a'_e + 1\right) \tag{8a}$$

$$b'_{ei} = \left(|b_{ei}| - b^{\max}_{\text{f}}\right) \Big/ \left(b^{\max}_{\text{f}} - b^{\min}_{\text{f}}\right) \tag{8b}$$

$$a'_e = \left(|a_e| - a^{\max}_{\text{f}}\right) \Big/ \left(a^{\max}_{\text{f}} - a^{\min}_{\text{f}}\right) \tag{8c}$$

where b_{ei} is the bending moment around the horizontal axis on the section at node i, and a_e is the axial force of frame element e. $b^{\max}_{\text{f}}$ and $b^{\min}_{\text{f}}$ are the maximum and minimum absolute values of bending moments at the element ends. $a^{\max}_{\text{f}}$ and $a^{\min}_{\text{f}}$ are the maximum and minimum absolute axial forces of frame elements. For the truss elements, weighted adjacency matrix ($\mathbf{P}_2 \in \mathbb{R}^{n\times n}$) is a normalized form of the truss axial force. The entry $p_{2_{ij}}$ in $\mathbf{P}_2$ corresponding to element e connecting nodes i and j is determined as follows:

$$p_{2_{ij}} = k^e_{\text{truss}} a'_e \tag{9a}$$

$$a'_e = \left(|a_e| - a^{\max}_{\text{q}}\right) \Big/ \left(a^{\max}_{\text{q}} - a^{\min}_{\text{q}}\right) \tag{9b}$$

where a_e is the axial force in the truss element e. $a^{\max}_{\text{q}}$ and $a^{\min}_{\text{q}}$ are the maximum and minimum absolute values of axial forces for truss elements, respectively. All values in these weighted adjacency matrices are in the range of $[0, 1]$, which helps avoiding numerical instabilities during training.

The degree matrices for frame, truss, and combined frame and truss elements are denoted by $\mathbf{D}_1 \in \mathbb{R}^{n\times n}$, $\mathbf{D}_2 \in \mathbb{R}^{n\times n}$, and $\mathbf{D}_3 \in \mathbb{R}^{n\times n}$, respectively. Entries in each degree matrix $\mathbf{D}_u$ ($u = 1, 2, 3$) are computed using the associated adjacency matrix $\mathbf{M}_u$ ($u = 1, 2, 3$) as $d_{\text{u}_{ij}} = \delta_{ij} \times \sum_{i=1}^{n} m_{\text{u}_{ij}}$, where δ_{ij} is the Kronecker delta which is 0 if $i \neq j$ and 1 if $i = j$. The normalized adjacency matrices of frame, truss, and combined frame and truss elements are $\tilde{\mathbf{M}}_1 \in \mathbb{R}^{n\times n}$, $\tilde{\mathbf{M}}_2 \in \mathbb{R}^{n\times n}$, and $\tilde{\mathbf{M}}_3 \in \mathbb{R}^{n\times n}$, respectively, which are computed using Eq. 5. The GCN-DDPG agent for bracing direction optimization uses $\tilde{\mathbf{M}}_1$, $\tilde{\mathbf{M}}_2$, $\tilde{\mathbf{M}}_3$, $\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{M}_2$, and $\mathbf{N}$ as a representation of the environment $\mathbf{S}_t$ which will be further explained in Section 5.2.

5.2 GCN-DDPG agent

Policy and value functions (i.e., Actor and Critic networks of a GCN-DDPG agent) consist of multiple GCN layers. The policy function takes state data described in Section 5.1 as input to compute the output denoted as $\mathbf{O} \in \mathbb{R}^{n\times h}$ which has the same number of rows as those of the node feature matrix $\mathbf{N} \in \mathbb{R}^{n\times g}$.

This output is used to determine bracing directions, explained in Section 5.3.

The inputs of the value function are state data and the output from the policy function to compute an estimation of the accumulated reward. In the value function, another matrix for prediction of the bracing directions denoted as $\mathbf{M}_o \in \mathbb{R}^{n\times n}$ is internally computed from the output of the policy function $\mathbf{O} \in \mathbb{R}^{n\times h}$, multiplied element-wise with $\mathbf{M}_2$ to exclude non-bracing elements, and normalized within the range [0,1] through dividing by h as $\mathbf{M}_o = (\mathbf{O}\cdot\mathbf{O}^{\mathrm{T}})\odot\mathbf{M}_2/h$, where $\odot$ is the element-wise multiplication.

When multiple GCN layers are connected, a node feature matrix is replaced with an output from the prior GCN layer. To represent internal forces in the structure, a normalized adjacency matrix in a GCN layer can be replaced with a weighted adjacency matrix. The output of a GCN layer is transformed by the Rectified Linear Unit (ReLU) activation function (Nair and Hinton, 2010) similar to the original GCN (Kipf and Welling, 2017) for all layers of both policy and value functions. ReLU is a nearly linear function that is computationally efficient for gradient-based optimization (i.e., SGD or Adam) (Chigozie et al., 2020). ReLU is applied to all layers of both policy and value functions except the last layer of the policy function which is processed by the Sigmoid activation function for predicting probability-based output (i.e., probability of taking action A_t^i in a state S_t) (Chigozie et al., 2020). Eq. 10a and Eq. 10b represent GCN layers with ReLU and Sigmoid activation functions, respectively, where $\mathbf{N}' \in \mathbb{R}^{n\times g}$ denotes a node feature matrix or an output from a prior GCN layer and $\mathbf{M}' \in \mathbb{R}^{n\times n}$ denotes $\tilde{\mathbf{M}}_1, \tilde{\mathbf{M}}_2, \tilde{\mathbf{M}}_3, \mathbf{P}_1, \mathbf{P}_2$, or $\mathbf{M}_o$ with size $\mathbb{R}^{n\times n}$. Eq. 10c represents multiple computing loops using the same GCN layer.

$$\mu(\mathbf{N}', \mathbf{M}') = \mathrm{ReLU}(\mathbf{N}'\mathbf{M}'\mathbf{w}_\mu) \tag{10a}$$

$$\sigma(\mathbf{N}', \mathbf{M}') = \mathrm{Sigmoid}(\mathbf{N}'\mathbf{M}'\mathbf{w}_\sigma) \tag{10b}$$

$$\mathrm{iter}^{\phi}[\mu(\mathbf{N}', \mathbf{M}')] = \underbrace{\mu(\mathbf{M}'(\mu(\mathbf{M}'(\ldots)\mathbf{w}_\mu))\mathbf{w}_\mu)}_{\phi\text{times}} \tag{10c}$$

where $\mathrm{ReLU}(\cdot) = \max(0, \cdot)$, $\mathrm{Sigmoid}(\cdot) = 1/(1+e^{-(\cdot)})$, and ϕ in $\mathrm{iter}^{\phi}[\,]$ indicates the number of computing loops.

Since the output of a GCN layer is a matrix but the value function output is a scalar representing the estimation of accumulated reward, two operations are used for transforming the output matrix of the last GCN layer of the value function into a scalar value. The first operation is a global sum pooling operation (GSP) (Aich and Stavness, 2018) which transforms the output matrix into a vector by summing up all entries in each column of the output matrix. Let $\mathbf{V} \in \mathbb{R}^{n\times g}$ be a matrix, the GSP operation to transform $\mathbf{V}$ into a vector can be represented as

$$\mathrm{Pool}(\mathbf{V}) = \left[\sum_{i=1}^{n} v_{i,1} \quad \cdots \quad \sum_{i=1}^{n} v_{i,g}\right] \in \mathbb{R}^{1\times g} \tag{11}$$

TABLE 1 Policy and value functions of GCN-DDPG for bracing direction optimization.

Policy function π	Value function Q
Inputs: $\tilde{\mathbf{M}}_1, \tilde{\mathbf{M}}_2, \tilde{\mathbf{M}}_3, \mathbf{P}_1, \mathbf{P}_2, \mathbf{N}$	Inputs: $\tilde{\mathbf{M}}_1, \tilde{\mathbf{M}}_2, \tilde{\mathbf{M}}_3, \mathbf{P}_1, \mathbf{P}_2, \mathbf{N}, \mathbf{M}_2, \mathbf{O}$
Computation:	Computation:
Step 1: $\mathbf{N}_{1.1} = \mu(\mu(\mathbf{N}, \mathbf{P}_1), \tilde{\mathbf{M}}_1)$ $\mathbf{N}_{2.1} = \mathrm{iter}^2[\mu(\mathbf{N}_{1.1}, \mathbf{P}_1)]$	Step 1: $\mathbf{N}_{1.1} = \mu(\mu(\mathbf{N}, \mathbf{P}_1), \tilde{\mathbf{M}}_1)$ $\mathbf{N}_{2.1} = \mathrm{iter}^2[\mu(\mathbf{N}_{1.1}, \mathbf{P}_1)]$
Step 2: $\mathbf{N}_{1.2} = \mu(\mu(\mathbf{N}, \mathbf{P}_2), \tilde{\mathbf{M}}_2)$ $\mathbf{N}_{2.2} = \mathrm{iter}^2[\mu(\mathbf{N}_{1.2}, \mathbf{P}_2)]$	Step 2: $\mathbf{N}_{1.2} = \mu(\mu(\mathbf{N}, \mathbf{P}_2), \tilde{\mathbf{M}}_2)$ $\mathbf{N}_{2.2} = \mathrm{iter}^2[\mu(\mathbf{N}_{1.2}, \mathbf{P}_2)]$
Step 3: $\mathbf{N}_3 = \mathbf{N}_{2.1} + \mathbf{N}_{2.2}$ $\mathbf{O} = \sigma(\mathbf{N}_3, \tilde{\mathbf{M}}_3)$	Step 3: $\mathbf{M}_o = (\mathbf{O}\cdot\mathbf{O}^{\mathrm{T}})\odot\mathbf{M}_2/h$
	Step 4: $\mathbf{N}_{1.3} = \mu(\mu(\mathbf{O}, \mathbf{M}_o), \tilde{\mathbf{M}}_2)$ $\mathbf{N}_{2.3} = \mathrm{iter}^2[\mu(\mu(\mathbf{N}_{1.3}, \mathbf{M}_o), \tilde{\mathbf{M}}_2)]$
	Step 5: $\mathbf{N}_3 = \mathbf{N}_{2.1} + \mathbf{N}_{2.2} + \mathbf{N}_{2.3}$ $\mathbf{N}_4 = \mu(\mathbf{N}_3, \tilde{\mathbf{M}}_3)$
	Step 6: $\mathbf{N}_5 = \mathrm{Pool}(\mathbf{N}_4)$
	Step 7: $Q = f_{NN}(\mathbf{N}_5)$
Output: $\mathbf{O} \in \mathbb{R}^{n\times h}$	Output: $Q \in \mathbb{R}^{1\times 1}$

The second operation is to compute the estimation of accumulated reward (i.e., Q-value $\in \mathbb{R}^{1\times 1}$) from the vector output of the GSP operation using a neural network which consists of approximation functions that have adjustable weight parameters and activation functions (Goodfellow et al., 2016). These approximation functions are connected together so that the output of the prior approximation function is the input of the next approximation function, similarly to how the GCN layers are connected, where every approximation function, except the last one, is called a hidden layer (Goodfellow et al., 2016). Eq. 12 represents an NN with two hidden layers, used in this work, for computing the estimation of accumulated reward from the vector output of the GSP operation $\mathbf{H} \in \mathbb{R}^{1\times g}$ with adjustable internal weight matrices $\boldsymbol{\mathcal{W}}_1$, $\boldsymbol{\mathcal{W}}_2$, $\boldsymbol{\mathcal{W}}_{\mathrm{out}}$, adjustable internal bias vectors $\boldsymbol{\mathcal{B}}_1$ and $\boldsymbol{\mathcal{B}}_2$, and adjustable internal bias scalar $\boldsymbol{\mathcal{B}}_{\mathrm{out}}$ as

$$f_{NN}(\mathbf{H}) = \boldsymbol{\mathcal{W}}_{\mathrm{out}}\Big(\mathrm{ReLU}\Big(\boldsymbol{\mathcal{W}}_2\left(\mathrm{ReLU}\left(\boldsymbol{\mathcal{W}}_1\mathbf{H}^{\mathrm{T}} + \boldsymbol{\mathcal{B}}_1\right)\right)^{\mathrm{T}} + \boldsymbol{\mathcal{B}}_2\Big)\Big)^{\mathrm{T}} + \boldsymbol{\mathcal{B}}_{\mathrm{out}} \in \mathbb{R}^{1\times 1} \tag{12}$$

Table 1 summarizes the computation processes of the policy and value functions of the GCN-DDPG agent for bracing direction optimization. In each column, the 1st row indicates if the computation belongs to the policy or the value function. The 2nd row denotes input data used for the computation. The 3rd row indicates the computation process using GCN layers, GSP operation, and NN in Eqs 10a–12a–Eqs 10a–12.

In the policy function, inputs from state data of frame and truss elements are separately processed in Steps 1 and 2, respectively. Output matrices in Steps 1 and 2 are combined to compute the output of the policy function in Step 3. In the value function, inputs from state data of frame and truss are also processed separately in Steps 1 and 2. Step3 computes the matrix for element direction prediction $\mathbf{M}_o$ which is used for processing

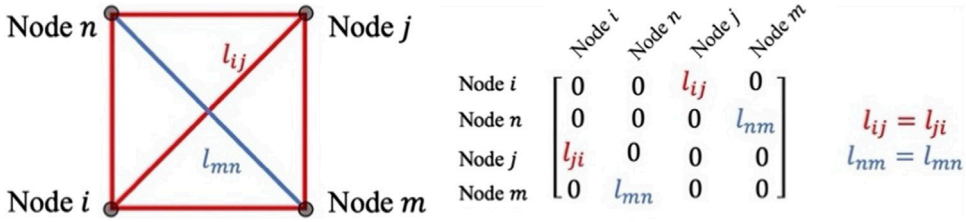


FIGURE 3
Bracing directions and associate link predictions in a grid cell.

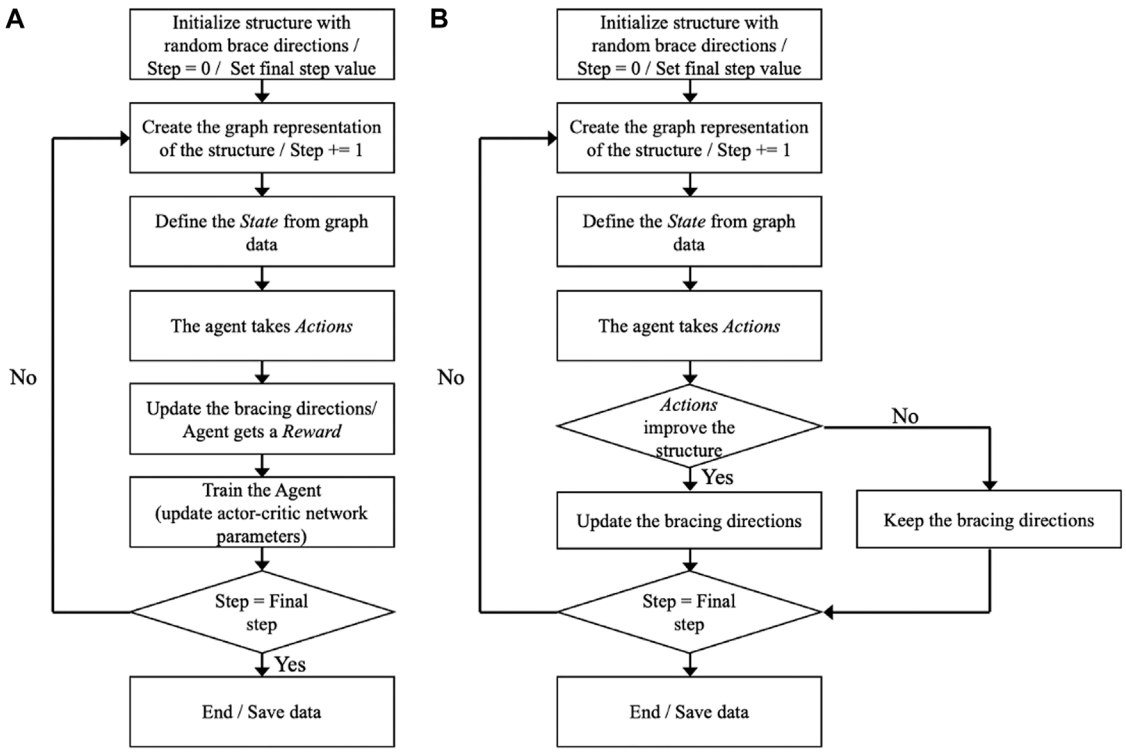


FIGURE 4
Algorithms of the proposed method; **(A)** Training phase, **(B)** Test phase.

action data in Step 4. Output matrices from the Steps 1, 2, and 4 are combined and re-processed by another GCN layer in Step 5. In Steps 6 and 7, the output matrix is converted into a vector using the GSP operation and the vector is finally converted to the Q-value by the NN. The last row denotes the output of the functions.

5.3 Action

The bracing direction in each grid cell is determined from a dot product of each row of $\mathbf{O} \in \mathbb{R}^{n \times h}$ from the policy function output, similarly to the link prediction using GCN (Kipf and Welling, 2016) for the existence of a connection between two nodes in the graph. In this paper, the structural nodes are represented as graph nodes. Therefore, the prediction of a link or connection between two structural nodes is equivalent to a structural element (bracing element) that connects those nodes. Figure 3 shows a 4-node structure in a grid cell with two possible diagonal braces which connect node i to node j and node n to node m, respectively. The dot product of each output matrix is used to predict the value of l_{ij} and l_{nm}. l_{ij} is equivalent to l_{ji} and l_{nm} is also equivalent to l_{mn} as shown in the matrix in

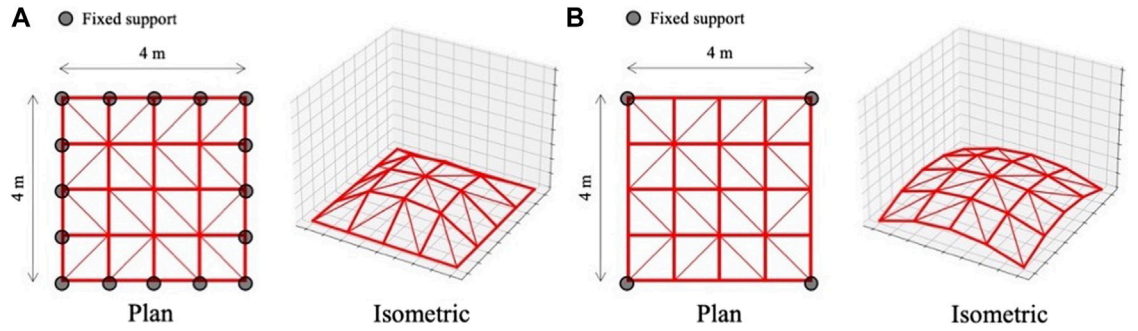


FIGURE 5
Structural models for bracing direction optimization during training phase; **(A)** Support condition 1, **(B)** Support condition 2.

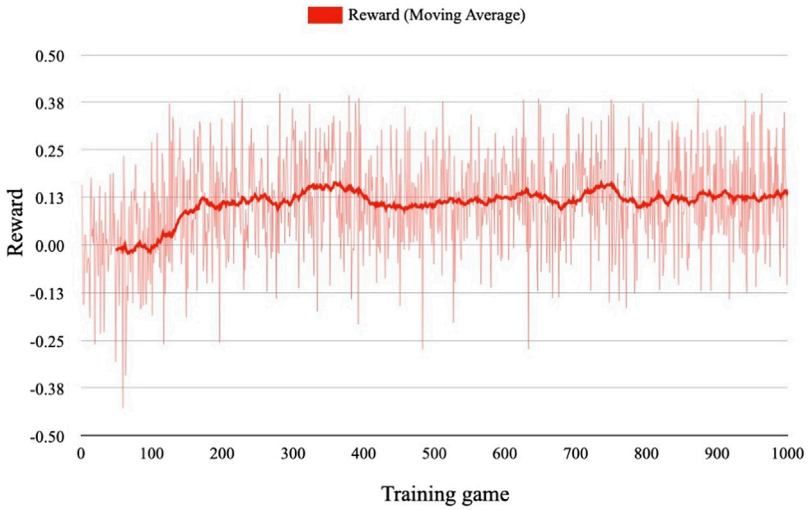


FIGURE 6
Variation of reward and its moving average in training phase.

Figure 3. The bracing direction optimization formulation in Eqs 3a–ca–cEqs 3a–c allows only one brace in each grid cell. Therefore, l_{ij} and l_{nm} are compared to determine the bracing direction in a grid cell.

The number of nodal embedding output dimensions is 100; i.e., $h = 100$. From the output of the policy function $\mathbf{O} \in \mathbb{R}^{n\times 100}$, the brace in grid cell k is determined as

$$\mathrm{A}_t^k = \begin{cases} (c_{1,k},\ c_{2,k}) = (1,\ 0) & (\text{if } l_{ij} > l_{mn}) \\ (c_{1,k},\ c_{2,k}) = (0,\ 1) & (\text{if } l_{ij} < l_{mn}) \end{cases} \tag{13a}$$

$$l_{ij} = l_{ji} = \sum_{u=1}^{100} o_{i,u} o_{j,u} \tag{13b}$$

$$l_{mn} = l_{nm} = \sum_{u=1}^{100} o_{n,u} o_{m,u} \tag{13c}$$

At each step, the agent can change any number of brace directions.

5.4 Reward

In RL, a reward signal is used for training an agent. The reward signal R_{t+1} that the agent receives after executing action A_t in a state S_t at step t is formulated from the change of the strain energy as follows:

$$\mathrm{R}_{t+1} = (E_t - E_{t+1})/E_0 \tag{14}$$

where E_t and E_0 are the strain energy of the structure at step t and the strain energy of the initial structure.

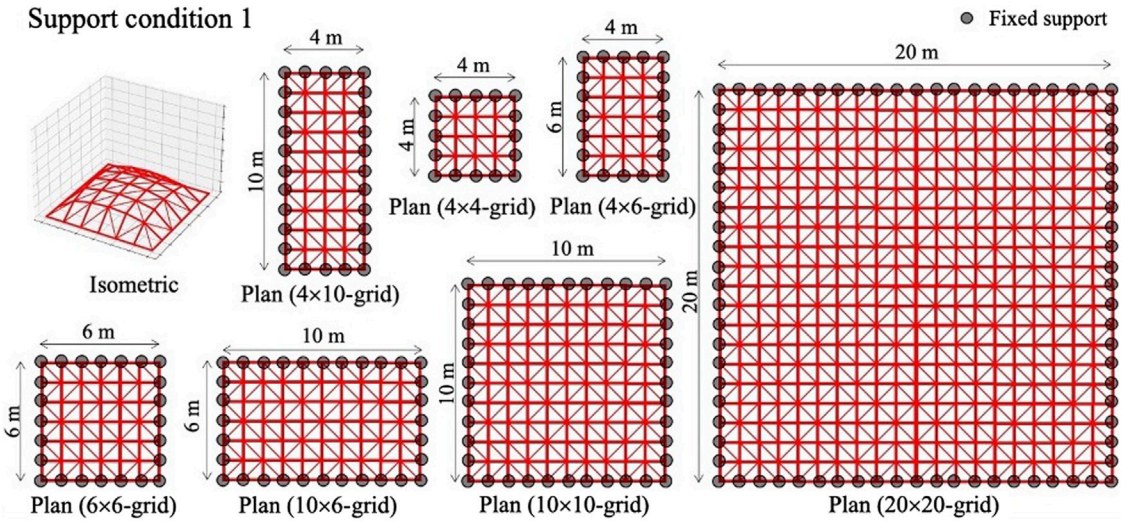


FIGURE 7
Structural models for test phase: Support condition 1.

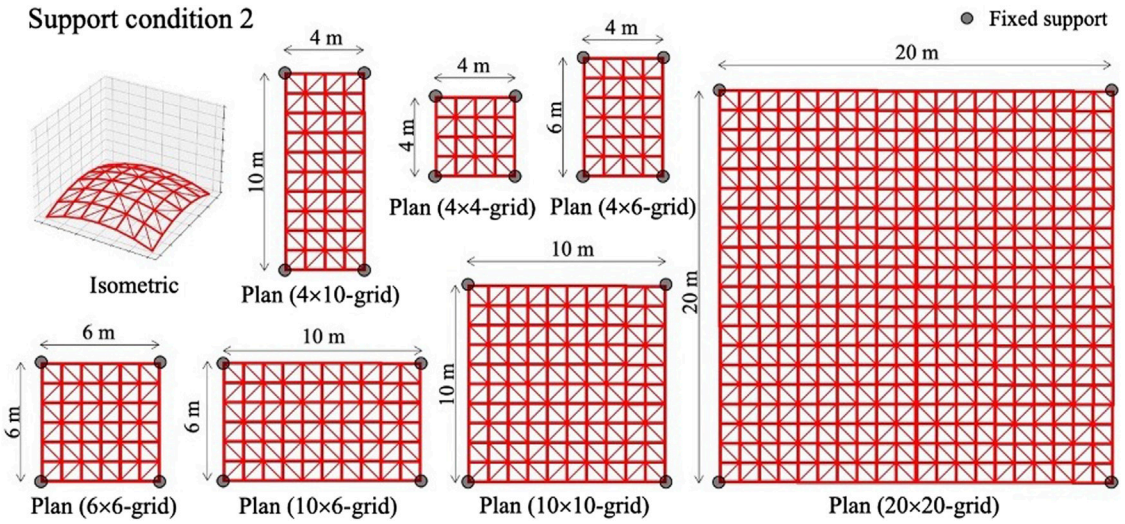


FIGURE 8
Structural models for test phase: Support condition 2.

6 Numerical examples

6.1 General experiment setting and structural model

The agent is trained to optimize the structure in the *training phase* during which the ability to improve performance is assessed. In the *test phase*, the agent performance is evaluated on structural configurations other than those used in training. The method is implemented using Python 3.6 environment. A PC with CPU Intel Core i5-6600 (3.3 GHz, 4 cores) and GPU AMD Radeon R9 M395 2 GB is employed for computation.

Training is carried out on a grid shell structure with 4 × 4 grids and diagonal truss braces. Each grid cell has dimensions of 1.0 m by 1.0 m. To simplify the problem, the 12-DoFs frame element has a hollow cylindrical section with an external diameter of 100 mm and an internal diameter of 90 mm. The 6-DoFs truss element has a solid circular section with a diameter of 43.6 mm. Both elements have Young's modulus of 205 kN/

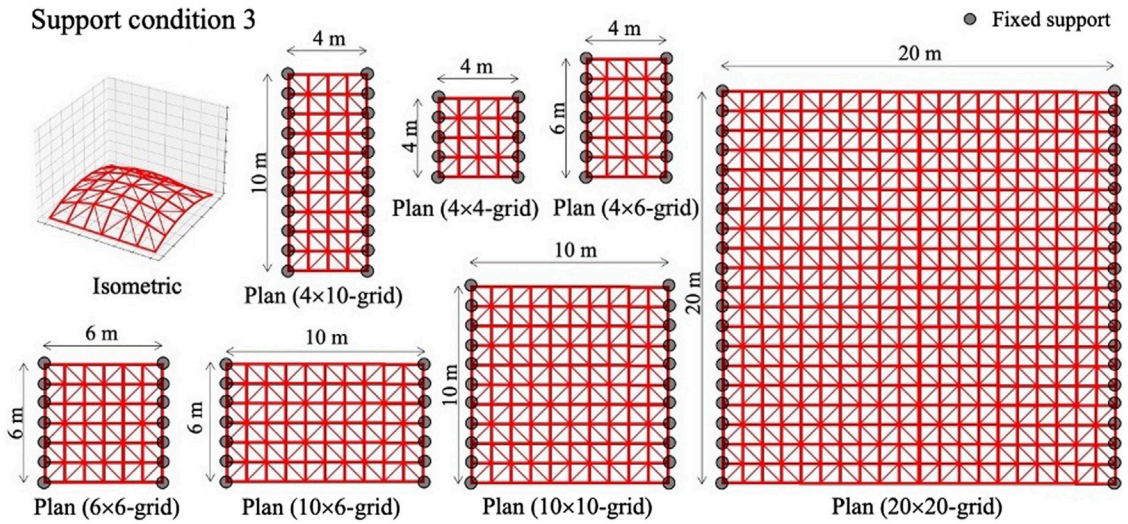


FIGURE 9
Structural models for test phase: Support condition 3.

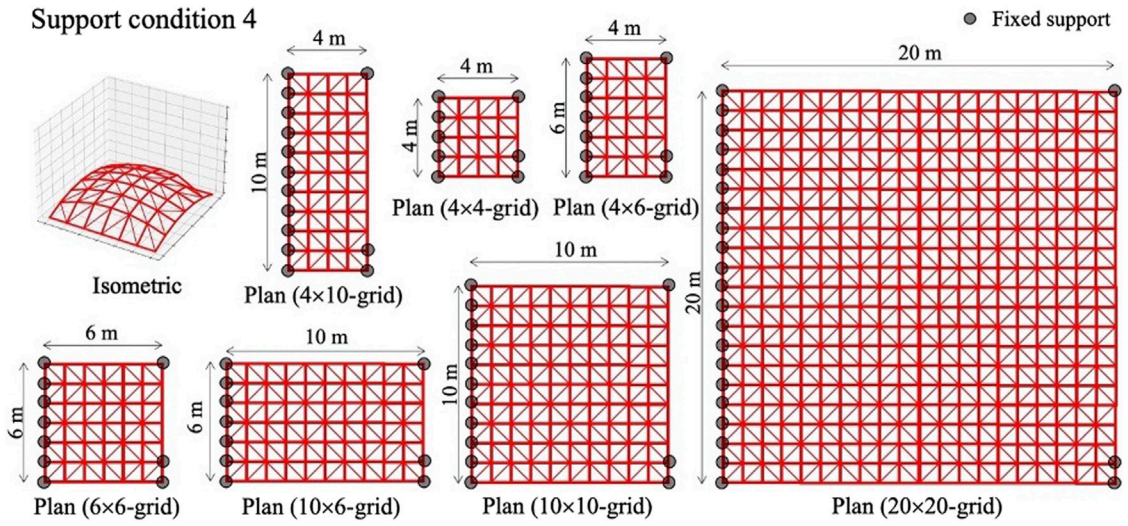


FIGURE 10
Structural models for test phase: Support condition 4.

mm^2 and a similar weight of 12 kg/m. All structural nodes are subjected to a vertical point load of 10 kN.

Results obtained in the test phase using the proposed method are compared with those of the enumeration method (EM) and the genetic algorithm (GA); EM is used for the benchmark when it is feasible to compute all possible solutions, and GA is used for the benchmark when computing all possible solution is not feasible due to the large search space.

The algorithm flowcharts for training and test phases are given in Figure 4. During the optimization, each MDP is denoted as a step. The loop of MDPs or a *game* is terminated when the final step is reached.

TABLE 2 Test results.

Support conditions	Grid size	Min. (N·m)	Mean (N·m)	Std. (N·m)	Reduction (%)
1	4 × 4	4.38	4.53	0.19	16.83
	4 × 6	15.71	16.01	0.38	13.21
	6 × 6	63.15	67.40	2.56	18.10
	4 × 8	72.49	73.72	0.69	5.55
	10 × 6	450.47	459.92	7.89	12.89
	10 × 10	2,242.26	2,288.05	40.79	19.80
	20 × 20	213,715.82	217,120.90	1742.84	10.08
2	4 × 4	77.88	89.35	10.02	19.94
	4 × 6	284.81	290.09	3.98	20.21
	6 × 6	818.14	896.06	46.93	19.71
	4 × 10	2,124.56	2,142.64	17.10	20.76
	10 × 6	4,642.28	4,713.70	37.63	19.79
	10 × 10	17,683.39	17,966.80	181.50	24.36
	20 × 20	1,275,569.06	1,312,975.69	42,317.86	19.66
3	4 × 4	32.11	32.78	0.53	6.43
	4 × 6	70.09	72.40	1.44	5.29
	6 × 6	345.73	356.54	6.92	12.59
	4 × 10	193.91	198.18	2.84	6.69
	10 × 6	2,841.17	2,869.52	18.50	10.95
	10 × 10	8,219.87	8,333.38	101.09	12.28
	20 × 20	589,997.91	596,066.09	4,460.72	9.38
4	4 × 4	40.81	41.34	0.46	17.71
	4 × 6	156.09	161.09	3.61	22.68
	6 × 6	527.56	535.43	5.95	17.30
	4 × 10	1,192.81	1,213.16	11.94	21.76
	10 × 6	3,466.03	3,542.59	47.95	11.63
	10 × 10	12,629.96	12,762.32	109.98	16.49
	20 × 20	920,936.67	931,040.36	8,426.17	16.35

6.2 Training phase

In each game of training, a dome-shaped structural model is initialized with supports assigned to two pre-determined structural shapes indicated in Figure 5. The maximum and minimum nodal height is 1 and 0 m, respectively. Note that structural shapes and size of structural elements are not adjusted during the optimization process. The brace directions are randomly initialized at every game.

The final step in each game is 200. At each step, the directions of braces are adjusted according to the agent actions. The agent surrogate functions are adjusted using the Adam optimizer. The mini-batch size is set to 32 and the learning rates are 10^{-7} and 10^{-6} for policy and value function, respectively. In the value function, the NN in Eq. 12 in Section 5.2 has two hidden layers, each consisting of 200 neurons. The mean square error is used as the loss function, and learning rates are reduced by a factor of $\beta = 0.1$ every 200 games (20,000 steps). Weights and biases of the surrogate functions are synchronized with those of the online functions every 100 steps using $\tau = 0.05$. The GCN-DDPG agent is trained for 1,000 games. In Figure 6, the vertical axis represents the cumulative reward obtained during training, and the horizontal axis is the game number. The thick line shows the moving average of the reward with a window size set to 50. From Figure 6, the cumulative reward increases during the first 200 games and then remains stable around a certain value. Because support and bracing directions are changed at every training game, the fact that the cumulative reward stabilizes indicates that the agent has the learning

10.3389/fbuil.2022.899072

TABLE 3 Total computational cost of structural analysis of each method.

Support conditions	Grid size	Size of the global stiffness matrix $n_D \times n_D$	Number of structural analyses (times)		
			GCN-DDPG (test phase)	Benchmarks	
				EM	GA
1	4 × 4	54 × 54	1,000	65,536	5,000
	4 × 6	90 × 90	1,000	—	5,000
	6 × 6	150 × 150	1,000	—	5,000
	4 × 10	198 × 198	1,000	—	5,000
	10 × 6	270 × 270	1,000	—	5,000
	10 × 10	486 × 486	1,000	—	5,000
	20 × 20	2,166 × 2,166	1,000	—	5,000
2	4 × 4	126 × 126	1,000	65,536	5,000
	4 × 6	186 × 186	1,000	—	5,000
	6 × 6	270 × 270	1,000	—	5,000
	4 × 10	306 × 306	1,000	—	5,000
	10 × 6	438 × 438	1,000	—	5,000
	10 × 10	702 × 702	1,000	—	5,000
	20 × 20	2,622 × 2,622	1,000	—	5,000
3	4 × 4	90 × 90	1,000	65,536	5,000
	4 × 6	138 × 138	1,000	—	5,000
	6 × 6	210 × 210	1,000	—	5,000
	4 × 10	234 × 234	1,000	—	5,000
	10 × 6	378 × 378	1,000	—	5,000
	10 × 10	594 × 594	1,000	—	5,000
	20 × 20	2,394 × 2,394	1,000	—	5,000
4	4 × 4	102 × 102	1,000	65,536	5,000
	4 × 6	150 × 150	1,000	—	5,000
	6 × 6	234 × 234	1,000	—	5,000
	4 × 10	246 × 246	1,000	—	5,000
	10 × 6	402 × 402	1,000	—	5,000
	10 × 10	642 × 642	1,000	—	5,000
	20 × 20	2,502 × 2,502	1,000	—	5,000
Total computational cost of structural analysis including the training phase			128,000		140,000

capability to optimize structural configurations of different topology and support conditions.

6.3 Test phase

In the test phase, 4 × 4, 4 × 6, 6 × 6, 4 × 10, 10 × 6, 10 × 10, and 20 × 20-grid shells with pre-determined geometries are employed to investigate the capability of the trained agent on configurations that have not been tested in the training phase. In Figures 7–10, four support conditions denoted as 1, 2, 3, and 4 are considered for each frame model. The bracing directions are initialized randomly.

Similarly to the training phase, the number of steps for the test phase is 100. However, in the test phase, only actions improving the objective function value are accepted at each step. The agent optimizes each structural model 10 times. Table 2 shows the minimum (Min.), mean, and standard deviation (Std.) for the strain energy, and mean energy reduction rate (Reduction) for each structural model.

10.3389/fbuil.2022.899072

TABLE 4 Comparison of results obtained by GCN-DDPG (test phase) and benchmarks.

Support conditions	Grid size	Min. (N·m)	Method	Benchmarks (N·m)	Diff
1	4 × 4	4.38	EM/GA	4.38/4.38	0.00
	4 × 6	15.71	GA	15.66	0.00
	6 × 6	63.15		61.55	0.03
	4 × 10	72.49		71.88	0.01
	6 × 10	450.47		422.35	0.07
	10 × 10	2,242.26		2,135.28	0.05
	20 × 20	213,715.82		202,566.99	0.06
2	4 × 4	77.88	EM/GA	77.88/77.88	0.00
	4 × 6	284.81	GA	279.67	0.02
	6 × 6	818.14		777.99	0.05
	4 × 10	2,124.56		2046.13	0.04
	6 × 10	4,642.28		4,342.47	0.07
	10 × 10	17,683.39		16,294.39	0.09
	20 × 20	1,275,569.06		1,170,860.54	0.09
3	4 × 4	32.11	EM/GA	31.59/31.59	0.02
	4 × 6	70.09	GA	66.69	0.05
	6 × 6	345.73		342.38	0.01
	4 × 10	193.91		182.56	0.06
	6 × 10	2,841.17		2,725.89	0.04
	10 × 10	8,219.87		7,697.87	0.07
	20 × 20	589,997.91		553,614.84	0.07
4	4 × 4	40.81	EM/GA	40.77/40.77	0.00
	4 × 6	156.09	GA	150.95	0.03
	6 × 6	527.56		485.53	0.09
	4 × 10	1,192.81		1,140.11	0.05
	6 × 10	3,466.03		3,150.67	0.10
	10 × 10	12,629.96		11,462.31	0.10
	20 × 20	920,936.67		850,282.94	0.08

In the test phase, the strain energy of the structure can be reduced using the proposed method by 5–25%, depending on the structure size and support conditions. For all cases, the minimum and mean of strain energy obtained from 10 tests are very similar, and the standard deviations are low compared with the mean. The trained agent is capable to optimize the bracing directions to reduce the strain energy on configurations that were not tested in the training phase.

6.4 Comparison of computation cost and performance with EM and GA

For the 4×4-grid structure, the global optimal solution can be obtained using the EM which generates 2^k combinations of bracing directions for k grid cells. For structures with a greater number of grids, it is not feasible to use EM. Therefore, the GA is employed to obtain optimal solutions (global optimum cannot be guaranteed). GA is a meta-heuristic method inspired by the process of natural evolution which can be used to solve combinatorial optimization problems. The key operations in GA are *selection* to transfers good solutions from one generation to the next generation, *crossover* to generate new solutions, and *mutation* to modify solutions with a certain probability, which can be helpful to avoid local minima. In this research, bracing directions in each grid cell are represented through binary strings and the GA algorithm is taken from the GA python library named Distributed Evolutionary Algorithms in Python (DEAP) (Fortin et al., 2012). The comparison is made to

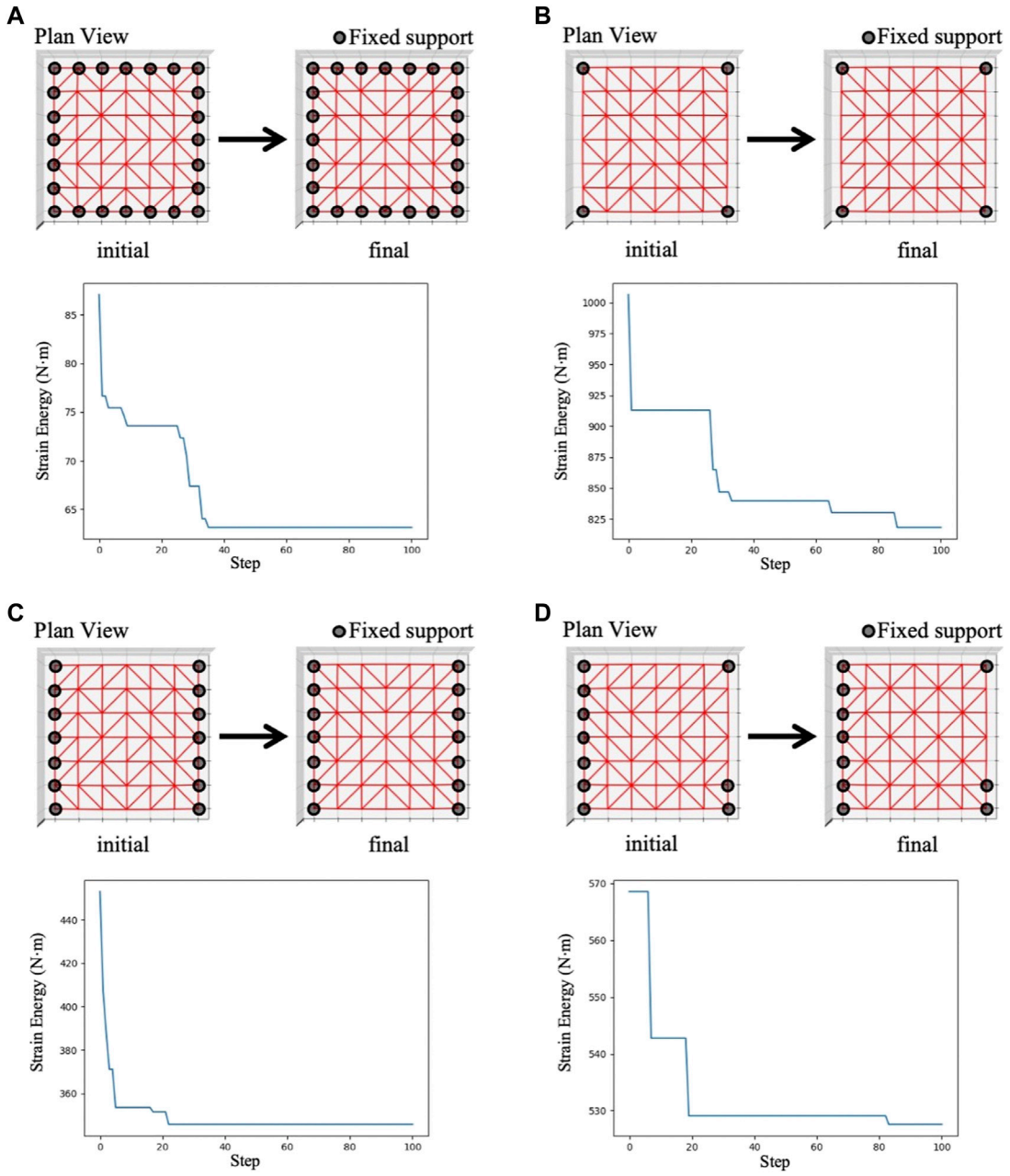


FIGURE 11
GCN-DDPG result of the 6×6-grid structure in the test phase; **(A)** Support condition 1, **(B)** Support condition 2, **(C)** Support condition 3, **(D)** Support condition 4.

benchmark both the performance and computational efficiency of GCN-DDPG and GA for the early-stage design of grid shells where several configurations are to be evaluated. Therefore, the number of population and generation of GA are determined based on feasibility of computational cost. In the following examples, the numbers of population and generations are 50 and 100, respectively.

The number of structural analyses carried out with GCN-DDPG in the training phase is equal to the number of training games multiplied by the number of steps in each training game, which is 100,000. A trained agent can be used in the test phase and in other problems without re-training. The computational cost of GCN-DDPG in the test phase, EM, and GA for each problem, and the size of the global stiffness

matrix are shown in Table 3 where the last row indicates the total computational cost. The total number of structural analyses required by GCN-DDPG is less than that required by GA. Since a significant computation time is required for the analysis of large-size structures, the GCN-DDPG is more efficient than the GA when applied to bracing direction optimization of grid shell structures that comprise many elements.

Benchmark results are compared in Table 4. The ratio of the difference between the minimum strain energy solution obtained by RL and that obtained by GA (and EM) is shown in the column labeled 'Diff', which is computed as follows:

$$\text{Diff} = (\min(\text{Result}_{\text{RL}}) - \text{Result}_{\text{GA}})/\text{Result}_{\text{GA}} \quad (15)$$

From Table 4, the solution quality and the efficiency of the GCN-DDPG agent can be verified. In most cases, results obtained by the GCN-DDPG agent are comparable to those obtained by EM and GA within a margin of 10% difference using less computational cost. The proposed method is useful in early-stage design, which typically requires testing several structures, and therefore an efficient computational process is needed. The RL agent could be further trained using other structural configurations to improve its performance.

Figure 11 shows initial brace topology, final brace topology, and the change of strain energy of GCN-DDPG best results for 6 × 6-grid structural models. Although the structural configurations differ considerably from those used in the training phase in terms of support conditions and size, the agent is capable to minimize the strain energy by adjusting the bracing directions. Therefore, it is possible to train the agent using small-size structural models and use it to optimize structures with different support conditions and sizes.

From Figures 11A,C where the support locations are symmetric, the agent obtains solutions with symmetrical layouts, despite the fact that a symmetrical feature is not explicitly represented.

7 Conclusion

A combined method of DDPG and GCN has been formulated for bracing direction optimization of grid shell structures to minimize the strain energy. The proposed DDPG framework allows the agent to modify the bracing direction in all grid cells simultaneously at each optimization step. The node feature matrix, adjacency matrices, and weighted adjacency matrices are formulated to encode the structural configuration and internal forces as graph representations. The agent is trained using Markov Decision Process (MDP) in the RL framework whereby training data are collected by interacting with the environment. The value function or critic network updates internal weights and biases to minimize the prediction loss for the accumulated reward or Q-value so that it can predict the long-term accumulated reward from state and action. The policy function or actor network updates weights and biases to maximize the equivalent reward calculated by the value function.

Numerical examples show that the trained agent can effectively optimize bracing directions to minimize the strain energy in the test phase. The agent is capable to optimize the bracing direction of structural configurations with size and support conditions different than those in the training phase. The proposed method produces solutions that compare, albeit of marginally lower quality, with those produced through the enumeration method (EM) and the genetic algorithm (GA). However, the trained agent can be employed for additional configurations to those tested in this work. The agent performs well for relatively large structural models without the re-training, thereby significantly reducing the computational cost of optimization. Future work should investigate whether the RL method can be applied without re-training to design significantly larger-size structural configurations (e.g., 200 × 200 grid size) compared to those employed for training. Therefore, The proposed method has good potential to be employed effectively in early-stage design, which typically requires testing several configurations.

Data availability statement

Experiment data from this research are available on the request to corresponding authors.

Author contributions

C-tK: Conceptualization, implementation, writing-original draft, data curation. KH: Conceptualization, writing-review and editing, data curation, resource. MO: Conceptualization, writing-review and editing, data curation, resource.

Funding

This work was supported by MEXT scholarship (Grant Number 180136); and JSPS KAKENHI (Grant Numbers JP 20H04467, JP 21K20461).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Aich, S., and Stavness, I. (2018). *Global sum pooling: A generalization trick for object counting with small datasets of large images*. Arxiv:1805.11123. [Online]. Available at: https://arxiv.org/abs/1805.11123.

Bellman, R. (1954). The theory of dynamic programming. *Bull. Amer. Math. Soc.* 60 (6), 503–515. doi:10.1090/s0002-9904-1954-09848-8

Bellman, R. (1957). A markovian decision process. *Indiana Univ. Math. J.* 6, 679–684. doi:10.1512/iumj.1957.6.56038

Berke, L., and Hajela, P. (1993). "Application of neural nets in structural optimization," in *Optimization of large structural systems. NATO ASI series (Series E: Applied sciences)*. Editor G. I. N. Rozvany (Dordrecht: Springer), Vol. 231. doi:10.1007/978-94-010-9577-8_36

Chigozie, N., Ijomah, W., Gachagan, A., and Stephen, M. (2020). *Activation functions: Comparison of trends in practice and research for deep learning*. Arxiv:1811.03378 [Online]. Available at: https://arxiv.org/abs/1811.03378.

Christensen, P. W., and Klarbring, A. (2009). *An introduction to structural optimization. Solid mechanics and its applications*, Vol. 153. Dordrecht: Springer. doi:10.1007/978-1-4020-8666-3

Dhingra, A. K., and Bennage, W. A. (1995). Topological optimization of truss structures using simulated annealing. *Eng. Optim.* 24 (4), 239–259. doi:10.1080/03052159508941192

Eltouny, K., and Liang, X. (2021). Bayesian-optimized unsupervised learning approach for structural damage detection. *Computer-Aided. Civ. Infrastructure Eng.* 36, 1249–1269. doi:10.1111/mice.12680

Fortin, F. A., De Rainville, F. M., Gardner, M. A., Parizeau, M., and Gagné, C. (2012). DEAP: Evolutionary algorithms made easy. *J. Mach. Learn. Res. Mach. Learn. Open Source Softw.* 13, 2171–2175.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning*. The MIT Press. 9780262035613. Available at: https://www.deeplearningbook.org.

Haarnoja, T., Zhou, A., Abbeel, P., and Levine, S. (2018). *Soft actor-critic: off-policy maximum entropy deep reinforcement learning with a stochastic actor*. ICML, 1856–1865. Arxiv:1801.01290. Available at: https://dblp.uni-trier.de/db/conf/icml/icml2018.html#HaarnojaZAL18.

Hayashi, K., and Ohsaki, M. (2021). Reinforcement learning and graph embedding for binary truss topology Optimization under Stress and Displacement Constraints. *Front. Built Environ.* 6, 59. doi:10.3389/fbuil.2020.00059

Hung, T. V., Viet, V. Q., and Thuat, D. V. (2019). A deep learning-based procedure for estimation of ultimate load carrying of steel trusses using advanced analysis. *J. Sci. Technol. Civ. Eng. (STCE) - HUCE* 13 (3), 113–123. doi:10.31814/stce.nuce2019-13(3)-11

Ivakhnenko, A. G. (1968). The group method of data handling – a rival of the of stochastic approximation. *Sov. Autom. Control* 13 (3), 43–55.

Jeong, M. J., and Yoshimura, S. (2002). "An evolutionary clustering approach to pareto solutions in multiobjective optimization," in Proceedings of the ASME 2002 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, 141–148. doi:10.1115/DETC2002/DAC-34048

Kanno, Y., and Fujita, S. (2018). Alternating direction method of multipliers for truss topology optimization with limited number of nodes: a cardinality-constrained second-order cone programming approach. *Optim. Eng.* 19 (2), 327–358. doi:10.1007/s11081-017-9372-3

Kawamura, H., Ohmori, H., and Kito, N. (2002). Truss topology optimization by a modified genetic algorithm. *Struct. Multidiscipl. Optim.* 23, 467–473. doi:10.1007/s00158-002-0208-0

Kiefer, J., and Wolfowitz, J. (1952). Stochastic estimation of the maximum of a regression function. *Ann. Math. Stat.* 23 (3), 462–466. doi:10.1214/aoms/1177729392

Kingma, D., and Ba, J. (2015). "Adam: a method for stochastic optimization," in 2014, Published as a conference paper at the 3rd International Conference for Learning Representations (San Diego). Arxiv:1412.6980.

Kipf, T. N., and Welling, M. (2016). *Variational graph auto-encoders*. Arxiv: 1611.07308. Available at: https://arxiv.org/abs/1611.07308.

Kipf, T. N., and Welling, M. (2017). "Semi-supervised classification with graph convolutional networks," in Proceedings of the 5th international conference on learning representations. Arxiv:1609.02907.

Kociecki, M., and Adeli, H. (2015). Shape optimization of free-form steel space-frame roof structures with complex geometries using evolutionary computing. *Eng. Appl. Artif. Intell.* 38, 168–182. ISSN 0952-1976. doi:10.1016/j.engappai.2014.10.012

Kupwiwat, C., and Yamamoto, K. (2020). Fundamental study on morphogenesis of shell structure using reinforcement learning. *Struct. I* 2020, 933–934. Architectural Institute of Japan. Available at: https://ci.nii.ac.jp/naid/200000462858/en/(Access March 17, 2022).

Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., et al. (2016). *Continuous control with deep reinforcement learning, ICLR*. Arxiv:1509.02971.

MacQueen, J. (1967). "Some methods for classification and analysis of multivariate observations," in 5-th Berkeley Symposium on Mathematical Statistics and Probability, 281–297.

Mai, T. H., Kang, J., and Lee, J. (2021). A machine learning-based surrogate model for optimization of truss structures with geometrically nonlinear behavior. *Finite Elem. Analysis Des.* 196, 103572. ISSN 0168-874X. doi:10.1016/j.finel.2021.103572

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., et al. (2013). *Playing Atari with deep reinforcement learning*. Arxiv:1312.5602. Available at: https://arxiv.org/abs/1312.5602.

Nair, V., and Hinton, G. E. (2010). *Rectified linear units improve restricted boltzmann machines*. Haifa, 807–814. Available at: https://dl.acm.org/citation.cfm.

Ohsaki, M., and Swan, C. (2002). "Topology and geometry optimization of trusses and frames," in *Recent advances in optimal structural design*.

Ohsaki, M. (1995). Genetic algorithm for topology optimization of trusses. *Comput. Struct.* 57 (2), 219–225. ISSN 0045-7949. doi:10.1016/0045-7949(94)00617-C

Ohsaki, M. (2010). *Optimization of finite dimensional structures*. 1st ed.. Boca Raton, FL: CRC Press. doi:10.1201/EBK1439820032

Puentes, L., Cagan, J., and McComb, C. (2021). Data-driven heuristic induction from human design behavior. *J. Comput. Inf. Sci. Eng.* 21 (2). doi:10.1115/1.4048425

Robbins, H., and Monro, S. (1951). A stochastic approximation method. *Ann. Math. Stat.* 22 (3), 400–407. doi:10.1214/aoms/1177729586

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol. Rev.* 65 (6), 386–408. doi:10.1037/h0042519

Ruder, S. (2016). *An overview of gradient descent optimization algorithms*. Arxiv: 1609.04747. Available at: https://arxiv.org/abs/1609.04747.

Sutton, R. S., and Andrew, G. B. (2018). *Reinforcement learning, an introduction*. 2nd ed. Cambridge, MA: The MIT Press. 9780262039246.

Topping, B. H. (1983). Shape optimization of skeletal structures: A review. *J. Struct. Eng. (N. Y. N. Y).* 109, 1933–1951. doi:10.1061/(asce)0733-9445(1983)109:8(1933)

Uhlenbeck, G. E., and Ornstein, S. L. (1930). On the theory of the Brownian motion. *Phys. Rev.* 36 (5), 823–841. doi:10.1103/PhysRev.36.823

van der Maaten, L., and Hinton, G. (2008). Visualizing data using t-SNE. *J. Mach. Learn. Res.* 9, 2579–2605. Available at: http://jmlr.org/papers/v9/vandermaaten08a.html (Access March 17, 2022).

Wang, D., Zhang, W. H., and Jiang, J. S. (2002). Truss shape optimization with multiple displacement constraints. *Comput. Methods Appl. Mech. Eng.* 191 (33), 3597–3612. ISSN 0045-7825. doi:10.1016/S0045-7825(02)00297-9

Yu, A., Palefsky-Smith, R., and Bedi, R. (2019). *Deep reinforcement learning for simulated autonomous vehicle control*. Stanford, CA, USA: Stanford University. Available at: http://vision.stanford.edu/teaching/cs231n/reports/2016/pdfs/112_Report.pdf (Access March 17, 2022).

Zhu, S., Ohsaki, M., Hayashi, K., and Guo, X. (2021). Machine-specified ground structures for topology optimization of binary trusses using graph embedding policy network. *Adv. Eng. Softw.* 159, 103032. ISSN 0965-9978. doi:10.1016/j.advengsoft.2021.103032

frontiers | Frontiers in Built Environment

TYPE Original Research
PUBLISHED 15 September 2022
DOI 10.3389/fbuil.2022.945944

OPEN ACCESS

EDITED BY
Husam Abu Hajar,
The University of Jordan, Jordan

REVIEWED BY
Bogdan Constantin Neagu,
Gheorghe Asachi Technical University of Iași, Romania
Manaf Al-Okaily,
Jadara University, Jordan
Olakunle Alao,
University College Dublin, Ireland
Adil Hammadi,
Curtin University, Australia

*CORRESPONDENCE
Bilal Abu-Salih,
b.abusalih@ju.edu.jo

SPECIALTY SECTION
This article was submitted to Computational Methods in Structural Engineering, a section of the journal Frontiers in Built Environment

RECEIVED 17 May 2022
ACCEPTED 23 August 2022
PUBLISHED 15 September 2022

CITATION
Coutinho K, Wongthongtham P, Abu-Salih B, Abu Saleh MA and Khairwal NK (2022), Carbon emission and cost of blockchain mining in a case of peer-to-peer energy trading. *Front. Built Environ.* 8:945944. doi: 10.3389/fbuil.2022.945944

Carbon emission and cost of blockchain mining in a case of peer-to-peer energy trading

Kevin Coutinho[1], Pornpit Wongthongtham[2], Bilal Abu-Salih[3]*, Mousa A. Abu Saleh[4,5] and Neeraj Kumari Khairwal[1]

[1]Alphabright Digital, Perth, WA, Australia, [2]Department of Mathematics and Data Analytics, The University of Notre Dame Australia, Fremantle, WA, Australia, [3]Department of Computer Science, The University of Jordan, Aljubeiha, Jordan, [4]Southern Association for the Preservation of the Environment and Society, Maan, Jordan, [5]Department of Environmental Studies, Greater Amman Municipality, Amman, Jordan

Blockchain is an emerging technology that has demonstrated great uptake potential in Peer-to-Peer (P2P) energy trading. The revolution of blockchain brings substantial benefits and innovation to sustainability energy transitions through P2P trading. Blockchain enables energy commodities to be traded. However, perceptions of this technology's impact on the environment and its associated costs have garnered recent adverse publicity. This paper aims to look at the linkages between blockchain technology and energy systems in terms of blockchain power consumption against blockchain advantage over renewable energy transitions *via* peer-to-peer energy trading. The amount of energy used and carbon released during the blockchain validation process is estimated, and the cost of blockchain is computed to assess its economic benefit in a peer-to-peer energy trading scenario. Real data from running peer-to-peer energy trading systems are used, and numerous insights on the transformation of peer-to-peer energy trading utilising various blockchain scaling methods are provided. Based on the analysis, this paper concludes that the cost of processing trading transactions is lower using blockchain than current coordination costs. Also, blockchain-based energy can be traded more frequently than current regulations allow in order to reap the full benefits of renewable energy. A secure blockchain-enabled P2P trading environment would lead to fair rates for energy providers and prosumers resulting in stimulating the renewable energy market.

KEYWORDS

blockchain, blockchain carbon emission, blockchain cost, peer-to-peer energy trading, renewable energy market

1 Introduction

As of March 2022, Australia had over 3.12 million solar photovoltaic (PV) installations with a total capacity of over 26 gigawatts (AustralianPVInstitute 2022). Since 2010, this sector has grown at a breakneck pace, and it has continued to do so. Because of Australia's large number of PV installations, distributed energy marketplaces

have sprung up to take advantage of the available energy. Peer-to-peer (P2P) energy trading, retail and wholesale energy, business-to-business (B2B) energy trading, power trading, and other marketplaces are being tested around the world. P2P trading, which allows people to trade electricity, is the most frequent market (Dudjak et al., 2021; Abu-Salih et al., 2022).

P2P energy trading has evolved as a new paradigm for energy management, allowing prosumers to sell excess energy and consumers to purchase inexpensive, locally generated renewable power (Wongthongtham, Morrison, and Liu 2019). The rising quantity of decentralized electricity production from solar installations opens the door to new electricity markets and allows consumers to share electricity (Christidis, et al., 2021a). P2P electricity markets give users the freedom to choose their energy source, such as investing in locally generated renewable energy.

Innovative technologies are being investigated in order to provide a safe and auditable P2P electricity market. Blockchain is a form of distributed ledger technology characterised as an encrypted, distributed, decentralised, transparent, trustless, and immutable database (Wongthongtham et al., 2021; Zheng et al., 2017; Kuo, Kim, and Ohno-Machado 2017a). The technology has demonstrated significant potential in P2P energy trading, with a rising number of commencement firms, projects, experiments, and research projects including blockchain as part of the operational plan. (Wu et al., 2022). Energy supply companies, technology developers, financial organisations, national governments, and academic institutions are among the more than 140 blockchain research projects and start-ups that are using blockchain for energy applications (Andoni et al., 2019; Shamsi and Paul 2020). The blockchain revolution promotes innovation and allows for a low-carbon transition and long-term sustainability (Juri et al., 2016). Furthermore, the use of blockchain technology in peer-to-peer electricity trading allows for a transition away from a highly centralised market dominated by a few businesses and towards a truly decentralised market dominated by microgrids according to Deloitte (Grewal-Carr and Marshall, 2016) and PWC (Hasse et al., 2016) research, by allowing energy commodities to be traded as digital assets.

Prosumers can sell their excess electricity directly to local customers *via* blockchain-enabled P2P energy trading, enabling mutually beneficial transactions. Customers spend less per kWh for their electricity and can use renewable energy sources without having to invest in the equipment. Consumers profit from this plan because they spend less on their electricity per kWh and benefit from renewable energy sources without owning the technology. Prosumers benefit from this plan because they make more than they otherwise would with feed-in tariffs. Because a battery can store untraded power, auctions for renewable electricity can create a dynamic market that benefits both prosumers and consumers. A more efficient market with reduced infrastructure costs also benefits electricity retailers and network providers. By eliminating market intermediaries, blockchain-based systems also enable anonymity and security to both prosumers and consumers. Agents with comparable energy demand profiles match supply and demand in real time, and trade is made through smart contracts (Damisa, Nwulu, and Siano 2022).

Blockchain technology and peer-to-peer trading have been extensively studied in the context of climate change policy, and have been implemented in a number of climate-related fields, including climate investment and carbon pricing (Abu-Salih et al., 2021; Alao and Paul 2021; Hua et al., 2022). As the cost of PV modules declines, the number of homes installing PV systems rises; P2P energy trading has become one of the most popular blockchain-based applications (Wang et al., 2021). However, questions about blockchain's energy use, carbon footprint, and cost have lately sparked debate. Blockchain's carbon footprint and cost come from its validation process, which necessitates specialised hardware with high computing power and large quantities of electricity. Public impressions of the technology's impact on the environment and its associated expenses have recently received critical attention. This distorts people's perceptions of blockchain's benefits. It is motivating to compare the amount of energy used to maintain blockchain with the amount of energy saved from the electrical grid as a result of P2P energy trading.

In this study, we look at how blockchain technology affects the environment in terms of energy usage, carbon emissions, and economic value. Additionally, numerous blockchain scalability approaches are examined in order to examine various options for the conversion of P2P energy trading platforms. We consider both the blockchain base layer model and the second layer solution, unlike other approaches that solely report on the base layer model. We examine blockchain mining energy consumption and compare it to household energy consumption as well as energy savings achieved through the implementation of peer-to-peer energy trading. The amount of carbon emitted by blockchain use is quantified and compared to the amount of carbon emitted by residential consumption, as well as the emission reductions achieved through peer-to-peer energy trading. Therefore, this paper aims to investigate the interconnections between blockchain technology and energy systems in terms of blockchain electricity consumption against blockchain benefit over renewable energy transitions through P2P energy trading. This research also offers numerous perspectives on the evolution of peer-to-peer energy trading through the use of various blockchain scalability solutions. The cost and advantages of blockchain-based P2P energy trading are calculated using real data from running P2P energy trading systems.

The remainder of the paper is organised as follows. Section 2 provides an overview background of the topic of this paper and the motivation to conduct this research. Section 3 summarises P2P energy trading case study (the RENeW Nexus project) and

discusses blockchain energy use and carbon footprint. Section 4 presents the results of the analysis along with a discussion. Finally, Section 5 concludes the work.

2 Background and motivation

This section furnishes a background on the topic of this paper. First, Blockchain and its benefits and impact in terms of P2P energy trading are discussed, followed by presenting Blockchain Trilemma and its impact on P2P energy trading. Lastly, relevant studies are examined and scrutinised, along with the motivation to carry out this research.

2.1 Blockchain for peer-to-peer energy trading

Blockchain is a type of distributed ledger technology that has been defined as a trust-less, distributed, decentralised, transparent, encrypted, and immutable database (Zheng et al., 2017; Kuo et al., 2017a). The ledger comprises a collection of all transactions (Christidis, Derbakova, et al., 2021). No central authority controls the transactions, which means that the control and influence are distributed through the network. The transactions are encrypted and stored in chronological order of blocks once a transaction is verified. In time, a chain of blocks is formed known as a "blockchain." A block cannot be altered once it is verified and added to the chain; hence, it is immutable. Blockchain uses a distributed and decentralised ledger for verifying and recording digital transactions, agreements, contracts, sales, and the like in a P2P network setting (Kakavand et al., 2017).

Blockchain-enabled P2P energy ensures prosumers and consumers to exchange energy excess and demand interchangeably on a virtual P2P basis. Energy is still delivered through a physical grid since demand and supply are regulated and controlled to maintain power system stability. Prosumers and consumers are linked directly in a trustless network *via* a trading platform that uses a distributed ledger. They have some mistrust towards one another. A consumer or prosumer is hesitant to enable others to alter trade transactions or to allow fraudulent transactions to take place. Prosumers are encouraged to send their surplus energy into the electrical grid, while consumers may buy energy from their peers *via* the grid at competitive prices. In blockchain-based peer-to-peer trading, specific criteria or conditions can be defined to allow transactions to be independently and automatically validated. To avoid fraudulent actions and duplicate spending, the amounts available in the ledger must be the same before and after each transaction. The flexibility of the power market model may be demonstrated by establishing a set of regulations.

Energy billing between prosumers and customers may be performed automatically using smart contracts in real time with blockchain and IoT support utilising smart meters. Without a central billing authority, payments can be sent directly to consumers and prosumers. In contrast to the standard 2-month billing period, payments can be made more often, such as every 5 s. Consumers and utility corporations both benefit from energy micropayments and pay-as-you-go alternatives (Mylrea and Gourisetti, 2017). It is verifiable and transparent to create a time-stamped trade transaction. Energy trade transactions are stored in multi-location data structures, giving the blockchain system stability and robustness. On the other side, boosting renewable energy quantities and developing flexible and efficient electricity trading platforms can help ensure energy supply security (Ometoruwa 2018).

Renewable energy is seldom gathered after it has been transported to the grid, where it is mixed with electricity generated by hydrocarbon power plants. In terms of energy blend procurement, a market share of renewable energy may be produced throughout the whole energy mix. Blockchain technology can be used to monitor the origins of the energy that is consumed or supplied. The transparency and provenance properties of blockchain technology may affect consumer behaviour, leading in a change in energy use between peak and off-peak hours when possible. Consumers might, for example, run energy-intensive appliances during peak solar energy production hours, incentivizing them to save energy.

In summary, possible blockchain effects and benefits for energy trading include 1) an effective, adaptable, and rigorous virtual P2P sustainable electricity trade system, 2) energy security, including cyber defence and electricity generation security, 3) energy lower cost through improving energy efficiency, and 4) sustainable development by enabling renewable energy production and reduced carbon solutions.

2.2 Blockchain trilemma and its implications in peer-to-peer energy trading

The challenge of P2P energy trading (also for the energy sector as a whole) could be around the sheer volume of data produced, e.g., the volume of P2P data transactions run concurrently. This is the case of information of kWh produced, kWh consumed, transactions, and so on being tracked in real-time. Energy generation or consumption every 30 min would create numerous ledger entries per second for the households (Wongthongtham et al., 2021).

The performance of blockchain technology is limited, as are the tradeoffs that surround the blockchain trilemma (scalability, security and decentralisation). According to (Chitchyan and Murkin, 2018; Ometoruwa 2018), the present blockchain technology can only solve two out of three of the trilemma's problems. The greatest long-term challenge to the viability of blockchain technology is regarded to be scalability (Abu-Salih

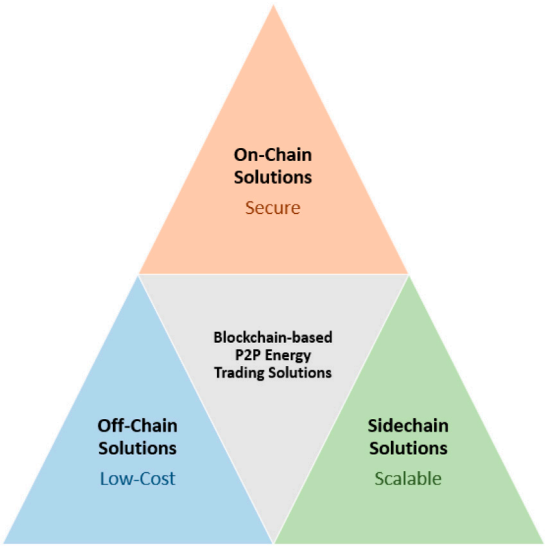


FIGURE 1
Blockchain-based P2P energy trading solutions.

2021). With great security and the benefit of decentralisation, the scalability issue for blockchain technology is compromised. Even though there are techniques to address the scalability problem, security and/or decentralisation may be harmed. In the trilemma view, it is critical to investigate the consequences of blockchain-based energy trade.

In this paper, a combination of scaling solutions, which are complimentary, is modelled for P2P energy trading consisting of on-chain, off-chain, and side-chain solutions as depicted in Figure 1. Other solutions are not covered in this paper which are sharding, alternative consensus models, and directed acyclic graph. The exclusion is due to its trade-offs being compromised. Sharding reduces the load on network nodes by partitioning the blockchain into smaller sub-chains, so-called shards. This would result in more transactions being processed and verified simultaneously. The throughput increases as the mining network expands, this the transactions complete faster. However, there are some challenges such as the high-security standard, shards communication method, shards synchronisation and consensus model (Andresen 2015). Theoretically, with sharding, there is a security risk for invalid transactions being recorded by a subset of malicious nodes which might then remain unchecked. There is also a risk for validator collusion in the consensus model. To mitigate the risks, each transaction should be processed by a sufficient number of nodes. The vulnerability depends on the sharding techniques and blockchain platforms used. Complications would be challenged over cross-shard communications and synchronisation over two or more shards. Thus, sharding, in general, traded off scalability issues for security risk.

In the subsequent sections, blockchain-based three approaches, viz., on-chain, off-chain and side-chain are discussed in detail. P2P electricity trading transactions can be expedited using blockchain technology. The P2P trading cost can also be reduced by off-chain and side-chain techniques. Further, the blockchain-based P2P trading solution can be customised by setting threshold on trading time period, number of trading transcations, etc. Hence, let us discuss each solution in detail.

2.2.1 On-chain solution

The Bitcoin blockchain was designed with a 1 megabyte block size restriction. Setting a block limit was designed to make sure that a malicious attacker could not effectively shut down the network by launching a denial of service attack. A denial of service attack without a block size restriction might hypothetically clog the network by sending very massive bogus data to be mined. Block sizes more than the block size limit are automatically rejected by the network when using a block size restriction.

When the blockchain requires many transactions per second, the transaction blocks may fill faster than the transactions can be processed. This causes transaction congestion and elevated transaction fees. It was then proposed to increase the block size (Andresen 2015; Garzik 2015) in order to bring down transaction fees and increase the transaction throughput. The proposed change to the block size required a software hard-fork. This meant that nodes on the network running the proposed changes would not be compatible with nodes that had not included the software upgrade. Another scaling issue arises with increasing the block size. For certain transactions, signature hash operations scale quadratically. This means that doubling the transaction size can double both the number of signature operations and the amount of data that needs to be verified. This is to be discussed further in the segregated witness protocol. The proposed solution was highly divisive and Bitcoin was forked into two different projects. The forked project, Bitcoin Cash, used the larger 8-megabyte block size as a solution to increase transaction speed and has risen since the block size to 32 megabytes.

Segregated witness (SegWit) (Shaolin, 2008) is a protocol that increased both the security and block size without requiring a hard-fork of the current Bitcoin Cash protocol (Wuille et al., 2015). All transactions on the Bitcoin Cash network are identified by a 64-digit transaction identifier. Without the SegWit protocol, it is possible for the transaction identifier to be altered by miners or relay nodes, without altering the transaction amount or wallet addresses. This is called scriptSig malleability (Song 2017). SegWit works by segregating the malleable witness parts of the identifier from the transaction identifier and introducing a new witness identifier.

Nodes not upgraded to use the SegWit protocol that still enforce the 1-megabyte block size limit can still confirm transactions by processing the transaction identifier and

completely ignoring the new witness identifier. As transaction identifiers produced by the SegWit protocol do not include any of the witness data structure, mining nodes enforcing the 1-megabyte block size can fit more transactions into the block effectively increasing the transaction throughput. SegWit2x has been proposed (Garzik 2015) to force the activation of SegWit and increase the block size to 2 Mb.

The quadratic scaling of the increased block size is targeted by changing the hash function that calculates the transaction signature. The SegWit protocol ensures that any byte in a transaction is only calculated, at most, twice. In summary, SegWit increases security by fixing malleability, increases block size without requiring a hard-fork and ensures backwards compatibility, allows greater transaction throughput with increased security, makes unconfirmed transactions being able to be spent, and makes it easier in creating a lightning network, smart contracts and scripts.

2.2.2 Off-chain solution

The Off-chain transactions allow for low-cost instantaneous transactions across a network of participants. Off-chain transactions use smart contracts and multi-signature wallets to provide P2P transactions between participants. Off-chain transactions are often referred to as a second layer solution. This is because no fundamental changes are made to the actual blockchain, i.e., base layer or the first layer. Rather the off-chain transactions work, on the top of the main blockchain, as second layer. Therefore, the transactions between participants are not recorded on the blockchain while the payment channel is open. Only the final settlement transaction between parties is recorded to the blockchain.

Off-chain transactions are made up of two parties who first bind a committed transaction to a multi-signature wallet before entering a smart contract. The transaction is signed and sent to the blockchain for consensus and confirmation. Either party or both can agree to enforce the contract at a later date. Both parties can execute transactions with each other by first committing to the blockchain with a "commitment payment." The parties can then conduct business as usual as long as the value of the pledge payment is not exceeded. Both parties must work together and agree on business transactions. As long as both parties have funds available within the payment channel, they can transmit as many payments to their counter-party as they like.

The parties can opt to safely end the channel after the last transaction, and the contract will pay the participants based on the final balance. The final transaction balance state is the only balance written to and broadcast on the blockchain, and the contract is completed. Transactions between participants not directly connected *via* a payment channel, but rather connected by proxy, are possible using off-chain networks. Off-chain networks (lightning network or Raiden network) are interconnected networks of bidirectional payment channels. Mediated payments are first secured by the sender by hashing the payment with an encrypted secret known as a preimage, known only to the recipient, locked into the contract and valid only after a particular date. This is known as a Hashed Timelock (Bowe and Hopwood 2017). Payments are securely routed by making multiple hops through payment channels connecting the participants of the payments. Intermediate network participants can claim a transaction fee and are rewarded as such by passing on the transaction to the next node along the route and by decrementing the time-lock. The final recipient can claim the transaction balance from their payment channel by producing the secret and passing it back along the chain so all nodes can claim their balance.

The Hashed Timelock Contract is revoked and the payment is claimed when the recipient can produce the secret preimage within a certain timeframe. If the balance of the transactions is not claimed by producing the preimage, the participants are refunded their bonded payment. The Bitcoin protocol is called the lightning network (Poon and Dryja 2016) and the Ethereum implementation is called the Raiden network. The benefits of off-chain transactions include 1) close to Instantaneous non-revokable transactions are possible; 2) Due to very low transaction fees, micropayments are possible; 3) Escrow services can be achieved without a third party; 4) Cross-chain payments, also known as atomic swaps are possible using Hashed Timelock Contract (Herlihy 2018); and 5) Transactions can be made anonymously (Green and Miers 2017). Limitations of off-chain transactions include 1) If a payment channel closes unexpectedly, participants are refunded their stake and the transactions are not processed; 2) Bitcoin requires SegWit to be deployed on participating nodes in order to use the Lightning Network; 3) Hardforks in the underlying blockchain may break second layer solutions; 4) Payment routing in second layer solutions are not part of the protocol and are still very early in development; and 5) Bonding payments in a smart contract ties up capital.

2.2.3 Side-chain solution

Side chains are an attempt to scale the enforcement and execution of smart contracts (Poon and Buterin 2017). The side chains are incentivised to continue working faithfully and autonomously by network fees enforced by the parent blockchain with bonded fraud proofs. Just like state channels, contracts are required to be bonded to the parent chain by staking a coin or token in an on-chain smart contract.

After the contract is updated or closed, participants must produce a Merkle tree proof in order for them to move their final transactions back to the main chain and settle their balance. Side chains can be used not only for payments, but extended to computation and, therefore, can be used for off-chain smart contracts. Side chains have the ability to form both nested chains and web-chains where transactions can take place across side chains (Todd 2014). The computation posed in the smart contract is broken up and distributed using mapReduce (Dean

and Ghemawat 2008). Side chains do not need to use PoW and are encouraged to use PoS.

In the event of fraudulent or erroneous behaviour, the blockchain will punish the offending party. Side chains are intended to work in tandem with other scaling methods such as Raiden and the Lightning Network and are seen as complementary. In essence, side-chain lowers computation expenses and allows smart contracts to be executed quickly. However, if a large number of participants attempt to quit the contract at the same time, the network may become overwhelmed and unable to process all of the contracts.

2.3 Research motivation

Depending on their energy demand and production balance, individuals input energy into the electrical grid or withdraw energy from the grid in the classic form of energy trading. At a feed-in-tariff rate, electricity retailers compensate households for surplus electricity produced by rooftop solar. Prosumers and consumers trade energy directly *via* P2P energy trading. Prosumers are rewarded for providing renewable energy to the grid by receiving compensation in the form of a crypto-currency or a crypto-utility token. Consumers are also motivated since they can use crypto-currency to acquire energy from their neighbors at a low tariff price. The tariff price is determined by the availability and demand of renewable energy sources. A cryptocurrency trading market or a set token supply and the changeable quantity of renewable energy could determine the tariff price. When the demand for energy on the P2P network exceeds the available supply, the energy distributor feeds the network at a higher price than the market price. This further incentivises network participants to generate excess renewable energy in order to lower energy prices and reward prosumers. It is vital to remember that prosumers can be both energy users and producers.

However, there are worries about blockchain's carbon footprint (Jiang et al., 2021; Erdogan, Ahmed, and Sarkodie 2022; Truby et al., 2022), which stems from its validation process (Mora et al., 2018a), which necessitates specialised hardware with high computing power and massive quantities of electricity. Challenges also arise from the transaction charges required to preserve the decentralised blockchain's integrity. Few studies have accepted mathematical proof and predicted that blockchain emissions could contribute to global warming (Mora et al., 2018a; Alofi et al., 2022; Dogan, Majeed, and Luni 2022) and that it uses more energy to produce an equivalent market value as mineral mining (Krause and Tolaymat 2018). Foteinis (2018) explored the blockchain carbon footprint. They estimated the global carbon emissions of 43.9 Mt in 2017 for Bitcoin and Ethereum mining using the life-cycle impact-assessment technique (Hileman and Rauchs 2017). According to Digiconomist's Bitcoin and Ethereum energy consumption indexes, Bitcoin consumes 55.76 TWh and Ethereum consumes 8.38 TWh of electricity yearly, equating to 0.25% and 0.04% of global electricity consumption, respectively (Digiconomist 2019a; Digiconomist 2019b).

Several studies attempted to conceptualise the linkage between Blockchain technology and carbon emission in P2P energy trading (Lu et al., 2022). For example, He et al. (2020) developed a Blockchain-based transactive energy and carbon market solutions for distributed solar systems. However, the detailed carbon market is limited and its mechanism is not addressed. Hua et al. (2020) presented an Ethereum blockchain-based decentralised energy trading platform, and its suggested pricing scheme achieves regional energy balance and carbon emission reduction. However, their approach focuses on a day-ahead scheduling whereby the outcomes may violate carbon emission limits and operating constraints in the intra-day phase (Zhong et al., 2022).

We assess the environmental impact of blockchain technology in terms of short-term energy consumption, carbon emissions, and economic value in this research. Furthermore, various blockchain scalability methods are investigated in order to explore multiple choices for the conversion of P2P energy trading platforms (Luu et al., 2016; Ledger 2018). However, salability has not been confirmed (Teng et al., 2021). Unlike many existing studies that only report on the base layer model (Jiang et al., 2021; Lu et al., 2022), we consider both the blockchain base layer model and the second layer solution. We assess blockchain mining energy consumption and compare it to residential energy usage as well as energy savings *via* P2P energy trading adoption. The amount of emissions emitted by blockchain use is calculated and compared to the amount of carbon emitted by residential use and the emission savings *via* peer-to-peer energy trading.

3 RENeW nexus project: Peer-to-peer energy trading case study

Through the RENeW Nexus project, the Western Australian government is the first to allow peer-to-peer energy trading across the grid. The Smart Cities and Suburbs Program of the Australian Government is funding the project (Sundararajan 2017). The project's goal is to look into the value and efficiency of peer-to-peer energy trading in Fremantle. As part of the RENeW Nexus initiative, electric smart metres have been installed in 50 homes throughout Fremantle for a P2P energy trading trial. Energy data is collected at regular intervals (down to 5 s) and sent *via* the Internet of Things. The energy received from the grid, the energy given to the grid, the energy produced by rooftop PV, the energy used in homes, and so on are all included.

Throughout the trial's research and development stage, the RENeW Nexus project employs a private consortium blockchain with no transaction fees. However, because the RENeW Nexus

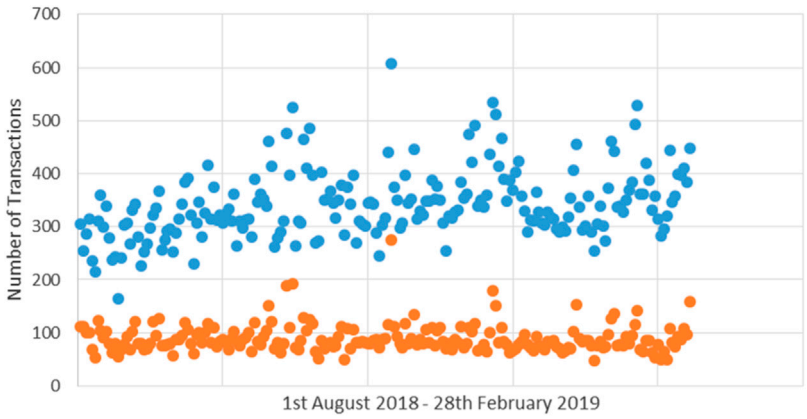


FIGURE 2
The minimum and highest number of transactions between 01/08/2019 and 28/02/2019.

may switch to a public network once the trial is over, the cost of using a public blockchain is estimated for the purposes of this study. Between 1 August 2018 and 28 February 2019, data on energy production and consumption were collected at 30-min intervals from 50 participants' residences. The study's purpose is to compare and contrast the amount of transactions, energy consumption, and carbon emissions connected with blockchain mining, as well as transaction costs, across the two blockchain architectures (i.e., base layer model and second layer model).

The energy balance between prosumers and consumers is determined by energy supply and demand, as shown by the number of transactions in the RENeW Nexus project. There is a high level of energy consumption when the amount of energy imported exceeds the amount of energy exported. Significant levels of energy supply are indicated when the amount of energy imported is less than the amount of energy exported. The number of transactions will be calculated based on demand and supply matching. The min and max number of transactions (No.TX) are shown in Eqs 1, 2, respectively:

$$min\ No.TX = \sum_{i=1}^{61} \underset{D==S}{min}\, k_t \tag{1}$$

$$max\ No.TX = \sum_{i=1}^{61} \underset{D==S}{max}\, k_t \tag{2}$$

Where D indicates total electricity demand, S indicates total electricity supply, k_t indicates the estimated transactions number, t denotes the day time slots, and i denotes number of days (i.e., 61 days).

In the event that energy demand exceeds supply, energy demand will be prioritised and matched with supply. When you rank energy demand in ascending order, you get the most transactions, and when you rank it in descending order, you get the least. In the situation of a demand gap between supply and demand, energy supply will be ranked and matched with demand. The highest number of transactions is obtained by ranking energy supply in descending order, while the smallest number of transactions is obtained by ranking energy supply in ascending order. Figure 2 depicts the minimum and greatest number of transactions during a period of 7 months. According to the distribution of transaction numbers, the lowest number of transactions is typically between 63 and 116, while the largest number of transactions is typically between 276 and 407 transactions.

3.1 Carbon footprint and energy use of blockchain

In a blockchain, miners are compensated for validating transactions and preserving the network's integrity, removing the need for a trusted third party. Blockchain is used to develop a decentralised administrative data protocol. On the other hand, the validation process demands a great deal of processing and electricity. To effectively calculate the carbon footprint of blockchains, we must first comprehend their power consumption and the carbon emissions produced by their mining activities. The carbon footprint of blockchain mining energy consumption is contrasted to the carbon footprint of electricity acquired from the grid as a result of peer-to-peer energy trading.

The technique used to calculate the amount of electricity and energy required to mine blockchain in this study follows the steps outlined in (Krause and Tolaymat 2018). To calculate the amount of power required by the entire blockchain network, the daily network hash rate was raised by the power efficiency of the mining rig (computer hardware), as shown in Eq. 3. The blockchain energy demand was then calculated by multiplying the blockchain energy requirement by the time it takes to generate a block, as shown in Eq. 4 (block completion time).

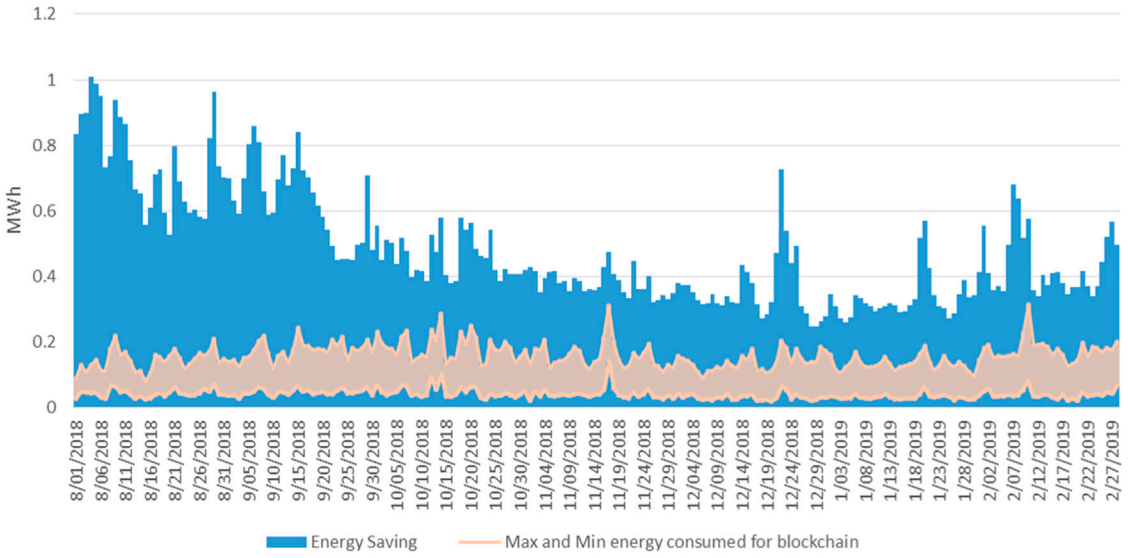


FIGURE 3
Power expended to mine blockchain compared with households' energy import (energy saving).

As shown in Eq. 5, the transactions in the RENeW Nexus project were multiplied by the energy required to mine blockchain, and then divided by the total transactions in the entire blockchain network (5).

$$Power\ Required\,(MW) = Network\ Hashrate\left(\frac{THash}{second}\right) \times Rig\ Power\ Efficiency\left(\frac{Joules}{MHash}\right) \tag{3}$$

$$Energy\ Required\,(MWh) = Power\ Required\,(MW) \times Block\ Time\,(hour) \tag{4}$$

$$Energy\ Consumption\,(MWh) = \frac{Households'\ No.TX \times Energy\ Required\,(MWh)}{Total\ blockchain\ network\ NO.TX} \tag{5}$$

In general, a blockchain network's energy consumption is the bare minimum. It is worth mentioning that the energy consumption estimates for blockchain mining do not include the energy used to cool and maintain mining equipment.

3.2 Blockchain cost

The Ethereum blockchain and its related fees were used to calculate the cost of blockchain transactions on a public blockchain. The Ethereum blockchain is a public network that uses Ether (ETH) as its cryptocurrency to pay for network transactions. The "gas" unit is used by the Ethereum blockchain network to measure the computing work involved in completing transactions or smart contracts. The price of power and processing hardware used to run the codes and complete the transactions is referred to as computational work. The cost per unit of gas, or gas price, is a critical component in the Ethereum transaction execution process. Based on the current exchange rate, the gas price in Wei units is converted into ETH units, and subsequently into USD. A total of 21,000 units of gas were used in this article for a single transaction. The expected number of transactions is then multiplied by the blockchain cost. Eqs 6, 7 demonstrate how to calculate the cost of a blockchain

$$Cost\,(USD) = Ether\ price\,(USD) \times Gas\ amount \times Gas\ price\ in\ ether\left(\frac{wei}{10^{18}}\right) \tag{6}$$

$$Total\ blockchain\ cost = \sum_{i=1}^{n} NO.TX \times Cost_i \tag{7}$$

Where i represents number of days (i.e., n days).

4 Results and discussion

Power is exchanged in a P2P fashion on any electricity trading platform, allowing customers to adjust their consumption to meet desired objectives such as lowering electricity bills, reducing grid imports, and lowering carbon emissions. Figure 3 compares the energy saved (households' energy import) to the energy consumed to mine blockchain grids on a daily basis.

The energy consumed to mine blockchain, on an average, is less than half the energy required for households' import, nearly 0.2 MWh. At few instances, the energy import for households' (~1 MWh) is more than five times the energy consumed for blockchain mining (<0.2 MWh). Not a single sample depicts more power consumption for blockchain mining in comparison to households' import. Thus, it can be inferred that blockchain-

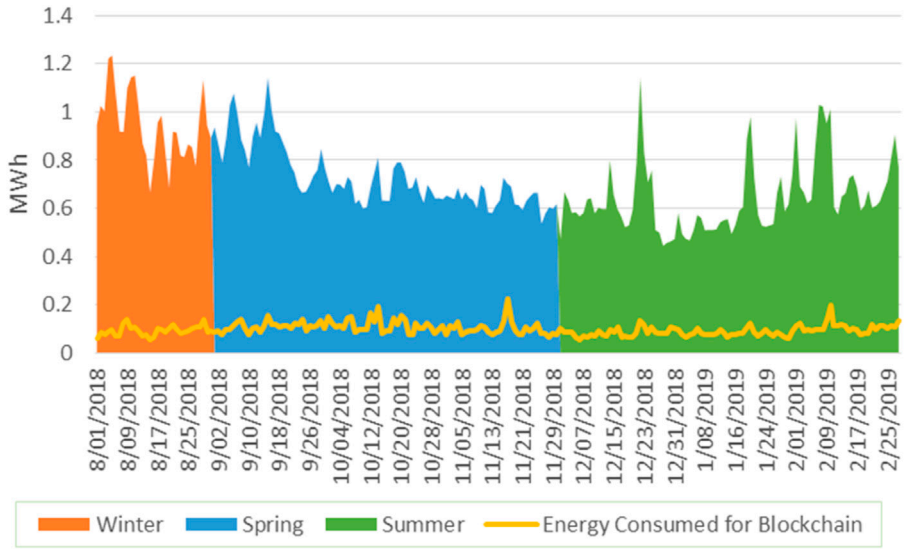


FIGURE 4
Blockchain's and households' energy consumption in different seasons.

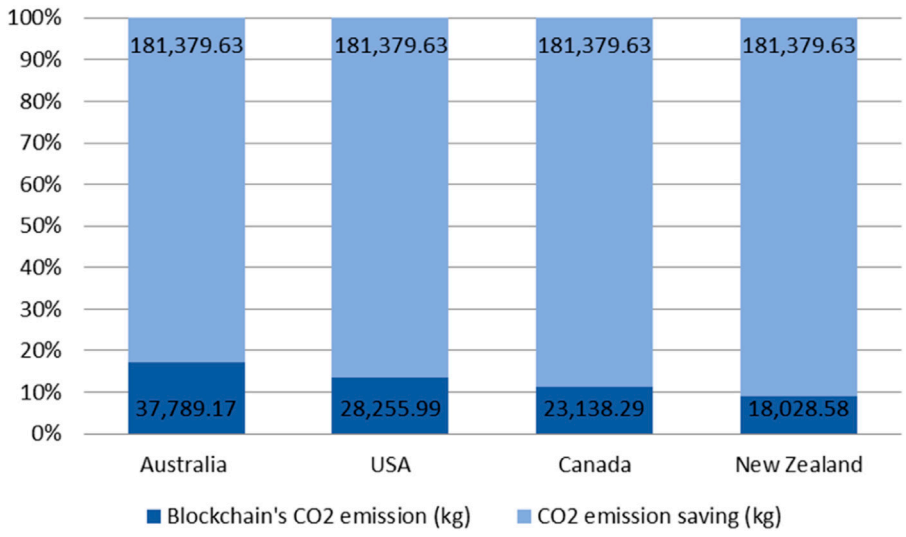


FIGURE 5
Carbon emissions from blockchain mining in several countries are compared to carbon emission savings from Australia in the P2P energy trading system.

based P2P trading is highly efficient as the households' energy import is comparatively larger than the power consumed for blockchain mining.

Figure 4 compares the average energy consumption of the blockchain to that of the ordinary household. While residential energy usage has seasonal patterns, blockchain's energy consumption is usually consistent. In comparison to home energy consumption, which is 1,234 kWh (average of 724 kWh) and a minimum of 446 kWh, blockchain consumers use a maximum of 226 kWh (average of 100 kWh) and a minimum of 54 kWh.

Mining can take place anywhere as long as there are electrical power available, an internet connection, and the appropriate

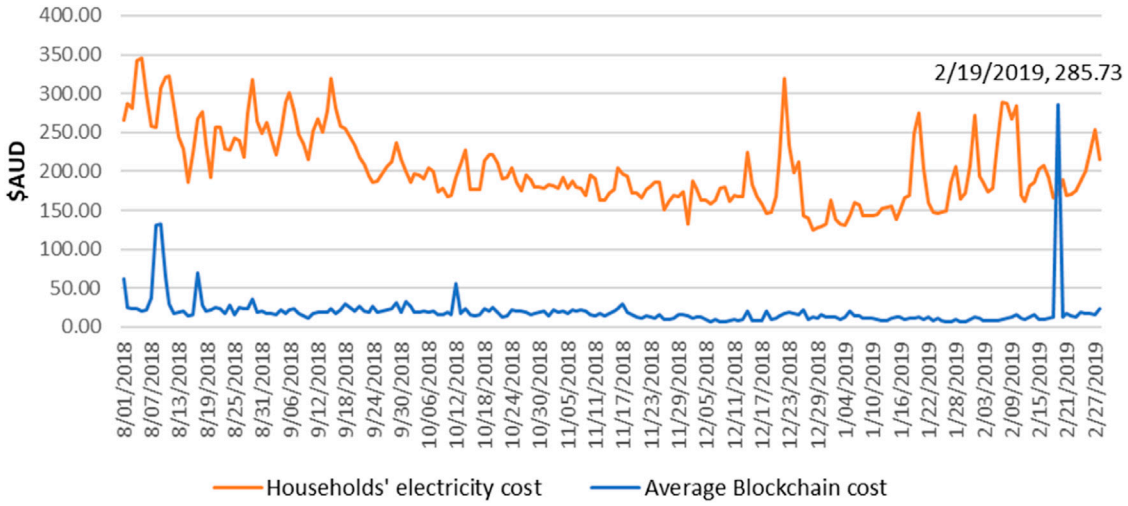


FIGURE 6
Average blockchain cost comparing with households' electricity cost.

mining rig machines. Hence miners can locate anywhere; ideally, they would be where electricity cost is cheap. It would be sustainable if miners locate where electricity is being generated by renewable energy, e.g., in Iceland, to take advantage of cleaner energy. As the miner location cannot be determined, the designation of emission in this paper is estimated using emission factors from selected countries (i.e., Australia, United States, Canada, and New Zealand). The amount of energy required is locality independent, but environmental impacts are country-specific can depend on the primary energy source.

Blockchain mined in Australia would generate approximately double the amount of carbon compared to the amount generated in New Zealand. Figure 5 shows the percentage of carbon emission from blockchain mining in the four countries compared with carbon emission saving from Australia in the P2P energy trading model.

Household power bills are computed using a normal Western Australia tariff rate of 28 cents per unit. Figure 6 compares the average blockchain cost to the cost of electricity in a home. It is worth noting that on 19 February 2019, the average Ethereum blockchain cost climbed at a rate that was even higher than the overall cost of residential power on that day. There were four transactions on the same day with a coding problem from the same wallet address, therefore, this could be a mistake by an Ethereum developer. One spent nearly half a million dollars in transaction costs in just 4 h, making it the most costly transaction fees ever paid for a cryptocurrency payment (Williams, 2019). It is thought that the developer mixed up the transaction value and the gas price (Sukhomlinova 2019).

Because the cost of blockchain can be quite high at times, the second layer technique is an excellent option. The block finality time (i.e., the time it takes to add transactions to the chain) can be set to avoid incurring the cost of computation for each transaction added to the blockchain. As a result, a significant amount of money is saved.

For monthly block finality time, Table 1 displays the minimum and maximum base layer blockchain costs against second layer solution costs. The second layer approach saves AUD $1,749.67 on the low end and AUD $5,682.00 on the high end.

Figure 7 shows the cost reductions from P2P trade with a base layer blockchain versus a second layer solution in terms of household electricity costs. When compared to the savings in family electricity from P2P trade, the second layer solution is virtually free.

Based on above graphical comparisons, blockchain-based P2P energy trading seems to be an efficient solution. However, for blockchain-based energy trade, two issues are a matter of grave concern, viz., the carbon footprint of blockchain technology and its accompanying costs. The cost and utility of implementing blockchain technology for peer-to-peer energy trading are evaluated. The energy saved in upscale P2P energy trade is compared to the energy consumed to support blockchain systems. Blockchain accounted for almost 20% of energy savings and 14% of residential energy consumption.

In addition, comparison among carbon emission savings from peer-to-peer trade and blockchain carbon emissions is accomplished. Mining activities can be carried out anywhere in the world, but it is preferable if they are carried out in locations where power is inexpensive or renewable energy sources are used to create electricity. If energy is mined in Australia, blockchain produces a 20% reduction in carbon emissions. In the United States, 15%, in Canada, 12%, and in New Zealand, 9%.

TABLE 1 A comparison in terms of the expenses of a base layer blockchain solution to the costs of a second layer blockchain solution.

Month	Min base layer cost ($AUD)	Max base layer cost ($AUD)	Min second layer cost ($AUD)	Max second layer cost ($AUD)	Min saving ($AUD)	Max saving ($AUD)
August	503.19	1,614.13	6.53	203.80	496.66	1,410.32
September	266.19	967.56	4.68	48.33	261.51	919.24
October	272.95	1,007.14	4.44	80.48	268.51	926.65
November	219.42	800.40	4.02	41.27	215.39	759.13
December	145.28	623.73	2.25	37.22	143.02	586.51
January	136.02	550.06	2.06	34.34	133.96	515.72
February	233.17	1,034.36	2.56	469.86	230.61	564.50

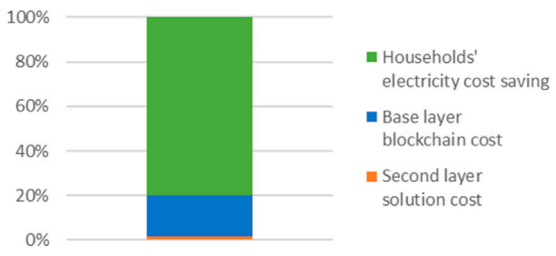


FIGURE 7
Comparison of the cost of electricity saved by homes, the cost of the base layer blockchain, and the cost of the second layer solution in percentages.

We evaluate the costs of two models to examine the cost of blockchain (i.e., base layer blockchain model and second layer solution model). When compared to the base layer blockchain, the second layer solution model delivers a large cost savings (almost 90%). As a result, the second layer solution costs about 1% of current family electricity prices, whereas the base layer blockchain costs about 23% of current household electricity expenses.

Thus, it can be inferred that blockchain energy consumption cost is substantially less compared to the cost of electricity saved using blockchain-based P2P energy trading. This gap can be further widened by employing second layer blockchain solutions. Carbon emission can also be significantly reduced by utilizing blockchain technology. Hence, blockchain-based P2P energy trading considerably reduces electricity cost while saving carbon emission.

5 Limitations and future work

Blockchain is a very efficient technology in reducing electricity cost and carbon emission for P2P energy trading. Yet few constraints of technology and domain hinder the huge potential of blockchain in P2P electricity trading. The privacy of the users on blockchain is preserved as the users are identified using pseudonyms, such as public key, instead of real identities. But, pattern of user's activity such as electricity consumption in different season, energy cost, energy import, etc. can be inferred to reveal more sensitive information (Wu et al., 2022). Blockchain trilemma of scalability, security and decentralization still stands as a challenge. While on-chain solutions are decentralized and secure, they lack scalability and speed to meet the demand of P2P electricity trading market. Also, on-chain transactions are costlier than off-chain transaction. Second layer solutions (off-chain and sidechain) provide scalability but are less decentralized. Concentration of power in blockchain can result in malicious activites such as inflated electricity prices. P2P electricity trading is not fully commercial and is restricted by the country regulations. Moreover, professional regulators dominate the B2B and B2C domain of the electricity market. For novice prosumers, it is a challenge to coordinate with distributed actors and well-established regulators to increase market participation while adhering to the regulations and competition in the traditional market. Overcoming these limitations can result in a secure, scalable, cost-efficient P2P energy trading. Further, it is worth indicating that Ethereum are computationally and energy-intensive with sometimes slow and expensive transactions (Buterin 2016; Chan and Olmsted 2017). In the future work we will incporporate newer and more computationally efficient blockchains which can solve this problem, such as Algorand (Gilad et al., 2017).

6 Conclusion

Blockchain is a major distributed ledger technology with a lot of promise for P2P energy trade. This paper investigates and examines blockchain-based P2P energy trade, given the growing interest in leveraging blockchain and distributed technology for P2P trading. P2P energy trading that is scalable, robust, and secure is still in its infancy, but blockchain technology promises a prospective breakthrough for energy trade and distribution.

Important economic, legal, and regulatory aspects must be examined to assure the viability of blockchain-based energy trade. In this paper, economic aspects of blockchain-based P2P trading have been examined. Power expended to mine blockchain is 20% of energy savings and accounts for almost 14% of residential energy consumption. Although carbon footprint of blockchain technology and its accompanying costs have been issues of concern, blockchain produces a 20% reduction in carbon emissions when energy is mined in Australia. This reduction can be further increased at locations with renewable energy resources. Also, the blockchain average cost is 23% compared to the households' electricity expenses. Further, this cost can be drastically diminished to 1% by employing second layer blockchain solutions, i.e., off-chain solutions and sidechain solutions. Thus, blockchain-based P2P energy trading is a secure, cost-efficient, environment-friendly solution resulting in economical electricity with reduced carbon emissions.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

PW contributed to conception and design of the study. KC and BA-S organized the database. PW performed the statistical analysis. PW, KC, BA-S, MA, and NK wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

The project is funded by the Australian Government through the Smart Cities and Suburbs Program.

Conflict of interest

KC and NK were employed by Alphabright Digital.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor HH declared a shared affiliation with the author BA at the time of review.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abu-Salih, B., Al-Tawil, M., Aljarah, I., Faris, H., Wongthongtham, P., Chan, K., et al. (2021). Relational learning analysis of social politics using knowledge graph embedding. *Data Min. Knowl. Discov.* 35, 1497–1536. doi:10.1007/s10618-021-00760-w

Abu-Salih, B. (2021). Domain-specific knowledge graphs: A survey. *J. Netw. Comput. Appl.* 185, 103076. doi:10.1016/j.jnca.2021.103076

Abu-Salih, B., Wongthongtham, P., Morrison, G., Coutinho, K., Al-Okaily, M., and Huneiti, A. (2022). 'Short-term renewable energy consumption and generation forecasting: A case study of western Australia. *Heliyon* 8, e09152. doi:10.1016/j.heliyon.2022.e09152

Alao, O., and Paul, C. (2021). "Towards a blockchain weather derivative financial instrument for hedging volumetric risks of solar power producers," in 2021 IEEE Madrid PowerTech, Madrid, Spain, 28 June 2021 - 02 July 2021 (Piscataway, New Jersey, United States: IEEE), 1–6.

Alofi, A., Bokhari, M., Bahsoon, R., and Hendley, B. (2022). 'Optimizing the energy consumption of blockchain-based systems using evolutionary algorithms: A new problem formulation. *IEEE Trans. Sustain. Comput.*, 1. doi:10.1109/tsusc.2022.3160491

Andoni, M., Robu, V., Flynn, D., Abram, S., Geach, D., Jenkins, D., et al. (2019). 'Blockchain technology in the energy sector: A systematic review of challenges and opportunities. *Renew. Sustain. Energy Rev.* 100, 143–174. doi:10.1016/j.rser.2018.10.014

Andresen, G. (2015). Increase maximum block size. Available at: https://github.com/bitcoin/bips/wiki/Comments:BIP-0101.

AustralianPVInstitute (2022). *'Australian PV market since april 2001', Australian PV institute (APVI) solar map*. Australia: funded by the Australian Renewable Energy Agency.

Bowe, S., and Hopwood, D. (2017). Hashed time-locked contract transactions. Available at: https://github.com/bitcoin/bips/wiki/Comments:BIP-0199.

Buterin, V. (2016). 'Ethereum: Platform review', opportunities and challenges for private and consortium blockchains.

Chan, W., and Olmsted, A. (2017). "Ethereum transaction graph analysis," in 2017 12th international conference for internet technology and secured transactions (ICITST), Cambridge, UK, 11-14 December 2017 (Piscataway, New Jersey, United States: IEEE), 498–500.

Chitchyan, R., and Murkin, J. (2018). Review of blockchain technology and its expectations: Case of the energy sector. arXiv preprint arXiv:1803.03567.

Christidis, K., Anna, D., Gaur, N., Jayachandran, P., and Srinivasan, M. (2021). "Using ledger sensors to enable contextual contracts across various enterprise blockchain applications,". Google Patents. United States;US11188883B2.

Christidis, K., Sikeridis, D., Wang, Y., and Devetsikiotis, M. (2021a). 'A framework for designing and evaluating realistic blockchain-based local energy markets. *Appl. Energy* 281, 115963. doi:10.1016/j.apenergy.2020.115963

Damisa, U., Nwulu, N., and Siano, P. (2022). 'Towards blockchain-based energy trading: A smart contract implementation of energy double auction and spinning reserve trading. *Energies* 15, 4084. doi:10.3390/en15114084

Dean, J., and Ghemawat, S. (2008). 'Mapreduce: Simplified data processing on large clusters. *Commun. ACM* 51, 107–113. doi:10.1145/1327452.1327492

Digiconomist (2019a). Bitcoin energy consumption index. Digiconomist.net. Available at: https://digiconomist.net/bitcoin-energy-consumption (Accessed April 24, 2022).

Digiconomist (2019b). Ethereum energy consumption index (beta). Digiconomist.net. Available at: https://digiconomist.net/ethereum-energy-consumption (Accessed April 24, 2019).

Dogan, E., Tariq Majeed, M., and Luni, T. (2022). 'Are clean energy and carbon emission allowances caused by bitcoin? A novel time-varying method. *J. Clean. Prod.* 347, 131089. doi:10.1016/j.jclepro.2022.131089

Dudjak, V., Neves, D., Alskaif, T., Khadem, S., Pena-Bello, A., Saggese, P., et al. (2021). Impact of local energy markets integration in power systems layer: A comprehensive review. *Appl. Energy* 301, 117434. doi:10.1016/j.apenergy.2021.117434

Erdogan, S., Ahmed, M., and Samuel, A. (2022). Analyzing asymmetric effects of cryptocurrency demand on environmental sustainability. *Environ. Sci. Pollut. Res.* 29, 31723–31733. doi:10.1007/s11356-021-17998-y

Foteinis, S. (2018). Bitcoin's alarming carbon footprint. *Nature* 554, 169–69. doi:10.1038/d41586-018-01625-x

Garzik, J. (2015). Block size increase to 2MB. *Bitcoin Improv. Propos.* 102.

Gilad, Y., Hemo, R., Micali, S., Vlachos, G., and Zeldovich, N. (2017). "Algorand: Scaling byzantine agreements for cryptocurrencies," in *Proceedings of the 26th symposium on operating systems principles* (New York, New York, United States: Association for Computing Machinery), 51–68.

Green, M., and Miers, I. (2017). "Bolt: Anonymous payment channels for decentralized currencies," in Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, Dallas Texas USA, 30 October 2017-3 November 2017 (ACM), 473–489.

Grewal-Carr, V., and Marshall, S. (2016). *Blockchain enigma paradox opportunity*. London, United Kingdom: Deloitte.

Hasse, F., Axel Hillebrand, T., Smole, E., Lay, L., and Charlet, M. (2016). *PwC global power & utilities, Blockchain - an opportunity for energy producers and consumers?* London, United Kingdom: PwC Glopal.

He, H., Zhao, L., Wang, Q., Chen, M., He, H., Gao, L., et al. (2020). 'Joint operation mechanism of distributed photovoltaic power generation market and carbon market based on cross-chain trading technology. *IEEE Access* 8, 66116–66130. doi:10.1109/access.2020.2985577

Herlihy, M. (2018). "Atomic cross-chain swaps," in *Podc'18: Proceedings of the 2018 acm symposium on principles of distributed computing* (New York, New York, United States: Association for Computing Machinery), 245–254.

Hileman, G., and Rauchs, M. (2017). Global cryptocurrency benchmarking study. *Camb. Centre Altern. Finance* 33.

Hua, W., Chen, Y., Qadrdan, M., Jiang, J., Sun, H., and Wu, J. (2022). 'Applications of blockchain and artificial intelligence technologies for enabling prosumers in smart grids: A review. *Renew. Sustain. Energy Rev.* 161, 112308. doi:10.1016/j.rser.2022.112308

Hua, W., Jiang, J., Sun, H., and Wu, J. (2020). 'A blockchain based peer-to-peer trading framework integrating energy and carbon markets. *Appl. Energy* 279, 115539. doi:10.1016/j.apenergy.2020.115539

Jiang, S., Li, Y., Lu, Q., Hong, Y., Guan, D., Xiong, Y., et al. (2021). 'Policy assessments for the carbon emission flows and sustainability of Bitcoin blockchain operation in China. *Nat. Commun.* 12, 1938–2010. doi:10.1038/s41467-021-22256-3

Juri, M., Timo, S., Catarina, N., Riitta, S., Marianne, T., Alexandra, B., et al. (2016). "Industrial blockchain platforms: An exercise in use case development in the energy industry," in ETLA Working Papers 43 (The Research Institute of the Finnish Economy).

Kakavand, H., Kost De Sevres, N., and Chilton, B. (2017). The blockchain revolution: An analysis of regulation and technology related to distributed ledger technologies. *SSRN Electron. J.* doi:10.2139/ssrn.2849251

Krause, M., and Tolaymat, T. (2018). 'Quantification of energy and carbon costs for mining cryptocurrencies. *Nat. Sustain.* 1, 711–718. doi:10.1038/s41893-018-0152-7

Kuo, T., Kim, H. E., and Ohno-Machado, L. (2017a). 'Blockchain distributed ledger technologies for biomedical and health care applications. *J. Am. Med. Inf. Assoc.* 24, 1211–1220. doi:10.1093/jamia/ocx068

Ledger, O. (2018). "A universal blockchain protocol enabling crossledger access through business modularization.

Lu, Y., Yue, L., Tang, X., Cai, B., Wang, H., Liu, L., et al. (2022). 'STRICTs: A blockchain-enabled smart emission cap restrictive and carbon permit trading system. *Appl. Energy* 313, 118787. doi:10.1016/j.apenergy.2022.118787

Luu, L., Narayanan, V., Zheng, C., Baweja, K., Gilbert, S., and Saxena, P. (2016). "A secure sharding protocol for open blockchains," in Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna Austria, October 24 - 28, 2016, 17–30.

Mora, C., Rollins, R. L., Taladay, K., Kantar, M. B., Chock, M. K., Shimada, M., et al. (2018a). Bitcoin emissions alone could push global warming above 2°C. *Nat. Clim. Chang.* 8, 931–933. doi:10.1038/s41558-018-0321-8

Mylrea, M., and Gourisetti, S. (2017). "Blockchain for smart grid resilience: Exchanging distributed energy at speed, scale and security," in *The resilience week (RWS) 2017* (Piscataway, New Jersey, United States: IEEE), 18–23.

Ometoruwa, T. (2018). Solving the blockchain trilemma: Decentralization, security & scalability. Available at: https://www.coinbureau.com/analysis/solving-blockchain-trilemma/.

Poon, J., and Buterin, V. (2017). Plasma: Scalable autonomous smart contracts. *White Pap.*

Poon, J., and Dryja, T. (2016). The bitcoin lightning network: Scalable off-chain instant payments. *draft version 0.5.9.2*, 14.

Shamsi, M., and Paul, C. (2020). "Towards the use of blockchain prediction markets for forecasting wind power," in *2020 6th IEEE international energy conference (ENERGYCon)* (Piscataway, New Jersey, United States: IEEE), 847–852.

Shaolin, F. (2008). Mandatory activation of segwit deployment. Available at: https://github.com/bitcoin/bips/blob/master/bip-0148.mediawiki.

Song, J. (2017). Transaction malleability explained. Available at: https://bitcointechtalk.com/transaction-malleability-explained-b7e240236fc7.

Sukhomlinova, L. (2019). The mystery of 2100 ETH gas transaction, mistake or something else? Available at: https://blogs.airdropalert.com/who-gives-away-2100-eth/.

Sundararajan, S. (2017). Australian government grants $8 million for blockchain energy pilot. Coindesk. Available at: https://www.coindesk.com/australian-government-grants-8-million-for-blockchain-energy-pilot.

Teng, F., Zhang, Q., Wang, G., Liu, J., and Li, H. (2021). 'A comprehensive review of energy blockchain: Application scenarios and development trends. *Int. J. Energy Res.* 45, 17515–17531. doi:10.1002/er.7109

Todd, P. (2014). Near-zero fee transactions with hub-and-spoke micropayments. Available at: https://lists.linuxfoundation.org/pipermail/bitcoin-dev/2014-December/006988.html.

Truby, J., Brown, R., Dahdal, A., and Ibrahim, I. (2022). 'Blockchain, climate damage, and death: Policy interventions to reduce the carbon emissions, mortality, and net-zero implications of non-fungible tokens and Bitcoin. *Energy Res. Soc. Sci.* 88, 102499. doi:10.1016/j.erss.2022.102499

Wang, T., Guo, J., Ai, S., and Cao, J. (2021). 'RBT: A distributed reputation system for blockchain-based peer-to-peer energy trading with fairness consideration. *Appl. Energy* 295, 117056. doi:10.1016/j.apenergy.2021.117056

Williams, M. (2019). Developer's mistake? Someone just paid $450K (3,150 ETH) for Ethereum transaction fees. cryptopotato.com. Available at: https://cryptopotato.com/developers-mistake-someone-just-paid-450k-3150-eth-for-ethereum-transaction-fees/.

Wongthongtham, P., Morrison, G., and Liu, X. (2019). The costs and benefits of blockchain based peer-to-peer energy trading: An evaluation from the perspective of carbon emission and economic value.

Wongthongtham, P., Marrable, D., Abu-Salih, B., Liu, X., and Morrison, G. (2021). 'Blockchain-enabled Peer-to-Peer energy trading. *Comput. Electr. Eng.* 94, 107299. doi:10.1016/j.compeleceng.2021.107299

Wu, Y., Wu, Y., Cimen, H., Vasquez, C., and Josep, M. (2022). 'P2P energy trading: Blockchain-enabled P2P energy society with multi-scale flexibility services. *Energy Rep.* 8, 3614–3628. doi:10.1016/j.egyr.2022.02.074

Wuille, P., Todd, P., Maxwell, G., and Russell, R. (2015). Version bits with timeout and delay. Available at: https://github.com/bitcoin/bips/wiki/Comments:BIP-0009.

Zheng, Z., Xie, S., Dai, H., Chen, X., and Wang, H. (2017). "An overview of blockchain technology: Architecture, consensus, and future trends," in *2017 IEEE international congress on big data (BigData congress)*. Editors G. Karypis and J. Zhang (Piscataway, New Jersey, United States: IEEE), 557–564.

Zhong, X., Zhong, W., Liu, Y., Yang, C., and Xie, S. (2022). Optimal energy management for multi-energy multi-microgrid networks considering carbon emission limitations. *Energy* 246, 123428. doi:10.1016/j.energy.2022.123428

frontiers | Frontiers in Built Environment

TYPE Original Research
PUBLISHED 04 October 2022
DOI 10.3389/fbuil.2022.966351

OPEN ACCESS

EDITED BY
Yasmin Murad,
The University of Jordan, Jordan

REVIEWED BY
Ayman S. Mosallam,
University of California, Irvine,
United States
Faris Matalkah,
Yarmouk University, Jordan

*CORRESPONDENCE
Nabil Al-Akhras,
alakhras@just.edu.jo

SPECIALTY SECTION
This article was submitted to Computational Methods in Structural Engineering,
a section of the journal
Frontiers in Built Environment

RECEIVED 10 June 2022
ACCEPTED 26 July 2022
PUBLISHED 04 October 2022

CITATION
Al-Akhras N and Othman O (2022), Bond behavior of NSM strips in corroded/cracked reinforced concrete. *Front. Built Environ.* 8:966351. doi: 10.3389/fbuil.2022.966351

Bond behavior of NSM strips in corroded/cracked reinforced concrete

Nabil Al-Akhras* and Osama Othman

Civil Engineering Department, Jordan University of Science and Technology, Irbid, Jordan

There is limited research on the bond behavior of NSM strips embedded in corroded/cracked reinforced concrete (RC). Therefore, this study was initiated to evaluate the bond behavior between corroded/cracked RC and NSM strips. The test specimens consist of two concrete prisms (reinforced and plain) mounted from two opposite sides with NSM strips. After 28 days of moisture curing, twenty-four specimens were immersed in saline solution and were exposed to three accelerated corrosion levels with steel mass loss of 2.2, 4.1 and 8.1%. The bond behavior was evaluated using double shear pullout test. The study showed that the corrosion cracks increased with the increase of corrosion level. The maximum crack width reached 0.91 mm at corrosion level 3. The mass loss of the corroded reinforcing steel increased from 2.1% for corrosion level 1–8.3% for corrosion level 3. The bond strength between NSM strips and corrosion cracked concrete was significantly decreased with maximum reduction of 48% at corrosion level 3. The other bond characteristics including maximum slip, bond toughness and bond stiffness were also decreased with corrosion with percent reduction of (31, 67 and 35%), respectively. The failure mode of the bond between NSM strips and corrosion cracked concrete was splitting of concrete. Theoretical models were proposed for predicting the bond strength between NSM strips and corrosion cracked concrete based on corrosion level and crack width. The proposed models showed excellent agreement with the experimental results from the current study and the literature.

KEYWORDS

concrete, corrosion, bond, CFRP composite, double pullout

Introduction

Many reinforced concrete (RC) structural elements suffer from degradation and deteriorations with time that reduce the structural capacity of the elements and may cause failure in severe situations. Therefore, repair and rehabilitation are needed to restore the structural capacity of the deteriorated elements. The use of fiber-reinforced polymer (FRP) is a popular practice for the repair of the deteriorated RC elements (Lee et al., 2000; Al-Allaf et al., 2016). The most commonly method used of FRP is externally bonded by wrapping FRP composite laminates to the external surface of the deteriorated RC elements. However, premature debonding between the FRP and concrete usually takes

place before reaching the full capacity of the FRP tensile strength (Mukhopadhyaya and Swamy, 2001; Nguyen et al., 2001).

Near surface mounted (NSM) fiber reinforced polymer technique was developed to overcome the premature debonding failure of FRP. This technique uses FRP composite strips inserted into slits inside the concrete cover of the RC member (Lorenzis and Nanni, 2015; Triantafyllou et al., 2018; Dias and Barros 2012; Mosallam et al., 2022). Recently, repair and strengthening using NSM technique showed promising results to increase the load-carrying capacity of the RC members without debonding. The high level of strength efficiency associated with repaired RC elements using NSM strips demonstrated this technique as a viable and promising repair procedure (Bilotta et al., 2011; Coelho et al., 2015).

Corrosion of reinforced concrete (RC) structures is a major problem all over the world especially in structures exposed to marine environment. Corrosion products possess larger volume than the original steel forming tensile stress that causes deterioration in RC structures (Peng et al., 2015; Fernandes et al., 2018; Kreit et al., 2011).

The corrosion of RC elements repaired with NSM strips produce negative impact on the bond strength between NSM strips and concrete. The intensity of the bond degradation is not well established and requires more investigation. Therefore, this study was initiated to evaluate the bond behavior between NSM strips and corrosion cracked concrete using different parameters including three levels of corrosion deterioration, number of NSM strips (Triantafyllou et al., 2018; Al-Akhras and Aleghnimat, 2020) and bond lengths of NSM strips (100 and 150 mm).

Self-compacting concrete (SCC) was considered in this study because SCC is relatively a new type of concrete and there are no research studies reported on the bond behavior between corrosion cracked SCC and NSM strips. The SCC was developed mainly to freely flow to any corner of the formwork without compaction. The SCC has higher compressive strength and lower stiffness than the normal vibrated concrete (Okamura et al., 2003).

Experimental work

Testing program

The testing program is composed of thirty-two concrete specimens. Each specimen consists of two concrete blocks (reinforced and plain). Each concrete block has dimensions of 150 × 150 × 200 mm. The two concrete blocks are mounted with NSM strips at the two opposite sides. The RC block is reinforced with four steel bars with diameter of 12 mm and two stirrups with diameter of 6 mm as shown in Figure 1. Twenty-four specimens were immersed in 3% NaCl solution and exposed to accelerated corrosion process. While eight specimens were kept in water bath at laboratory temperature of 25□C as reference (uncorroded) specimens. The testing program investigated is shown in Table 1.

Concrete mixing and casting

The SCC mixture was prepared at w/c ratio of 0.45 using Type I cement, crushed limestone coarse aggregate, crushed limestone fine aggregate, silica sand, and high-performance superplasticizer. The coarse aggregate has a maximum aggregate size of 12 mm, bulk specific gravity of 2.61 and absorption of 2.2%. The fine aggregate has a bulk specific gravity of 2.63, absorption of 1.9% and fineness modulus of 2.68. The SCC mixture possesses a compressive strength of 47 MPa and splitting strength of 3.4 MPa after 28 days of moist curing. Table 2 shows the proportions of the SCC mixture.

A tilting drum mixer with a capacity of 0.1 m^3 was used for mixing the SCC. Twelve concrete blocks (150 × 150 × 200 mm) and six cylinders (100 × 200 mm) were casted from each SCC batch. The SCC mixing procedure was as follows. The inside surface of the mixer was wetted with little amount of water. The coarse aggregate was added and followed by some amount of water with continuous mixing. The cement, fine aggregate, silica sand and water were added gradually and alternatively. Finally, the superplasticizer was added to achieve flowing concrete. After 3 minutes of blending the ingredients, the cohesive SCC was ready for casting in the molds (Figure 2A). The fresh SCC properties including slump flow, V-funnel and L-Box were conducted and found to be 710 mm, 11 s, and 0.9. The fresh concrete was casted into the molds (Figure 2C) without vibration. The casted concrete specimens were covered with wet burlap for 1 day then transferred to a water tank for another 27 days curing (Figure 2D).

NSM strips

The NSM strips were Sika CarboDur S1.525 having a width of 15 mm and a thickness of 2.5 mm. The NSM strips were embedded into slit in the concrete cover (7 mm in width and 25 mm in depth) and bonded with epoxy. The spacing between the slits for the three NSM strips was 30 mm. The mechanical properties of the NSM strips as supplied from the manufacturer were, longitudinal elastic modulus of 160 GPa and longitudinal tensile strength of 2,800 MPa. The adhesive epoxy was Sikadur-30. The mechanical properties of the epoxy as supplied from the manufacturer were: compressive strength of 85 MPa, tensile strength of 17 MPa, shear strength of 7 MPa, and elastic modulus of 10 GPa.

The NSM strips were installed on the SCC specimens using the following procedure (Al-Akhras and Aleghnimat, 2020). The concrete cover was cut into grooves using concrete saw (Figure 3A). The groove length of the plain concrete block

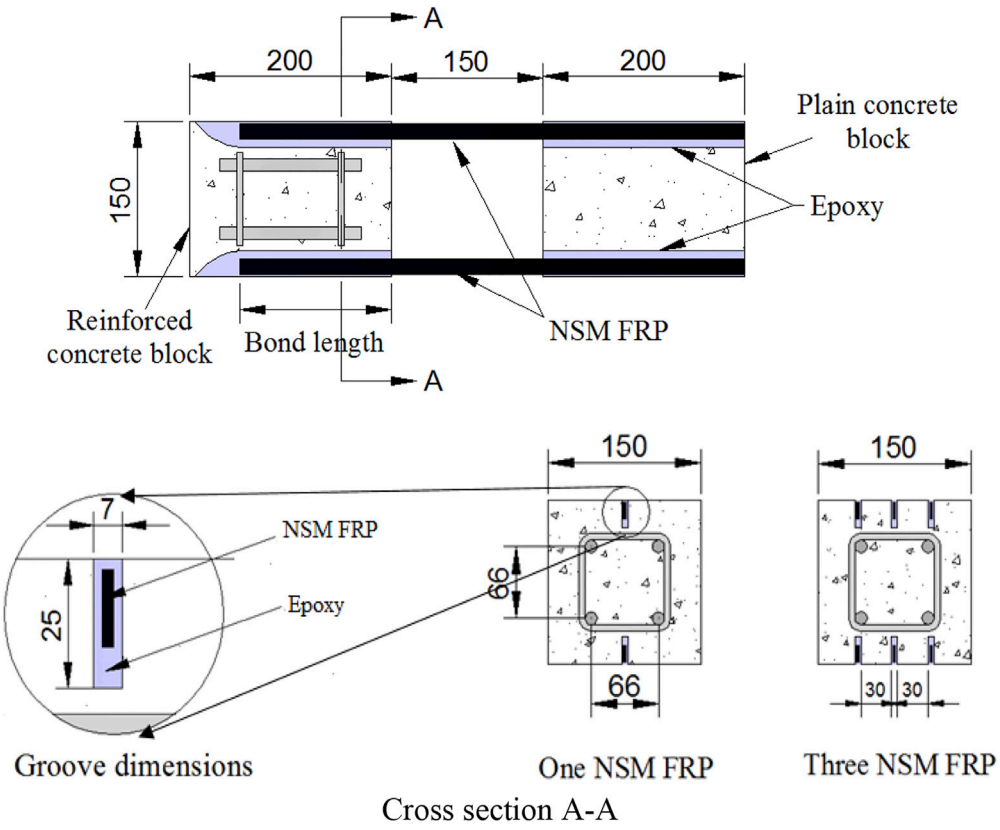


FIGURE 1
Schematics of specimens mounted with NSM (dimensions are in mm).

TABLE 1 Testing program adopted in the study.

Specimen	Number of NSM strips	Length of pullout strips (mm)	Number of specimens	Exposure
S1-10L0	1	100	2	No corrosion
S3-10L0	3	100	2	
S1-10L1	1	100	2	Corrosion Level 1
S3-10L1	3	100	2	
S1-10L2	1	100	2	Corrosion Level 2
S3-10L2	3	100	2	
S1-10L3	1	100	2	Corrosion Level 3
S3-10L3	3	100	2	
S1-15L0	1	150	2	No corrosion
S3-15L0	3	150	2	
S1-15L1	1	150	2	Corrosion Level 1
S3-15L1	3	150	2	
S1-15L2	1	150	2	Corrosion Level 2
S3-15L2	3	150	2	
S1-15L3	1	150	2	Corrosion Level 3
S3-15L3	3	150	2	
Total specimens			32	

TABLE 2 Proportions of the SCC mixture.

Mixture	Content (kg/m^3)
Cement	470
Water	204
Coarse aggregate	778
Fine aggregate	543
Silica sand	366
Superplasticizer	7

was 200 mm and that for the RC block was 100 or 150 mm (Al-Akhras and Mashaqbeh, 2020). The NSM strips were cut into two lengths of 450 and 500 mm for bond lengths of 100 and 150 mm using an electric saw (Triantafyllou et al., 2018). The grooves' surface was cleaned using brush, vacuum and isopropanol (Figure 3B). (Al-Allaf et al., 2016) The epoxy adhesive was prepared by mixing quantities of parts A and B with ratio of 3:1 using electric mixer. The quantities were mixed for about 3 minutes until the mixture becomes uniform and light brown in color (Figure 3D). (Bilotta et al., 2011) The concrete grooves were filled with epoxy using a trowel and the FRP strips (without surface pretreatment) were pressed firmly into the epoxy of the grooves (Coelho et al., 2015). The grooves' surface was squeezed and smoothed using a trowel to expel the air and to provide smooth and clean surface.

Accelerated corrosion process

Twenty-four specimens (RC and plain concrete blocks) were immersed in a circular plastic tank (600 by 1,200 mm) containing 3% NaCl solution for the accelerated corrosion process. The reinforcing steel were connected on parallel to the positive pole of the power supply to act as the anode. The

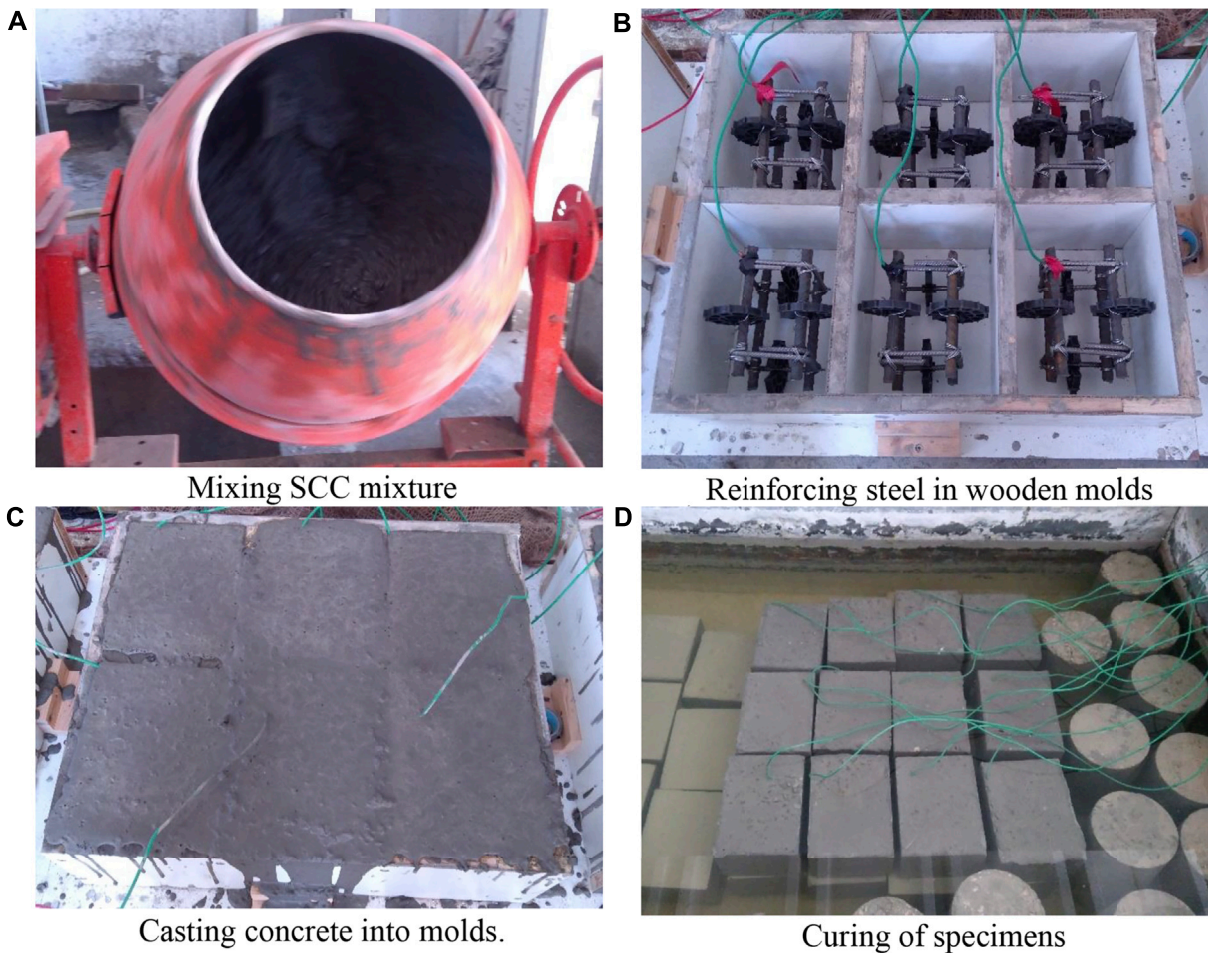


FIGURE 2
Mixing, casting and curing of concrete specimens. **(A)** Mixing SCC mixture. **(B)** Reinforcing steel in wooden molds. **(C)** Casting concrete into molds. **(D)** Curing of specimens.

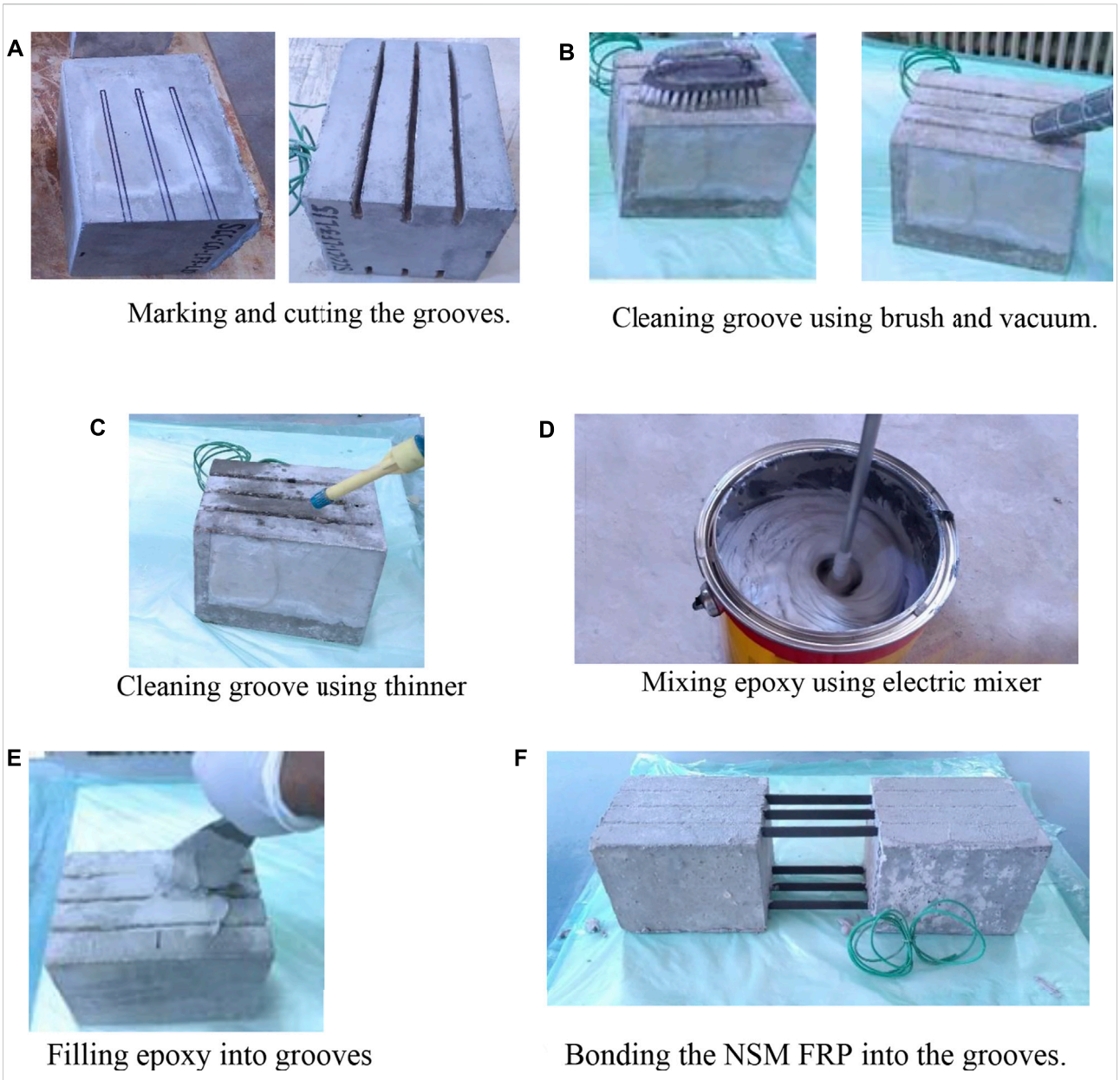


FIGURE 3
Installation of the FRP strips. **(A)** Marking and cutting the grooves. **(B)** Cleaning groove using brush and vacuum. **(C)** Cleaning groove using thinner. **(D)** Mixing epoxy using electric mixer. **(E)** Filling epoxy into grooves. **(F)** Bonding the NSM FRP into the grooves.

negative pole was connected to rectangular steel plate (cathode) immersed in the tank (Figure 4). Each specimen was exposed to an electrical current of 0.15 A and 12 V for time-periods of 5 (Lee et al., 2000; Mukhopadhyaya and Swamy, 2001; Nguyen et al., 2001; Okamura et al., 2003; Kreit et al., 2011; Lorenzis and Nanni, 2015; Mosallam et al., 2022), days for the first, second and third level of corrosion, respectively. The electrical current values were chosen based on previous work of Al-Akhras and Mashaqbeh (2020) and Al-Akhras and Aleghnimat (2020).

Pullout test

The specimens were removed from the corrosion tank, dried with a cloth to become saturated surface dry and

FIGURE 4
Accelerated corrosion process using power supply.

subjected to pullout test using universal testing machine with capacity of 600 kN. The two concrete blocks (plain and reinforced) were pulled from each other using steel fasteners at rate of 0.44 mm/min until failure of the bond of NSM strips. Linear variable differential transducers (LVDT) were placed at each bonded side to measure the slip. The load and displacement measurements were collected using data acquisition system connected to computer. Figure 5 illustrates the pullout test.

Test result and discussion

Crack pattern

The corrosion crack pattern was monitored for different corrosion levels as shown in Figure 6. The corrosion crack pattern varied with different corrosion level. The number of corrosion cracks increased with the increase of corrosion level. Two small corrosion cracks were observed after the first corrosion level. However, six corrosion cracks were detected for the second corrosion level. Moreover, a network of 10 corrosion cracks was noticed at the third corrosion level.

Crack width and mass loss

The crack width of the corrosion cracks and mass loss of reinforcing steel were evaluated for different corrosion levels. The

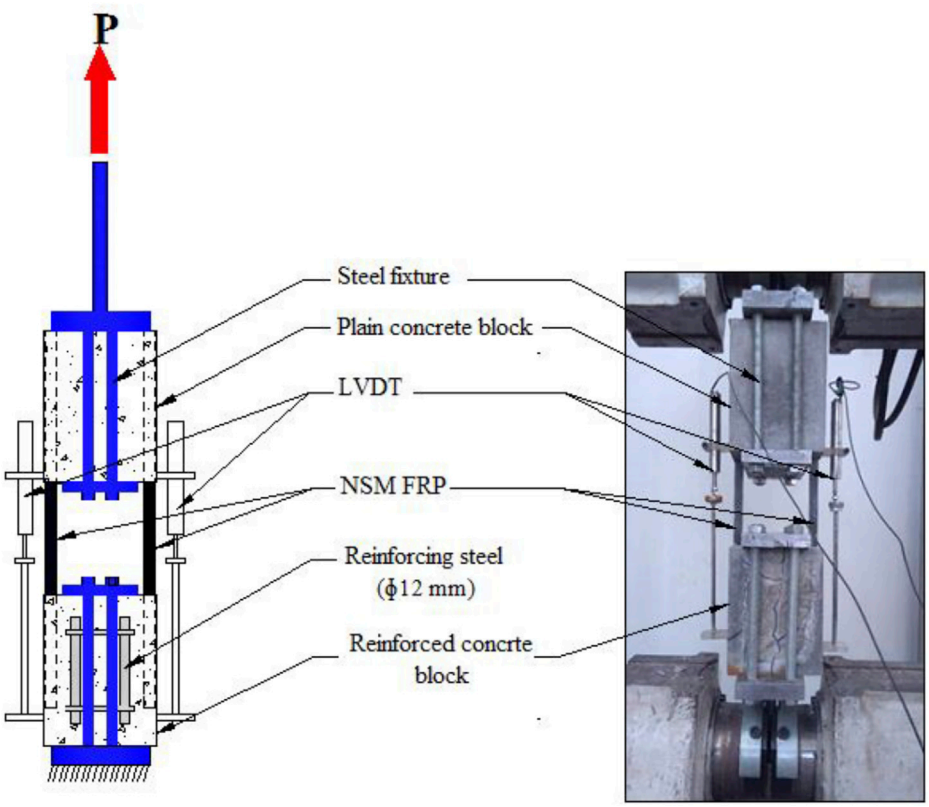


FIGURE 5
Test setup of the double pullout test.

FIGURE 6
Typical crack paterns for different corrosion levels.

TABLE 3 Maximum crack width and steel mass loss.

Corrosion level	NSM FRP	Bond length (mm)	Maximum crack width (mm)	Steel Mass loss (%)
1	1	100	0.22	2.2
		150	0.25	2.5
	3	100	0.20	2.2
		150	0.28	2.0
2	1	100	0.58	3.9
		150	0.62	4.1
	3	100	0.55	3.8
		150	0.63	4.2
3	1	100	0.91	8.1
		150	0.90	7.9
	3	100	0.85	8.0
		150	0.88	8.2

crack width was measured using an optical microscope with an accuracy of 0.005 mm. Table 3 shows the maximum crack width and steel mass loss for different corrosion levels. The maximum crack width was increased from an average of 0.25 mm for corrosion level 1 to an average of 0.88 mm for corrosion level 3. Additionally, the mass loss was increased with corrosion level from an average of 2.2% for corrosion level 1 to an average of 8.1% for corrosion level 3.

Effect of corrosion level on bond behavior

Figures 7–10 depict the effect of corrosion on the bond behavior between NSM strips and concrete. The bond stress

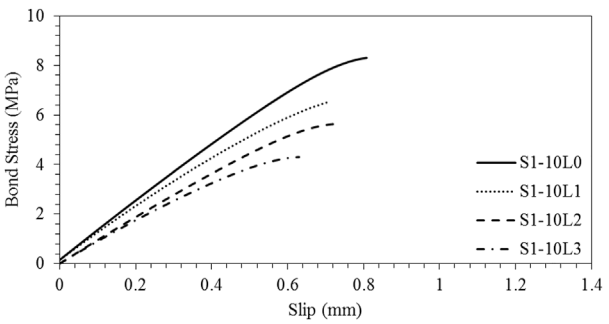


FIGURE 7
Relation between bond stress and slip for S1 strip and bond length of 100 mm.

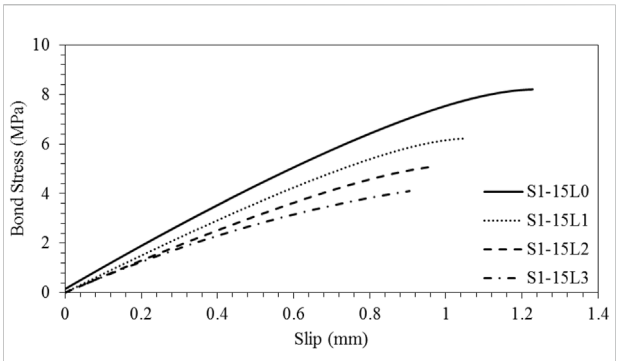


FIGURE 8
Relation between bond stress and slip for S1 strip and bond length of 150 mm.

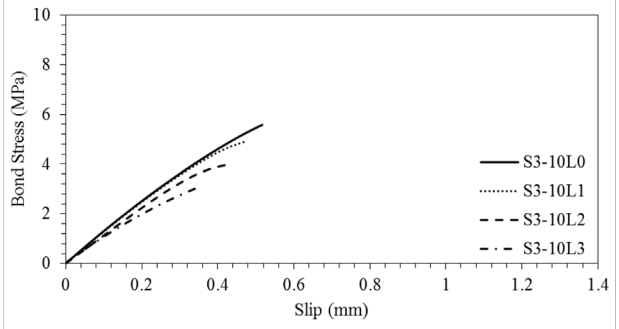


FIGURE 9
Relation between bond stress and slip for S3 strip and bond length of 100 mm.

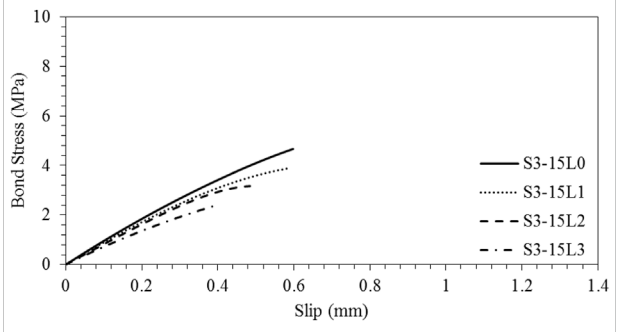


FIGURE 10
Relation between bond stress and slip for S3 strip and bond length of 150 mm.

was determined based on the average of two specimens using MS Excel. The scatter of the data was small with coefficient of variation of less than 2%. The bond-slip behavior followed similar behavior that consists of linear trend in the beginning then nonlinear behavior up to failure. The post peak behavior did not appear because the epoxy bonding is brittle material that failed suddenly after reaching the ultimate stage.

The results indicated that the bond strength and maximum slip decreased with increasing the corrosion level. This behavior is due to the fact that the corrosion of reinforcing steel cause significant deterioration in the concrete that degrade the bond between the NSM and concrete.

Table 4 provides the bond characteristics (bond strength, maximum slip, toughness and stiffness) of the NSM strips. The bond strength is determined from the maximum stress of the stress-slip curve. The maximum slip is determined from the maximum slip of the stress-slip curve. The toughness is defined as the total area under the stress-slip curve. The stiffness is defined as the slope of the initial line of the stress-slip curve.

The bond characteristics decreased significantly with the increase of the corrosion level. The ultimate load and maximum slip of the one NSM strip at bond length of 100 mm were decreased from (27 kN and 0.81 mm) for uncorroded specimens to (14 kN and 0.63 mm) for corroded specimens at corrosion level 3. The percent reduction in the ultimate load and maximum slip were 48 and 22%.

The toughness and stiffness for one NSM strip with 100 mm bond length were also decreased from (3802 J/m^2 and 13.1 GPa/m) for uncorroded NSM strips to (1580 J/m^2 and 8.9 GPa/m) for corroded NSM strips at corrosion level 3. The percent reduction in the toughness and stiffness were 58 and 32%.

The bond characteristics were increased with the increase of the bond length from 100 to 150 mm. The ultimate load and maximum slip at corrosion level 2 were increased from (18.3 kN and 0.73 mm) for bond length of 100 mm to (24.8 kN and 0.97 mm) for bond length of 150 mm. The toughness at corrosion level 3 was increased from (1580 J/m^2) for bond length of 100 mm to (2154 J/m^2) for bond length of 150 mm.

The bond stress for NSM strips with length of 100 mm is observed higher than that for NSM with length of 150 mm. However, the pullout load for NSM strips with length of 100 mm is lower than that for NSM with length of 150 mm. This behavior is explained as follows. The bond stress is determined by dividing the pullout load by the area of the NSM strip. The area of the NSM strip increases with the increase of the NSM bond length. Therefore, the bond stress decreased as the bond length increased (Figure 11).

The slip of the 150 mm NSM strips was higher compared to that of the 100 mm strips. This behavior is attributed to the fact that as the bond length of the NSM increase, the length exposed to slip increases and consequently the slip increase.

The ultimate load and maximum slip for one NSM strips (S1) were higher compared to those for 1 strip of three adjacent NSM strips (S3). This behavior may be explained because the overlapped area between the three adjacent NSM strips cause stress concentration on each NSM strip. Therefore, causing premature failure of the three adjacent NSM strips prior to reaching the maximum capacity of the bond strength (Figure 12).

TABLE 4 Bond characteristics for one NSM strip.

Specimen	Ultimate load (kN)	Bond strength (MPa)	Maximum slip (mm)	Toughness (J/m^2)	Stiffness (GPa/m)
S1-10L0	27.0	8.30	0.81	3,802	13.1
S1-15L0	40.0	8.20	1.23	5,978	9.8
S3-10L0	18.1	5.58	0.517	1,576	12.8
S3-15L0	22.7	4.66	0.60	1,526	9.4
S1-10L1	21.3	6.54	0.71	2,638	12.1
S1-15L1	30.3	6.22	1.05	3,756	7.6
S3-10L1	15.8	4.87	0.47	1,281	12.6
S3-15L1	18.9	3.87	0.58	1,305	8.8
S1-10L2	18.3	5.63	0.73	2,317	9.5
S1-15L2	24.8	5.09	0.97	2,776	6.5
S3-10L2	12.8	3.95	0.42	938	11.2
S3-15L2	15.4	3.17	0.48	891	8.2
S1-10L3	14.0	4.31	0.63	1,580	8.9
S1-15L3	20.0	4.10	0.91	2,154	6.4
S3-10L3	10.1	3.12	0.36	614	10.4
S3-15L3	11.5	2.35	0.39	497	7.1

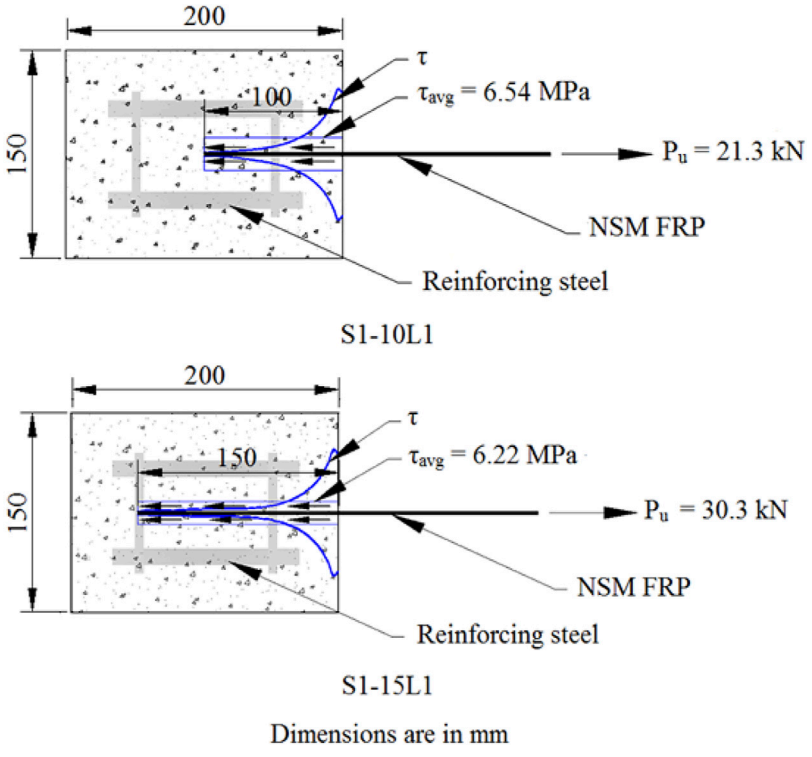


FIGURE 11
Schematics of the stress distribution for S1-15L1.

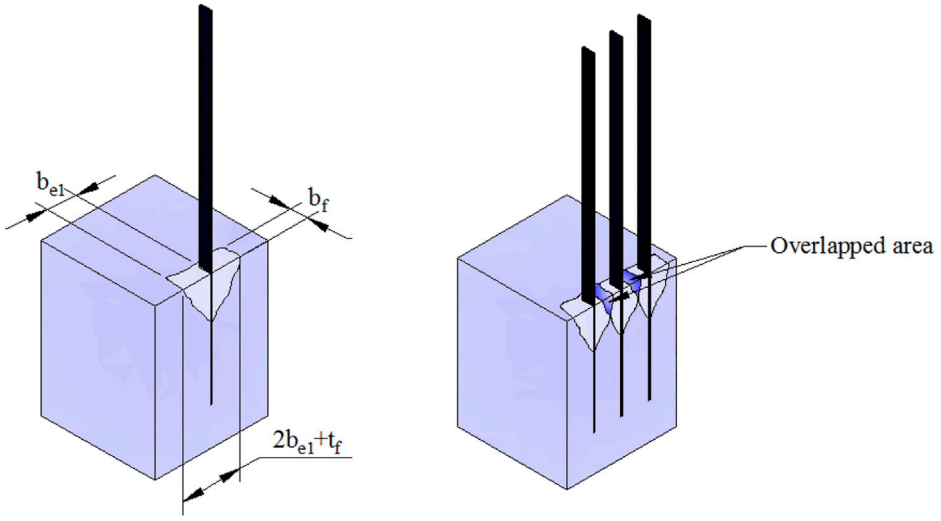


FIGURE 12
Failure mode of one and three NSM strips (After Seo et al., 2013)

The residual values of the bond characteristics are introduced and defined as follows:

$$R_i = \frac{Corroded\ Bond\ Characteristic}{Uncorroded\ Bond\ Characteristic} \times 100\,\% \qquad (1)$$

Where, R is residual bond characteristic (ultimate bond strength, maximum slip, toughness or stiffness); *i* is corrosion level of 1, two or 3. The residual bond characteristics were determined for all specimens and presented in Figure 13.

The residual bond characteristics decreased with corrosion level. The minimum residual bond characteristics (strength, slip, toughness and stiffness) were observed for one NSM strip (S1) at corrosion level 3. The minimum residual strength, slip, toughness and stiffness for one NSM strip with bond length of 100 and 150 mm at corrosion level 3 were (52, 78, 42 and 68%) and (50, 74, 36 and 65%), respectively. The corresponding minimum residual values for 1 strip of the three NSM strips at bond length of 100 and 150 mm at corrosion level 3 were (56, 69, 39% 81%) and (50, 65, 33% 75%), respectively.

Failure mode

The failure mode of the bond between NSM and concrete was splitting of concrete as shown in Figure 14. This behavior is explained by the fact that the tensile strength of the epoxy bond (17 MPa) is significantly higher than the splitting strength of the concrete (3.4 MPa). Therefore, the concrete failed and split before the epoxy bond failed. Small concrete fractures of the split concrete remained attached to the NSM strips after failure. These concrete fractures increased in size with the increase of corrosion level. Large parts of split concrete remained attached to the NSM strips after failure at corrosion level 3. This behavior is due to the fact that corrosion cracks significantly weakened the concrete.

Theoretical models

Theoretical models for the effect of corrosion cracks of concrete on the bond strength of NSM were developed using the following procedure (Al-Akhras and Aleghnimat, 2020). Models of the bond between NSM strips and intact concrete were developed (Al-Akhras and Mashaqbeh, 2020). Reduction factors for the effect of corrosion cracks on the bond between NSM strips and concrete were proposed based on corrosion crack width and steel mass loss (Triantafyllou et al., 2018). The bond strengths between NSM and corrosion cracked concrete were determined by multiplying the bond strength between NSM and intact concrete by the reduction factor (Al-Allaf et al., 2016). The theoretical bond strength were compared and validated with the experimental ones.

The bond strengths between NSM strips and intact concrete are modelled using the following equations based on (Seo et al. 2011):

$$T_{f1} = \emptyset b_f t_f f_{yf}\ (\text{Tension failure of FRP, kN}) \qquad (2)$$

$$T_{f2} = \lambda \tau_f \left(2 b_f l_d\right) (\text{Epoxy failure, kN}) \qquad (3)$$

$$T_{f3} = 0.57 \beta \sqrt{f_{ck}} A_{cf}\ (\text{Splitting failure of concrete, kN}) \qquad (4)$$

$$T_f = \min\left\{T_{f1}, T_{f2}, T_{f3}\right\} (\text{Total tension failure, kN}) \qquad (5)$$

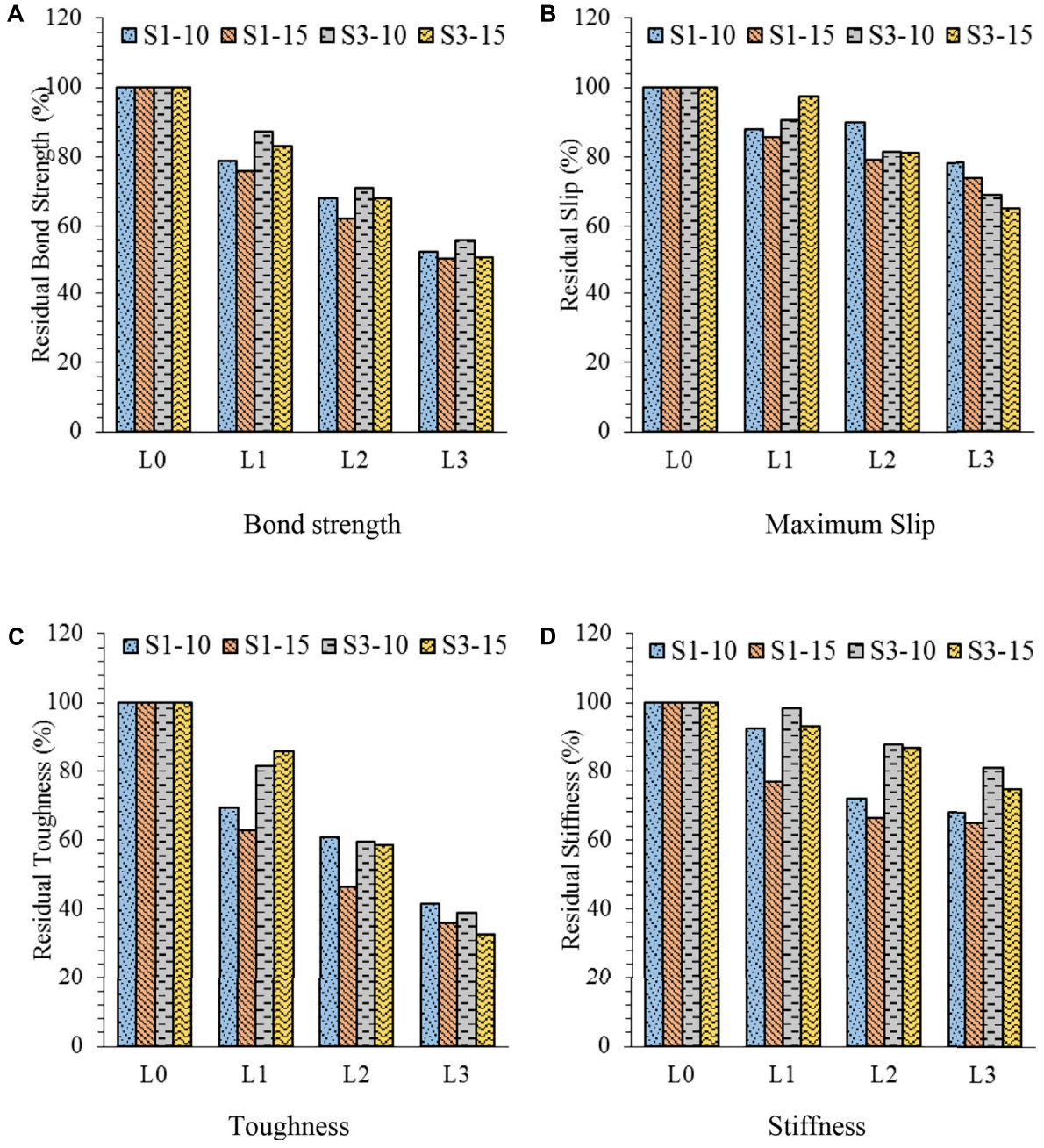


FIGURE 13
Residual bond characteristics for different corrosion levels. **(A)** Bond strength. **(B)** Maximum Slip. **(C)** Toughness. **(D)** Stiffness.

Where; b_f and t_f are width and thickness of NSM strip, f_{yf} is yield strength of FRP, λ is reduction factor of bond, τ_f is shear strength of epoxy, l_d is bond length, β is experimental coefficient (equal to 1), f_{ck} is concrete compressive strength as determined by authors, and A_{cf} is surface area of concrete split failure. The A_{cf} and reduction factor of bond length (k) were determined using the following equation based on Seo et al. (2013).

$$A_{cf} = 2b_f\sqrt{(kh_e)^2 + b_{e1}^2} + (kh_e)\left(t_f + b_{e1}\right) \quad (6)$$

Where; h_e is length of failed concrete. Based on observations of the experimental results, the effective width (b_{e1}) is found approximately equal to h_e. The theoretical bond strengths were determined and presented in Table 5 along with the experimental bond strengths. Excellent agreement is observed

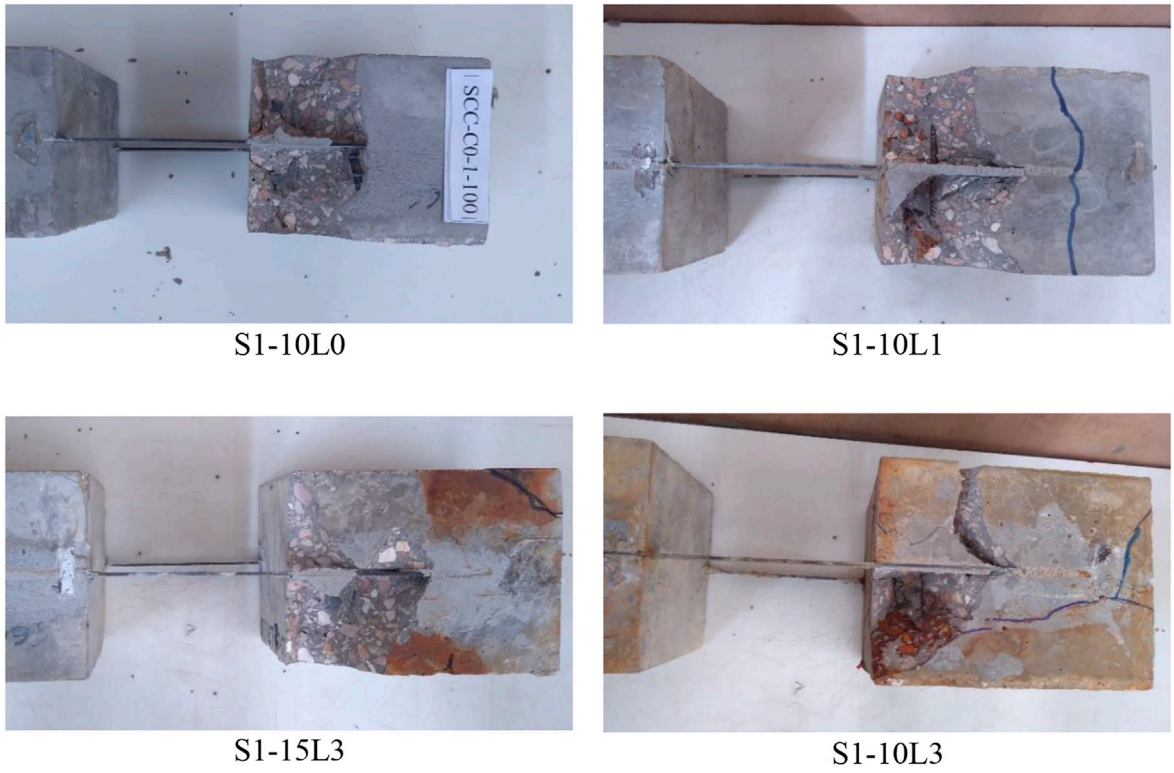


FIGURE 14
Failure mode of the NSM bond with concrete.

TABLE 5 Theoretical and experimental bond strengths between NSM and intact concrete.

Specimen	Bond strength (MPa)		Difference (%)
	Experimental	Theoretical	
S1-10L0	8.30	8.48	2.2
S3-10L0	5.58	5.33	4.5
S1-15L0	8.20	7.81	4.8
S3-15L0	4.66	4.58	1.7

between the theoretical and experimental bond strengths with maximum difference of 4.8%.

Correlation models between the residual bond strength of NSM and (steel mass loss and maximum crack width) were developed. Quadratic equation models were considered since they showed the best fit compared to other models. The correlation models showed excellent fit between the residual bond strength and steel mass loss with R-square of 0.94 and between residual bond strength and maximum crack width with R-square of 0.95 as shown in Figures 15, 16.

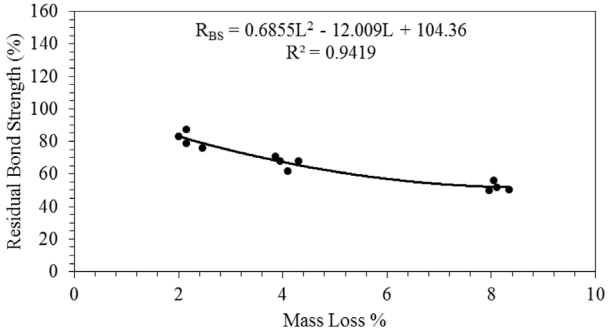


FIGURE 15
Correlation model between residual bond strength (R_{BS}) and mass loss (L).

The bond strengths between NSM strips and corroded cracked concrete was evaluated using the following equation:

$$\tau_{Corroded} = R \times \tau_{uncorroded} \quad (7)$$

Where; $\tau_{Corroded}$ is bond strength between NSM and corrosion cracked concrete, $\tau_{uncorroded}$ is bond strength between NSM and

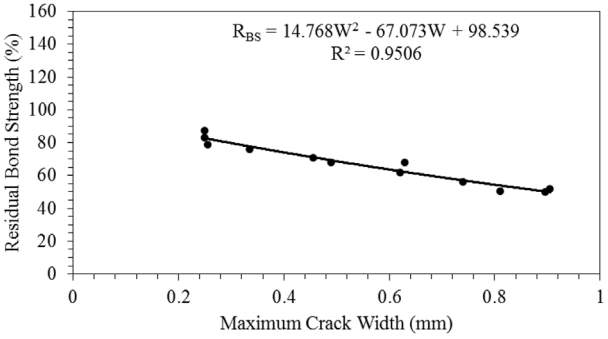


FIGURE 16
Correlation model between residual bond strength (R_{BS}) and crack width (W).

intact concrete, *R* is reduction factor (residual) determined from the correlation models. The theoretical bond strengths between NSM and corrosion cracked concrete were determined and tabulated in Tables 6, 7 along with the experimental bond strengths.

Excellent agreement is observed between the theoretical and experimental bond strengths between NSM strips and corrosion cracked concrete with maximum difference of 7.4%.

The proposed models were also validated with the experimental results from the literature (Haddad and Al-Dalou, 2018) who investigated the bond behavior between CFRP laminates and corrosion cracked normal vibrated concrete. The validation of the proposed models with experimental results from the literature based on steel mass loss and crack width are tabulated in Tables 8, 9. Excellent

TABLE 6 Validation of the proposed model based on steel mass loss.

Specimen	Mass loss (%)	*R* (%)	Uncorroded bond strength (MPa)	τ_{theo} (MPa)	τ_{exp} (MPa)	Difference (%)
S1-10L1	2.2	81.7	8.3	6.78	6.54	3.54
S3-10L1	2.2	81.7	5.6	4.56	4.87	6.80
S1-10L2	4.0	67.6	8.3	5.61	5.63	0.36
S3-10L2	3.9	68.3	5.6	3.81	3.95	3.67
S1-10L3	8.1	52.1	8.3	4.32	4.31	0.23
S3-10L3	8.1	52.1	5.6	2.91	3.12	7.22
S1-15L1	2.5	79.1	8.2	6.48	6.22	4.01
S3-15L1	2.0	83.1	4.7	3.87	3.87	0.00
S1-15L2	4.1	66.7	8.3	5.47	5.09	6.95
S3-15L2	4.3	65.4	4.7	3.05	3.17	3.93
S1-15L3	8.0	52.2	8.3	4.28	4.10	4.21
S3-15L3	8.3	51.9	4.7	2.42	2.35	2.89

TABLE 7 Validation of the proposed model based on maximum crack width.

Specimen	Maximum crack width (mm)	*R* (%)	Uncorroded bond strength (MPa)	τ_{theo} (MPa)	τ_{exp} (MPa)	Difference (%)
S1-10L1	0.26	82.2	8.3	6.82	6.54	4.11
S3-10L1	0.25	82.8	5.6	4.62	4.87	5.41
S1-10L2	0.49	69.0	8.3	5.73	5.63	1.75
S3-10L2	0.45	71.2	5.6	3.97	3.95	0.50
S1-10L3	0.91	49.7	8.3	4.12	4.31	4.61
S3-10L3	0.74	56.8	5.6	3.17	3.12	1.58
S1-15L1	0.34	77.4	8.3	6.34	6.22	1.89
S3-15L1	0.25	82.8	4.7	3.86	3.87	0.26
S1-15L2	0.62	62.4	8.3	5.11	5.09	0.39
S3-15L2	0.58	64.3	4.7	3.00	3.17	5.67
S1-15L3	0.90	50.1	8.3	4.11	4.10	0.24
S3-15L3	0.81	53.7	4.7	2.50	2.35	6.00

TABLE 8 Validation of the proposed model with Haddad and Al-Dalou (2018) based on steel mass loss.

Mass loss (%)	*R* (%)	Uncorroded bond strength (MPa)	τ_{theo} (MPa)	τ_{exp} (MPa)	Difference (%)
2.4	79.55	2.97	2.36	2.23	5.51
3.5	70.80	2.97	2.10	1.93	8.10
5	61.54	2.97	1.83	1.74	4.92
2.3	80.43	2.83	2.28	2.53	10.96
4	67.37	2.83	1.91	2.18	14.14
6.2	56.35	2.83	1.59	1.83	15.09
2.5	78.68	2.48	1.95	2.2	12.82
3.7	69.38	2.48	1.72	1.8	4.65
5.4	59.59	2.48	1.48	1.58	6.76
2.3	80.43	3.1	2.49	2.64	6.02
4	67.37	3.1	2.09	2.18	4.31
6.2	56.35	3.1	1.75	1.75	0.00
2.3	80.43	2.24	1.80	2.1	16.67
4	67.37	2.24	1.51	1.88	24.50
6.2	60.30	2.24	1.35	1.68	24.44

TABLE 9 Validation of the proposed model with Haddad and Al-Dalou (2018) based on maximum crack width.

Crack width (mm)	*R* (%)	Uncorroded bond strength (MPa)	τ_{theo} (MPa)	τ_{exp} (MPa)	Difference (%)
0.25	82.68	2.97	2.46	2.23	9.35
0.63	62.16	2.97	1.85	1.93	4.32
0.98	47.05	2.97	1.40	1.74	24.29
0.15	88.78	2.83	2.51	2.53	0.80
0.59	64.12	2.83	1.81	2.18	20.44
0.93	53.60	2.83	1.52	1.83	20.39
0.2	85.69	2.48	2.13	2.2	3.29
0.64	61.68	2.48	1.53	1.8	17.65
0.88	51.00	2.48	1.26	1.58	25.40
0.17	87.54	3.1	2.71	2.64	2.58
0.53	67.15	3.1	2.08	2.18	4.81
0.96	47.82	3.1	1.48	1.75	18.24
0.16	88.16	2.24	1.97	2.1	6.60
0.59	69.50	2.24	1.56	1.88	20.51
0.85	62.30	2.24	1.40	1.68	20.00

agreement is observed between the proposed model and experimental results based on steel mass loss with difference of 15%. Only two specimens showed a difference of 24%.

Also, very good agreement is observed between the proposed models and literature based on crack width with maximum difference of 25%. Although, the work of Haddad and Al-Dalou (2018) was different from the current study where they investigated the bond behavior between CFRP laminate and normal vibrated concrete.

Conclusion

Little information is reported in the literature on the bond behavior between NSM strips and corrosion cracked SCC. Therefore, this study fills the gap in literature on that subject. Based on the results of the study, the following conclusion are drawn:

1) The crack pattern of the corroded specimens varied with different corrosion levels. The number and width of corrosion

cracks increased with the increase of corrosion level. The maximum crack width was 0.91 mm at corrosion level 3. The reinforcing steel mass loss increased from 2.1% for corrosion level 1–8.3% for corrosion level 3.

2) The bond strength characteristics between NSM strips and corrosion cracked concrete showed progressive degradation with corrosion level. The minimum residuals of bond strength, maximum slip, toughness and stiffness reached 52, 69, 33 and 65% at corrosion level 3.
3) The pullout load of NSM strips with bond length of 150 mm was higher than that of NSM strips with bond length of 100 mm. The ultimate load and maximum slip for one single NSM strip were observed higher compared to those for one NSM strip of three adjacent NSM strips.
4) The failure mode of the bond between NSM and corrosion cracked concrete was splitting of concrete. Small concrete fractures were remained attached to the NSM strips after failure. These concrete fractures increased in size with the increase of corrosion level.
5) Theoretical models were proposed for predicting the bond strength between NSM strips and corrosion cracked concrete based on crack width and steel mass loss. The proposed models showed good agreement with the experimental results obtained from the current study and from the literature.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

NA-A supervise the experimental program that was conducted for the paper and revise thoroughly the draft of the paper OO, write the initial draft of the paper.

Funding

The authors acknowledge the financial support received from the Deanship of Scientific Research at Jordan University of Science and Technology (Grant 367/2017).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Al-Akhras, N., and Aleghnimat, R. (2020). Evaluating corrosion deterioration in self-compacted reinforced concrete beams and prisms using different tests. *Constr. Build. Mat.* 256, 119347. doi:10.1016/j.conbuildmat.2020.119347

Al-Akhras, N. M., and Mashaqbeh, Y. (2020). Potential use of Eucalyptus leaves as green corrosion inhibitor of steel reinforcement" journal of building engineer. *Elsevier* 35, 101848. doi:10.1016/j.jobe.2020.101848

Al-Allaf, M., Weekes, L., Augusthus-Nelson, L., and Leach, P. (2016). An experimental investigation into the bond-slip behavior between CFRP composite and lightweight concrete. *Constr. Build. Mater.* 113, 15–27. doi:10.1016/j.conbuildmat.2016.03.032

Bilotta, A., Ceroni, F., Di Ludovico, M., Nigro, E., Pecce, M., and Manfredi, G. (2011). Bond efficiency of EBR and NSM FRP systems for strengthening concrete members. *J. Compos. Constr.* 15 (5), 757–772. doi:10.1061/(asce)cc.1943-5614.0000204

Coelho, M. R. F., Sena-Cruz, J. M., and Neves, L. A. C. (2015). A review on the bond behavior of FRP NSM systems in concrete. *Constr. Build. Mater.* 93, 1157–1169. doi:10.1016/j.conbuildmat.2015.05.010

Dias, S. J. E., and Barros, J. A. O. (2012). NSM shear strengthening technique with CFRP laminates applied in high-strength concrete beams with or without pre-cracking. *Compos. Part B Eng.* 43 (2), 290–301. doi:10.1016/j.compositesb.2011.09.006

Fernandes, P., Sena-Cruz, J., Xavier, J., Silva, P., Pereira, E., and Cruz, J. (2018). Durability of bond in NSM CFRP-concrete systems under different environmental conditions. *Compos. Part B Eng.* 138, 19–34. doi:10.1016/j.compositesb.2017.11.022

Haddad, R., and Al-Dalou, A. (2018). Experimental study on bond behavior between corrosion-cracked reinforced concrete and CFRP sheets" Journal of Adhesion Science and Technology. *Taylor Francis* 32 (6), 590–608. doi:10.1080/01694243.2017.1371912

Kreit, A., Al-Mahmoud, F., Castel, A., and François, R. (2011). Repairing corroded RC beam with Near-Surface Mounted CFRP rods. *Mat. Struct.* 44 (7), 1205–1217. doi:10.1617/s11527-010-9693-6

Lee, C., Bonacci, J., Thomas, M., Maalej, M., Khajehpour, S., Hearn, N., et al. (2000). Accelerated corrosion and repair of reinforced concrete columns using carbon fibre reinforced polymer sheets. *Can. J. Civ. Eng.* 27 (5), 941–948. doi:10.1139/l00-030

Lorenzis, L., and Nanni, A. (2015). Strengthening of reinforced concrete structures with near surface mounted FRP rods. *ACI Struct. J.* 1948 (5), 1–8.

Mosallam, A. S., Ghabban, N., Mirnateghi, E., and Agwa, A. A. K. (2022). Nonlinear numerical simulation and experimental verification of bondline strength of CFRP strips embedded in concrete for NSM strengthening applications. *Struct. Concr.* 23, 1794–1815. doi:10.1002/suco.202100537

Mukhopadhyaya, P., and Swamy, N. (2001). Interface shear stress: A new design criterion for plate debonding. *J. Compos. Constr.* 5 (1), 35–43. doi:10.1061/(asce)1090-0268(2001)5:1(35)

Nguyen, D., Chan, D., and Cheong, H. (2001). Brittle failure and bond development length of CFRP-concrete beams. *J. Compos. Constr.* 5 (1), 12–17. doi:10.1061/(asce)1090-0268(2001)5:1(12)

Okamura, H., Ouchi, M., Schutter, G., Bartos, P., Domone, P., Gibbs, J., et al. (2003). Self-compacting concrete. *J. Adv. Concr. Technol.* 1 (1), 5–15. doi:10.3151/jact.1.5

Peng, H., Hao, H., Zhang, J., Liu, Y., and Cai, C. S. (2015). Experimental investigation of the bond behavior of the interface between near-surface-mounted CFRP strips and concrete. *Constr. Build. Mater.* 96, 11–19. doi:10.1016/j.conbuildmat.2015.07.156

Seo, S.-Y., Yoon, S.-J., Kwon, Y.-S., and Choi, K.-B. (2011). Bond behavior between near-surface-mounted fiber reinforced polymer plates and concrete in structural strengthening. *J. Korea Concr. Inst.* 23 (5), 675–682. doi:10.4334/jkci.2011.23.5.675

Seo, S. Y., Feo, L., and Hui, D. (2013). Bond strength of near surface-mounted FRP plate for retrofit of concrete structures. *Compos. Struct.* 95, 719–727. doi:10.1016/j.compstruct.2012.08.038

Triantafyllou, G. G., Rousakis, T. C., and Karabinis, A. I. (2018). Effect of patch repair and strengthening with EBR and NSM CFRP laminates for RC beams with low, medium and heavy corrosion. *Compos. Part B Eng.* 133, 101–111. doi:10.1016/j.compositesb.2017.09.029

frontiers | Frontiers in Built Environment

ORIGINAL RESEARCH
published: 21 October 2022
doi: 10.3389/fbuil.2022.895210

Application of Soft Computing for Estimation of Pavement Condition Indicators and Predictive Modeling

Shadi Hanandeh[1]*, *Ahmad Hanandeh*[2], *Mohammad Alhiary*[3] and *Mohammad Al Twaiqat*[3]

[1]Civil Engineering Department, Al-Balqa Applied University, Al-Salt, Jordan, [2]MIS Department, Applied Science Private University, Amman, Jordan, [3]Consultation Office, Jordan University, Amman, Jordan

The pavement management system is recognized as an assertive discipline that works on pavement indices to predict the pavement performance condition. This study used soft computing methods such as genetic algorithms and artificial intelligence to propose a modern generation of pavement indices for road networks in Jordan. The datasets used in this study were collected from multiple roads in Jordan, and 128 data points were used in this study. The input variables are the pavement condition index (PCI) and the international roughness index (IRI) in the artificial neural network (ANN) and gene expression programming (GEP) models. The output variable is the pavement serviceability rate (PSR). The results show an efficient performance benefit of using these techniques. In addition, the ANN and GEP models were able to predict the output variable with a reasonable accuracy, where the ANN model has an R^2 value of 0.95, 0.87, and 0.98 for the PCI, IRI, and PSR, respectively. The (R^2) values of the GEP model are 0.94, 0.89, and 0.99 for PCI, IRI, and PSR, respectively.

Keywords: pavement condition index, pavement management system, pavement serviceability rate, genetic algorithm, international roughness index, neural network

OPEN ACCESS

Edited by:
Yasmin Murad,
The University of Jordan, Jordan

Reviewed by:
Mansour Fakhri,
K.N. Toosi University of Technology, Iran
Ahmad N. Tarawneh,
Hashemite University, Jordan

***Correspondence:**
Shadi Hanandeh
hanandeh@bau.edu.jo

Specialty section:
This article was submitted to Computational Methods in Structural Engineering, a section of the journal Frontiers in Built Environment

Received: *13 March 2022*
Accepted: *19 May 2022*
Published: *21 October 2022*

Citation:
Hanandeh S, Hanandeh A, Alhiary M and Al Twaiqat M (2022) Application of Soft Computing for Estimation of Pavement Condition Indicators and Predictive Modeling. Front. Built Environ. 8:895210. doi: 10.3389/fbuil.2022.895210

INTRODUCTION

The pavement management system is determined by AASHTO and defined as a collection of instruments or techniques that can help decision makers in determining excellent approaches for implementing, assessing, and sustaining pavement layers in a practical situation over a specific time period. The maintenance and rehabilitation (M&R) can be designed for the different kinds of roads in pavement management systems. The pavement requires improvement and maintenance due to distress caused by traffic load and environmental conditions, so the first step to evaluate any road is performed by human measurement: this method has low efficiency and consumes a lot of time and labor. Moreover, modern tools were utilized to evaluate pavement distress by employing computerized programs to analyze data and inspect the pavement surface condition; although it is more efficient and faster, it is also difficult to automate the pavement rating. There are many pavement condition evaluation parameters used to measure and rate pavement distresses which reflect their overall condition. Some of these parameters are the pavement condition index (PCI), international roughness index (IRI), and pavement serviceability rate (PSR). Many studies proposed correlations between IRI and PCI as shown in **Eq. 1**.

$$\mathrm{IRI}_{(p)} = \left(\frac{79.933}{14.061 + \exp(0.048^{*}\mathrm{PCI})}\right), \tag{1}$$

where IRI is presented in m/km (Elhadidy et al., 2019).

The pavement condition index has been developed by the U.S. Army Corps of Engineers (1982). The PCI value indicates the type, quantity, and severity of the distress. This value can be used to identify the rate of deterioration, and it can be also be used to determine the appropriate time for maintenance and rehabilitation. The PCI value is a numerical value that ranges from 0 to 100, where the value equal to 100 represents the best pavement condition. The functional assessment of roads mostly involves the ride condition or surface texture of the road segment. The surface roughness reflects the longitudinal profile disturbances as the paving surface adversely affects car-driving quality, thus supplying the road users with this ease. Roughness is represented by calculating the international roughness index displayed in units of meters/kilometers. The present serviceability rate (PSR) measures the pavement serviceability which is the mean rating of each user panel rating for ride-ability; PSR ranges from 0 to 5, which 0 for very poor and 5 for very good. AASHTO recommended that 2.5 or more is acceptable for the highway segment.

Several studies have focused on the pavement management system (PMS) to obtain a correlation between PCI, PSR, and IRI. A powerful correlation that determines the relationship between PCI and IRI with a high value of the coefficient of determination (Dewan and Smith, 2002). This study found a reasonable relationship between IRI and PCI as presented in **Eq. 1**. The surface roughness can be conducted from the current roughness model when data are available on pavement surface distress; they created an equation between IRI and PCI for data given in this area. The limitation of this study is the absence of some distresses type and severity that may significantly influence the value of IRI; finally, the agencies provided recommendations to deal with the model and improve the proposed model, as shown in **Eq. 2**.

$$\mathrm{IRI} = 0.0171\,(153 - \mathrm{PCI}), \tag{2}$$

where IRI is presented in m/km.

Park et al. (2007) established the relationship between PCI (pavement surface distresses) and IRI (surface roughness). The surface roughness was investigated using Paver software for different North Atlantic highway pavement sections. They investigated another significant correlation between IRI and PCI for 63 data points from 20 pavement sections in different states of America. The collected data from nine states gave a poor model with $R^2 = 0.59$.

$$\mathrm{Log}\,(\mathrm{PCI}) = 2 - 0.436 \log\left(\frac{\mathrm{IRI}}{0.727}\right). \tag{3}$$

However, a variance analysis shows that there is a close association between IRI and PCI. Arhin et al. (2015) studied the pavement in the District of Columbia to obtain the correlation between IRI and PCI. IRI was collected using advanced machines that measure the regularity of pavement; on the other hand, the PCI was obtained by the subjective rate for the pavement condition and observing the distresses in the pavement. This study utilized the historical data (using 895 data points) of IRI–PCI to develop the model that is based on regression analysis, the method of forecasting IRI based on PCI through practical and pavement grouping.

The findings of the model indicated that the composite pavement was better than asphalt pavement, followed by concrete pavement. The proposed model was intended to be statistically significant with a 5% significance level; The R^2 amounts range between 56% and 82%. The presented equation is shown in **Eq. 4**.

$$\mathrm{PCI} = -0.224\,(\mathrm{IRI}) + 120.02, \tag{4}$$

where IRI is presented in miles).

Suprapto et al. (2017) studied another relationship between the international roughness index and structural numbers. They found that IRI is used for assessing the functional condition of the pavement, while the structural number (SN) is a parameter for assessing the structural condition of the pavement. The pavement condition data were observed using the Android application on smartphones; it is an efficient and low-cost application to collect the data. The main goal of this study is to determine the functional and structural conditions of the road. It was found that when the value of IRI was less than 4, the pavement is functionally in a good condition. Furthermore, when the value of structural number effective (SN_{eff}) is less than the value of the structural number future (SNF), then the structural condition of the road segments has not been able to serve traffic with a design life of 20 years. The previous model will be used to determine the best road maintenance strategies for the pavement using an overlay material when the value of IRI exceeds 4. For structural road maintenance carried out in various scenarios, the phased construction, and direct construction, direct construction gives better results than the phased construction scenarios for the design life of 5–20 years and 10–20 years. Elhadidy et al. (2019) developed a simplified regression model using the long-term pavement performance (LTPP) database to make the relationship between IRI and PCI. They used a total of 1,208 sections with 10,868 data from seven countries (Dubai, Australia, Belarus, Hungary, Norway, Russia, and Slovakia) that were used for model development, while 240 sections with 1,876 data points were used for the model validation. They suggested predicting PCI based on IRI using different highly accurate databases ($R^2 = 0.995$); the mean error for this model is very small with a small standard deviation.

$$\mathrm{IRI_P} = \left(\frac{79.933}{14.061 + \exp\,(0.048^{*}\mathrm{PCI})}\right), \tag{5}$$

where IRI is presented in m/km.

Al-Omari and Darter (1994) developed the relationship between IRI and presented serviceability rating. The main objectives of this research are to develop a predictive model for PSR as a function of profile IRI for composite, flexible, and rigid pavement. The second goal is to determine the level of rehabilitation based on the IRI and type of distress, derived from the long-term pavement performance (LTPP) database. This study was developed for each pavement type in the states of Louisiana, Michigan, New Jersey, New Mexico, Ohio, and Indiana, and all six states together. A nonlinear relationship

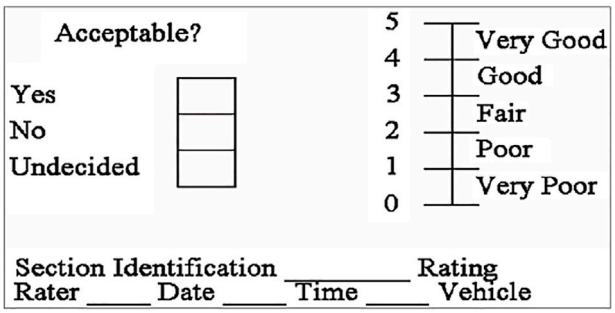


FIGURE 1 | Individual present serviceability rating form.

between IRI and PSR was developed for all types of pavement with an R^2 value = 0.97 by using the statistical analysis system (SAS) data set. In their model, they included IRI, PSR, and pavement type for every pavement section in each state as presented in **Eq. 6**.

$$PSR = 5^{*}e^{-0.26^{*}IRI}, \quad (6)$$

where IRI is presented in mm/m.

They found that there were no significant differences between the models for different states and pavement types. Gulen et al. (1994) developed another relationship between IRI and present serviceability index rating (PSI); a specially equipped car was used to measure the IRI. This study was conducted for Taiwan highway; the analysis is based on a backpropagation neural network methodology that evaluates the applicability of IRI to represent the pavement performance. Moreover, IRI can be predicted from pavement distresses rating using video images, which are recorded by a camera, and this is what we call a backpropagation neural network. They developed a correlation between mean present serviceability index rating and IRI from data that was given by the Indiana Department of Transportation (INDOT) for 10 randomly selected sections in Indiana with 1mi for each section at three roughness levels for all types of pavement that are presented in **Eqs 7–9**.

$$PSI = 7.21e^{-0.471IRI} \quad (R^2 = 0.84) \text{ for asphalt pavement}, \quad (7)$$

$$PSI = 14.05e^{-0.74IRI} \quad (R^2 = 0.93) \text{ for concrete pavement, and} \quad (8)$$

$$PSI = 9.0e^{-0.56IRI} \quad (R^2 = 0.84) \text{ for composite pavement}, \quad (9)$$

where IRI is presented in m/km.

The neural networks are used as an agent for evaluating pavement conditions. They improved the ANN scheme for the measurement of pavements' situation grade with regard to several cracks and distress. Neural network investigation confirms the ability to ascertain pavement evaluation indices (Eldin and Senouci, 1995). Lima et al. (2019) assessed the abilities of

TABLE 1 | Descriptive statistical analysis for three models' parameters.

Parameter	Number of data	Mean	Standard deviation	Median	Max	Min
PCI	128	73.66	27.54	80	100	10
IRI (mm/m)	128	2.33	1.66	1.9	9.6	0.44
PSR	128	3.02	0.95	3.17	4.5	0.5

FIGURE 2 | Different locations of the study area in Amman city.

ANN in discovering the visual condition index of different types of pavements. Several studies recommended genetic algorithms (GA) and artificial neural network paradigms to estimate the pavement evaluation indices, pavement material characterization, and structural engineering fields such as prediction of layer modulus and other fields such as using ANN to estimate subgrade and base course stabilization (Gopalakrishnan et al., 2017; Hanandeh et al., 2020a, 2020b; You et al., 2020; Piryonesi and El-Diraby, 2021; Ye et al., 2021; Zuhair Murad, 2021; Al Bodour et al., 2022; Hanandeh, 2022).

Barzegaran et al. (2021) used IRI prediction models from pavement distresses which could be an alternative solution. In this study, 507 km of asphalt pavements in Kermanshah, Iran, were studied utilizing IRI and pavement surface and evaluation rating (PASER) as a quick and cost-effective index. PASER developed IRI prediction models utilizing regression (R^2 = 0.66) and artificial neural networks (ANNs) (R^2 = 0.69). The model development (R^2 0.97) and validation (R^2 = 0.85) findings suggested that the ANN model performed well.

Fakhri et al. (2021) used the laser crack measurement system (LCMS) to identify and measure surface distresses and roughness on 268 km of key highways in Iran. LCMS gives the highest detection and measurement accuracy when compared to manual surveys. The results show that the ANN outperforms the other models for the 100-m sections, with a coefficient of determination (R^2) of 0.82. With a 0.94 correlation, the best case was utilizing ANN in 100-m chunks for places with a mild climate and medium traffic intensities.

Fakhri and Dezfoulian (2019) used artificial neural networks and regression models to develop a relationship between deflection bowl parameters derived from a falling weight deflectometer (FWD) and two pavement performance indices, the international roughness index (IRI) and pavement surface evaluation and rating index (PASER). Our model provides a satisfactory connection between IRI, PASER, and structural indices based on deflection measurements, as shown by the results. The superiority of ANN performance over non-intelligent models may be seen when comparing the outcomes of ANN with regression models. The results of this research show that combining the IRI and PASER indices results in a more accurate structural pavement evaluation.

This study studied maintenance and rehabilitation strategies for multilane roads in the Irbid governorate. Also, we developed proposed models using artificial neural networks and genetic algorithms to find the relationship between the parameters (PCI, IRI, and PSR) that measure the pavement performance, and this was achieved by using 61 section samples from different locations. These methods reduced the number of pavement tests; therefore, they saved cost and time. The pavement engineers can use one index to evaluate the pavement performance.

TABLE 2 | GEP setting parameter.

Function set	+, –, *, and /
Genes	3
Chromosomes	40
Head size	6
Linking function	Addition
Constant per gene	5
Mutation rate	0.06
Inversion rate	0.09
Transposition rate	0.15
One-point recombination rate	0.4
Two-point recombination rate	0.3
Two-point recombination rate	0.3
Gene transportation rate	0.1
Gene recombination rate	0.1

DATA COLLECTION METHODS

The data were collected from different locations in Jordan by conducting field measurements for suburban and rural divided multilane highways located in the Irbid governorate road network (Abdulrahim, 2018). Many different types of equipment were used to develop 128 data points for 61 segments. These segments were selected at different locations in Jordan and then coordinated using the android GPS application called "share my GPS location." The PAVER technique used visual inspection for 40–50 m for each section with an area ranging from 140 to 326 m^2, based on the width of the corridor to assess distresses' kind, amount, and severity. Before beginning the examination, the safety concerns were addressed by closing the part under investigation with reflective cones. The visual data were categorized on a specific PAVER sheet, and the density of each distress was calculated using **Eq. 10**.

$$\text{Density} = \frac{\text{Distress Amount}}{\text{Sample Unit Area}} * 100. \tag{10}$$

The present serviceability rating was calculated on the identical sections using the method shown in **Figure 1** and following the Federal Highway Administration's (FHWA) instructions. To limit the difference between the evaluation groups, the average individual rating was obtained for each four-person evaluation panel. The data for the international roughness index were collected in the field for the same sections that were used to calculate PCI. The results were recorded in millimeters per meter (mm/m) using Roughometer III equipment with sensors mounted on a vehicle. IRI was calculated from an observed longitudinal road profile by multiplying the output of a quarter-car model by the length of the profile to yield a summary roughness index with units of slope.

A comprehensive statistical analysis was conducted on the variables of the models used in this work, including mean, standard deviation, average, maximum, and minimum values displayed in **Table 1**.

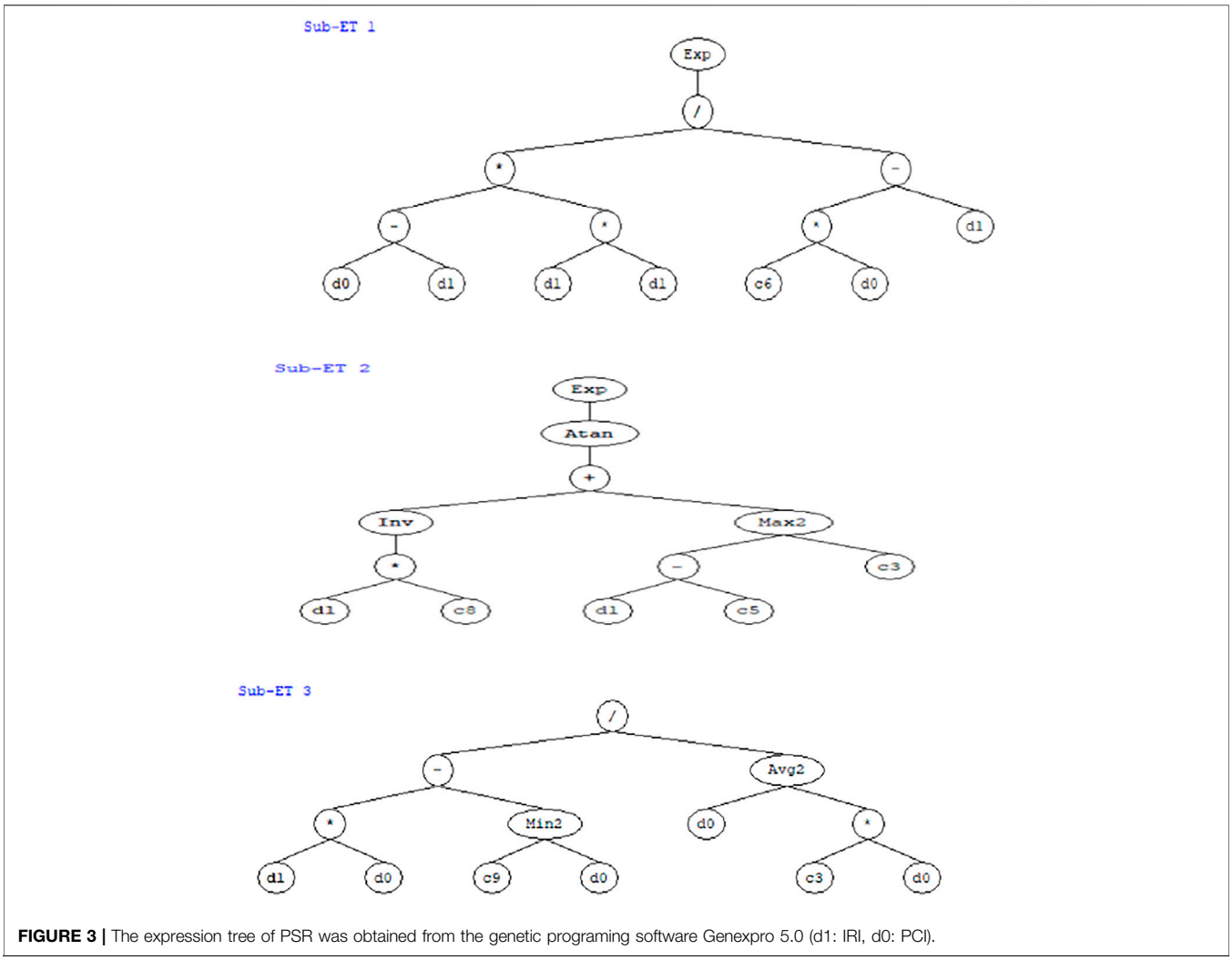


FIGURE 3 | The expression tree of PSR was obtained from the genetic programing software Genexpro 5.0 (d1: IRI, d0: PCI).

LOCATION OF DISTRESSES

Using the coordinates collected during the field study, the maps for each street in the study area were prepared as shown in **Figure 2**.

METHODOLOGY

In this study, two machine learning methods were used to analyze the collected data genetic algorithm and an artificial neural network. A genetic algorithm is a theoretical algorithm that depends on historical information; it is based on a survival-of-the-fittest concept. This is used as a search point with improved performance. In the case of maintenance management, this technique shows high performance with problems that have huge variable characteristics. ANN is a high-performance model that is used to correlate the component of the system together using the numerical load that is turned based on input, processing, output, and experience. The ANN model is used by pavement engineers as a tool to solve and predict complex system problems and models. To maintain the ANN at an acceptable level and as a prospect for the future, it must be combined with engineering techniques. ANNs, like a human brain, use experience, examples, and practice to learn.

GENETIC ALGORITHM

A genetic algorithm (GA) is a technique for search based on natural selection and genetics. GA is based on Darwin's evolution theory, which simulates the fundamental method of evolution to reach the most optimal form of the objective function. The genetic algorithm has three processes that included selection, crossing, and mutations. The individual's improvement by GA was shown by Holland (1975), generalized by his students, which allows a population composed to reach the maximum fitness under specified rules. GAs use historical information to solve problems and perform better solutions for the new generation (Goldberg and Holland, 1988). Gene expression programming (GEP) software is a branch of GP that develops software programs

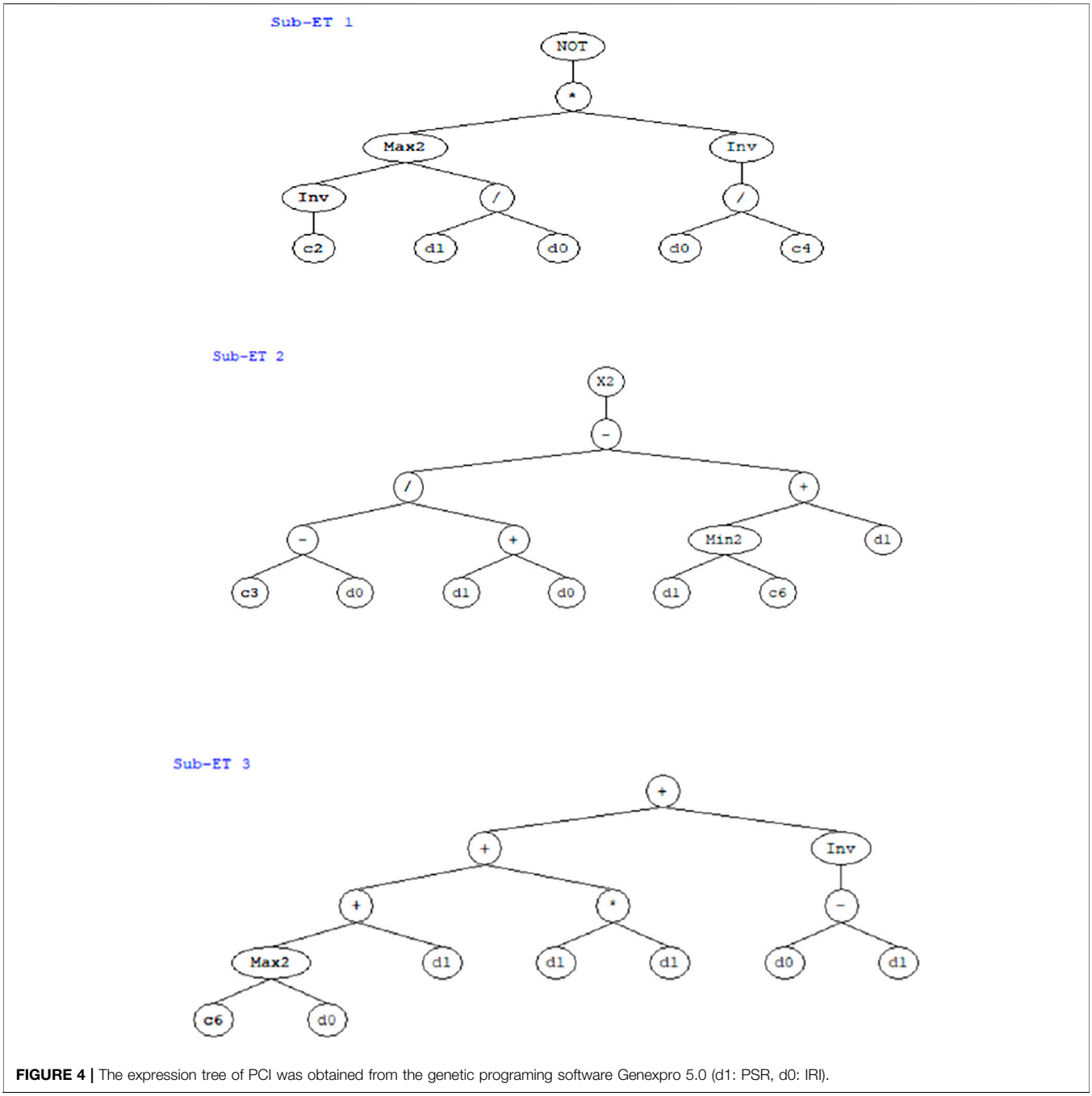


FIGURE 4 | The expression tree of PCI was obtained from the genetic programing software Genexpro 5.0 (d1: PSR, d0: IRI).

of different sizes and shapes encoded in linear chromosomes of fixed length. The chromosomes are built from many genes. Each gene is encoding a smaller subprogram; therefore, the two main parameters of GEP are the chromosomes and expression trees (ETs). Decoding information (from the chromosomes to the ETs) is called a translation process. In this research, GeneXpro5.0 software based on the Darwinian principle to solve problems was used for 86 data points to predict correlations between PCI, IRI, PSR, and the 42 data points for testing. There are five main steps to create the model on the software as follows:

1. Determine the fitness function.
2. Choose the number of terminals needed for the beats model (50,000 terminals were chosen in this research).
3. Determine the number of genes
4. Prepare to use gene expression programming by choosing the function length
5. Choose the set of genetic operations that cause variation and their rate.

After the trials arrived at 50,000 iterations, the run to get the optimum model has been stopped based on clear parameters

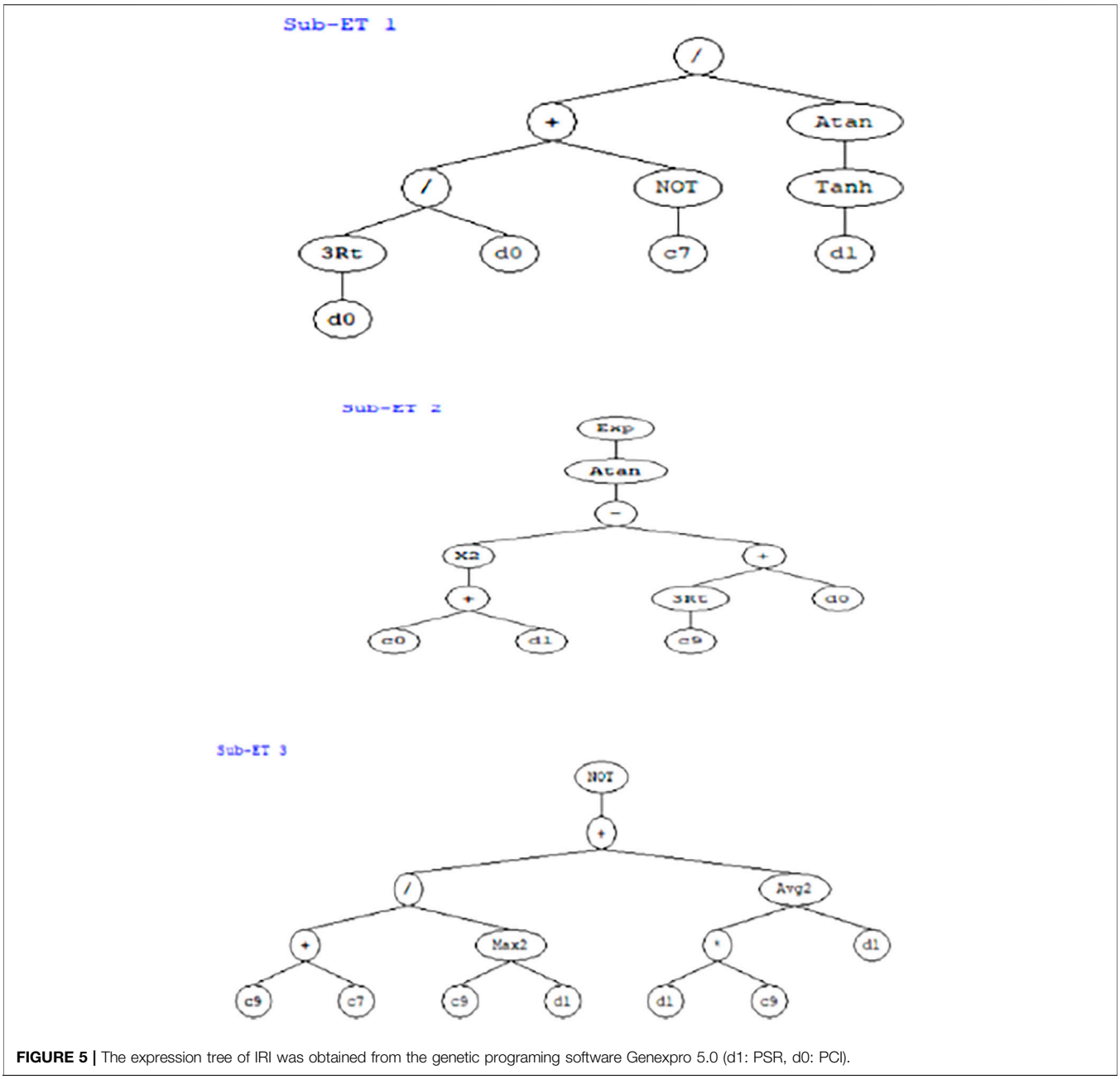


FIGURE 5 | The expression tree of IRI was obtained from the genetic programing software Genexpro 5.0 (d1: PSR, d0: PCI).

(R^2 value > 0.80), and the correlation coefficient was chosen and illustrated ETs that indicated a perfect solution with 3 generations as shown in **Figures 2–4**, respectively.

TREE MAPPING TO PREDICT (PCI, IRI, AND PSR) BY USING GA

The chosen GEP model was generated utilizing three genes with addition as a connecting role. The GEP criterion for the three proposed models is presented in **Table 2**, and the expression trees for the three models were presented in **Figures 3–5**.

The illustrated ETs in C++ language were translated to the mathematical equations stated in **Eqs 10–12**, respectively. The produced genetics algorithm pattern is performed in the manner of the representation tree as presented in **Figures 2–4**. Three sub-trees were developed (Sub-ET-1, 2, and, 3) and converted into a scientific equation model as given in **Eqs 11–13**, respectively.

$$\mathrm{PSR} = e^{\frac{[\mathrm{PCI}-\mathrm{IRI}]*\mathrm{IRI}^2}{[-5.539\mathrm{PCI}]-\mathrm{IRI}}} + e^{\tan^{-1}\left[\frac{1}{3.419*\mathrm{IRI}}+\max\{(\mathrm{IRI}-3.899),2.464\}\right]} + \frac{2[\mathrm{IRI}^{*}\mathrm{PCI}-3.877]}{\mathrm{PCI}-6.589\mathrm{PCI}}, \tag{11}$$

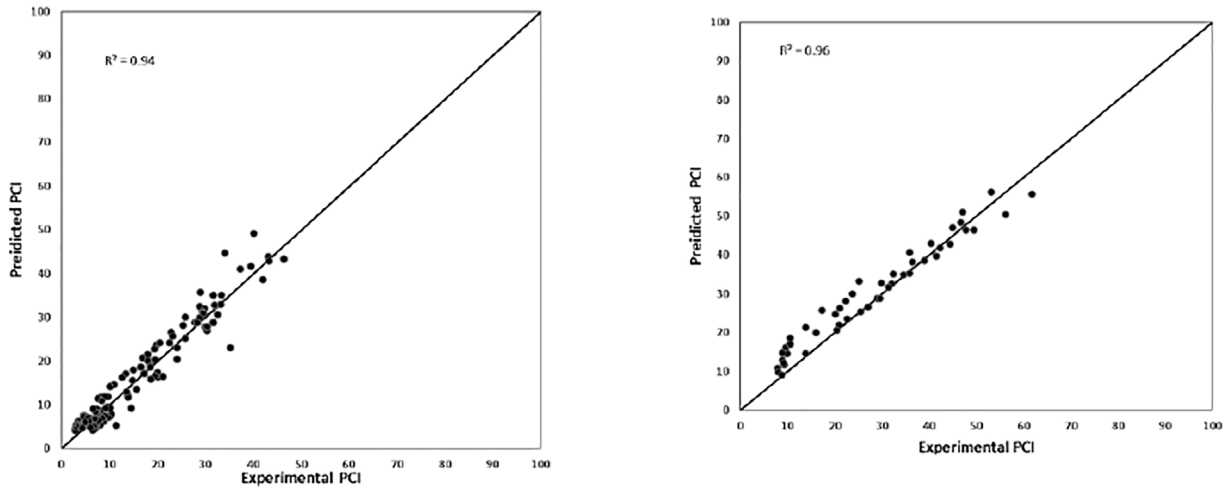


FIGURE 6 | Correlation of the predicted PCI versus measured PCI for testing and training data.

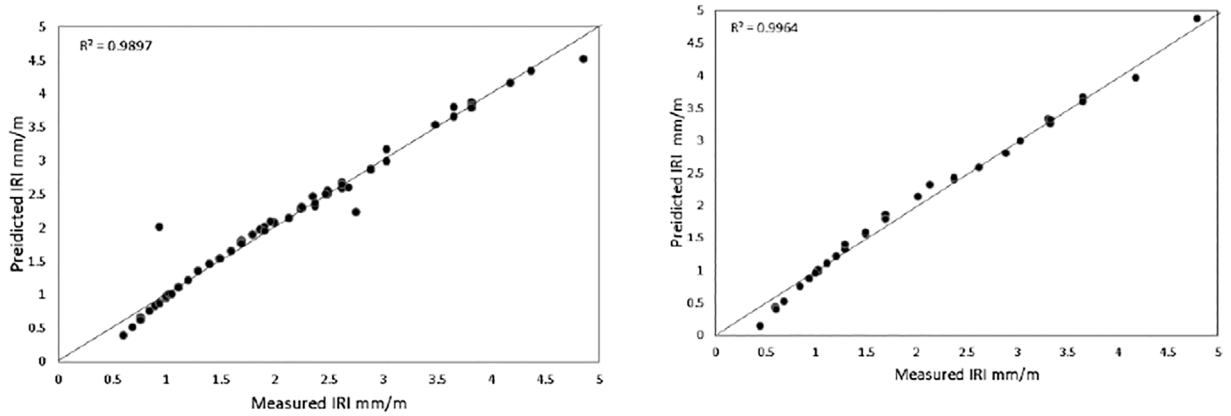


FIGURE 7 | Correlation scheme of predicted IRI against IRI for training and testing data.

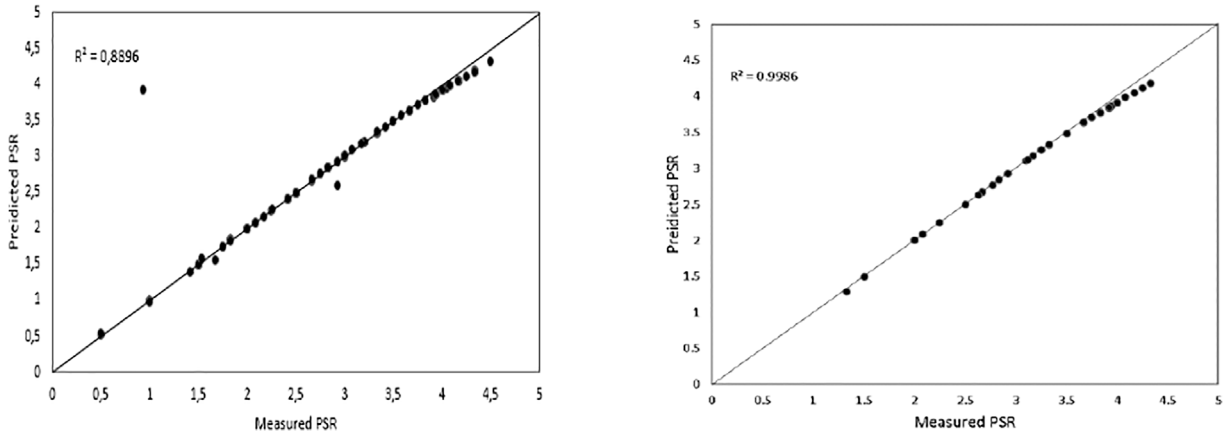


FIGURE 8 | Correlation scheme of predicted PSR against PSR for training and testing data.

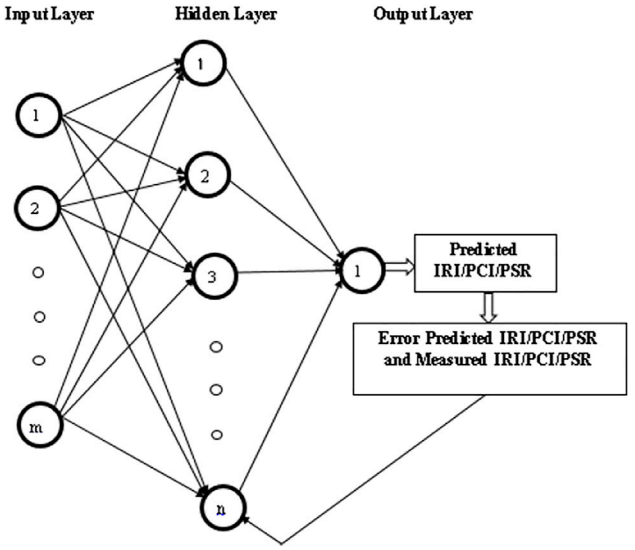


FIGURE 9 | Structure of backpropagation neural network.

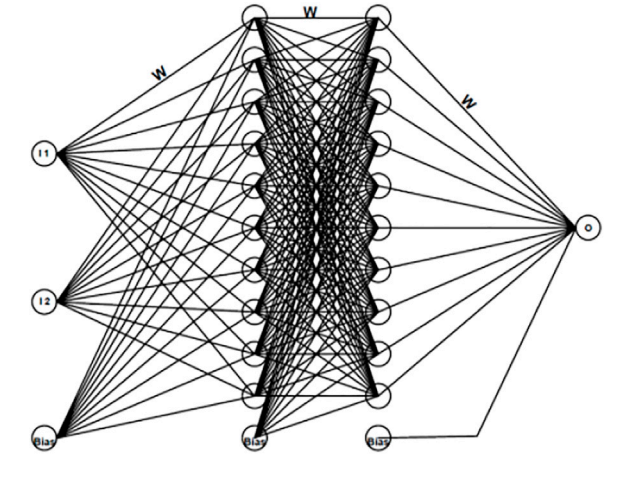


FIGURE 10 | ANN structure.

TABLE 3 | Output and input for developed ANN model.

Model number	Output (O)	Input 1 (I_1)	Input 2 (I_2)
1	PCI	IRI	PSR
2	IRI	PSR	PCI
3	PSR	IRI	PCI

$$\mathrm{PCI} = 1 - \max\left\{1.358, \frac{\mathrm{PSR}}{\mathrm{IRI}}\right\} * \frac{3.982}{\mathrm{IRI}} + \left[\frac{-7.292 - \mathrm{IRI}}{\mathrm{PSR} + \mathrm{IRI}} - \min\{\mathrm{PSR}, 4.251\} + \mathrm{PSR}\right]^2 + \max\{4.791, \mathrm{IRI}\} + \mathrm{PSR} + \mathrm{PSR}^2 + \frac{1}{\mathrm{IRI} - \mathrm{PSR}}, \quad (12)$$

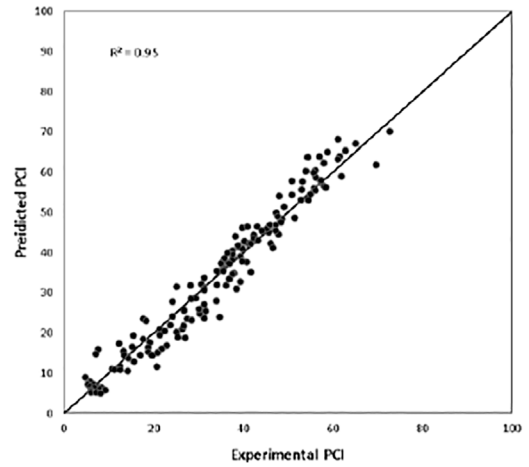
$$\mathrm{IRI} = \frac{\mathrm{PCI}^{\frac{1}{3}} + 2.876}{\tan^{-1}(\tan \mathrm{PSR})} + e^{\tan^{-1}\left[(8.623+\mathrm{PSR})^2 - (1.372+\mathrm{PCI})\right]} + \left[1 - \left[\frac{5.6332}{\max\{2.5832, \mathrm{PSR}\}} + \frac{3.5832^{*}\mathrm{PSR}}{2}\right]\right]. \quad (13)$$

Here, R^2 ranges from 0 to 1 which indicates the proposed model; if the model does not improve prediction, the value is zero, otherwise, the value is 1 for the perfect prediction model. The model is exhibited in **Figures 6–8**, which show the regression plot for predicted PCI, IRI, and PSR, respectively, versus measured for training and testing data.

ARTIFICIAL NEURAL NETWORK

An artificial neural network is a valuable computational tool based on a mathematical simulation model, ANNs derive the output data by processing the information, and this is achieved by precisely training the network. Ceylan et al. (2002) used ANNs as pavement structural analysis tools for the rapid and accurate prediction of critical responses and deflection profiles of flexible pavements subjected to typical highway loadings. Solanki (2013) developed a database for stabilized subgrade soils in Oklahoma and developed artificial neural network (ANN) models to estimate the resilient modulus (Mr). ANNs consist of multiple layers of the analogous processing component; it also contains hidden layers connected to the adjacent layer that can be modified and improve the performance during training by changing the weighting factor according to the predetermined algorithm. ANNs treat the (input–output) data frequently, and they can be trained by repeating the process. Many types of algorithms can be used to train the network: the training algorithms can be divided into two types: supervised and unsupervised. The supervised algorithms adjust the weights and the thresholds using the input and target output values, while the unsupervised algorithms only use the input values. The supervised training algorithms include backpropagation which is the most popular algorithm for training ANNs. The backpropagation ANNs, developed by Meier, consisted of frequently processing training examples to the network. ANN contains several highly interconnected processing components known as "nodes." The net input is identified using weighting value, input data, and bias value in the processing element. Thus, the input data will be converted to output data by providing a function of these net inputs. The structure of the ANN design consists of a collection of nodes, hidden layers, and output data. This study introduces ten nodes, two hidden layers, and one output variable data. The nodes of input data are acquired by field tests, and then these data will be transmitted to a hidden layer; in this layer, the data are treated and obtained as the output data. The trained neural network models will eventually enable pavement engineers to easily solve complex resource-intensive problems instead of using traditional techniques.

Figure 9 displays the arrangement of the neural network pattern, as noticed in several input parameters. By altering the number of hidden layers and achieving the most suitable model.



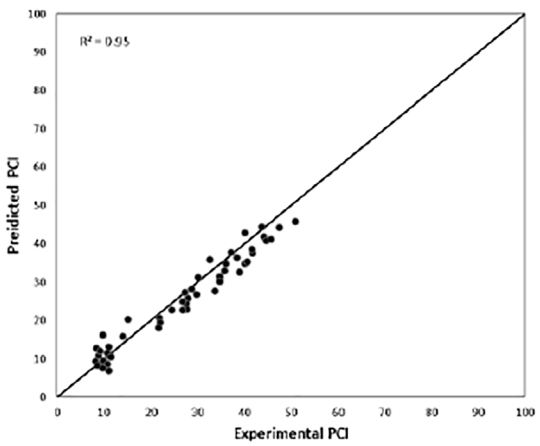
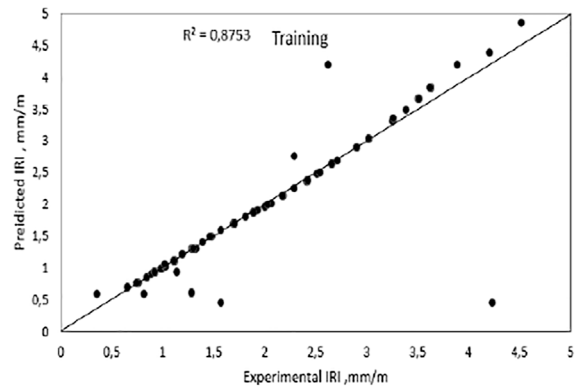


FIGURE 11 | The regression plot of predicted model 1 PCI against experimental PCI for testing and training data.



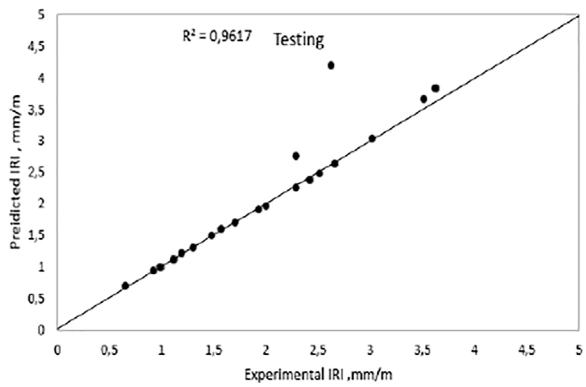
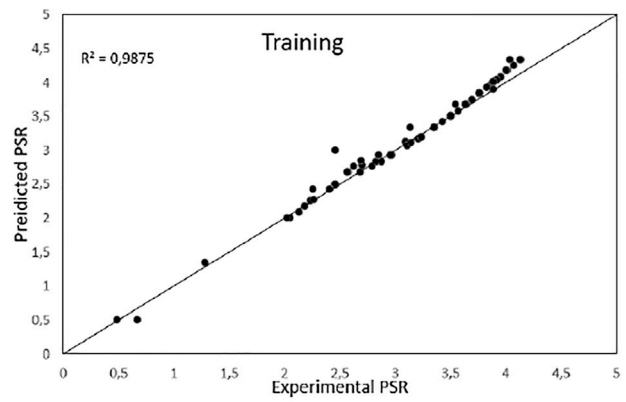


FIGURE 12 | The regression plot of predicted model 2 IRI against experimental IRI for testing and training data.



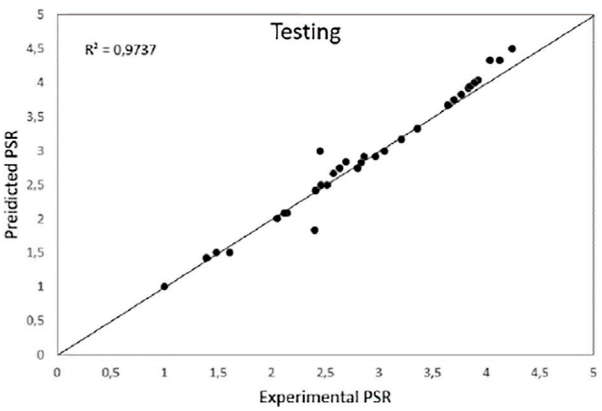


FIGURE 13 | The regression plot of predicted model 3 PSR against experimental PSR for testing and training data.

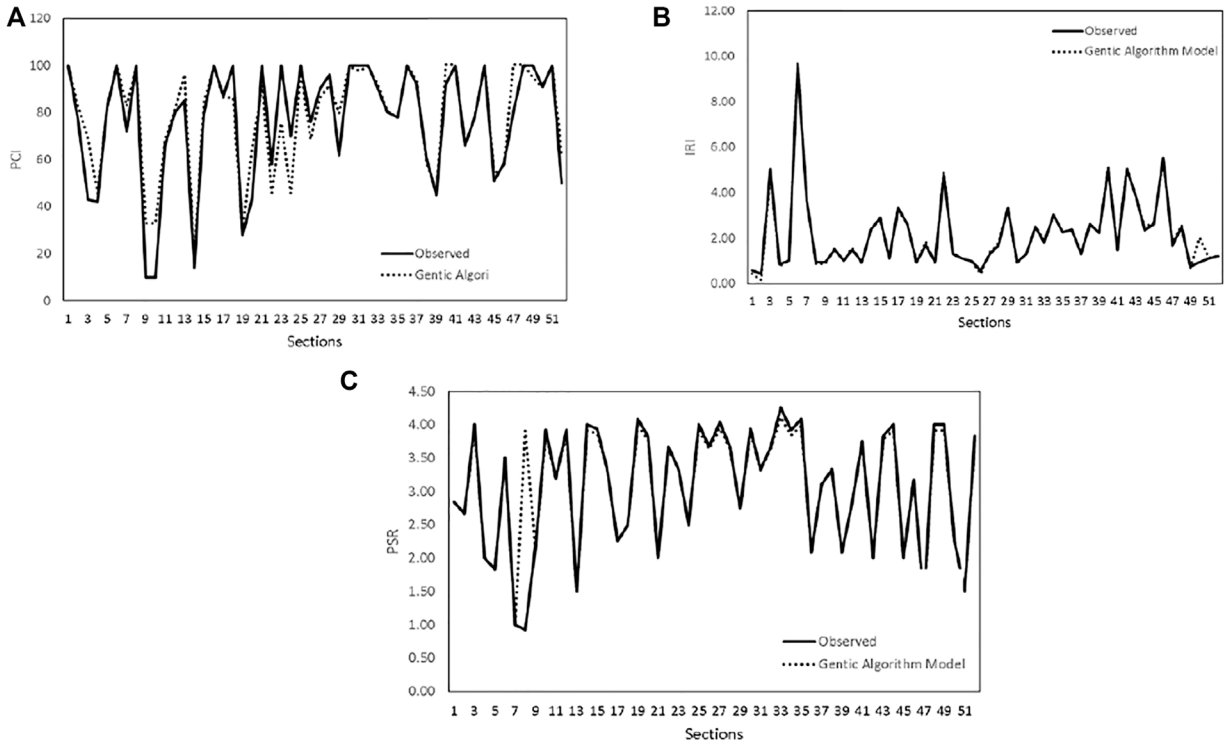


FIGURE 14 | Predicted PCI **(A)**, IRI **(B)**, and PSR **(C)**, respectively, using GEP.

In addition, the last layer called the output of the backpropagation process is shown, in **Figure 10**, to get a minimum error.

There are five main steps to create the ANN model on software as follows: The dependent and independent data need to be determined for each model, as shown in **Table 3**. The inputs and outputs at each node in the first hidden layer need to be measured (H1 to H10). $\sum WI + H$ measures inputs and outputs at each node in the second hidden layer (H1 to H10). The inputs and outputs at the nodes in the output layer O are measured.

In this study, three models were developed by ANN as presented in Table 5. The structure of developed models contains one output (O) and two inputs (I_1 and I_2) with two hidden layers that include 10 neurons for each layer and bias illustrated in **Figure 9**. The input leaf node receives the data from the experimental work and passes it to the neural node of the hidden layer without evaluation. The hidden layer then processes the data from the input layer and retrieves the valuable values to form the output layer. The neurons in the last layer are called the output layer, which can provide network predictions for external work through PSR/PCI/IRI mode values. RELU is used as an activation function, and the number of periods is 1,000, which generates a minor error. Therefore, the neural network establishes 10 × 10 weight (W) and 10 bias values for connecting the input layer and the hidden layer, and a 10 × 10 weight (W_2) and 2 bias values for the PSR/PCI/IRI connecting the hidden layer. The structure is shown in **Figure 10**. The input leaf node receives the data from the experimental work and passes it to the neural node of the hidden layer without performing the evaluation. The hidden layer then processes the data from the input layer, retrieves valuable values to form the input, and links each input. Others press pesos (w). The last layer of neurons is called the output layer, which can provide network predictions for external work on PSR/PCI/IRI mode values. RELY acts as a trigger function, and the number of epochs is 1,000, which produces a minor error. Therefore, the neural network establishes 10 × 10 weight (W) and 10 bias values for connecting the input layer and the hidden layer, and a 10 × 10 weight (W_2) and 2 bias values for the PSR/PCI/IRI connecting the hidden layer.

MODEL TO PREDICT PCI, IRI, AND PSR BY USING ANN

This model utilized IRI and PSR as the input parameters and the output is IRI; the model includes two hidden layers and each hidden layer includes 10 neurons as shown in **Figures 11–13**.

MODEL VERIFICATION AND APPLICATION

To confirm the three proposed models, obtained from the genetic algorithm method in this research, 52 sections were selected to confirm the efficiency of the proposed model PCI, PSR, and IRI in different locations in Jordan as shown in **Figure 14**. The genetic algorithm models generated in this study were employed to estimate the PCI, PSR, and IRI. The relative deviation, **Eq. 14**,

was employed to discover the deviation within the measured and proposed model to verify the efficiency of the model.

$$\text{Relative deviation} = \left|\frac{\text{observed value} - \text{model value}}{\text{model value}}\right| * 100\%. \quad (14)$$

CONCLUSION

This study proposed innovative methods to predict the pavement performance indices PSR, PCI, and IRI. Soft computing techniques were used to evaluate the pavement performance conditions. Two methods of machine learning artificial neural networks were employed to evaluate the pavement evaluation indices. Genetics algorithms and methods were performed to predict PSR/PCI and IRI values. The artificial neural network with two hidden and 10 nodes shows a reliable method to predict PSR/PCI and IRI. The model was simulated in Python software to decide their confidence; the coefficient of determination value was our standard in the diagnosis of the model performance. Using trial and error to achieve the most reliable hidden layer numbers for the best regression, the ANN has an adequate evaluation ability to determine the output PSR, PCI, and IRI, the value of PSR obtained by Python software for the first model was 0.98. On the other hand, the GN model developed better results than ANN and the generated equations from GA are applicable for practical use; it may also save the effort and cost for site inspection. In addition, additional measured data were used in this study to verify the proposed models.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material; further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

SH, conceptualization, programming methodology, and writing—original draft. AH: methodology, form analysis, data analysis, and writing. SH: data collection and analysis. MA and MT: data interpretation and writing—style editing and formatting.

REFERENCES

Abdulrahim, M. (2018). *Estimation of Free Flow Speed at Suburban and Rural Divided Multilane Highways*.

Al Bodour, W., Hanandeh, S., Hajij, M., and Murad, Y. (2022). Development of Evaluation Framework for the Unconfined Compressive Strength of Soils Based on the Fundamental Soil Parameters Using Gene Expression Programming and Deep Learning Methods. *J. Mat. Civ. Eng.* 34 (2), 4087. doi:10.1061/(ASCE)mt.1943-5533.0004087

Al-Omari, B., and Darter, M. I. (1994). Relationships between International Roughness Index and Present Serviceability Rating. *Transp. Res. Rec.* (1435).

Arhin, S. A., Williams, L. N., Ribbiso, A., and Anderson, M. F. (2015). Predicting Pavement Condition Index Using International Roughness Index in a Dense Urban Area. *J. Civ. Eng. Res.* 5 (1), 10–17. doi:10.5923/j.jce.20150501.02

Barzegaran, J., Shahni Dezfoulian, R., and Fakhri, M. (2021). Estimation of IRI from PASER Using ANN Based on K-Means and Fuzzy C-Means Clustering Techniques: a Case Study. *Int. J. Pavement Eng.*, 1–15. doi:10.1080/10298436.2021.2000988

Dewan, S. A., and Smith, R. E. (2002). Estimating International Roughness Index from Pavement Distresses to Calculate Vehicle Operating Costs for the San Francisco Bay Area. *Transp. Res. Rec.* 1816, 65–72. doi:10.3141/1816-08

Eldin, N. N., and Senouci, A. B. (1995). A Pavement Condition-Rating Model Using Backpropagation Neural Networks. *Comput. Civ. Infrastruct. Eng.* 10, 433–441. doi:10.1111/j.1467-8667.1995.tb00303.x

Elhadidy, A. A., El-Badawy, S. M., and Elbeltagi, E. E. (2019). A Simplified Pavement Condition Index Regression Model for Pavement Evaluation. *Int. J. Pavement Eng.* 22, 643–652. doi:10.1080/10298436.2019.1633579

Fakhri, M., Karimi, S. M., and Barzegaran, J. (2021). Predicting International Roughness Index Based on Surface Distresses in Various Climate and Traffic Conditions Using Laser Crack Measurement System. *Transp. Res. Rec.* 2675, 397–412. doi:10.1177/03611981211017906

Fakhri, M., and Shahni Dezfoulian, R. (2019). Pavement Structural Evaluation Based on Roughness and Surface Distress Survey Using Neural Network Model. *Constr. Build. Mater.* 204, 768–780. doi:10.1016/j.conbuildmat.2019.01.142

Goldberg, D. E., and Holland, J. H. (1988). Guest Editorial Genetic Algorithms and Machine Learning. *Mach. Learn* 3, 95–99.

Gopalakrishnan, K., Khaitan, S. K., Choudhary, A., and Agrawal, A. (2017). Deep Convolutional Neural Networks with Transfer Learning for Computer Vision-Based Data-Driven Pavement Distress Detection. *Constr. Build. Mater.* 157, 322–330. doi:10.1016/j.conbuildmat.2017.09.110

Gulen, S., Woods, R., Weaver, J., and Anderson, V. L. (1994). Correlation of Present Serviceability Ratings with International Roughness Index. *Transp. Res. Rec.* 1435, 27.

Hanandeh, S., Alabdullah, S. F., Aldahwi, S., Obaidat, A., and Alqaseer, H. (2020a). Development of a Constitutive Model for Evaluation of Bearing Capacity from CPT and Theoretical Analysis Using Ann Techniques. *Geomate* 19, 229–235. doi:10.21660/2020.74.36965

Hanandeh, S., Ardah, A., and Abu-Farsakh, M. (2020b). Using Artificial Neural Network and Genetics Algorithm to Estimate the Resilient Modulus for Stabilized Subgrade and Propose New Empirical Formula. *Transp. Geotech.* 24, 100358. doi:10.1016/j.trgeo.2020.100358

Hanandeh, S. (2022). Introducing Mathematical Modeling to Estimate Pavement Quality Index of Flexible Pavements Based on Genetic Algorithm and Artificial Neural Networks. *Case Stud. Constr. Mater.* 16, e00991. doi:10.1016/j.cscm.2022.e00991

Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems : An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*.

Lima, D., Santos, B., and Almeida, P. (2019). Methodology to Assess Airport Pavement Condition Using GPS, Laser, Video Image and GIS. In Pavement and Asset Management. 301–307. doi:10.1201/9780429264702-36

Park, K., Thomas, N. E., and Lee, K. W. (2007). Applicability of the International Roughness Index as a Predictor of Asphalt Pavement Condition. *J. Transp. Eng.* 133 (12), 706–709. doi:10.1061/(ASCE)0733-947X(2007)133:12(706)

Piryonesi, S. M., and El-Diraby, T. E. (2021). Examining the Relationship between Two Road Performance Indicators: Pavement Condition Index and International Roughness Index. *Transp. Geotech.* 26, 100441. doi:10.1016/j.trgeo.2020.100441

Solanki, P. (2013). Artificial Neural Network Models to Estimate Resilient Modulus of Cementitiously Stabilized Subgrade Soils. *Int. J. Pavement Res. Technol.* 6 (3), 155–164. doi:10.6135/ijprt.org.tw/2013.6(3).155

Suprapto, M., and Setyawan, A. (2017). The Use of International Roughness Index and Structural Number for Rehabilitation and Maintenance Policy of Local

Highway. *In IOP Conf. Ser. Mat. Sci. Eng.* 176 (1), 012031. doi:10.1088/1757-899X/176/1/012031

Ye, W., Jiang, W., Tong, Z., Yuan, D., and Xiao, J. (2021). Convolutional Neural Network for Pothole Detection in Asphalt Pavement. *Road Mater. Pavement Des.* 22, 42–58. doi:10.1080/14680629.2019.1615533

You, L., Yan, K., and Liu, N. (2020). Assessing Artificial Neural Network Performance for Predicting Interlayer Conditions and Layer Modulus of Multi-Layered Flexible Pavement. *Front. Struct. Civ. Eng.* 14, 487–500. doi:10.1007/s11709-020-0609-4

Zuhair Murad, Y. (2021). Predictive Model for Bidirectional Shear Strength of Reinforced Concrete Columns Subjected to Biaxial Cyclic Loading. *Eng. Struct.* 244, 112781. doi:10.1016/j.engstruct.2021.112781

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor YM declared a shared affiliation with the authors MA, MT at the time of review.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

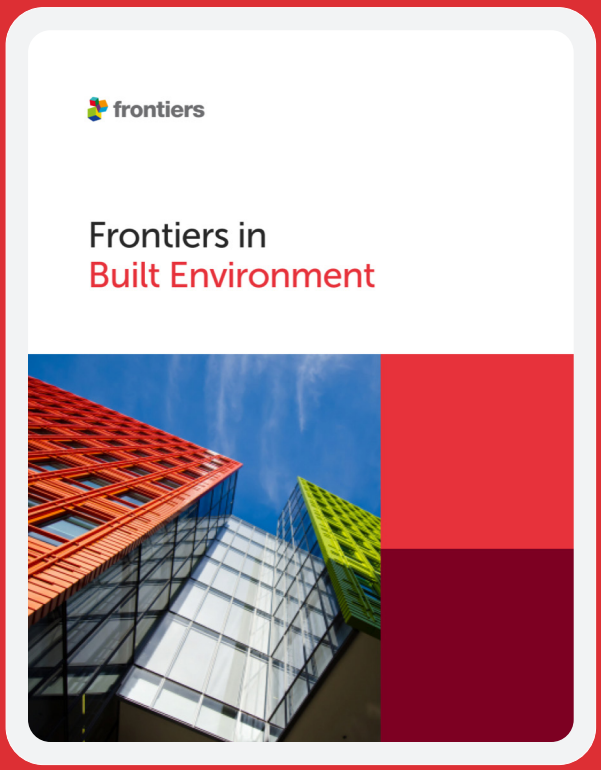
Frontiers in
Built Environment

Innovations in the engineering of sustainable buildings, cities, and urban spaces

An innovative journal that advances our knowledge of civil engineering. It focuses on the development of sustainable methodologies for the design and management of resilient buildings and infrastructure.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact