



Neural network models for the critical bending moment of uniform and tapered beams

Carlos Couto *

RISCO – Department of Civil Engineering, University of Aveiro, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal

ARTICLE INFO

Keywords:
 Neural network
 Critical moment
 Machine learning
 Steel beams
 Tapered members
 Mono-symmetric sections

ABSTRACT

Most design standards require the calculation of the elastic critical bending moment for the design and verification of steel beams. Formulae exists for uniform beams with double or mono-symmetric cross-sections but, for tapered beams, simple design formulae are yet to be developed because of the complexity associated with the non-uniform geometry of these members.

This work proposes a neural network model to calculate the critical moment. The model is developed using the Backpropagation and Levenberg–Marquardt algorithms and considering 60549 data samples for training and 8526 samples for validation. The samples are generated by a numerical model using the Finite Element Method (FEM). An innovative methodology to reduce the number of samples necessary to train the model is implemented based on the concept of random sample generation with constrained geometric proportions.

The accuracy of the developed model is further illustrated on some particular cases and against the FEM results of other authors. For the uniform beams, the results of the proposed model are compared against those from existing formulae for uniform members showing its improved accuracy.

Finally, it is expected that this investigation demonstrates the benefits of the use of neural networks based solutions as fast assessment tools during the search of optimal structural solutions in the early design stages.

1. Introduction

The critical bending moment of beams is a limit value for which the deformation remains in the plane of the loads since for higher values, the beam suffers a lateral deflection accompanied by the twisting of the cross-section. Because of such behaviour, this limit state is named as lateral-torsional buckling and is included in steel design standards to calculate the ultimate strength of the beams. In addition to this buckling phenomena, other limit states such as yielding or local buckling in the section, web crippling or failure due to shear stresses may occur simultaneously and influence the lateral-torsional buckling resistance, limiting the beam capacity to carry loads. Underestimation of these limit states or insufficient measures to avoid them, such as the inclusion of bracing systems, may cause partial or total collapse of the structures. Fig. 1 shows a structural collapse of an industrial building made of tapered members where it may be observed some of these failure modes.

The critical moment can be calculated either using analytical formulae, when available, or alternatively using time-consuming Finite Element Method (FEM) models. According to Bleich [1], the subject of lateral-torsional buckling was firstly reported on [2,3] for uniform beams. These studies were then followed by important contributions of

other authors, namely Timoshenko [4], Vlasov [5], Winter [6], Hill [7], Clark and Hill [8] and Galambos [9] (*vide* [10]). As the literature on the subject is extensive following these developments, the reader may also further refer to Baláž and Koleková [11], Koleková and Baláž [12], Serna et al. [13] and Braham [14] for a more recent overview. In Baláž and Koleková [11], it is reported that a formula to calculate the elastic critical moment of uniform beams was firstly proposed by Mrázik and Djubek [15] in 1958 and later modified by Djalaly [16] in 1974. This formula was given in the draft version of Eurocode 3 [17], for uniform beams with double and mono-symmetric section. However, this formula led to some inconsistent results, particularly for the case of beams with mono-symmetric sections (see for example Koleková and Baláž [12] or Braham [14]) and was later removed from the final version of the Eurocode 3 [18], and published in the ECCS publication [19], instead (see also Section 5.2). More recently, the lateral-torsional buckling of beams with mono-symmetric sections was studied by Kitipornchai et al. [20], Mohri et al. [21], Mohri et al. [22], Mohammadi et al. [23] and Achref et al. [24]. For the case of tapered beams, a comprehensive literature review is provided in Andrade and Camotim [25]. Several authors used the Finite Element method to study this problem [26–32], yet analytical solutions and analytical expressions are yet to be

* Corresponding author.

E-mail address: ccouto@ua.pt.



Fig. 1. Example of collapsed portal frame made of tapered members (courtesy of René Oly).

developed. Nevertheless, simplified methods exist based on the equivalent moment approach developed by Serna et al. [33] or the equivalent member by Galéa [34] or later by Braham and Hanikenne [28]. The concept underlying these methods is that the response of a tapered beam can be approximated to that of an equivalent uniform member which, despite its notable merits, provides a conservative approach. The design of web tapered members according to the North-American Specification AISC360-16 [35] and its accompanying design guide for web tapered members [36] is somehow based on this concept, as well. Eurocode 3 [18] and the Australian Standard AS4100 [37] include the concept of the critical moment for the stability design of tapered beams and allow its calculation using more advanced methods such as FEM.

Recent developments on design models for tapered beams [38–42] are also based on the calculation of the critical moment or the corresponding critical amplifier.

In this study, we develop neural network models using machine learning techniques to efficiently calculate the critical moment of uniform and tapered beams, simply supported and subject to bending loads in the strong-axis applied at the beam ends and covering double and mono-symmetric sections.

2. An overview of the literature on neural networks applied to steel members

Machine learning and, in particular, neural network models have been successfully applied by different researchers to solve complex problems related to the analysis and design of structures [43]. This section presents an overview of the literature grouped into different categories and with the indication of the number of samples, when available, used to develop the models.

2.1. Cellular and castellated beams

A significant amount of research on steel beams and the application of machine learning techniques has been put into the study of cellular or castellated members. Gholizadeh et al. [44] studied the load carrying capacity of simply supported castellated steel beams, susceptible to web-post buckling. They developed a neural network model and an Adaptive Neuro-Fuzzy Inference System (ANFIS) using 140 samples from FEM and demonstrated that the best accuracy was associated with the ANFIS and the neural network models in comparison with their proposed equations. Abambres et al. [45] developed a neural network-based formula for the critical moment of simply supported cellular beams subject to uniformly distributed transverse loads. A dataset with 3645 samples was generated using FEM. Sharifi et al. [46,47] employed neural network models for simulating and predicting the ultimate capacities of cellular steel beams against lateral-torsional buckling. The models were trained on a dataset of 99 samples. The ultimate capacity and, particularly, the resistance against lateral-distortional buckling was also investigated by Hosseinpour et al. [48] using neural networks. A study on the elastic and inelastic buckling of cellular beams was presented by Rajana et al. [49]. They used an impressive amount of 102600 FE analyses of samples which were then used to develop and validate a neural network-based formula. Nguyen et al. [50] developed also a neural network model to predict the load-carrying capacity of castellated beams. The authors collected 150 samples from the literature, from which 140 are those in [44]. Limbachiya and Shamass [51] investigated neural networks to predict the web post-buckling shear strength of cellular beams made of normal strength steel. The authors used 304 samples from a FEM model to develop their models. In addition, they concluded that using geometric parameters as the input of the models led to increased accuracy instead of using geometric ratios. Ferreira et al. [52] developed a neural network formula to predict the lateral-torsional buckling (LTB) resistance of slender steel cellular beams. Their models were trained on 768 samples obtained from a parametric study from FEM. In addition, they compared their neural network model with analytical expressions showing that it can accurately predict the LTB resistance of cellular beams as well as the LTB resistance combined with web post-buckling or web distortional buckling modes.

2.2. Distortional buckling of steel members

Research on the application of neural networks to the distortional buckling of steel members is also a substantial research field. Pala [53], Pala and Caglar [54] proposed a neural network based formula for the determination of the elastic distortional buckling stress of cold-formed steel C-sections with both end sections pinned; the results were obtained from another author and the number of samples considered was not reported. As mentioned by the authors, the results of the neural network model were satisfactory and proved consistent with the literature. Dias and Silvestre [55] presented closed-form expressions to calculate the distortional critical stress and half-wavelength of elliptic hollow sections under compression, using neural networks. The authors considered a total of 396 samples. Tohidi and Sharifi [56] used a neural network model to predict the restrained distortional buckling strength of half-through I-section bridge girders. They developed a FEM model to build a dataset of 396 samples used to train and develop their model.

2.3. Beams

In addition to the study of cellular and castellated beams, as well as the study of distortional buckling, researchers have also successfully employed neural networks to address other topics related to beams. Neural networks for optimization of cold-formed steel beams were applied by Adeli and Karim [57], Karim and Adeli [58], and Tashakori and Adeli [59] for space structures using the model developed in [60,61]. In Guzelbey et al. [62], the authors proposed a neural network approach for the estimation of

available rotation capacity of wide flange beams, using a total of 75 experimental samples from the literature and then developed an explicit formulation for this problem. In D'Aniello et al. [63], a soft-computing based study aimed to estimate the available rotation capacity of cold-formed rectangular and square hollow section (RHS-SHS) steel beams was developed using neural network and genetic expression programming with 64 experimental data collected from the literature. The flexural behaviour of steel beams and the proposal of a flexural overstrength factor was investigated in [64,65] using machine learning techniques. A model based on neural networks was developed by Güneyisi et al. [64] for I-H and RHS-SHS steel beams using 141 experimental samples collected from the literature. In D'Aniello et al. [65], the authors used neural networks and gene expression programming to develop a predictive model for circular hollow section beams based on 128 experimental samples from the literature. The authors of both works demonstrated the improved accuracy of their models in comparison with existing analytical formulae. Tohidi and Sharifi [66] developed a neural network model with a reliable modelling method for simulating and predicting the ultimate moment capacities for intermediate doubly-symmetric steel l-beams. Their dataset consisted of 174 data samples generated from a developed FEM model. Tohidi and Sharifi [67] developed a three-layer Backpropagation neural network to predict the residual buckling strength of corroded steel plate girders. The samples were obtained using FEM but only 37 data points were used to develop the models. The prediction of ultimate resistance of beams subjected to patch-load using neural networks was studied in [68,69]. In these studies, the authors collected 161 samples from the available experiments in the literature.

2.4. Columns

Apart from the research on distortional buckling of cold-formed members, the application of neural networks to columns remains a less explored field. Kassas et al. [70] developed a neural network to aid in the selection of the most economical cold-formed section for a particular application; in their study a total of 493 samples were considered. In Lyu et al. [71], the comparison analysis of strength prediction on the perforated cold-formed steel columns using FEM, regression analysis and neural network methods was investigated. The dataset contained 90 samples from stub columns tests and FEM results. A deep learning-based axial capacity prediction for cold-formed steel channel sections was developed using Deep Belief Network (DBN) in Fang et al. [72] using 10500 data samples generated from FEM. In Fang et al. [73], a database of 50000 data samples to train and develop DBN for structural performance of cold-formed steel channel sections with edge-stiffened/un-stiffened web holes, under axial compression. The capacity prediction of cold-formed stainless steel tubular columns using machine learning methods, including the use of neural networks, was studied in [74] considering a total of 322 test data collected from the literature as the database. Nguyen et al. [75] investigated the neural network and Adaptive Neural Fuzzy Inference System (ANFIS) models to compute the axial compression capacity of cold-formed steel oval hollow section columns. Their dataset comprised 128 samples collected from the literature. The results demonstrated that the developed models were more accurate than the existing formulae. Nguyen et al. [76] developed a neural network model for predicting the critical buckling load of the web tapered columns. Their database consisted of 269 FEM simulations. In comparison, the results of their model demonstrated to be more accurate than the existing formulae.

3. Data generation

For the problem addressed in this work, a host of samples can be generated for training and developing our model and the more samples the better. Therefore, a compromise must be sought for the definition of a realistic number of samples that are required to develop a model, targeting a broad application context, and offering a sound

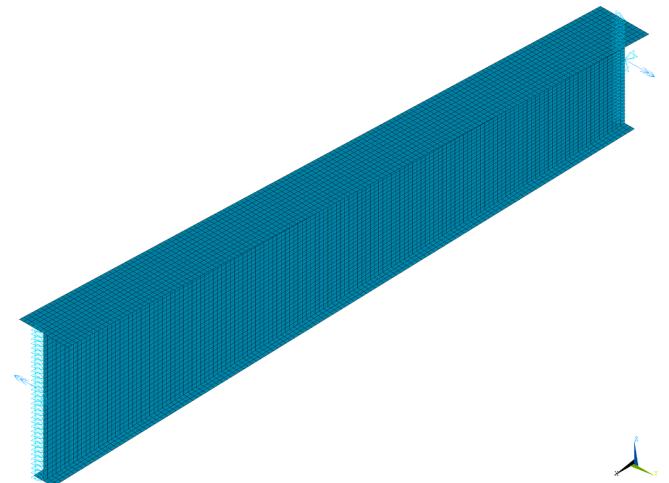


Fig. 2. Numerical model details to calculate the critical moment of beams.

generalization ability.

The numerical model used to obtain the target value of each sample, i.e., the critical moment of steel beams, is described next, while the samples considered in the dataset, as well as the strategy for its definition and characterization is detailed afterwards.

3.1. Numerical model

A numerical model based on the FEM was created in the software ANSYS [77], using the shell finite element SHELL181 available in the program. SHELL181 is suitable for analysing thin to moderately-thick shell structures and it is a four-node element with six degrees of freedom at each node, corresponding to the translations in the x , y , and z directions, and rotations about the x , y , and z -axes.

The simply supported beams were modelled by considering the mid-line of the web and flange plates and a preliminary study on the mesh sensitivity was carried out. The mesh size was defined as a function of the beam geometry and respecting an aspect ratio of 1.0 for the elements. Generally, for beam lengths less than 20 times the height of the (tallest) cross-section, 16 divisions on the largest flange were adopted and, for longer lengths, 12 divisions were defined instead to reduce the computational cost.

Fork-supports were considered at the beam ends by imposing fixed lateral displacements at the web in the y -axes direction and fixed vertical displacements in the z -axes direction at the centroid. To prevent the rigid-body movement in the longitudinal direction, i.e., in the x -axes direction, an additional node at the centroid of the starting section (at $x = 0$) was fixed. Because thin-walled beams are considered, to prevent the occurrence of either local or distortional buckling modes, multi-point constraints were added to the model to obtain the global mode and the corresponding critical moment. These additional equations were added to the model using the CERIG command along the length of the beam. Using this command, the rotation (R_x) of the nodes that define the cross-section was constrained to the rotation (R_x) of the corresponding node at the cross-section centroid. At the beam ends, kinematic constraints were added to the model to ensure correct loading distribution following the assumptions in Snijder et al. [78]. Accordingly, in the web nodes the five degrees of freedom U_x , U_y , U_z , R_x and R_y were constrained to the centroid of the section, and in the flanges, all the degrees of freedom were constrained to the flange middle node.

To apply the loading, bending moments were defined at the centroid of the end sections. For non-uniform bending situations, the γ corresponds to the ratio between the end-moments.

A linear elastic material model with Young modulus of $E = 210$ MPa and poisson ratio of $\nu = 0.3$ was considered.

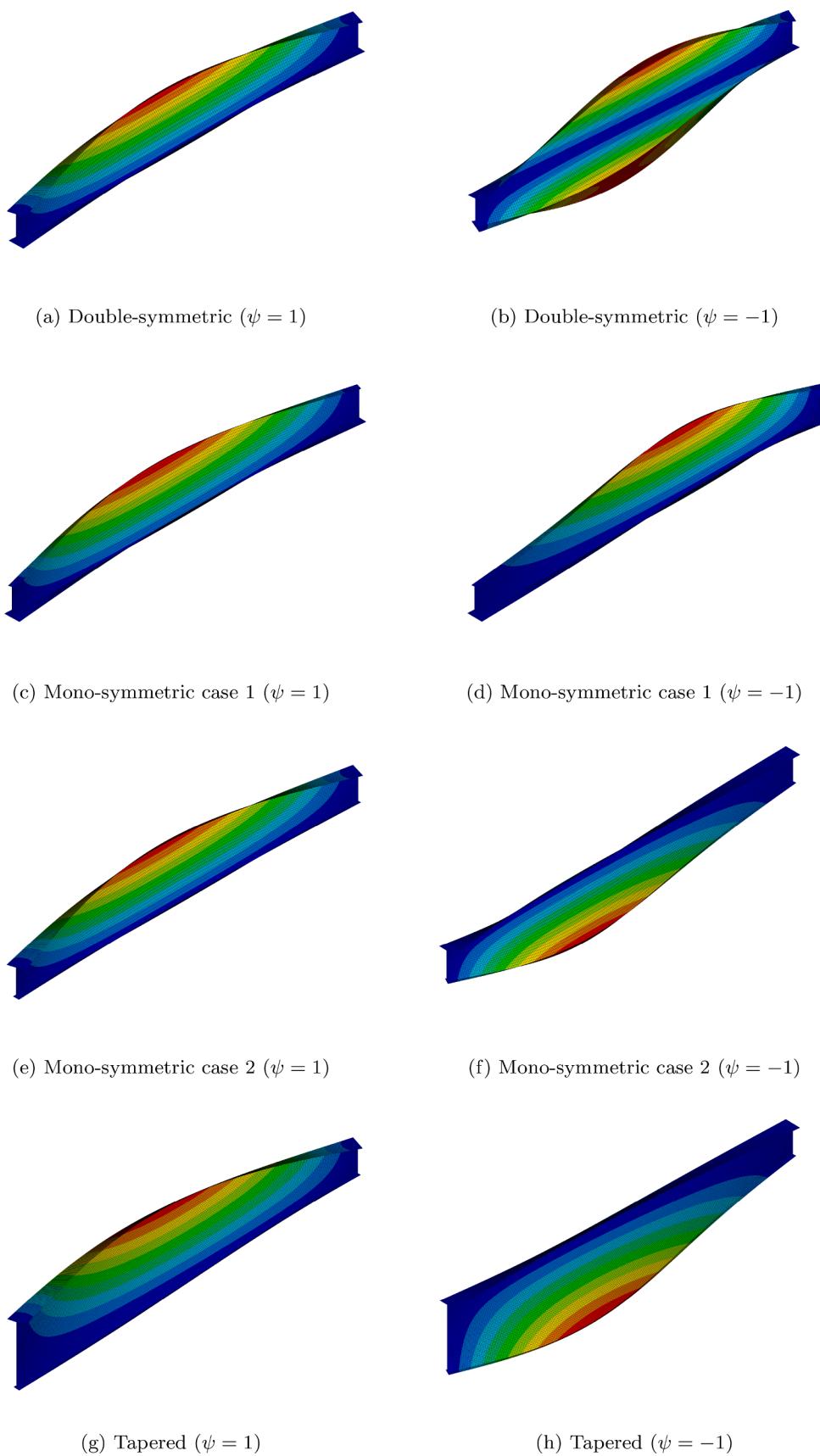


Fig. 3. Examples of buckling mode shapes of beams with different geometries and loading.

Table 1
Features considered in the models.

| Feature | Input values | Min. [mm] | Max. [mm] |
|---------|--------------|-----------|-----------|
| x_1 | h_{w1} | 75 | 1000 |
| x_2 | h_{w2} | 75 | 1000 |
| x_3 | t_w | 1 | 40 |
| x_4 | b_{f1} | 75 | 1000 |
| x_5 | t_{f1} | 1 | 80 |
| x_6 | b_{f2} | 75 | 1000 |
| x_7 | t_{f2} | 1 | 80 |
| x_8 | ψ | -1 | 1 |
| x_9 | L | 600 | 40000 |

Linear buckling analysis (LBA) were then performed to obtain the elastic critical moment for lateral-torsional buckling of the beam, M_{cr} , which corresponds to the first positive eigenvalue calculated by the software.

The numerical model used in this study is depicted in Fig. 2 and examples of the obtained global buckling modes are presented in Fig. 3.

3.2. Dataset

3.2.1. Definition and description

The Finite Element model described in the previous section was used to generate the samples in the dataset. Each sample is then defined by 9 features based on its geometrical dimensions and the ratio between the end-moments (ψ), with the minimum and maximum values of each feature indicated in Table 1 and the notation given in Fig. 4. The values for the geometric parameters of each data sample were designed to cover a wide range of cases in practice and since the neural networks were developed using this dataset, their scope of application is limited to the beams within the range of these features and the corresponding notation and are limited to simply supported beams subject to end-moments at extremities, accordingly to the numerical model used to obtain the samples.

The target output, y , of each sample in the dataset was defined as,

$$y = \frac{M_{cr}}{M_{cr,u}} \quad (1)$$

where M_{cr} is the elastic critical moment for lateral-torsional buckling of the beam calculated with the Finite Element model presented in the Section 3.1; and $M_{cr,u}$ is the elastic critical moment for lateral-torsional buckling of a beam of uniform symmetrical cross-section with equal flanges and subject to a constant bending moment. For tapered and mono-symmetric beams, it was considered one equivalent section with a web height equal to the average of the web heights at the beam edges, a flange width corresponding to the average of the widths from the top and the bottom flanges, and the thickness of the flanges equal to the average thickness of both flanges. $M_{cr,u}$, is calculated as (see also Eq. (17) and [19]):

$$M_{cr,u} = \frac{\pi^2 EI_z}{L^2} \sqrt{\left[\frac{I_w}{I_z} + \frac{L^2 GI_t}{\pi^2 EI_z} \right]} \quad (2)$$

where E and G are the steel Young modulus and shear modulus,

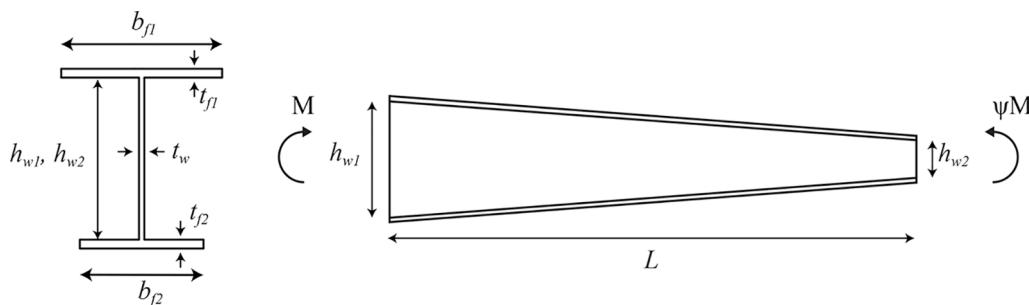


Fig. 4. Cross-section and beam notations.

respectively; I_t is the torsion constant; I_w is the warping constant; I_z is the second moment of area about the minor axis; and L is the length of the beam.

3.2.2. Number of samples

A strategy was implemented to reduce the number of samples to a reasonable amount and, thus, lower the computational cost necessary for creating the dataset because the simple consideration of 4 values for each of the 9 features defined earlier would result in a number of cases of $4^9 = 262144$. In the literature, this concept is often referred as smart sampling, which is an active research topic that the reader may further reference, for example, to [79].

The two main ideas for the smart sampling strategy adopted in the present study are i) a random number generator for each feature consisting in a minimum and maximum values and a certain number of points in the interval or step size; and ii) imposing the constrained geometric proportions defined in Table 2 to the samples. It should be noted that the geometric limits defined in this table seek to follow common values used in practice, in particular, the limits for the flange and the web proportions defined in [80].

Accordingly, an example of this strategy is that the features x_1 and x_2 , corresponding to the beam web height at both extremities, can have values between 75 and 1000 (see Table 1), however, the ratio between the two, i.e., the taper ratio of the beam, must remain within the [1, 4] interval. This is implemented by defining a certain web height for one edge, $h_{w,min}$, and the web height on the other edge, $h_{w,max}$, is a function of a random taper ratio chosen, for example, from the values [1, 2, 3, 4]. Finally, $h_{w,min}$ and $h_{w,max}$ are randomly defined as h_{w1} and h_{w2} .

The samples were then grouped into a training set and a validation (test) set, used respectively to train and validate the neural networks. The neural network models were not exposed to the samples of the validation set during learning. This allows to identify whether the model has effectively learned the patterns in the data and is capable of generalization to new samples instead of memorizing the patterns of the data. In the latter case, the neural network would have a poor generalization performance and would be unreliable to predict the results of new samples. Figs. 5 and 6 plot the histograms of the features for the training set with 60549 samples and the validation (test) set with 8526 samples, respectively. In both cases, a higher number of samples for the

Table 2
Imposed geometric constraints in the generated samples.

| Constraint | Input values ^a | Min. | Max. |
|-----------------------|--------------------------------|------|---------------|
| Taper ratio | $h_{w,max}/h_{w,min}$ | 1 | 4 |
| Flange width ratio | b_{f1}/b_{f2} | 0.25 | 4 |
| Height to width ratio | $h_{w,max}/b_{f,max}$ | 1 | 4 |
| Beam length | $L/h_{w,max}$ | 2 | 40 |
| Plate thickness ratio | $t_{f1}/t_w; t_{f2}/t_w$ | 1 | ≈ 24 |
| Flange slenderness | $b_{f1}/t_{f1}; b_{f2}/t_{f2}$ | 6.25 | 100 |
| Web slenderness | $h_{w,max}/t_w$ | 25 | ≈ 300 |

^a $h_{w,min} = \min(h_{w1}, h_{w2})$, $h_{w,max} = \max(h_{w1}, h_{w2})$ and $b_{f,max} = \max(b_{f1}, b_{f2})$.

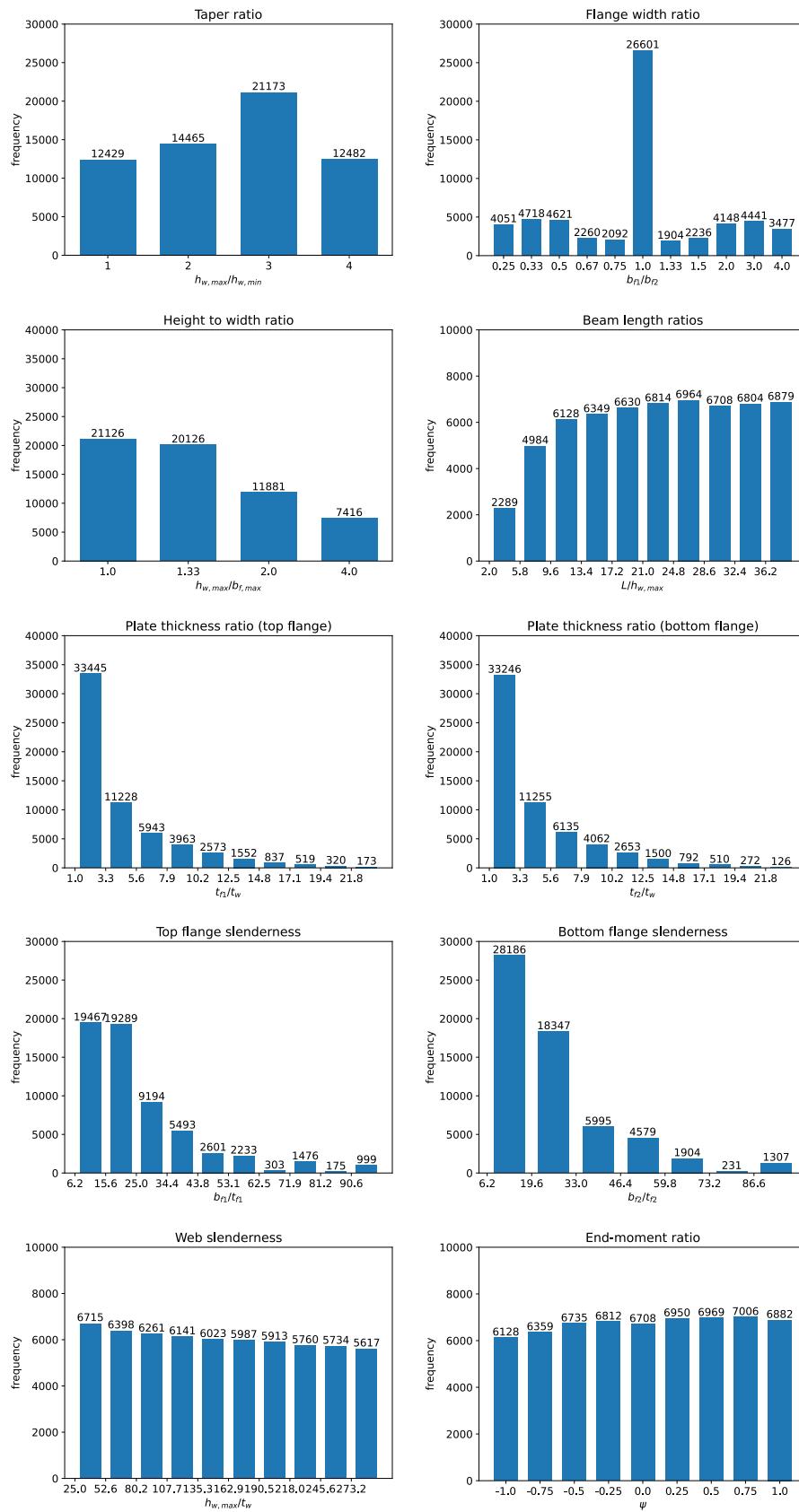


Fig. 5. Frequency distribution in the training set with 60549 samples.

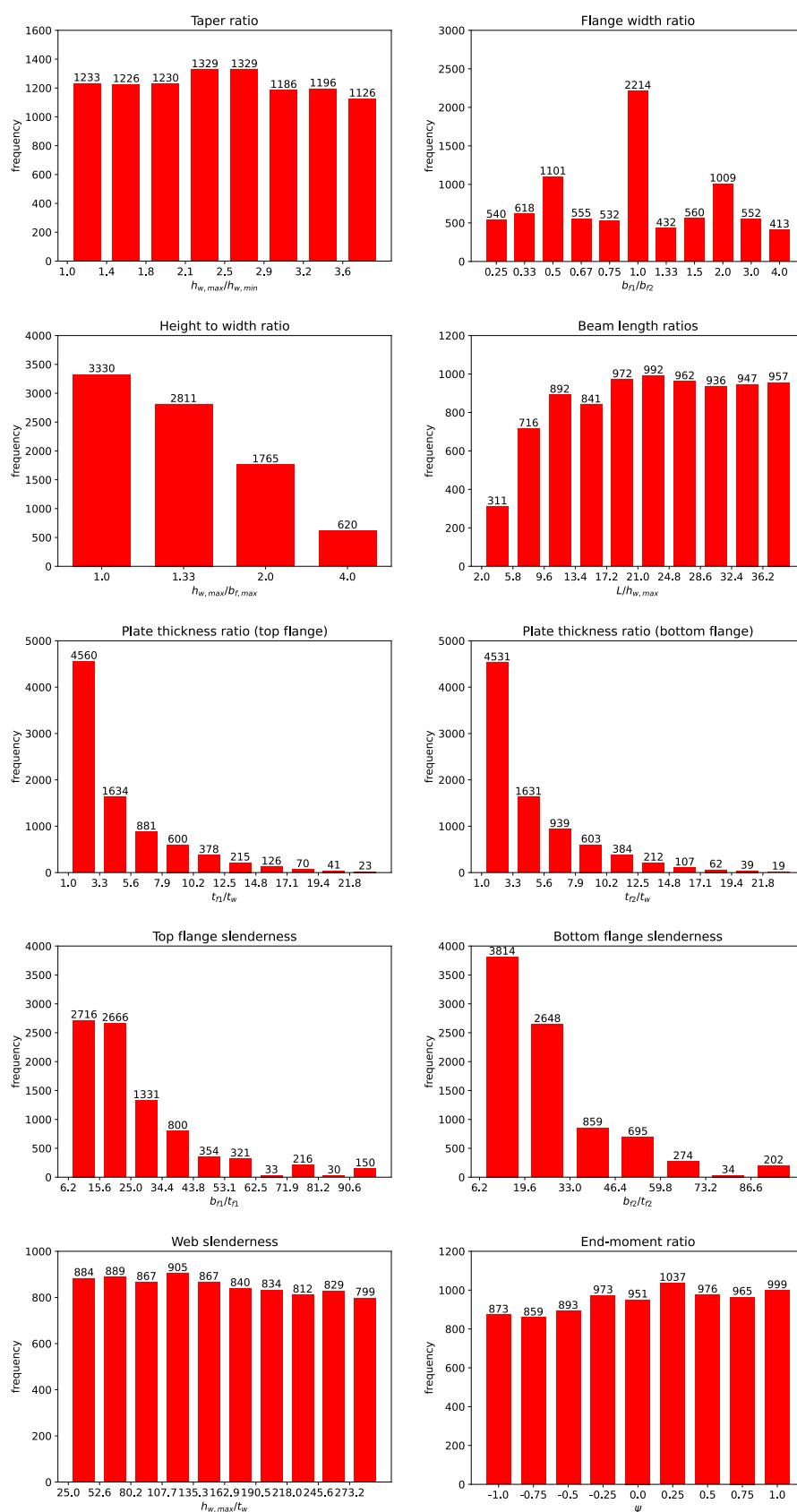


Fig. 6. Frequency distribution in the validation set with 8526 samples.

flange width ratio equal to 1.0 was considered to increase the number of beams with doubly symmetric cross-sections. Some additional considerations and results regarding the dataset size are presented in the Section 5.1.

4. Neural networks

The neural network (NN) are mathematical models described as a group of neurons arranged in layers and their connections that form a network. The ANN can map a certain range of input values (features) to a specific target value (result). The feed-forward Multilayer Perceptron (MLP) network was used in this study with an input layer, one or two hidden layers, and an output layer. Each neuron of a certain layer is only connected to the neurons in the next layer, and includes an extra, independent, value called the bias. In the forward pass, the weighed sums of a layer k can be calculated as,

$$z^{(k)} = W^{(k)} a^{(k-1)} + b^{(k)} \quad (3)$$

where $a^{(k-1)}$ denotes the outputs of the neurons from the previous layer, $W^{(k)}$ is the weight matrix and $b^{(k)}$ are the bias values. The neurons in the first layer receive the components of each sample feature given in Section 3.2 as inputs,

$$a^0 = x_i \quad (4)$$

while the output of each neuron in a certain layer k can be calculated as:

$$a^{(k)} = f(z^{(k)}) \quad (5)$$

The $f(x)$ is the activation function. In this study the hyperbolic tangent was used,

$$f(x) = \tanh(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)} \quad (6)$$

and for regression problems, such as the one studied herein, the linear activation function is used as the output for the final layer M , that becomes,

$$z^{(M)} = \hat{y} = W^{(M)} a^{(M-1)} + b^{(M)} \quad (7)$$

Finally, for the i th sample presented to the network the \hat{y}_i corresponds to its predicted value calculated with Eq. (7).

The training of the ANN models is achieved by adjusting the coefficients of the weight matrix W and bias b for a range of samples. During the training process, these values are adjusted to minimize the errors between the network output \hat{y} and the desired target output, defined by Eq. (1). For that purpose, the Backpropagation algorithm [81,82] was used together with the Levenberg–Marquardt optimization algorithm [83,84] (see also [85]), described as follows.

The Backpropagation algorithm uses the mean squared error as the performance index to be minimized, which is given as,

$$V = \frac{1}{2} \sum_{i=1}^Q (y_i - \hat{y}_i)^2 = \frac{1}{2} \sum_{i=1}^Q e_i^T e_i \quad (8)$$

and that can be approximated using a steepest descent rule as,

$$\hat{V} = \frac{1}{2} e_i^T e_i \quad (9)$$

where the total sum of squares is replaced by the squared errors for a single data sample. According to the gradient descent algorithm (see [85]), the updates to the weights and bias are calculated as,

$$\Delta W^k = -\alpha \delta^k a^{k-1} \quad (10)$$

$$\Delta b^k = -\alpha \delta^k \quad (11)$$

where α is the learning rate and δ^k is,

$$\delta^k = \frac{\partial \hat{V}}{\partial z^k} \quad (12)$$

The following recurrence relation can be established [85],

$$\delta^k = f'(z^k) W^{k+1T} \delta^{k+1} \quad (13)$$

which is initialized at the final layer, M , as,

$$\delta^M = -f'(z^M) (y_i - \hat{y}_i) \quad (14)$$

The Levenberg–Marquardt algorithm is a modification of the Gauss–Newton method for the weight updates Δx resulting in [85],

$$\Delta x = [J^T(x) J(x) + \mu I]^{-1} J^T(x) e(x) \quad (15)$$

where J is the Jacobian matrix [85], I is the identity matrix, $e(x)$ is the error vector between the network output \hat{y}_i and target value y_i , and μ is a parameter that is multiplied by β whenever the loss increases and divided by β when loss decreases, being adopted $\beta = 10$ and $\mu = 3$ in this work. This algorithm corresponds to the Gauss–Newton method for small μ and, when μ is large, it becomes the gradient descent method.

These methods are implemented into the *pyrenn toolkit* [86] which is developed in Python programming language. *Pyrenn* was used in this work to train and develop the neural networks. For training, 500 epochs were considered corresponding this number to the times the training samples are presented to the network to adjust the weights and bias. This number is reasonable as the Levenberg–Marquardt algorithm converges rapidly.

The accuracy of the models is evaluated using the coefficient of determination R^2 , defined as,

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad (16)$$

where y_i corresponds to the actual output of the i th sample in the dataset given by Eq. (1) and \bar{y} is the mean value of all the outputs.

This metric gives the proportion of the variance between the predicted values and the given values, and the higher the R^2 is, the greater is the model capacity to identify the patterns in the data, thus the higher is its accuracy.

5. Results

This section shows the results obtained in this study for different neural network models against the results of the Finite Element model from Section 3.1. The results of the network architecture and dataset size are presented first as they were used to decide the number of samples considered as well as to define the network architecture used to then investigate the accuracy when compared to numerical results for uniform and non-uniform beams, presented after. In addition, for uniform beams, comparisons for existing analytical methods are also provided.

5.1. Network architecture and dataset size

The definition of the network architecture and the number of samples requires an iterative procedure since the accuracy of the network models depends on both. In addition, recall that the generation of samples followed a randomized methodology with some geometric constraints as highlighted before and thus a reasonable number of samples is necessary. Consequently, the models were trained several times with increasing number of samples until a satisfactory generalization ability was attained while the network architecture was simultaneously being varied to establish a reasonable number of samples in

Table 3

Accuracy and training times of different neural network architectures (using 60549 samples for training).

| Architecture | R^2 (train) | R^2 (test) | Training time ^a [min.] |
|-----------------------------------|---------------|--------------|-----------------------------------|
| $9 \times 18 \times 1$ | 0.98765 | 0.98697 | 71 |
| $9 \times 32 \times 1$ | 0.99338 | 0.99333 | 77 |
| $9 \times 32 \times 8 \times 1$ | 0.99878 | 0.99859 | 114 |
| $9 \times 32 \times 16 \times 1$ | 0.99908 | 0.99897 | 121 |
| $9 \times 32 \times 32 \times 1$ | 0.99952 | 0.99945 | 134 |
| $9 \times 64 \times 1$ | 0.99715 | 0.99694 | 92 |
| $9 \times 64 \times 8 \times 1$ | 0.99951 | 0.99938 | 132 |
| $9 \times 64 \times 16 \times 1$ | 0.99974 | 0.99965 | 141 |
| $9 \times 64 \times 32 \times 1$ | 0.99982 | 0.99966 | 162 |
| $9 \times 64 \times 64 \times 1$ | 0.99982 | 0.99974 | 207 |
| $9 \times 128 \times 1$ | 0.9985 | 0.9983 | 124 |
| $9 \times 128 \times 16 \times 1$ | 0.99987 | 0.99969 | 190 |
| $9 \times 128 \times 64 \times 1$ | 0.99997 | 0.99968 | 380 |
| $9 \times 256 \times 1$ | 0.99906 | 0.99884 | 240 |
| $9 \times 512 \times 1$ | 0.99934 | 0.99906 | 513 |

^a These values are merely indicative.

the dataset and to choose a proper network architecture.

Different network architectures were considered, with a single hidden layer ($9 \times m \times 1$) and with two hidden layers ($9 \times m \times n \times 1$). The choice of the parameters m and n was done by trial and error, starting by doubling the number of inputs (18) and then considering multiples of 8 for the first hidden layer, as well as for the second hidden layer. The accuracy results and the corresponding training times are highlighted in Table 3, in this case, considering 60549 and 8526 samples, respectively, for the train and validation sets with the distribution given in Figs. 5 and 6.

The accuracy of the models increases in the presence of a second hidden layer but increasing the number of neurons in the second layer has a reduced impact on the results. The computational cost, i.e., the training time, is more influenced by the number of neurons in the first hidden layer. The recorded training time may vary in two consecutive runs but for reference the models were trained on Intel® Xeon® Gold 6148 CPUs @ 2.40 GHz. The calculations were performed in a single core since *pyrenn toolkit* is not yet implemented to take profit of multi-core architectures and, while the training times may be influenced by the available RAM allocated for each run, this particular point was not exploited in this study.

For these number of samples and some single and double hidden layer architectures the results are also plotted in Figs. 7 and 8, respectively, further highlighting these observations.

Regarding the number of samples, using the $9 \times 128 \times 16 \times 1$ neural network model, the influence of the training set size on the accuracy of the validation (test) set is shown in Fig. 9.

In this figure, the number of samples in the training is varied from 10000 to 20000, and then to 40000. The samples in each training set are included in the next set, i.e., the 40 k set contains the samples from the 20 k and 10 k sets. During this experimentation the number of samples in the validation set is kept constant (8526) to ensure that the results are comparable.

It is observed that the prediction accuracy on the training set does not significantly change irrespective of the number of samples considered while for the validation set it increases from $R^2 = 0.99148$ to $R^2 = 0.99892$ and then to $R^2 = 0.99965$ for the 10 k, 20 k and 40 k training samples, respectively. For the 60549 samples, the accuracy obtained for the validation set was $R^2 = 0.99969$ (see Fig. 8). These figures also show a less scattering on the validation set results when the number of samples used to train the models is increased, supporting the notion that the more samples the better the models will perform.

These results show the converging nature of the methodology to define the number of samples and that considering 60549 is reasonable. Moreover, the neural network with the architecture $9 \times 128 \times 16 \times 1$ is considered next to further illustrate its accuracy for uniform and tapered

beams, the choice being based simultaneously on the accuracy and respective training time presented by this neural network in comparison with the other considered models.

5.2. Uniform beams

As mentioned before, in the draft version of Eurocode 3 [17], a formula was given to calculate the critical moment of uniform beams with double and mono-symmetric section. However, this formula led to some inconsistent results (see for example Koleková and Baláz [12] or Braham [14]) and was later removed from the final version of the Eurocode 3 [18], and published in the ECCS publication [19], instead. Ignoring the transverse loading related terms, not covered in the present study, the ECCS formula becomes:

$$M_{cr} = C_1 \frac{\pi^2 EI_z}{(k_z L)^2} \times \left\{ \sqrt{\left[\left(\frac{k_z}{k_w} \right)^2 I_w + \frac{(k_z L)^2 GI_t}{I_z} + (C_3 z_j)^2 \right]} + C_3 z_j \right\} \quad (17)$$

where C_1 and C_3 are factors depending on the loading and end restraint conditions; E and G are the steel Young modulus and shear modulus, respectively; I_t is the torsion constant; I_w is the warping constant; I_z is the second moment of area about the minor axis; L is the length of the beam; k_z and k_w are effective length factors; and z_j is a parameter to account for beams with unequal flanges.

Later, resulting from the works of Baláz and his colleagues [12] an improved formulae was proposed that, here ignoring the terms related to the transverse loading, reads,

$$M_{cr} = \mu_{cr} \frac{\pi \sqrt{EI_z GI_t}}{L} \quad (18)$$

where

$$\mu_{cr} = \frac{C_1}{k_z} \left[\sqrt{1 + \kappa_{wt}^2 + C_3 \zeta_j^2} + C_3 \zeta_j \right] \quad (19)$$

and κ_{wt} is a non-dimensional parameter; C_1 and C_3 are, as previously, factors depending on the loading and end restraint conditions; and ζ_j is a parameter similar to z_j , defined previously, to account with unequal flanges.

Although the same factors C_1 and C_3 are included in Eqs. (17) and (19), their values are different and given in their respective references and, for the sake of brevity, are not detailed here.

Figs. 10–12, plot the results of the critical moment of uniform beams for different bending moment diagrams (ψ) as a function of the length. The predicted results from the neural network (NN) with the architecture $9 \times 128 \times 16 \times 1$ are compared to the analytical ones from Eq. (17) (ECCS) and Eq. (18) (Baláz et al.), as well as with the results from the Finite Element Method (FEM) described in Section 3.1. In these figures, the section dimensions are given in the title following the $h_w \times t_w + b_{f1} \times t_{f1}$ notation, and for unequal flanges a third term is included with the dimensions of the lower flange ($b_{f2} \times t_{f2}$).

As expected, the critical moment of the beams decreases with increasing length. With respect to the relationship between the moments applied at the extremities, the value of the critical moment increases from the uniform case ($\psi = 1$) to the non-uniform distribution cases given by $\psi = 0$ and $\psi = -1$, being higher in the latter case. The analytical formulations and the neural network model follow these trends and show a very good accuracy in comparison to the FEM results.

In the Fig. 13 the results are compared for beams with a mono-symmetric section with unequal width flanges, $278.6 \times 7.1 + 150 \times 10.7 + 72 \times 10.7$. The FEM results depicted are those provided in Braham [14] to allow a comparison with results from other authors.

For this particular case, it is observed that the bi-triangular bending moment distribution ($\psi = -1$) does not correspond to the one with higher critical moment as the previous analysed cases. This is a result of

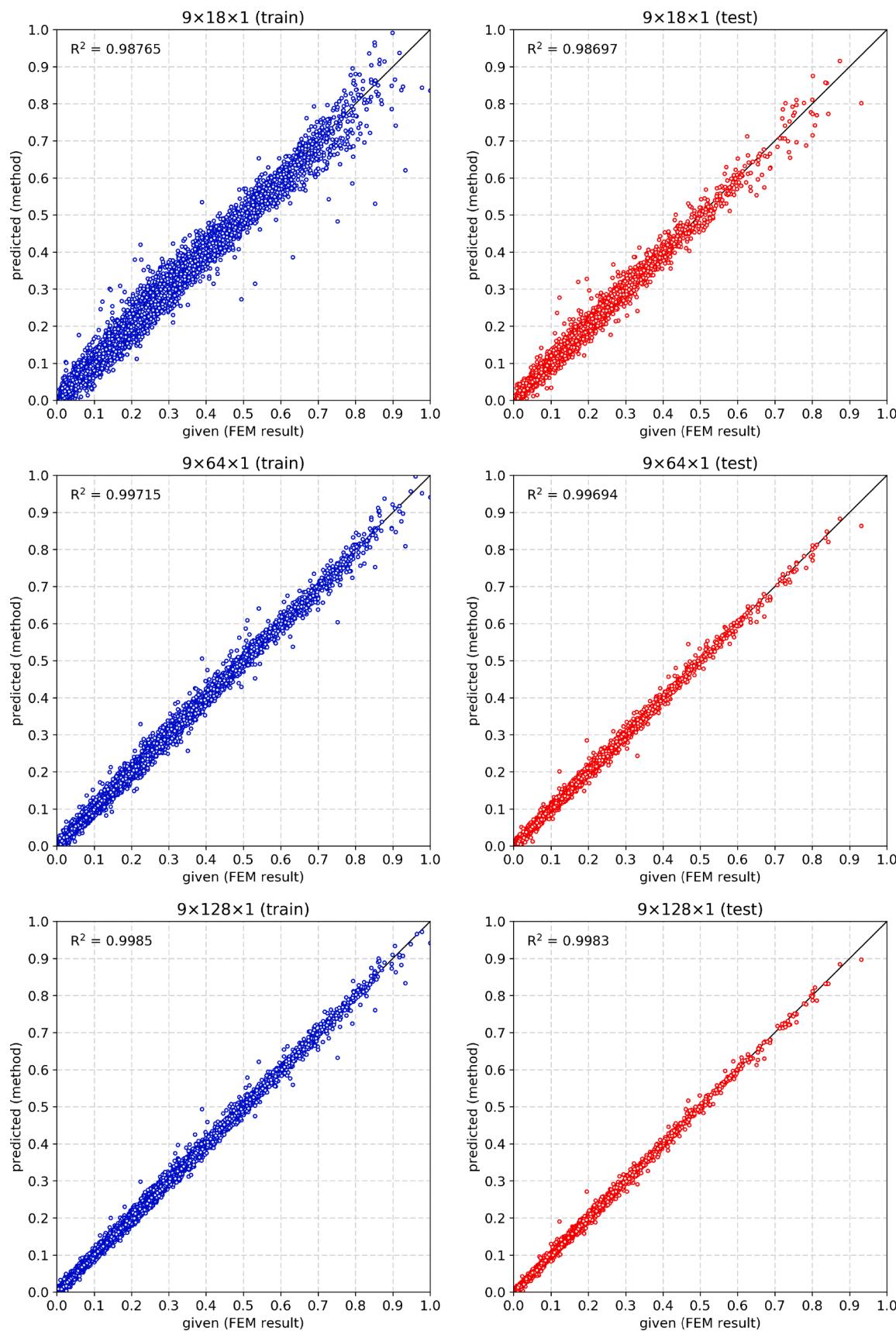


Fig. 7. Accuracy of single hidden layer neural networks.

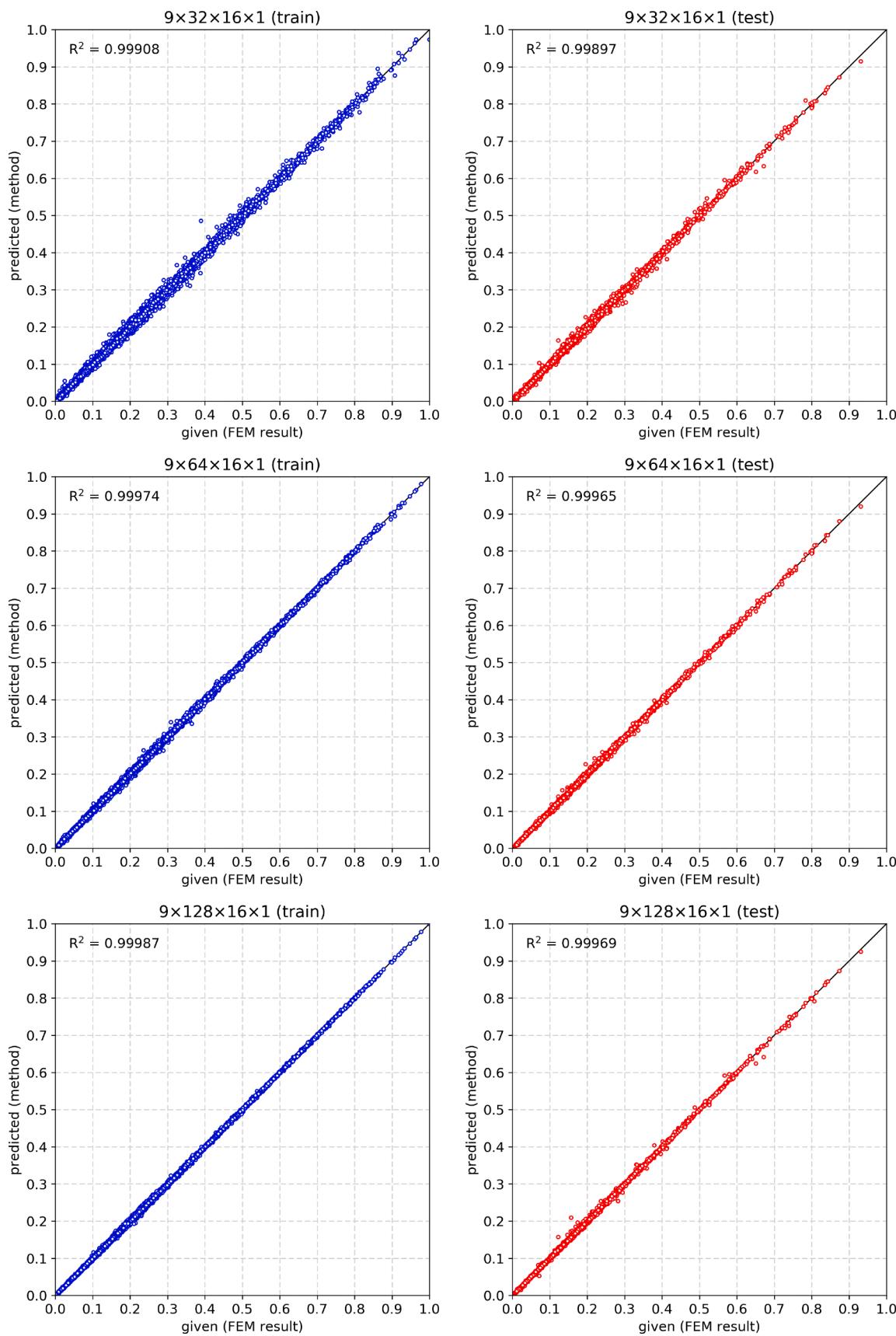


Fig. 8. Accuracy of two hidden layer neural networks.

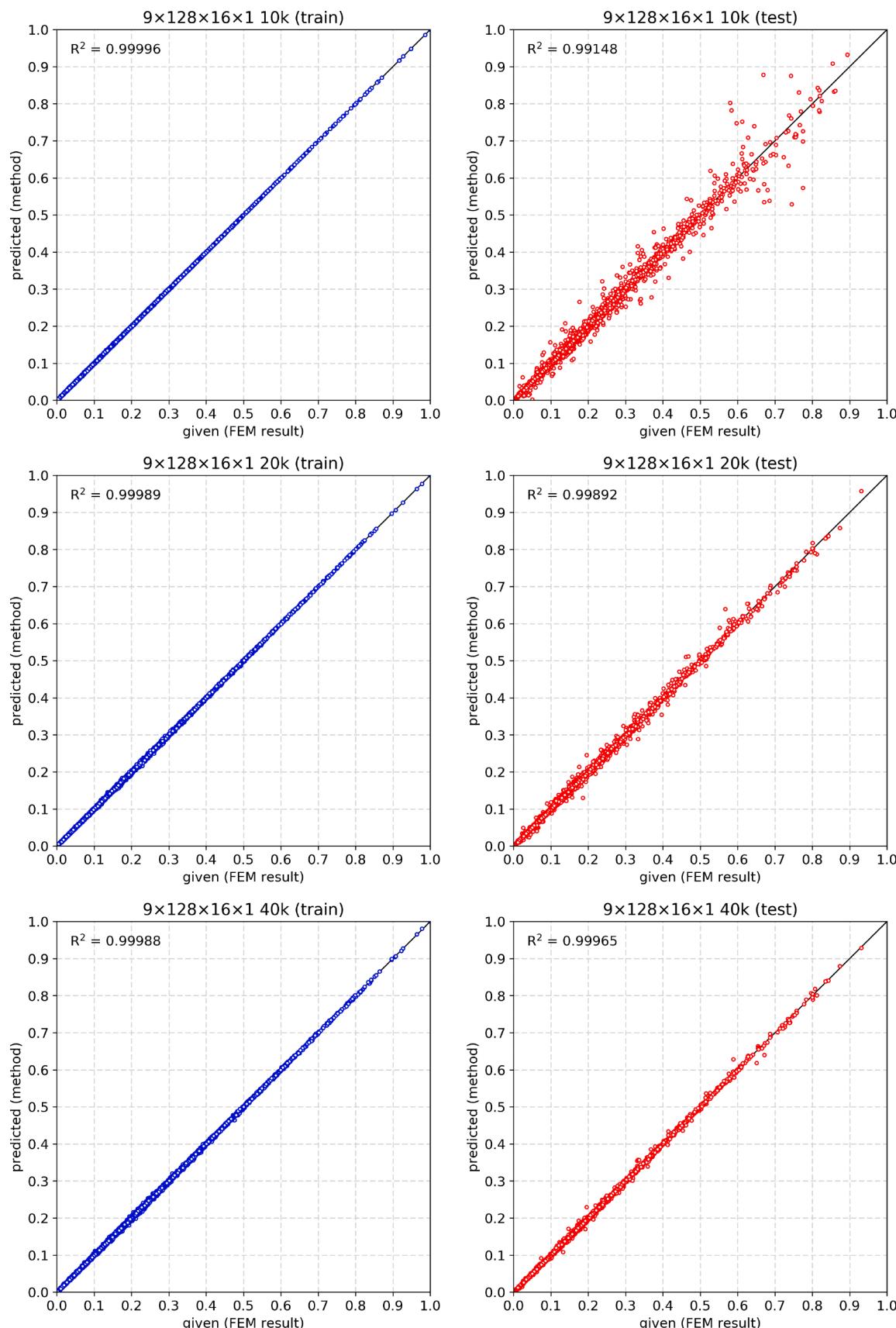


Fig. 9. Accuracy of the neural network $9 \times 128 \times 16 \times 1$ for different sizes of training set. Note that the validation set contains the same number of samples (8526) as used in all parts of this study.

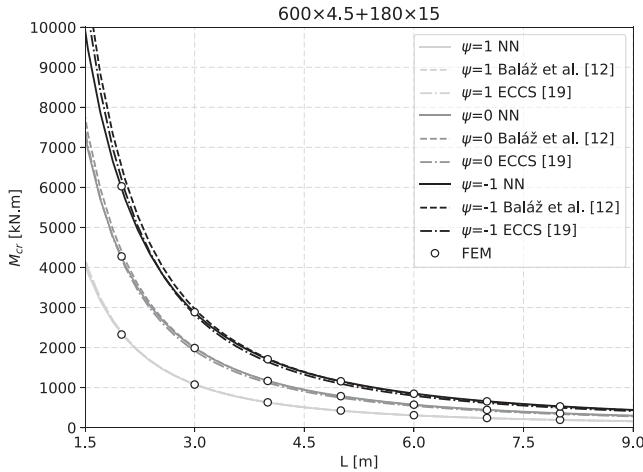


Fig. 10. Results for beams with double-symmetric section.

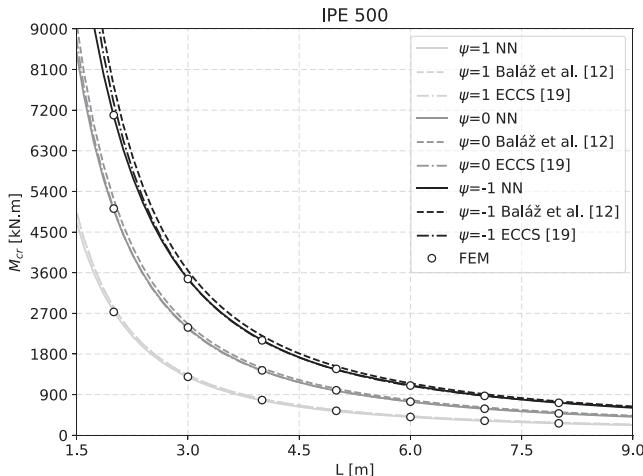


Fig. 11. Results for IPE 500 beams.

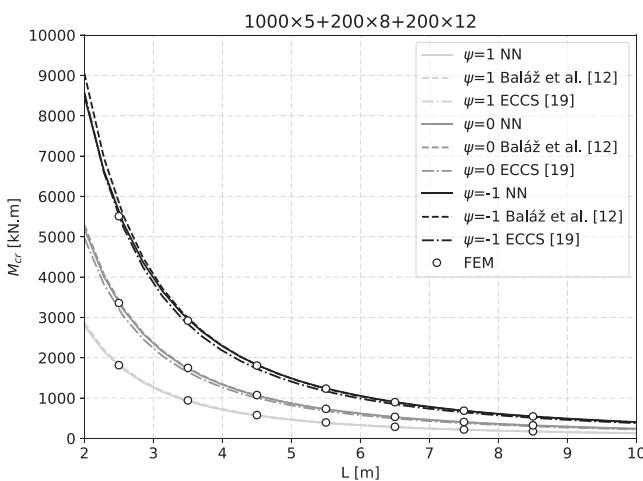


Fig. 12. Results for mono-symmetric beams with equal width flanges.

the unequal width of the flanges. Also stemming from this particularity, the ECCS formulation underestimates the critical moment of the beams for the cases $\psi = 0$ and $\psi = -1$, while the analytical formula from Baláz et al. provides a better agreement with the results from the FEM. This is further discussed in more detail in Refs. [11,12]. Again, the predicted

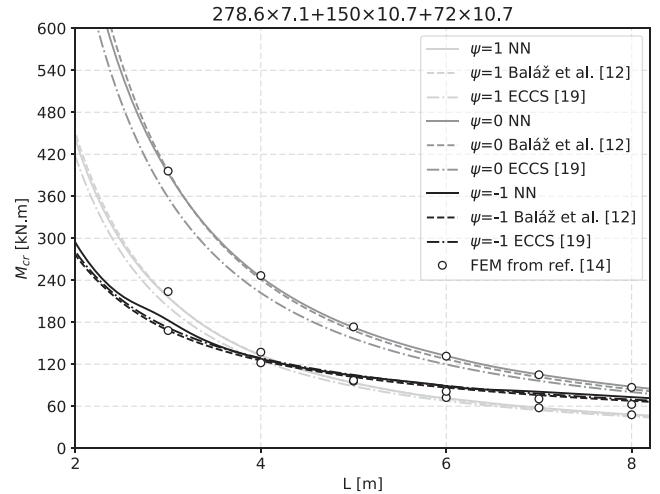


Fig. 13. Results for mono-symmetric beams with unequal width flanges and FEM results from Braham [14].

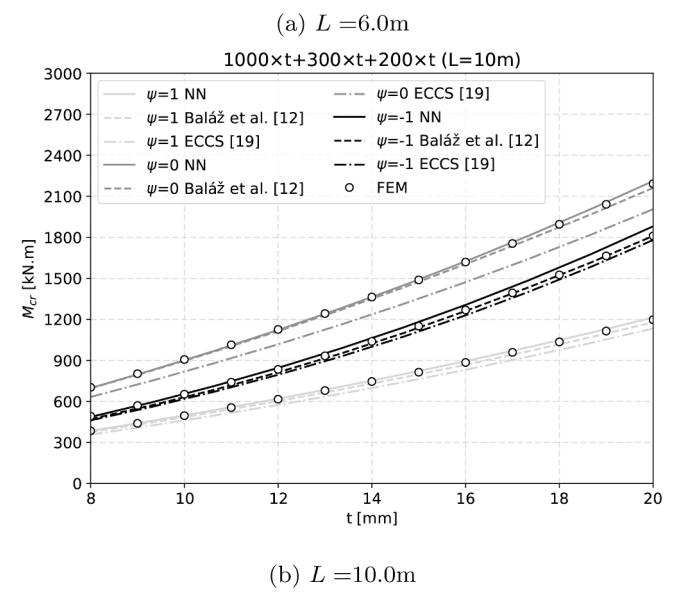
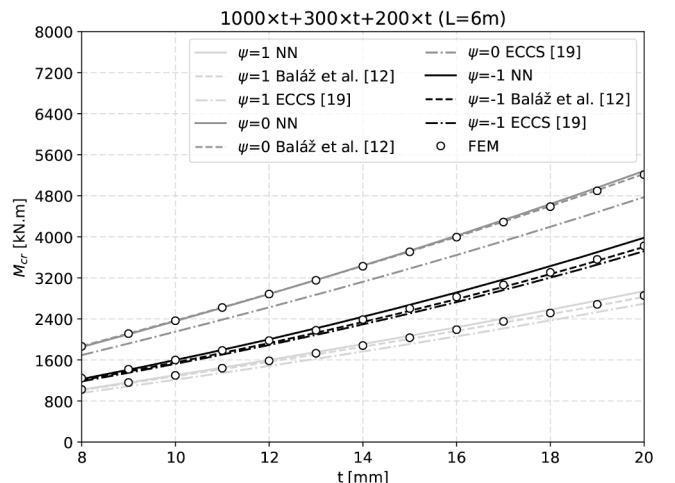


Fig. 14. Results for mono-symmetric beams with varying thickness of web and flanges.

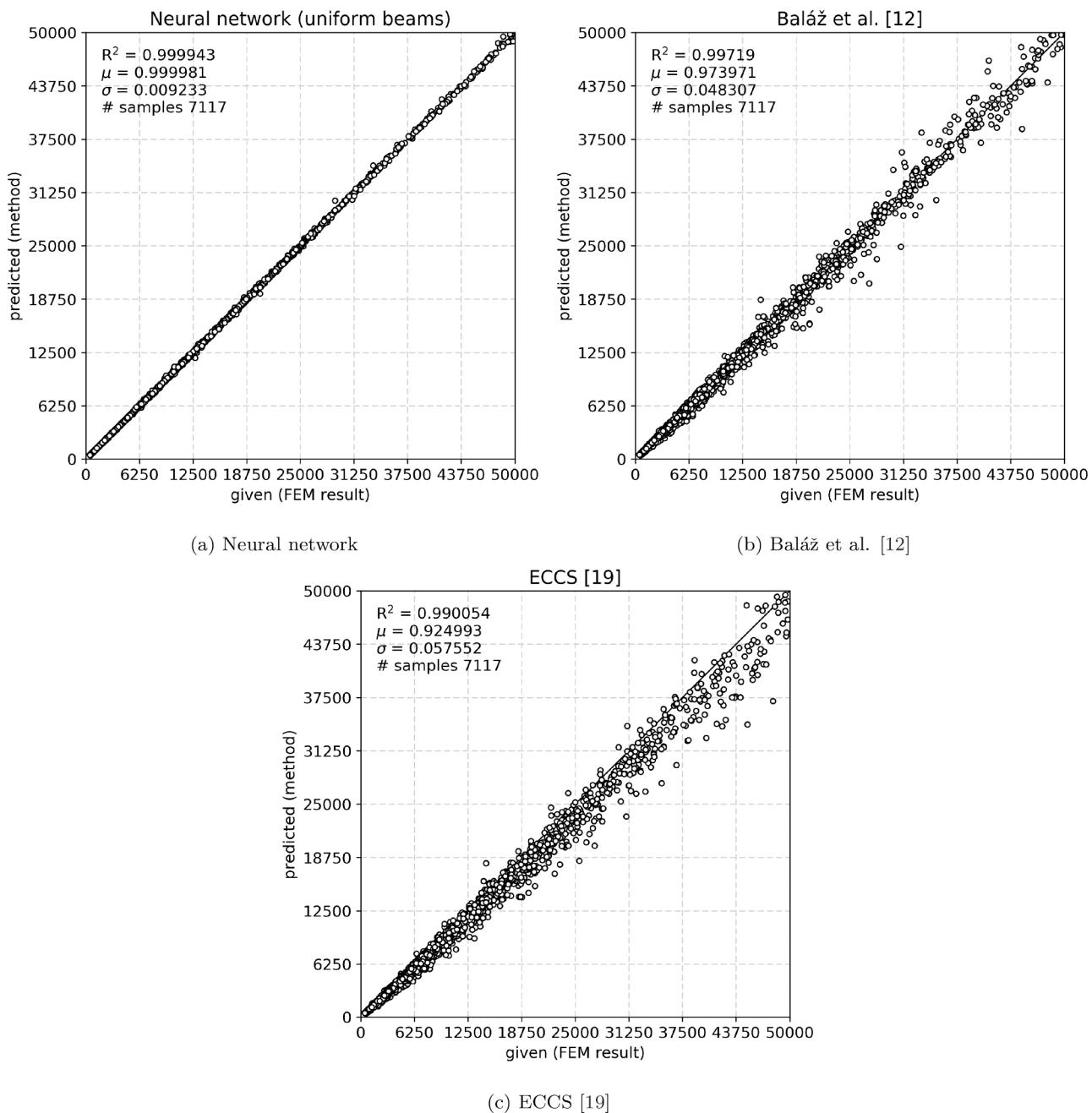


Fig. 15. Accuracy of the neural network and analytical models for beams with uniform geometry. Results of M_{cr} in [kN.m].

results of the neural network model are in good agreement with those obtained with FEM.

To further illustrate this point, Fig. 14 shows the results of two unequal width flange mono-symmetric sections where the thickness of the web and flanges was varied between 8 mm and 20 mm for beams with $L = 6.0$ and $L = 10$ m.

Again, it can be seen that the case $\psi = -1$ does not always lead to the situation where the critical moment is the highest. The ECCS formula underestimates the critical moment of beams with $\psi = 0$ for both lengths, while the Baláz et al. formula leads to a good accuracy for all the cases considered. The neural network follows the trends of both the analytical and FEM results and thus is capable of identifying the continuous pattern for this problem, predicting higher values for the critical moment as the thickness t increases.

In Fig. 15, additional statistical measurements are provided. The μ is the average of the ratio between the predicted and given values for each

sample, and σ is the corresponding standard deviation.

As expected, the neural network is the model with the best accuracy with $R^2 = 0.999943$, highest $\mu = 0.999981$ and lowest $\sigma = 0.009233$. The model proposed by Baláz et al. [12] shows a better accuracy in comparison to the ECCS one [19], but with lower accuracy than the neural network model, in line with the above mentioned observations.

5.3. Tapered beams

For tapered beams, as mentioned, no simple analytical formula exists for the elastic critical moment. As an alternative, predictive models based on neural networks can provide a fast and accurate methodology for that purpose. The neural network model developed in this study is compared against the results from the FEM to assess its reliability for the calculation of the critical moment of tapered beams. Specifically, because the training and testing datasets were obtained from a random

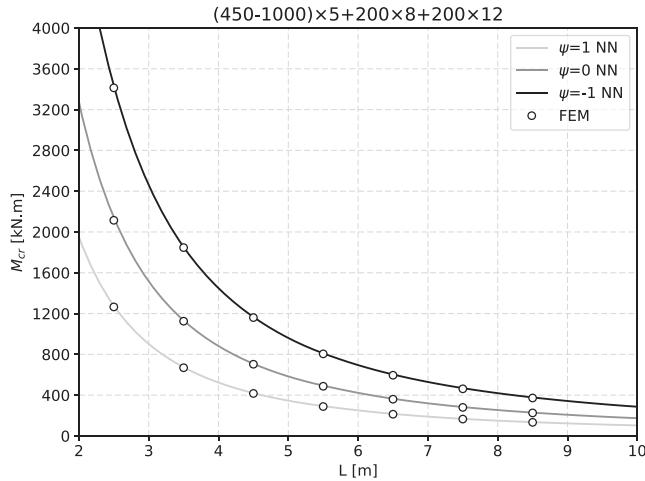


Fig. 16. Results for tapered beam with mono-symmetric section with equal width flanges.

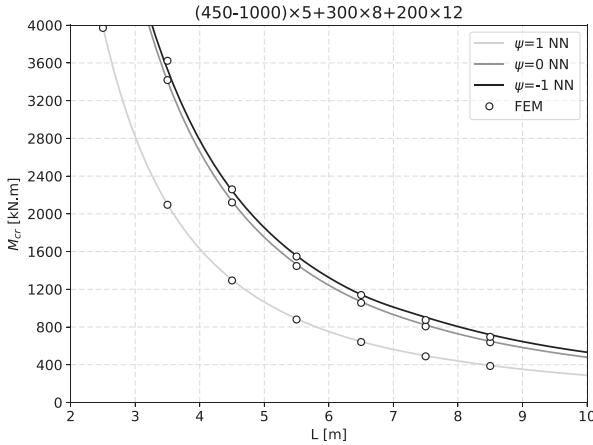


Fig. 17. Results for tapered beam with mono-symmetric section with unequal width flanges.

procedure, it is necessary to investigate the neural network accuracy when some parameters are linearly varied, for example the length of a beam and the thickness of the plates, as done in the previous section for uniform beams. Considerations regarding the taper ratio, i.e., the ratio between the height of the end-sections of the beam, were also investigated herein.

Figs. 16 and 17 plot the results obtained with the FEM model described in Section 3.1 and the neural network for different tapered beams and considering different bending moment diagrams (ψ). The critical moment of the beams M_{cr} results are plotted against the beam length.

In general, the results of the critical moment of non-uniform beams are similar to those of uniform beams, with higher values obtained for shorter lengths and different bending moment distributions. The behaviour observed for the particular case of unequal flange widths is also consistent with that observed for the uniform case where the results obtained for the $\psi = -1$ case are closer to those obtained for $\psi = 0$. This is further illustrated in Fig. 18, which shows the comparison of the results obtained with the FEM model and the proposed neural network for two tapered beam lengths with mono-symmetric cross-section. The results are plotted as a function of varying plate thickness t in the flanges and web.

For the shorter length, $L = 6.0\text{m}$, and irrespective of the thickness of the plates, the critical moment of the beams with $\psi = -1$ is lower than the one obtained for $\psi = 0$, while for the beams with $L = 10.0\text{m}$ it is

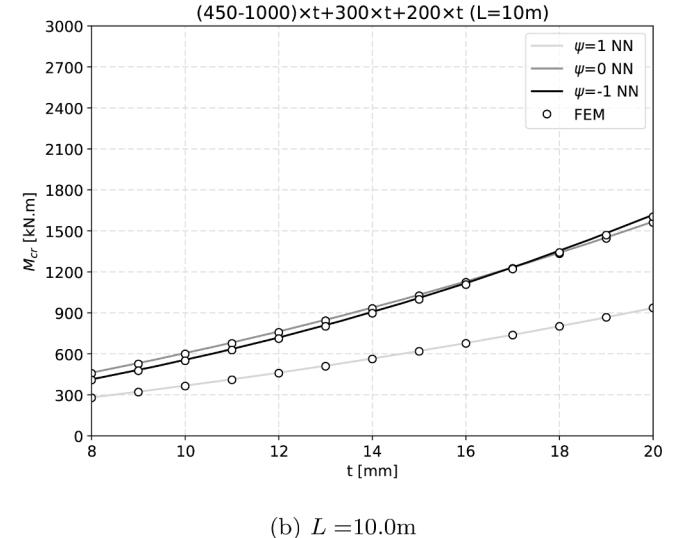
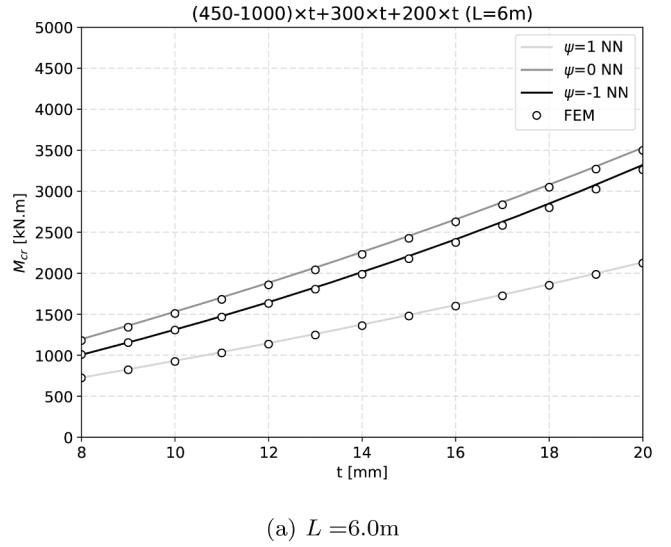


Fig. 18. Results for tapered mono-symmetric beams with varying thickness of web and flanges.

observed that for a thickness around $t = 16\text{ mm}$, the critical moment becomes higher for the $\psi = -1$ case. Despite this non-linear behaviour, the neural network model is capable of predicting the critical moment of these beams including the mentioned singular results obtained for the 10.0m beam with plate thicknesses around $t = 16\text{ mm}$ and $\psi = 0$ and $\psi = -1$.

Finally, in Fig. 19, the predictions from the neural network are compared against the Finite Element results of Asgarian et al. [32].

The average of the ratio between the predicted and given values for each sample was $\mu = 1.02568$ and the corresponding standard deviation $\sigma = 0.021726$, yielding $R^2 = 0.997825$. These results demonstrate that the proposed neural network model is accurate and provide a valid methodology to estimate the critical moment of tapered beams. These accurate results together with the fact that neural networks are extremely fast to calculate a case (in the order of milliseconds), open the door for new calculation methodologies to be adopted in the future for the design of steel structures. As for the particular case of tapered beams where no analytical exists, the neural networks can be adopted in the early design stages to design optimal solutions that can be later confirmed using most widely validated strategies such as the Finite Element method.

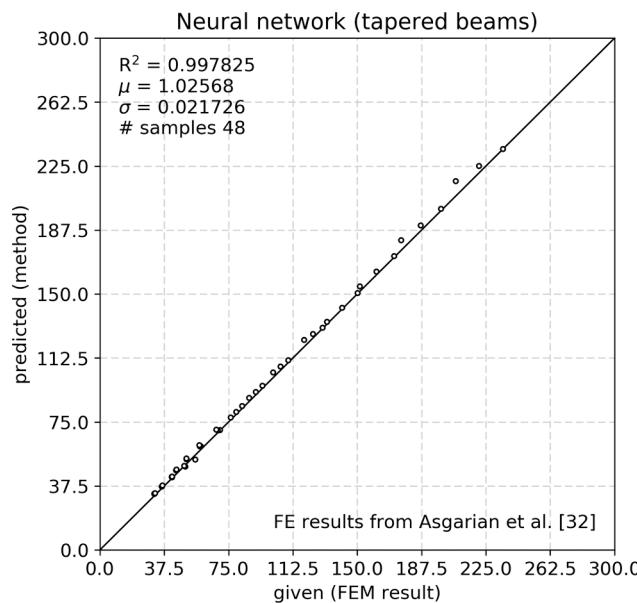


Fig. 19. Accuracy of the neural network predictions for beams with tapered geometry against the results from Asgarian et al. [32]. Results of M_{cr} in [kN.m].

6. Conclusions

The challenging problem of the critical moment of beams was addressed in this study using machine learning techniques to develop neural network models for simply supported beams. The Finite Element Method and Linear Buckling Analysis were performed to create a large database of results. These results were divided into a training and a validation dataset, each of them used to train and validate the developed neural networks, respectively. To create a robust dataset, a smart sampling strategy was implemented based on the random generation of samples with constraints to their geometric proportions. This innovative technique allowed to substantially decrease the number of samples needed to attain a very good accuracy and generalization ability of the model.

The training of the neural networks was accomplished by the Levenberg–Marquardt and Backpropagation algorithms. Several network architectures were trained and validated. Shallow and deep neural networks were considered, respectively, with one or two hidden layers, and it was concluded that the prediction accuracy increased when in the presence of a second hidden layer, while the training cost (time) was mostly influenced by the number of neurons in the first hidden layer. A neural network with 9 input neurons, 128×16 neurons in the (double) hidden layer and 1 neuron on the output layer was elected as the most appropriate model.

The accuracy of the selected model was compared against existing analytical methods for uniform beams with double or mono-symmetric section. This comparison confirmed that the proposed model surpasses the analytical ones when compared to the results of the Finite Element Method. For tapered beams, the nonexistence of simple analytical formulae may be overcome by the use of the developed model. In comparison with high-accuracy Finite Element Method results, the neural network led to good correlation providing an accurate methodology to estimate the critical moment.

Future developments to the proposed neural network model shall include *i*) beams with transverse loading such as uniformly distributed or point loads, *ii*) beams with other than fork-support boundary conditions, *iii*) including lateral supports along the beam and *iv*) composed loading including axial compression and combined end-moments with transverse loads.

Finally, this work confirms that the neural network models are

capable of predicting the critical moment of uniform and tapered beams, with double or mono-symmetric cross-section. Given the accuracy of these models together with their fast calculation times, it is likely that hybrid structural design methodologies are adopted in the near future. Machine learning models, such as neural networks, may be adopted during the early stage (preliminary) design to search for the optimal structural solutions, while the final design checks are made by traditional methods such as the Finite Element Method or analytical methods codified into design standards. This final design step is likely to remain necessary for years to come, until more appropriate design methods exist and, particularly, the safety and reliability of using machine learning design based methods are thoroughly investigated by the scientific community.

Data availability

The data required to reproduce these findings will be made available on request.

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.

Acknowledgements

This work is funded by national funds through FCT – Fundação para a Ciéncia e a Tecnologia, I.P., under the Scientific Employment Stimulus – Institutional Call – CECINST/00026/2018 and Rede Nacional de Computação Avançada (RNCA) exploratory research project CPCIA/1/6717/2020. The author wishes to express his sincere gratitude towards Professor Paulo Vila Real and Mr. Marc Braham for their invaluable sharing of knowledge that contributed to improve this work.

References

- [1] Bleich F. Buckling strength of metal structures, Mc Graw-Hill Book Company, Inc., Cardnr. 51-12588; 1952.
- [2] Prandtl L. Kipperscheinungen (phd thesis), Nürnberg; 1899.
- [3] Michell AG. Elastic stability of long beams under transverse forces. *Lon Edinb Dublin Philos Mag J Sci* 1899;48:298–309.
- [4] Timoshenko S. Sur la stabilité des systèmes élastiques, A. Dumas; 1913.
- [5] Vlasov VZ. Tonkostennye uprugie sterzhni (in Russian). Gosstroizdat; 1940.
- [6] Winter G. Lateral stability of unsymmetrical I-beams and trusses in bending. *Trans Am Soc Civ Eng* 1943;108:247–60.
- [7] Hill HN. Lateral buckling of channels and z-beams. *Trans Am Soc Civ Eng* 1954; 119:829–41.
- [8] Clark J, Hill H. Lateral buckling of beams. *J Struct Div* 1960;86:175–96.
- [9] Galambos TV. Inelastic lateral buckling of beams. *J Struct Div* 1963;89:217–42.
- [10] Yoo CH, Lee S. Stability of structures: principles and applications. Elsevier; 2011.
- [11] I. Baláz, Y. Koleková, Stability of monosymmetric beams, in: Proc. of the 6th International Colloquium “Stability and ductility of steel structures, pp. 57–64.
- [12] Koleková Y, Baláz I. LTB resistance of beams influenced by plastic reserve or local buckling. *Eng Mech* 2012:235.
- [13] Serna MA, López A, Puente I, Yong DJ. Equivalent uniform moment factors for lateral-torsional buckling of steel members. *J Constr Steel Res* 2006;62:566–80.
- [14] Braham M. Le déversement élastique des poutres en I à section monosymétrique soumises à un gradient de moment de flexion. *Construction Métallique* 2001;1: 17–28.
- [15] Mrázik A, Djubek K. Stability of Thin-Walled Member Structures, volume Research Report IX.3 S/58, ÚSTARCH - Slovak Academy of Sciences, 1958.
- [16] Djalaly H. calcul de la resistance ultime au déversement, *Construction Metallique*, CTICM 11; 1974.
- [17] CEN, ENV 1993-1-1, Eurocode 3: Design of steel Structures – Part 1-1: General Rules and Rules for Buildings; 1992.
- [18] CEN, EN 1993-1-1, Eurocode 3: Design of steel structures - Part 1-1: General rules and rules for buildings; 2005.
- [19] ECCS, Publication No. 119 - Rules for Member Stability in EN 1993-1-1: Background documentation and Design Guidelines; 2006.
- [20] Kitipornchai S, Wang CM, Trahair NS. Buckling of monosymmetric I-beams under moment gradient. *J Struct Eng* 1986;112:781–99.
- [21] Mohri F, Brouki A, Roth J. Déversement des poutres en I sous chargements asymétriques. *Revue Construction Métallique* 2000;2:43–52.

- [22] Mohri F, Brouki A, Roth J. Theoretical and numerical stability analyses of unrestrained, mono-symmetric thin-walled beams. *J Constr Steel Res* 2003;59: 63–90.
- [23] Mohammadi E, Hosseini SS, Rohanmanesh MS. Elastic lateral-torsional buckling strength and torsional bracing stiffness requirement for monosymmetric I-beams. *Thin-Walled Struct* 2016;104:116–25.
- [24] Achref H, Foudil M, Cherif B. Higher buckling and lateral buckling strength of unrestrained and braced thin-walled beams: Analytical, numerical and design approach applications. *J Constr Steel Res* 2019;155:1–19.
- [25] Andrade A, Camotim D. Lateral-torsional buckling of singly symmetric tapered beams: theory and applications. *J Eng Mech* 2005;131:586–97.
- [26] Yang Y-B, Yau J-D. Stability of beams with tapered I-sections. *J Eng Mech* 1987; 113:1337–57.
- [27] Bradford M. Stability of tapered I-beams. *J Constr Steel Res* 1988;9:195–216.
- [28] Braham M, Hanikenne D. Lateral buckling of web tapered beams: An original design method confronted with a computer simulation. *J Constr Steel Res* 1993;27: 23–36.
- [29] Gupta P, Wang S, Blandford G. Lateral-torsional buckling of nonprismatic I-beams. *J Struct Eng* 1996;122:748–55.
- [30] Andrade A, Camotim D, Dinis PB. Lateral-torsional buckling of singly symmetric web-tapered thin-walled I-beams: 1D model vs. shell FEA. *Comput Struct* 2007;85: 1343–59.
- [31] Zhang L, Tong GS. Lateral buckling of web-tapered I-beams: a new theory. *J Constr Steel Res* 2008;64:1379–93.
- [32] Asgarian B, Soltani M, Mohri F. Lateral-torsional buckling of tapered thin-walled beams with arbitrary cross-sections. *Thin-Walled Struct* 2013;62:96–108.
- [33] Serna M, Ibáñez J, López A. Elastic flexural buckling of non-uniform members: closed-form expression and equivalent load approach. *J Constr Steel Res* 2011;67: 1078–85.
- [34] Galéa Y. Deversement des barres à section en I bissymétriques et hauteur d'âme bilinéaire variable. *Construction Métallique* 1986;23:50–4.
- [35] AISC, AISc 360-16 Specifications for Structural Steel Buildings, Chicago; 2016.
- [36] Kaehler RC, White DW, Kim YD. Frame design using web-tapered members. American Institute of Steel Construction; 2011.
- [37] Standards Australia, AS4100-1998, Steel structures; 1998.
- [38] Marques L, Simões da Silva L, Greiner R, Rebelo C, Taras A. Development of a consistent design procedure for lateral-torsional buckling of tapered beams. *J Constr Steel Res* 2013;89:213–35.
- [39] Tankova T, Simões da Silva L, Marques L. Buckling resistance of non-uniform steel members based on stress utilization: General formulation. *J Constr Steel Res* 2018; 149:239–56.
- [40] Kucukler M, Gardner L. Design of laterally restrained web-tapered steel structures through a stiffness reduction method. *J Constr Steel Res* 2018;141:63–76.
- [41] Szalai J, Papp F, Hajdú G. Validation of the overall stability design methods (OSDM) for tapered members, SDSS 2019 - International Colloquium on Stability and Ductility of Steel Structures; 2019.
- [42] Maia É, Couto C, Vilà Real P, Lopes N. The General Method for the fire design of I-section web-tapered beams. *Thin-Walled Struct* 2021;169:108377.
- [43] Waszczyzny Z, et al. Neural networks in the analysis and design of structures. Springer; 1999.
- [44] Gholizadeh S, Pirmoz A, Attarnejad R. Assessment of load carrying capacity of castellated steel beams by neural networks. *J Constr Steel Res* 2011;67:770–9.
- [45] Abambres M, Rajana K, Tsavdaridis KD, Ribeiro TP. Neural network-based formula for the buckling load prediction of I-section cellular steel beams. *Computers* 2018; 8.
- [46] Sharifi Y, Moghbeli A, Hosseinpour M, Sharifi H. Neural networks for lateral torsional buckling strength assessment of cellular steel I-beams. *Adv Struct Eng* 2019;22:2192–202.
- [47] Sharifi Y, Moghbeli A, Hosseinpour M, Sharifi H. Study of Neural Network Models for the Ultimate Capacities of Cellular Steel Beams. *Iran J Sci Technol-Trans Civ Eng* 2020;44:579–89.
- [48] Hosseinpour M, Sharifi Y, Sharifi H. Neural network application for distortional buckling capacity assessment of castellated steel beams, in: *Structures*, vol. 27, Elsevier, p. 1174–83.
- [49] Rajana K, Tsavdaridis KD, Kotsakis E. Elastic and inelastic buckling of steel cellular beams under strong-axis bending. *Thin-Walled Struct* 2020;156.
- [50] Nguyen T-A, Ly H-B, Tran VQ. Investigation of ANN Architecture for Predicting Load-Carrying Capacity of Castellated Steel Beams. *Complexity* 2021;2021.
- [51] Limbachiya V, Shamass R. Application of artificial neural networks for web-post shear resistance of cellular steel beams. *Thin-Walled Struct* 2021;161:107414.
- [52] Ferreira FPV, Shamass R, Limbachiya V, Tsavdaridis KD, Martins CH. Lateral-torsional buckling resistance prediction model for steel cellular beams generated by artificial neural networks (ANN). *Thin-Walled Struct* 2022;170: 108592.
- [53] Pala M. A new formulation for distortional buckling stress in cold-formed steel members. *J Constr Steel Res* 2006;62:716–22.
- [54] Pala M, Caglar N. A parametric study for distortional buckling stress on cold-formed steel using a neural network. *J Constr Steel Res* 2007;63:686–91.
- [55] Dias J, Silvestre N. A neural network based closed-form solution for the distortional buckling of elliptical tubes. *Eng Struct* 2011;33:2015–24.
- [56] Tohidí S, Sharifi Y. A new predictive model for restrained distortional buckling strength of half-through bridge girders using artificial neural network. *KSCE J Civ Eng* 2016;20:1392–403.
- [57] Adeli H, Karim A. Neural network model for optimization of cold-formed steel beams. *J Struct Eng* 1997;123:1535–43.
- [58] Karim A, Adeli H. Global optimum design of cold-formed steel hat-shape beams. *Thin-Walled Struct* 1999;35:275–88.
- [59] Tashakori A, Adeli H. Optimum design of cold-formed steel space structures using neural dynamics model. *J Constr Steel Res* 2002;58:1545–66.
- [60] Adeli H, Park HS. A neural dynamics model for structural optimization—theory. *Comput Struct* 1995;57:383–90.
- [61] Adeli H, Park HS. Neurocomputing for design automation. CRC Press, Inc.; 1998.
- [62] Guzelbey IH, Cevik A, Gögiş MT. Prediction of rotation capacity of wide flange beams using neural networks. *J Constr Steel Res* 2006;62:950–61.
- [63] D'Aniello M, Güneyisi EM, Landolfo R, Mermerdaş K. Analytical prediction of available rotation capacity of cold-formed rectangular and square hollow section beams. *Thin-Walled Struct* 2014;77:141–52.
- [64] Güneyisi EM, D'Aniello M, Landolfo R, Mermerdaş K. Prediction of the flexural overstrength factor for steel beams using artificial neural network. *Steel Compos Struct* 2014;17:215–36.
- [65] D'Aniello M, Güneyisi EM, Landolfo R, Mermerdaş K. Predictive models of the flexural overstrength factor for steel thin-walled circular hollow section beams. *Thin-Walled Struct* 2015;94:67–78.
- [66] Tohidí S, Sharifi Y. Neural networks for inelastic distortional buckling capacity assessment of steel I-beams. *Thin-Walled Struct* 2015;94:359–71.
- [67] Tohidí S, Sharifi Y. Load-carrying capacity of locally corroded steel plate girder ends using artificial neural network. *Thin-Walled Struct* 2016;100:48–61.
- [68] Fonseca ET, Vellasco MM, Pacheco MAC, Vellasco PCDS, De Andrade SA. A neural network system for patch load prediction. *J Intell Rob Syst* 2001;31:185–200.
- [69] Fonseca ET, Vellasco PCDS, de Andrade SA, Vellasco MM. A patch load parametric analysis using neural networks. *J Constr Steel Res* 2003;59:251–67.
- [70] El-Kassis E, Mackie R, El-Sheikh A. Using neural networks in cold-formed steel design. *Comput Struct* 2001;79:1687–96.
- [71] Lyu Z, Zhang J, Zhao N, Xiang Q, Song Y, Li J. A comparative study on the performance of FEM, RA and ANN methods in strength prediction of pallet-rack stub columns. *Int J Steel Struct* 2020;20:1509–26.
- [72] Fang Z, Roy K, Mares J, Sham C-W, Chen B, Lim JB. Deep learning-based axial capacity prediction for cold-formed steel channel sections using deep belief network, in: *Structures*, vol. 33, Elsevier, p. 2792–802.
- [73] Fang Z, Roy K, Chen B, Sham C-W, Hajirasouliha I, Lim JB. 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* 2021; 166:108076.
- [74] Xu Y, Zheng B, Zhang M. Capacity prediction of cold-formed stainless steel tubular columns using machine learning methods. *J Constr Steel Res* 2021;182:106682.
- [75] Nguyen T-H, Tran N-L, Nguyen D-D. Prediction of Axial Compression Capacity of Cold-Formed Steel Oval Hollow Section Columns Using ANN and ANFIS Models. *Int J Steel Struct* 2021.
- [76] Nguyen T-H, Tran N-L, Nguyen D-D. Prediction of critical buckling load of web tapered I-section steel columns using artificial neural networks. *Int J Steel Struct* 2021;1:1–23.
- [77] ANSYS, Academic Research Mechanical, Release 18.2; 2018.
- [78] Snijder H, Van der Aa R, Hofmeyer H, van Hove B. Lateral torsional buckling design imperfections for use in non-linear FEA. *Steel Constr* 2018;11:49–56.
- [79] Loyola DG, Pedregnana RM, Gimeno García S. Smart sampling and incremental function learning for very large high dimensional data, *Neural Networks* 2016;78: 75–87. Special Issue on Neural Network Learning in Big Data.
- [80] AASHTO, LRFD Bridge Design Specifications; 2016.
- [81] Werbos PJ. Beyond regression: New tools for prediction and analysis in the behavior science, Ph. D. thesis, Harvard University; 1974.
- [82] Werbos PJ. The roots of backpropagation: from ordered derivatives to neural networks and political forecasting, vol. 1. Wiley; 1994.
- [83] Levenberg K. A method for the solution of certain non-linear problems in least squares. *Q Appl Math* 1944;2:164–8.
- [84] Marquardt DW. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *J Soc Ind Appl Math* 1963;11:431–41.
- [85] Hagan M, Menhaj M. Training feedforward networks with the Marquardt algorithm. *IEEE Trans Neural Netw* 1994;5:989–93.
- [86] Atabay D. pyrenn: a recurrent neural network toolbox for Python and Matlab, Institute for Energy Economy and Application Technology, Technische Universität, München,<http://pyrenn.readthedocs.io/en/latest/>; 2018.