Machine Learning-Driven Web-Post Buckling Resistance Prediction for High-Strength Steel Beams with elliptically-based web openings

Musab Rabi [a], Yazeed S. Jweihan [b], Ikram Abarkan [c], Felipe Piana Vendramell Ferreira[d], Rabee Shamass [e], Vireen Limbachiya [f], Konstantinos Daniel Tsavdaridis [g], Luis Fernando Pinho Santos [f]

[a] Dept of Civil Engineering, Jerash University, Jerash, Jordan, 26150.

[b] Civil and Environmental Engineering Department, College of Engineering, Mutah University, Mutah, Karak, Jordan, 61710, P.O. BOX 7

[c] Department of Physics, Faculty of Sciences, Abdelmalek Essaâdi University, 93002 Tetouan, Morocco

[d] Faculty of Civil Engineering, Federal University of Uberlândia, – Campus Santa Mônica, Uberlândia, Minas Gerais, Brazil

[e] Department of Civil and Environmental Engineering, Brunel University London, London, UK

[f] Division of Civil and Building Services Engineering, School of Build Environment and Architecture, London South Bank University, UK

[g] Department of Engineering, School of Science and Technology, City, University of London, Northampton Square, EC1V 0HB, London, UK

Abstract

The use of periodical elliptically-based web (EBW) openings in high strength steel (HSS) beams has been increasingly popular in recent years mainly because of the high strength-to-weight ratio and the reduction in the floor height as a result of allowing different utility services to pass through the web openings. However, these sections are susceptible to web-post buckling (WPB) failure mode and therefore it is imperative that an accurate design tool is made available for prediction of the web-post buckling capacity. Therefore, the present paper aims to implement the power of various machine learning (ML) methods for prediction of the WPB capacity in HSS beams with (EBW) openings and to assess the performance of existing analytical design model. For this purpose, a numerical model is developed and validated with the aim of conducting a total of 10764 web-post finite element models, considering S460, S690 and S960 steel grades. This data is employed to train and validate different ML algorithms including Artificial Neural Networks (ANN), Support Vector Machine Regression (SVR) and Gene Expression Programming (GEP). Finally, the paper proposes new design models for WPB resistance prediction. The results are discussed in detail, and they are compared with the numerical models and the existing analytical design method. The proposed design models based on the machine learning predictions are shown to be powerful, reliable and efficient design tools for capacity predictions of the WPB resistance of HSS beams with periodical (EBW) openings.

Keywords: Finite Element Modelling, Web-post buckling resistance, Elliptically-based web openings, High strength steel beams, Artificial Neural Network, Gene Expression Programming, Support Vector Machine Regression

1. Introduction

Steel beams with periodical web openings (i.e., castellated, cellular and Angelinas) can be used in multi-story building designs, since they have many advantages such as the reduction in the structure's self-weight and the floor height since as a result of allowing different utility services to pass through the web openings [1, 2]. The present study focuses on S460, S690, S960 grade high strength steel beams (HSS) beams comprising periodical elliptically-based web (EBW) openings. The manufacturing and castellation process include three main steps: thermal cutting, shifting and welding [3]. These procedures result in an increase in inertia about the strong axis leading to improve flexural stiffness. Different from circular and hexagonal shapes, the elliptically-based web openings benefit from the stress redistribution around the neutral axis leading to increase the bearing capacity of the beam [4].

Although steel beams with periodical (EBW) openings have many attractive characteristics, as exemplified previously, new buckling mode can be developed, depending mainly on the geometric characteristics of these beams. In this context, these structures can achieve the following buckling modes: local flange and web, lateral-torsional, lateral-distortional and web-post (WPB) [5–9]. The present paper focuses on the latter, WPB, which is a local failure mode characterized by S-shape double curvature owing to the horizontal shear stresses developed in the web-post region [10, 14]. The key influential parameters that affect the WPB resistance are the web thickness, the web-post width and, the opening height [15, 16].

The use of HSS (i.e. with yield stress (f_y) is greater than 460 MPa) has been increasing in structural systems owing to the exceptional benefits, compared with normal strength, including higher strength/weight ratio, longer spans and lighter sections as well as less carbon footprint [4, 17-20]. Mela and Heinisuo [21] reported that designing in HSS may achieve a 34% savings in materials as function of lightweight structures meeting sustainability criteria [22–24]. Other HSS advantages can also be highlighted as greater corrosion resistance and consequently

improving the durability and reducing costs associated with regular maintenance and inspections [25–27].

There are several studies focus on steel beams with (EBW) openings including normal strength steel [6, 15, 28–32] and high strength steel [4]. For normal strength steels, Tsavdaridis and D'Mello [29, 30, 32] and Tsavdaridis et al. [31] presented various optimization techniques considering several opening shapes (i.e. circular, elongated, elliptical, and hexagonal). These studies showed that the elliptically-based opening shape presented greater resistance to plasticization mechanisms and lower deflections compared with other shapes. In Tsavdaridis and D'Mello [6] three-point bending tests were performed on steel beams comprising various web opening shapes. It was shown that steel beams with elliptically-based opening shape had greater WPB resistance when compared to the others one. Ferreira et al. [15] carried out a parametric study in web-post finite element models. In Shamass et al. [16], the capacity of WPB resistance in normal strength beams with EBW openings was studied using Artificial Neural Network (ANN) model. The authors showed that geometrical parameters of the steel profile such as the web thickness and height had positive influence on the WPB capacity, while the geometrical parameters of the EBW opening including the radius, width and height had negative impact reflecting inverse relationship. Very limited research is presented on the HSS beams with EBW openings by Ferreira et al. [4]. In this paper, an analytical design approach was proposed for WPB capacity prediction based on the strut model given in Eurocode 3 (EC3) [33]. This procedure will be presented in Section 2 of this study. With the presentation of this background considering HSS steel beams with EBW opening, it is noted that the studies are scarce.

Currently, machine learning (ML) algorithms are widely used in solving engineering problems of structural members, mainly for design and verification issues, considering steel beams with web openings [5, 12, 16, 34–43].

Machine learning models can analyse large and intricate datasets, accommodating nonlinear relationships, and adapt to various complexities in beam

geometries. Furthermore, machine learning can provide more precise predictions, reducing the need for excessive overdesign and potentially resulting in cost savings without compromising safety. The present work focuses on Artificial Neural Networks (ANN), Support Vector Machine Regression (SVR) and Gene Expression Programming (GEP) algorithms. The ANN has become one of the most popular ML algorithms, and it was invented by Rosenblatt [44]. An ANN model consists of interlinked nodes, like the human brain, displayed in three main layers including input, hidden, and output layers [45-47]. An ANN projected with more than three hidden layers is known as deep learning. The SVR was developed by Vapnik [48]. According to Tinoco et al. [49], the SVR implements nonlinear kernel function to find an optimal hyper-plane that most accurately fits the training data while enabling a certain margin of error [50, 51]. The GEP was originally proposed by Ferreira [52] for solving complex mathematical problems. The GEP is constructed of different individuals known as chromosomes that form a simple expression tree (ETs) of several genes or "sub-ETs". These genes are linked together by an assigned mathematical functions such as subtraction, addition, division, or multiplication [53]. A review of machine learning for structural engineering is found in Huu-Tai [54]. In this context, the present paper aims to develop machine learning techniques for predicting the WPB capacity of HSS beams with EBW openings. For this task, a database of 10764 web-post finite element models, which are consider S460, S690 and S960 steel grades, is used to train ANN, SVR and GEP algorithms. The results are discussed in detail, and compared with the results from the numerical models and the analytical method proposed by Ferreira et al. [4].

2. Ferreira et al. based on EC3 calculations

A new design approach was proposed by Ferreira et al. [4] to predict the WPB capacity of HSS beams with EBW openings. This approach is basically an extension of the previously developed model for normal strength steel beams [15]. The model considers the web-post effective length as a compressed bar. The procedure is based

on EC3 [33] considering the buckling curve *c*, similar to that presented in SCI P355 [55]. The methodology is described in Eqs. (1-10).

$$l_{eff} = k\sqrt{\left(\frac{d_o - 2R}{2}\right)^2 + \left(\frac{s}{2} - R\right)^2} \quad (1)$$

$$k = 0.516 - 0.288\left(\frac{H}{d_o}\right) + 0.062\left(\frac{s}{s-w}\right) + 2.384\left(\frac{s}{d_o}\right) - 2.906\left(\frac{w}{d_o}\right) \quad (2)$$

$$\lambda_w = \frac{l_{eff}\sqrt{12}}{t_w} \quad (3)$$

$$f_{cr,w} = \frac{\pi^2 E}{{\lambda_w}^2} \quad (4)$$

$$\lambda_0 = \sqrt{\frac{f_y}{f_{cr,w}}} \quad (5)$$

$$\phi = 0.5\left[1 + 0.49(\lambda_0 - 0.2) + {\lambda_0}^2\right] \quad (6)$$

$$\chi = \frac{1}{\phi + \sqrt{\phi^2 - {\lambda_0}^2}} \leq 1.0 \quad (7)$$

$$\sigma_{Rk} = K_{HSS}\chi f_y \quad (8)$$

$$K_{HSS} = -1.45 + 1.61\left(\frac{H}{d_o}\right) + 0.33\left(\frac{s}{s-w}\right) - 0.90\left(\frac{s}{d_o}\right) + 0.21\left(\frac{w}{d_o}\right) - 0.004\left(\frac{d_o}{t_w}\right) + 0.49\lambda_0 \quad (9)$$

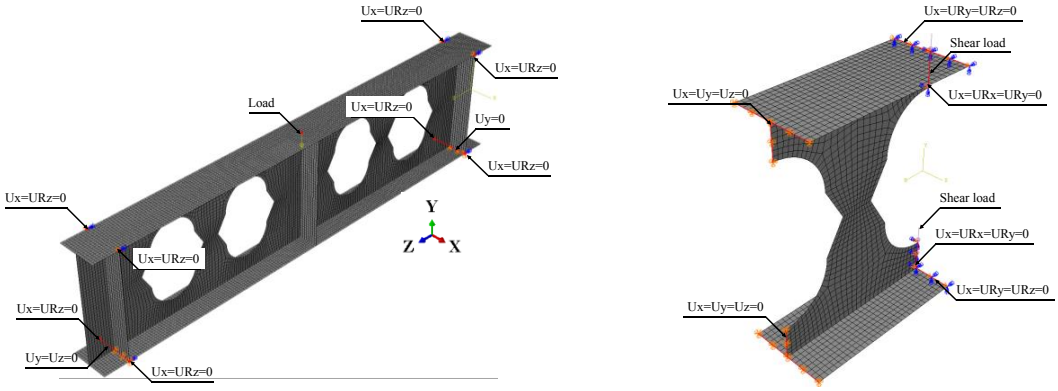
$$V_{EC3} = \sigma_{Rk} t_w (s - w) \quad ((0)$$

In these expressions, d_o, R and w are the height, radius and width of the opening, respectively, s, λ_w, l_{eff} and t_w are the web-post width, the web-post slenderness factor, the web-post effective length and the web thickness, respectively. H is the distance measured from the centres of flanges after castellation process, $f_{cr,w}$ is the critical stress in the web-post, λ_0 and χ are a reduced slenderness factor and a reduction factor, respectively.

3. Finite element method

This section discusses the development of the numerical model conducted previously by Ferreira et al. [15] for steel beams with elliptically-based web openings using the ABAQUS software [56]. A similar approach of the previously validated model is employed in this study. The model demonstrated very good depiction of the experimental ultimate strength capacity with the mean and standard deviation are 1 and 6.9%, respectively. Similar observation is found in terms of the load-mid span displacement curve, maximum loading capacity, failure mode and vertical shear resistance [15]. The validation was conducted for both the full and web-post models similar to that previously used in several research [4, 9, 11, 13-16, 57-60]. The development and validation of the numerical model is concisely presented herein.

In order to conduct a robust numerical model, allowing for the identification of the failure mechanism of the WPB, full and single web-post models are established. The models are developed by performing a Buckling and post-buckling analyses with initial geometric imperfection of $d_g/500$, (dg denotes for the beam height). This value was found to be appropriate for steel beams with periodical perforated web openings [13, 15, 16], given the complexity of estimating the geometric imperfection owing to the manufacturing and castellation process. Full and web-post models are discretised with S4R shell element and reduced integration [15, 16, 61, 62]. Following the mesh sensitivity study, a 10 mm element size is shown to be the most appropriate element size for all beams, providing an accurate predictions of the experimental results. The boundary conditions for both the full and web-post models are illustrated in Fig. 1a and Fig. 1b, respectively. The constitutive behaviour of steel in the beam model is represented using a multi-linear constitutive model [63, 64]. A value of 200 GPa and 0.3 are taken for the Young's modulus and Poisson's ratio, respectively.



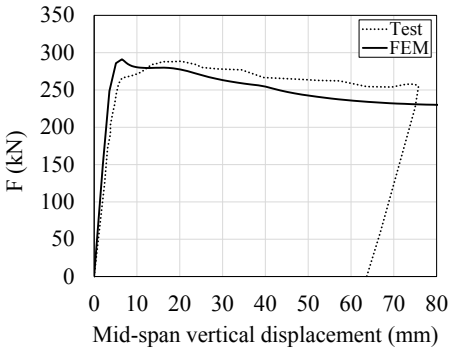
(a) Full model (b) Web-post model

Fig. 1: Discretization and boundary conditions

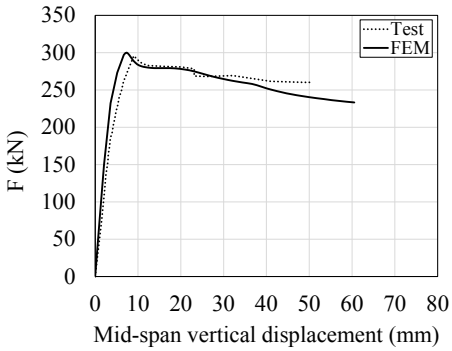
3.1 Model validation

The finite element model is validated using experimental tests reported in [6], particularly specimens A1, A2, B1, B2 and B3. It is noteworthy to indicate that only normal steel beams is employed in the validation process given the scarce of the experimental tests on high-strength perforated steel beams with EBW openings.

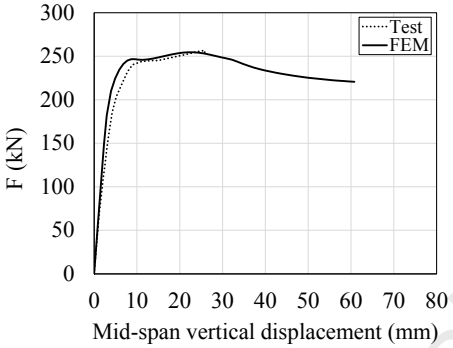
The validation results for the full and web-post models are presented in Fig. 2 and Table 1, respectively. Regarding the full models, the results are presented considering load-displacement relationships. The maximum and minimum relative errors of the test-to-finite element ratio were 0.8% and -5.1%, respectively. Whereas, the results of the web-post models are shown by global shear. All models had the load-bearing capacity governed by web-post buckling, similar to the response observed in tests [6]. Based on the results presented in this section, it is possible to conclude that the FEM is validated and is capable of providing accurate predictions.



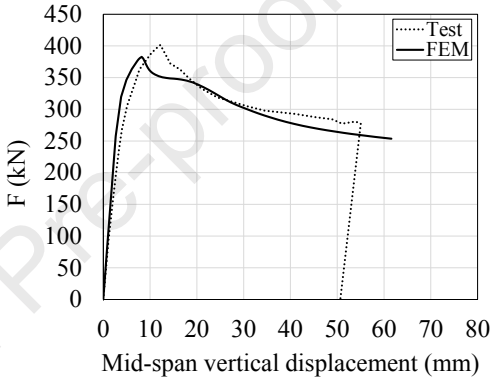
(a) A1



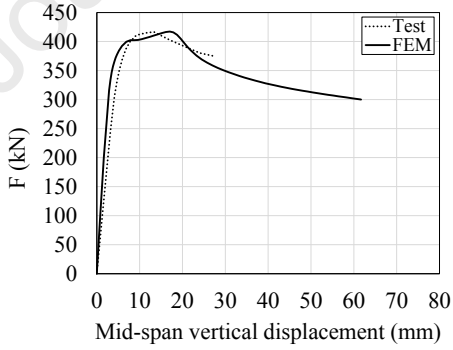
(b) A2



(c) B1



(d) B2



(e) B3

Fig. 2: Load-displacement curves obtained from the numerical models and corresponding tests, considering full models

Table 1: WPB resistance, considering the web-post models

Model	V_{TEST} (kN)	V_{FE} (kN)	V_{TEST}/V_{FE}
A1	144.4	157	0.92
A2	149	159	0.94
B1	127.5	121	1.05
B2	201.2	200.5	1.00
B3	207.5	188	1.10
		Average	1.00
		S.D	6.93%
		CoV	6.90%

3.2 Parametric study

Building on the discussion presented previously, the web-post models can be used to study the WPB of perforated HSS beams with EBW openings. The parametric study includes a various range of the geometric parameters of the steel beam such as t_f (flange thickness), b_f (flange width), H, t_w, d_o, w, R., as show in Fig. 3. To automate the pre-processing, processing and post-processing, a Python script is developed. At total, 10764 web-post models are used considering three different HSS grades, S460, S690 and S960. The parameters investigated are shown in Fig. 4.

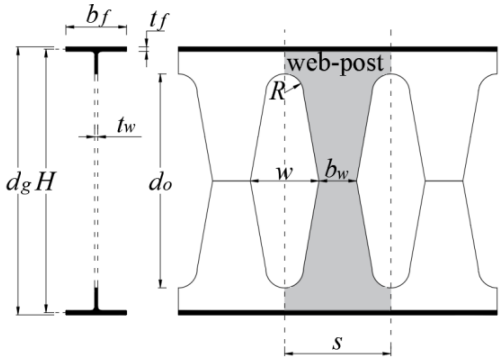


Fig. 3: Geometrical parameters of perforated steel beams with EBW openings

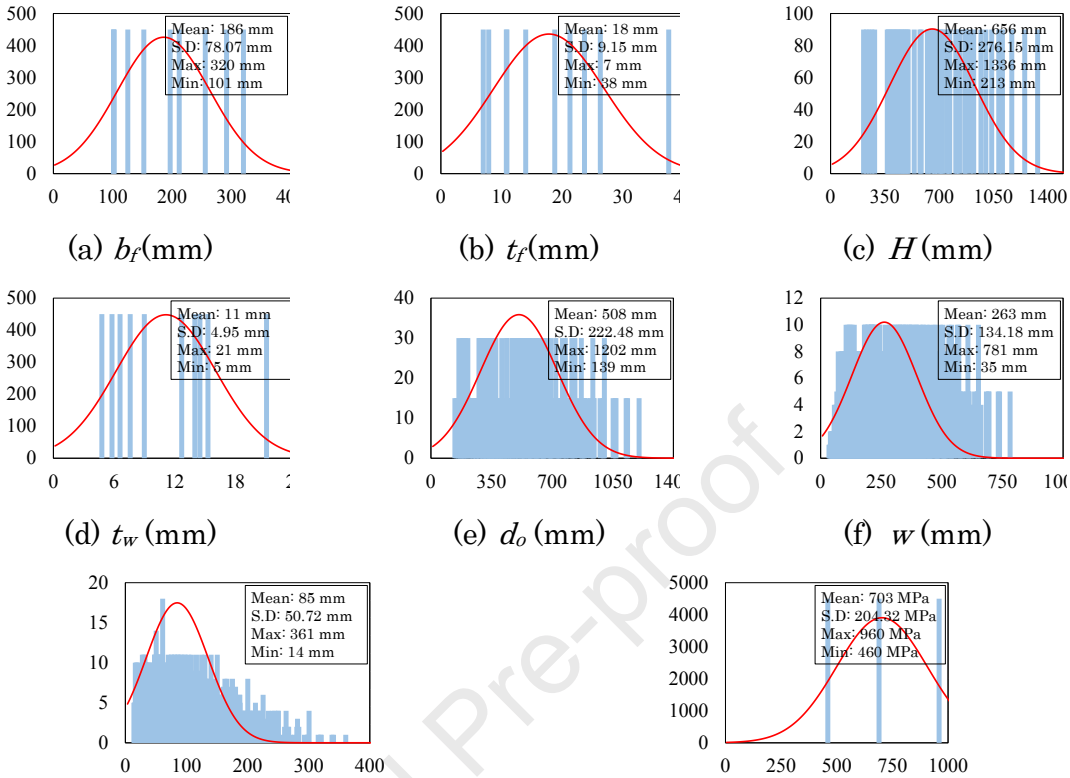


Fig. 4: Normal distribution considering number of models per parameter investigated

4. Machine learning methods

In this section, three machine learning models will be presented, considering Artificial Neural Network (ANN), Support Vector Machines (SVR) and Gene Expression Programming (GEP).

4.1. Artificial Neural Network

4.1.1. Neural Network Architecture

In this study, a Multi-Layer Perceptron Network (MLPN) that solved an input-output fitting problem was employed to predict WPB load of the HSS beams with elliptical web opening. A typical neural network architecture with three neurones

is presented in Fig. 5. In this paper, the ANN model is developed with a shallow structure consisting of a single hidden layer in order to promote generalization and avoid over fitting issues [65-68]. In this study, the input parameters used are d_o, H, R, t_w, w and f_y, while the output is the WPB load (V_{ANN}). The primary aim of the neural network is to allocate weights to the hidden layer neurons and introduce bias values in both the hidden and output layers. This process is crucial for establishing the connections between the input and output parameters.

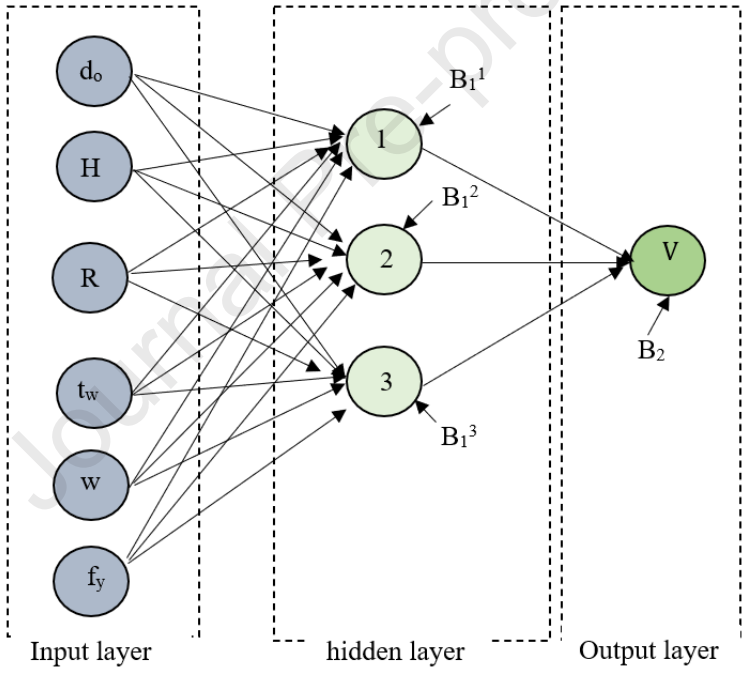


Fig. 5: A typical neural network architecture with three neurones

4.1.2. Setting up Artificial Neural Network

The ANN initially categorizes the data into three groups. The training data (70%) is utilized to train the model and refine it based on errors. Validation data (15%) serves to evaluate the network's generalization and halt the training when no

further improvements are observed. The testing data (15%) remains independent, having not been used in training or validation, and is employed to assess the model's accuracy separately [69]. To address the high accuracy requirements for small to medium-sized problems, the Levenberg-Marquardt back-propagation training algorithm was selected [36, 39, 40].

All data variables were subjected to normalization using the mapminmax method, which was employed to range the data from -1 to 1. This process involved applying Eq. (11) to all inputs and outputs.

$$x_i^n = \frac{(y_{max} - y_{min})(x_i - x_i^{min})}{x_i^{max} - x_i^{min}} + y_{min} \tag{11}$$

where x_i and x_i^n are the actual value for inputs or output and the corresponding normalized value respectively. The minimum and maximum values for the inputs or output are denoted as x_i^{min} and x_i^{max}, the y_{min} and y_{max} are taken by default as -1 and +1 for each row of X, respectively. Table 2 summarizes the range of the data employed from the parametric study for he input and output parameters.

Table 2: the range of the data employed from the parametric study for The range for the input and output parameters

	$\boldsymbol{H}$ (mm)	$\boldsymbol{t_w}$ (mm)	$\boldsymbol{d_o}$(mm)	$\boldsymbol{w}$(mm)	$\boldsymbol{R}$(mm)	$\boldsymbol{f_y}$ (Mpa)	$\boldsymbol{V}$(kN)
Upper limit	1335.8	21.1	1202.3	781.5	360.7	960	4212.8
Lower limit	213.4	4.8	138.7	34.7	13.9	460	48.7

In order to find the optimum number of neurons in the hidden layer, four ANN models were developed with 4, 6, 8 and 10 nuerons. The performance of these models is then assessed. The hyperbolic tangent transfer function [70] needed to predict the output parameter based normalised input values is given in as follows:

$$0_s = B_2^s + \sum_{k=1}^{r} w_{k,l}^{ho} \cdot \left(\frac{2}{1 + e^{-2H_k}} - 1\right) \tag{12}$$

$$H_k = B_1^k + \sum_{j=1}^{q} w_{j,k}^{ih} . I_j \quad (13)$$

Where, O_s and q denote the normalized output value and the number of input parameters; s and r are number of output parameters and the number of hidden neurons; $w_{j,k}^{ih}$ is the weights of the connection between I_j and H_k; $w_{k,l}^{oh}$ are the weights of the connection between H_k and O_s; B_1^s and B_2^k are the biases of s^{th} output neuron and k^{th} hidden neuron (H_k), respectively.

4.1.3. Assessing Accuracy of Neural Network Output

The accuracy of the ANN model is assessed by comparing the predicted values with the target values using different statistical measures including the coefficient of determination (R^2), Root Mean Square Error ($RMSE$) and Mean Absolute Error (MAE). These measures are determined as follows::

$$R^2 = 1 - \frac{\sum_{i=1}^{m}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{m}(y_i - \bar{y})^2} \quad (14a)$$

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2} \quad (14b)$$

$$MAE = \frac{1}{m}\sum_{i=1}^{m}|y_i - \hat{y}_i| \quad (14c)$$

where, m is the number of data point, y_i and $\hat{y}_i$ denote the i^{th} actual and predicted output respectively, $\bar{y}$ represents the average of y_i. To achieve excellent ANN model accuracy, the value of (R^2) should be close to 1, and ($RMSE$) and (MAE) must be minimal.

To further evaluate the model's accuracy, the influence of each input parameter was examined using the Garson Algorithm and the Connection Weight Approach. The later considers raw connection weights, identifying for the direction and the contribution that inputs might have on the output [71]. A positive impact indicates that increasing the input parameter will increase the output parameter's value, while a negative impact implies the opposite relationship. Eq. (15) outlines the

calculation process for the Connection Weight Approach, where $Input_x$ signifies importance, Hidden denotes the hidden-output connection weights and XY represents the input-hidden connection weights.

The contributions of inputs are also calculated through Garson's algorithm. It worth noting that Garson's algorithm does not identify the direction of the relationship between the input and output since it uses absolute values of connection weights, Eq. (16) illustrates the Garson Algorithm's calculation process [72].

In these equations, the subscripts n, k and m refer to the output, input and hidden neurons, whereas o, h and I refer to output, hidden and input layers, respectively. N_h and N_i and denote the numbers of neurons in the hidden and input layers, respectively; w is connection weights.

$$\mathrm{Input}_x = \sum_{Y=A}^{E} \mathrm{Hidden}_{XY} \tag{15}$$

$$I_j = \frac{\sum_{m=1}^{m=Nh}\left(\frac{w_{jm}^{ih}}{\sum_{k=1}^{Ni} w_{km}^{ih}} w_{mn}^{ho}\right)}{\sum_{k=1}^{k=Ni}\left[\sum_{m=1}^{m=Nh}\left(\frac{w_{km}^{ih}}{\sum_{k=1}^{Ni} w_{km}^{ih}} w_{mn}^{ho}\right)\right]} \tag{16}$$

4.2. Support Vector Machines

4.2.1. Overview

Support Vector Machine (SVM) is a machine learning method developed by Vapnik [48] and has gained popularity due to its outstanding performance [73]. It was initially implemented for classification purpose and known as Support Vector Classification (SVC), then used to handle regression problems under the name Support Vector Regression (SVR). The primary concern in SRV is to identify an optimal hyper plane that matches the training data while allowing a certain degree of error "ε", as shown in Fig. 6. All predictions inside the ε-insensitive tube are expected to have a tolerable error relative to the target, while any deviation outside this tube is penalized [74]. Support vector is a subset of the training data that is

crucial for determining and making accurate predictions. However, it's important to note that having more support vectors can improve accuracy but also increases computational time.

In many cases, the relationship between inputs and the output is not linear in the original input space. Therefore, SVR employs the kernel trick to map the data into a higher-dimensional space where linearity can be achieved. This is accomplished using kernel functions (i.e. sigmoid, polynomial and Radial Basis Function (RBF), to enable SVR to handle complex non-linear data. A general overview of the mathematical background of SVR is presented. Further information is available in [75].

The prediction model used in the SVR should be based on the following function:

$$f(x) = w^T \phi(x) + b \tag{17}$$

By minimizing the subsequent objective function:

$$\min_{w,b} \quad \frac{1}{2}\|w\|^2 \qquad \text{Subject to : } \begin{matrix} y_i - w^T\psi(x_i) - b \le \varepsilon \\ w^T\psi(x_i) + b - y_i \le \varepsilon \end{matrix} \tag{18}$$

Here, w and b represent the parameters of the regression function, while $\phi(x)$ denotes a non-linear function that maps the input data into a higher-dimensional feature space

In certain scenarios, the optimization problem could not be solved (i.e. Eq. (18)) due to the strict constraints, which means that all data points must be entirely within the ε-tube. Consequently, any violation beyond the margin makes it infeasible [75, 76]. To address this challenge, Cortes and Vapnik [77] introduced the concept of a 'soft margin,' allowing for a degree of error tolerance within the model. This is achieved by introducing slack variables ξ_i and ξ_i', representing upper and lower training deviations, respectively, outside the ε-insensitive tube, as illustrated in Fig.6. Consequently, the optimization problem can be reformulated as follows:

$$\min_{w,b,\xi_i,\xi_i'} \ \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i') \tag{19}$$

$$\text{Subjected to}: \begin{matrix} y_i - w^T\psi(x_i) - b \le \varepsilon + \xi_i \\ w^T\psi(x_i) + b - y_i \le \varepsilon + \xi_i' \end{matrix}$$

$$\xi_i, \xi_i' \ge 0; \qquad i = 1,2,3\ldots, N$$

In this equation, C is a positive hyper parameter that provides excellent balance compromise between the flatness of the regression function f and the error tolerance ε [75, 76]. These two hyper parameters must be optimized during the training process using well-established tuning techniques like grid search or random search to achieve optimal performance for the SVR model.

Alternatively, the optimization problem can be redefined using Lagrange multipliers α_i and α_i^* as follows:

$$\max_{\alpha_i,\alpha_i^*} \sum_{i=1}^{N} y_i(\alpha_i - \alpha_i^*) - \varepsilon\sum_{i=1}^{N}(\alpha_i + \alpha_i^*) - \frac{1}{2}\sum_{i,j=1}^{N}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)\phi(x_i)^T\phi(x_j) \tag{20}$$

$$\text{Subjected to: } \sum_{i=1}^{N}(\alpha_i - \alpha_i^*) = 0$$

$$\alpha_i, \alpha_i' \in [0, C]; \qquad i = 1,2,3\ldots, N$$

The prediction model can be reformulated as follows:

$$f(x) = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*)\phi(x_i)^T.\phi(x) + b \tag{21}$$

$$\text{where;} \quad w = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*)\phi(x_i)^T$$

The solution for the Lagrange multipliers $(\alpha_i - \alpha_i^*)$ can be either zero or non-zero. The non-zero terms correspond to the support vectors, which play a fundamental role in defining the final regression function and can be re-expressed as:

$$f(x) = \sum_{k=1}^{nSVs}(\alpha_k - \alpha_k^*)K(x_k, x) + b \tag{22}$$

Where x is the input vector, x_k represents the support vectors, $nSVs$ is the number of support vectors,$K(x_k, x) = \phi(x_k)^T.\phi(x)$ represents the kernel function and b is the bias Eqs. (23 a-d) represent the main kernel function:

Linear kernel:

$$K(x_k, x) = x_k^T x \tag{23-a}$$

Polynomial kernel:

$$K(x_k, x) = (x_k^T x + r)^d \tag{23-b}$$

RBF kernel:

$$K(x_k, x) = exp(-\gamma \|x_k - x\|^2) \tag{23-c}$$

Sigmoid kernel:

$$K(x_k, x) = \tanh(\gamma x_k^T x + r) \tag{23-d}$$

Where γ is a coefficient, r is an offset, and d indicates the degree of the polynomial kernel. This study exclusively employs the (RBF) kernel due to its proven success in numerous applications, as evidenced by references [78-80].

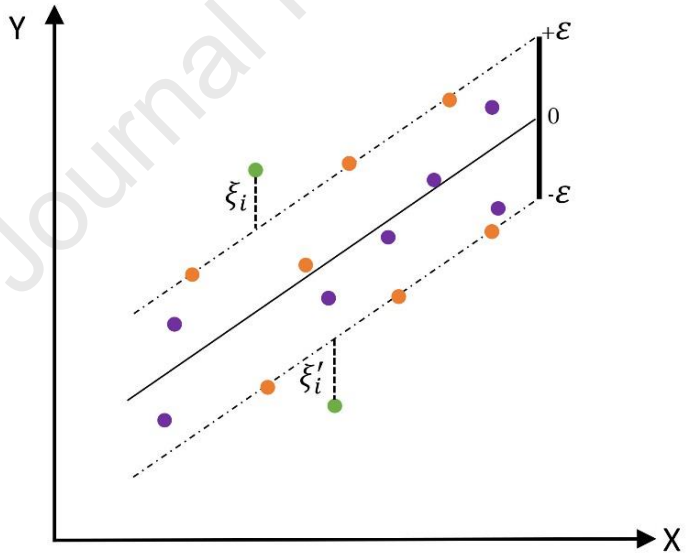


Fig. 6: ε-insensitive tube for linear SVR.

4.1.1. Data preparation

The SVR model is built using a total of 10764 data points extracted from the parametric study utilizing the *fitrsvm* function from the Statistics and Machine Learning Toolbox in MATLAB [81]. The data are divided into three groups including

342 training, validation and testing data. 70% of the data is set to train the SVR model,
343 while 15% of the data is allocated for validation purpose, and the remaining 15% is
344 reserved for testing with an unseen dataset. To address differences in units and
345 quantity limits, both input and output data have been normalized in a range of -1 to
346 1, using Eq. (11). The input parameters used are d_o, H, R, t_w, w, and f_y, while the
347 output is the WPB load (V_{SVR}).

348 *4.1.1.1. Grid Search*

349 In this study, the grid search method with 10-fold cross-validation is employed
350 to optimize hyper parameters for the RBF kernel, specifically C,ε and γ. While this
351 tuning technique is exhaustive and computationally time-consuming, requiring
352 substantial resources. Nevertheless, it remains widely adopted due to its established
353 accuracy [82]. To optimize computational efficiency, a two-step grid search approach
354 is implemented, following the practical guide [83]. The first step involves a coarse
355 grid search with a large interval range and step size. The second step refines the
356 search within the interval where the optimal hyper parameters are identified during
357 the coarse grid search. The initial range considered for C spans from 10^{-3} to 10^{5}, ε
358 ranges from 10^{-9} to 10^{-1}, and γ varies from 10^{-7} to 10^{1}, all with a step size of 1 in
359 logarithmic scale. In the finer grid search, this step size is further reduced to 0.25.
360 For each set of hyper parameters, a 10-fold cross-validation is applied, which involves
361 partitioning the training dataset into 10 equal folds. In each iteration, one fold serves
362 as the test set while the remaining nine folds are used for training, as depicted in Fig.
363 7. This operation is repeated for 10 iterations, for each hyper parameter combination
364 the average performance measures are collected. Subsequently, the combination that
365 provides the lowest $RMSE$ and the highest R^2 is selected as the optimal solution.

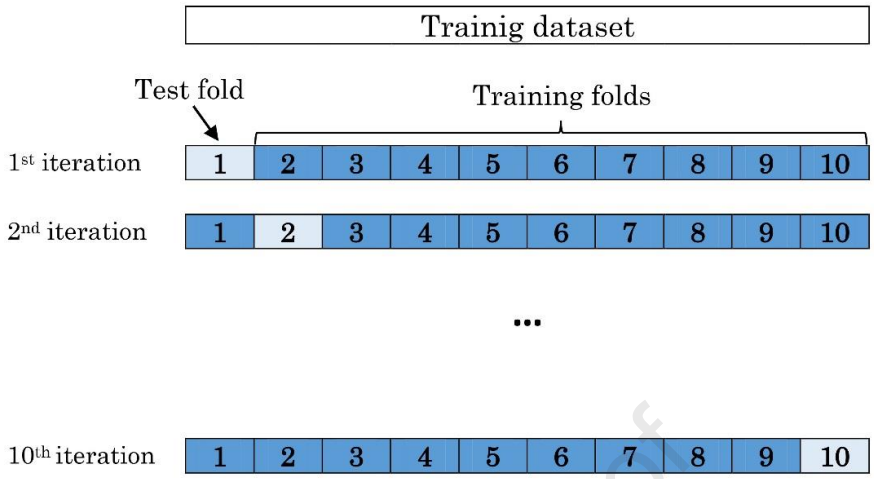


Fig. 7: An illustration of the concept of 10-fold cross-validation.

4.1.3 Accuracy assessment

The performance of SVR model has been assessed using statistical key measures, including the coefficient of determination (R^2), the Root Mean Square Error ($RMSE$) and the Mean Absolute Error (MAE), as expressed in Eqs. (14a), (14b) and (14c), respectively. The correlation between inputs and outputs is also assessed using Pearson's method [84]. This provides a clear understanding of the strength and direction of the linear interdependence between inputs and their influence on the output. Pearson's correlation coefficient (ρ) can be calculated as follows:

$$\rho = \frac{\sum_i (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_i (X_i - \bar{X})^2 \sum_i (Y_i - \bar{Y})^2}} \tag{24}$$

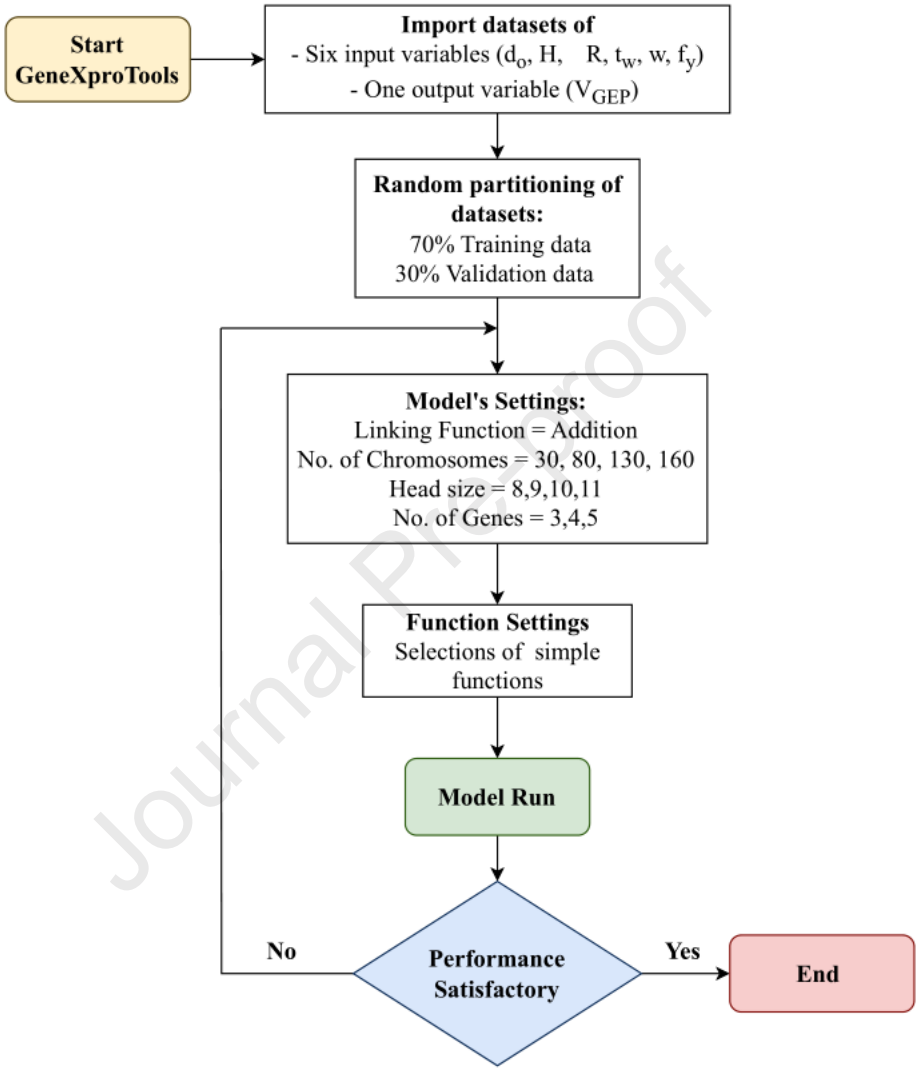
Where X_i and Y_i represent the i^{th} values of the variable X and Y, respectively, $\bar{X}$ and $\bar{Y}$ are the averages of the variables X and Y, respectively. Pearson's correlation coefficient falls within the range of -1 to 1. In other words, when ρ is equal to 1, it suggests a strong positive linear relationship between X and Y, implying that as X increases, Y increases. Conversely, when ρ is equal to -1, it indicates a strong negative linear relationship, meaning that as X increases, Y decreases. When ρ is 0, it indicates no linear correlation between the two variables, X and Y. In this case,

there is no consistent linear pattern between the variables, and they are considered uncorrelated.

4.3. Gene Expression Programming (GEP)

4.1.1 Overview

GEP is an artificial intelligence-based technique that was originally developed by Ferreira [52] for solving complex mathematical problems. A GEP model is constructed of different individuals known as chromosomes that form a simple expression tree (ETs) of several genes or "sub-ETs". These genes are linked together by an mathematical functions such as subtraction, addition, division, or multiplication [52, 81, 82]. In this study, GeneXproTools 5.0 software was utilized to develop the GEP model. There are several parameters that need a careful calibration in order to develop an optimal GEP model and these include the number of chromosomes (N_c), number of genes (N_g), and head size (H_s) [83, 84]. The calibration of these parameters require basically performing several trials by interchanging the main hyper-parameter settings in the software. Therefore, a systematic approach has been followed in this study for developing the best-fitted and most accurate predictive GEP model, as elaborated in Fig. 8.



400

401 **Fig. 8**. Flow chart of the GEP model.

402 4.1.2 Data preparation and accuracy assessment

403 A total of 10764 data points generated from the parametric study is employed
404 to train the GEP model. The data was categorised into training and validation data
405 sets with proportion of 70% and 30%, respectively. Similar to the ANN and SVR
406 model, the input parameters used are d_o, H, R, t_w, w, and f_y, while the output is the
407 WPB load (V_{GEP}). The model includes several trials to select the most appropriate
408 values for the setting parameters (i.e. N_c, N_g, and H_s) and for the different linking
409 functions (i.e. addition, multiplication and division). The range employed for the
410 setting parameters are 30-160, 3-5, and 8-11 for N_c, N_g, and H_s, respectively. In total,
411 nine models (T1- T9) were established to select the best-performed GEP model. The
412 accuracy of the GEP model is assessed through comparing the predicted values with
413 the corresponding targeted values using the statistical measures including R^2, *RMSE*
414 and *MAE* presented previously in Eqs (14a), (14b) and (14c), respectively.

415 # 4. Results and analysis

416 A comprehensive analysis and discussion for the results of the various machine
417 learning methods employed in this study is presented in this section. This includes a
418 careful examination of the different predicted model in terms of the optimization,
419 validation and accuracy. The results from the proposed machine learning methods
420 are then compared with the corresponding targeted values from the numerical model
421 and those calculated from the analytical expression proposed by Ferreira et al. [4]

422 ## 4.1. Prediction-based ANN

423 The accuracy data for all ANN models produced is presented in Table 3. The
424 results show that there is a clear correlation between accuracy and number of
425 neurons until 8 neurons, at which some results stagnant and some decrease in the
426 level of accuracy. The increase in neurons results in more complicated formula (which
427 lacking behind practicality) and potentially yields ANN model to faces overtraining
428 issue. The R^2, *MAE* and *RMSE* for 8-neuron model with all data are 0.9984, 17.33,
429 and 27.17, respectively. In conclusion, given that the ANN model with eight neurons

shows high level of accuracy with the influence of the inputs on the resistance is similar to what is physically expected. Hence, the model with 8-neurons is selected to be used in further analysis. A comparison of the predicted WPB resistance obtained from the ANN model (V_{ANN}) and the corresponding targeted values from the parametric study (V_{FE}) is given in Fig. 15 (a). Overall, the results show that ANN is a powerful tool in accurately predicting the web-post buckling load of the HSS beams with elliptical web openings.

Table 3: Performance metrics of the ANN models

No. of Neurons	R^2			All data		
	Training	Validation	Testing	R^2	MAE	RMSE
4	0.9959	0.9959	0.9957	0.9959	29.34	44.04
6	0.9975	0.9973	0.9972	0.9974	22.42	34.75
8	0.9984	0.9984	0.9986	0.9984	17.33	27.17
10	0.9984	0.9984	0.9984	0.9984	17.74	27.51

Fig. 9 provides the impact of each input parameter using The Connection Weight Approach. All models show that as the opening height (d_o) and the opening radius (R) increases there is a negative impact on the web-post buckling capacity. On the other hand, there is a positive impact on the web-post buckling capacity with the increase in the distance between geometric centres of flanges (H), the web thickness (t_w), opening width (w), and the yield stress (f_y). . This is in line with what it is expected to see from each parameter, as a thicker web leads to a slenderer web-post and increased resistance, aligning with expectations. Conversely, raising the opening height and radius reduces the tee section's height, resulting in reduced resistance. Additionally, as the height H increases, both the web-post's slenderness and the tee section's height increase, leading to higher vertical shear resistance. This

holds true as long as the geometric ratios (d_o/H, w/d_o, and R/d_o) and web thickness remain constant. Finally, as the yield stress of the HSS increase the capacity of the beams also increase. The parameter with the largest positive impact on the web-post buckling capacity of HSS beams with elliptically-based openings was the beam's height (H) and the parameter with the largest negative impact on the capacity was the opening height (d_o).

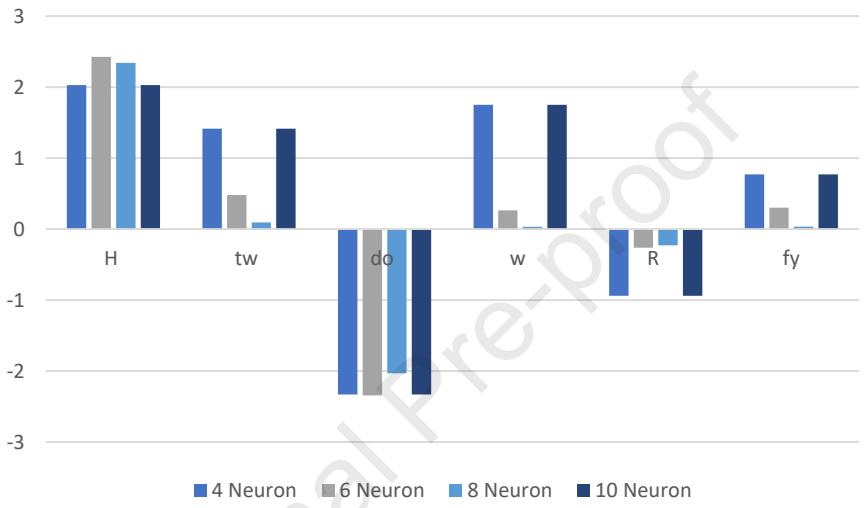


Fig. 9 Impact of input parameters- Connection Weight Approach

The significance of the six input parameters in terms of its contribution value to the output as determined from Garson algorithm is illustrated in Fig. 10. It can be observed that the beam's height (H) and opening height (d_o) are the most significant parameters on the capacity. The percentage contribution of these parameters towards the ultimate capacity is 32.5%, and 25.1%, respectively. It is found that the opening radius (R) and opening width (w) have the lowest effect on the capacity, and percentage contribution of these parameters on the capacity is 3.8% and 9.9%, respectively. The web-thickness (t_w) and yield stress (f_y) of the HSS have intermediate effect on the capacity with percentage contributions of 16.0% and 12.6%, respectively.

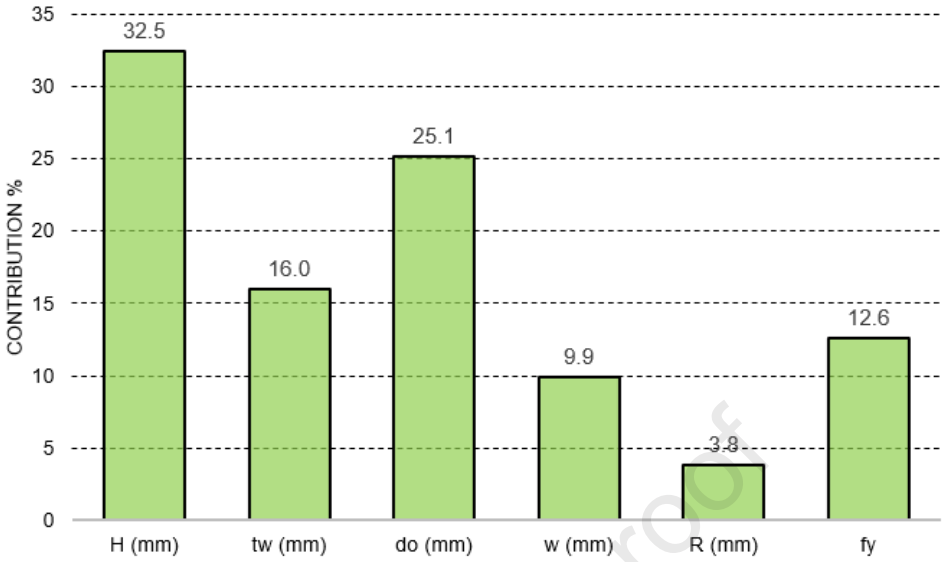


Fig. 10: Contribution (%) of input parameters to the resistance (8 neurons)

The proposed ANN design formula for predicting the WPB capacity for HSS beams with elliptically-based web openings is giving in Eq. (25). It is indicated that normalisation and demoralization of the inputs and the output, respectively, are necessary when Eq. (25) is used, as discussed previously in Eq. 11. The values of w1(i, j), w2(i) and B1(i) corresponding to each neuron i are listed in Table 4. The value of the output bias B_2 is equal to -0.1398.

$$(V)_n = B_2 + \sum_{i=1}^{n=8} W_2 \left(\frac{2}{1 + e^{-2H_i}} - 1 \right)$$

$$H_i = B_1(i) + W_1(i,1)(H)_n + W_1(i,2)(t_w)_n + W_1(i,3)(d_o)_n + W_1(i,4)(w)_n + W_1(i,5)(R)_n + W_1(i,6)(f_y)_n \quad (25)$$

Table 4: The connection weight and the bias values.

Neuron	$w_1(i,j)$						$w_2(i)$	$B_1(i)$
	H	t_w	d_o	w	R	f_y	V_{ANN}	
1	-7.681	2.351	4.945	1.966	-0.208	0.151	-0.436	-1.411
2	4.553	-2.314	-3.031	-0.398	0.403	0.173	-0.706	2.599

3	-3.874	-2.553	3.535	3.009	1.265	-5.126	-0.017	1.422
4	0.304	-0.087	-0.242	0.009	0.004	0.179	36.724	0.364
5	0.240	-0.215	-0.139	-0.018	0.004	0.145	-21.815	0.246
6	1.979	0.446	-1.805	-0.792	-0.527	0.440	0.679	-2.315
7	-0.569	0.875	0.480	-0.909	-0.293	0.125	-1.345	-1.932
8	0.323	0.065	-0.317	0.040	0.005	0.185	-18.269	0.492

The ANN model with its parameters has been implemented in user-friendly excel sheet. The user is prompted to enter the required input parameters within the range specified in Table 2 to ensure the accuracy of the results. The sheet can be found at: https://github.com/Rabee-Shamass/Web-Post-Buckling-Resistance-Prediction-of-HSS-Beams-with-Elliptical-Web-Openings

4.2. Prediction-based SVR

The correlation between the inputs and outputs implemented in the SVR model is shown in Fig. 11. There is a strong positive correlation between certain input variables, such as (w) and (R) with a coefficient of 0.81, and between (H) and (w) with a coefficient of 0.83. Additionally, a strong correlation is found between (w) and (d_0) with a coefficient of 0.86. There are also correlations around 0.7 between (H) and (t_w), (H) and (R), and (R) and (d_0). However, it's important to note that no significant correlation is observed between (f_y) and the remaining input variables. This is expected since the geometric parameters of the web-opening are not affected by the material property (f_y). On the other hand, strong correlation coefficients are observed between certain inputs, particularly (t_w), (H), and (d_0), and the output (V_{SVR}), with coefficients of 0.89, 0.63, and 0.51, respectively. This suggests that these inputs have a significant influence on the output variable. Conversely, (R) and (f_y) appear to have a weak linear correlation and relatively less impact on (V_{SVR}). These correlations collectively imply that the relationship between the inputs and output tends to be

non-linear. Therefore, it is important to consider the use of non-linear kernel functions such as RBF when employing the SVR model to accurately capture the non-linear patterns in the data.

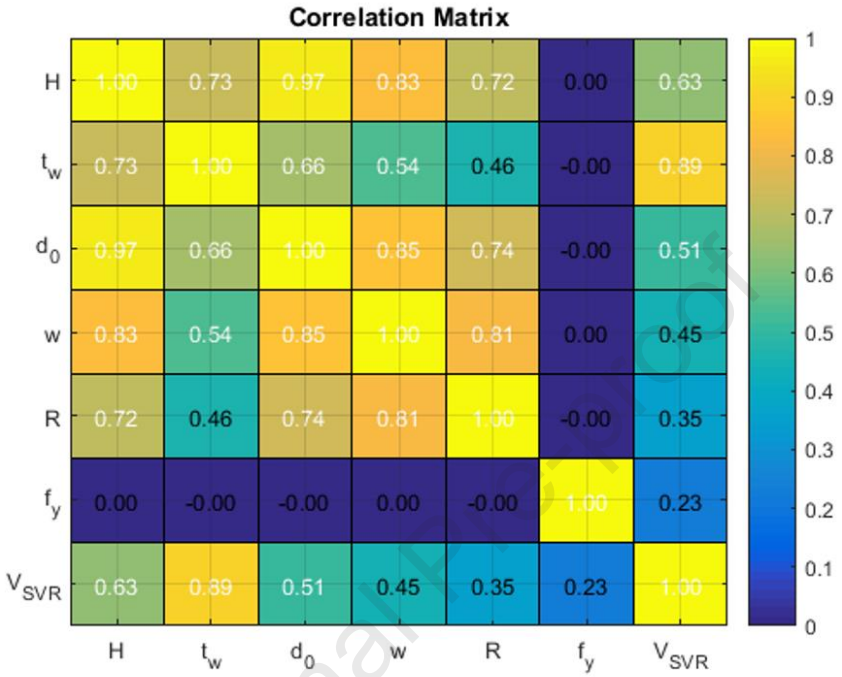
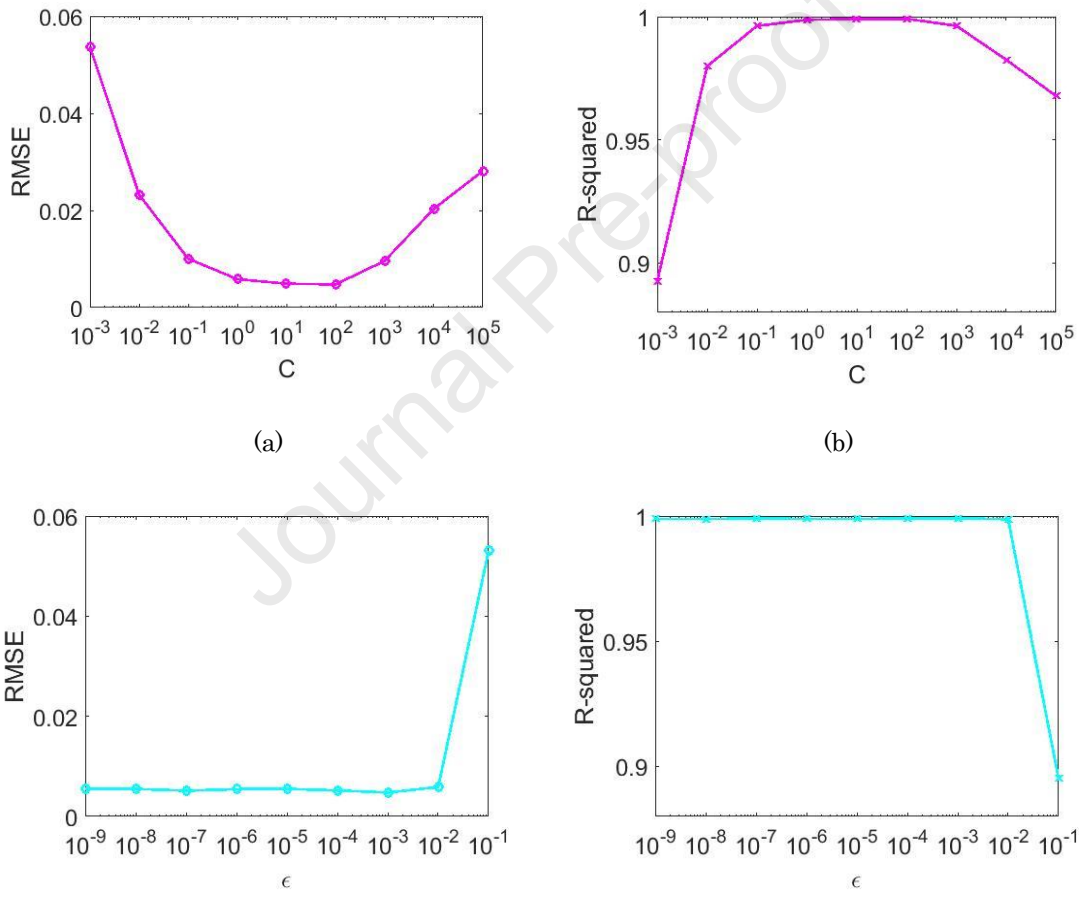


Fig. 11: Pearson's correlation matrix results for the present data.

To achieve accurate prediction with the SVR model, the grid search technique with 10-fold cross-validation is applied, considering 729 possible combinations of hyper parameters for coarse and refined grid search. Fig. 12 displays the variation of each considered hyper parameter during the coarse grid search as a function of R^2 and $RMSE$, with the remaining two hyper parameters being set to their optimal values. The optimal combination of hyper-parameters, which results in the lowest $RMSE$ and the highest R^2 of 0.0048 and 0.9991 respectively, falls within the range of 10^1 to 10^3 for C and 10^{-1} to 10^1 for γ. Therefore, a finer grid search with a step size of 0.25 in log scale is conducted within these intervals to identify the optimal combinations.

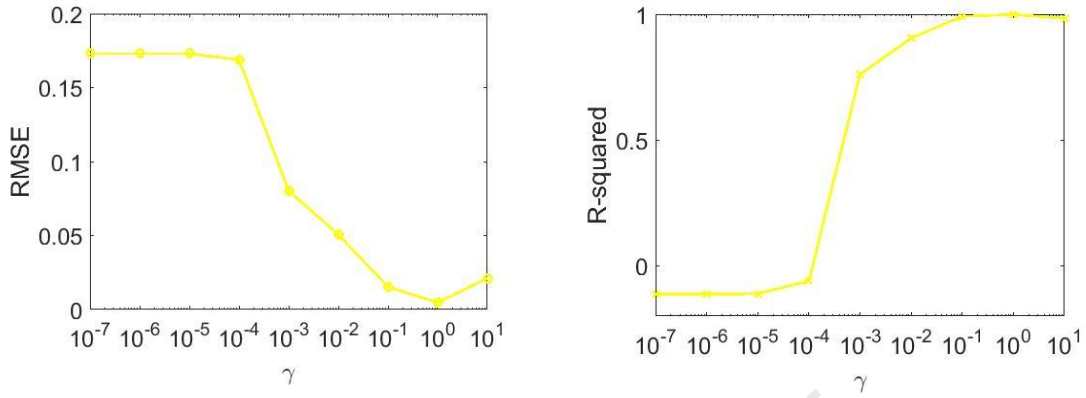
However, both R^2, and $RMSE$ with respect to ε remain relatively stable from 10^{-9} up to 10^{-2}. Fig. 13 demonstrates a strong relationship between ε and the

percentage of support vectors. Specifically, as ε decreases (while keeping the other hyperparameters fixed), the number of support vectors increases, leading to a more complex model. Consequently, a finer grid search with a 0.25 step size in log scale is employed in the interval from 10^{-4} to 10^{-2} and the optimal value for this hyper parameter is chosen to minimize $RMSE$, maximize R^2, and reduce the number of support vectors. Hence, the best combination of hyper parameters obtained is $(C, \varepsilon, \gamma) = (10^{1.5}, 10^{-2.75}, 10^{0.5})$, with $RMSE$ 0.0035 (normalised value) of and an R^2 of 0.9995.



(a) (b)

(c) (d)



(e) (f)

Fig. 12: Identification of hyper parameters through coarse grid search

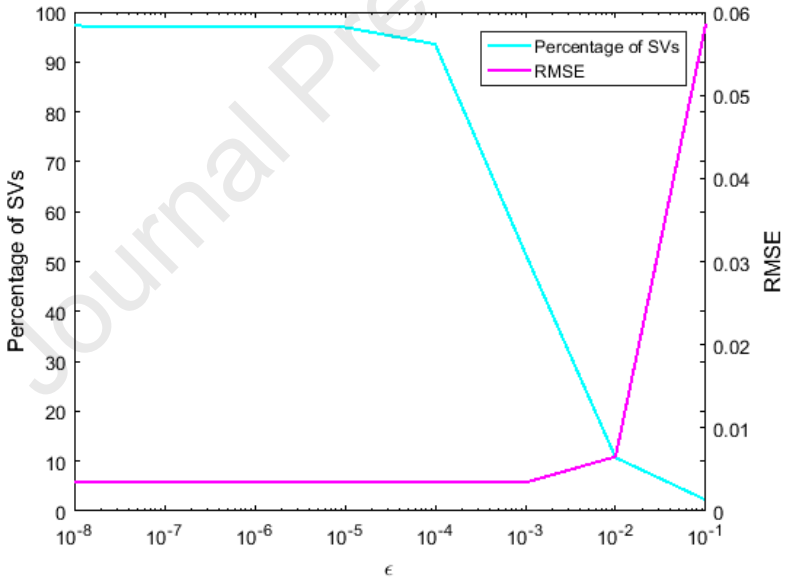


Fig. 13: The influence of ε on percentage of SVs and RMSE.

Fig. 15 (b) displays the predicted ultimate load capacity results with the SVR model (V_{SVR}) compared to those obtained from the parametric study (V_{FE}). Additionally, Table 5 provides the performance metrics for demoralised training, validation, and test data. The presented results indicate the SVR provides excellent depictions of the corresponding actual values, with an R^2 of 99.97% and $RMSE$ of

11.72. Therefore, this model is considered to be a reliable choice for estimating the ultimate capacity, as it provides accurate predictions.

Table 5: Performance metrics of SVR model.

Training		Validation		Testing		All data		
R^2	*RMSE*	R^2	*RMSE*	R^2	*RMSE*	R^2	*RMSE*	*MAE*
0.9999	7.43	0.9993	17.63	0.9993	18.63	0.9997	11.72	6.55

5.3 Prediction-based GEP

In order to achieve the best-performed GEP model with the highest R^2 value and lowest RMSE and MAE values for both the training and validating datasets, a total of nine models (T1 to T9) were performed. Table 6 summarizes the primary setting parameters used in several models with their main statistical measures (i.e. RMSE, MAE and R^2). It is clearly observed that the number of chromosomes (N_c) has significant influence on the accuracy of the model performance, while the other setting parameters, including the number of genes (N_g) and head size (H_s), have shown a negligible impact on the model performance. It is worth noting that in some scenarios the GEP does not account all the input variables (i.e. T1, T2, T5 and T9) and therefore only the models with 6 inputs were considered in the sensitivity analysis. In conclusion, model (T4) is found to provide the most accurate predictions with best performance measures for both training and validation datasets, and therefore it has been selected for further work.

A comparison of the predicted web-post buckling resistance (V_{GEP}) obtained from the GEP model and the corresponding targeted values from the parametric study (V_{FE}) is shown in Fig. 15 (c). The statistical measures of R^2, *MAE*, and *RMSE* values for this data are 0.977, 103.8, and 67.5, respectively. The closeness between these statistical values and those obtained individually for the training and validation datasets indicates the generalization performance of the model.

555 Table 6: Performance metrics of the GEP models

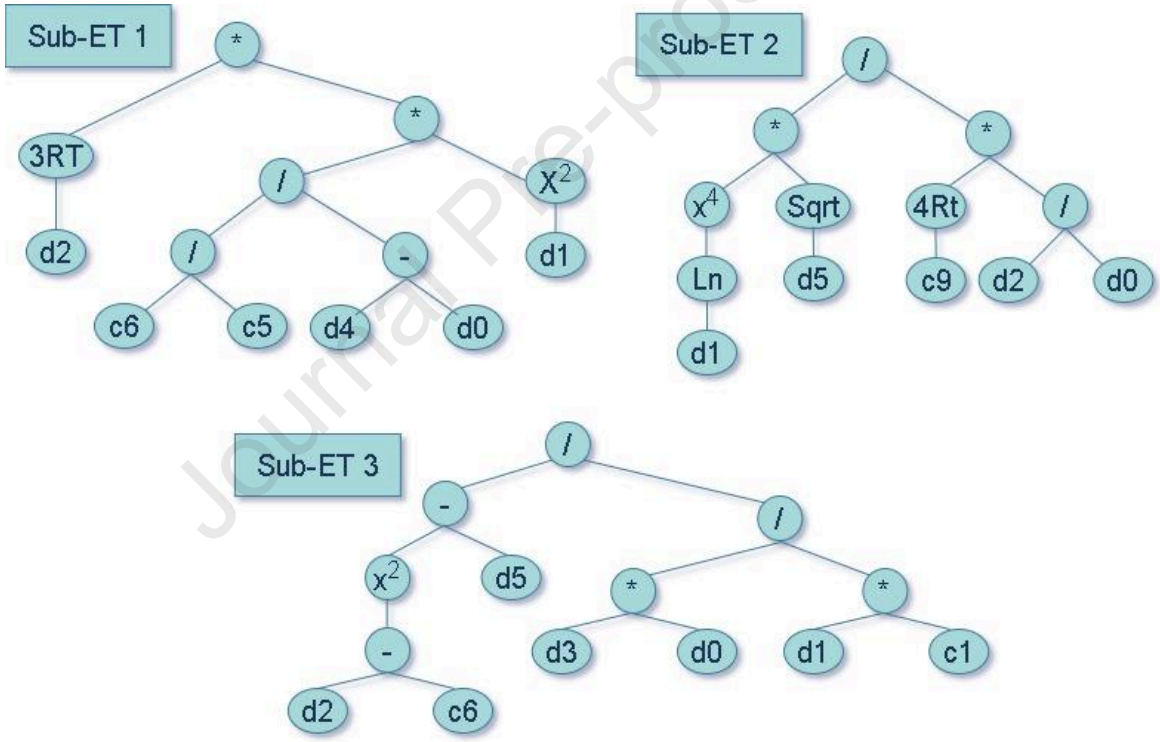
Models	Setting Parameters			GEP Modelling Results						
				Training dataset (70%)			Validation dataset (30%)			No. of used variables
	N_c	H_s	N_g	RMSE	MAE	R^2	RMSE	MAE	R^2	
T1	30	8	3	158.43	103.83	0.948	179.98	115.4	0.933	4
T2	80	8	3	145.98	87.07	0.952	152.72	91.61	0.952	5
T3	130	8	3	121.17	87.64	0.969	131.4	86.18	0.963	6
T4	160	8	3	104.32	67.88	0.977	102.59	66.56	0.977	6
T5	160	9	3	163.21	97.39	0.949	165.28	100.73	0.942	5
T6	160	10	3	126.54	84.78	0.965	139.17	95.62	0.962	6
T7	160	11	3	162.38	106.41	0.946	159.17	102.61	0.946	6
T8	160	8	4	155.07	95.41	0.951	147.16	89.06	0.955	6
T9	160	8	5	143.31	99.23	0.957	150.06	102.48	0.952	5

556 The expression tree of the proposed GEP model is shown in Fig. 14. The
557 notations $d0, d1, d2, d3, d4, \text{and } d5$ represent sequentially the input variables
558 $H, t_w, d_0, w, R, \text{and } f_y$. The constants involved in the model are as follows: $C6 = 116.130$
559 and $C5 = 0.975$ in the first gene (Sub-ET1), $C9 = 1.025$ in the second gene (Sub-ET2),
560 and $C1 = -9.350$ and $C6 = 18.087$ in the third gene (Sub-ET3). This tree, however, is
561 simplified mathematically in Eq. (26). This expression can be used straightforward
562 for prediction of the WPB of the HSS steel beam with elliptically-based web openings
563 without the need for normalising the inputs, reflecting a more practical design tool.
564 It should be mentioned that the unite for the geometric inputs parameters (H, t_w, d_o, w
565 and R) is in mm, whereas f_y and V_{GEP} are in MPa and kN, respectively. The

566 expression tree is usually read from left to right and from bottom to up. This process
567 is done individually for each gene (Sub-ETs) in the tree. After that, the transformed
568 sub-equations are linked together using the linking function in the model (Addition)
569 to obtain the GEP equation. To simplify the result equation for direct use of
570 prediction, the notations (d0 to d5) were expressed in the equation with their
571 corresponding variables, whereas the constant values are substituted numerically.

$$V_{GEP} = \sqrt[3]{d_0}\left(\frac{119.107}{R-H}\right){t_w}^2 + \frac{[ln(t_w)]^4\sqrt{f_y}}{1.0062\left(\frac{d_0}{H}\right)} + \frac{(d_0 - 18.087)^2 - f_y}{\left(\frac{wH}{-9.350t_w}\right)} \quad (26)$$

572



573

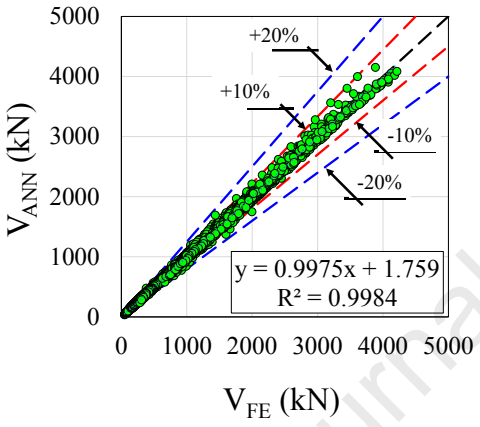
574

575 Fig. 14: Expression tree of the proposed GEP model

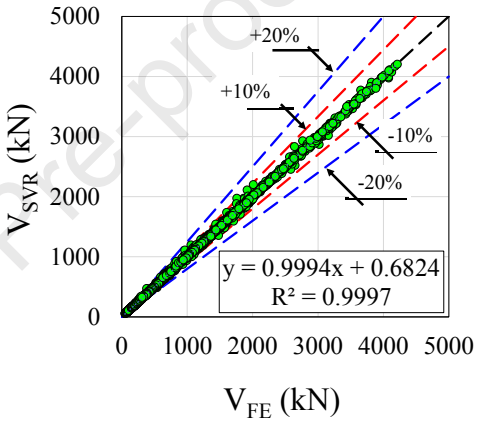
5.4 Comparison between ANN, SVR, GEP and Ferreira et al. based on EC3 calculations

Fig. 15 shows the comparison of the WPB predictions of the different machine learning methods (i.e. ANN, SVR and GEP), as well as the procedure proposed by Ferreira et al. [4]based on EC3, and the corresponding targeted values from the FE models. For the ANN model (Fig. 15a), it was found that the values of mean, standard deviation, and coefficient of variation equal to 1.00, 4.06% e 4.06%, respectively, considering V_{ANN}/V_{FE} ratio. The maximum and minimum relative error values (V_{ANN}/V_{FE}-1) were equal to 25% and -30%, respectively. Regarding the statistical analysis of SVR model (Fig. 15b), the mean, standard deviation, and coefficient of variation were equal to 1.00, 1.81% and 1.81%, respectively, considering V_{SVR}/V_{FE} ratio. The maximum and minimum relative error values (V_{SVR}/V_{FE}-1) were equal to 24% and -10%, respectively. Fig. 15c shows the results of GEP model, and it was verified values of mean, standard deviation, and coefficient of variation were equal to 0.96, 20.41% and 21.27%, respectively, considering V_{GEP}/V_{FE} ratio. The GEP model presented values equal to 49% and -241%, considering the maximum and minimum relative errors (V_{GEP}/V_{FE}-1), respectively. The model proposed by Ferreira et al. [4] is shown in Fig. 15d. The mean, standard deviation, and coefficient of variation were equal to 1.02, 8.47% and 8.29%, respectively, considering V_{EC3}/V_{FE} ratio. The maximum and minimum relative errors (V_{EC3}/V_{FE}-1) were found equal to 61% and 23%, respectively. Table 7 shows all the results of statistical analyses. Based on the presented results and discussion, it can be noted that the ANN and SVR model offers

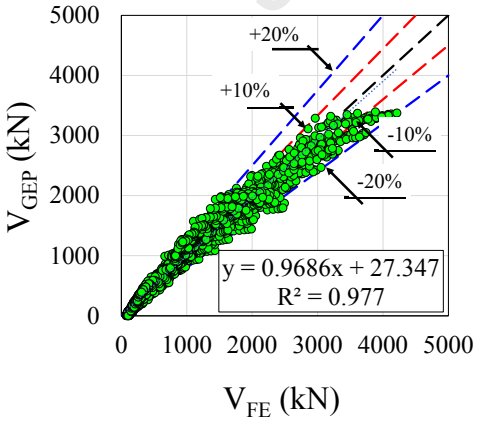
598 the most accurate predictions with the least relative errors. Although, the GEP model
599 exhibits more conservative predictions with lower level of accuracy, their predictions
600 are relatively similar to the analytical model proposed by Ferreira et al. [4] based on
601 EC3 approach. In addition, one of the key advantages for GEP is that it offers a more
602 practical and straightforward design equation with a very simple mathematical
603 functions (i.e. addition, multiplication and division) with no requirement for
604 normalization of the inputs.



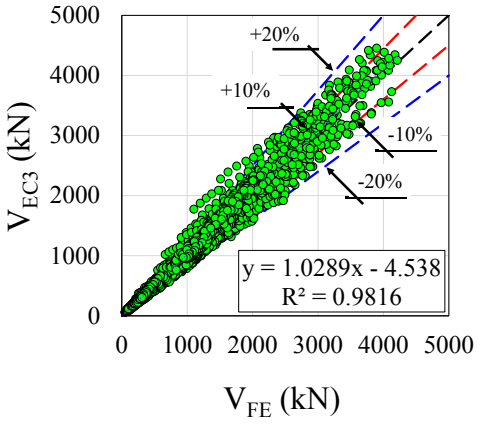
(a) ANN



(b) SVR



(c) GEP



(d) Ferreira et al. [4] based on EC3

605 Fig. 15: Comparison between ANN, SVR, GEP and Ferreira et al. [4] based on EC3
606 predictions with finite element models.

607 Table 7: Statistical analysis

Analysis	ANN	SVR	GEP	EC3*
R^2	0.9984	0.9997	0.9770	0.9816
RMSE (kN)	27.17	11.72	103.80	100.19
MAE (kN)	17.33	6.55	67.48	59.27
Minimum relative error	25%	24%	49%	61%
Maximum relative error	-30%	-10%	-100%	-23%
Mean	1.00	1.00	0.96	1.02
S.D.	4.06%	1.81%	20.41%	8.47%
CoV	4.06%	1.81%	21.27%	8.29%

608 *According to Ferreira et al. [4]

609 Conclusions

610 This paper has provided a thorough investigation of the WPB resistance for HSS
611 beams with elliptically-based web opening. For this purpose, a numerical model was
612 developed and validated, and then a comprehensive parametric study has been
613 conducted including various influential parameters, resulting in a total of 10764 data
614 points. A various machine learning methods were developed, trained and validated
615 using the data generated from the parametric study, and these include Artificial
616 Neural Networks (ANN), Support Vector Machine Regression (SVR) and Gene
617 Expression Programming (GEP) algorithms. The development and the assessment of
618 the performance of these models were presented. A comparison of the predictions of
619 the different machine learning methods and the corresponding targeted values
620 obtained from the FE model was conducted. The performance of the analytical model
621 proposed by Ferreira et al. [4] was also assessed. The proposed machine learning

methods are shown to be a powerful, reliable and efficient design tools for predicting the WPB resistance of HSS beams with periodical elliptically-based openings. A summary of the key findings and conclusions are presented as follows:

- The predictions of the ANN model shows excellent depiction of the corresponding actual values with the R^2, *MAE* and *RMSE* values are 0.9984, 17.33, and 27.17, respectively.
- The ANN model is further validated through the evaluation of the impact of each input parameter that has on the WPB resistance using the Garson Algorithm and the Connection Weight Approach. The impact of these inputs is found to be as physically expected.
- The SVR model provides excellent predictions of the corresponding targeted values, exhibiting outstanding accuracy with R^2, *MAE* and *RMSE* values are 0.9997, 11.72, and 6.55, respectively.
- The results of the GEP model indicates a sufficient level of accuracy R^2, *MAE*, and *RMSE* values for this data are 0.977, 103.8, and 67.5, respectively.
- Similar to the GEP model, the analytical model proposed by Ferreira et al. [4] reflects acceptable and sufficient level of accuracy with R^2, *MAE* and *RMSE* values are 0.9816, 59.3, and 100.2, respectively.
- Based on the presented results, the ANN and SVR model offers the most accurate predictions with the least relative errors.
- Although, the GEP model and the analytical approach proposed by Ferreira et al. [4] provide less accurate predictions compared with ANN and SVR, they

644 offer more practical and straightforward design equations with a very simple
645 mathematical functions.

646 References

647 [1] F.P.V. Ferreira, C.H. Martins, S. De Nardin, Advances in composite beams
648 with web openings and composite cellular beams, J Constr Steel Res. 172
649 (2020) 106182. https://doi.org/10.1016/j.jcsr.2020.106182.

650 [2] R.M.M. Lawson, J. Lim, S.J.J. Hicks, W.I.I. Simms, Design of composite
651 asymmetric cellular beams and beams with large web openings, J Constr
652 Steel Res. 62 (2006) 614–629. https://doi.org/10.1016/j.jcsr.2005.09.012.

653 [3] K.D. Tsavdaridis, C. D'Mello, Structural beam, GB 2492176, 2012.

654 [4] F.P.V. Ferreira, R. Shamass, L.F.P. Santos, K.D. Tsavdaridis, V.
655 Limbachiya, Web-post buckling resistance calculation of perforated high-
656 strength steel beams with elliptically-based web openings for EC3,
657 Structures. 55 (2023) 245–262. https://doi.org/10.1016/j.istruc.2023.05.139.

658 [5] F.P.V. Ferreira, R. Shamass, V. Limbachiya, K.D. Tsavdaridis, C.H.
659 Martins, Lateral–torsional buckling resistance prediction model for steel
660 cellular beams generated by Artificial Neural Networks (ANN), Thin-
661 Walled Structures. 170 (2022) 108592.
662 https://doi.org/10.1016/j.tws.2021.108592.

663 [6] K.D. Tsavdaridis, C. D'Mello, Web buckling study of the behaviour and
664 strength of perforated steel beams with different novel web opening

665 shapes, J Constr Steel Res. 67 (2011) 1605–1620.
666 https://doi.org/10.1016/j.jcsr.2011.04.004.
667 [7] D. Kerdal, D.A. Nethercot, Failure modes for castellated beams, J Constr
668 Steel Res. 4 (1984) 295–315. https://doi.org/10.1016/0143-974X(84)90004-
669 X.
670 [8] E. Ellobody, Interaction of buckling modes in castellated steel beams, J
671 Constr Steel Res. 67 (2011) 814–825.
672 https://doi.org/10.1016/j.jcsr.2010.12.012.
673 [9] E. Ellobody, Nonlinear analysis of cellular steel beams under combined
674 buckling modes, Thin-Walled Structures. 52 (2012) 66–79.
675 https://doi.org/10.1016/j.tws.2011.12.009.
676 [10] L.F. Grilo, R.H. Fakury, A.L.R. de Castro e Silva, G. de S. Veríssimo,
677 Design procedure for the web-post buckling of steel cellular beams, J
678 Constr Steel Res. 148 (2018) 525–541.
679 https://doi.org/10.1016/j.jcsr.2018.06.020.
680 [11] S. Durif, A. Bouchaïr, O. Vassart, Experimental and numerical
681 investigation on web-post specimen from cellular beams with sinusoidal
682 openings, Eng Struct. 59 (2014) 587–598.
683 https://doi.org/10.1016/j.engstruct.2013.11.021.
684 [12] V. Limbachiya, R. Shamass, Application of Artificial Neural Networks for
685 web-post shear resistance of cellular steel beams, Thin-Walled Structures.
686 161 (2021) 107414. https://doi.org/10.1016/j.tws.2020.107414.

687 [13] P. Panedpojaman, T. Thepchatri, S. Limkatanyu, Novel design equations
688 for shear strength of local web-post buckling in cellular beams, Thin-
689 Walled Structures. 76 (2014) 92–104.
690 https://doi.org/10.1016/j.tws.2013.11.007.
691 [14] F.P.V. Ferreira, K.D. Tsavdaridis, C.H. Martins, S. De Nardin, Composite
692 action on web-post buckling shear resistance of composite cellular beams
693 with PCHCS and PCHCSCT, Eng Struct. 246 (2021) 113065.
694 https://doi.org/10.1016/j.engstruct.2021.113065.
695 [15] F.P.V. Ferreira, R. Shamass, L.F.P. Santos, V. Limbachiya, K.D.
696 Tsavdaridis, EC3 design of web-post buckling resistance for perforated
697 steel beams with elliptically-based web openings, Thin-Walled Structures.
698 175 (2022) 109196. https://doi.org/10.1016/j.tws.2022.109196.
699 [16] R. Shamass, F.P.V. Ferreira, V. Limbachiya, L.F.P. Santos, K.D.
700 Tsavdaridis, Web-post buckling prediction resistance of steel beams with
701 elliptically-based web openings using Artificial Neural Networks (ANN),
702 Thin-Walled Structures. 180 (2022) 109959.
703 https://doi.org/10.1016/j.tws.2022.109959.
704 [17] Rabi, M., Shamass, R. and Cashell, K.A., 2022. Experimental investigation
705 on the flexural behaviour of stainless steel reinforced concrete beams.
706 Structure and Infrastructure Engineering, pp.1-13.

[18] Rabi, M., Cashell, K.A. and Shamass, R.J.E.S., 2019. Flexural analysis and design of stainless steel reinforced concrete beams. Engineering Structures, 198, p.109432.

[19] Rabi, M., Cashell, K.A. and Shamass, R., 2021. Ultimate behaviour and serviceability analysis of stainless steel reinforced concrete beams. Engineering Structures, 248, p.113259.

[20] Rabi, M., Shamass, R. and Cashell, K.A., 2022. Structural performance of stainless steel reinforced concrete members: A review. Construction and Building Materials, 325, p.126673.

[21] K. Mela, M. Heinisuo, Weight and cost optimization of welded high strength steel beams, Eng Struct. 79 (2014) 354–364. https://doi.org/10.1016/j.engstruct.2014.08.028.

[22] C. Miki, K. Homma, T. Tominaga, High strength and high performance steels and their use in bridge structures, J Constr Steel Res. 58 (2002) 3–20. https://doi.org/10.1016/S0143-974X(01)00028-1.

[23] M. Veljkovic, B. Johansson, Design of hybrid steel girders, J Constr Steel Res. 60 (2004) 535–547. https://doi.org/10.1016/S0143-974X(03)00128-7.

[24] R. Bjorhovde, Development and use of high performance steel, J Constr Steel Res. 60 (2004) 393–400. https://doi.org/10.1016/S0143-974X(03)00118-4.

[25] A. WHEELER, B. RUSSELL, Behaviour and design of webs in high strength steel under flexural loading, in: Fourth International Conference

729 on Advances in Steel Structures, Elsevier, 2005: pp. 137–142.
730 https://doi.org/10.1016/B978-008044637-0/50018-4.
731 [26] G. Shi, X. Zhu, H. Ban, Material properties and partial factors for
732 resistance of high-strength steels in China, J Constr Steel Res. 121 (2016)
733 65–79. https://doi.org/10.1016/j.jcsr.2016.01.012.
734 [27] B. Karabulut, G. Ferraz, B. Rossi, Lifecycle cost assessment of high
735 strength carbon and stainless steel girder bridges, J Environ Manage. 277
736 (2021) 111460. https://doi.org/10.1016/j.jenvman.2020.111460.
737 [28] K.D. Tsavdaridis, Structural performance of perforated steel beams with
738 novel web openings and with partial concrete encasement, Doctoral Thesis,
739 City University London, 2010.
740 https://openaccess.city.ac.uk/id/eprint/11660/.
741 [29] K.D. Tsavdaridis, C. D'Mello, FE Investigation of Perforated Sections with
742 Standard and Non-Standard Web Opening Configurations and Sizes, in:
743 S.L. Chan (Ed.), 6th International Conference on Advances Is Steel
744 Structures, Hong Kong Institute of Steel Construction, Hong Kong, China,
745 2009: pp. 213–220. http://www.hkisc.org.
746 [30] K.D. Tsavdaridis, C. D'Mello, Vierendeel Bending Study of Perforated
747 Steel Beams with Various Novel Web Opening Shapes through Nonlinear
748 Finite-Element Analyses, Journal of Structural Engineering. 138 (2012)
749 1214–1230. https://doi.org/10.1061/(asce)st.1943-541x.0000562.

[31] K.D. Tsavdaridis, J.J. Kingman, V. V. Toropov, Application of structural topology optimisation to perforated steel beams, Comput Struct. 158 (2015) 108–123. https://doi.org/10.1016/j.compstruc.2015.05.004.

[32] K.D. Tsavdaridis, C. D'Mello, Optimisation of novel elliptically-based web opening shapes of perforated steel beams, J Constr Steel Res. 76 (2012) 39–53. https://doi.org/10.1016/j.jcsr.2012.03.026.

[33] European committee for standardization, EN 1993-1-1: Eurocode 3 – Design of steel structures – Part 1-1: General rules and rules for buildings, (2002).

[34] S. Gholizadeh, A. Pirmoz, R. Attarnejad, Assessment of load carrying capacity of castellated steel beams by neural networks, J Constr Steel Res. 67 (2011) 770–779. https://doi.org/10.1016/j.jcsr.2011.01.001.

[35] Y. Sharifi, S. Tohidi, Lateral-torsional buckling capacity assessment of web opening steel girders by artificial neural networks — elastic investigation, Frontiers of Structural and Civil Engineering. 8 (2014) 167–177. https://doi.org/10.1007/s11709-014-0236-z.

[36] S. Tohidi, Y. Sharifi, Inelastic lateral-torsional buckling capacity of corroded web opening steel beams using artificial neural networks, The IES Journal Part A: Civil & Structural Engineering. 8 (2015) 24–40. https://doi.org/10.1080/19373260.2014.955139.

770 [37] S. Tohidi, Y. Sharifi, Load-carrying capacity of locally corroded steel plate
771 girder ends using artificial neural network, Thin-Walled Structures. 100
772 (2016) 48–61. https://doi.org/10.1016/j.tws.2015.12.007.
773 [38] Y. Sharifi, A. Moghbeli, M. Hosseinpour, H. Sharifi, Neural networks for
774 lateral torsional buckling strength assessment of cellular steel I-beams,
775 Advances in Structural Engineering. 22 (2019) 2192–2202.
776 https://doi.org/10.1177/1369433219836176.
777 [39] Y. Sharifi, A. Moghbeli, M. Hosseinpour, H. Sharifi, Study of Neural
778 Network Models for the Ultimate Capacities of Cellular Steel Beams,
779 Iranian Journal of Science and Technology, Transactions of Civil
780 Engineering. 44 (2020) 579–589. https://doi.org/10.1007/s40996-019-
781 00281-z.
782 [40] M. Hosseinpour, Y. Sharifi, H. Sharifi, Neural network application for
783 distortional buckling capacity assessment of castellated steel beams,
784 Structures. 27 (2020) 1174–1183.
785 https://doi.org/10.1016/j.istruc.2020.07.027.
786 [41] T.-A. Nguyen, H.-B. Ly, V.Q. Tran, Investigation of ANN Architecture for
787 Predicting Load-Carrying Capacity of Castellated Steel Beams,
788 Complexity. 2021 (2021) 1–14. https://doi.org/10.1155/2021/6697923.
789 [42] M. Abambres, K. Rajana, K. Tsavdaridis, T. Ribeiro, Neural Network-
790 Based Formula for the Buckling Load Prediction of I-Section Cellular Steel
791 Beams, Computers. 8 (2018) 2. https://doi.org/10.3390/computers8010002.

792 [43] V. V. Degtyarev, K.D. Tsavdaridis, Buckling and ultimate load prediction
793 models for perforated steel beams using machine learning algorithms,
794 Journal of Building Engineering. 51 (2022) 104316.
795 https://doi.org/10.1016/j.jobe.2022.104316.
796 [44] F. Rosenblatt, The perceptron: A probabilistic model for information
797 storage and organization in the brain, Psychol Rev. 65 (1958) 386–408.
798 [45] Rabi, M., Abarkan, I. and Shamass, R., 2023. Buckling resistance of hot-
799 finished CHS beam-columns using FE modelling and machine learning.
800 Steel Construction.
801 [46] Rabi, Musab, Felipe Piana Vendramell Ferreira, Ikram Abarkan, Vireen
802 Limbachiya, and Rabee Shamass. "Prediction of the cross-sectional
803 capacity of cold-formed CHS using numerical modelling and machine
804 learning." Results in Engineering 17 (2023): 100902.
805 [47] Rabi, Musab. "Bond prediction of stainless-steel reinforcement using
806 artificial neural networks." Proceedings of the Institution of Civil
807 Engineers-Construction Materials (2023): 1-11.
808 [48] V.N. Vapnik, The Nature of Statistical Learning Theory, Springer New
809 York, New York, NY, 2000. https://doi.org/10.1007/978-1-4757-3264-1.
810 [49] J. Tinoco, A. Gomes Correia, P. Cortez, Support vector machines applied to
811 uniaxial compressive strength prediction of jet grouting columns, Comput
812 Geotech. 55 (2014) 132–140.
813 https://doi.org/10.1016/j.compgeo.2013.08.010.

814 [50] Y. Xu, B. Zheng, M. Zhang, Capacity prediction of cold-formed stainless
815 steel tubular columns using machine learning methods, J Constr Steel Res.
816 182 (2021) 106682. https://doi.org/10.1016/j.jcsr.2021.106682.
817 [51] M. Zarringol, H.-T. Thai, M.Z. Naser, Application of machine learning
818 models for designing CFCFST columns, J Constr Steel Res. 185 (2021)
819 106856. https://doi.org/10.1016/j.jcsr.2021.106856.
820 [52] C. Ferreira, Gene Expression Programming: A New Adaptive Algorithm
821 for Solving Problems, Complex Systems. 13 (2001) 87–127.
822 [53] Y.S. Jweihan, Predictive model of asphalt mixes' theoretical maximum
823 specific gravity using gene expression programming, Results in
824 Engineering. 19 (2023) 101242.
825 https://doi.org/10.1016/j.rineng.2023.101242.
826 [54] H.-T. Thai, Machine learning for structural engineering: A state-of-the-art
827 review, Structures. 38 (2022) 448–491.
828 https://doi.org/10.1016/j.istruc.2022.02.003.
829 [55] R.M. Lawson, S.J. Hicks, Design of composite beams with large web
830 openings. SCI P355., The Steel Construction Institute, 2011.
831 [56] Simulia, A.B.A.Q.U.S., 2009. Abaqus/CAE. Theory manual. Dassault
832 Systèmes, Providence, Rhode Island.
833 [57] W. Zaarour, R. Redwood, Web Buckling in Thin Webbed Castellated
834 Beams, Journal of Structural Engineering. 122 (1996) 860–866.
835 https://doi.org/10.1061/(ASCE)0733-9445(1996)122:8(860).

836 [58] K.D. Tsavdaridis, G. Galiatsatos, Assessment of cellular beams with
837 transverse stiffeners and closely spaced web openings, Thin-Walled
838 Structures. 94 (2015) 636–650. https://doi.org/10.1016/j.tws.2015.05.005.
839 [59] F.P.V. Ferreira, K.D. Tsavdaridis, C.H. Martins, S. De Nardin, Ultimate
840 strength prediction of steel–concrete composite cellular beams with
841 PCHCS, Eng Struct. 236 (2021) 112082.
842 https://doi.org/10.1016/j.engstruct.2021.112082.
843 [60] F.P.V. Ferreira, K.D. Tsavdaridis, C.H. Martins, S. De Nardin, Buckling
844 and post-buckling analyses of composite cellular beams, Compos Struct.
845 262 (2021). https://doi.org/10.1016/j.compstruct.2021.113616.
846 [61] F.P.V. Ferreira, A. Rossi, C.H. Martins, Lateral-torsional buckling of
847 cellular beams according to the possible updating of EC3, J Constr Steel
848 Res. 153 (2019) 222–242. https://doi.org/10.1016/j.jcsr.2018.10.011.
849 [62] F.P.V. Ferreira, C.H. Martins, LRFD for Lateral-Torsional Buckling
850 Resistance of Cellular Beams, International Journal of Civil Engineering.
851 18 (2020) 303–323. https://doi.org/10.1007/s40999-019-00474-7.
852 [63] R. Shamass, F. Guarracino, Numerical and analytical analyses of high-
853 strength steel cellular beams: A discerning approach, J Constr Steel Res.
854 166 (2020) 105911. https://doi.org/10.1016/j.jcsr.2019.105911.
855 [64] X. Yun, L. Gardner, Stress-strain curves for hot-rolled steels, J Constr
856 Steel Res. 133 (2017) 36–46. https://doi.org/10.1016/j.jcsr.2017.01.024.

[65] MATLAB and Statistics Toolbox Release 2019a, The MathWorks, Inc., Natick, Massachusetts, United States, 2019.

[66] Tarawneh, A., Saleh, E., Almasabha, G. and Alghossoon, A., 2023. Hybrid Data-Driven Machine Learning Framework for Determining Prestressed Concrete Losses. Arabian Journal for Science and Engineering, pp.1-15.

[67] Tarawneh, A., Almasabha, G. and Murad, Y., 2022. ColumnsNet: neural network model for constructing interaction diagrams and slenderness limit for FRP-RC columns. Journal of Structural Engineering, 148(8), p.04022089.

[68] Jweihan, Y.S., Al-Kheetan, M.J. and Rabi, M., 2023. Empirical Model for the Retained Stability Index of Asphalt Mixtures Using Hybrid Machine Learning Approach. Applied System Innovation, 6(5), p.93.

[69] Jweihan, Y.S., Alawadi, R.J., Momani, Y.S. and Tarawneh, A.N., 2022. Prediction of Marshall Test Results for Dense Glasphalt Mixtures Using Artificial Neural Networks. Frontiers in Built Environment, 8, p.949167.

[70] Gupta, T., Patel, K. A., Siddique, S., Sharma, R. K. and Chaudhary, S. (2019) Prediction of mechanical properties of rubberised concrete exposed to elevated temperature using ANN, Measurement, 147 , p. 106870.

[71] Olden, J. D., Joy, M. K. and Death, R. G. (2004) An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data, Ecological Modelling, 178 (3), pp. 389-397.

[72] Garson DG. Interpreting neural network connection weights. 1991.

878 [73] T. Kari, W. Gao, A. Tuluhong, Y. Yaermaimaiti, Z. Zhang, Mixed Kernel
879 Function Support Vector Regression with Genetic Algorithm for
880 Forecasting Dissolved Gas Content in Power Transformers. Energies 11
881 (2018) 2437. https://doi.org/10.3390/en11092437
882 [74] Awad, M., Khanna, R., Awad, M., & Khanna, R. (2015). Support vector
883 Regression. Efficient learning machines: Theories, concepts, and
884 applications for engineers and system designers, 67-80.
885 [75] Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector
886 regression. Statistics and computing, 14, 199-222.
887 [76] Fletcher, T. (2009). Support vector machines explained. Tutorial paper, 1-
888 19.
889 [77] C. Cortes, V. Vapnik, Support-vector 857 networks, Mach. Learn. 20 (3)
890 (1995) 273– 297.
891 [78] Lafdani, E. K., Nia, A. M., & Ahmadi, A. (2013). Daily suspended sediment
892 load prediction using artificial neural networks and support vector
893 machines. Journal of Hydrology, 478, 50-62.
894 [79] Nguyen, H., Vu, T., Vo, T. P., & Thai, H. T. (2021). Efficient machine
895 learning models for prediction of concrete strengths. Construction and
896 Building Materials, 266, 120950.
897 [80] Chen, S. T., & Yu, P. S. (2007). Pruning of support vector networks on flood
898 forecasting. Journal of hydrology, 347(1-2), 67-78.
899 [81] MATLAB and Statistics Toolbox Release 2019a, The MathWorks, Inc.,

[82] Natick, Massachusetts, United States, 2019.

[83] Syarif, I., Prugel-Bennett, A., & Wills, G. (2016). SVM parameter optimization using grid search and genetic algorithm to improve classification performance. TELKOMNIKA (Telecommunication Computing Electronics and Control), 14(4), 1502-1509. http://doi.org/10.12928/telkomnika.v14i4.3956

[84] Hsu, C. W., Chang, C. C., & Lin, C. J. (2003). A practical guide to Support vector classification.

[85] K. Pearson, Note on regression and inheritance in the case of two parents, Proceedings of the Royal Society of London 58 (1895) 240–242. https://doi.10.1098/rspl.1895.0041.

[86] Ferreira, C., 2002. Gene expression programming in problem solving. Soft computing and industry: recent applications, pp.635-653.

[87] Jweihan, Y.S., 2023. Predictive model of asphalt mixes' theoretical maximum specific gravity using gene expression programming. Results in Engineering, p.101242.

[88] Raheel, M., Iqbal, M., Khan, R., Alam, M., Azab, M. and Eldin, S.M., 2023. Application of gene expression programming to predict the compressive strength of quaternary-blended concrete. Asian Journal of Civil Engineering, pp.1-14.

[89] Khan, K., Iqbal, M., Raheel, M., Amin, M.N., Alabdullah, A.A., Abu-Arab, A.M. and Jalal, F.E., 2022. Prediction of Axial Capacity of Concrete Filled

922 Steel Tubes Using Gene Expression Programming. Materials, 15(19),
923 p.6969.

924

Enhancing Web-Post Buckling Resistance Prediction in High-Strength Steel Beams with Elliptical Web Openings

Musab Rabi [a], Yazeed S. Jweihan[b], Ikram Abarkan[c], Felipe Piana Vendramell Ferreira[d], Rabee Shamass[e], Vireen Limbachiya[e], Konstantinos Daniel Tsavdaridis[f], Luis Fernando Pinho Santos[e]

Highlights

- The behaviour of high strength steel beams with elliptically-based web openings is examined.
- A total of 10764 web-post finite element models are established
- The paper includes various machine learning methods
- Machine learning-based design formulas are proposed for predicting the web-post buckling resistance
- A comparison with results from the finite element models and existing analytical model is presented

Enhancing Web-Post Buckling Resistance Prediction in High-Strength Steel Beams with Elliptical Web Openings

Musab Rabi [a], Yazeed S. Jweihan[b], Ikram Abarkan[c], Felipe Piana Vendramell Ferreira[d], Rabee Shamass[e], Vireen Limbachiya[e], Konstantinos Daniel Tsavdaridis[f], Luis Fernando Pinho Santos[e]

Dear Editor,

The authors wish to confirm that there are no known conflicts of interest associated with this research. We understand that the Corresponding Author is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs. We confirm that we have provided a current, correct email address which is accessible by the Corresponding Author.

Dr. Musab Rabi

Musab.rabi@jpu.edu.jo