



This is a repository copy of *Predicting the buckling behaviour of thin-walled structural elements using machine learning methods*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/196054/>

Version: Published Version

Article:

Mojtabaei, S.M. orcid.org/0000-0002-4876-4857, Becque, J., Hajirasouliha, I. orcid.org/0000-0003-2597-8200 et al. (1 more author) (2023) Predicting the buckling behaviour of thin-walled structural elements using machine learning methods. *Thin-Walled Structures*, 184. 110518. ISSN 0263-8231

<https://doi.org/10.1016/j.tws.2022.110518>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

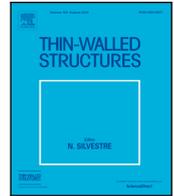
<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



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



Full length article

Predicting the buckling behaviour of thin-walled structural elements using machine learning methods



Seyed Mohammad Mojtabaei ^{a,c,*}, Jurgen Becque ^b, Iman Hajirasouliha ^a, Rasoul Khandan ^c

^a Department of Civil and Structural Engineering, The University of Sheffield, Sheffield S1 3JD, UK

^b Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, UK

^c College of Engineering and Physical Sciences, Aston University, Birmingham B4 7ET, UK

ARTICLE INFO

Keywords:

Thin-walled members
Machine learning
Artificial Neural Network (ANN)
K-fold cross-validation
Buckling resistance
Modal decomposition

ABSTRACT

The design process of thin-walled structural members is highly complex due to the possible occurrence of multiple instabilities. This research therefore aimed to develop machine learning algorithms to predict the buckling behaviour of thin-walled channel elements subjected to axial compression or bending. Feed-forward multi-layer Artificial Neural Networks (ANNs) were trained, in which the input variables comprised the cross-sectional dimensions and thickness, the presence/location of intermediate stiffeners, and the element length. The output data consisted of the elastic critical buckling load or moment, while also providing an immediate modal decomposition of the buckled shape into the traditionally defined 'pure' buckling mode categories (i.e. local, distortional and global buckling). The sample output for training was prepared using a combination of the Finite Strip Method (FSM) and the Equivalent Nodal Force Method (ENFM). The ANN models were subjected to a K-fold cross-validation technique and the hyperparameters were tuned using a grid search technique. The results indicated that the trained algorithms were capable of predicting the elastic critical buckling loads and carrying out the modal decomposition of the critical buckled shapes with an average accuracy (R^2 -value) of 98%. The influence of the various channel parameters on the output was assessed using the SHapley Additive exPlanations (SHAP) method.

1. Introduction

Cold-formed steel (CFS) structural elements, manufactured near room temperature from thin steel plate, possess tangible advantages, such as high strength-to-weight and stiffness-to-weight ratios, ease of handling and transportation, a flexible manufacturing process capable of producing a variety of cross-sectional shapes, and recyclability without loss of quality, which in turn promotes sustainability [1,2]. However, their limited wall thickness results in a heightened susceptibility to instabilities and generates a need to account for these through a design process of increased complexity.

Traditional standard-prescribed design methods are typically based on the effective width approach, pioneered by von Karman [3], in their treatment of cross-sectional instability. The plate elements constituting the cross-section are thereby treated as hinged along their adjoining lines, and interaction between plate elements is conveniently ignored. The Eurocode EN1993-1-3 [4] follows this approach. Additionally, distortional buckling is dealt with using a column buckling model of a relevant cross-sectional subassembly, with the restraint exerted by the remainder of the cross-section represented by a Winkler foundation. The North-American design specifications (AISI-S100) [5] also use the

effective width concept in their fundamental approach, but specify separate strength curves for distortional buckling. These traditional code-specified design rules are quite prescriptive, however, and tend to become cumbersome and tedious when applied to geometrically more advanced cross-sections, which may contain features such as multiple intermediate web/flange stiffeners and complex lip stiffeners. The fact that these design rules rely on traditional distinctions between, for instance, 'flanges' and 'webs' even pre-empts application to some cross-sections with non-traditional, novel and innovative geometries. For these reasons, they are sometimes seen as an impediment to further development, optimization and innovation in the field.

To a large extent because of the above issues, the Direct Strength Method (DSM) has enjoyed a steady rise in popularity as an alternative design method since it was first proposed by Schafer and Peköz in 1998 [6]. The DSM relies on the determination of the individual (elastic) local, distortional and global buckling stresses of the member, and combines this information with the yield stress of the material to define a slenderness value corresponding to each type of instability. Statistically calibrated strength curves are then used to determine the ultimate capacity. The DSM has historically been closely linked to the

* Corresponding author.

E-mail address: smmojtabaei@sheffield.ac.uk (S.M. Mojtabaei).

Finite Strip Method (FSM), where the latter is used as an analysis tool to determine the elastic buckling stresses corresponding to the various modes from the 'signature diagram', which plots the buckling stress against the buckle half-wavelength. This is further explained in Section 2.1. However, a number of difficulties may arise in practical application [7]. An 'indistinct minimum' may be present, meaning that the local mode minimum in the signature curve is obscured by the distortional buckling curve, or vice versa. The minima in the signature curve also more often than not do not correspond to the 'pure' modes, but to coupled instabilities. The distortional minimum in particular usually includes a non-negligible contribution of the local mode. These issues have spurred research into the 'modal decomposition problem' with the aim of determining the buckled shapes and buckling stresses of the pure local, distortional and global modes, as well as their contributions in a randomly deformed shape. To generate the capability of a complete decomposition where all possible deformations within the deformation space are accounted for, two more types of pure modes are typically added: shear modes and (transverse) extension modes. Groundbreaking work in this area was conducted by Adany and Schafer [8], who achieved modal decomposition by importing the mechanical descriptions of the buckling modes established in Generalized Beam Theory (GBT) into the FSM. However, GBT is based on a number of idealized assumptions (e.g. Vlasov's assumptions), which are fundamentally incompatible with the more generally formulated mechanical framework of the FSM, leading to an inelegant solution which lacks full orthogonality between the modes and needs several 'patch-up' solutions. More recent solutions to the modal decomposition problem [9–11] do not suffer from these shortcomings. Among them, the method of the equivalent nodal forces (ENFM) [11] is the most robust and general one. It was therefore employed in this research and a short description is provided in Section 2.2.

The main goal of this paper is to investigate whether Machine Learning can be used to predict the buckling behaviour of CFS members and, additionally, provide a viable and robust solution to the modal decomposition problem. The field of Artificial Intelligence (AI), and specifically Machine Learning and Deep Learning, have seen significant development in recent years and are increasingly finding their way into structural engineering applications. It has previously been demonstrated that machine learning techniques are capable of providing accurate predictions in highly nonlinear problems with large numbers of parameters [12,13]. Machine learning [14] can be defined as a family of methods that evaluate the relationship between input and output parameters by detecting latent patterns in data, and consequently use the uncovered relationships or patterns to predict future data (i.e. supervised learning). Machine learning can be also employed to carry out decision-making under uncertainties (e.g. by reinforcement learning to find the optimal behaviour in an environment). Compared to the rule-based predictive analytics on which conventional structural design is based, machine learning algorithms can be more efficient and powerful tools by automatically extracting relationships and patterns from large-volume high-dimensional data without relying on data engineering and domain knowledge. The following paragraphs aim to give an overview of previous research where these approaches have been specifically employed to CFS.

A first group of studies [15–18] have applied AI to the manufacturing/rolling process of CFS profiles, mostly with the aim of predicting the properties of the finished product. Furthermore, a number of research studies have investigated the design and optimization of CFS elements for different applications, using AI methods. El-Kassas et al. [19] presented an optimization framework using an Artificial Neural Network (ANN) to find the optimum cross-sectional shape for columns. The same researchers also developed neural networks which were trained based on the predictions of the BS-5950 Part 5 [20] design standard to predict the failure load of CFS lipped channel sections [21]. Pala [22] employed an ANN method to estimate the elastic distortional buckling stresses of CFS C-sections under pure compression and pure

bending, and subsequently proposed ANN-based predictive equations. These equations were then used in a follow-up study by Pala and Caglar [23] to investigate the effects of the geometric parameters (including the web height, flange width, flange thickness, and inclination and length of the lips) on the distortional buckling stress. In another study, an ANN algorithm was trained by Guzelbey et al. [24] using an experimental dataset, with the aim of predicting the web crippling strength of CFS trapezoidal decks. The results indicated that the ANN could provide considerably more accurate predictions compared to those obtained from the current design codes.

Due to the complexity of calculating distortional buckling stresses using classical shell theories, Dias and Silvestre [25] trained an ANN on an analytically developed dataset, and consequently presented closed-form expressions to estimate the critical distortional buckling stress and associated half-wavelength of elliptical hollow sections in compression. In another relevant study, Tohidi and Sharifi [26] developed an ANN model to predict the ultimate moment capacities of steel I-beams and demonstrated that the proposed ANN-based formula is more accurate than the existing design codes.

A soft-computing technique using ANN and Genetic Expression Programming (GEP) was developed by D'Aniello et al. [27] to predict the rotational capacity of CFS steel beams with rectangular and square hollow sections. Subsequently, predictive models were proposed and verified against experimental data. Degtyarev [28] trained ANN models to predict the elastic shear buckling load and the ultimate shear strength of CFS channels with slotted webs, and proved that this led to more accurate predictions compared to the code-prescribed design equations. In a follow-up study, Degtyarev and Naser [29] compared the results of five different machine learning boosting algorithms, including gradient boosting regressor (GBR), extreme gradient boosting (XGBoost), light gradient boosting machine learning (LightGBM), gradient boosting with categorical features support (CatBoost), and adaptive boosting (AdaBoost), in predicting the elastic shear buckling loads and the shear strength. It was reported that the CatBoost algorithm was capable of providing the most accurate predictions compared to other boosting algorithms.

Fang et al. [30,31] trained Deep Belief Network (DBN) algorithms, using the results of experimentally validated FE models, to predict the axial compressive capacity of CFS channel sections with and without holes. Additionally, the predictions obtained from the DBN were used to propose enhancement/reduction factors on the axial capacity of such cross-sections. Zarringol et al. [13] and Xu et al. [32] employed machine learning algorithms to predict the ultimate strength of concrete-filled CFS tubular columns and stainless steel tubular columns, respectively, under various loading conditions. More recently, Couto [33] trained an ANN model based on the results of FE simulations to predict the critical buckling moment of tapered beams and compared the ANN-based predictions with those given by the available design guidelines.

Contrary to previous research, this paper focuses on developing ANN algorithms to investigate the elastic stability of CFS structural elements by predicting the elastic critical buckling load and decomposing the associated critical buckled shape into its modal contributions. The training datasets for the elastic critical buckling loads and moments were compiled based on FSM output, while the ANN models to achieve modal decomposition were trained on results obtained from the Equivalent Nodal Force Method (ENFM).

The results of this study will potentially prove useful in creating a practical design tool for engineers and practitioners and provide them with a largely intuition-based insight into the mechanical behaviour of thin-walled elements, without necessarily requiring any background knowledge of programming or complex mechanical concepts. A future extension of the presented research could capitalize on the potential to further train the ANN into a method of near universal versatility by invoking the strengths of other decomposition methods (e.g. the 'polarization method' [10]) in areas where the ENFM lacks applicability (e.g. for cross-sections with rounded corners).

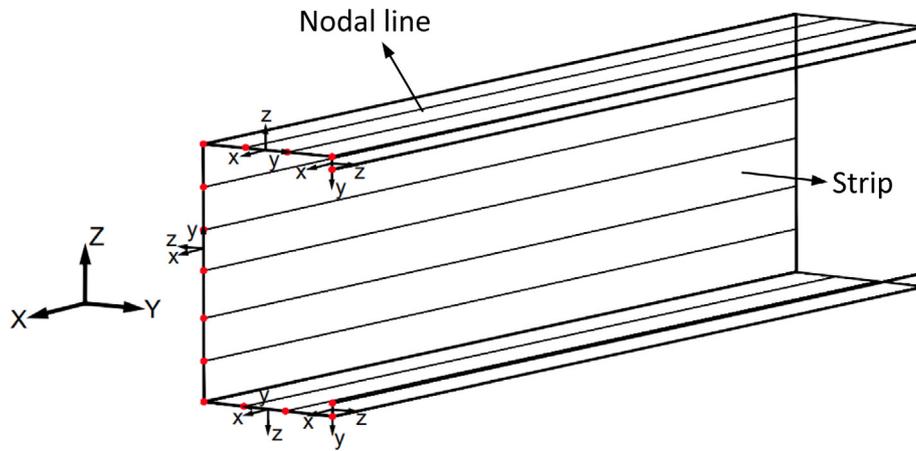


Fig. 1. Discretization of a thin-walled member into strips.

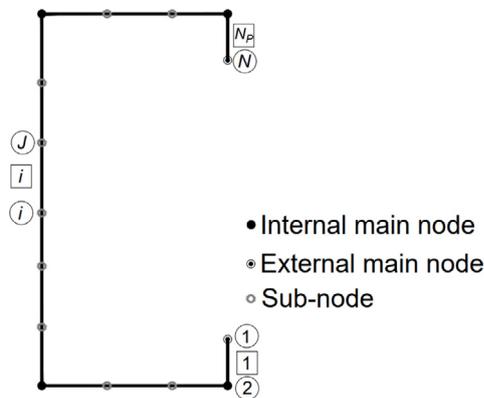


Fig. 2. Main nodes and sub-nodes in a lipped channel section.

2. Theoretical background

This section aims to summarize the basic principles of the FSM and the ENFM.

2.1. The Finite Strip Method (FSM)

In the FSM [34] thin-walled elements are divided into a number of longitudinal strips, the longitudinal boundaries of which are called the ‘nodal lines’ (Fig. 1). In the remainder of this paper, these nodal lines will simply be referred to as ‘nodes’. As also illustrated in Fig. 1, a local (x, y, z) coordinate system associated with each strip was defined, as well as a global (X, Y, Z) coordinate system. The total number of nodes and strips are indicated by N and N_p , respectively. The nodes of a cross-section can be categorized into three groups, as shown in Fig. 2: (i) N_{mi} internal main nodes, which connect adjacent strips with non-aligned local y -axes - the number of main nodes depends on the geometry of the cross-section, (ii) N_{me} external main nodes, which are positioned along the free edges of a cross-section (for a lipped channel section the number of external main nodes $N_{me} = 2$), and (iii) N_s sub-nodes which connect adjacent strips with aligned local y -axes. The number of sub-nodes can be arbitrarily chosen, with a higher number of subnodes typically resulting in more accurate results [34].

As shown in Fig. 3, each nodal line has four degrees of freedom, namely the longitudinal x -displacements (u) measured at the ends of the nodal line, and the in-plane y -displacement (v), the out-of-plane z -displacement (w) and the rotation about the x -axis (θ), measured at

mid-length. While the degrees of freedom (u_i, v_i) and (u_j, v_j) determine the in-plane membrane displacements of a strip located between nodal lines i and j , the deformations caused by plate bending are determined by (w, θ_i) and (w_j, θ_j) . For a strip with pinned boundary conditions, the shape functions in the longitudinal direction are assumed to be sinusoidal for v and w , and co-sinusoidal for u . In the transverse direction, on the other hand, they are linear for u and v , and cubic for w .

The elastic and geometric stiffness matrices of the member, \mathbf{K} and \mathbf{G} , are assembled using the local elastic and geometric stiffness matrices of each strip [34], which leads to the formulation of the following stability eigenvalue problem:

$$(\mathbf{K} - \lambda \mathbf{G}) \cdot \mathbf{v} = 0 \tag{1}$$

where \mathbf{v} is an eigenvector revealing the buckled shape of the element, and λ is the eigenvalue indicating the corresponding elastic buckling stress. The critical buckling stress (σ_{cr}) for a given buckle half-wavelength (L) is obtained as the lowest eigenvalue (λ_{cr}). The plot showing (σ_{cr}) as a function of (L) is usually called the ‘signature curve’ of the member, and an example is provided in Fig. 4.

2.2. The Equivalent Nodal Force Method (ENFM)

The most straightforward and robust way to achieve modal decomposition of buckled shapes is the ‘method of the equivalent forces’ proposed by Becque et al. [11]. In this method, the pure local, distortional and global buckling modes, with the added shear and transverse extension modes, form a full orthogonal set of basis vectors of the complete deformation space. This implies that any deformation can be expressed as a linear combination of these basis vectors. By enforcing orthogonality of the basis vectors, the decomposition becomes mathematically unique and its interpretation therefore unambiguous. The method hinges on the determination of sets of nodal forces which produce the pure buckled shapes in a first-order linear elastic problem.

2.2.1. Local modes

The local buckling modes are characterized by the corners of the cross-section (coinciding with the internal main nodes as defined in Section 2.1) remaining in place during buckling. According to this definition, the local buckling modes can be completely described by a subset of degrees of freedom: (i) the rotations θ of all nodes, and (ii) the out-of-plane displacements w of all nodes except the internal main nodes. Therefore, a set of basis vectors of the local space $\mathbf{v}_{L,i}$ can be constructed by assembling all possible $4N \times 1$ vectors in which all elements are zero, except for a single element corresponding to one of the above-mentioned degrees of freedom, which is set equal to 1. The

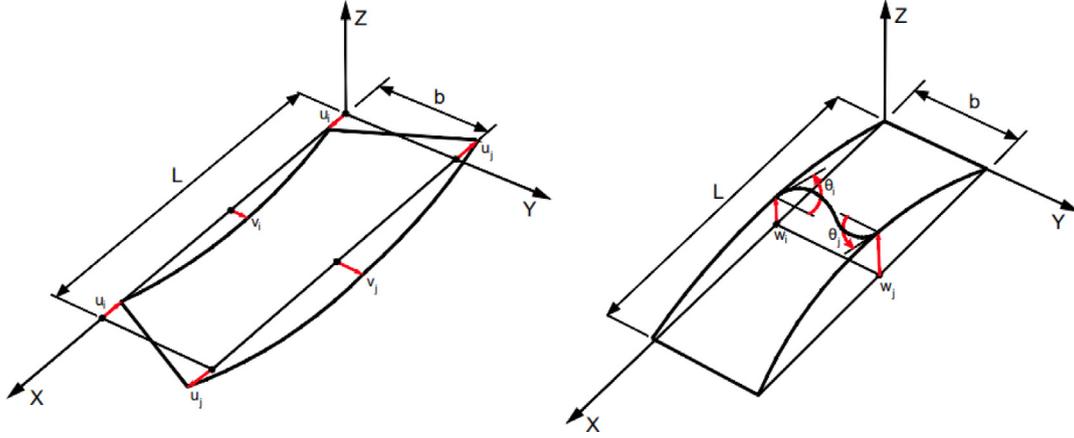


Fig. 3. Degrees of freedom in the FSM.

number of the basis vectors of the local space (equal to the number of local modes) is equal to $2N - N_{mi}$. The $v_{L,i}$ vectors can then be arranged as columns in a matrix H_L with dimensions $4N \times (2N - N_{mi})$, and any local mode shape (d_L) can be expressed as a linear combination of the basis vectors (v_L) [8]:

$$d_L = H_L a \tag{2}$$

In the above equation, a is a vector of unknown coefficients, which can be determined for a given loading by solving the following eigenvalue problem:

$$(H_L^T K H_L - \lambda_L H_L^T G H_L) a = 0 \tag{3}$$

The eigenvalues (λ_L) resulting from Eq. (3) are the local buckling stresses corresponding to the buckled shapes determined from (2) once the (a) vectors are known.

2.2.2. Distortional modes

To uniquely define the distortional modes, four criteria are imposed on the cross-section [11]: (D_I) the nodal forces (f) maintain cross-sectional equilibrium, (D_{II}) the longitudinal forces acting on the member are zero, (D_{III}) the transverse membrane stresses are zero, and (D_{IV}) the distortional modes are orthogonal to the local modes.

The following set of equations is used to ensure cross-sectional equilibrium:

$$\left\{ \begin{array}{l} \sum_{i=1}^N f_Y^i = 0 \\ \sum_{i=1}^N f_Z^i = 0 \\ \sum_{i=1}^N (f_Y^i Z_i - f_Z^i Y_i) = 0 \end{array} \right. \tag{4}$$

where f_Y^i and f_Z^i are the nodal force components in the global Y and Z directions at node i , respectively, and Y_i and Z_i are the global coordinates of node i relative to an arbitrary origin. It is worth noting that the final equation in Eq. (4) contains a typo in the original formulation [25]. The matrix form of Eq. (4) is written as $C_1 f = 0$, where C_1 is a $3 \times 4N$ matrix.

Criterion (D_{II}) can also be expressed in matrix form with the help of a reduced $3N \times 1$ vector (\hat{f}) which contains all nodal forces except the longitudinal end forces:

$$f = \begin{bmatrix} I \\ 0 \end{bmatrix} \hat{f} = T \hat{f} \tag{5}$$

In Eq. (5) I is the $3N \times 3N$ identity matrix and 0 is an $N \times 3N$ matrix containing zeros. Consequently, T is a matrix of size $4N \times 3N$.

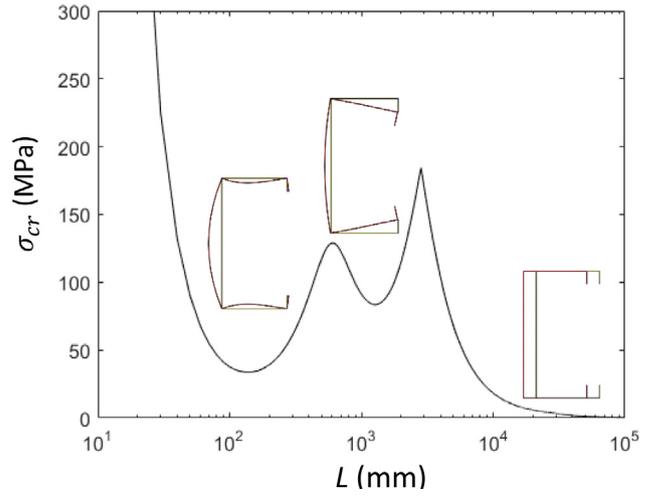


Fig. 4. Typical signature curve.

In the FSM, criterion (D_{III}) can only be imposed in an averaged form over the strip width (b), which leads to the following set of equations [25]:

$$\frac{(V_{1,i} - V_{2,i}) \cos \alpha_i - (W_{1,i} - W_{2,i}) \sin \alpha_i}{b_i} + \frac{(U_{1,i} + U_{2,i})}{2} \left(\frac{\nu \pi}{L} \right) = 0 \tag{6}$$

($i = 1, \dots, N_p$)

where ν is Poisson's ratio and the subscripts '1' and '2' refer to both nodal lines at the boundaries of the strip. Furthermore, α_i is the angle measured from the local y-axis of strip i to the global Y-axis, measured positive in the counterclockwise direction. The coefficients in Eq. (6) can be assembled into a matrix C_σ , while the degrees of freedom are contained in a vector d , allowing Eq. (6) to be translated into:

$$C_\sigma d = C_\sigma K^{-1} f = C_2 f = 0 \tag{7}$$

Finally, the orthogonality criterion (D_{IV}) is expressed mathematically as:

$$H_L^T K d = H_L^T f = 0 \tag{8}$$

The forces generating the distortional modes (F_D) can thus be determined as the null space of the matrix C_D :

$$\widehat{F}_D = null(C_D) = null \begin{bmatrix} C_1 \\ C_2 \\ H_L^T \end{bmatrix} \tag{9}$$

where: $F_D = T \widehat{F}_D$

The matrix C_D has dimensions $(2N - N_{mi} + N_p + 3) \times 3N$. A constrained eigenvalue problem can then be formulated using the matrix $H_D = K^{-1}F_D$, comprising a set of basis vectors of the distortional space:

$$(H_D^T K H_D - \lambda_D H_D^T G H_D) \mathbf{a} = \mathbf{0} \quad (10)$$

This determines the buckling stresses of the distortional modes (λ_D) and the associated modal shapes $\mathbf{d}_D = H_D \mathbf{a}$ under a given loading. It is noted that for an open unbranched cross-section (e.g. a lipped channel) the number of distortional modes is always equal to $N_{mi} - 2$.

2.2.3. Global modes

To define the pure global modes, similar criteria to those maintained for the distortional modes are applied, except that cross-sectional equilibrium of the nodal forces is no longer required: (G_I) no longitudinal forces are necessary to generate the global modes, (G_{II}) the transverse membrane stresses are zero, and (G_{III}) the global modes are orthogonal to both local and distortional modes. Based on this, the forces generating the global modes (F_G) can be determined as the null space of the matrix C_G :

$$\widehat{F}_G = \text{null}(C_G) = \text{null} \begin{pmatrix} C_2 \\ H_L^T \\ H_D^T \end{pmatrix} \quad (11)$$

where: $F_G = T \widehat{F}_G$

Subsequently, the matrix $H_G = K^{-1}F_G$, containing a set of basis vectors of the global space, is employed to formulate a constrained eigenvalue problem:

$$(H_G^T K H_G - \lambda_G H_G^T G H_G) \mathbf{a} = \mathbf{0} \quad (12)$$

This yields the global mode shapes $\mathbf{d}_G = H_G \mathbf{a}$ under a given loading and the corresponding global buckling stresses (λ_G). It is noted that the number of global modes is always three.

2.2.4. Transverse extension modes

The transverse extension modes are determined by imposing that: (TE_I) no longitudinal nodal forces are necessary to generate the transverse extension modes, and (TE_{II}) the transverse extension modes are orthogonal to the local, distortional and global modes. These criteria can be mathematically expressed as:

$$\widehat{F}_{TE} = \text{null}(C_{TE}) = \text{null} \begin{pmatrix} H_L^T \\ H_D^T \\ H_G^T \end{pmatrix} \quad (13)$$

where: $F_{TE} = T \widehat{F}_{TE}$

The matrix $H_{TE} = K^{-1}F_{TE}$, comprising a set of basis vectors of the transverse extension space, can then be used to formulate an eigenvalue problem in a similar fashion to Eqs. (3), (10) and (12) to obtain the transversely extended modal shapes ($\mathbf{d}_{TE} = H_{TE} \mathbf{a}$) under a given loading.

2.2.5. Shear modes

As a final category of buckling modes, the shear modes are determined by their orthogonality with all previously obtained modes:

$$F_S = \text{null}(C_S) = \text{null} \begin{pmatrix} H_L^T \\ H_D^T \\ H_G^T \\ H_{TE}^T \end{pmatrix} \quad (14)$$

The shear modes ($\mathbf{d}_S = H_S \mathbf{a}$) under a given loading can then be obtained by constraining the solutions of Eq. (1) to a linear combination of the basis vectors of the shear space, contained in $H_S = K^{-1}F_S$. The number of shear modes is equal to the total number of nodes N .

2.2.6. Modal contributions

Once the pure modes (\mathbf{d}_i) (i.e. the local (\mathbf{d}_L), distortional (\mathbf{d}_D), global (\mathbf{d}_G), transverse extension (\mathbf{d}_{TE}) and shear modes (\mathbf{d}_S)) have been determined, their participations in a random deformed shape \mathbf{v} (expressed through the local (c_L), distortional (c_D), global (c_G), transverse extension (c_{TE}) and shear mode (c_S) participation factors) can be calculated as:

$$c_i = \sum (0.5 \mathbf{v}^T K \mathbf{d}_i) \quad (15)$$

where the sum is carried out over all the basis vectors of the considered subspace.

A particularly relevant application is the case where \mathbf{v} is a vector belonging to the FSM output obtained from Eq. (1).

3. Dataset and parameter space

A dataset was compiled pertaining to 4608 CFS elements with lipped channel sections. The data cover various lengths, cross-sectional dimensions and thicknesses, and account for the possible presence of intermediate stiffeners in the web and/or flanges. More specifically, four cross-sectional shapes were considered, namely an unstiffened lipped channel, a lipped channel with intermediately stiffened flanges, a lipped channel with a stiffened web, and a lipped channel with stiffened flanges and web, as shown in Table 1. For each cross-sectional shape an identical number of 1170 data points were generated. The input data consisted of seven independent parameters, including six cross-sectional parameters (see Table 1): 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). The remaining parameter was the element length (L). It is noted that the intermediate stiffeners always consisted of two 10 mm legs with an intersecting angle of 60°. The ranges of the selected parameters are given in Table 1. The output data was generated by performing FSM and ENFM analyses on the selected CFS elements, and comprised: (i) the elastic critical buckling load (P_{cr}), (ii) the modal contributions to the critical buckling shape associated with (P_{cr}), expressed through the local ($c_{L,C}$), distortional ($c_{D,C}$) and global ($c_{G,C}$) participation factors, (iii) the elastic critical moment (M_{cr}), and (iv) the modal contributions to the critical buckling shapes associated with (M_{cr}), expressed through the local ($c_{L,F}$), distortional ($c_{D,F}$) and global ($c_{G,F}$) participation factors. It should be noted that the modal decomposition results across the dataset showed negligible contributions from the transverse extension and shear modes (less than 1%). Therefore, only the local, distortional and global instabilities were taken into account in this study.

Fig. 5 shows the histograms of the input and output parameters of the dataset. For illustrative purposes, the distributions of the compressive ($\lambda_{cr,C}$) and flexural ($\lambda_{cr,F}$) cross-sectional slenderness parameters are also shown in Fig. 5. These slenderness parameters are defined as:

$$\lambda_{cr,C} = \sqrt{\frac{P_y}{P_{cr}}} \quad (16)$$

$$\lambda_{cr,F} = \sqrt{\frac{M_y}{M_{cr}}} \quad (17)$$

where P_y and M_y are the compressive yield load and the yield moment, respectively. The yield stress of material was assumed to be 350 MPa. It is noted that the ranges of the input parameters were selected to be representative of commercially available channel sections. Additional advice in this respect was sought from the industrial project partner [35]. For instance, the member lengths in the dataset were kept within the practically encountered span lengths for CFS ($500 \text{ mm} \leq L \leq 3000 \text{ mm}$), which in turn led to lower modal contributions from the global instabilities. This is illustrated by the typical modal decomposition for a compressed lipped channel element shown in Fig. 6.

Table 1
Cross-sectional shapes and dimensions considered in the dataset (in mm).

Unstiffened lipped channel	Lipped channel with stiffened flanges	Lipped channel with stiffened web	Lipped channel with stiffened flanges and web
$120 \leq h \leq 360$ $50 \leq b \leq 120$ $10 \leq c \leq 35$ $1 \leq t \leq 4$ $500 \leq L \leq 3000$	$120 \leq h \leq 360$ $50 \leq b \leq 120$ $10 \leq c \leq 35$ $1 \leq t \leq 4$ $500 \leq L \leq 3000$ $0.2 \leq r_1 \leq 0.8$ $d = 10$ $\theta = 60^\circ$	$120 \leq h \leq 360$ $50 \leq b \leq 120$ $10 \leq c \leq 35$ $1 \leq t \leq 4$ $500 \leq L \leq 3000$ $0.1 \leq r_2 \leq 0.9$ $d = 10$ $\theta = 60^\circ$	$120 \leq h \leq 360$ $50 \leq b \leq 120$ $10 \leq c \leq 35$ $1 \leq t \leq 4$ $500 \leq L \leq 3000$ $0.2 \leq r_1 \leq 0.8$ $0.1 \leq r_2 \leq 0.9$ $d = 10$ $\theta = 60^\circ$

Fig. 7 present the correlations between the input parameters and (i) the elastic critical buckling loads (P_{cr}), and (ii) the corresponding modal contributions ($c_{L,C}$, $c_{D,C}$, $c_{G,C}$). Fig. 8 explores this correlation for the elastic critical buckling moments (M_{cr}) and their corresponding modal contributions ($c_{L,F}$, $c_{D,F}$, $c_{G,F}$). Based on the coefficients of determination (R^2) generally poor correlations are observed between the input and output parameters, especially when it comes to the participation factors. An exception can be found for the plate thickness (t) where considerably higher correlations exist with P_{cr} ($R^2 = 0.853$) and M_{cr} ($R^2 = 0.749$). This poor correlation points to machine learning as a potentially more viable avenue to obtain reliable predictions.

4. Machine learning methods

4.1. Overview of Artificial Neural Networks (ANN)

A particular kind of machine learning, the feedforward multilayer ANN, was employed in this study. The main advantage of the ANN approach is that the training process is carried out on a collection of representative examples without requiring a well-defined process to algorithmically convert the input to the output data. The ANN is inspired by the biological structure of the human brain and has a parallel-distributed architecture with a number of interconnected nodes, commonly referred to as neurons. The neurons are arranged in input, hidden and output layers, and each neuron is connected to all of the neurons in the next layer via weighted connections. In the ANN computational process, the data is first fed into the neural network through the input layer which communicates with the hidden layers. Nodes in the hidden layer combine data from various neurons in the input layer with appropriate weights. In the next stage, these weighted inputs are summed up for each neuron of the hidden layer and then passed to the next layer through an activation function, along with a bias. This process is continued until the last hidden layer is reached, which is linked to the output layer where the outputs are retrieved. It should be noted that the connections between the nodes do not form closed loops and information flows unidirectionally within the network. The calculation process of an ANN is schematically represented in Fig. 9.

The ANN procedure can mathematically be expressed as follows:

$$a_i^j = f^j \left(\sum_k w_{ik}^j a_k^{j-1} + b_i^j \right) \quad (18)$$

In the above equation, the a values are called the activations, and f is the activation function which decides whether a neuron should be activated or not within the network (see Section 4.4). The w and b parameters represent the weights and biases of the ANN model, respectively. The subscripts j and i denote the j^{th} layer and i^{th} node, while k represents the number of nodes in the $(j-1)^{\text{th}}$ layer which are connected to the i^{th} node in the j^{th} layer. In the input layer, each a_i^1 value is essentially equal to the i^{th} input parameter (x_i). In the output layer, where $j = n$ (n is the total number of layers in the network), the calculated activation values (a_i^j) constitute the predicted output values (y'). The weights and biases are initially assumed at the beginning of the training process and then learned by the network using an algorithm based on the “backward propagation of errors” [36].

Following the calculation of the activations (a_i^j) for all layers of the network, a cost function $J(y, y')$ is determined for the model 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) and the Mean Squared Error (MSE). The training process of an ANN is aimed at searching for those values of the weights and biases which minimize the cost function. In this study the Gradient Descent method [37] was used for this purpose, where the following gradients are numerically calculated:

$$G_w = \frac{\partial J(w, b)}{\partial w} \quad (19)$$

$$G_b = \frac{\partial J(w, b)}{\partial b} \quad (20)$$

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

$$w_t = w_{t-1} - \alpha G_{w,t-1} \quad (21)$$

$$b_t = b_{t-1} - \alpha G_{b,t-1} \quad (22)$$

where α is the chosen learning rate, most often in the range between 0.0 and 1.0. The influence of α on the performance of the model

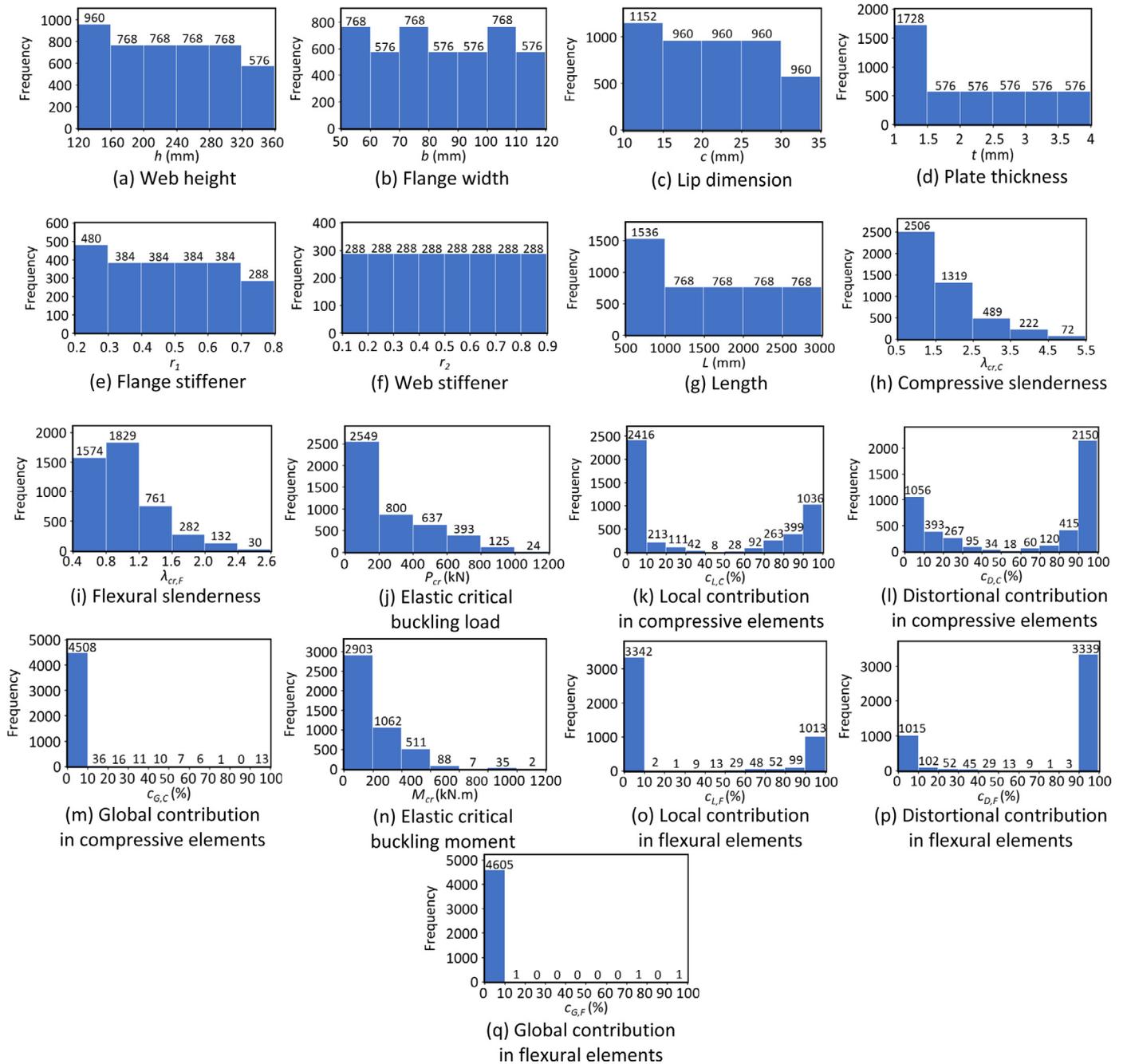


Fig. 5. Histograms of: (a–g) input parameters, (h–i) cross-sectional slenderness and (j–q) output parameters.

was investigated in Section 5.1. It is noted that this ANN process has previously been successfully implemented in several AI-related studies with deep-learning frameworks [30,38,39].

4.2. K-fold cross-validation

K-fold cross-validation is mainly employed in applied machine learning to estimate the accuracy of the model for unseen data. While implementing K-fold cross-validation into the network can be computationally expensive, it can provide significantly more information about the performance of the model and consequently help to increase its accuracy. In addition, the K-fold cross-validation technique can be especially beneficial when the network is required to predict multiple outputs (as is here the case when predicting the modal contributions in the buckled shapes).

In traditional ANN methods the available dataset is divided into training, validation and test sets to avoid instances of overfitting. However, this can noticeably reduce the number of data points available for training and make the model greatly dependent on the selection of the samples into training, validation and test sets. In the K-fold cross-validation approach, on the other hand, the dataset is randomly shuffled and then divided into K folds. K–1 folds are used to train the model and the remaining fold (the test set) is employed for the evaluation. In a permutative approach, each of the K folds is then used, in turn, as the test set, as illustrated in Fig. 10. This implies that each sample is given an opportunity to be used in the test set once and is used to train the model K–1 times. In each iteration, an evaluation score is calculated and retained. The performance of the model is then taken as the average of the evaluation scores. In this study, the network was subjected to 5-fold cross-validation, as shown

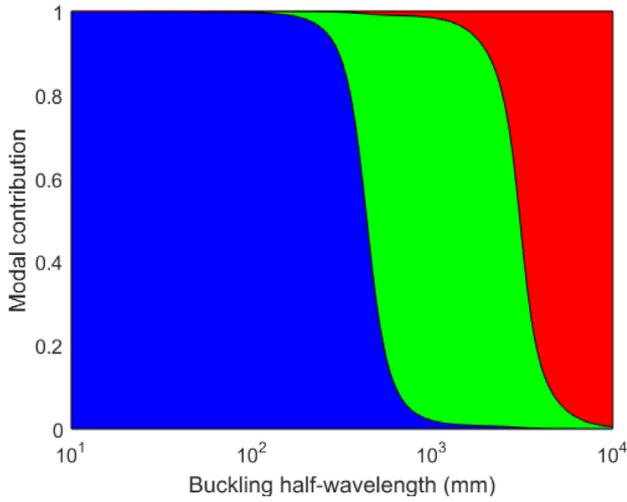


Fig. 6. Typical modal decomposition of FSM output for a compressed lipped channel element.

in Fig. 10. In each iteration the dataset was divided into training and test sets in an 80%–20% proportion.

4.3. Data preparation

4.3.1. Feature standardization and output transformation

As previously discussed in Section 3, the input parameters (x) in this study consisted of the seven geometric features of the CFS element (h, b, c, t, r_1, r_2, L), while the output parameters (y) consisted

of (i) a single value representing the elastic critical buckling load (P_{cr}) or bending moment (M_{cr}), and (ii) a vector containing the three 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 improve the performance of the machine learning algorithms, the dataset was preprocessed. In particular, the input parameters of the dataset were standardized using the following equation:

$$\hat{x} = \frac{x - \mu}{\sigma} \quad (23)$$

where \hat{x} and x are the standardized and original values of the input parameter, and μ and σ denote the mean and the standard deviation of x . In addition, it was observed that the histograms of the output data exhibited skewed distributions. This was especially evident for the modal contributions. Therefore, the logarithmic values of the output data were used to provide a more uniform distribution:

$$\hat{y} = \log(1 + y) \quad (24)$$

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

4.3.2. Performance metrics

Machine learning algorithms can be evaluated using various performance metrics. In this study, the ANNs were 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:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y - y')^2 \quad (25)$$

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y - y'}{y} \right| (\%) \quad (26)$$

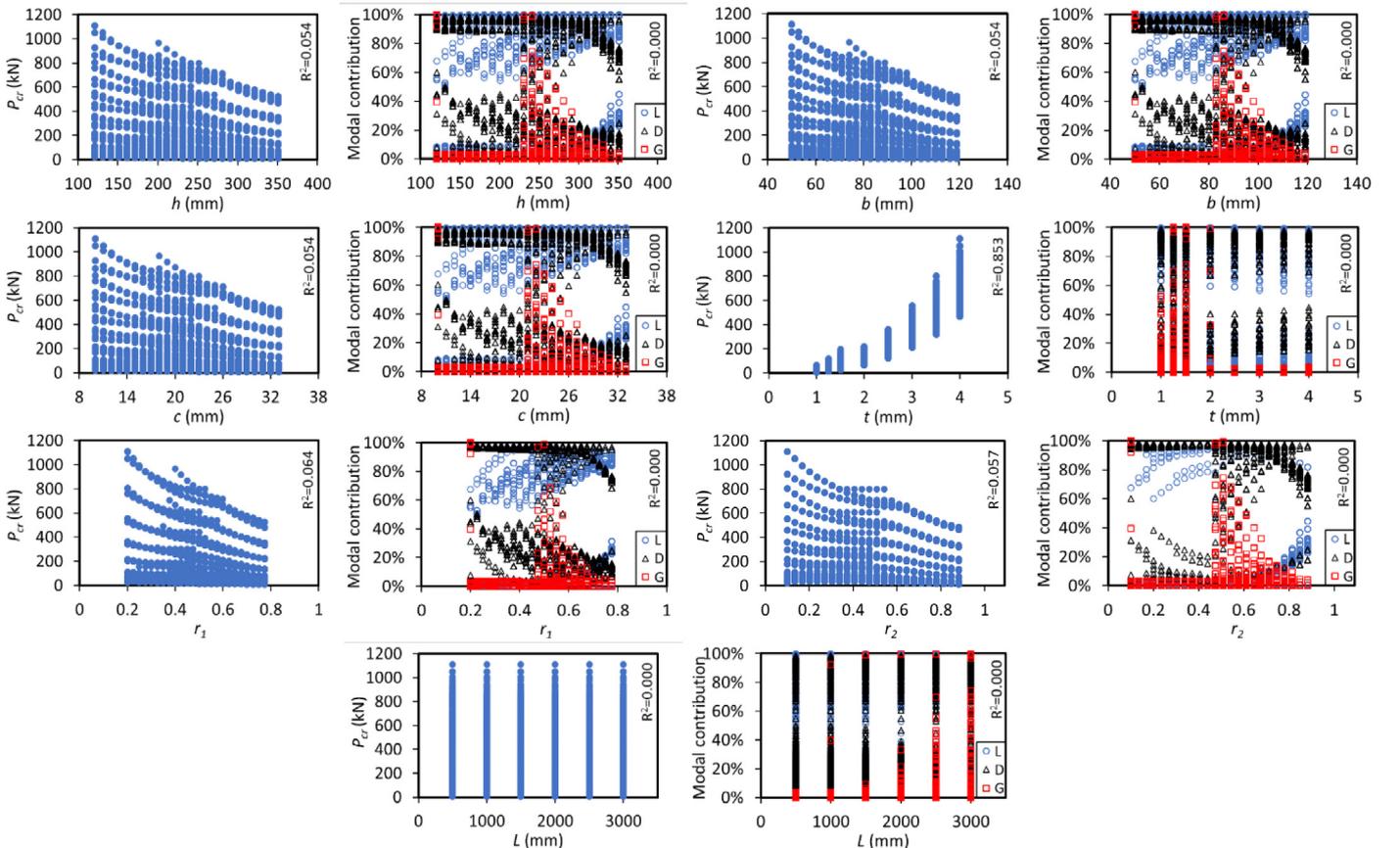


Fig. 7. Correlations between the input parameters and (i) the elastic critical buckling load (P_{cr}), and (ii) the modal contributions for compressive elements (L: Local, D: Distortional and G: Global).

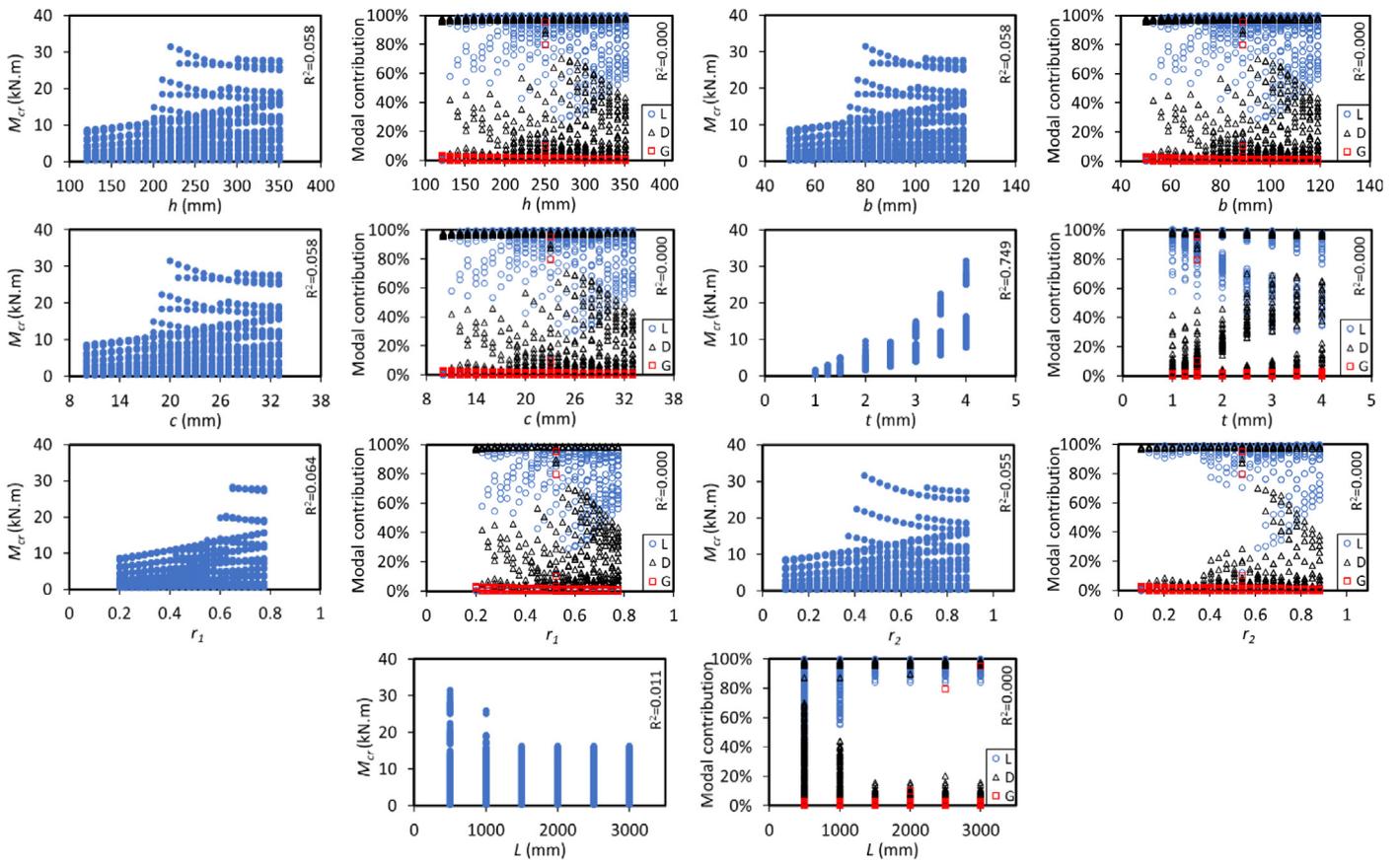


Fig. 8. Correlations between the input parameters and (i) the elastic critical buckling moment (M_{cr}) and (ii) the modal contributions in flexure (L: Local, D: Distortional and G: Global).

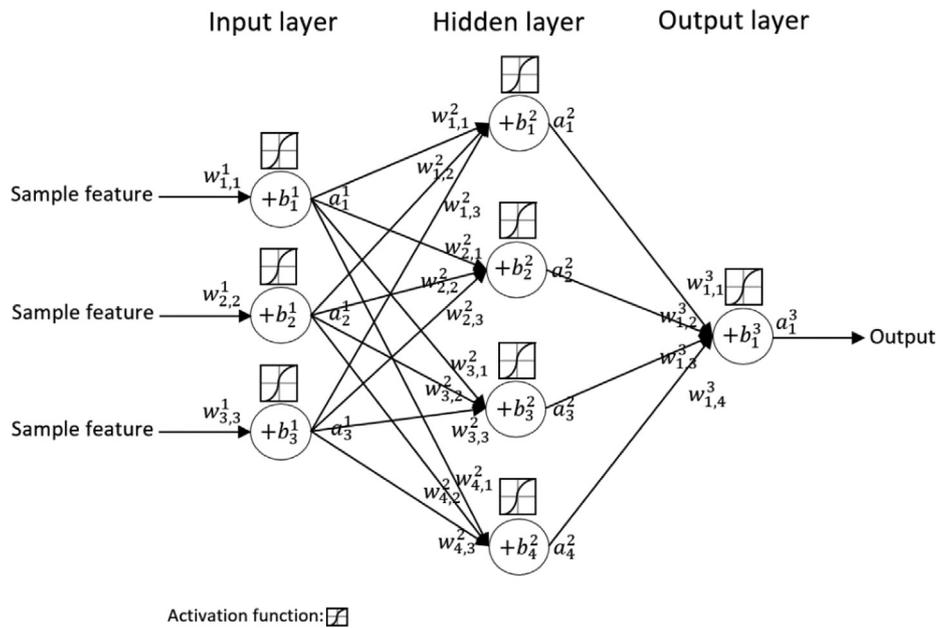


Fig. 9. Calculation process of ANN.

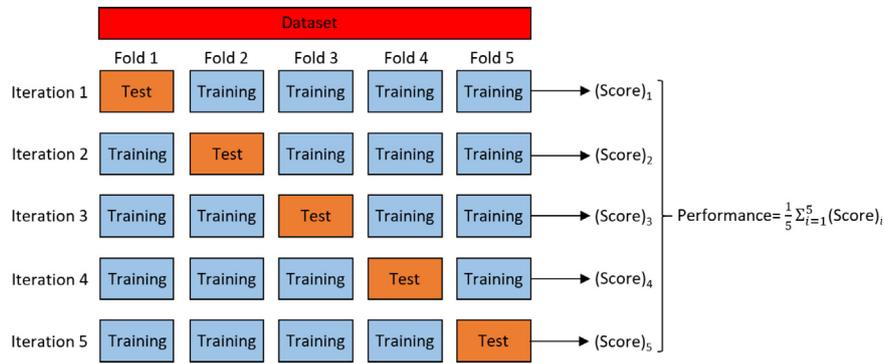


Fig. 10. 5-fold cross-validation.

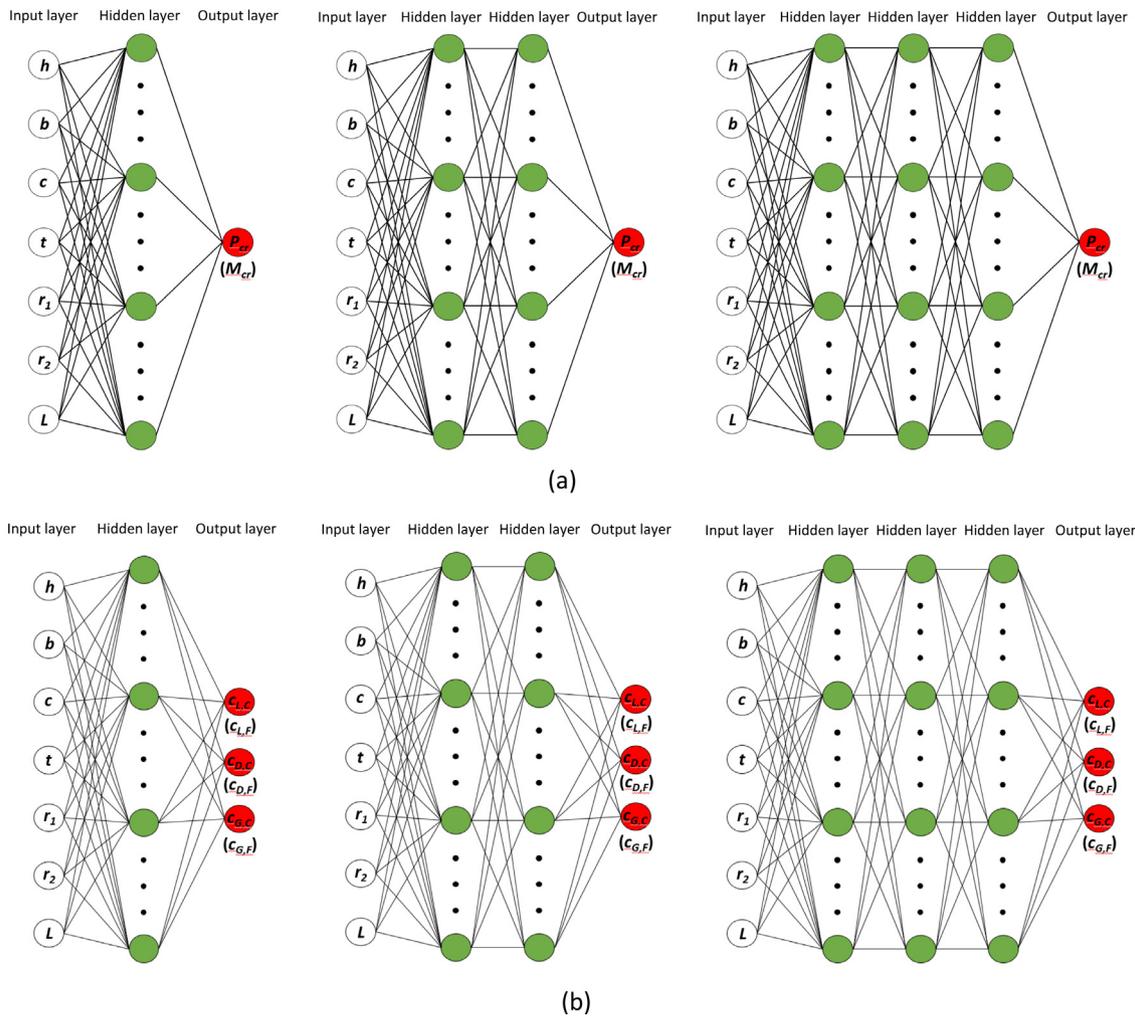


Fig. 11. Architecture of the ANNs with one, two and three hidden layers for the predictions of (a) elastic buckling resistances and (b) modal contributions.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y - y')^2}{\sum_{i=1}^n (y - \bar{y})^2} \quad (27)$$

where n is the number of samples and \bar{y} is the mean value of the y -values.

4.4. Tuning of hyperparameters

The performance and accuracy of the ANN model highly depends on the network parameters (the so-called ‘hyperparameters’) which are

set before training, and include the learning rate, the number of neurons, the number of layers, the activation functions and the optimizer. In this study, the optimal hyperparameters were determined using a well-known hyperparameter tuning method called ‘grid search’ [40], which is based on a trial-and-error process. The following ranges of hyperparameters were studied:

- One, two and three hidden layers were examined, as shown in Fig. 11.
- The number of neurons was varied from 10 to 100 in intervals of 11.

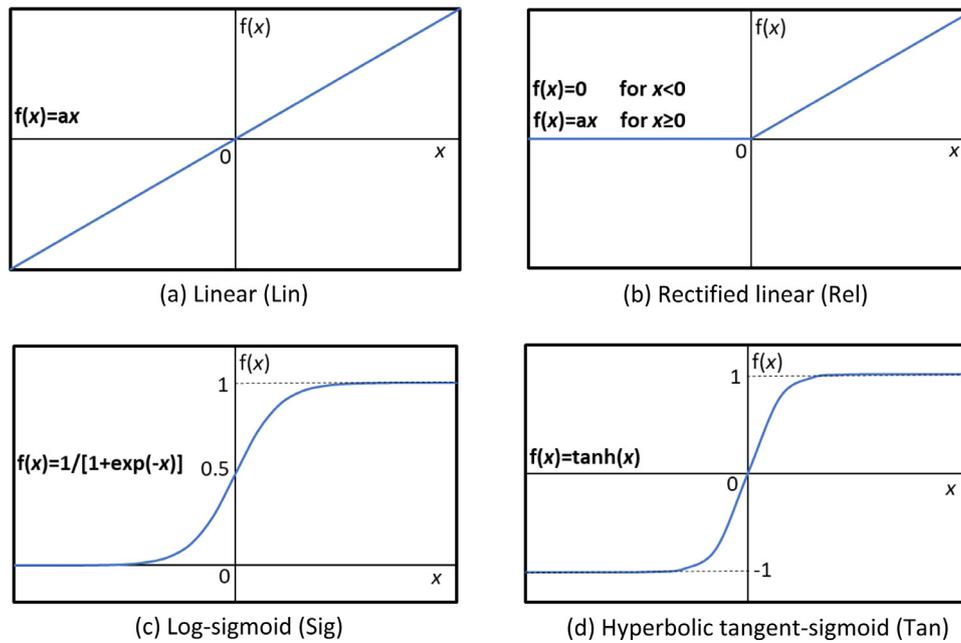


Fig. 12. Activation functions.

- The learning rate was set to 0.1, 0.2 and 0.3.
- A wide variety of activation functions are available for implementation into the ANN. In this study, four different activation functions were examined, including linear (Lin), rectified linear (Rel), log-sigmoid (Sig) and hyperbolic tangent-sigmoid (Tan). These functions are illustrated in Fig. 12. While all four activation functions were examined for the hidden layers, only the Sig and the Tan activation functions were studied in the nodes of the output layer. Using the Sig function can be an appropriate option in terms of computational costs and accuracy where output values should always be positive (Fig. 12). The Tan function, on the other hand, has a steeper slope compared to the other activation functions and therefore results in higher updates in the weights and a potential for faster learning.
- As discussed in Section 4.1, a cost function can be used to represent the performance of the machine learning algorithm. In this study, the MSE and the MAPE cost functions were employed to assess the accuracy of the ANN models.

Following a sensitivity analysis, the number of epochs was set to 1000 in all ANN models. This represents the number of passes of the entire training dataset that the machine learning algorithm completes in order to update the biases and weights. In addition, the default weight and bias initializers were implemented in all ANN models using the “Glorot” and “Zeros” functions, respectively [41].

5. Results and discussions

In this study, the ANN models were developed and tuned in the MATLAB software package [42]. In total, 1440 different ANN models were evaluated for each prediction target in order to tune the hyperparameters. Based on these results the most efficient models with the highest performance were identified. The relative importance of the various input parameters on the predicted output values was also assessed using the Shapley (SHAP) method [43].

5.1. Effects of hyperparameters

In a first step, an ANN model with one hidden layer, using the Sig and Tan activation functions for the hidden and output layers,

respectively, was considered, and the number of neurons and the learning rates were varied in order to study their influence. As shown in Fig. 13, the cost function converged at approximately 40 neurons when the elastic critical buckling load (P_{cr}) or bending moment (M_{cr}) was predicted, while convergence was achieved at about 50 neurons when modal decomposition was the target of the ANN model. A further increase in the number of neurons resulted in negligible changes in the cost functions. It can also be seen from Fig. 13 that the value of the learning rate generally had a negligible effect on the performance of the network. Using either MAPE or MSE as the performance metric for the models also did not result in any substantial difference in the convergence rate or model performance.

Fig. 14 illustrates the influence of the activation functions and the number of hidden layers on the overall performance of ANN models with 40 neurons and a learning rate of 0.3. The labels in Fig. 14 (e.g. ‘Lin-Sig’) indicate the type of activation function used in the hidden layers, followed by the type of activation function in the output layer. It can be seen that the types of activation function significantly affected the performance of the models, as measured by the MSE and MAPE. Using the Sig activation function in the output layer always resulted in ANN models with significant remaining errors. It was also concluded that the ANN models with the Lin activation function within the input layers were incapable of providing accurate solutions. On the other hand, the best performance was achieved when the Tan activation function was utilized in both the hidden and the output layers. The results also demonstrate that training the ANN models with one hidden layer may lead to relatively high error levels compared to those with two and three hidden layers. However, using a larger number of hidden layers considerably increases the computational cost of the training process.

5.2. Selected networks

Following the tuning of the hyperparameters, the ANN model with the best performance was selected for each prediction target. A summary of the selected ANN models is provided in Table 2, including the features of the hyperparameters and the model performance in terms of MSE and MAPE. While the number of hidden layers was set to two for the predictions of the elastic critical buckling loads (P_{cr})

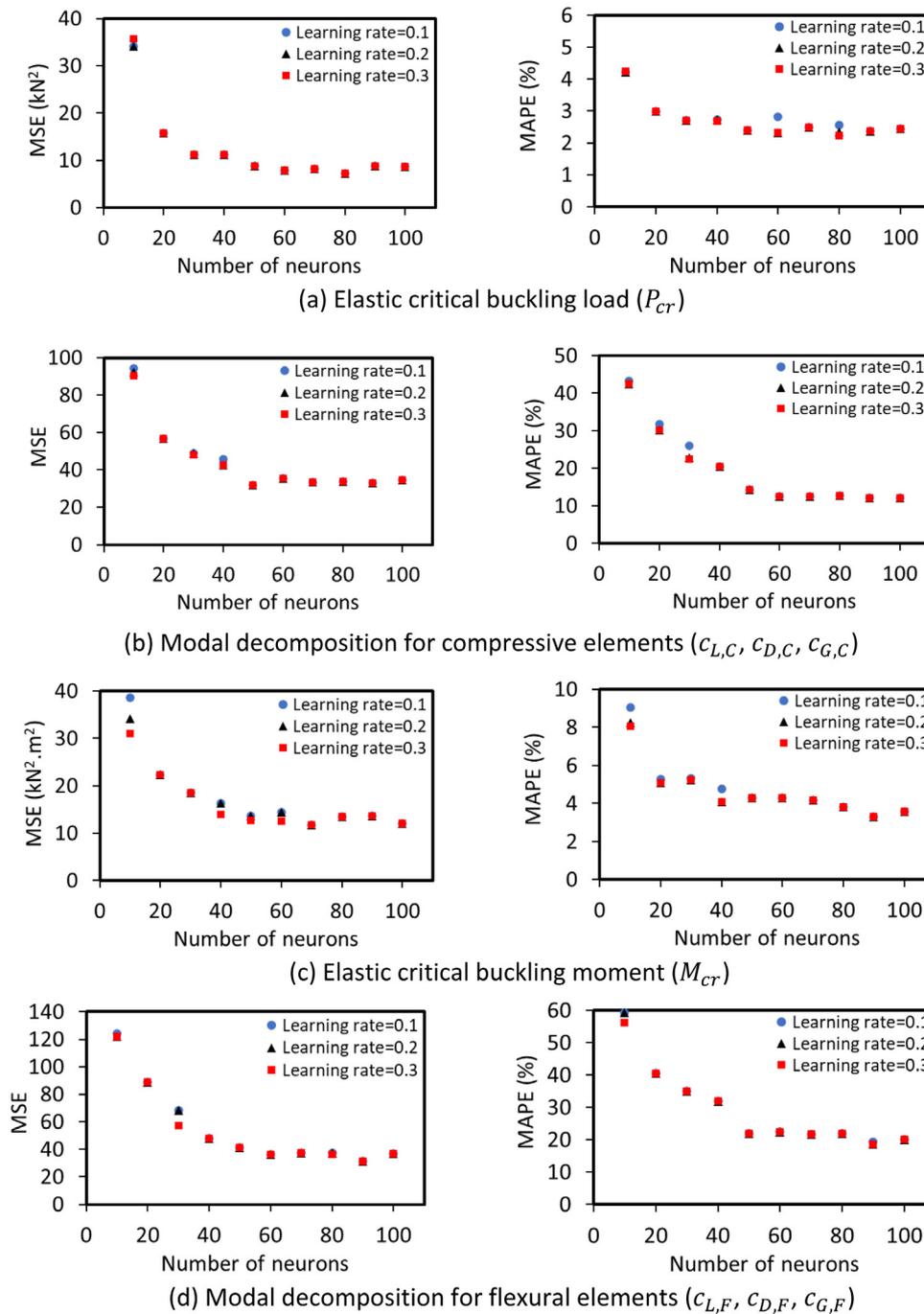


Fig. 13. Influence of number of neurons and learning rate on cost functions.

and bending moments (M_{cr}), the selected ANN models for 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 contained three hidden layers. The optimum number of neurons was found to be 40 and 50 for the ANN models predicting the elastic buckling resistance and the modal decomposition, respectively. For all selected ANN models the learning rate was 0.3, and the Tan activation function was implemented for both the hidden and the output layers. The MAPE was used as the cost function during the training process.

In general, the elastic critical buckling loads and bending moments were more accurately predicted by the ANN models (with MAPEs of 2.75% and 2.98%) than the modal decompositions of the compressive and flexural buckled shapes (with MAPEs of 19.93% and 28.05%,

respectively). Figs. 15 to 18 explore the relationship between the network predictions and the actual responses obtained from the FSM and ENFM by means of a linear regression analysis. The statistical indicators for the ratios of the ANN predictions to the actual responses are also presented in these figures, including the coefficient of determination (R^2), the coefficient of variation (COV) and the mean. For a perfect fit, all data should fall along a 45° line, as the network outputs would be equal to the actual responses. It should be noted that for each prediction problem the model was independently re-trained about 5 times to obtain the most accurate results. This was done because each training starts from different initial weights and biases and this randomized aspect can produce networks of slightly different performance. The best performing model was retained. An

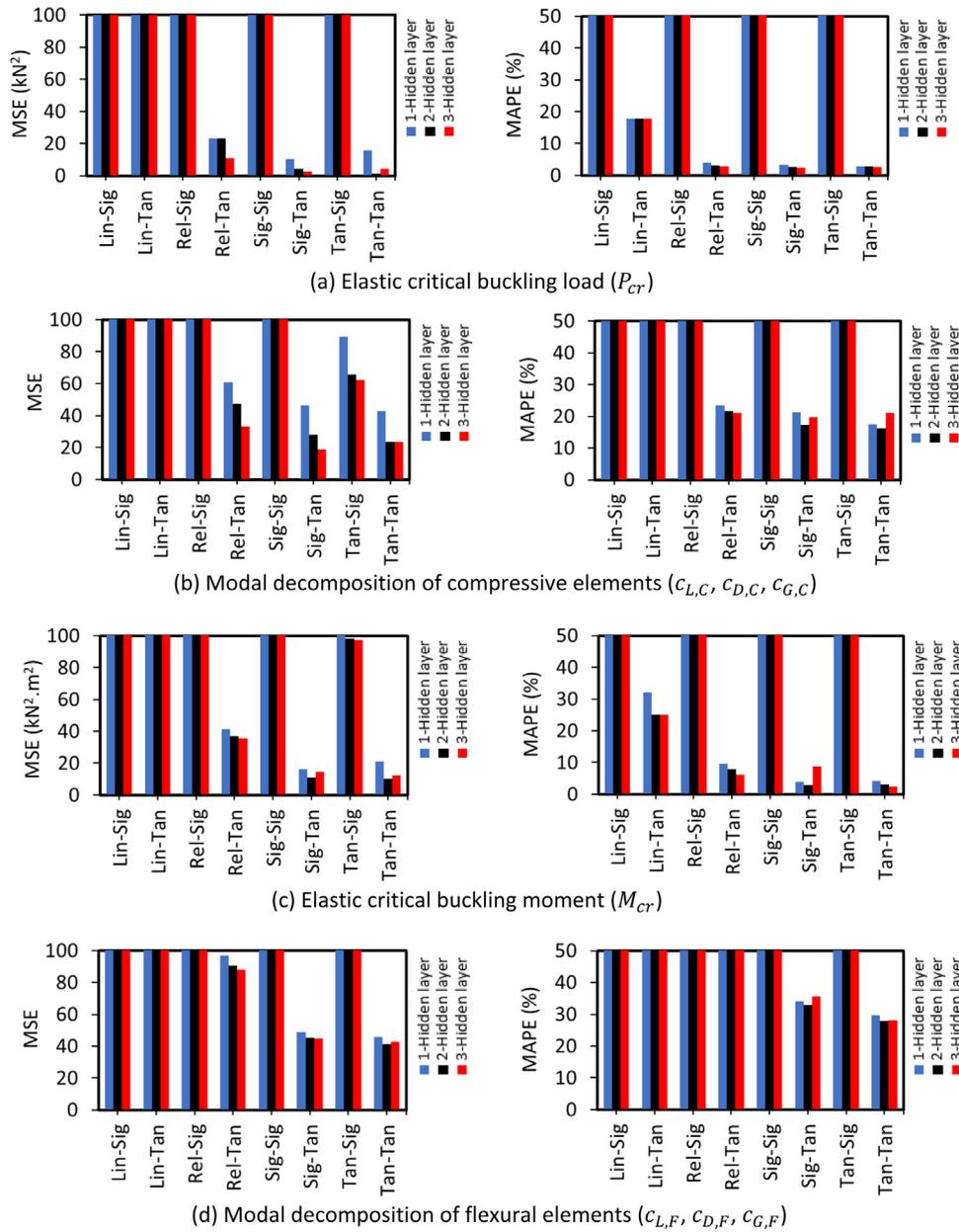


Fig. 14. Influence of activation functions and number of hidden layers on cost functions.

excellent fit was obtained for the elastic critical buckling loads (P_{cr}) and bending moments (M_{cr}), with $R^2 > 0.99$. On the other hand, the predictions for 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 showed a slightly lower level of accuracy, with $R^2 > 0.95$. This is attributed to the fact that predicting the modal decomposition is a multi-output regression problem, while only a single-output ANN model is required for the elastic buckling resistances. In addition, as shown in the histograms of the output data (Fig. 5), the distributions of the modal contribution results were sometimes skewed, which in turn resulted in less accurate predictions. This was especially evident for the contributions of the global (i.e. lateral-torsional) buckling mode in the critical buckled shapes of the flexural elements, which were less than 10% across the whole training dataset (Fig. 5q). This was consciously accepted, since CFS beam elements are typically used in applications where they are laterally supported by floor/roof diaphragms and consequently restrained against lateral-torsional buckling. However, it did result in

slightly inaccurate predictions of the global participation factors in bending (Fig. 18). Finally, it should be noted that ANN models are generally expected to be less accurate for data outside the ranges of the selected training dataset.

5.3. Impact of input parameters on critical buckling resistance and modal decomposition

The Shapley value is based on a solution concept in game theory, which is used to determine the contribution of each player in a coalition or a cooperative game [44]. The Shapley value represents the relative importance or influence of a specific feature (i.e. an input parameter) on the model predictions (F). To compute the Shapley value associated with the i^{th} feature, this feature is first excluded from the feature set (x), and predictions are obtained for all possible subsets S of features that can be formed: $S \subseteq x \setminus \{i\}$. If we indicate the prediction of a model with a subset S of features by F_S , and the prediction of another model

Table 2
Summary of the selected ANN models for buckling behaviour of CFS elements.

Element type	Target	Hyperparameter features					Performance metrics	
		No. hidden layers	No. neurons	Activation function	Learning rate	Cost function	MSE	MAPE (%)
Compressive	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
Flexural	M_{cr}	2	40	Tan	0.3	MAPE	10.15 (kN m) ²	2.98
	$C_{L,f}$, $C_{D,f}$, $C_{G,f}$	3	50	Tan	0.3	MAPE	42.54	28.05

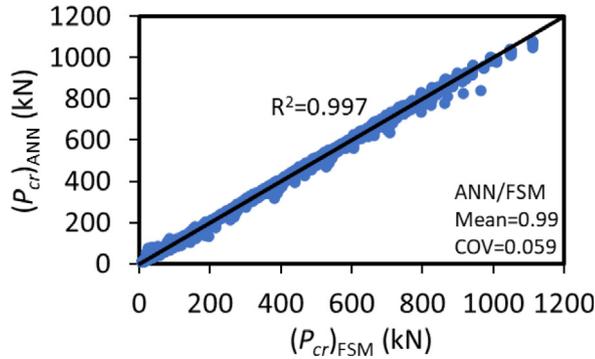


Fig. 15. Performance of the selected ANN model for the prediction of the elastic critical buckling load (P_{cr}).

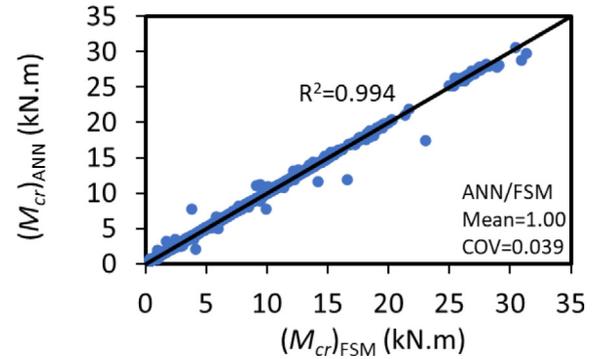


Fig. 17. Performance of the selected ANN model for the prediction of the elastic critical buckling moment (M_{cr}).

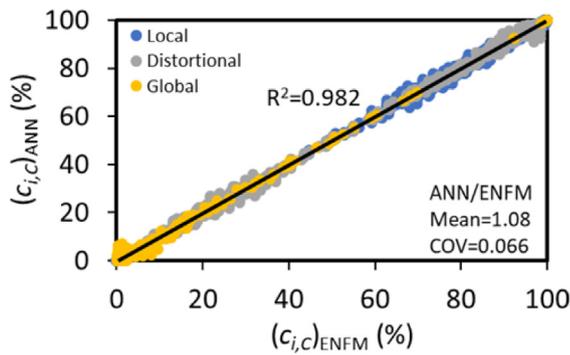


Fig. 16. Performance of the selected ANN model for the prediction of the modal contributions in the critical buckled shape of thin-walled compressive members ($c_{L,c}$, $c_{D,c}$, $c_{G,c}$).

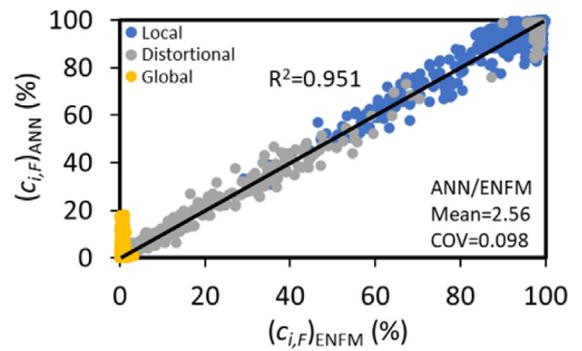


Fig. 18. Performance of the selected ANN model for the prediction of the modal contributions in the critical buckled shape of thin-walled flexural members ($c_{L,f}$, $c_{D,f}$, $c_{G,f}$).

with the i^{th} feature added to this subset as $F_{S \cup \{i\}}$, then the marginal contribution of the i^{th} feature can be quantified as: $[F_{S \cup \{i\}}(x_{S \cup \{i\}}) - F_S(x_S)]$ (where x_S represents the input features in the subset S , where the i^{th} feature is absent). The Shapley value is then computed using a weighted average over all possibilities:

$$\phi_i = \sum_{S \subseteq X \setminus \{i\}} \frac{|S|!(|X| - |S| - 1)!}{|X|!} [F_{S \cup \{i\}}(x_{S \cup \{i\}}) - F_S(x_S)] \quad (28)$$

In the above equation, $|S|$ and $|X|$ represent the number of elements in the subset S and the total number of features, respectively. The procedure is supported in MATLAB by the 'shapley' function.

Fig. 19 shows the mean absolute Shapley values calculated over the whole data set. It can be seen that the element thickness (t) always has the highest influence on the buckling behaviour of the CFS members compared to other input parameters. Other influential input parameters were the cross-sectional dimensions (h , b , c), which had an important effect on the elastic critical buckling load/moment and the local and distortional modal contributions, while the length of the element (L) logically had the second-highest impact on the contributions of the global modes.

6. Summary and conclusions

Machine learning algorithms were developed using Artificial Neural Networks (ANN) 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 shape. The selected dataset consisted of 4608 samples of thin-walled channels, considering various cross-sectional geometries, locations of intermediate stiffeners and element lengths as the input data. The output data for training, i.e. the elastic critical buckling loads and the modal contributions in the buckled shapes, were prepared based on the results of the Finite Strip Method (FSM) and the Equivalent Nodal Force Method (ENFM). The study showed that the tuning of the hyperparameters significantly affects the performance of the ANN models. However, properly tuned ANN models for the prediction of the elastic buckling loads and the modal decomposition demonstrated a high level of accuracy with coefficients of determination of over 0.99 and 0.95, respectively, revealing machine learning as a viable solution technique for this highly non-linear problem. The slightly lower accuracy of the ANN models in predicting modal decompositions was partially attributed to the skewed distribution of the global buckling

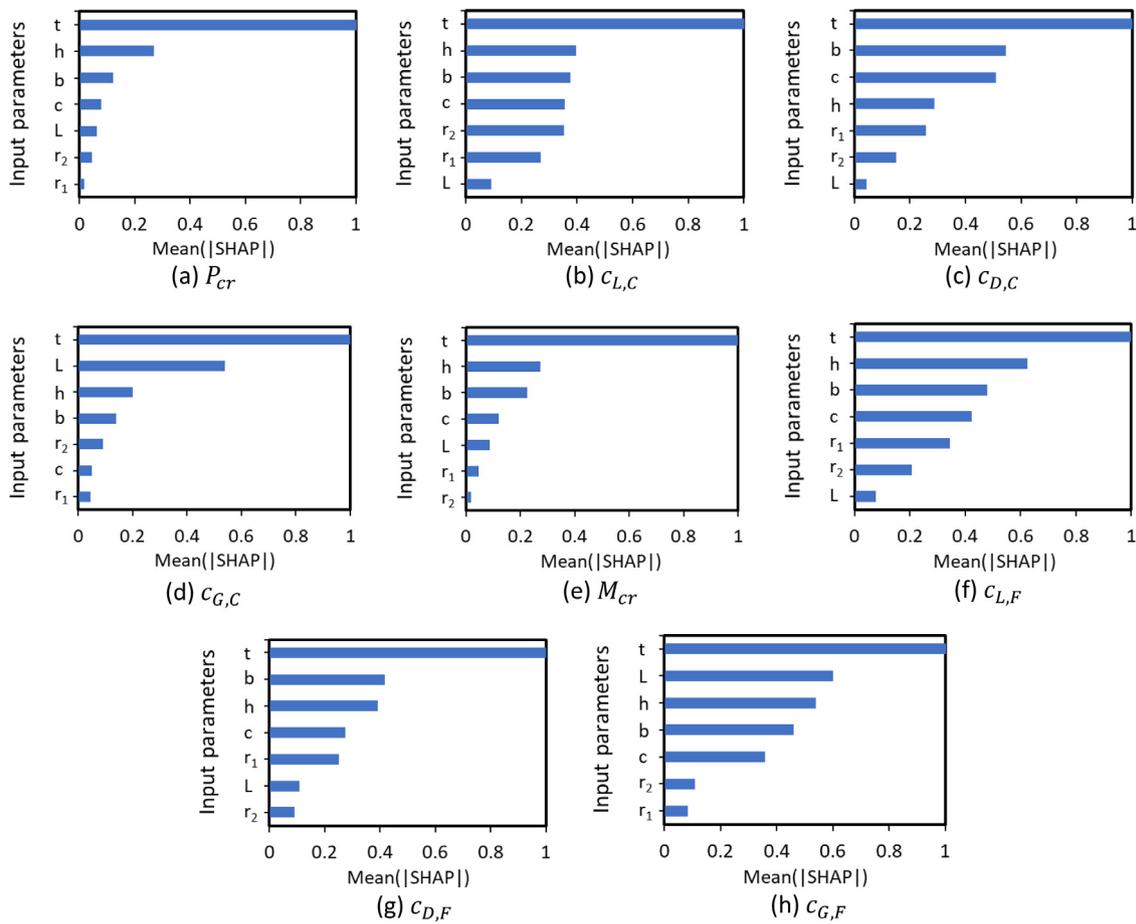


Fig. 19. Importance of the input parameters in the prediction of output values.

mode participations within the dataset, which itself was a consequence of only considering practical lengths. In addition, it was shown that the prediction of modal contributions, which is a multi-regression problem, is more sensitive to the size and quality of the training dataset compared to single-regression problems. The above stated accuracies can be expected within the practical ranges of geometric parameters listed in Table 1.

CRedit authorship contribution statement

Seyed Mohammad Mojtabaei: Writing – original draft, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jurgen Becque:** Writing – review & editing, Validation, Methodology, Conceptualization. **Iman Hajirasouliha:** Validation, Supervision, Resources, Project administration, Conceptualization. **Rasoul Khandan:** Supervision, Resources, Project administration, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors would like to thank the EPSRC for their financial support for this work via the EPSRC Fellowship Award.

References

- [1] D.T. Phan, S.M. Mojtabaei, I. Hajirasouliha, J. Ye, J.B.P. Lim, Coupled element and structural level optimisation framework for cold-formed steel frames, *J. Construct. Steel Res.* (2019) 105867.
- [2] S.M. Mojtabaei, J. Becque, I. Hajirasouliha, Local buckling in cold-formed steel moment-resisting bolted connections: Behavior, capacity, and design, *J. Struct. Eng.* 146 (2020) 04020167.
- [3] T. von Karman, E.E. Sechler, L.H. Donnell, The strength of thin plates in compression, *ASME Appl. Mech. Trans.* 54 (1932) 53–70.
- [4] CEN, Eurocode 3: Design of Steel Structures, Part 1.3: General Rules—Supplementary Rules for Cold Formed Members and Sheeting, European Committee for Standardization, Brussels, 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] B.W. Schafer, T. Peköz, Computational modeling of cold-formed steel: characterizing geometric imperfections and residual stresses, *J. Construct. Steel Res.* 47 (1998) 193–210.
- [7] B.W. Schafer, Review: The Direct Strength Method of cold-formed steel member design, *J. Construct. Steel Res.* 64 (2008) 766–778.
- [8] S. Ádány, B.W. Schafer, Buckling mode decomposition of single-branched open cross-section members via finite strip method: Derivation, *Thin-Walled Struct.* 44 (2006) 563–584.
- [9] V. Karakonstantis, J. Becque, Modal classification and decomposition of instabilities in thin-walled structural members, in: J.K. Paik, S.J. Kim, C.H. Song (Eds.), Proceedings of the 7th International Conference on Thin-Walled Structures, The Korea Ship and Offshore Research Institute, Pusan National University, South-Korea, 2014.
- [10] J. Becque, X. Li, A new approach to modal decomposition of buckled shapes, *Structures* 4 (2015) 2–12.
- [11] J. Becque, X. Li, B. Davison, Modal decomposition of coupled instabilities: The method of the equivalent nodal forces, *Thin-Walled Struct.* 143 (2019) 106229.
- [12] Y. Xu, M. Zhang, B. Zheng, Design of cold-formed stainless steel circular hollow section columns using machine learning methods, *Structures* 33 (2021) 2755–2770.

- [13] M. Zarringol, H.-T. Thai, M.Z. Naser, Application of machine learning models for designing CFCFST columns, *J. Construct. Steel Res.* 185 (2021) 106856.
- [14] T.M. Mitchell, *Machine Learning*, McGraw-Hill, New York, 2010.
- [15] J. Ghaisari, H. Jannesari, M. Vatani, Artificial neural network predictors for mechanical properties of cold rolling products, *Adv. Eng. Softw.* 45 (2012) 91–99.
- [16] A. Brahme, M. Winning, D. Raabe, Prediction of cold rolling texture of steels using an Artificial Neural Network, *Comput. Mater. Sci.* 46 (2009) 800–804.
- [17] M.R. Jamli, N.M. Farid, The sustainability of neural network applications within finite element analysis in sheet metal forming: A review, *Measurement* 138 (2019) 446–460.
- [18] Y. Wang, C. Li, L. Peng, R. An, X. Jin, Application of convolutional neural networks for prediction of strip flatness in tandem cold rolling process, *J. Manuf. Process.* 68 (2021) 512–522.
- [19] E.M.A. El-Kassas, R.I. Mackie, A.I. El-Sheikh, Using neural networks in cold-formed steel design, *Comput. Struct.* 79 (2001) 1687–1696.
- [20] BS5950, British Standards Institution, *Structural Use of Steelwork in Building, Part 5: Code of Practice for Design of Cold-Formed Sections*, BSI, London, UK, 1987.
- [21] E.M.A. El-Kassas, R.I. Mackie, A.I. El-Sheikh, Using neural networks to predict the design load of cold-formed steel compression members, *Adv. Eng. Softw.* 33 (2002) 713–719.
- [22] M. Pala, A new formulation for distortional buckling stress in cold-formed steel members, *J. Construct. Steel Res.* 62 (2006) 716–722.
- [23] M. Pala, N. Caglar, A parametric study for distortional buckling stress on cold-formed steel using a neural network, *J. Construct. Steel Res.* 63 (2007) 686–691.
- [24] I.H. Guzelbey, A. Cevik, A. Erklig, Prediction of web crippling strength of cold-formed steel sheetings using neural networks, *J. Construct. Steel Res.* 62 (2006) 962–973.
- [25] J.L.R. Dias, N. Silvestre, A neural network based closed-form solution for the distortional buckling of elliptical tubes, *Eng. Struct.* 33 (2011) 2015–2024.
- [26] S. Tohidi, Y. Sharifi, Neural networks for inelastic distortional buckling capacity assessment of steel I-beams, *Thin-Walled Struct.* 94 (2015) 359–371.
- [27] M. D’Aniello, E.M. Güneyisi, R. Landolfo, K. Mermerdaş, Analytical prediction of available rotation capacity of cold-formed rectangular and square hollow section beams, *Thin-Walled Struct.* 77 (2014) 141–152.
- [28] V.V. Degtyarev, Neural networks for predicting shear strength of CFS channels with slotted webs, *J. Construct. Steel Res.* 177 (2021) 106443.
- [29] V.V. Degtyarev, M.Z. Naser, Boosting machines for predicting shear strength of CFS channels with staggered web perforations, *Structures* 34 (2021) 3391–3403.
- [30] Z. Fang, K. Roy, J. Mares, C.-W. Sham, B. Chen, J.B.P. Lim, Deep learning-based axial capacity prediction for cold-formed steel channel sections using Deep Belief Network, *Structures* 33 (2021) 2792–2802.
- [31] Z. Fang, K. Roy, B. Chen, C.-W. Sham, I. Hajirasouliha, J.B.P. Lim, Deep learning-based procedure for structural design of cold-formed steel channel sections with edge-stiffened and un-stiffened holes under axial compression, *Thin-Walled Struct.* 166 (2021) 108076.
- [32] Y. Xu, B. Zheng, M. Zhang, Capacity prediction of cold-formed stainless steel tubular columns using machine learning methods, *J. Construct. Steel Res.* 182 (2021) 106682.
- [33] C. Couto, Neural network models for the critical bending moment of uniform and tapered beams, *Structures* 41 (2022) 1746–1762.
- [34] Y.K. Cheung, L.G. Tham, *Finite Strip Method*, CRC Press, Boca Raton, USA, 1998.
- [35] BW Industries, Carnaby Industrial Estate, Lancaster Rd, Carnaby, Bridlington, UK, YO15 3QY. <http://www.bw-industries.co.uk>.
- [36] S.E. Dreyfus, Artificial neural networks, back propagation, and the Kelley-Bryson gradient procedure, *J. Guid. Control Dyn.* 13 (1990) 926–928.
- [37] S. Ruder, An overview of gradient descent optimization algorithms, 2017, arXiv:1609.04747.
- [38] A. Parvat, J. Chavan, S. Kadam, S. Dev, V. Pathak, A survey of deep-learning frame-works, in: *International Conference on Inventive Systems and Control, ICISC, 2017*, pp. 1–7, arXiv:1609.04747.
- [39] Z. Fang, K. Roy, Q. Ma, A. Uzzaman, J.B.P. Lim, Application of deep learning method in web crippling strength prediction of cold-formed stainless steel channel sections under end-two-flange loading, *Structures* 33 (2021) 2903–2942.
- [40] P. Liashchynskiy, P. Liashchynskiy, Grid search, random search, genetic algorithm: a big comparison for NAS, 2019, arXiv:1912.06059.
- [41] X. Glorot, Y. Bengio, Understanding the difficulty of training deep feedforward neural networks, in: *Proceedings of the 13th International Conference on Artificial Intelligence and Statistics (AISTATS)*, Vol. 9, Sardinia, Italy, 2010, JMLR: W & CP 9.
- [42] Mathworks, *Matlab R2021b*, Mathworks Inc, 2021.
- [43] S.M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, *Adv. Neural Inf. Process. Syst.* (2017) 4765–4774.
- [44] E. Štrumbelj, I. Kononenko, Explaining prediction models and individual predictions with feature contributions, *Knowl. Inf. Syst.* 41 (2014) 647–665.