Article:

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Machine learning-based predictions of buckling behaviour of cold-formed steel structural elements

Seyed Mohammad Mojtabaei[1,2], Jurgen Becque[3], Rasoul Khandan[1], Iman Hajirasouliha[2]

2

Department of Civil and Structural Engineering, The University of Sheffield, Sheffield S1 3JD, UK
[3] Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, UK

1   Introduction

Cold-formed steel (CFS) structural elements are produced from thin steel plate near room temperature. They offer numerous benefits, such as high strength-to-weight and stiffness-to-weight ratios, ease of handling, a flexible manufacturing process, and recyclability without any loss of quality [1-3]. However, due to their limited wall thickness, they are susceptible to instabilities, which necessitates a more complex design process.

The traditional standard-prescribed design methods for cross-sectional instability, in particular the effective width approach, are based on ignoring the interaction between plate elements and treating them as hinged along adjoining lines. The Eurocode (EN1993-1-3) [4] and North-American design specifications (AISI-S100) [5] follow this approach but become cumbersome when applied to more complex cross-sections. These rules are seen as an impediment to further development, optimization and innovation in the field due to their reliance on traditional distinctions between flanges and webs, and their inability to accommodate non-traditional geometries.

The Direct Strength Method (DSM) has therefore become

a popular alternative design method [6]. The DSM requires the determination of the local, distortional, and global buckling stresses of a member and combines those with the yield stress of the material to define a slenderness value for each type of instability. The DSM has historically been linked to the Finite Strip Method (FSM), which can be used in the design process to determine the elastic buckling stresses. However, difficulties can arise in practical application, such as the occurrence of indistinct minima and coupled modes [7]. Research into the *modal decomposition problem* aims to determine the buckled shapes and buckling stresses of pure local, distortional, and global modes, as well as their contributions in a randomly deformed shape. Two more types of pure modes, namely shear modes and (transverse) extension modes, are typically added to enable a complete decomposition. The equivalent nodal forces method (ENFM) is the most robust and general solution to the modal decomposition problem and is employed in this research [8].

The primary objective of this paper is to explore the feasibility of utilizing Machine Learning to predict the buckling behavior of CFS members. Furthermore, this study aims

to offer a reliable and effective solution to the modal decomposition problem.

It is noted that several studies have previously used AI methods to predict the strength of CFS profiles and generate optimized shapes [9-16]. However, this study focused on the development of Artificial Neural Network (ANN) models, which is a subset of machine learning techniques, to predict the elastic critical buckling load and the modal contributions in the buckled shape of CFS structural elements. The ANN models were trained on datasets compiled using Finite Strip Method (FSM) output to determine elastic critical buckling loads and moments, and Equivalent Nodal Force Method (ENFM) output to determine modal decompositions. The study aimed to develop a practical design tool for engineers and practitioners to understand the buckling behavior of CFS elements without nessesarily requiring programming skills or complex mechanical concepts.

2 Dataset and parameter space

This study involved compiling a dataset of 4608 CFS elements with lipped channel cross-sections. The dataset included various lengths, cross-sectional dimensions and thicknesses, and accounted for the possible presence of intermediate stiffeners in the web and/or flanges. Four cross-sectional shapes were considered, each corresponding to 1170 data points. The input data consisted of seven independent parameters, including six cross-sectional parameters (the web height (h), the flange width (b), the lip length (c), the plate thickness (t), and the locations of the intermediate stiffeners in the flanges (r_1) and the web (r_2)), and the element length (L). Table 1 shows the details of the four cross-sectional shapes considered, which include a lipped channel (LP), a lipped channel with intermediately stiffened flanges (LPF), a lipped channel with a stiffened web (LPW), and a lipped channel with stiffened flanges and web (LPFW). The intermediate triangular stiffeners consisted of two 10 mm legs at a 60° angle. It should be noted that the selected parameters were kept within the practical ranges of commercially avialable cross-sections, as listed in Table 1.

The output data was obtained through FSM and ENFM analyses of the CFS elements, and included: (i) the elastic critical buckling load (P_{cr}), (ii) the modal contributions in the critical buckled shape associated with (P_{cr}), expressed through the local ($C_{L,C}$), distortional ($C_{D,C}$) and global ($C_{G,C}$) participation factors [8], (iii) the elastic critical moment (M_{cr}), and (iv) the modal contributions in the critical buckled shapes associated with (M_{cr}), expressed through the local ($C_{L,F}$), distortional ($C_{D,F}$) and global ($C_{G,F}$) participation factors.The present study only considered local, distortional and global instabilities, due to the typically negligible contributions from other modes (i.e. transverse extension and shear modes).

The material yield stress was taken as 350 MPa. Advice was sought from an industrial project partner to ensure the dataset's practicality. The member lengths in the dataset were limited to the practical span lengths of CFS members ($500\,mm \leq L \leq 3000\,mm$), typically resulting in

lower modal contributions from global instabilities. A typical modal decomposition for a compressed lipped channel element is illustrated in Fig. 2.

Table 1 Cross-sectional shapes and ranges of dimensions selected considered in the dataset

LP	LPF	LPW	LPFW
$120 \leq h \leq 360$ $50 \leq b \leq 120$ $10 \leq c \leq 35$ $1 \leq t \leq 4$ $500 \leq L \leq 3000$	$120 \leq h \leq 360$ $50 \leq b \leq 120$ $10 \leq c \leq 35$ $1 \leq t \leq 4$ $500 \leq L \leq 3000$ $0.2 \leq r_1 \leq 0.8$ $d=10$ $\theta=60^0$	$120 \leq h \leq 360$ $50 \leq b \leq 120$ $10 \leq c \leq 35$ $1 \leq t \leq 4$ $500 \leq L \leq 3000$ $0.1 \leq r_2 \leq 0.9$ $d=10$ $\theta=60^0$	$120 \leq h \leq 360$ $50 \leq b \leq 120$ $10 \leq c \leq 35$ $1 \leq t \leq 4$ $500 \leq L \leq 3000$ $0.2 \leq r_1 \leq 0.8$ $0.1 \leq r_2 \leq 0.9$ $d=10$ $\theta=60^0$

function with a bias, and this process continues until the output layer is reached. Information flows unidirectionally in the network, and the calculation process is illustrated in Fig. 2.

The mathematical expression for the ANN procedure can be written as:

$$a_i^j = f^j\left(\sum_k w_{ik}^j a_k^{j-1} + b_i^j\right) \tag{1}$$

where the activations (a values) are determined by the activation function (f) and the weighted sum of the activations from the previous layer, along with the bias term. The weights (w) and biases (b) are parameters of the ANN model, with the subscripts j and i denoting the corresponding layer and node, respectively. In the input layer, each activation value (a_i^1) corresponds to the ith input parameter (x_i), while in the output layer, where $j = n$ (n is the total number of layers in the network), the activation values (a_i^j) represent the predicted output values (y'). The weights and biases are initialized at the beginning of the training process and adjusted using the 'backward propagation of errors' algorithm [18].

After computing the activations (a_i^j) for all network layers, the model determines a cost function $J(y, y')$ based on the original output data (y) and the predicted output values (y') of all training samples. This cost function can take on various forms, such as the Mean Absolute Error (MAE), the Mean Absolute Percentage Error (MAPE), or the Mean Squared Error (MSE). The training process of an ANN aims to identify the weights and biases that minimize the cost function. This study employed the Gradient Descent method [18] to achieve this objective, which involves numerically calculating the following gradients:

$$G_w = \frac{\partial J(w,b)}{\partial w} \tag{2}$$

$$G_b = \frac{\partial J(w,b)}{\partial b} \tag{3}$$

The weights and biases are then updated in each iteration (t) using the following equations:

$$w_t = w_{t-1} - \alpha G_{w,t-1} \tag{4}$$

$$b_t = b_{t-1} - \alpha G_{b,t-1} \tag{5}$$

where α is the chosen learning rate, most often in the range between 0.0 and 1.0.
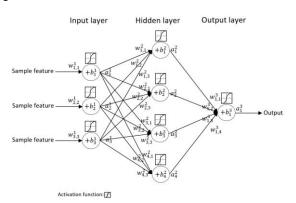

Figure 2 Calculation process of ANN

3.2 Data preparation

The input parameters (x) consisted of seven geometric features of the CFS element (h, b, c, t, r_1, r_2, L), while the output parameters (y) included a single value representing the elastic critical buckling load (P_{cr}) or bending moment (M_{cr}), and a vector containing the participation factors of the critical compressive ($C_{L,C}, C_{D,C}, C_{G,C}$) or flexural ($C_{L,F}, C_{D,F}, C_{G,F}$) buckled shapes. To enhance the performance of the algorithms, the input parameters were preprocessed by standardizing them using the following equation:

$$\hat{x} = \frac{x - \mu}{\sigma} \tag{6}$$

where x represents the original input, while $\hat{x}$ is the standardized input. The mean and standard deviation of x are represented by μ and σ, respectively. Furthermore, it is likely that the output parameters exhibit skewed distributions. To address this issue and achieve a more uniform distribution, a logarithmic transformation was applied to the output data:

$$\hat{y} = \log(1 + y) \tag{7}$$

where $\hat{y}$ and y are the logarithmic and original values of the output data, respectively.

In this study, the performance of the ANNs was evaluated based on the values of the MSE (mean squared error), the MAPE (mean absolute percentage error) and R^2 (coefficient of determination), defined as:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y - y')^2 \tag{8}$$

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y-y'}{y}\right| \tag{9}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y-y')^2}{\sum_{i=1}^{n}(y-\bar{y})^2} \tag{10}$$
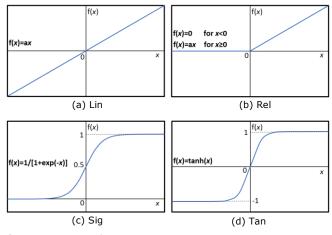
where n is the number of samples and $\bar{y}$ is the mean value of the y-values.

3.3 K-fold cross-validation

The K-fold cross-validation technique was implemented in the ANNs to evaluate the accuracy of the models for unseen data. K-fold cross-validation can be particularly useful for predicting multiple outputs (i.e. the modal decomposition problem). In traditional ANN methods, the available dataset is divided into training, validation, and test sets, which reduces the number of data points available for training and can make the model heavily dependent on sample selection. In K-fold cross-validation, the dataset is randomly shuffled and divided into K folds, with K-1 folds used for training, and one fold used for testing. Each sample is given the opportunity to be used in the test set once and is used to train the model K-1 times. The performance of the model is then taken as the average of all evaluation scores. In this study, the network was subjected to 5-fold cross-validation with an 80%-20% proportion of training and test set data.

3.4 Tuning of hyperparameters

The accuracy of the ANN model depends on its hyperparameters, such as the learning rate, the number of layers and neurons, the activation functions, and the optimizer. The optimal hyperparameters were determined through a trial-and-error process called grid search, which examined the effects of using one to three hidden layers, 10 to 100 neurons, and learning rates of 0.1, 0.2, and 0.3. Four activation functions, including linear (Lin), rectified linear (Rel), log-sigmoid (Sig), and hyperbolic tangent-sigmoid (Tan), were studied, with only the latter two being used in the output layer (Fig. 3). Two cost functions, mean squared error (MSE) and mean absolute percentage error (MAPE), were employed to assess the model accuracy.

Figure 3 Activation functions

4 Results and discussions

Following the tuning of the hyperparameters, the ANN model with the best performance for each prediction target was identified (more information can be found in [19]). Table 2 summarizes these models, including their hyperparameters and performance in terms of Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE). While the elastic critical buckling loads (P_{cr}) and bending moments (M_{cr}) were predicted using a two-layer ANN model, the modal decomposition of the compressive ($C_{L,c}$, $C_{D,c}$, $C_{G,c}$) and flexural ($C_{L,f}$, $C_{D,f}$, $C_{G,f}$) buckled shapes required three hidden layers. The optimal number of neurons for the ANN models predicting the elastic buckling load and the modal decomposition was found to be 40 and 50, respectively. The Tan activation function was used for both the hidden and output layers, in combination with a learning rate of 0.3. The MAPE was chosen as the cost function during the training process of all ANN models.

The ANN models predicting elastic critical buckling loads and bending moments outperformed those predicting the modal decomposition of the compressive and flexural buckled shapes. The former ANN models achieved mean absolute percentage errors (MAPEs) of 2.75% and 2.98%, respectively, compared to MAPEs of 19.93% and 28.05% for the modal decompositions. Linear regression analyses were used to compare the network predictions to the actual responses obtained from the FSM and ENFM, and the statistical indicators such as R^2, COV, and mean are presented in Figs. 4 to 7. Each ANN was independently re-

trained about five times to obtain the most accurate results, and the best performing model was retained. An excellent fit was obtained for the elastic critical buckling loads and bending moments, with $R^2 > 0.99$. However, predictions for the modal decomposition of the compressive and flexural buckled shapes showed a slightly lower level of accuracy, with $R^2 > 0.95$. This is because predicting the modal decomposition is a multi-output problem, while only a single-output ANN model is required for the elastic buckling resistances. The distributions of the modal contribution results were sometimes skewed, which also resulted in less accurate predictions, especially for the global participation factors in bending. It is noted that ANN models are generally less accurate for data outside the ranges of the selected training dataset.

Table 2 Summary of the selected ANN models

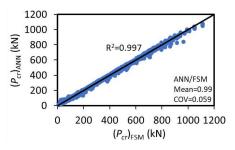| Element type	Target	Hyperparameter features					Performance metrics	
		No. hid. layers	No. neurons	f	α	Cost function	MSE	MAPE (%)
Comp.	P_{cr}	2	40	Tan	0.3	MAPE	1.19 (kN)^2	2.75
	$C_{L,c}$, $C_{D,c}$, $C_{G,c}$	3	50	Tan	0.3	MAPE	23.77	19.93
Flex.	M_{cr}	2	40	Tan	0.3	MAPE	10.15 (kNm)^2	2.98
	$C_{L,f}$, $C_{D,f}$, $C_{G,f}$	3	50	Tan	0.3	MAPE	42.54	28.05

Figure 4 Performance of the selected ANN model for the prediction of the elastic critical buckling load (P_{cr})
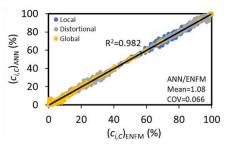

Figure 5 Performance of the selected ANN model for the prediction of the modal contributions in the critical buckled shape of thin-walled compression members ($C_{L,c}$, $C_{D,c}$, $C_{G,c}$)
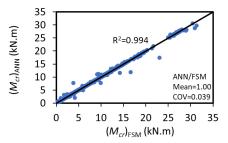

Figure 6 Performance of the selected ANN model for the prediction of the elastic critical buckling moment (M_{cr})
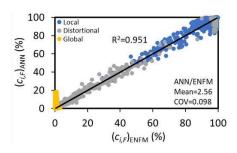
Figure 7 Performance of the selected ANN model for the prediction of the modal contributions in the critical buckled shape of thin-walled flexural members ($C_{L,f}$, $C_{D,f}$, $C_{G,f}$)

5   Conclusions

In this study, ANN models were utilized to predict the elastic critical buckling loads and bending moments of thin-walled structural elements, as well as the contributions of the various pure buckling modes in the critical buckled shapes. It was demonstrated that well-tuned ANN models achieved a high level of accuracy in predicting the elastic buckling loads and the modal decomposition, with coefficients of determination of over 0.99 and 0.95, respectively. This underscores the potential of machine learning as a solution technique for this highly non-linear problem. The slightly lower accuracy of the ANN models in predicting modal decompositions was attributed to the skewed distribution of the global buckling mode participations within the dataset, itself a result of considering only practical member lengths. Besides, it was observed that the prediction of modal contributions, it being a multi-regression problem, is more sensitive to the size and quality of the training dataset in comparison to single-regression problems.

References

[1]   Hasanali, M.; Roy, K.; Mojtabaei, S.M.; Hajirasouliha, I.; Clifton, G.C.; Lim, J.B.P. (2022) A critical review of cold-formed steel seismic resistant systems: Recent developments, challenges and future directions. Thin-Walled Structures 180, 109953.

[2]   Papargyriou, I.; Mojtabaei, S.M.; Hajirasouliha, I.; Becque, J; Pilakoutas, K. (2022) Cold-formed steel beam-to-column bolted connections for seismic applications. Thin-Walled Structures 172, 108876.

[3]   Hasanali, M.; Mojtabaei, S.M.; Clifton, G.C.; Hajirasouliha, I.; Torabian, S.; Lim, J.B.P. (2022) Capacity and design of cold-formed steel warping-restrained beam-column elements. Journal of Constructional Steel Research, 190, 107139.

[4]   CEN, Eurocode 3: design of steel structures, part 1.3: general rules—supplementary rules for cold formed members and sheeting, in, Brussels: European Comittee for Standardization, (2005).

[5]   AISI S100-16, North American specification for the design of cold-formed steel structural members. American Iron and Steel Institute (AISI), Washington, DC, USA, (2016).

[6]   Schafer, B.W.; Peköz, T. (1998) Computational modeling of cold-formed steel: characterizing geometric imperfections and residual stresses. Journal of Constructional Steel Research, 47 193-210.

[7]   Schafer, B.W. (2008) Review: The Direct Strength Method of cold-formed steel member design, Journal of Constructional Steel Research, 64 766-778.

[8]   Becque, J. (2015) A new approach to modal decomposition of buckled shapes, Structures, 4 2-12.

[9]   Xu, Y.; Zhang, M.; Zheng, B. (2021) Design of cold-formed stainless steel circular hollow section columns using machine learning methods, Structures, 33 2755-2770.

[10]   Zarringol, M.; Thai, H.-T.; Naser, M.Z. (2021) Application of machine learning models for designing CFCFST columns, Journal of Constructional Steel Research, 185 106856.

[11]   Ghaisari, J.; Jannesari, H.; Vatani, M. (2012) Artificial neural network predictors for mechanical properties of cold rolling products, Advances in Engineering Software, 45 91-99.

[12]   Brahme, A.; Winning, M.; Raabe, D. (2009) Prediction of cold rolling texture of steels using an Artificial Neural Network, Computational Materials Science, 46 800-804.

[13]   Jamli, M.R.; Farid, N.M. (2019) The sustainability of neural network applications within finite element analysis in sheet metal forming: A review, Measurement, 138 446-460.

[14]   Dias, J.L.R.; Silvestre, N. (2011) A neural network based closed-form solution for the distortional buckling of elliptical tubes, Engineering Structures, 33 2015-2024.

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

[16]   El-Kassas, E.M.A.; Mackie, R.I.; El-Sheikh, A.I. (2021) Using neural networks in cold-formed steel design, Computers & Structures, 79 1687-1696.

[17]   Mitchell, T.M. (2010) Machine Learning, New York : McGraw-Hill.

[18]   Dreyfus, S.E. (1990) Artificial neural networks, back propagation, and the Kelley-Bryson gradient procedure, Journal of Guidance, Control, and Dynamics, 13 926-928.

[19]   Mojtabaei, S.M; Becque, J.; Hajirasouliha, I.; Khandan, R. (2023) Predicting the buckling behaviour of thin-walled structural elements using machine learning methods. Thin-Walled Structures 184, 110518.