scientific reports

Check for updates

OPEN

Prediction of buckling damage of steel equal angle structural members using hybrid machine learning techniques

Nang Xuan Ho[1,3], Tien-Thinh Le[2,3 ✉], The-Hung Dinh[2,3] & Van-Hai Nguyen[2,3]

This article deals with prediction of buckling damage of steel equal angle structural members using a surrogate model combining machine learning and metaheuristic optimization technique. In particular, a hybrid Artificial Intelligence (AI)-based model involving Artificial Neural Network (ANN) and Particle Swarm Optimization (PSO) was developed and calibrated for the problem at hand. For this purpose, a database concerning compression tests of steel equal angle structural members was constructed from available resources with geometry variables such as length, width, thickness, mechanical properties of materials such as yield strength and initial imperfections (i.e. residual peak stress and initial geometric imperfections) and critical buckling load of columns. The hybrid PSOANN model was adopted because its prediction capability is higher than the traditional technique – i.e. scaled conjugate gradient (SCG). Indeed, ANN trained by PSO delivered better performance in terms of RMSE, MAE, ErrorStD, R^2 and Slope in comparison to ANN trained by SCG, for instance. RMSE decreases from 0.141 to 0.055; MAE decreases from 0.108 to 0.042; R^2 increases from 0.749 to 0.959, when switching from ANN alone to hybrid PSOANN, respectively. Moreover, a Partial Dependence (PD) investigation was performed to interpret the "black-box" PSOANN model.

Structural elements under compression are widely used in diverse projects due to the efficiency of exploitation: truss systems[1] or reinforced concrete columns[2–4]. In the case of instability of structural members, there is a significant risk of damage to the structure, see[2] for an example of buckling damage of the steel bar in a reinforced concrete column. Iasinski's work[5] and Johnson's formula[6] enabled us to look at with critical loads for structural elements with low and medium slenderness. However, these formulas apply only for isotropic, homogeneous materials and there are many assumptions about the structural elements: cross-section geometry, length, boundary conditions, concentric axial loads, lack of initial stress, etc. In reality, though, these conditions are not verified. The presence of these factors reduces the stability force-resistance of structural members under compression, compared to the original design[7–9].

In terms of modeling, theoretical investigations mainly focus on simply supported boundary conditions[10]. Recently, with the development of numerical analysis, iterative algorithms have been widely applied to nonlinear problems such as instability in the field of computational mechanics, for instance, Arc-Length method[11–13]. These methods could be applied to the numerical finite element scheme[14] to investigate the buckling behavior of structural members. However, there are a range of drawbacks in using finite element software including mesh scheme[15], iteration convergence[16], implementation of initial imperfections[17], etc. For example, for intricate shapes, automatic mesh generation using highly skewed tetrahedral elements may create distorted or poorly-shaped elements in thin-walled structures[18]. Iterative solvers in FEA may struggle to converge, particularly for problems with non-linearities such as material plasticity, contact mechanics, or large deformations[19]. In addition, simulating pre-existing stresses in welds or formed components requires advanced techniques like importing results from thermal analysis[20].

[1]Faculty of Vehicle and Energy Engineering, Phenikaa University, Yen Nghia, Ha Dong, Hanoi 12116, Vietnam. [2]Faculty of Mechanical Engineering and Mechatronics, Phenikaa University, Yen Nghia, Ha Dong, Hanoi 12116, Vietnam. [3]Phenikaa Research and Technology Institute (PRATI), A&A Green Phoenix Group JSC, No. 167 Hoang Ngan, Trung Hoa, Hanoi 11313, Cau Giay, Vietnam. ✉email: thinh.letien@phenikaa-uni.edu.vn

Recently, soft computing techniques, such as surrogate models, have begun to contribute solutions to various structural and mechanical engineering problems in a significant manner. As proved in the literature, soft computing techniques provide more reliable designs[21], evaluations[22] and predictions[23] than conventional methods. For instance[24], successfully developed a surrogate model for the design of self-compacting concrete based on various mixtures of ingredients, which could be extremely time-consuming in laboratory experiments. In another study[25], improved the design of square concrete steel tube columns by using an artificial neural network (ANN) model to explore the nonlinear relationship between geometry, mechanical properties and axial capacity of structural elements[26]. have proposed an ANN tool for prediction and design of an orthotropic steel-deck bridge, but the study has not yet been carried out, even using professional software tools. This computational model helped the authors to evaluate many different cases that would be time-consuming, effort-intensive and costly to investigate experimentally. In addition, the work of other researchers, including[27], has confirmed the efficiency of artificial intelligence techniques for the prediction of related mechanical engineering problems.

Within this context, the current work presents the development of a surrogate model based on the combination of Particle Swarm Optimization (PSO) and ANN to predict buckling damage to steel equal angle structural members. The hybrid PSOANN model was proposed because of its higher prediction capability than the single ANN model. The database and selection of variables are presented in Sect. 2.1, while Sect. 2.2 introduces the ANN and PSO algorithms. Section 3 presents the results and a discussion thereof, including the PSO optimization process and Partial Dependence (PD) investigation to explore the relationship between variables and output response. Finally, the dependence of the output response on input variables is shown through sensitivity analysis.

Materials and methods
Database and selection of variables
The selected geometry variables (length, width, thickness) were chosen because they directly influence the structural behavior under compression. These parameters are critical in defining the cross-sectional properties and slenderness ratio of steel equal angle members, which are known to significantly affect buckling performance[28]. Besides, yield strength was selected as a key material property because it defines the elastic limit of the material, directly influencing the transition to plastic buckling behavior[29]. Last but not least, initial imperfections, such as residual stress peak and initial geometric deviations, were included because they are well-documented to reduce the buckling capacity of members[30].

In this paper, 66 configurations of compression test on steel equal angle structural members were collected from the available literature[31]. In Ban et al.[31], an experimental program was conducted to test 420 MPa high-strength steel equal angle (see Fig. 1a) columns under axial loading. The boundary conditions of the tests were pinned-pinned, as shown in Fig. 1c. The min, max, average, standard deviation, coefficient of variation and quantile values at 25 and 75% of input variables are summarized in Table 1. It should be noticed that our
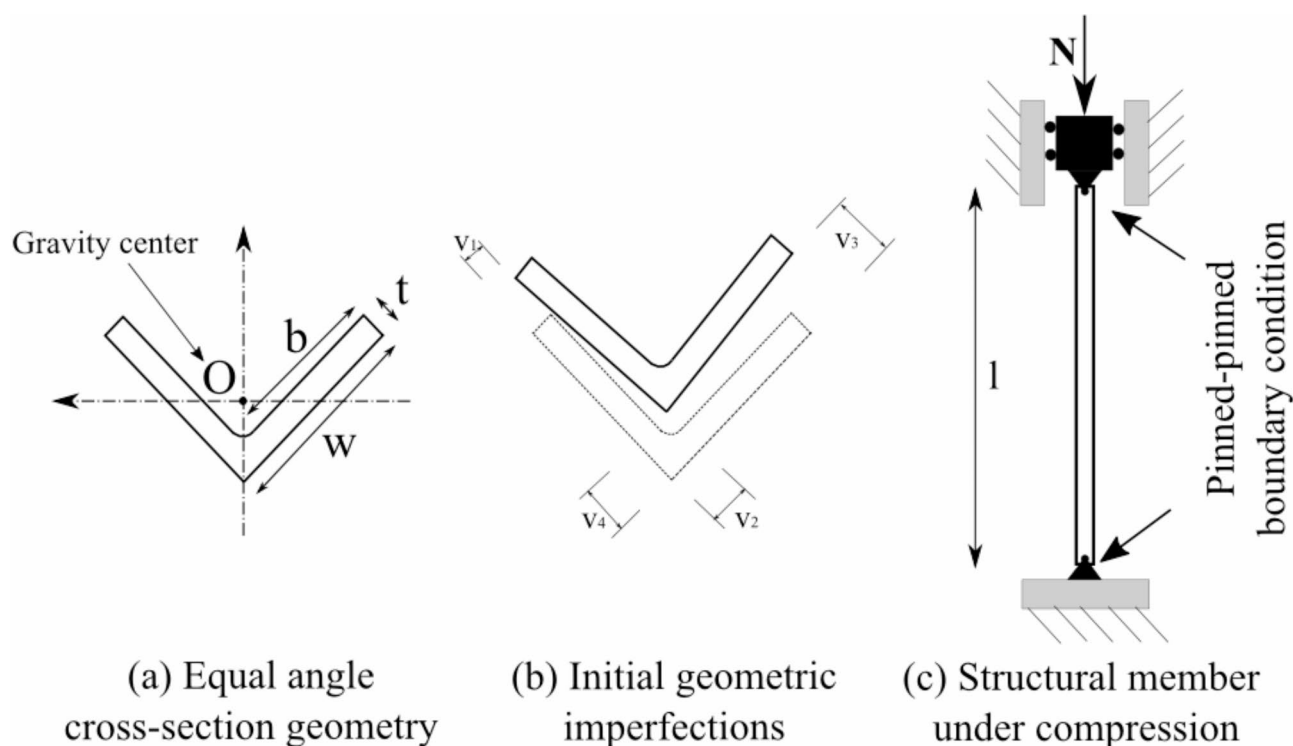

Fig. 1. Schematization including geometry parameters for (**a**) cross-section, (**b**) initial imperfections and (**c**) testing configuration.

Parameter	Notation	Unit	Role	Min	Q_{25}	Q_{50}	Q_{75}	Max	Average	Standard deviation	Coefficient of variation (%)
Length	l	mm	Input	749.2	1194.7	1546.9	2146.8	3838	1718.2	730.38	42.51
Thickness	t	mm	Input	7.77	9.79	10.04	11.76	13.79	10.31	1.75	16.96
w/t ratio	w_t	–	Input	11.96	14.81	15.55	15.98	16.48	15.26	1.04	6.82
b/t ratio	b_t	–	Input	9.5	12	12.78	13.4	13.4	12.57	0.9	7.16
Yield strength	f_y	MPa	Input	442.1	448.8	459.4	460.7	542.4	457.23	20.19	4.42
Residual stress peak	σ_r	MPa	Input	35.93	35.93	43.75	53.86	105.77	47.47	15.01	31.62
Initial imperfection	v_1	mm	Input	0.38	1.42	1.78	2.3	4.3	1.93	0.74	38.27
Initial imperfection	v_2	mm	Input	0.32	0.84	1.02	1.34	2.84	1.16	0.47	41.06
Initial imperfection	v_3	mm	Input	0.18	1.42	1.78	2.38	8.82	2.01	1.19	59
Initial imperfection	v_4	mm	Input	0.4	0.8	1.12	1.52	2.72	1.18	0.48	40.77
Critical buckling load	N_u	kN	Output	529.1	876.1	1081.35	1491.4	2177.2	1141.8	415.51	36.39

Table 1. Initial statistical analysis of the database used in this study.

Correlation coefficient r	l	t	w_t	b_t	f_y	σ_r	v_1	v_2	v_3	v_4	N_u
l	1.00	0.19	0.01	0.02	0.28	0.04	−0.01	0.46	0.19	0.24	−0.03
t		1.00	−0.43	−0.35	0.34	0.35	−0.01	−0.02	−0.04	0.14	0.96
w_t			1.00	0.97	−0.58	−0.95	0.32	−0.17	0.31	−0.02	−0.32
b_t				1.00	−0.59	−0.98	0.38	−0.16	0.34	−0.01	0.25
f_y					1.00	0.72	−0.06	0.21	0.00	0.15	0.17
σ_r						1.00	−0.35	0.18	−0.30	0.03	0.22
v_1							1.00	−0.10	0.49	−0.05	−0.02
v_2			sym.					1.00	0.16	0.35	−0.17
v_3									1.00	0.10	−0.06
v_4										1.00	0.09
N_u											1.00

Table 2. Correlation matrix of the database used in this study.

proposed model is valid for the range of variables shown in Table 1. High-strength steel was selected because it exhibits better mechanical behavior than normal strength steel, and therefore higher durability[32]. The main geometry parameters of the structural elements are the length l, the thickness t, the width w, the free outstanding width b of angle legs, and four initial geometric imperfections v_1, v_2, v_3, v_4 as shown in Fig. 1b. Also, the residual stress was measured and reported in the analysis[33]. The mechanical properties of the constituent materials are expressed by yield strength, also measured for each configuration, as indicated in Table 1. It should be noted that the columns were designed to exhibit slenderness ranging from 30 to 80, which covered a wide range, from medium to long columns. Further, the correlation matrix within the database is also given in Table 2, showing an initial evaluation (linear mode) of the relationship between input variables and output response.

The entire dataset was divided into training (70%) and testing (30%) subsets. Data points were randomly assigned to the training and testing sets to ensure an even distribution of variable ranges across both subsets. This randomness avoids bias toward specific parameter values in either set, therefore reduce the risk of overfitting. Moreover, to avoid bias in the training process of the ANN model, data variables were made to range from 0 to 1. This also reduces the risk of overfitting. It is interesting to notice that after training, the model's performance was assessed on the independent testing set, which it had never used during training.

Methods used
Artificial neural network

In this work, as the database contained a large number of input variables (i.e. 10), ANN was selected as the soft computing technique. As proved in the literature, ANN is highly efficient with large dimensional problems – for instance[25], in structural engineering. Figure 2 shows a representation of ANN, including the input, hidden and output layers, which are three basic elements of an ANN model. The three above layers are inter-connectable via artificial neurons (i.e. computational nodes), whose objective is to compute weight parameters of the ANN model. For a single-output prediction problem, ANN model generalizes the following non-linear function:

$$f : X \in \mathbb{R}^N \mapsto Y \in \mathbb{R}^1, \tag{1}$$

where X and Y represent the input and output vectors, respectively. The function f can be expressed as the following:
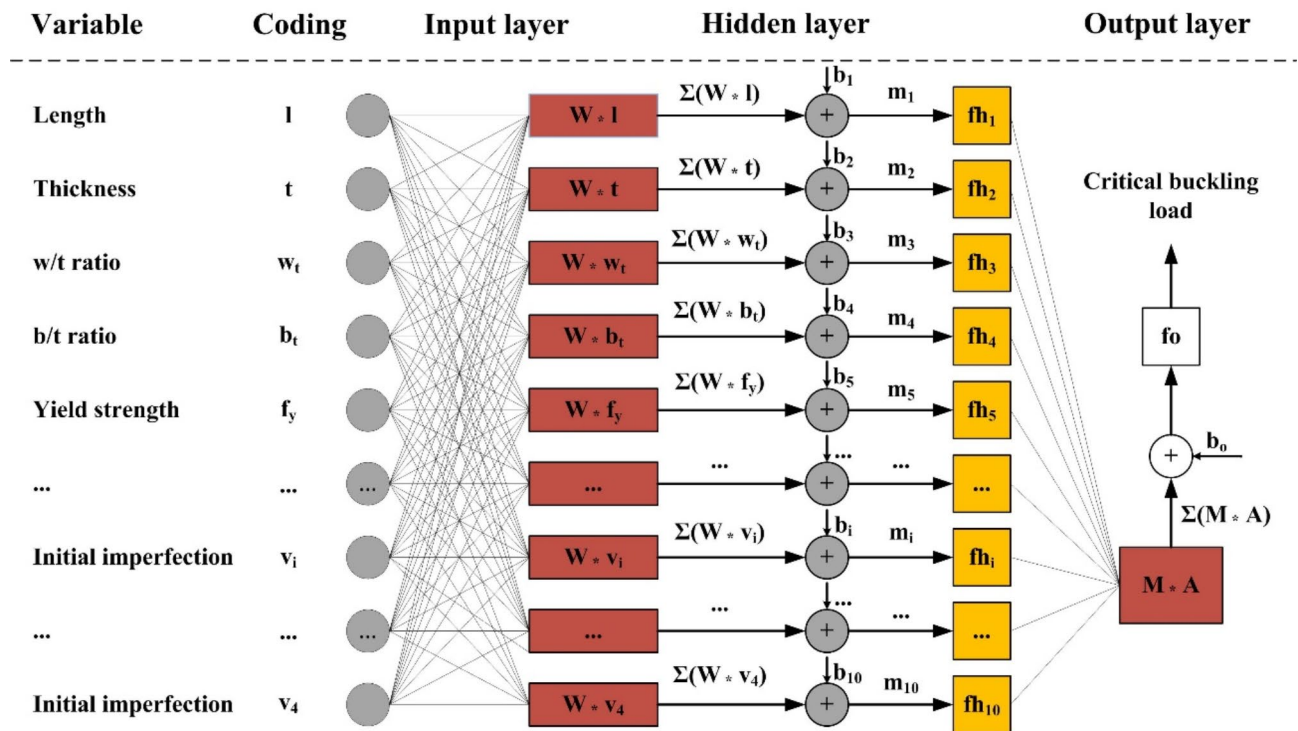
Fig. 2. The schematization of the ANN algorithm including 10 inputs used in this study.

$$X \mapsto f(X) = fo\left(M \times \left(fh\left(b + W \times X\right)\right) + b_o\right), \tag{2}$$

where W and M represent the weight matrix of hidden and output layers, respectively. fh and f_0 represent the activation functions of hidden and output layers, respectively. b and b_0 represent the bias vectors of hidden and output layers, respectively. Such a relationship is illustrated in Fig. 2.

$$m = W \times X + b \tag{3}$$

$$A = fh\left(W \times X + b\right) \tag{4}$$

$$\lambda = fo\left(M \times A + b_o\right) \tag{5}$$

There are numerous algorithms available for training ANN. Some, such as Levenberg-Marquardt backpropagation, require a large amount of memory to answer numerous classification problems[34], and the convergence rate in this approach is slow[35]. To tackle these challenges, other approaches have been developed, such as ANN models based on the principle of gradient descent[36], i.e. the first-order conjugate gradient algorithm. However, because the line search is performed at each iteration, this strategy is still costly. In this study, we employed the scaled conjugate gradient algorithm (SCG)[36] in conjunction with the PSO approach to train the ANN model. This type of evaluation enables us to investigate the performance of the PSO metaheuristic optimization technique.

Particle swarm optimization

Kennedy and Eberhart developed this efficient-swarm intelligence technique for addressing difficult optimization problems based on the social behavior of animals (e.g., a flock of birds or a school of fish)[37]. The basic idea is to move a swarm of particles iteratively to discover the global best position in a search space. The particles go across the search space, looking at various ordinary expressions for the optimum placements. The particles' final placements are optimum answers to the challenge at hand.

Assume that the population size, or number of particles, in a D-dimensional search space is m. The particle's location and velocity i ($i = 1, 2, \ldots, m$) are represented by x_i and v_i, respectively. Individual particles are associated with the best location in the swarm, pBest, and the best position of all particles, gBest. Figure 3 depicts the PSO method.

Figure 4 introduces the flowchart methodology of the current study. The raw data points are analyzed statistically for deducing correlation matrix. Then the entire dataset is divided into training (70%) and testing (30%) subsets. The ANN model is trained by optimizing weights and bias using PSO global optimization technique. The model prediction performance is then validated by using different quality metrics such as R^2, RMSE and MAE. Finally, sensitivity analysis is carried out for the proposed model and variables.
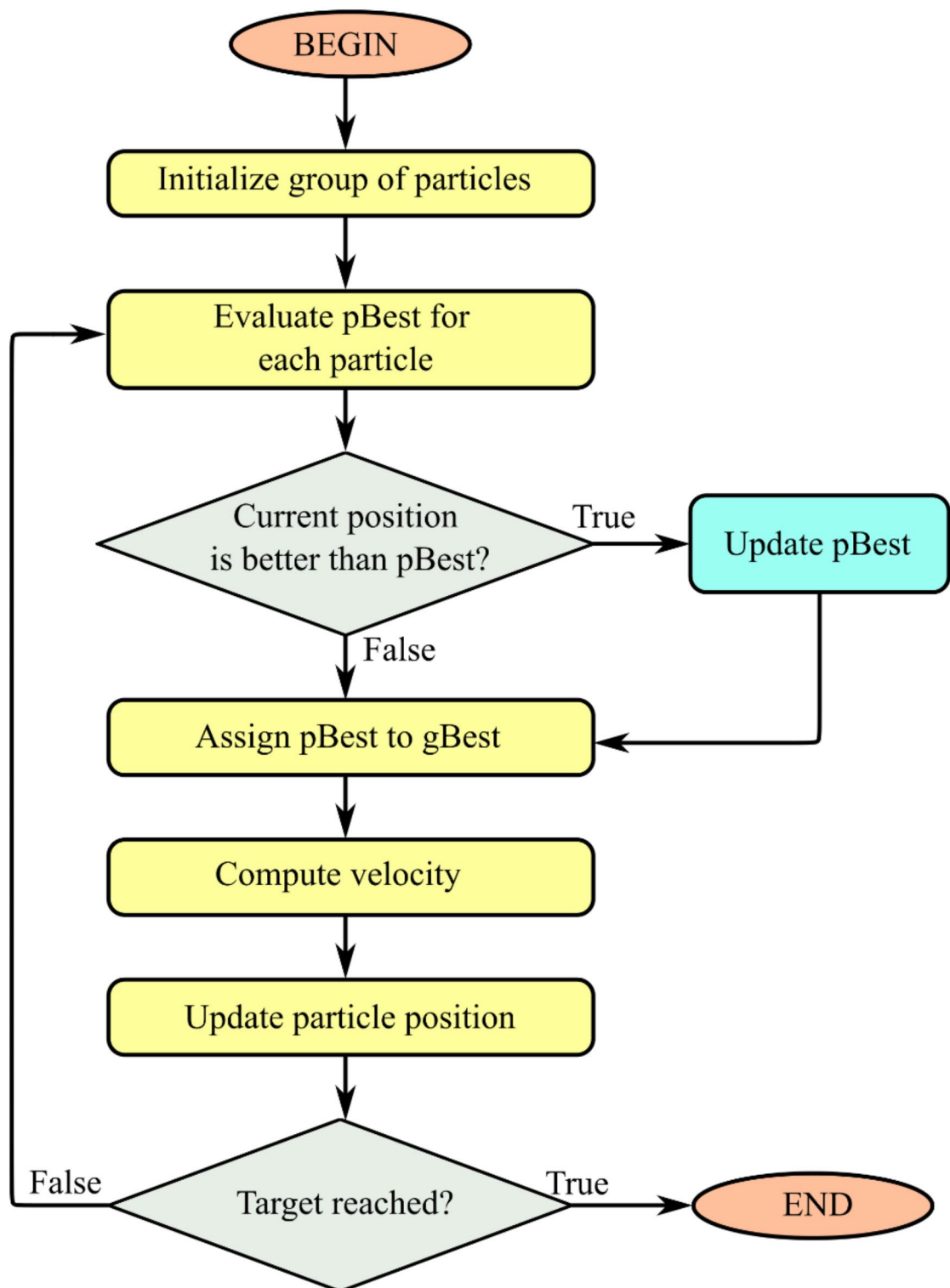
Fig. 3. The procedure of the PSO algorithm.

Performance indicators

In this work, several performance indicators including the Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and coefficient of determination (R^2) have been used to assess the quality of the prediction model, compared to experimental data. Details of these performance indicators and their formulas can be found in[38,39].
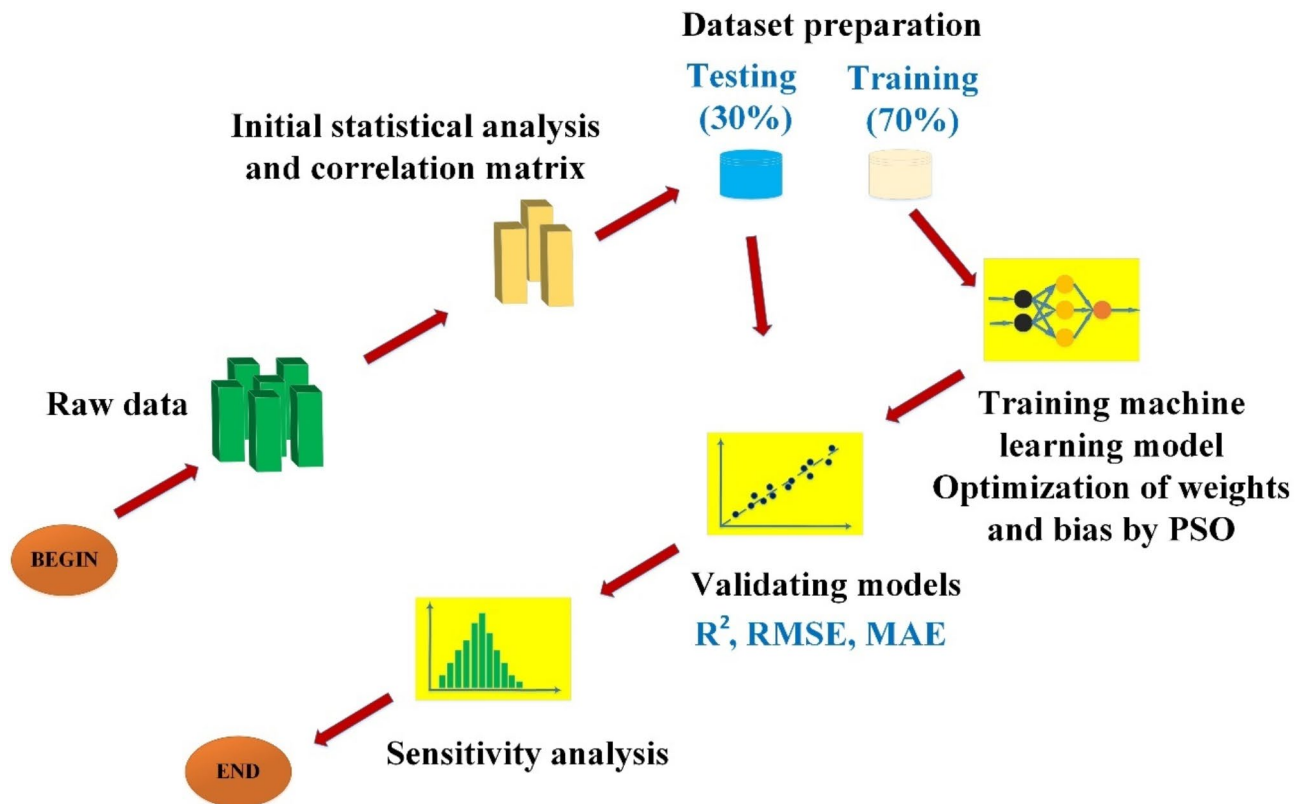
Fig. 4. Flowchart of the current study.

Results and discussion
Optimization of ANN's weight parameters using PSO

In this section, the weight parameters of the proposed ANN model are optimized using the PSO technique. It is worth noting that the prediction results are very much dependent on the architecture of the ANN model[40]. With a fixed quantity of inputs and outputs, the sought parameters of the ANN model are the number of hidden layers and neurons in each of them[41]. As demonstrated by many research works, one hidden layer might be enough to deal with complex non-linear relationships in the prediction problem. For example, in[42], the authors have employed an ANN model with a single hidden layer to predict the axial capacity of square concrete-filled steel tubular columns;[43] used such a model when investigating earthquake slope stability. For these reasons, a single-hidden-layer ANN model has been adopted in this work to reduce computational resources. In terms of number of computational nodes in each hidden layer, it has been chosen to be 2 times higher than the number of inputs, as recommended by several research works[44]. Consequently, the selected architecture of the ANN model in this study was 1 hidden layer containing 20 neurons. In terms of activation functions, a sigmoid activation function has been employed for the hidden layer and a linear activation function has been chosen for the output layer[45]. A mean square error function has been employed to be the cost function. In the current study, Matlab was employed for the training and post-processing of the model, using a Dell computer of 16 Gb of RAM and Intel Core i5, 2.9 GHz.

On the other hand, the key purpose of applying evolutionary algorithms when training AI models is to calibrate the connection between population size and the problem dimension[46]. In many cases of evolutionary algorithms – for instance, Differential Evolution – it is strongly suggested that the population should be 5 to 10 times greater than the number of predictors[47]. However, it is worth noting that a large population size does not always mean a good performance of the model[48]. In this work, the population size for PSO was chosen as 50. Inertia weight was chosen as 0.1, the personal learning coefficient was set as 2, the global learning coefficient 4 and the velocity limit 10%. It should be noted that these types of parameter ranges are often used to train AI models using PSO – for example[49], .

In this work, a maximum number of 100 iterations was employed as the stopping criterion during optimization by PSO. Figure 5 presents the cost functions (RMSE, MAE and R^2, respectively) during training and also testing. The choice of 100 iterations is shown to be relevant to obtain optimized results for all three performance indicators. During the training phase, the proposed model exhibited a good performance in terms of three quality assessments. It should be noted that the testing data were new at the time of application. This remark enables us to make the observation that no overfitting (i.e. performance indicators of testing data lead to an incorrect direction) occurred during the training phase. The robustness and efficiency of PSO during the training phase were then confirmed.
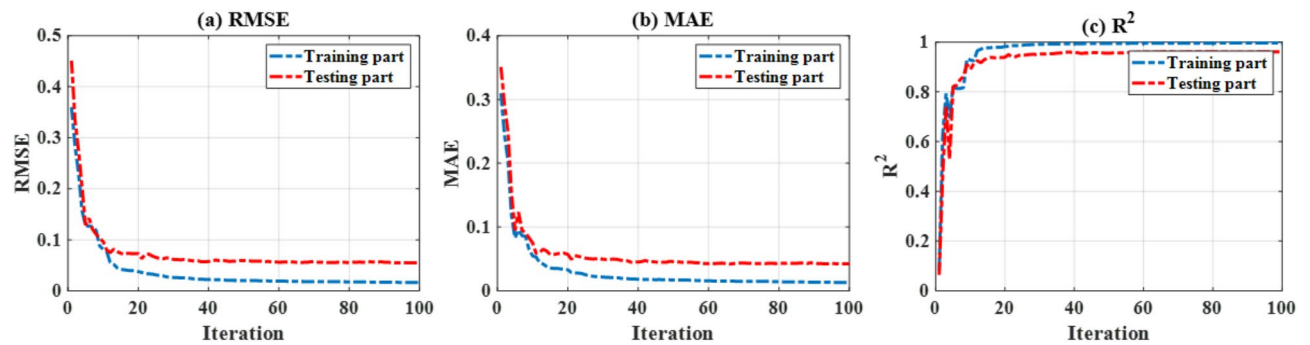
Fig. 5. Cost functions in terms of (**a**) RMSE, (**b**) MAE and (**c**) R^2 during the optimization of ANN's weight parameters by using PSO, for training and testing phases, respectively.
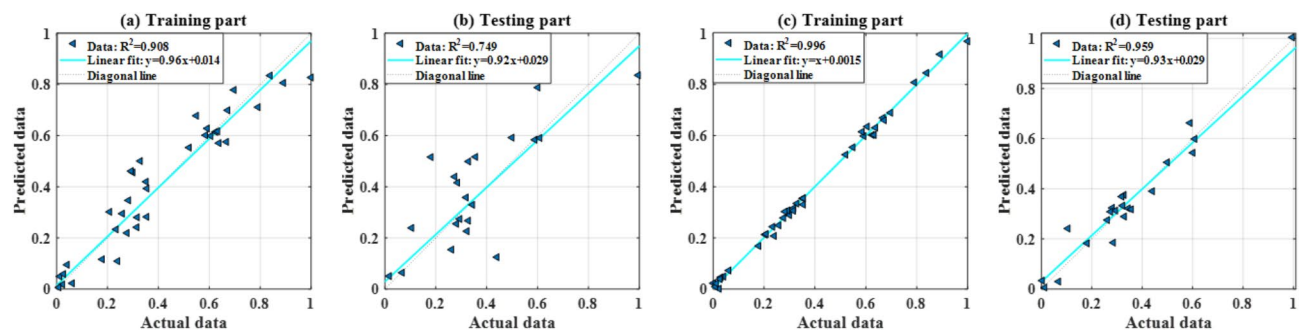

Fig. 6. Prediction capability in terms of regression analysis using training data for (**a**) ANN alone and (**c**) PSOANN, using testing data for (**b**) ANN alone and (**d**) PSOANN.

Dataset	Model	RMSE	MAE	ErrorMean	ErrorStD	R^2	Slope
Training	ANN	0.084	0.064	0.002	0.085	0.908	0.956
	PSOANN	0.016	0.013	−0.001	0.017	0.996	0.998
Testing	ANN	0.141	0.108	−0.003	0.144	0.749	0.922
	PSOANN	0.055	0.042	−0.004	0.056	0.959	0.926

Table 3. The prediction capability of ANN and PSOANN.

Prediction capability

In this section, the prediction capability of two ANN models trained by scaled conjugate gradient and PSO, respectively, is presented for comparison purposes. Figure 6 presents the expected and output data in a regression mode, when training ANN with SCG and PSO, respectively. All quantitative information was summarized in Table 3, including the values of RMSE, MAE, ErrorMean, ErrorStD, R^2, and Slope of training and testing phases. Analysis of the results shows that for the training dataset, ANN trained by PSO yielded better performance in terms of RMSE, MAE, ErrorMean, ErrorStD, R^2 and Slope in comparison to ANN trained by SCG. Similar results were obtained for the testing part: the PSOANN model exhibited the best prediction results for five statistical criteria: RMSE, MAE, R^2, ErrorStD and Slope (RMSE decreases from 0.141 to 0.055; MAE decreases from 0.108 to 0.042; R^2 increases from 0.749 to 0.959; ErrorStD decreases from 0.144 to 0.056; Slope increases from 0.922 to 0.926, using lone ANN and hybrid PSOANN, respectively). Thus, based on both error analysis and assessment of prediction quality, we can say that PSOANN is the most efficient model for predicting the buckling of a steel equal angle.

Sensitivity analysis and discussion

In this section, PD[50] was applied to investigate the marginal effect of input variables on the predicted result of the PSOANN prediction model, as presented in the previous section. As proved in many investigations, PD can determine the nature of the relationship between output and input (linear, monotonic or more complex)[51]. In other words, PD shows how the average prediction changes when the input is changed[52]. In this project, PD code was directly implemented in Matlab, as it is favorable for matrix computation. However, it should be noticed that PD technique has several drawbacks as below. PD assumes that the input variables being analyzed are
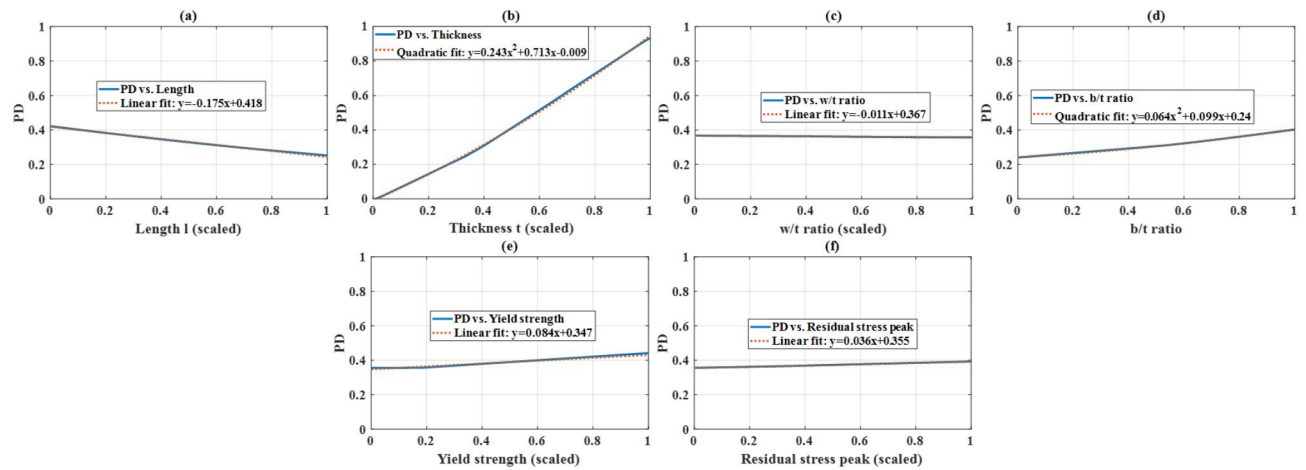
Fig. 7. PD investigation and most appropriate fit of input variable: (**a**) length l, (**b**) thickness t, (**c**) w/t ratio, (**d**) b/t ratio, (**e**) yield strength and (**f**) residual stress peak.
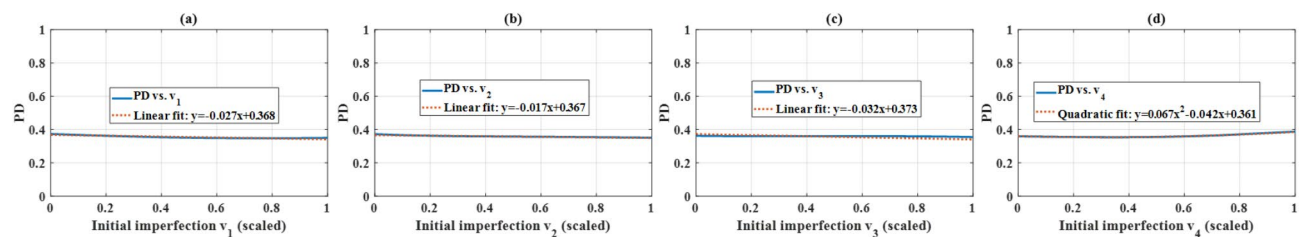

Fig. 8. PD investigation and most appropriate fit of input variable: (**a**) initial imperfection v_1, (**b**) initial imperfection v_2, (**c**) initial imperfection v_3, (**d**) initial imperfection v_4 variable.

independent of others, which may not always be true in real-world datasets where variables are often correlated. This can lead to misleading results, particularly for structural mechanics problems where geometry and material properties are interdependent. Most recently, SHAP (Shapley Additive Explanations) has received attention because this technique provides a complementary approach to PD, addressing many of its limitations[53,54]. For instance, SHAP could highlight how initial imperfections and thickness interact to affect the critical buckling load.

Based on the PSOANN model developed previously (weights and bias), PD allows us to interpret the "black-box" PSOANN model as shown in Figs. 7 and 8, respectively. Figure 7a shows that the relationship between output and length l can be fitted by a linear equation, with a negative slope of −0.175. That means the buckling capacity of the columns decreases with increasing column length. Figure 7b presents the relationship between output and thickness t. However, a nonlinear positive effect is obtained, i.e. the buckling capacity increases with increasing thickness, following a quadratic equation. It is also seen that within the range of thickness values considered in this study, the variation of PD is the highest. For other cases of b/t ratio, yield strength, residual peak stress, and initial imperfection v_4, a positive effect was obtained. However, a negative effect was observed in the rest of the cases – i.e. w/t ratio, initial imperfections v_1, v_2 and v_3. Nonetheless, the variation of PD in the case of w/t ratio and five initial imperfections was not huge compared to the cases of length, thickness and b/t ratio, for instance.

The percentage of sensitivity (i.e. level of influence) of an input variable is calculated as the integral of its respective PD curve. Ten obtained values of the area were scaled into the range of [0, 1], sorted and plotted in Fig. 9b in a bar graph, together with the linear correlation coefficient obtained directly from the database shown in Fig. 9a (also see Table 2). Thickness exhibits the most influence (positive effect) on the buckling capacity of steel equal angles, as identified by both the linear correlation coefficient and PSOANN model (through PD investigation). We can also see that the influence of thickness by far surpasses the contribution of other inputs.

As obtained by PD for the PSOANN model, l, b_t and f_y are the variables that had most influence on the buckling capacity of steel equal angles. However, a maximum of 12% sensitivity was observed for the case of length l. On the other hand, b_t, w_t, f_y, σ_r and v_2 exhibit linear correlation coefficients from 0.17 to 0.25, which are not relevant, especially in view of the number of entries in the database. Nonetheless, PSOANN and linear correlation coefficient are in close agreement in demonstrating the effect of input variables on the buckling capacity of columns.
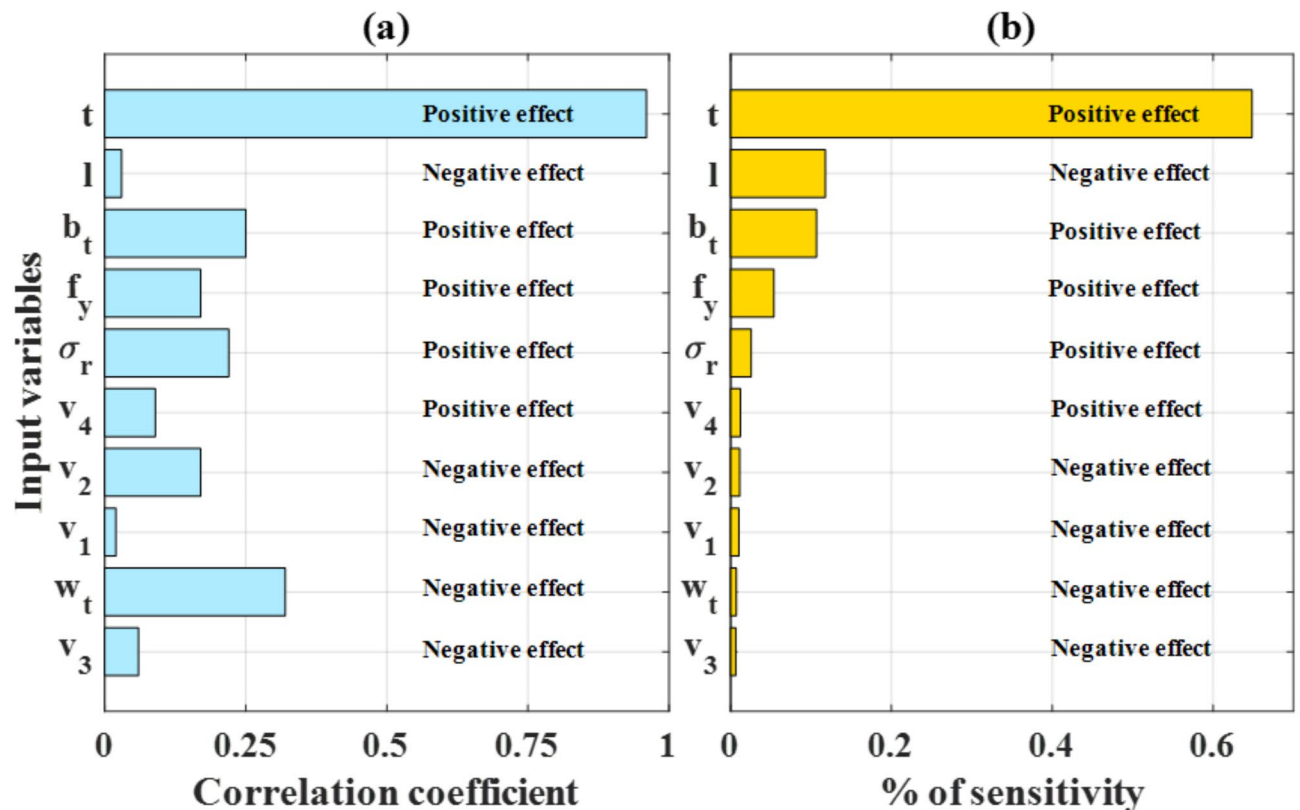
Fig. 9. Percentage of sensitivity on input variable compared to linear correlation coefficient (direct from the database), as well as positive and negative effect of each input.

The integration of the model's predictions into current design codes or standards depends on several factors, including the model's validation, its alignment with regulatory requirements, and the comprehensiveness of the underlying data. It should be noticed that design codes typically require extensive real-world validation across diverse scenarios to ensure reliability and consistency. Besides, the model proposed in the present study was trained on a specific dataset, which may not cover all possible structural configurations, loading conditions, or material types. For integration, these variables must align with the formats and parameters used in design codes (e.g., safety factors, design equations).

In short, the present study introduces a hybrid PSOANN model, where PSO significantly improves the training process by avoiding local minima and enhancing model performance metrics like RMSE and R^2, as shown in Fig. 6; Table 3. The application of PSO provides a robust global optimization approach compared to single ANN model. Moreover, unlike several techniques such as Scaled Conjugate Gradient, PSO searches the solution space globally, ensuring better model generalization. Next, by performing PD analysis, the current work reveals the influence of individual input variables on the buckling load prediction. This step enhances user confidence in deploying machine learning models for structural design and assessment.

Conclusions

In this work, an ANN model whose weights and bias are optimized by PSO is proposed to the critical buckling load of steel equal-angle structural members. Such a model was demonstrated to be efficient when compared to ANN trained by traditional techniques such as SCG. The hybrid PSOANN was a potential surrogate model for the estimation of the buckling capacity of columns, reducing cost and time in laboratory experiments. In addition, PSOANN model allows us to investigate the dependence of output on input variables by using PD. This information can be helpful in structural engineering; for example, some materials whose properties or presence in the structure do not significantly influence the performance of the final design, such as buckling resistance, strength, or durability, can be minimized to reduce production cost. For instance, reducing the need for additional coatings, reinforcements, or filler materials can speed up fabrication and reduce the cost of additional processing or handling.

Despite its achievements, the study has some limitations. First, the dataset used for training may not represent all possible geometric and material configurations. Additionally, the proposed Machine learning model poses challenges for integration into design codes and engineering practices that prioritize simplicity and transparency. Furthermore, the computational intensity of the PSOANN framework may limit its applicability for real-time or large-scale applications without further optimization.

In future works, a larger database should be investigated (to cover a broader range of values). Nonetheless, the feasibility of using surrogate modeling (e.g. ANN-based model) to study nonlinear buckling of columns should also be evaluated, together with other types of damage of structural members. Besides, SHAP analysis should be employed for further investigation of the interaction between variables. Finally, a statistical context should be applied in subsequent studies to exploit the dependence of prediction output on input variables.

Data availability
The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

References
1. Wang, W., Kodur, V., Yang, X. & Li, G. Experimental study on local buckling of axially compressed steel stub columns at elevated temperatures. *Thin-Walled Struct.* **82**, 33–45 (2014).
2. Giamundo, V., Lignola, G. P., Prota, A. & Manfredi, G. Analytical evaluation of FRP wrapping effectiveness in restraining reinforcement bar buckling. *J. Struct. Eng.* **140**, (2014).
3. Megahed, K., Mahmoud, N. S. & Abd-Rabou, S. E. M. Application of machine learning models in the capacity prediction of RCFST columns. *Sci. Rep.* **13**, 20878 (2023).
4. Wu, F., Tang, F., Lu, R. & Cheng, M. Predicting compressive strength of RCFST columns under different loading scenarios using machine learning optimization. *Sci. Rep.* **13**, 16571 (2023).
5. Kollar, L. *Structural Stability in Engineering Practice* CRC. (1999).
6. Beer, F. P. & Johnston, E. R. *Mechanics of Materials* McGraw-Hill. (1992).
7. Kim, D. K. et al. Strength and residual stress evaluation of stub columns fabricated from 800 MPa high-strength steel. *J. Constr. Steel Res.* **102**, 111–120 (2014).
8. Shi, G., Zhou, W. J., Bai, Y. & Lin, C. C. Local buckling of 460 MPa high strength steel welded section stub columns under axial compression. *J. Constr. Steel Res.* **100**, 60–70 (2014).
9. Liu, Y. & Liang, Y. Satin Bowerbird optimizer-neural network for approximating the capacity of CFST columns under compression. *Sci. Rep.* **14**, 8342 (2024).
10. Saetiew, W. & Chucheepsakul, S. Post-buckling of linearly tapered column made of nonlinear elastic materials obeying the generalized Ludwick constitutive law. *Int. J. Mech. Sci.* **65**, 83–96 (2012).
11. Crisfield, M. A. A faster modified newton-raphson iteration. *Comput. Methods Appl. Mech. Eng.* **20**, 267–278 (1979).
12. Chang, T. Y. & Sawamiphakdi, K. Large deflection and post-buckling analysis of shell structures. *Comput. Methods Appl. Mech. Eng.* **32**, 311–326 (1982).
13. Stoykov, S. Buckling analysis of geometrically nonlinear curved beams. *J. Comput. Appl. Math.* **340**, 653–663 (2018).
14. Baydoun, I., Savin, É., Cottereau, R., Clouteau, D. & Guilleminot, J. Kinetic modeling of multiple scattering of elastic waves in heterogeneous anisotropic media. *Wave Motion.* **51**, 1325–1348 (2014).
15. Jiang, D., Bechle, N. J., Landis, C. M. & Kyriakides, S. Buckling and recovery of NiTi tubes under axial compression. *Int. J. Solids Struct.* **80**, 52–63 (2016).
16. Valeš, J. & Kala, Z. Mesh convergence study of solid FE model for buckling analysis. *AIP Conf. Proc.* 150005. (1978).
17. Stoffel, M., Bamer, F. & Markert, B. Artificial neural networks and intelligent finite elements in non-linear structural mechanics. *Thin-Walled Struct.* **131**, 102–106 (2018).
18. Nguyen, P., Doškář, M., Pakravan, A. & Krysl, P. Modification of the quadratic 10-node tetrahedron for thin structures and stiff materials under large-strain hyperelastic deformation. *Int. J. Numer. Methods Eng.* **114**, 619–636 (2018).
19. Peña, J. M., LaTorre, A. & Jérusalem, A. SoftFEM: the Soft Finite element Method. *Int. J. Numer. Methods Eng.* **118**, 606–630 (2019).
20. Karalis, D. G. Increasing the efficiency of computational welding mechanics by combining solid and shell elements. *Mater. Today Commun.* **22**, 100836 (2020).
21. Sebaaly, H., Varma, S. & Maina, J. W. Optimizing asphalt mix design process using artificial neural network and genetic algorithm. *Constr. Build. Mater.* **168**, 660–670 (2018).
22. Ye, F. et al. Artificial neural network based optimization for hydrogen purification performance of pressure swing adsorption. *Int. J. Hydrog. Energy.* **44**, 5334–5344 (2019).
23. Sadowski, Ł., Piechówka-Mielnik, M., Widziszowski, T., Gardynik, A. & Mackiewicz, S. Hybrid ultrasonic-neural prediction of the compressive strength of environmentally friendly concrete screeds with high volume of waste quartz mineral dust. *J. Clean. Prod.* **212**, 727–740 (2019).
24. Asteris, P. G. & Kolovos, K. G. Self-compacting concrete strength prediction using surrogate models. *Neural Comput. Appl.* **31**, 409–424 (2019).
25. Tran, V. L., Thai, D. K. & Kim, S. E. Application of ANN in predicting ACC of SCFST column. *Compos. Struct.* **228**, 111332 (2019).
26. Fahmy, A. S. & El-Madawy, M. E. T. Atef Gobran, Y. using artificial neural networks in the design of orthotropic bridge decks. *Alex. Eng. J.* **55**, 3195–3203 (2016).
27. Le, B. A., Yvonnet, J. & He, Q. C. Computational homogenization of nonlinear elastic materials using neural networks. *Int. J. Numer. Methods Eng.* **104**, 1061–1084 (2015).
28. Kaveh, A., Eskandari, A. & Movasat, M. Buckling resistance prediction of high-strength steel columns using Metaheuristic-trained Artificial neural networks. *Structures* **56**, 104853 (2023).
29. Mojtabaei, S. M., Becque, J., Hajirasouliha, I. & Khandan, R. Predicting the buckling behaviour of thin-walled structural elements using machine learning methods. *Thin-Walled Struct.* **184**, 110518 (2023).
30. Hou, Z., Hu, S. & Wang, W. Interpretable machine learning models for predicting probabilistic axial buckling strength of steel circular hollow section members considering discreteness of geometries and material. *Adv. Struct. Eng.* https://doi.org/10.1177/13694332241289175 (2024).
31. Ban, H. Y., Shi, G., Shi, Y. J. & Wang, Y. Q. Column buckling tests of 420 mpa high strength steel single equal angles. *Int. J. Str. Stab. Dyn.* **13**, 1250069 (2013).
32. Ma, T. Y., Liu, X., Hu, Y. F., Chung, K. F. & Li, G. Q. Structural behaviour of slender columns of high strength S690 steel welded H-sections under compression. *Eng. Struct.* **157**, 75–85 (2018).
33. Huiyong, B. & Gang, S. Wang Yuanqing. Residual stress tests of high-strength steel equal angles. *J. Struct. Eng.* **138**, 1446–1454 (2012).
34. Narayanakumar, S. & Raja, K. A BP Artificial neural network model for Earthquake Magnitude Prediction in Himalayas, India. *Circuits Syst.* **7**, 3456–3468 (2016).
35. Zweiri, Y. H., Whidborne, J. F. & Sceviratne, L. D. A three-term Backpropagation Algorithm. *Neurocomputing* **50**, 305–318 (2002).
36. Schmidhuber, J. Deep learning in neural networks: an overview. *Neural Netw.* **61**, 85–117 (2015).

37. Shi, Y. & Eberhart, R. A modified particle swarm optimizer. in *IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence (Cat. No.98TH8360)* https://doi.org/10.1109/ICEC.1998.699146 69–73. (1998).
38. Chai, T. & Draxler, R. R. Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature. *Geosci. Model Dev.* **7**, 1247–1250 (2014).
39. Willmott, C. J. & Matsuura, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* **30**, 79–82 (2005).
40. Sarir, P., Chen, J., Asteris, P. G., Armaghani, D. J. & Tahir, M. M. Developing GEP tree-based, neuro-swarm, and whale optimization models for evaluation of bearing capacity of concrete-filled steel tube columns. *Eng. Comput.* https://doi.org/10.1007/s00366-019-00808-y (2019).
41. Sheela, K. G. & Deepa, S. N. Review on methods to fix number of hidden neurons in neural networks. *Math. Probl. Eng.* https://doi.org/10.1155/2013/425740 https://www.hindawi.com/journals/mpe/2013/425740/ (2013).
42. Ren, Q., Li, M., Zhang, M., Shen, Y. & Si, W. Prediction of Ultimate Axial Capacity of Square Concrete-Filled Steel Tubular Short Columns using a Hybrid Intelligent Algorithm. *Appl. Sci.* **9**, 2802 (2019).
43. Gordan, B., Jahed Armaghani, D., Hajihassani, M. & Monjezi, M. Prediction of seismic slope stability through combination of particle swarm optimization and neural network. *Eng. Comput.* **32**, 85–97 (2016).
44. Kuri-Morales, A. Closed determination of the number of neurons in the hidden layer of a multi-layered perceptron network. *Soft Comput.* **21**, 597–609 (2017).
45. Abambres, M., Rajana, K., Tsavdaridis, K. D. & Ribeiro, T. P. Neural network-based formula for the buckling load prediction of I-Section Cellular Steel beams. *Computers* **8**, 2 (2019).
46. Javed, H. et al. On the efficacy of ensemble of Constraint Handling techniques in Self-Adaptive Differential Evolution. *Mathematics* **7**, 635 (2019).
47. Piotrowski, A. P. Review of Differential Evolution population size. *Swarm Evol. Comput.* **32**, 1–24 (2017).
48. Chen, T., Tang, K., Chen, G. & Yao, X. A Large Population Size Can Be Unhelpful in Evolutionary Algorithms. *arXiv* (2012).
49. Momeni, E., Jahed Armaghani, D., Hajihassani, M. & Mohd Amin, M. F. Prediction of uniaxial compressive strength of rock samples using hybrid particle swarm optimization-based artificial neural networks. *Measurement* **60**, 50–63 (2015).
50. Zhao, Q. & Hastie, T. Causal interpretations of Black-Box models. *J. Bus. Econ. Stat.*, 1–10 (2019).
51. Qi, C. et al. Towards Intelligent Mining for Backfill: a genetic programming-based method for strength forecasting of cemented paste backfill. *Miner. Eng.* **133**, 69–79 (2019).
52. Goldstein, A., Kapelner, A., Bleich, J. & Pitkin, E. Peeking inside the black box: visualizing statistical learning with plots of individual conditional expectation. *J. Comput. Graph. Stat.* **24**, 44–65 (2015).
53. Nguyen, V. H. et al. Optimization of milling conditions for AISI 4140 steel using an integrated machine learning-multi objective optimization-multi criteria decision making framework. *Measurement* **242**, 115837 (2025).
54. Sun, Y., Zhang, L., Yao, M. & Zhang, J. Neural network models and shapley additive explanations for a beam-ring structure. *Chaos Solitons Fractals*. **185**, 115114 (2024).

Author contributions

Nang Xuan Ho, Tien-Thinh Le are responsible for analysis and preparing the first draft of the manuscript. The-Hung Dinh and Van-Hai Nguyen are responsible for material preparation and data collection. All authors contributed to the study conception and design, reviewed all previous versions of the manuscript, and read and approved the final manuscript.

Funding

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to T.-T.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.