Proceedings of the
Annual Stability Conference
Structural Stability Research Council
San Antonio, Texas, March 19-22, 2024

Machine Learning for predicting the cross-sectional elastic buckling strengths of cold-formed steel members

Shaokuai Wang[1], Justin Mo[2], Lingfeng Yin[3]. Zhanjie Li[4]

Abstract

The elastic buckling strength is a key parameter in predicting the design strength in cold-formed steel member design. Particularly, for cold-formed steel sections, there are several buckling modes that are typically categorized: local-plate, distortional, and/or global buckling. As the necessary first step in design, finding the elastic buckling solution generally necessitates numerical analysis for elastic buckling prediction due to the complexity nature of cross-sectional instabilities though some simplified analytical solutions are available. Even the most commonly used tool such as finite strip method (FSM) that has the least modeling effort has its own challenges and complexities in buckling prediction. The paper aims to explore Machine Learning (ML) for predicting the elastic buckling strength of cold-formed steel members - shifting away from the mechanic-based numerical analysis. The Artificial Neural Network (ANN), a subgroup of machine learning algorithms, will be employed for the model training. The ML model will start with the commonly used sections, lipped channel section (C), where the dataset includes those from standard industrial shapes and parametrically randomized shapes. Then, those data will be categorized into geometrical data and material data as inputs and elastic buckling strengths and half-wavelengths of local and distortional buckling as the outputs. The developed ML models' efficiency and accuracy will be evaluated. Further extension of the ML models to more generalized sections will be further explored.

1. Introduction

Cold-formed steel (CFS) members, characterized by their open cross-sections, present complex buckling behaviors with commonly knowns modes such as global, local, and distortional buckling. The Direct Strength Method (DSM) (Schafer, 2008), increasingly favored over the Effective Width Method (EWM) (Von Kármán et al., 1932; Winter, 1947), offers a direct approach to assessing load-bearing capacity with consideration of these buckling modes. DSM utilizes the full section properties, reduced material strength rather than calculating effective

---

[1] Graduate Student, School of Civil Engineering, Southeast University, Nanjing 211189, China, wangshaokuai@foxmail.com

[2] Senior HS student, Bethlehem Central High School, Delmar, New York, jmo230615@gmail.com

[3] Professor, School of Civil Engineering, Southeast University, Nanjing 211189, China, eking@seu.edu.cn

[4] Professor, Department of Engineering, SUNY Polytechnic Institute, Utica, NY 13502, Zhanjie.Li@sunypoly.edu

section characteristics. This method relies on the premise that a member's ultimate strength can be determined through its elastic buckling and yield stresses based on the observed experimental correlation.

The critical elastic buckling stresses of CFS member can be determined through numerical or analytical solutions. The semi-analytical Finite Strip Method (FSM), a numerical approach, discretizes a CFS cross-section in strips. This method, particularly effective for prismatic sections with simply supported end conditions, yields the member's signature curve, elucidating local, distortional, and global elastic buckling loads or moments as required in the specification. Each buckling mode correlates with a buckling half-wavelength, providing a comprehensive mode description. Pioneered by Hancock and his team, the Finite Strip Method's use in stability analysis of CFS members was further elaborated by (Cheung & Tham, 1998; Hancock et al., 2001; Schafer & Adany, 2006). This method's significant potential in CFS design and behavior analysis is underlined by its successful application in various studies. In particularly, CUFSM (Li & Schafer, 2010), an open-source software tool implementing the seme-analytical FSM, has been widely used in analysis and design of CFS members. Meanwhile, shell Finite element models (Cook, 2007; Zienkiewicz & Taylor, 2005), a more powerful method for elastic buckling of CFS members, could tackle more complex geometry, boundary, and loading conditions. However, a significant amount of visual judgment is often required to categorize these buckling solutions into local, distortional, or global buckling for use in specifications. The instability modes in finite element models typically exhibit coupling, adding to the complexity of mode identification. To enable a formal definition of buckling modes, there are two methods recently developed in the past two decades: Generalized Beam Theory (GBT) (Schardt, 1989) and constrained Finite Strip Method (cFSM) (Schafer & Adany, 2006) including a variety of extension or enhancement (Elzein, 2013; Harik et al., 1991). Formal mode decomposition and identification can be performed through these methods, which enables a deeper digest of the buckling behaviors including the potential interactive behaviors. On the other hand, analytical solutions for the elastic buckling of cold-formed steel members based on simplification of the classic theories are often available in the design specifications, such as those for local and distortional buckling in AISI (AISI, 2020). Improving the accuracy of the analytical equations are also consistently pursued, such as the local buckling for lipped channels in (Ding & Schafer, 2023). In addition, for commonly used sections, these elastic buckling solutions are conveniently tabulated in resources such as the AISI Cold-Formed Steel Design Manual (2017) (Manual, 2017) and the CFSEI Tech Note G103-11 (Li & Schafer, 2011), providing accessible references for engineers.

While numerical tools and analytical equations based on the mechanics offer foundational insights into elastic buckling, the integration of machine learning applications in CFS and broader structural contexts marks a significant leap forward, bridging traditional methods with advanced predictive capabilities (Pitton et al., 2019; Tohidi & Sharifi, 2015; Wang et al., 2023; Wu et al., 2019; Zarringol et al., 2020). As one of the machine learning methods, Artificial Neural Networks (ANNs), simulating the learning mechanisms of biological organisms, have been increasingly applied in structural engineering for buckling behavior prediction. Resembling the human brain structure, ANNs consist of numerous processing elements and excel at establishing accurate relationships between inputs and outputs in datasets lacking specific solutions. Pioneering this application in the field were researchers like Kaveh (Iranmanesh &

Kaveh, 1999; Kaveh et al., 2018; Kaveh et al., 2001; Kaveh & Iranmanesh, 1998), who conducted extensive studies on the predictive capabilities of neural networks in structural domains. Successful implementations include Waszczyszyn and Bartczak's (Waszczyszyn & Bartczak, 2002) use of Backpropagation Neural Networks (BPNN) for predicting buckling loads in cylindrical shells under axial compression and Markopoulos et al (Markopoulos et al., 2007) .'s application of ANNs with Levenberg-Marquardt and backpropagation algorithms to predict failure modes in PVC cylindrical shells. The application of ANNs in cold-formed steel (CFS) is also expanding. Pala (Pala, 2006) utilized ANNs to estimate elastic torsional buckling stresses in CFS C-sections under pure compression and bending, leading to the development of ANN-based predictive equations. Pala and Caglar (Pala & Caglar, 2007) further studied the impact of geometrical parameters on torsional buckling stresses using these equations. In another study, Guzelbey et al. (Guzelbey et al., 2006) trained an ANN with experimental data to predict failures in cold-formed steel webs, demonstrating greater accuracy compared to current design codes. Degtyarev (Degtyarev, 2021) trained an ANN model to predict elastic shear buckling loads and ultimate shear strengths in CFS C-sections, proving its superiority over code-prescribed design equations. Furthering this research, Degtyarev and Naser (Degtyarev & Naser, 2021) compared five different machine learning enhancement algorithms for predicting elastic shear buckling loads and shear strengths in cold-formed thin-walled steel components, with the CatBoost algorithm emerging as the most accurate.

Hence, the aim of this study is to determine the cross-sectional buckling strength of cold-formed steel using a machine learning model. This approach seeks to leverage the advanced capabilities of machine learning to enhance the accuracy and efficiency of predicting buckling strength in these CFS members. In this study, an Artificial Neural Network (ANN) algorithm is employed to predict the elastic critical buckling load of cold-formed lipped channel steel members. Utilizing a dataset based on 260,000 sectional parameters and elastic buckling calculations from CUFSM, the data is divided into training, testing, and validation sets for neural network model training. The model applies early stopping (Morgan & Bourlard, 1989) to control overfitting and uses random grid search (Bergstra & Bengio, 2012) for hyperparameter selection. Various ANN models with different inputs, outputs, and hidden layer counts were constructed. After finalizing the hyperparameters, each model type was trained ten times to identify the most effective model, resulting in a total of 180 trained models.

2. Elastic Buckling prediction methodology and machine learning data development

2.1 Two-step approaching using finite strip method and constrained finite strip method

Local and distortional buckling loads of lipped channel sections are determined using the semi-analytical finite strip method under simply supported end boundary conditions as implemented in CUFSM software (Li & Schafer, 2010). Both buckling loads are determined from the signature curve. Typically, the signature curve has two distinct minima, unique minima, corresponding to local and distortional buckling. However, the signature curve is not always as unambiguous with unique minima. It could have either or both minima "indistinct", characterized as non-unique minima. For cross-sections with non-unique minima in the signature curve, the two-step procedure for identifying unique minimum is employed (Li & Schafer, 2010) involving the cFSM solutions.

2.2 Machine learning data development

In accordance with the product specification range for cold-formed steel (CFS) members as outlined by the SFIA (shown in Figure 1(c)), 260,000 different CFS lipped channel sections were randomly generated. The dimensions of these sections, including the out-out member depth (H), flange width (B), lip length (D), and material thickness (t), are depicted in Figure 1 (a), while the centerline dimensions are presented in Figure 1 (b). All the Finite Strip Method (FSM) models are designed using the centerline dimensions.

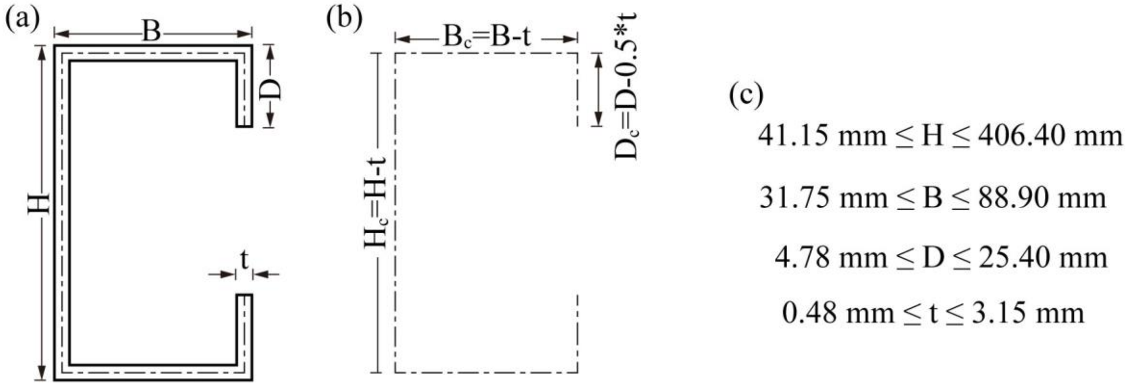


Figure 1 (a) CFS lipped channel section (b) Centerline dimensions of lipped channel sections (c) dimension limits

Boxplots (Figure 2) and histograms with kernel density estimation (Figure 3) have been constructed. The boxplot provides a visual summary of data distribution, showing the median, quartiles, and outliers. The boxplot illustrates the distribution of four variables: H_c, B_c, Dc, and t. The median of each variable is marked by the red line within the box, indicating the central tendency. The blue boxes extend from the first quartile to the third quartile, representing the interquartile range which encapsulates the middle 50% of the data. The 'whiskers' extend to the minimum and maximum values within a 1.5 interquartile range from the lower and upper quartiles, respectively. There are not points outside this range, which would be considered outliers. These visual metrics reflect the variability and symmetry of the data distribution.

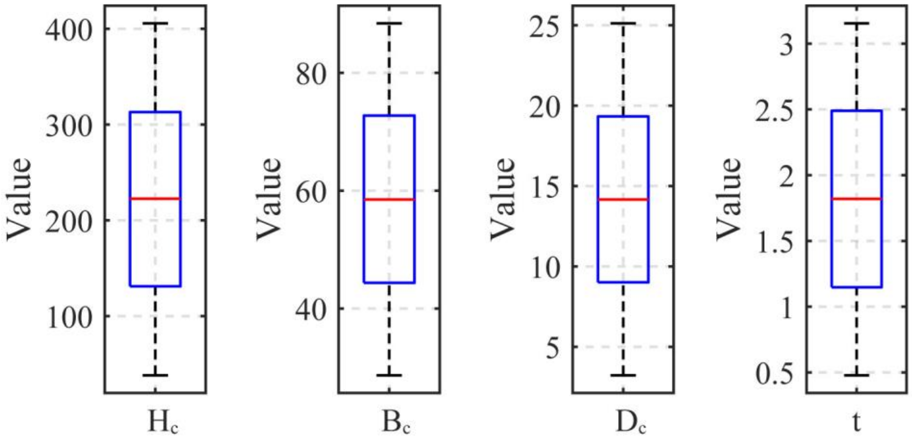


Figure 2 Box plot for dimensions of lipped channel sections

The histogram, enhanced with kernel density estimation, offers a smooth curve representing the data's density or frequency distribution. Normalizing the histogram ensures that its area sums to one, allowing for the comparison of distributions across different scales or datasets. This normalization is crucial for accurately interpreting the data's underlying distribution, regardless of sample size or unit differences. Figure 3 depicts histograms with overlaid kernel density estimates for four variables. Each histogram is color-coded to visually represent frequency distribution. The kernel density estimation, outlined in red, provides a smoothed curve to highlight the central tendency and distribution spread.

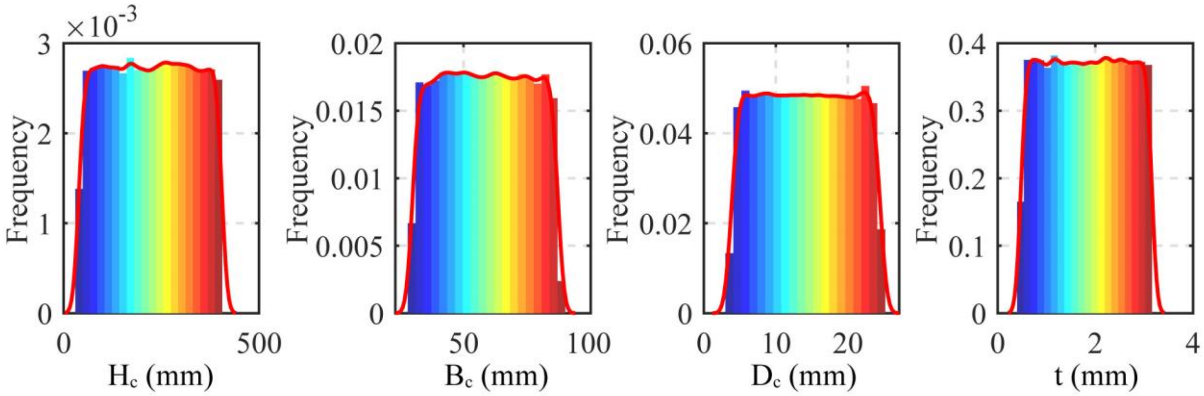


Figure 3 Histograms of generated sections

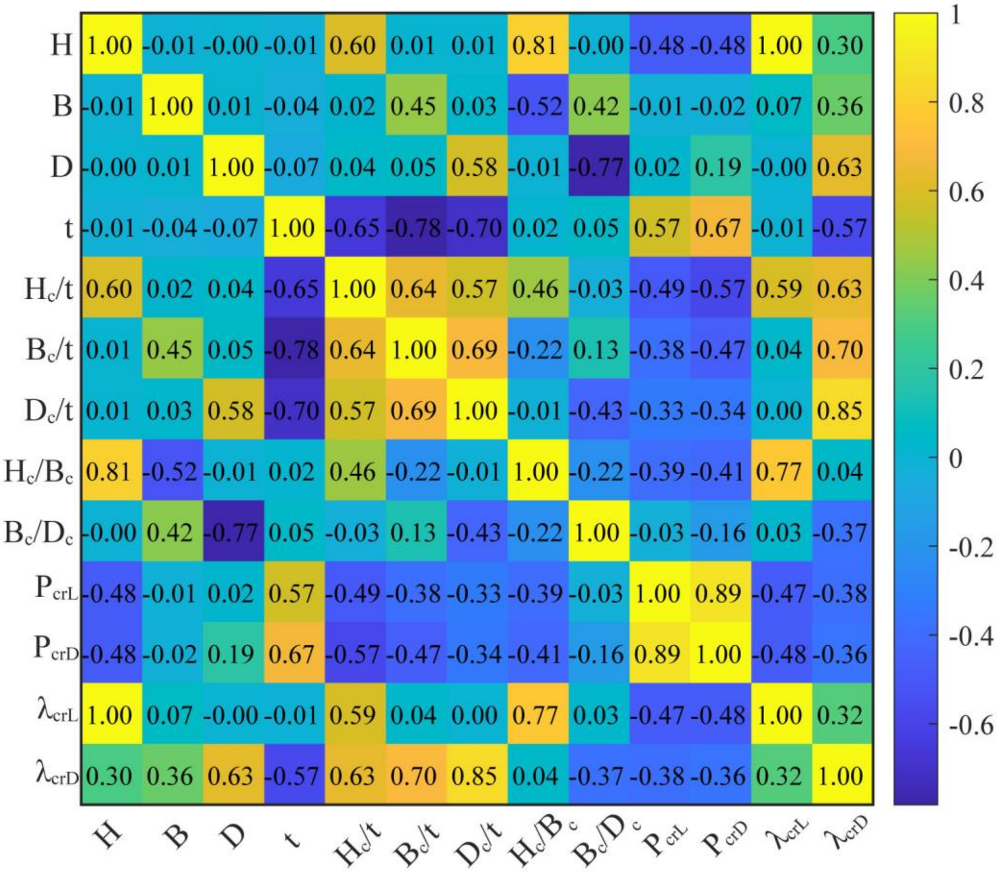
As observed from Figure 2 and Figure 3, the 260,000 randomly generated samples are uniformly distributed across the entire dimensional space. This uniform distribution is essential for ensuring a comprehensive analysis across the spectrum of possible geometrical configurations of the lipped channel cold-formed steel sections.

All FSM models are computed within CUFSM to determine the critical forces for elastic local buckling (P_{crL}) and elastic distortional buckling (P_{crD}), along with their corresponding half-wavelengths (λ_{crL} and λ_{crD}). In addition, dimensionless parameters such as H_c/t, b_c/t, d_c/t, h_c/b_c, and b_c/d_c are calculated to provide further insight into the geometric properties of the cold-formed steel sections relative to P_{crL}, P_{crD}, λ_{crL} and λ_{crD}.

The correlation coefficient is a statistical measure that calculates the strength of the relationship between two variables. The correlation coefficient, often denoted as *r*, is a numerical measure that expresses the degree of linear relationship between two variables. It is commonly defined by the Pearson correlation coefficient Eq. (1):

$$r=\frac{\sum\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)}{\sqrt{\sum\left(x_i-\bar{x}\right)^2\sum\left(y_i-\bar{y}\right)^2}} \tag{1}$$

Where x_i and y_i are the values of the two variables, and $\bar{x}$, $\bar{y}$ are the means of those variables, respectively. The value of *r* ranges from -1 to 1, with -1 indicating a perfect negative linear relationship, 0 indicating no linear relationship, and 1 indicating a perfect positive linear relationship. The heatmap (Figure 4) is a visual representation of the correlation matrix, illustrating the Pearson correlation coefficients between various parameters of the section. This heatmap is essential for identifying relationships between dimensions and buckling parameters, guiding the feature selection for predictive models.



	H	B	D	t	H_c/t	B_c/t	D_c/t	H_c/B_c	B_c/D_c	P_{crL}	P_{crD}	λ_{crL}	λ_{crD}
H	1.00	-0.01	-0.00	-0.01	0.60	0.01	0.01	0.81	-0.00	-0.48	-0.48	1.00	0.30
B	-0.01	1.00	0.01	-0.04	0.02	0.45	0.03	-0.52	0.42	-0.01	-0.02	0.07	0.36
D	-0.00	0.01	1.00	-0.07	0.04	0.05	0.58	-0.01	-0.77	0.02	0.19	-0.00	0.63
t	-0.01	-0.04	-0.07	1.00	-0.65	-0.78	-0.70	0.02	0.05	0.57	0.67	-0.01	-0.57
H_c/t	0.60	0.02	0.04	-0.65	1.00	0.64	0.57	0.46	-0.03	-0.49	-0.57	0.59	0.63
B_c/t	0.01	0.45	0.05	-0.78	0.64	1.00	0.69	-0.22	0.13	-0.38	-0.47	0.04	0.70
D_c/t	0.01	0.03	0.58	-0.70	0.57	0.69	1.00	-0.01	-0.43	-0.33	-0.34	0.00	0.85
H_c/B_c	0.81	-0.52	-0.01	0.02	0.46	-0.22	-0.01	1.00	-0.22	-0.39	-0.41	0.77	0.04
B_c/D_c	-0.00	0.42	-0.77	0.05	-0.03	0.13	-0.43	-0.22	1.00	-0.03	-0.16	0.03	-0.37
P_{crL}	-0.48	-0.01	0.02	0.57	-0.49	-0.38	-0.33	-0.39	-0.03	1.00	0.89	-0.47	-0.38
P_{crD}	-0.48	-0.02	0.19	0.67	-0.57	-0.47	-0.34	-0.41	-0.16	0.89	1.00	-0.48	-0.36
λ_{crL}	1.00	0.07	-0.00	-0.01	0.59	0.04	0.00	0.77	0.03	-0.47	-0.48	1.00	0.32
λ_{crD}	0.30	0.36	0.63	-0.57	0.63	0.70	0.85	0.04	-0.37	-0.38	-0.36	0.32	1.00

Figure 4 Correlation matrix for the 9 inputs and 4 outputs parameters

Figure 5 displays a series of scatter plots, each depicting the density of data points correlating parameters related to the structural characteristics of cold-formed steel members with their elastic buckling properties. The color gradient represents the concentration of data points, with warmer colors indicating higher density regions. Normalizing the density in the scatter plot matrix helps understand the distribution of data points within the parameter space, regardless of the absolute number of observations, thus demonstrates a holistic view of the data's density.

Based on the analysis of the scatter plots in Figure 5, we observe that the critical local buckling load (P_{crL}) and the critical distortional buckling load (P_{crD}) exhibit a negative correlation with the geometric ratios of height-to-thickness (H_c/t), width-to-thickness (B_c/t), depth-to-thickness (D_c/t), and height-to-width (H_c/B_c). In contrast, the half-wavelength for local buckling (λ_{crL}) shows a positive correlation with the height-to-thickness (H_c/t) and height-to-width (H_c/B_c) ratios. Similarly, the half-wavelength for distortional buckling (λ_{crD}) is positively correlated with the height-to-thickness ($H_{c/t}$), width-to-thickness ($B_{c/t}$), and depth-to-thickness ($D_{c/t}$) ratios.

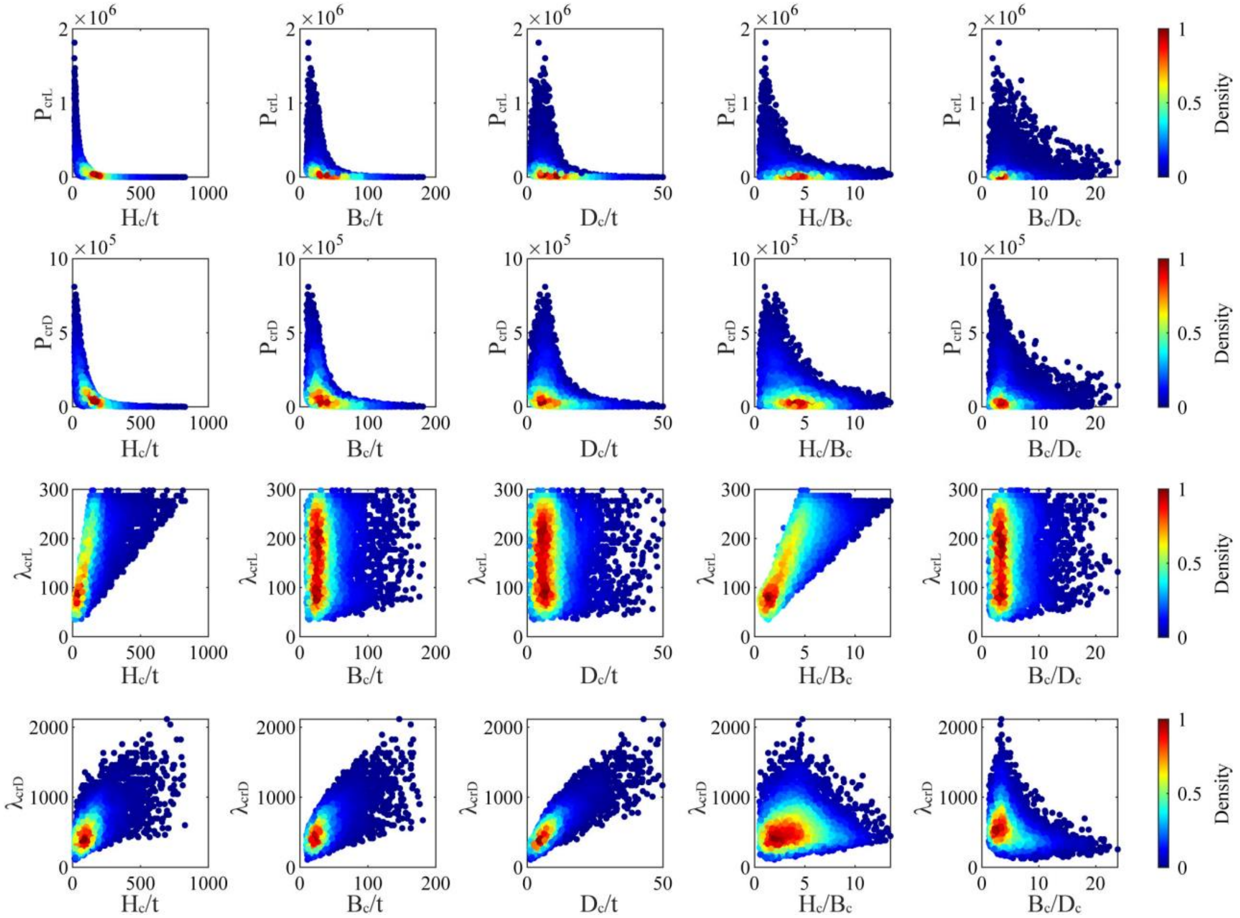


Figure 5 Scatter plot matrix

3. Machine learning framework

Neural networks mimic the functioning of the biological nervous system, with elements operating in parallel. These elements' interconnections are crucial for the network's function. Training involves adjusting the weights of these connections to achieve desired outputs from given inputs, as shown in Figure 6. Through iterative training with many input and target pairs, the network learns to approximate the target output closely. This learning process is akin to finding the optimal pathways in a complex landscape based on feedback.

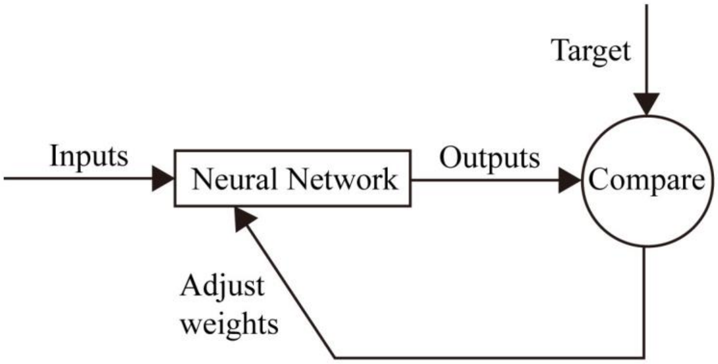


Figure 6 Basic methodology used in neural network training

The flowchart in Figure 7 delineates a structured approach to developing an Artificial Neural Network (ANN) model for computational analysis. It initiates with ‘Data Preparation,’ where data is generated using CUFSM (i.e., computational analysis) and then partitioned into distinct training, validation, and testing subsets. Subsequently, the ‘Determining ANN Architecture’ phase involves specifying the input and output nodes, the number of nodes in hidden layers, and the learning rate. ‘Training ANN’ encompasses selecting activation and loss functions, followed by the actual construction and training of the ANN model. The final phase, ‘Evaluation of ANN,’ entails the assessment of the ANN model, extraction of parameters, and the preservation of the trained model for future application. This diagram offers a comprehensive roadmap for the development and optimization of ANN models within a scientific research framework.

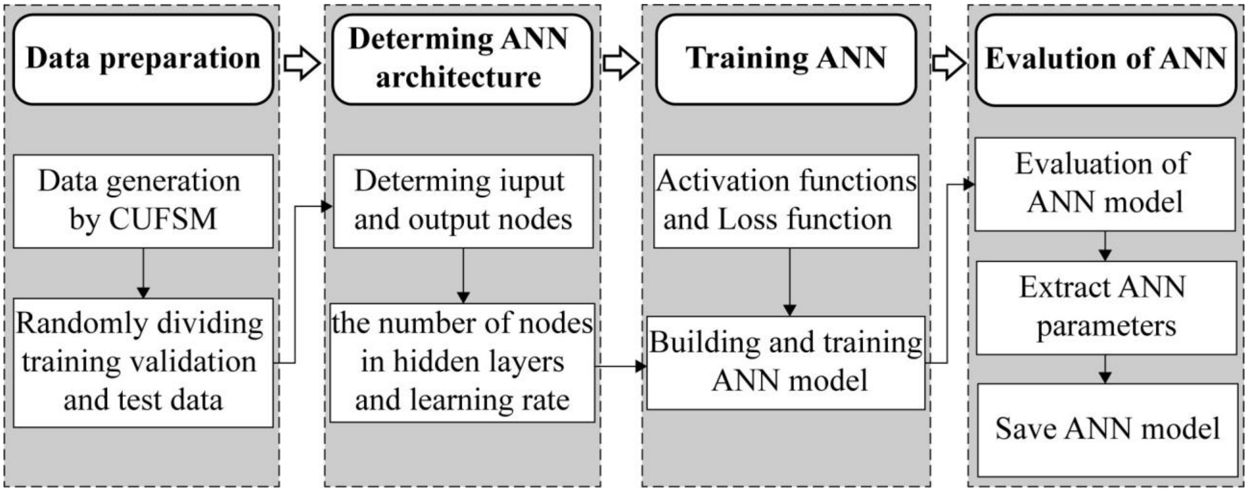


Figure 7 ANN construction, training, and evaluation process

The schematic chart in Figure 8 illustrates a multilayer perceptron neural network architecture, detailing the flow from input to output. In the input layer, predictors are denoted as p_i, which connect through weighted paths (weights $w^{l}_{j,k}$) to the hidden layer. Each neuron in the hidden layer sums its weighted inputs and bias (b^{1}_{j}) before applying an activation function (Figure 9 (a)) f^{1} to produce an output a^{1}_{j}. The process repeats from the hidden to the output layer, with outputs a^{2}_{j} generated after applying a second activation function (Figure 9 (b)) f^{2}.

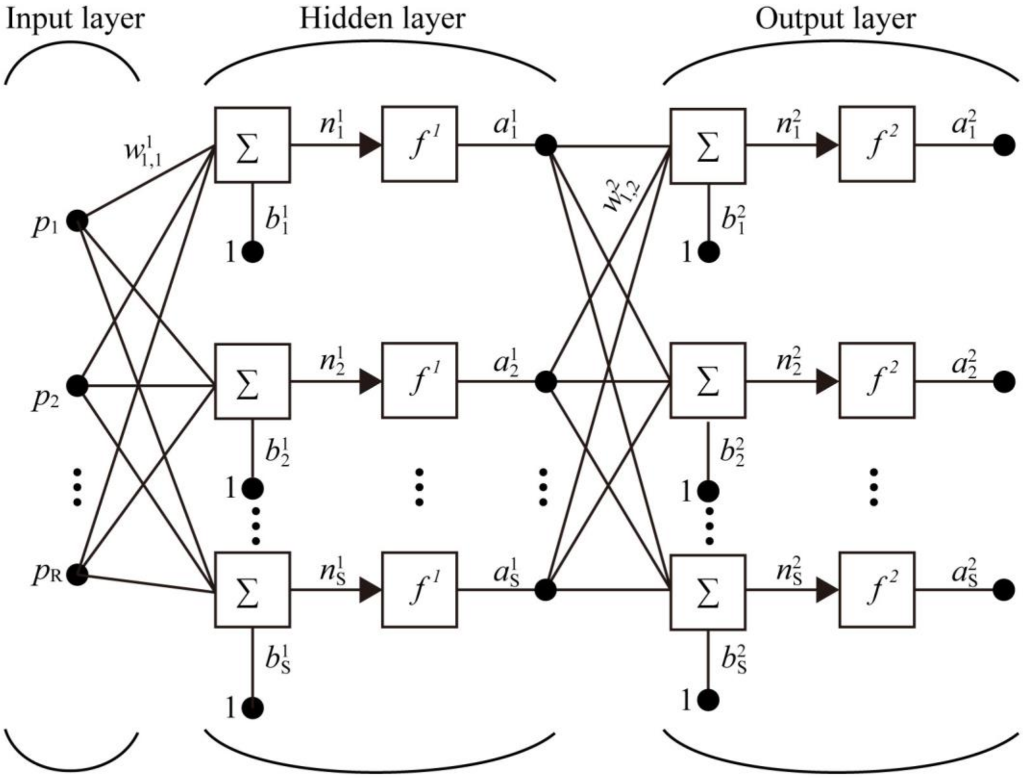


Figure 8 Architecture of ANN

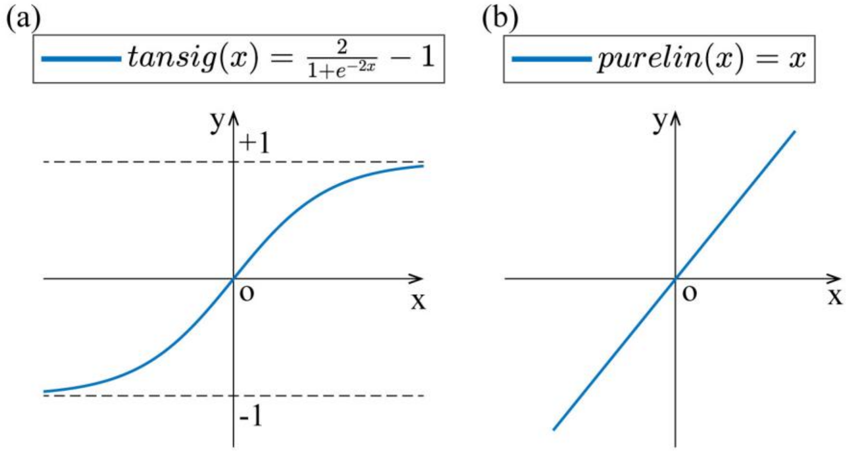


Figure 9 Activation functions (a) tansig; (b) purelin

The ANN Forward propagation process can mathematically be expressed as follows:

$$a_j^l = f\left(\sum_k w_{jk}^l a_k^{l-1} + b_j^l\right) \tag{2}$$

Where $w_{j,k}^l$ is the weight for the connection from k^{th} neuron in the $(l-1)^{th}$ layer to the j^{th} neuron in l^{th} layer. b_j^l is the bias of the j^{th} neuron in the l^{th} layer. a_j^l is the activation of the j^{th} neuron in l^{th} layer.

Backpropagation in neural networks is a method used for refining the model’s weights and biases in response to the error observed between the predicted and targeted outputs. It involves calculating the gradient of the loss function with respect to each weight by the chain rule, propagating the error backward from the output layer to the input layer. This gradient informs how the weights should be adjusted to minimize the loss. The process iteratively updates the weights and biases to reduce prediction error, thereby improving the model's performance over successive training epochs. The cost function is manifested in various incarnations, notably Mean Absolute Error (MAE), Mean Squared Error (MSE), among others. The study employs Levenberg-Marquardt algorithm for optimization, numerically deriving the requisite gradients for this minimization.

The Levenberg-Marquardt algorithm was designed to approach second-order training speed without having to compute the Hessian matrix. When the performance function has the form of a sum of squares, then the Hessian matrix can be approximated as:

$$H = J^T J \tag{3}$$

and the gradient can be computed as:

$$g = J^T e \tag{4}$$

where J is the Jacobian matrix that contains first derivatives of the network errors with respect to the weights and biases, and e is a vector of network errors. The Levenberg-Marquardt algorithm uses this approximation to the Hessian matrix in the following Newton-like update:

$$\begin{aligned} w_{k+1} &= w_k - \left[J^T J + \mu \mathrm{I} \right]^{-1} J^T e \\ b_{k+1} &= b_k - \left[J^T J + \mu \mathrm{I} \right]^{-1} J^T e \end{aligned} \tag{5}$$

When the scalar μ is zero, this is just Newton's method, using the approximate Hessian matrix. When μ is large, this becomes gradient descent with a small step size. Newton's method is faster and more accurate near an error minimum, so the aim is to shift toward Newton's method as quickly as possible. Thus, μ is decreased after each successful step (reduction in performance function) and is increased only when a tentative step would increase the performance function. In this way, the performance function is always reduced at each iteration of the algorithm.

Figure 10 represents a typical data partitioning strategy for machine learning workflows, where the dataset is divided into subsets for training, validation, and testing. It shows that 80% of the total dataset is designated for training, which includes a subset of 20% used for validation to mitigate overfitting. The remaining 20% of the overall dataset is earmarked for the test set, which is utilized to evaluate the model's efficacy on data it has not previously encountered, ensuring the assessment of the model's predictive power and generalization ability.

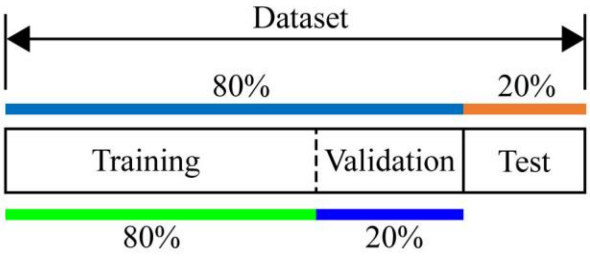


Figure 10 Dataset splitting for training, validation, and test.

To neutralize the scale differences among data points, normalization is applied to the dataset. The common formula for normalization brings the values into a consistent range, typically [0, 1] or [-1, 1]. This is achieved by subtracting the minimum value and dividing by the range of the dataset. The formula (Eq. (6)) for scaling to [-1, 1] is:

$$x' = \frac{2\left(x - \min(x)\right)}{\max\left(x\right) - \min\left(x\right)} - 1 \tag{6}$$

where x is the original value, x' is the normalized value, and min(x) and max(x) are the minimum and maximum values in the data, respectively. This transformation enhances the algorithm's convergence and performance by treating all variables equally.

Random Grid Search (Bergstra & Bengio, 2012) is a method for hyperparameter optimization where a grid of hyperparameter values is set up and random combinations are selected and evaluated. This process offers a pragmatic alternative to exhaustive grid search, particularly when the search space is large. By randomly sampling the space, it provides a cost-effective way of approximating the best hyperparameters for a given model. This method balances the thoroughness of exploring the hyperparameter space with the computational efficiency required in practice.

In the outlined approach, the hyperparameter space for a neural network is defined, with the number of neurons in the hidden layer ranging from 2 to 128 in increments of 2, and learning rates set at 0.0001, 0.001, 0.01, and 0.1. A random grid search is employed to select combinations of neurons and learning rates, ensuring a minimum coverage of 25% of the total hyperparameter space. To prevent overfitting, training ceases if the Mean Squared Error (MSE) on the validation set increases for six consecutive iterations. An epoch consists of a full forward and backward pass of all samples through the network, with the maximum number of epochs set at 1000. The activation functions are 'tansig' (Figure 9 (a)) for input to hidden layers and 'purelin' (Figure 9 (b)) for hidden to output layers. The optimal hyperparameters are those that yield the best validation set performance, and these are designated as the best hyperparameters for the model.

Acknowledging the limitations posed by random initial weights and biases in neural network training, it's a common practice to train the network multiple times using the optimal hyperparameter settings. This approach helps to mitigate the impact of randomness in the initial conditions, thereby providing a more reliable assessment of the network's performance. By repeating the training process with the same optimal hyperparameters, the model's results with best-performing instance can be selected, which improves the robustness and generalizability of the neural network model.

4. Machine Learning Training results

Several machine learning models are developed with a variety of inputs and outputs in this study.

4.1 Model I: 4 dimensional inputs and 2 critical load outputs

Figure 11 illustrates this neural network architecture with the input, hidden, and output layers. The input layer nodes, denoted by H_c, B_c, D_c, and t, represent the features fed into the network. These inputs are connected to the first hidden layer, which processes the inputs through a series of weighted connections. The output layer, depicted with nodes P_{crL} and P_{crD}, signifies the network's final predictions.

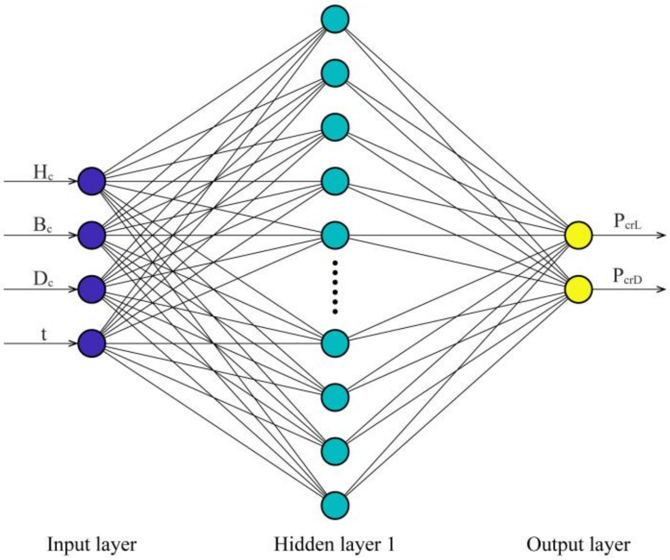


Figure 11 Architecture of the ANN with 4 inputs, 2 outputs and 1 hidden layer

The optimal hyperparameter configuration, as determined by random grid search, encompasses a neural network with 124 neurons in the hidden layer, a learning rate of 0.01, and the employment of 'tansig' and 'purelin' as activation functions for the respective layers. The Levenberg-Marquardt (L-M) training algorithm was adopted, with a stopping criterion set at 1000 epochs, or if the validation set error increased for six consecutive epochs, indicating potential overfitting. This parameterization is poised to balance model complexity with computational efficiency and generalization performance.

Figure 12 illustrates the training progression of the neural network model, represented by the Mean Squared Error (MSE) over the course of 1000 epochs. The training and validation error rates are plotted on a logarithmic scale, demonstrating a rapid decrease in MSE as the number of epochs increases, and then plateauing, indicating the convergence of the model. The close proximity of the training and validation lines suggests that the model is generalizing well without overfitting.

Figure 13 presents two histograms displaying the frequency distribution of errors for two different predicted variables, P_{crL} and P_{crD}, in a neural network model. The histograms are overlaid with a 'Zero Error' line to indicate the point of perfect prediction. The distribution of errors across different magnitudes suggests the model's accuracy in predicting the critical buckling loads for local (P_{crL}) and distortional (P_{crD}) buckling. The concentration of data around the zero-error line indicates high prediction accuracy, while the spread provides insights into the model's error variance.

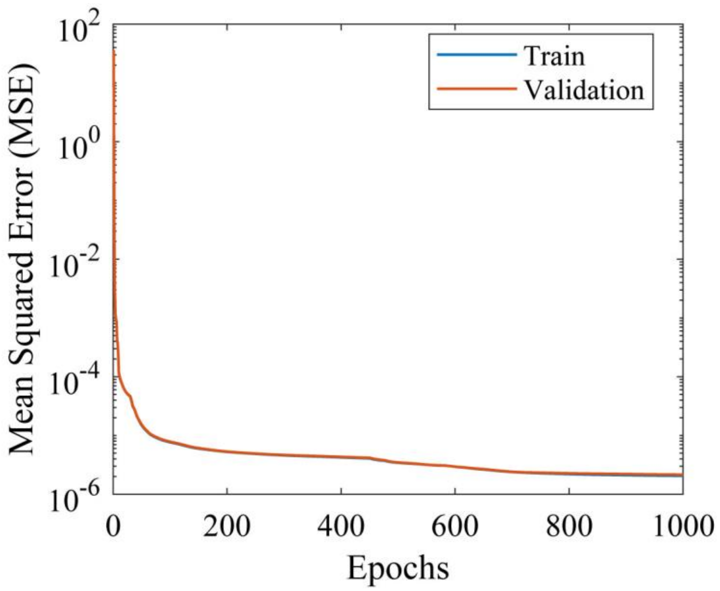


Figure 12 Performance of the ANN model

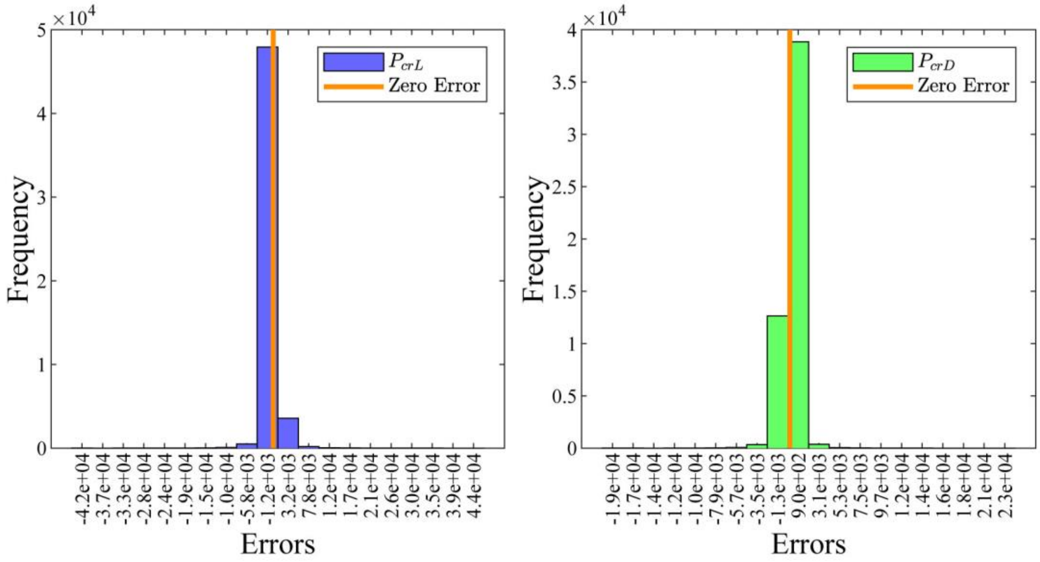


Figure 13 Error histogram for network

Figure 14 exhibits two scatter plots comparing the predicted critical buckling loads, P_{crL} and P_{crD}, from a Backpropagation Neural Network (BPNN) against the calculated values from CUFSM. Figure 14 (a) correlates P_{crL} and Figure 14 (b) for P_{crD}. The color gradient represents density, indicating the concentration of data points. Both plots include a fit line and a 1:1 line, with the fit line demonstrating the regression equation, coefficient of determination (R^2) (Eq. (7)), correlation coefficient (R) (Eq. (1)), and Mean Absolute Percentage Error (MAPE) (Eq. (8)). The close alignment of data points along the 1:1 line and high R^2 values close to unity suggest a strong agreement between the BPNN predictions and CUFSM calculations, signifying high model accuracy.

The coefficient of determination, denoted as R^2 , is calculated using the Eq. (7):

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(y_i - \hat{y}_i \right)^2}{\sum_{i=1}^{n} \left(y_i - \bar{y} \right)^2} \tag{7}$$

where y_i are the observed values, $\hat{y}$ are the predicted values, and $\bar{y}$ is the mean of the observed data. The Mean Absolute Percentage Error (MAPE) is calculated as Eq. (8):

$$MAPE = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \tag{8}$$

where y_i are the actual values and $\hat{y}$ are the forecasted values, and n is the number of observations. MAPE expresses accuracy as a percentage of the error.

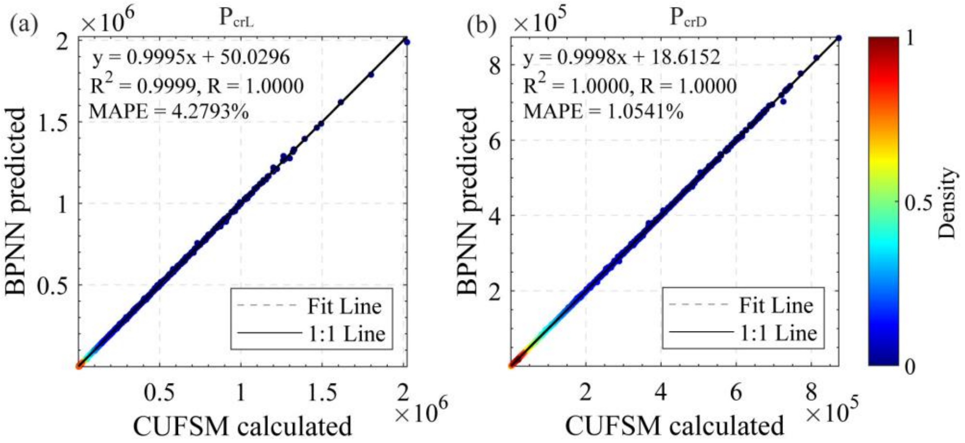


Figure 14 Comparisons of P_{crL} (a) and P_{crD} (b) predicted by BPNN and CUFSM

4.2 Model II: 4 dimensional inputs and 2 critical load and 2 half-wavelength outputs

Figure 15 represents a feedforward neural network architecture, specifically designed for the prediction of structural parameters. The input layer nodes, signified by H_c, B_c, D_c, and t, correspond to the structural features, P_{crL}, P_{crD}, λ_{crL}, and λ_{crD}.

The optimal hyperparameter configuration, as determined by random grid search, encompasses a neural network with 120 neurons in the hidden layer, a learning rate of 0.0001, and the employment of ‘tansig’ and ‘purelin’ as activation functions for the respective layers. The Levenberg-Marquardt (L-M) training algorithm was utilized, with a stopping criterion set at 1000 epochs, or if the validation set error increased for six consecutive epochs, indicating potential overfitting.

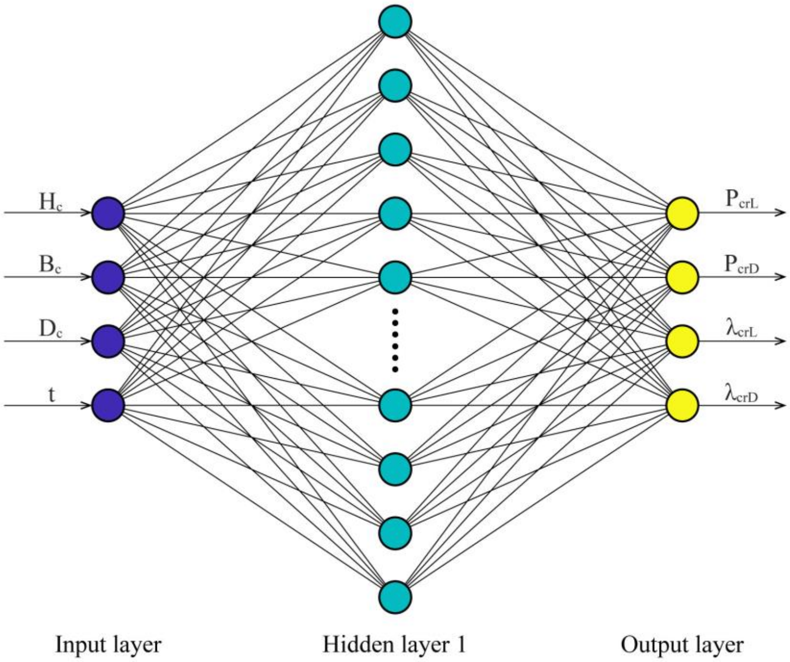


Figure 15 Architecture of the ANN with 4 inputs, 4 outputs and 1 hidden layer

Figure 16 represents a training curve of a neural network, with the Mean Squared Error (MSE) plotted against the number of epochs during training and validation phases. The MSE is on a logarithmic scale to better visualize the rate of decrease. Both training and validation errors decrease sharply and then plateau, indicating the model has reached convergence. The convergence suggests that the model has learned the underlying patterns without overfitting, as indicated by the validation error stabilizing alongside the training error.

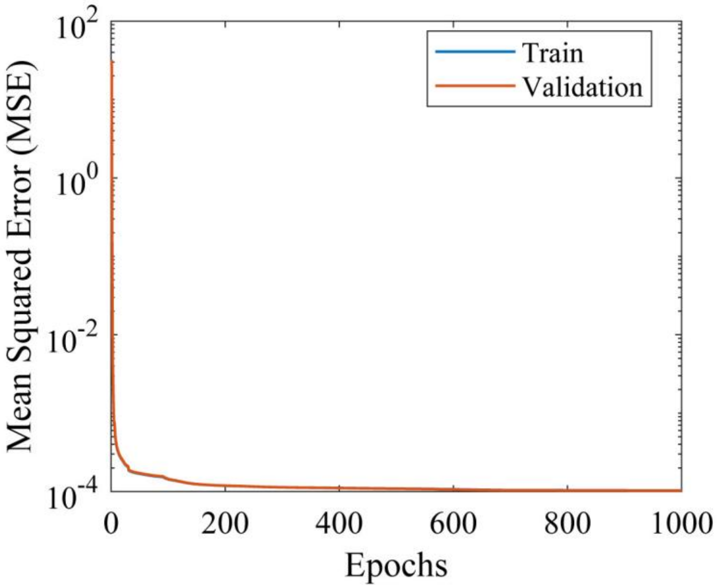
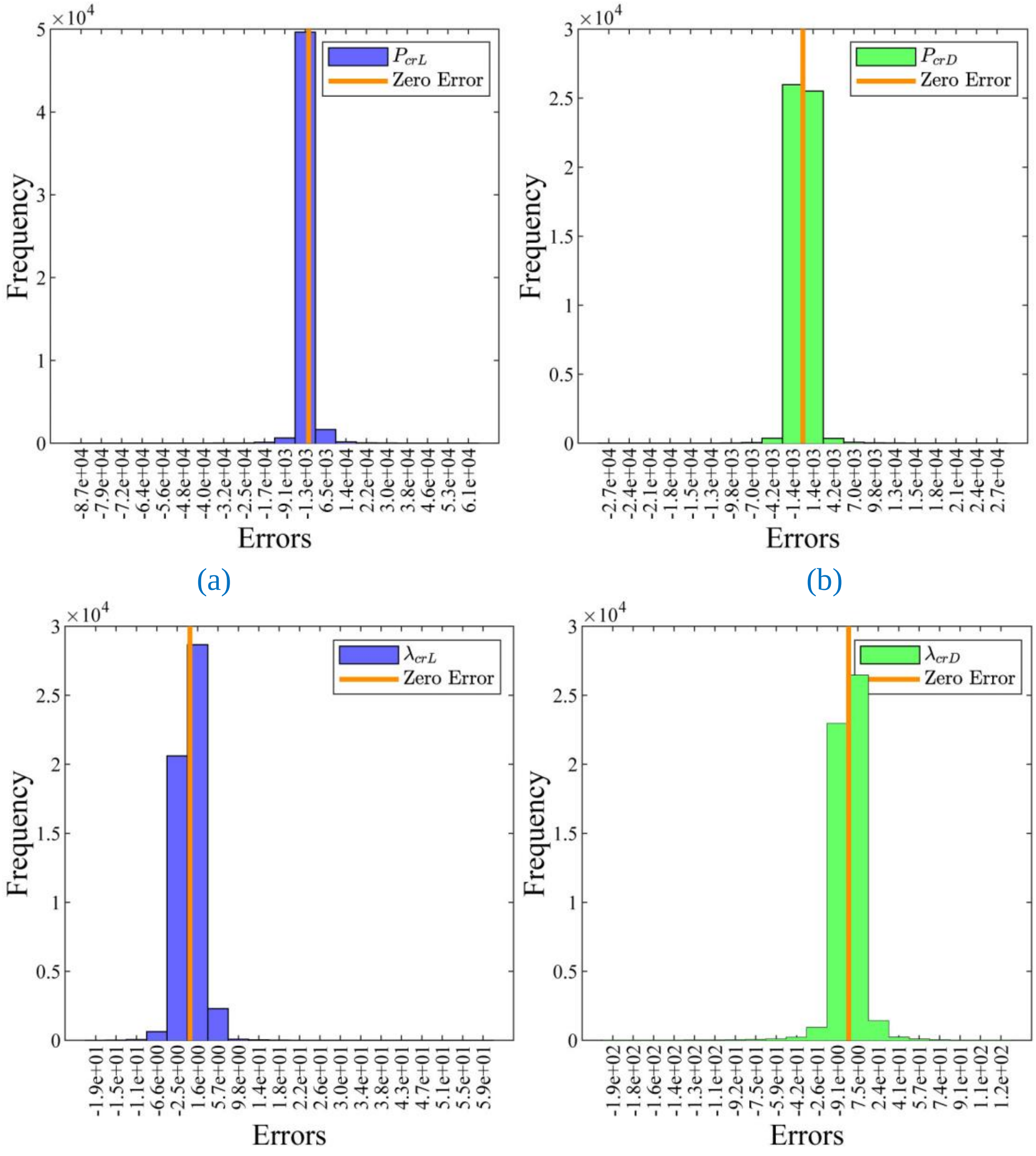


Figure 16 Performance of the ANN model

Figure 17 present histograms of prediction errors from a neural network model. In the four subgraphs, the x-axis signifies the error magnitude, and the y-axis indicates the frequency of these errors. Figure 17 (a) shows the errors for the predicted critical local buckling load (P_{crL}), and the Figure 17 (b) for the predicted critical distortional buckling load (P_{crD}). Figure 17 (c) showcases the frequency of errors for the critical local buckling load (λ_{crL}), while Figure 17 (d) illustrates the frequency of errors for the critical distortional buckling load (λ_{crD}). A 'Zero Error' line is included to benchmark perfect predictions. The dense accumulation of data around this line in both histograms indicates a high accuracy of the neural network model, with the majority of predictions closely aligning with the true values computed by CUFSM.



(a) (b)

(c) (d)

Figure 17 Error histogram of P_{crL}, P_{crD}, λ_{crL} and λ_{crD} for network: (a) P_{crL}; (b) P_{crD}; (c) λ_{crL}; (d) λ_{crD}.

The provided figures represent scatter plots correlating the predicted results from a Backpropagation Neural Network (BPNN) with calculated values using the CUFSM method for different buckling loads. Figures (a) and (b) show the relationships for critical local (P_{crL}) and distortional (P_{crD}) buckling loads, respectively, whereas Figures (c) and (d) correspond to the half-wavelengths for local (λ_{crL}) and distortional (λ_{crD}) buckling. The plots include regression fit lines, the line of perfect prediction (1:1 line), and display the regression equation, coefficient of determination (R^2) (Eq. (7)), correlation coefficient (R) (Eq. (1)), and Mean Absolute Percentage Error (MAPE) (Eq. (8)). The color gradient signifies the density of data points around the fit line, with denser regions indicating higher concentrations of data points. The near-perfect R and R^2 values and low MAPE percentages suggest an excellent predictive performance of the BPNN model. The observation indicates that the predictive performance of the model, as measured by Mean Absolute Percentage Error (MAPE), shows larger discrepancies in estimating P_{crL} and P_{crD}, with MAPE values of 6.6815% and 2.5956% respectively, compared to the half-wavelength parameters λ_{crL} and λ_{crD}. This may suggest differences in model sensitivity or complexity in capturing the phenomena governing P_{crL} and P_{crD}.

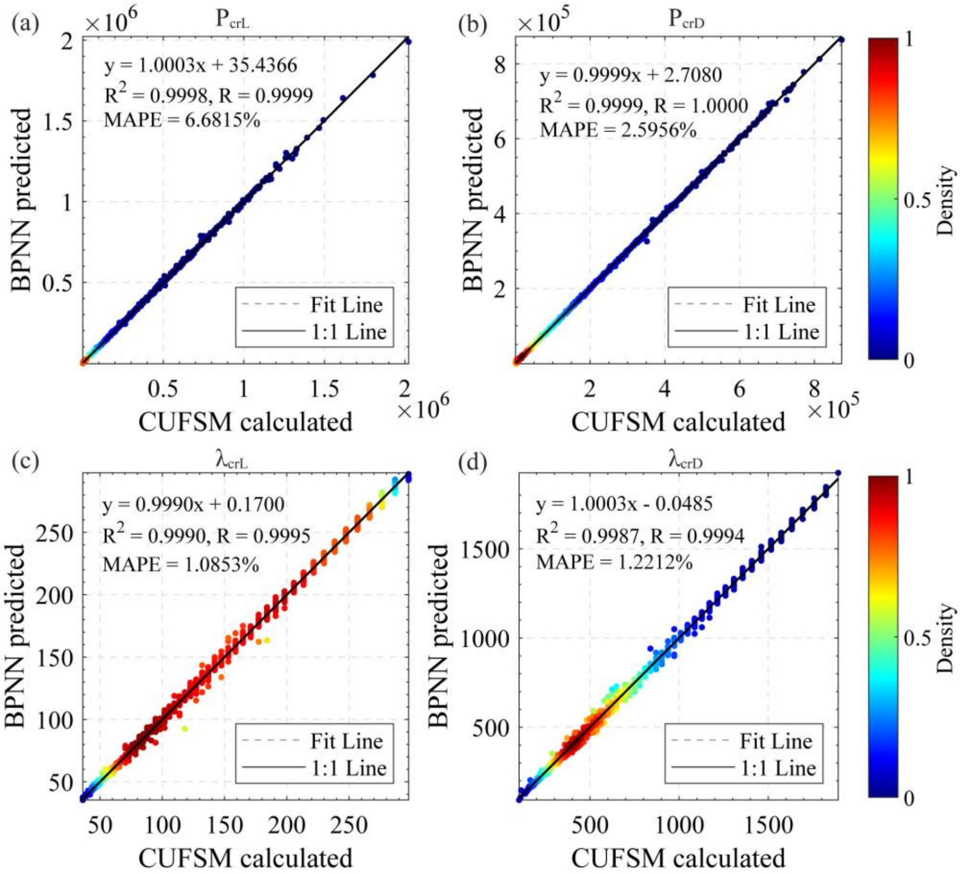


Figure 18 Comparisons of P_{crL}, P_{crD}, λ_{crL} and λ_{crD} predicted by BPNN and CUFSM

4.3 Model III: more inputs

Figure 19 exhibits a neural network structure delineating the data flow from the input to the output layer. The input layer consists of structural features including geometric dimensions and their ratios: H_c, B_c, D_c, t, H_c/t, B_c/t, D_c/t, H_c/B_c, and B_c/D_c. These input neurons are fully connected to a hidden layer, which in turn connects to the output layer. The output neurons correspond to the critical local buckling load (P_{crL}), critical distortional buckling load (P_{crD}), and the associated half-wavelengths for local (λ_{crL}) and distortional (λ_{crD}) buckling.

The optimal hyperparameter configuration, as determined by random grid search, encompasses a neural network with 124 neurons in the hidden layer, a learning rate of 0.001, and the employment of ‘tansig’ and ‘purelin’ as activation functions for the respective layers. The Levenberg-Marquardt (L-M) training algorithm was utilized, with a stopping criterion set at 1000 epochs, or if the validation set error increased for six consecutive epochs, indicating potential overfitting.

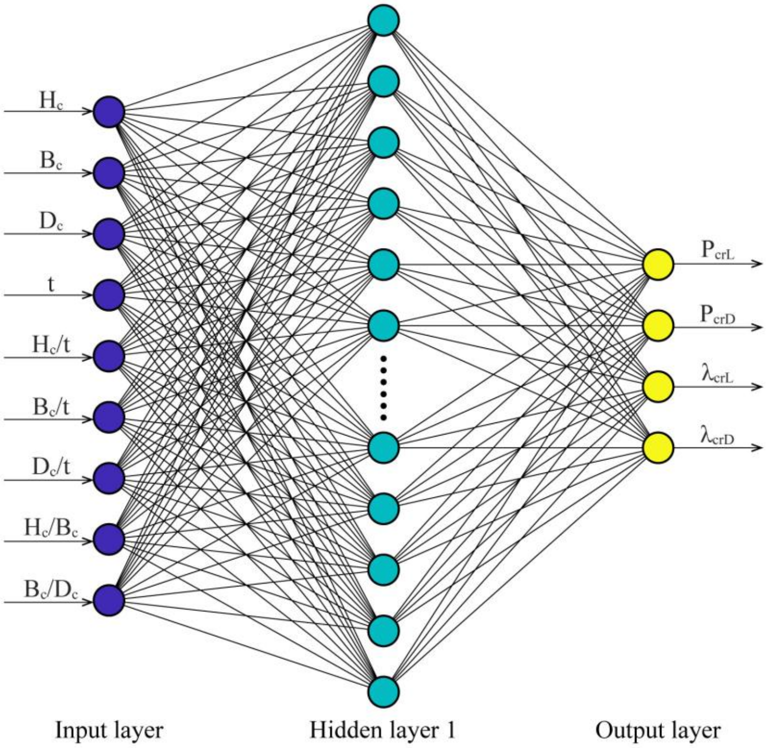


Figure 19 Architecture of the ANN with 9 inputs, 4 outputs and 1 hidden layer

Figure 20 depicts the Mean Squared Error (MSE) in a logarithmic scale over the number of epochs during the training and validation of a neural network model. Both training and validation errors rapidly decrease and plateau early in the training process, indicating quick convergence and suggesting that the model is learning efficiently. The convergence of the training and validation errors suggests a well-fitted model with good generalizability. The absence of significant divergence between the training and validation lines towards the latter epochs implies that the model is not overfitting to the training data.

Figure 21, showcases a series of histograms that illustrate the frequency of prediction errors for a neural network model. Subfigures (a) and (b) display the error distributions for the predicted critical local buckling load (P_{crL}) and critical distortional buckling load (P_{crD}), respectively. Subfigures (c) and (d) detail the error frequencies for the predicted half-wavelengths associated with local (λ_{crL}) and distortional (λ_{crD}) buckling loads. A ‘Zero Error’ line is present in each histogram to benchmark the model’s prediction accuracy. The histograms are instrumental for evaluating the model’s precision, as they reveal the distribution and concentration of errors relative to perfect predictions.

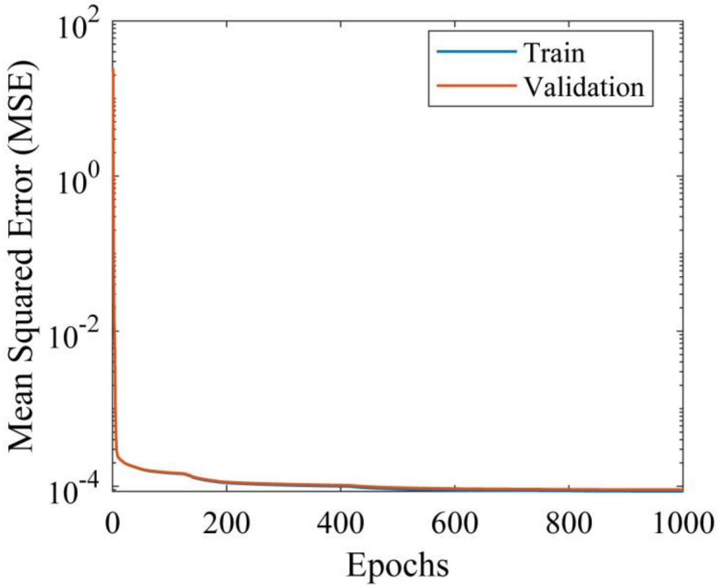
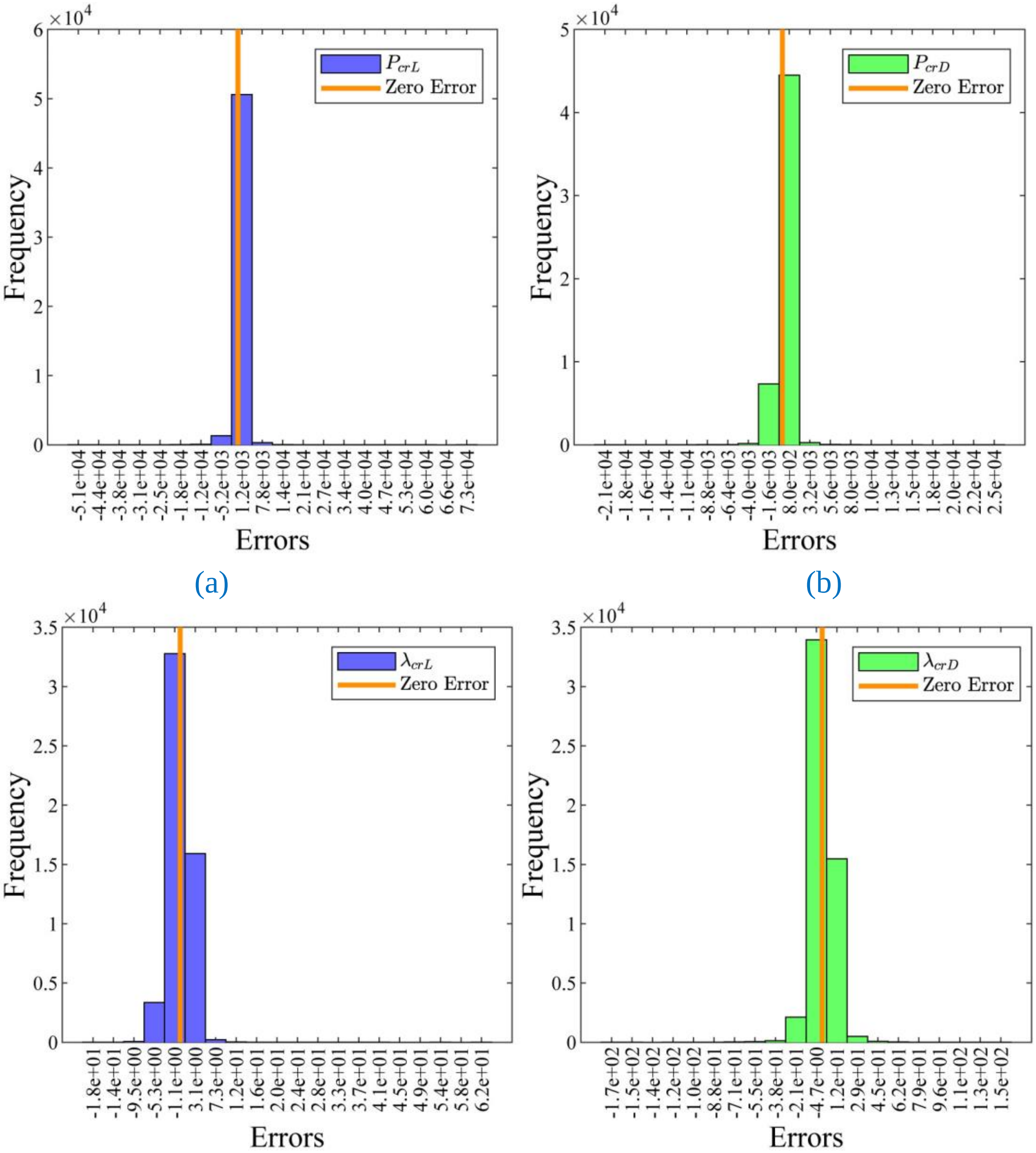


Figure 20 Performance of the ANN model

Figure 22 displays scatter plots with regression lines and density coloration that compare the output of a Backpropagation Neural Network (BPNN) against CUFSM-calculated values for structural parameters. Specifically, subfigures (a) and (b) depict the predicted versus calculated values for critical local (P_{crL}) and distortional (P_{crD}) buckling loads, respectively. Subfigures (c) and (d) showcase the predicted versus calculated half-wavelengths for local (λ_{crL}) and distortional (λ_{crD}) buckling. Each plot includes a fit line and a 1:1 line to indicate perfect predictions, with the regression equation, coefficient of determination (R^2) (Eq. (7)), correlation coefficient (R) (Eq. (1)), and Mean Absolute Percentage Error (MAPE) (Eq. (8)) provided for quantitative assessment. The density gradient within the plots reflects the concentration of data points around the fit line, demonstrating the predictive accuracy of the model across different data densities. While the Mean Absolute Percentage Error (MAPE) is relatively low for all four parameters, it is noted that the MAPE for P_{crL} is comparatively higher at 5.3436%. This suggests that while the model's overall predictive performance is robust, there is a slightly greater average deviation in the prediction of P_{crL} from the actual values when compared to the other parameters.



(a) (b)

(c) (d)

Figure 21 Error histogram of P_{crL}, P_{crD}, λ_{crL} and λ_{crD} for network

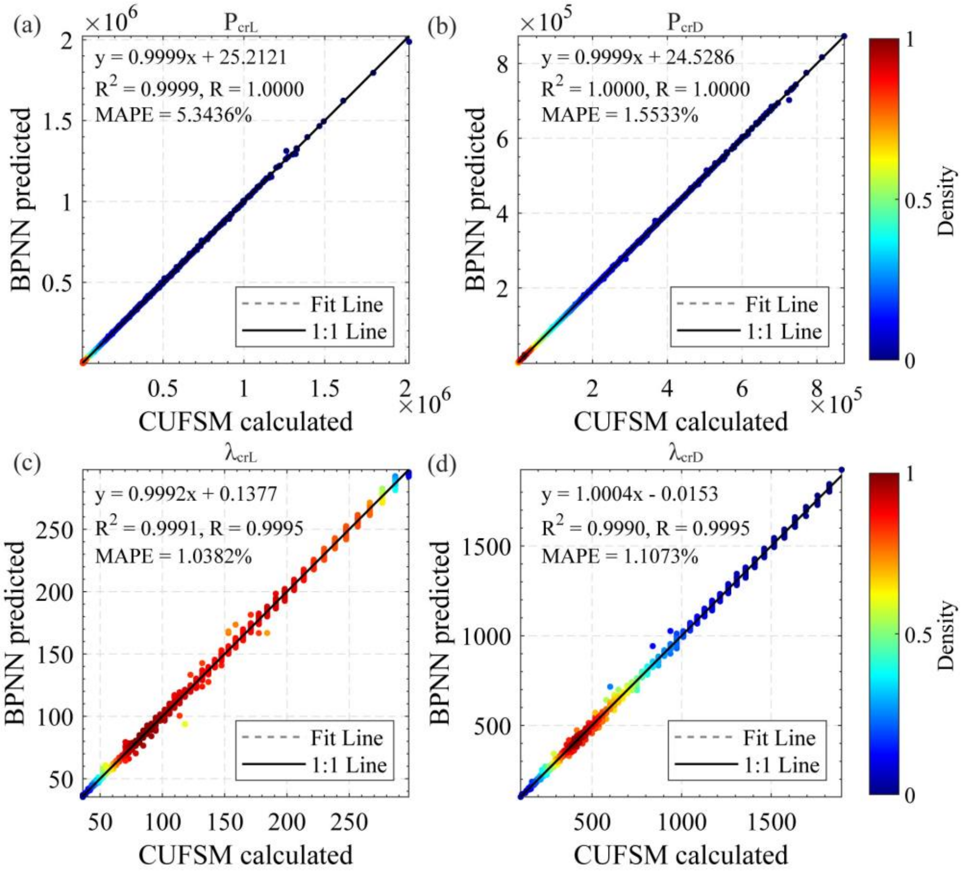


Figure 22 Comparisons of P_{crL}, P_{crD}, λ_{crL} and λ_{crD} predicted by BPNN and CUFSM

4.4 Model IV: ML model with 2 hidden layers

To enhance the predictive capabilities of the neural network model, an investigation into the addition of a second hidden layer was conducted. This architectural modification, as depicted in the referenced Figure 23, involves the implementation of a neural network with two hidden layers. The adjustment aims to capture more complex patterns and interactions within the data, potentially improving the accuracy and generalization of the model's predictions.

Figure 23 illustrates a neural network architecture with two hidden layers. The input layer contains neurons representing both geometric dimensions of structural elements and their derived ratios, signifying the model's inputs. These inputs are connected to the first hidden layer, which then connects to the second hidden layer, adding a level of abstraction and capability to capture more complex relationships within the data. The second hidden layer connects to the output layer, which predicts several structural parameters: the critical local buckling load (P_{crL}), critical distortional buckling load (P_{crD}), and their respective half-wavelengths (λ_{crL} and λ_{crD}).

The optimal hyperparameter configuration, as determined by random grid search, encompasses a neural network with 124-64 neurons in the hidden layers, a learning rate of 0.1, and the employment of 'tansig' and 'purelin' as activation functions for the respective layers. The Levenberg-Marquardt (L-M) training algorithm was utilized, with a stopping criterion set at 1000 epochs, or if the validation set error increased for six consecutive epochs, indicating potential overfitting. This parameterization is poised to balance model complexity with computational efficiency and generalization performance.

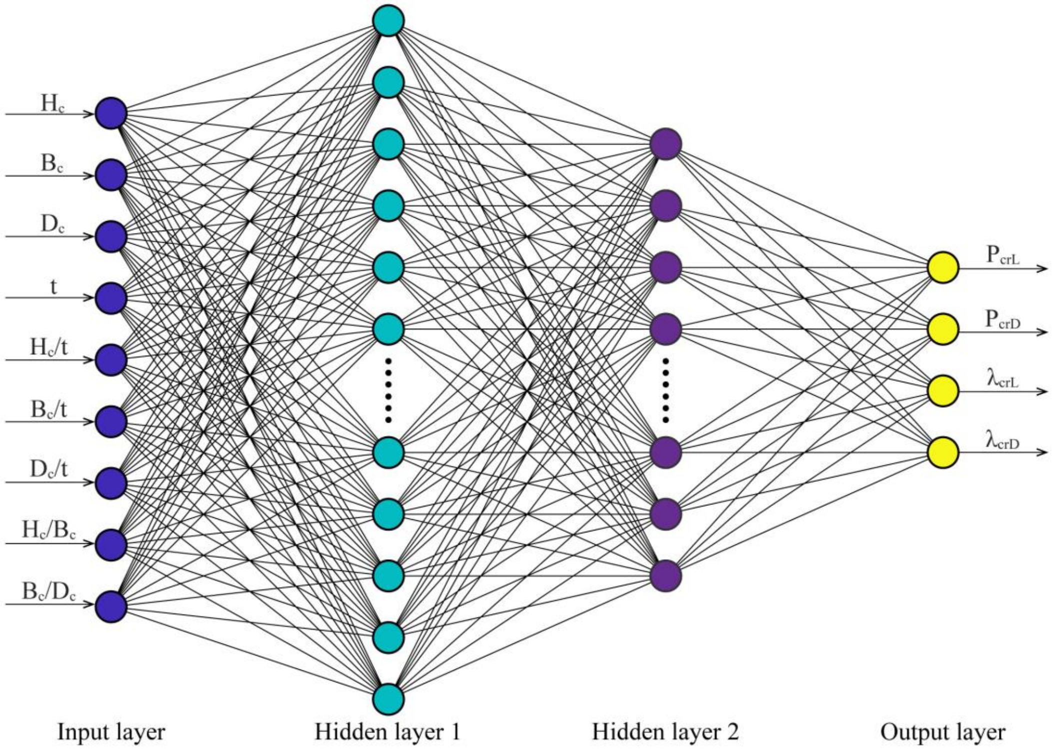


Figure 23 Architecture of the ANN with 9 inputs, 4 outputs and 2 hidden layer

Figure 24 presents the Mean Squared Error (MSE) during the training and validation phases of a neural network over 361 epochs. The MSE is plotted on a logarithmic scale to better visualize the range of error magnitudes. Both the training (blue) and validation (orange) curves show a rapid decrease in error, suggesting that the model is learning effectively from the data. The convergence of these lines indicates that the model generalizes well without overfitting, as the validation error remains low and closely follows the training error. This is a positive indication of the model's robustness.

Figure 25 exhibits a series of histograms detailing the frequency of errors between predicted and actual values for different structural parameters within a neural network model. The histograms compare errors in predictions of critical local (P_{crL}) and distortional (P_{crD}) buckling loads, as well as their corresponding half-wavelengths (λ_{crL} and λ_{crD}). Each histogram is juxtaposed with a 'Zero Error' line to indicate the ideal outcome where the predicted values exactly match the actual ones. The distribution of errors across these histograms provides insights into the accuracy and precision of the neural network's predictive capabilities.

- Figure 25 (a) displays the error distribution for P_{crL}, with the majority of the frequency bars clustered close to the Zero Error line, which suggests high accuracy in prediction.
- Figure 25 (b) shows the error distribution for P_{crD}, where the histogram bars are also concentrated near the Zero Error line, indicating a similarly high predictive accuracy.

- Figure 25 (c) represents the error distribution for λ_{crL}, once again demonstrating a tight grouping near the Zero Error line, suggesting that the predictions are generally accurate with minimal deviation.
- Figure 25 (d) presents the error distribution for λ_{crD}, with the histogram bars showing a similar pattern to the other parameters, clustered near the Zero Error, which implies precision in the neural network's predictive performance.

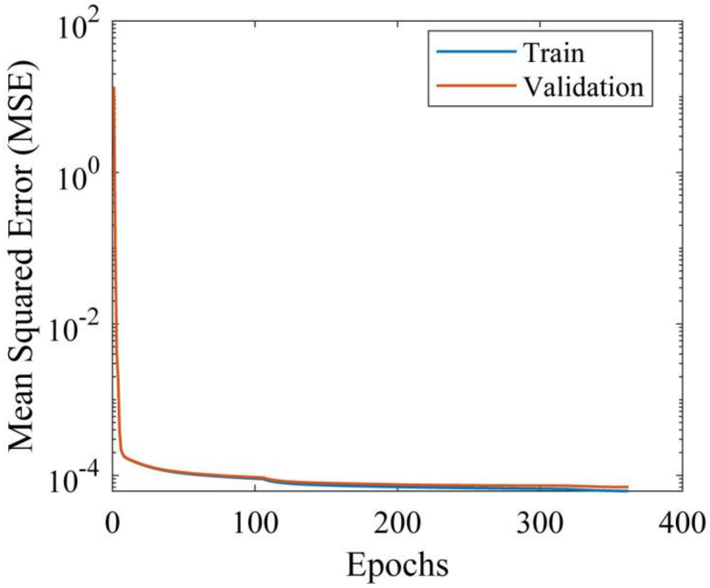
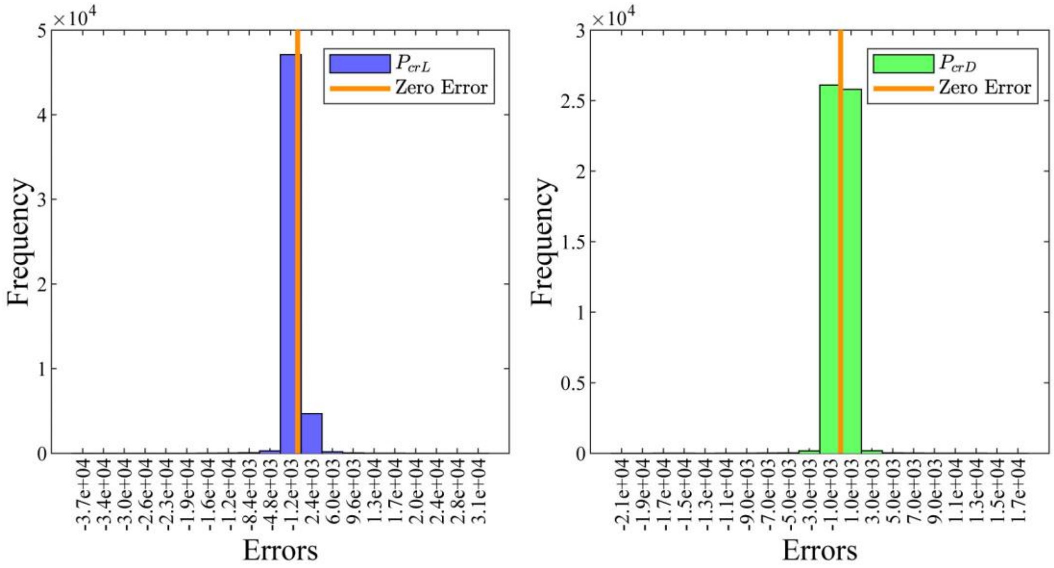
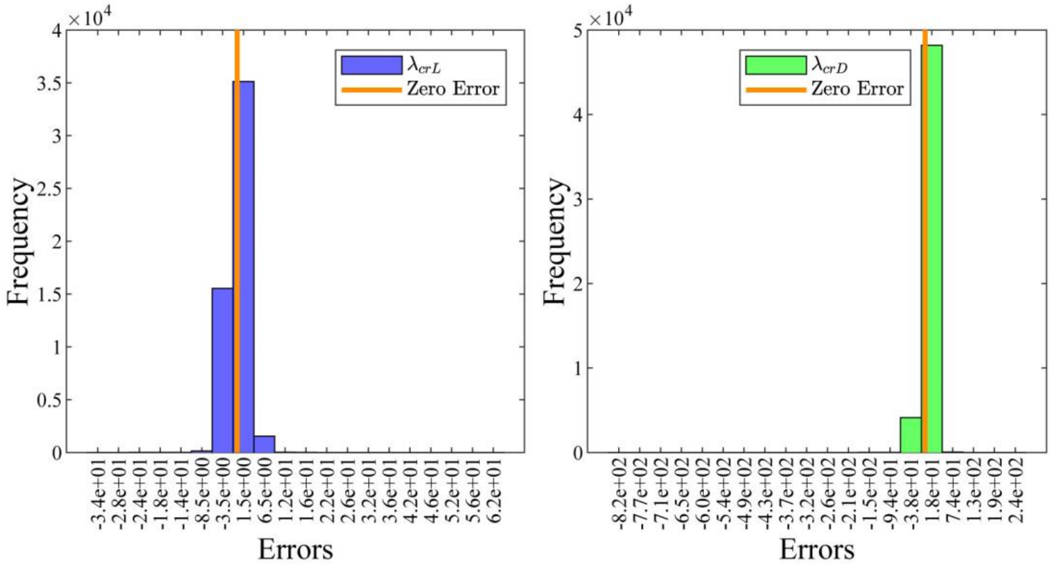


Figure 24 Performance of the ANN model



(a) (b)



(c) (d)

Figure 25 Error histogram of P_{crL}, P_{crD}, λ_{crL} and λ_{crL} for network

Figure 26 exhibits a dense clustering of data points along the line of perfect agreement (the 1:1 line), which indicates a high level of precision in the neural network's predictive capability. This is further substantiated by the regression equations and the statistical indices provided: the coefficient of determination (R^2), the Pearson correlation coefficient (R), and the Mean Absolute Percentage Error (MAPE). In all cases, the R^2 and R values are extremely close to 1, and the MAPE values are low, signaling excellent model performance. The color bar to the right of each plot represents the density of data points, with warmer colors indicating a higher concentration of points.

- Figure 26 (a) shows a plot for P_{crL} with a nearly perfect R^2 and R value of 1.0000, and a MAPE of 3.2423%, indicating very high predictive accuracy.
- Figure 26 (b) displays results for P_{crD}, again with R^2 and R values of 1.0000, and a slightly lower MAPE of 1.1052%, suggesting a similarly high level of prediction accuracy.
- Figure 26 (c) introduces the λ_{crL} parameter, where the R^2 is 0.9989, R is 0.9995, and MAPE is 1.0146%, showing a marginal decrease in prediction precision but still within an excellent range.
- Figure 26 (d) presents λ_{crD} data, with a R^2 of 0.9990, R of 0.9995, and MAPE of 1.0391%, demonstrating a consistency in the high predictive ability of the BPNN for this parameter as well.

The dashed lines represent the fit line from the regression analysis, which is used to generate the equations and statistical indices. These plots are a strong visual representation of the neural network's ability to accurately predict structural parameters.

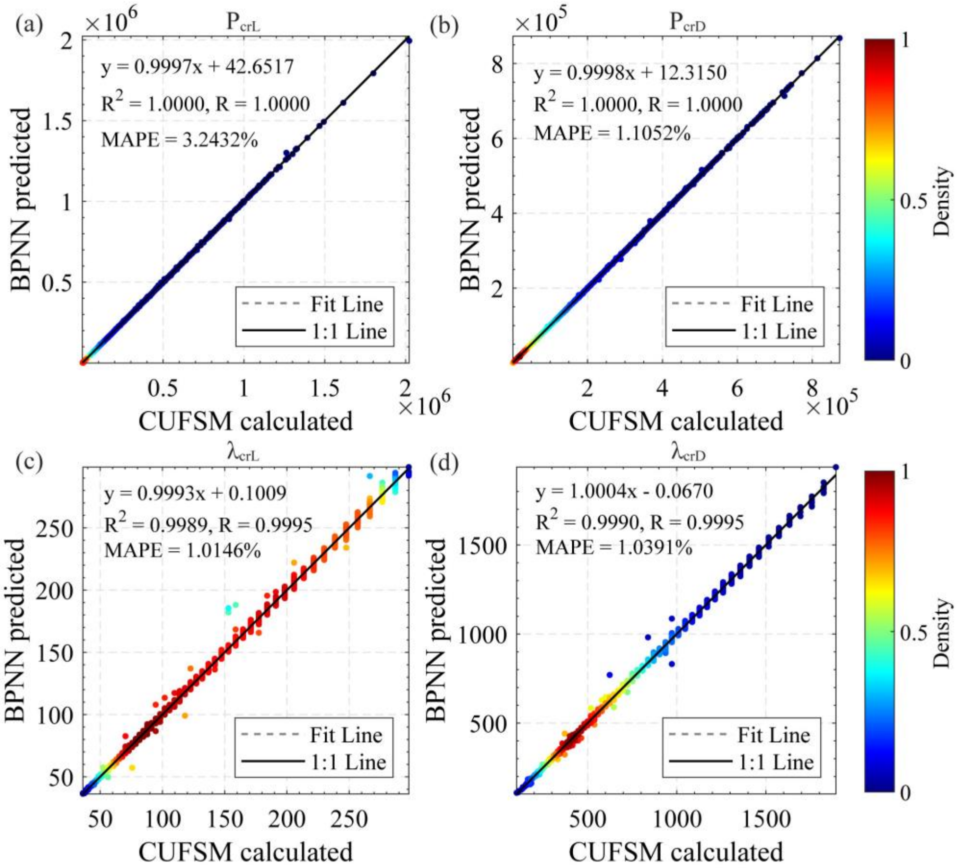


Figure 26 Comparisons of P_{crL}, P_{crD}, λ_{crL} and λ_{crD} predicted by BPNN and CUFSM

5. Comparison with AISI analytical equations

The current AISI S100 (2020) Appendix 2 provides analytical expressions of critical local buckling stresses P_{crL} of individual plate elements of a cross section. The general equation for critical local buckling stress in AISI S100 (2020) is shown in Eq. (9) and Eq. (10).

$$P_{crL} = A_g F_{crL} \tag{9}$$

where A_g is gross cross-sectional area.

$$F_{crL} = k \frac{\pi^2 E}{12\left(1-\mu^2\right)} \left(\frac{t}{w}\right)^2 \tag{10}$$

where k is the plate buckling coefficient, t is the plate thickness, w is the plate flat width, E is Young's modulus and μ is Poisson's ratio.

- For stiffened element under uniform compression from AISI S100 Section 1.1: $k = 4$
- For unstiffened element under pure compression per AISI S100 Section 1.2.1: $k = 0.43$

The elastic distortional buckling load in AISI S100 (2020), P_{crd}, shall be calculated as follows Eq. (11) and Eq. (12):

$$P_{crD} = A_g F_{crD} \tag{11}$$

where A_g is gross cross-sectional area.

$$F_{crd} = \frac{\bar{k}_{ffe} + \bar{k}_{fwe} + \bar{k}_{f}}{\tilde{k}_{ffg} + \tilde{k}_{fwg}} \qquad (12)$$

See AISI S100 (2020) Eq. 2.3.3.1-1 to Eq. 2.3.3.1-7 for detailed formulas.

The comparison between the (AISI S100 (2020)) formulas and the Backpropagation Neural Network (BPNN) models is conducted using an identical test set. This direct comparison allows for an assessment of the predictive accuracy and generalization capability of the BPNN against the established AISI empirical formulations.

- Figure 27 consists of four scatter plots (Figure 27 (a), (b), (c), (d)), which are used to assess the accuracy of two different predictive models: an ANN model, presumably a Back Propagation Neural Network (BPNN), and an analytical model based on the (AISI S100 (2020)) standards.
- Figure 27 (a) contrasts the AISI analytical predictions for P_{crL} with those calculated by the CUFSM. The plot shows a fair correlation, as indicated by the coefficient of determination (R^2) of 0.9542 and the Pearson correlation coefficient (R) of 0.9768. However, the Mean Absolute Percentage Error (MAPE) is relatively high at 27.6046%, suggesting moderate discrepancies between the methods.
- Figure 27 (b) demonstrates the BPNN predicted values for PcrL versus the CUFSM calculations. This model exhibits an excellent fit, with both R^2 and R reaching the ideal value of 1.0000 and a low MAPE of 3.2432%, indicating very high predictive accuracy.
- Figure 27 (c) shows the AISI analytical predictions for P_{crD} against CUFSM calculations, with a R^2 of 0.9510 and R of 0.9752. The MAPE is also high at 27.3183%, similar to Figure 26 (a), reflecting significant predictive variance.
- Figure 27 (d) displays the BPNN predicted values for P_{crD}, closely aligning with the CUFSM calculations, as seen by the ideal R^2 and R values of 1.0000 and a minimal MAPE of 1.1052%, showcasing the superior predictive performance of the BPNN model.

The dashed line in each plot represents the regression fit, while the solid line depicts the ideal 1:1 correspondence between predicted and calculated values. A color bar on the right side of each plot indicates the density of data points, with warmer colors signifying a higher concentration of closely matching values. Overall, these plots suggest that the BPNN model has a significantly higher predictive accuracy compared to the AISI analytical predictions when evaluated against the CUFSM calculations. The consistent high performance of the BPNN across both P_{crL} and P_{crD} parameters indicates its potential as a reliable predictive tool in structural engineering applications.

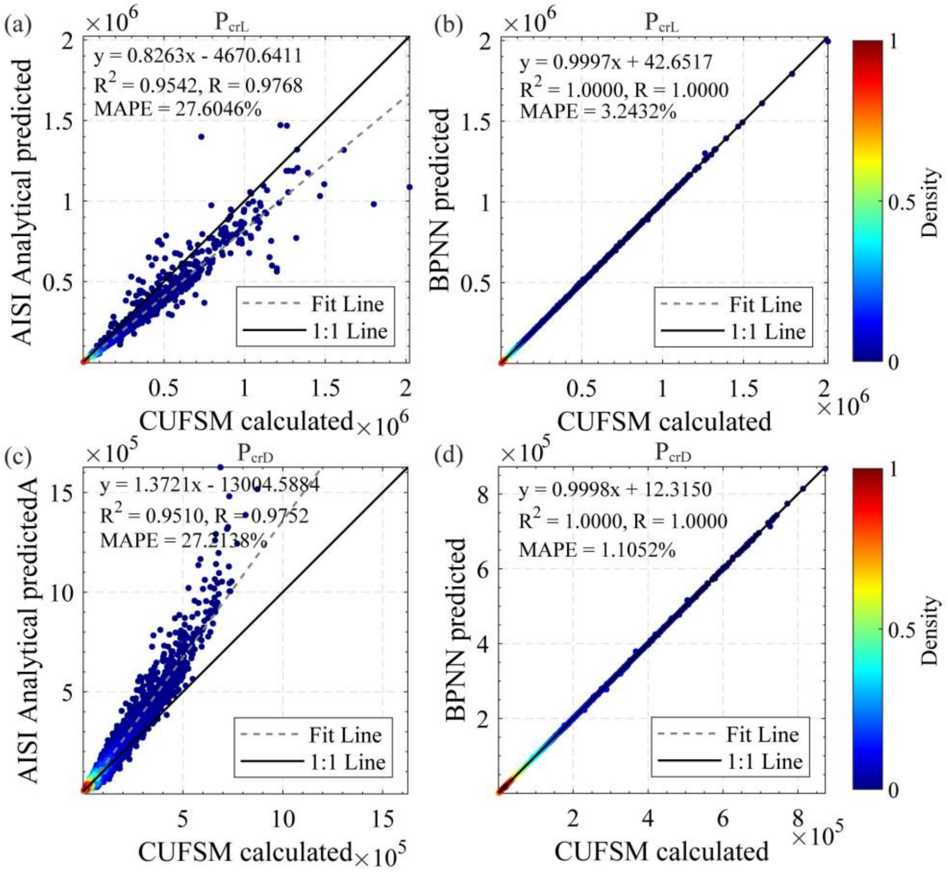


Figure 27 Comparisons of PcrL, PcrD, λcrL and λcrD predicted by BPNN and AISI Equation

Meanwhile, Ding and Schafer (2023) have proposed a new formula (Eq. (13)) for calculating the critical local buckling load (P_{crL}) for lipped channels without punchouts.

$$F_{crl,nh} = k_h \frac{\pi^2 E}{12\left(1\text{-}v^2\right)}\left(\frac{t}{h}\right)^2$$
$$k_h = 4 + \frac{1.2\eta_h}{1+0.22\eta_h+0.05\eta_h^2} \tag{13}$$

where $\eta_h = h/b$, applicable for $1.2 \le \eta_h \le 22$ and $r/t \ge 1.5$

Figure 28 illustrates two scatter plots (Figure 28 (a) and (b)) that compare the predictions of critical load (PcrL) for lipped channels without punchouts using two different predictive methods against values calculated by the CUFSM.

- Figure 28 (a) presents a comparison of the new analytical formula proposed by Ding and Schafer (2023) with the CUFSM calculated values. The linear regression equation provided (y=1.0469x−1632.9642), along with the coefficient of determination (R^2=0.9913) and the Pearson correlation coefficient (R=0.9956), suggest a strong linear relationship with a high degree of fit to the calculated data.
- Figure 28 (b) compares the predictions from a Back Propagation Neural Network (BPNN) with the CUFSM calculated values. The predictive performance of the BPNN is depicted as near-perfect with a R^2 and R of 1.0000. The regression equation (y=1.0002x−12.6004)

almost perfectly matches the 1:1 line, and the MAPE is very low at 3.0492%, indicating an exceptional level of accuracy in the BPNN predictions.

In both plots, the dashed line represents the regression fit to the data, while the solid line indicates the ideal 1:1 relationship where predicted values perfectly match the calculated ones. The color gradient represents the density of the data points, with warmer colors indicating a higher concentration of points. Overall, the figure suggests that both the new analytical formula by Ding and Schafer (2023) and the BPNN model provide accurate predictions of P_{crL} for lipped channels without punchouts, with the BPNN model demonstrating a marginally higher predictive accuracy as indicated by the statistical measures provided.

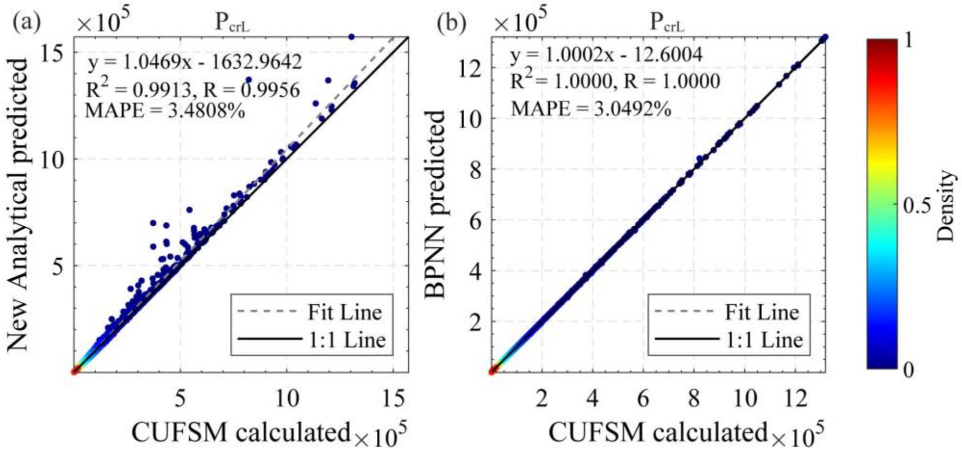


Figure 28 Comparisons of PcrL, PcrD, λcrL and λcrD predicted by BPNN and New analytical

6. Discussions

In this study, the predictive performance of the ANN model, AISI S100 equation and the newly proposed analytical equation was quantitatively assessed using the MAPE for various structural parameters: P_{crL}, P_{crD}, λ_{crL}, and λ_{crD}. The ANN models, with architectures denoted as 4-124-2, 4-120-4, 9-124-4, and 9-124-64-4, exhibited a progressive improvement in prediction accuracy for both critical load parameters (P_{crL} and P_{crD}) as the complexity of the model increased. The most intricate ANN model (9-124-64-4) achieved the lowest MAPE values of 3.2432% for P_{crL} and 1.1052% for P_{crD}, indicating a superior predictive capability compared to its simpler counterparts.

Table 1 MAPE of ANN model and Analytical Equation (%)

Predictive model	PcrL	PcrD	λcrL	λcrD
4-124-2	4.2793	1.0541	-	-
4-120-4	6.6815	2.5956	1.0853	1.2212
9-124-4	5.3436	1.5533	1.0382	1.1073
9-124-64-4	3.2432	1.1052	1.0146	1.0391
AISI S100 equation	27.6046	27.2138	-	-
New Analytical equation	3.4808	-	-	-

*4-124-2: 4 inputs, 124 neurons in hidden layer and 2 outputs

The discrepancies in the MAPE values for the λ_{crL} and λ_{crD} were minimal across different ANN models, with all values being relatively low, indicating high precision in the prediction of these parameters. It is noteworthy that the half-wavelength predictions did not exhibit as substantial a variation with model complexity as the critical load predictions, which may be attributed to the inherent properties of the half wave length that are less sensitive to the modeling approach.

The AISI predictions yielded the highest MAPE values for both critical load parameters, which underscores the potential limitations of conventional methods in capturing the complex behavior of lipped channels without punchouts. The data unequivocally supports the adoption of the new analytical formula and the ANN approach, particularly the most complex model, in predicting the critical load capacities with greater accuracy. Compared to the new analytical equation, the Artificial Neural Network (ANN) models developed in this study demonstrate superior predictive accuracy and robustness. Furthermore, they possess the capability to simultaneously predict additional structural parameters with high precision. This underlines the efficacy of ANN methodologies in capturing complex relationships within the data, which may not be fully represented by traditional analytical equations.

These findings have important implications for the design and analysis of cold-formed steel structures, where predictive accuracy is paramount. The new analytical formula offers a simplified yet effective alternative to computational methods, while the ANN models provide a robust tool for capturing non-linear relationships in structural behavior. Future research should aim to explore the integration of ANN models into design software to facilitate their use in engineering practice.

7. Conclusions

In this study, a comprehensive dataset comprising 260,000 different cross-sectional parameters of cold-formed steel structures, including PcrL, PcrD, λcrL, and λcrD, was calculated and subdivided into training, validation, and test sets. The training set was employed to adjust the weights and biases within the neural network model, the validation set to monitor the model's generalization error, and the test set to evaluate the model's performance on unseen data. A total of 180 neural network models with varying hyperparameters were constructed, with the optimal hyperparameters identified through random grid search. To counter the influence of random initial weights and biases on loss function minimization, the models underwent ten training iterations under the selected hyperparameters. The findings demonstrate that a neural network architecture with two hidden layers yielded the most precise fitting results, indicating a substantial improvement over traditional AISI analytical equations. This suggests a promising direction for design practices in the industry, advocating for more efficient, data-driven methodologies that accommodate the complexities of different structural scenarios. However, it is important to acknowledge that while the two-layer hidden neural network models predict structural parameters with high accuracy, the computational cost, particularly when employing the Levenberg-Marquardt (L-M) algorithm, is substantial. The average duration to train a model once was approximately 65 hours. Additionally, the use of approximate Hessian matrices precluded the exploitation of GPU acceleration, unlike gradient descent methods. Therefore, future research should seek to explore a training algorithm that strikes a balance between predictive accuracy and computational efficiency, potentially harnessing the power of parallel processing to reduce training times while maintaining or enhancing the predictive capabilities of the models.

Acknowledgments

This paper is based in part upon work supported by the National Natural Science Foundation of China (No. 52278150). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the funding agency.

References

AISI. (2020). North American specification for the design of cold-formed steel structural members. AISI S100-20. Washington, DC: American Iron and Steel Institute.

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. Journal of machine learning research, 13(2).

Cheung, Y., & Tham, L. (1998). Finite Strip Method. CRC Press.

Cook, R. D. (2007). Concepts and applications of finite element analysis. John wiley & sons.

Degtyarev, V., & Naser, M. (2021). Boosting machines for predicting shear strength of CFS channels with staggered web perforations. Structures,

Degtyarev, V. V. (2021). Neural networks for predicting shear strength of CFS channels with slotted webs. Journal of Constructional Steel Research, 177, 106443.

Ding, C., & Schafer, B. W. (2023). Analytical Equations for Critical Local Buckling Stress of Lipped Channels.

Elzein, A. (2013). Plate stability by boundary element method (Vol. 64). Springer Science & Business Media.

Guzelbey, I. H., Cevik, A., & Erklig, A. (2006). Prediction of web crippling strength of cold-formed steel sheetings using neural networks. Journal of Constructional Steel Research, 62(10), 962-973.

Hancock, G. J., Murray, T., & Ellifrit, D. S. (2001). Cold-formed steel structures to the AISI specification. CRC Press.

Harik, I., Liu, X., & Ekambaram, R. (1991). Elastic stability of plates with varying rigidities. Computers & structures, 38(2), 161-169.

Iranmanesh, A., & Kaveh, A. (1999). Structural optimization by gradient ‐ based neural networks. International journal for numerical methods in engineering, 46(2), 297-311.

Kaveh, A., Bakhshpoori, T., & Hamze-Ziabari, S. M. (2018). GMDH-based prediction of shear strength of FRP-RC beams with and without stirrups. Computers and Concrete, an International Journal, 22(2), 197-207.

Kaveh, A., Elmieh, R., & Servati, H. (2001). Prediction of moment-rotation characteristic for semi-rigid connections using BP neural networks.

Kaveh, A., & Iranmanesh, A. (1998). Comparative study of backpropagation and improved counterpropagation neural nets in structural analysis and optimization. International Journal of space structures, 13(4), 177-185.

Li, Z., Schafer, B.W. (2010), Application of the finite strip method in cold-formed steel member design, Journal of Constructional Steel Research, Volume 66, Issues 8–9, Pages 971-980

Li, Z., & Schafer, B. (2011). Local and distortional elastic buckling loads and moments for SSMA stud sections. Cold-Formed Steel Engineers Institute, 5.

Li, Z., & Schafer, B. W. (2010). Buckling analysis of cold-formed steel members with general boundary conditions using CUFSM conventional and constrained finite strip methods.

Manual, D. (2017). 2017 AISI Cold-Formed Steel Design Manual.

Markopoulos, A. P., Manolakos, D. E., & Vaxevanidis, N. M. (2007). Prediction of the collapse modes of PVC cylindrical shells under compressive axial loads using Artificial Neural Networks. IFIP International Conference on Artificial Intelligence Applications and Innovations,

Morgan, N., & Bourlard, H. (1989). Generalization and parameter estimation in feedforward nets: Some experiments. Advances in neural information processing systems, 2.

Pala, M. (2006). A new formulation for distortional buckling stress in cold-formed steel members. Journal of Constructional Steel Research, 62(7), 716-722.

Pala, M., & Caglar, N. (2007). A parametric study for distortional buckling stress on cold-formed steel using a neural network. Journal of Constructional Steel Research, 63(5), 686-691.

Pitton, S. F., Ricci, S., & Bisagni, C. (2019). Buckling optimization of variable stiffness cylindrical shells through artificial intelligence techniques. Composite Structures, 230, 111513.

Schafer, B. W. (2008). The direct strength method of cold-formed steel member design. Journal of Constructional Steel Research, 64(7-8), 766-778.

Schafer, B. W., & Adany, S. (2006). Buckling analysis of cold-formed steel members using CUFSM: conventional and constrained finite strip methods. Eighteenth international specialty conference on cold-formed steel structures, Orlando, FL,

Schardt, R. (1989). Verallgemeinerte Technische Biegetheorie.

Tohidi, S., & Sharifi, Y. (2015). Neural networks for inelastic distortional buckling capacity assessment of steel I-beams. Thin-Walled Structures, 94, 359-371.

Von Kármán, T., Sechler, E. E., & Donnell, L. (1932). The strength of thin plates in compression. Transactions of the American Society of Mechanical Engineers, 54(2), 53-56.

Wang, T., Zha, Z., & Pan, C. (2023). Prediction for elastic local buckling stress and ultimate strength of H-section beam. Heliyon, 9(4).

Waszczyszyn, Z., & Bartczak, M. (2002). Neural prediction of buckling loads of cylindrical shells with geometrical imperfections. International journal of non-linear mechanics, 37(4-5), 763-775.

Winter, G. (1947). Strength of thin steel compression flanges. Transactions of the American Society of Civil Engineers, 112(1), 527-554.

Wu, T.-Y., El-Tawil, S., & McCormick, J. (2019). Effect of cyclic flange local buckling on the capacity of steel members. Engineering Structures, 200, 109705.

Zarringol, M., Thai, H.-T., Thai, S., & Patel, V. (2020). Application of ANN to the design of CFST columns. Structures,

Zienkiewicz, O. C., & Taylor, R. L. (2005). The finite element method for solid and structural mechanics. Elsevier.