# ANN-Based Model for Predicting the Nonlinear Response of Flush Endplate Connections

Gregory Georgiou[1] and Ahmed Elkady[2*]

[1]Department of Civil, Maritime and Environmental Engineering, University of Southampton, United Kingdom. Address: Burgess road, Boldrewood campus, Southampton SO17 7QF, UK. e-mail: gmg1g20@soton.ac.uk

[2]Department of Civil, Maritime and Environmental Engineering, University of Southampton, United Kingdom ([*]*corresponding author*). Address: Burgess road, Boldrewood campus, Southampton SO17 7QF, UK. e-mail: a.elkady@soton.ac.uk

## Abstract

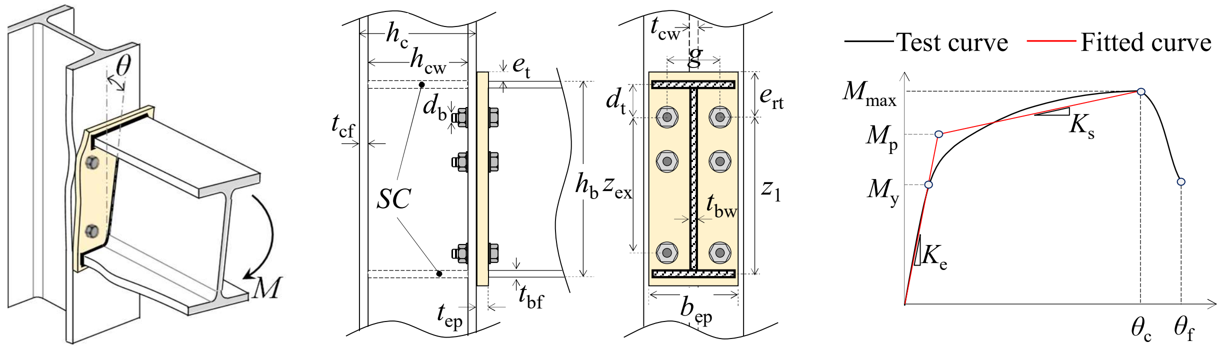
Predicting the moment-rotation response parameters of semi-rigid steel connections can be challenging given the many components contributing to the connection's elastic and plastic deformations. This is the case for the popular flush endplate beam-to-column connections (FEPCs). The literature has highlighted the limitations of current analytical, mechanical, and -traditional- empirical models in providing accurate predictions of the FEPCs' moment-rotation response. Considering this limitation, machine-learning methods have gained wide attention recently in structural engineering applications to address problems associated with complex structural deformation and damage phenomena. To that end, the superior nonlinearity of artificial neural networks (ANN) is employed herein in to predict the response characteristics of FEPCs. A large dataset of about 200 specimens, collected from past experimental programs, is utilized to train the ANN for predicting the bilinear response of FEPCs including strain hardening. The paper describes the deduction of the response parameters from test data using data fitting, the determination of significant geometric, material, and layout features, the ANN architecture and algorithms, and the accuracy metrics of the new model. The SHAP algorithm is used to explain the innerworkings of the model. A computer tool as well as a descriptive guide to the mathematical construction of the ANN are provided to aid with model implementation in practice.

**Keywords:** Steel joints, machine learning, artificial neural networks, moment-rotation response, endplate connections

## Introduction

Steel bolted flush endplate connections (FEPCs), shown in Figure 1(a), are popular in construction practice worldwide. These connections are characterized by a power-shaped nonlinear moment-rotation response. This response is governed by the different elastic and plastic deformation modes of the various connection components such as the endplate, the bolts, the column flange, and the column web. FEPCs response generally falls within the semi-rigid classification (Elkady 2022). It is common however to represent them as pinned or fixed in design or in numerical models, for simplicity.

Due to the multitude of deforming components, predicting a semi-rigid connection's response can be quite challenging. There has been a relatively large effort in the literature to develop reliable models to predict the full moment-rotation response or the key response parameters (elastic rotation stiffness and plastic strength) of FEPCs. This includes empirical (Frye and Morris 1975; Kukreti et al. 1987; Benterkia 1991; Abolmaali et al. 2005; Kozlowski et al. 2008; Ostrowski and Kozłowski 2017), semi-empirical (Rölle 2013; Kong et al. 2020), analytical (Brown et al. 2001; Murray and Shoemaker 2002), and mechanical (CEN 2005) models. Evaluations of the available models have shown that their accuracy is limited, especially for the elastic rotation stiffness. This has been conclusively corroborated recently using a comprehensive experimental-based review and evaluation of these models (Elkady and Mak 2022; Ding and Elkady 2023b). Specifically, errors exceeding 100% can be easily detected in the predictions. This is strongly observed in predictions of the elastic rotational stiffness and ductility. This inaccuracy can be attributed to several reasons including the underlying physical assumptions in analytical and mechanical models, the limited number and/or quality of experimental data used in regressing empirical models, and the limited nonlinearity of traditional empirical models (more details can be found in Ding and Elkady (2023b)).



(a) (b) (c)

Figure 1. (a) General deformation profile of an FEPC; (b) layout of an FEPC showing key geometric features; (c) bilinear fitting of moment-rotation data showing key response parameters

Within the past decade, there has been an increasing adoption of non-parametric regression techniques and machine learning (ML) models, such as decision trees (e.g., XGBoost), artificial neural networks (ANN), and multivariate adaptive regression splines (MARS), within the field of structural engineering. Compared to commonly used empirical models that employ multi-variate linear regression, ML models can capture the high nonlinearity associated with complex physical phenomena and handle the inter-dependency between a large number of predictors. On the other hand, two major criticisms are made against ML models: 1) the complexity of the mathematical model for adoption in practice, and 2) the opacity of the Blackbox model. Concerning the former, the availability and usage of computer tools (subroutines, software, etc.) in everyday structural design and analysis lessen the need for manual computations. In addition, some ML methods such as ANNs and MARS can be expressed in a manageable mathematical form on paper so that others can simply code it. Concerning the latter (i.e., model opacity), the relatively recent emergence of algorithms for model interpretation has alleviated this issue. Lastly, although the utilization of ML models within the structural engineering field has been driven –in part– by the research community's interest in practising with a niche approach, there are many other cases for which using ML models for engineering problems involving classification and regression is properly justified, if not necessary.

For steel connections, several researchers attempted to employ ML models; mainly ANNs (Abdalla and Stavroulakis 1995; De Lima et al. 2005; Faridmehr et al. 2021; Kueh 2021). Abdalla and Stavroulakis (1995) trained deep ANNs using 11 specimens to predict the moment-rotation response of single web angle and shear-tab connections. The model used only three features (predictors). De Lima et al. (2005) trained an ANN using 26 specimens to predict the elastic stiffness and plastic strength of bolted extended endplate connections. Ghassemieh and Nasseri (2012) developed an ANN model to predict the trilinear response of 8-bolt extended endplate connections with plate rib stiffeners. The model was trained using a total of 25 data point generated by 3-dimensional finite element (FE) simulations; the FE model was validated against two test specimens. Faridmehr et al. (2021) trained an ANN model using data from 77 test specimens of connections with top, seat and web angles. The model performed better compared to Eurocode 3 component method with respect to elastic stiffness and plastic strength, where the observed errors were mostly with 27%. Kueh (2021) developed an ANN model for predicting the elastic stiffness and ultimate strength of flush endplate connections. The model was trained using a dataset of 52 physical and FE-simulated FEPC specimens. Although these models were found to be of better performance compared to other empirical models, their performance remained limited. This is because a limited amount of data was used in the models' development. This in turn affects the quality of model training and the model's ability to capture the effect of all significant response predictors. This is particularly detrimental for sensitive response parameters such as the elastic rotational stiffness. Another issue with many of the existing ML models is that they are not made available to end users, whether through reporting the model's mathematical parameters or providing a computer tool.

Considering the aforementioned background on semi-rigid connections, a strong argument can be made towards the utilization of ML models to overcome the limitations of existing models and predict the complex response of semi-rigid connections. This paper attempts to examine such applications for FEPCs. To the best of the author's knowledge, this is the first study to do so for FEPCs using a large dataset made up of only experimental data. Specifically, three ANNs are trained to predict three of the

FEPCs' key response parameters that are sufficient to characterize the connection behaviour as a bilinear one with hardening. This was done using a comprehensive and curated digital database of past laboratory tests. In this paper, we describe the architecture of the ANN, the determination of significant features, the performance and inner workings of the developed ANNs, the model's limitations, and recommendations for future developments.

## Experimental Database

A digital multi-attribute database was recently compiled for past experimental research on FEPCs (Mak and Elkady 2021). The database currently comprises close to 600 specimens. The database also includes the digital moment-rotation response of each test specimen. A systematic procedure was utilized to fit the test data with a bilinear curve as shown in Figure 1(c) based on the equal-area fitting approach. Details of the fitting procedure can be found in Elkady (2022). Several response parameters are deduced including the joint's elastic rotation stiffness $K_e$, the plastic moment $M_p$, and the post-yield stiffness $K_s$. The database attributes were carefully curated and checked to make sure they did not include input mistakes. In this study, a subset of this database is used. This subset represents tests on bare steel beam-to-column connections with I-shaped columns and a major-axis orientation. Tests involving splice connections, rigid column sections, beam axial load, or irregular bolt layouts are ignored. This subset contains 198 specimens.

It should be noted that generally, larger data sets are better for training any regression or ML model. However, there is no lower limit for the dataset size. For ANNs, the dataset size is dependent on the ANN architecture (number of features and neurons), training algorithm, the data spread/quality, and the nature of the problem being modeled. The adequacy of the dataset size can accordingly be assessed based on the model accuracy, generalizability, and the absence of overfitting. Those are assessed later to demonstrate the model robustness. Furthermore, the dataset involves high quality data deduced from well curated tests. The dataset is also well distributed covering the practical design space with no obvious data gaps as will be demonstrated in the subsequent sections.

## Target Response Parameters

In this study, emphasis is placed on predicting three response parameters that are sufficient to characterize the connection response as a bilinear curve with strain hardening. These are the elastic rotational stiffness ($K_e$), the post-yield stiffness ($K_s$), and the plastic moment ($M_p$). The latter ($M_p$) is often referred to as the equivalent yield moment ($M_{ye}$) in the literature (Lignos et al. 2019; Elkady 2022). Figure 2 shows the histogram distribution of the three parameters based on the collected dataset.

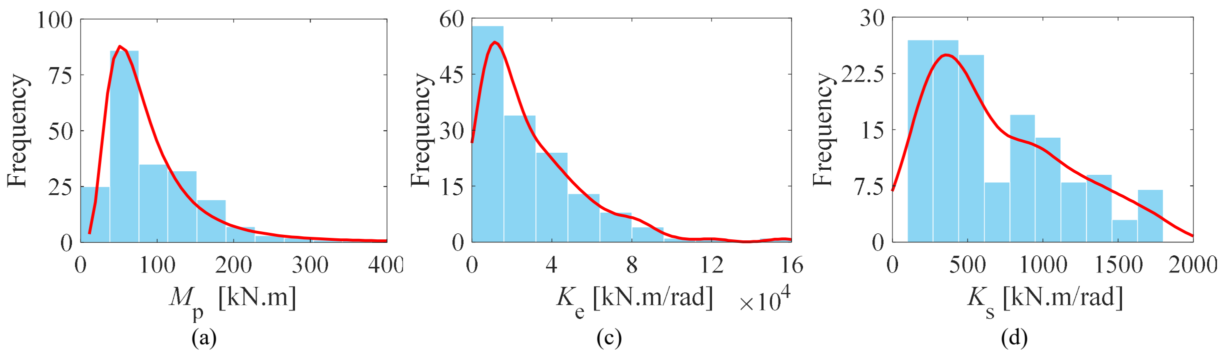



Figure 2. Distribution of the response parameters based on the collated database: (a) plastic moment, (b) elastic rotational stiffness, and (c) post-yield stiffness

## Determination of Significant Features

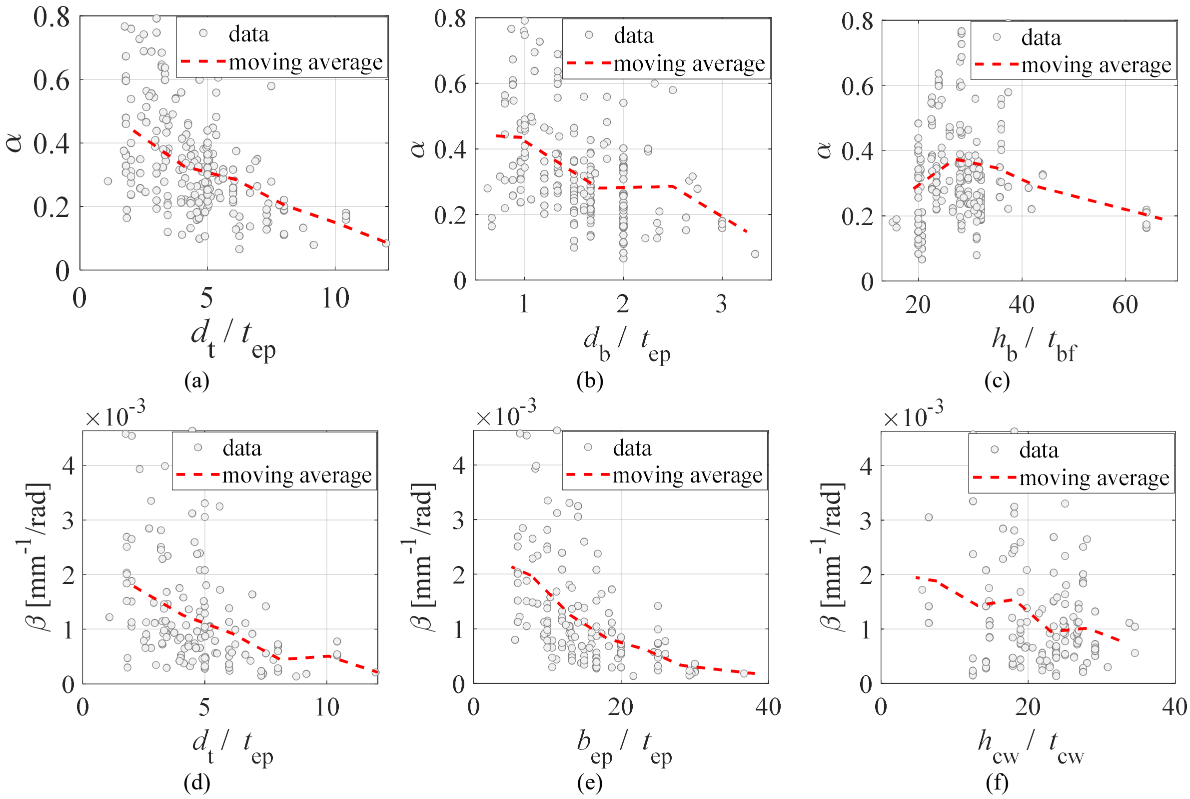
As a first step, the significance of different geometric and material features in predicting the target response parameters is investigated. This is guided by key features that control the underlying physics and expected deformation mechanisms in the connection, as demonstrated in Figure 1(a). A summary of the features used in some of the past analytical and empirical models is provided in Table 1. Various researchers used different numbers of features in their model, ranging from 3 to 12 features. Since the connection's plastic strength is mostly controlled by either endplate or column flange bending, common geometric features included: 1) the lever arm of bending, which can be represented by either the beam depth ($h_b$) or the distance of the extreme tension bolt row to the center of compression (e.g., $z_1$ or $z_{ex}$), 2) the endplate bending length, which can be represented by the distance of the extreme tension bolt row to the beam flange center ($d_t$) or alternatively to the flange's exterior or interior edge, 3) the thickness of the endplate ($t_{ep}$) and the column flange ($t_{cf}$), and 4) the bolt gauge length ($g$) or alternatively the endplate width ($b_{ep}$). Material features are also important when it comes to plastic strength prediction. This includes the yield stress of the column, beam, endplate, and the bolt ($f_{y,C}$, $f_{y,B}$, $f_{y,P}$, and $f_{y,b}$, respectively). For the elastic rotational stiffness, geometric parameters are only relevant (excluding the modulus of elasticity which does vary significantly steel). Notably, for $K_e$, past models included features to capture the influence of the column web panel zone flexibility, such as the column height ($h_c$) and its web thickness ($t_{cw}$), or collectively, the column's web shear area ($A_{vc}$).

Table 1. Summary of key features used in existing predictive methods for FEPCs

| Reference | Model type | Predictions | Features |
|---|---|---|---|
| Frye and Morris (1975) | Empirical | Full $M$-$\theta$ | $t_{ep}$, $t_{cf}$, $z_{ex}$ |
| Kukreti et al. (1987) | Empirical | Full $M$-$\theta$ | $t_{ep}$, $b_{ep}$, $h_b$, $t_{bw}$, $t_{bf}$, $d_t$, $d_b$, $f_{y,B}$, $f_{y,b}$ |
| Benterkia (1991) | Empirical | Full $M$-$\theta$ | $t_{ep}$, $t_{cf}$, $h_b$, $g$, $d_t$, $d_{t2}$, $f_{y,B}$, $f_{y,C}$, $f_{y,P}$, $P_{y,b}$, $P_t$ |
| Brown and Anderson (2001) | Analytical | $K_e$ | $t_{ep}$, $t_{cf}$, $g$, $t_{bw}$, $z_1$, $A_{vc}$ |
| Murray and Shoemaker (2002) | Analytical | $M_p$ | $t_{ep}$, $b_{ep}$, $h_b$, $g$, $z_1$, $z_2$, $m$, $d_b$, $f_{y,P}$, $f_{u,b}$, … |
| CEN (2005) | Mechanical | $K_e$ and $M_p$ | $t_{ep}$, $b_{ep}$, $h_b$, $g$, $z_1$, $z_2$, $m$, $d_b$, $f_{y,P}$, $f_{u,b}$, … |
| Abolmaali et al. (2005) | Empirical | Full $M$-$\theta$ | $t_{ep}$, $b_{ep}$, $h_b$, $t_{bw}$, $t_{bf}$, $g$, $d_b$, $d_t$, $f_{y,P}$ |
| Kozlowski et al. (2008) | Empirical | $K_e$ and $M_p$ | $t_{ep}$, $h_b$, $h_c$, $d_b$ |
| Rölle (2013) | Semi-Empirical | $K_e$ and $M_p$ | $t_{ep}$, $t_{cf}$, $h_b$, $t_{bw}$, $g$, $z_1$, $m$, $m_2$, $d_b$, $f_{y,P}$, $f_{u,b}$ |
| Kong et al. (2020) | Semi-Empirical | $K_e$ and $M_u$ | $t_{ep}$, $b_{ep}$, $h_b$, $h_c$, $t_{cw}$, $g$, $d_1$, $d_b$ |

Using the collated experimental database, correlation plots were further used to confirm the significance of the above features. Figures 3(a) and 3(b) show several sample correlation plots of the response parameters with respect to selected geometric and material features. In these plot, $M_p$ is normalized by the beam's plastic strength $M_{p,b}$ (noted as $\alpha$) and $K_e$ is the normalized by the beam's rigidity $EI_x$ (noted as $\beta$). Similarly, the features are expressed as dimensionless quantities; for example, the ratio $d_t$ to $t_{ep}$ is used to represent the combined effect of the two geometric features. This is done to better visualize these correlations, considering the dependency of the response parameters on several features. A curve representing the moving average is superimposed in the same figure to represent the data tendency. Figure 3 demonstrates the strong correlation between the different response parameters and the selected features.

For example, all three parameters are inverse proportionality with $d_t/t_{ep}$, which represents the relative rigidity of the endplate portion experiencing bending beyond the extreme bolt row in tension. Figure 3(b) shows the inverse proportionality between $\alpha$ and $d_b/t_{ep}$, indicating that FEPCs with thinner endplates and stronger bolts develop lower moment capacity while those with thick endplates and weaker bolts develop larger moment capacity. This is rational since higher resistance is expected by the high strength bolts, if they are the main deformation component. Figure 3(f) also demonstrates the inverse proportionality between $\beta$ and $h_{cw}/t_{cw}$, which represents the column web panel zone slenderness, where a compact panel zone results in higher stiffness and vice versa. The post-yield stiffness is strongly correlated with the endplate material hardening slope, represented by $f_{u,P}/f_{y,P}$ (see Figure 3(i)). Even with these clear correlations, large variability is observed around the moving average. This is attributed to the fact that the FEPCs response is affected by many other parameters that are not represented in Figure 3.



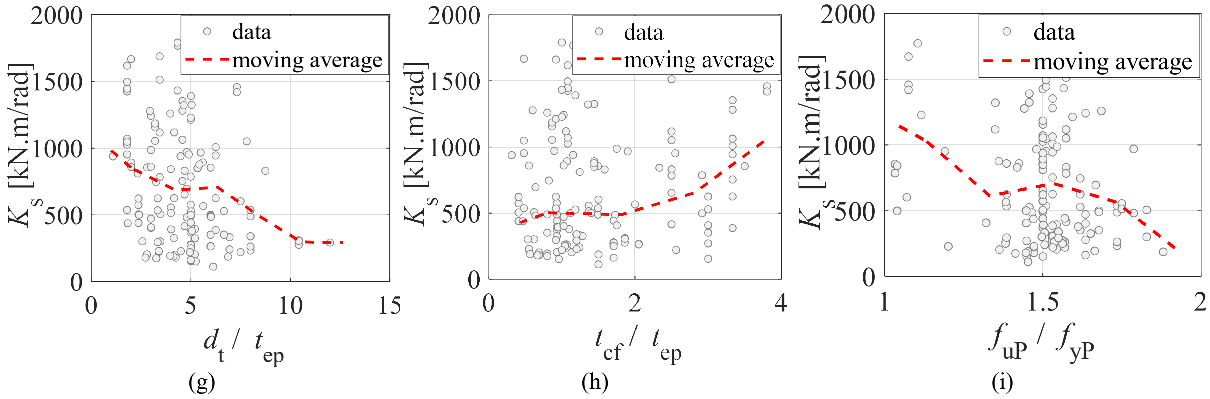



Figure 3. Correlation between the response parameters and selected geometric and material features

Based on these plots, candidate geometric and material features are selected for the ANN development of each response parameter. These features are summarized in Table 2. Other than the features already defined in Figure 1(a), two material features are considered; the endplate material's measured yield and ultimate stress $f_{y,P}$ and $f_{u,P}$, respectively, as well as the bolt measured ultimate stress $f_{u,b}$. The material properties of the column and beam are not considered since deformations within these two components did not control the plastic hinge formation in the majority of the specimens. Also, beams and columns were mostly fabricated from S355 steel (or equivalent grades) within the collated test specimens.

Two categorical features were considered; *SC* which represent the presence/absence of column web stiffeners (i.e., continuity plates), and *Joint* which identifies whether the joint is cantilever (exterior) or cruciform (interior). The *SC* feature controls the extent of the column flange deformation; hence it affects both the strength and the stiffness. The *Joint* feature controls the deformation of the column web panel zone; hence, it affects the elastic stiffness. Within the studied dataset, 54% of the specimens are exterior joints, 67% had unstiffened column, 6% had a column stiffener at the beam compression flange only, and 27% had a column stiffened at both beam flanges.

Label encoding is used to transform the categorical features (*SC* and *Joint*) into numerical ones. Specifically, the values 0, 1, and 2 are assigned to specimens with column stiffeners at both beam flanges, with one stiffener at the compression flange, and no stiffeners, respectively. Similarly, a value of 1 is assigned to interior joint specimens (i.e., cruciform test setup with symmetric loading) and a value of 2 is assigned to exterior joint specimens (i.e., cantilever test setup) or interior joint specimens with asymmetric beam loading.

One should note that the bolt's gauge length ($g$) is not selected here. Although this feature is important and appears in most predictive models, its value does not change much in practice (typically, between 80mm and 160mm). Instead, the endplate width ($b_{ep}$) is chosen as a stronger predictor of the endplate rigidity and strength. This was observed through multiple trial runs with various ANN models with different combination of features.

The post-yield stiffness, $K_s$, is commonly ignored in available models that tend to prioritize the bilinear perfectly-plastic response. For FEPCs, $K_s$ is about 5% $K_e$ on average. Larger $K_s$ values are observed for specimens controlled by bolt elongation. For the post-yield stiffness, $K_s$, the material strain hardening plays a major role. Therefore, both the yield and ultimate yield stresses of the endplate (generally, the main deforming component) are considered as significant features. Finally, it is worth noting that for the ANN training, the selected features are not used in a normalized form (e.g., $d_t/t_{ep}$) since using them separately yielded better fit for the networks. For the same reason, $\beta$ is used as a target response parameter instead of $K_e$, as this yielded better results and eliminated the need to include the beam rigidity parameters $EI_x$ as features.

Table 2. Significant features identified for each response parameter

| ANN | Significant features | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Joint | SC | $d_t$ | $b_{ep}$ | $t_{ep}$ | $h_b$ | $t_{bf}$ | $h_{cw}$ | $t_{cw}$ | $b_{cf}$ | $t_{cf}$ | $d_b$ | $f_{u,b}$ | $f_{y,P}$ | $f_{u,P}$ |
| $M_p$ | | ● | ● | ● | ● | ● | ● | | ● | ● | ● | ● | ● | ● | |
| $\beta$ | ● | ● | ● | ● | ● | ● | ● | ● | | | ● | ● | | | |
| $K_s$ | | ● | ● | ● | ● | | ● | ● | ● | | ● | ● | | ● | ● |

Figure 4 shows the kernel density distribution for the significant features. Table 3 summarizes the statistical distribution of the significant features and the target response parameters, including the mean ($\mu$), standard deviation ($\sigma$), minimum, and maximum values. Note that these statistics differ slightly since different datasets were used to train the ANN of each response parameter. The data spans a wide range of the design space of FEPCs, covering those with shallow and deep beams as well as those with thin and thick endplates. Connections fabricated from conventional carbon steel, high-strength steel and stainless-steel materials are also included. Figure 5 shows the correlation matrix for the significant features. In general, the features are not correlated as indicated by the dominant blue colour of the heat map. Strong correlations (>0.7) are observed between $t_{cw}$ and $t_{cf}$ as well as $t_{bf}$ and $b_{cf}$. The latter is expected given that the columns are mostly hot-rolled sections with proportional dimensions. Nonetheless, these correlated parameters were still used within the same network as indicated in Table 2 since removing one of them would result in a lower model performance.

Table 3. Statistical summary of the features and target response parameters for the different datasets

| Feature | Type | $M_p$ ANN dataset | | | | $\beta$ ANN dataset | | | | $K_s$ ANN dataset | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | μ | σ | min | max | μ | σ | min | max | μ | σ | min | max |
| $d_t$ | input | 55 | 13 | 30 | 118 | 56 | 15 | 30 | 118 | 56 | 13 | 30 | 118 |
| $b_{ep}$ | input | 201 | 50 | 120 | 320 | 202 | 52 | 120 | 320 | 203 | 54 | 120 | 320 |
| $t_{ep}$ | input | 15 | 5 | 6 | 32 | 15 | 5 | 6 | 32 | 15 | 5 | 6 | 32 |
| $h_b$ | input | 327 | 86 | 180 | 678 | 327 | 94 | 180 | 678 | - | - | - | - |
| $t_{bf}$ | input | 12 | 3 | 6 | 19 | 12 | 3 | 7 | 19 | 11 | 3 | 6 | 19 |
| $h_{cw}$ | input | - | - | - | - | 225 | 70 | 98 | 398 | 225 | 78 | 98 | 398 |
| $t_{cw}$ | input | 12 | 5 | 5 | 30 | - | - | - | - | 12 | 6 | 5 | 30 |
| $b_{cf}$ | input | 231 | 60 | 120 | 407 | - | - | - | - | - | - | - | - |
| $t_{cf}$ | input | 20 | 10 | 7 | 40 | 20 | 10 | 7 | 40 | 20 | 11 | 7 | 40 |
| $d_b$ | input | 21 | 3 | 16 | 30 | 56 | 3 | 16 | 30 | 21 | 3 | 16 | 30 |
| $f_{u,b}$ | input | 955 | 139 | 440 | 1413 | - | - | - | - | - | - | - | - |
| $f_{y,P}$ | input | 351 | 141 | 221 | 1045 | - | - | - | - | 349 | 141 | 221 | 1045 |
| $f_{u,P}$ | input | - | - | - | - | - | - | - | - | 507 | 127 | 350 | 1079 |
| SC | input | 1.333 | 0.907 | 0 | 2 | 1.368 | 0.899 | 0 | 2 | 1.462 | 0.850 | 0 | 2 |
| Joint | input | - | - | - | - | 1.444 | 0.499 | 1 | 2 | - | - | - | - |
| $M_p$ | output | 96 | 61 | 18 | 377 | - | - | - | - | - | - | - | - |
| β | output | - | - | - | - | 1.2e-3 | 9.7e-4 | 1.4e-4 | 46e-4 | - | - | - | - |
| $K_s$ | output | - | - | - | - | - | - | - | - | 698 | 449 | 112 | 1790 |

Units: geometric parameters [mm], material parameters [MPa], moment [kN.m], stiffness [kN.m/rad], and normalized stiffness [mm-1/rad]

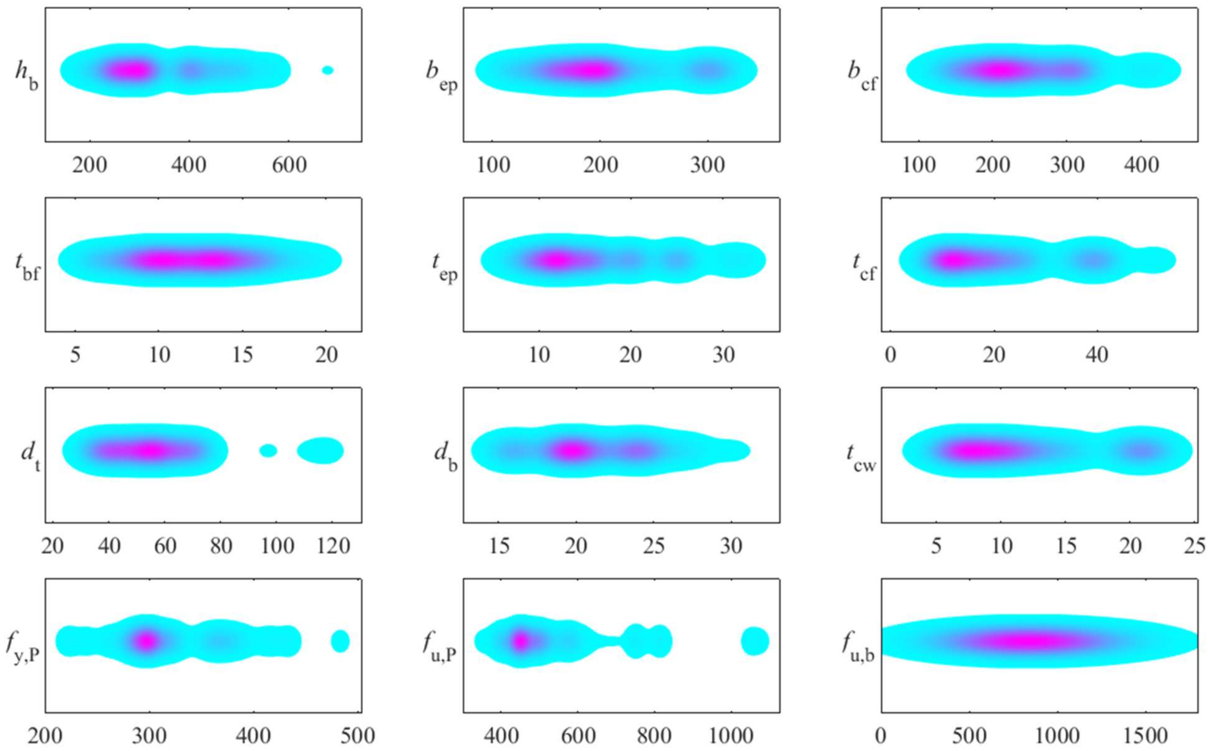



Figure 4. Kernell distribution plot of the significant features [units: mm and MPa]

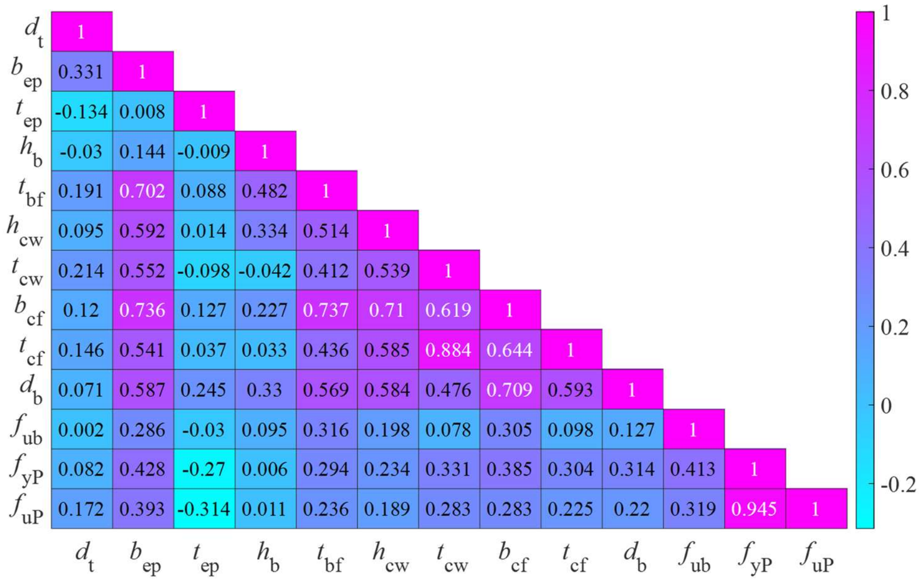



Figure 5. Correlation matrix of significant features

## ANN Architecture and Training

Three separate ANNs were trained to fit each of the three response parameters. The ANNs had the same simple layout as shown in Figure 6 with a single hidden layer. This was done intentionally, rather than utilizing deep ANNS, to simplify implementation in practice. Also, utilizing deep ANNs did not yield improved performance for the problem in question. The ANNs were developed within MATLAB® statistics and machine Learning toolbox (The MathWorks 2022) and utilized the Levenberg-Marquardt (LM) algorithm (Levenberg 1944; Marquardt 1963) for the back-propagation training. This algorithm is regularly used for regression analysis, due to its relatively fast computational speed for small and medium networks. The back-propagation method had a gradient target of $10^{-7}$. The training algorithm employed the mean squared error as a performance metric with a target value (fitness) of zero and a maximum number of 1000 epochs. All ANNs used the *hyperbolic tangent sigmoid* function as the activation function in the hidden layer. For the output layer, the $M_p$ network utilized the *linear transfer function* while the $\beta$ and $K_s$ networks utilized the *standard sigmoid transfer function*. The stiffness parameters are generally sensitive to input changes; therefore, the latter activation function was chosen to prevent their values from becoming negative.

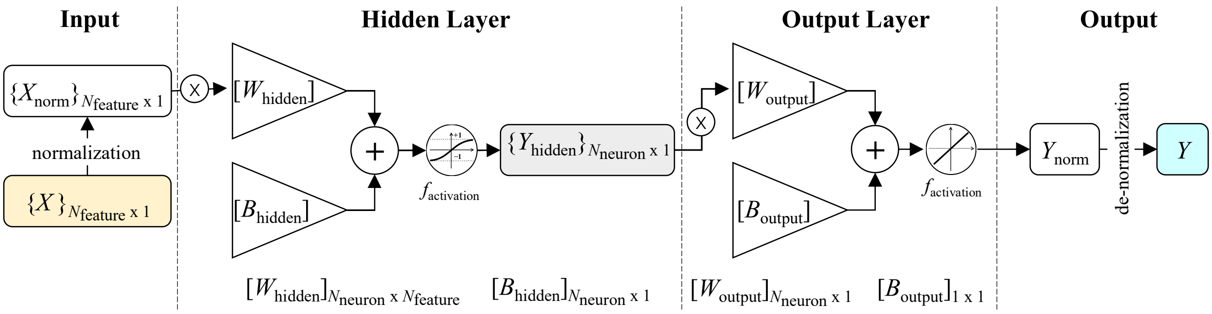


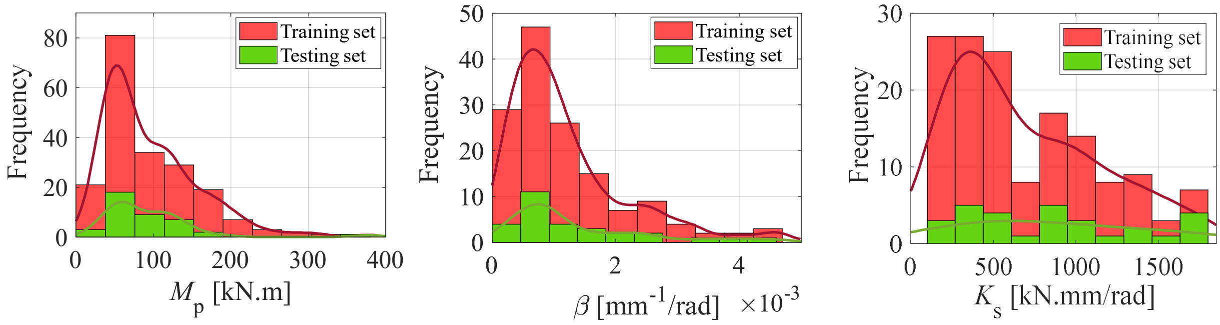

Figure 6. Outline of the feed-forward ANN archeticure

The number of neurons within the hidden layer was optimized, by targeting the least number of neurons that satisfies an acceptable performance threshold across the training and testing datasets. The hidden layer of the trained ANN models for $M_p$, $\beta$, and $K_s$ have 11, 11, and 12 neurons, respectively, as summarized in Table 4.

To help the training algorithm and avoid biased network parameters, due to the various scales from the different input parameters, both the input (features) and output (response parameter) values were normalized. There are two main approaches to data scale normalization as given by Equation 1: 1) the scaling to a range method (also referred to as the *MinMax* method), where $X_{\mathrm{min}}$ and $X_{\mathrm{max}}$ are the minimum and maximum values of input/output parameter $X$, respectively, and 2) the *z-score* method, where $\mu_X$ and $\sigma_X$ are the mean and standard deviation of input/output parameter $X$, respectively. The former scales the feature values between 0 and 1 while the later transforms the data to a new distribution with a mean equal to 0.0 and a standard deviation equal to 1. Both methods have advantages when it comes to training the model; hence, both are used for the different response parameters. Another reason for normalizing both the input and output parameters is to aid with the gradient descent step, by providing small weights and biases to the network. Because the *Tansig* function is bound between -1 and 1, if large weights and biases were trained, this could lead to the output of the units being predicted near the saturation regions of the function. The *z-score* scaling was used for the $M_{\mathrm{p}}$ network while the *MinMax* scaling was used for the $\beta$ and $K_{\mathrm{s}}$ networks. The latter was essential given the usage of the *Logsig* activation function for the stiffness networks.

$$X_{\mathrm{norm}} = \begin{cases} \dfrac{X - \mu_{\mathrm{X}}}{\sigma_{\mathrm{X}}} & \text{for } z - score \\ \dfrac{X - X_{\mathrm{min}}}{X_{\mathrm{max}} - X_{\mathrm{min}}} & \text{for } MinMax \end{cases} \quad (1)$$

The data were split randomly into training and testing subsets using an 80-20 split. A 20% of the training data was set for validation to help tuning the model's hyperparameters. To further improve the model training and minimize the risk of overfitting, care was taken to ensure that the target response parameter distribution is consistent between the training and testing datasets as demonstrated in Figure 7.




(a) (b) (c)

Figure 7. Consistent distribution of the training and testing datasets for each response parameter

Table 4. ANN parameters and training settings

| Network | $N_{specimens}$ | $N_{feature}$ | $N_{neuron}$ | Normalization | Activation function | |
|---|---|---|---|---|---|---|
| | | | | | Hidden layer | Output layer |
| $M_p$ | 198 | 12 | 11 | z-score | Tansig | Purelin |
| $\beta$ | 144 | 10 | 11 | MinMax | Tansig | Logsig |
| $K_s$ | 145 | 11 | 12 | MinMax | Tansig | Logsig |

Several studies in the literature suggested the employment of hybrid ANN models where the weights and biases of the network are optimized using available genetic- or bio-inspired computational algorithms, rather than utilizing gradient decent. This was tried in this study, and it was concluded that using such hybrid models did not achieve better model performance compared to traditional gradient decent. However, it was observed that employing such algorithms on the already-trained ANN can help in further improving the performance of the network. Specifically, once the networks were trained, the Particle Swarm Optimization (PSO) algorithm (Kennedy and Eberhart 1995) was applied to further optimize the weights and biases of the network. This generally resulted in about 5~10% improvement in the quality-of-fit metrics. The PSO used a swam of 75 particles, where each particle position represents a combination of the network weights and biases. One particle position was pre-set based on the trained network while the remaining 74 particle positions were random generated. A total of 2000 iterations were carried out where the weights and biases were bounded between -1.5 to 1.5 and -3 to 3 respectively, implementing the stochastic behaviour. This meant that the starting global best position of all the particles was the trained network's weights and bias vector. With each subsequent iteration, the particles converge towards this global best value with some randomness. By implementing this randomness, if a particle's personal best was found to be better than the current global best, this global best was updated for all particles, and subsequently improving the network.

## Model Performance and Interpretation

Figure 8 shows the predicted versus the observed values of the three response parameters based on the training, testing, and full datasets. The plot reflects the good agreement between the ANN model predictions and the observed data. This is particularly the case for the plastic strength predictions. The stiffness parameters ($\beta$ and $K_s$) show relatively larger, but not significant, variability which is understandable given the sensitivity of these response parameters. The plot also demonstrates the absence of overfitting given the consistent quality-of-fit metrics between the training and testing datasets. The same applies to other two response parameters $\beta$ and $K_s$. The high quality-of-fit metrics of the testing dataset confirm the generalizability and applicability of the model to new/unseen data that falls within the model features' applicability range.

Quantitatively, three model performance metrics are computed to evaluate the accuracy of the ANN model. These are the root mean square error (*RMSE*), the coefficient of determination ($R^2$), and the percentage of data falling within an error margin of ±15% ($P_{15}$). Employing multiple error metrics is key, when evaluating regression models, to assess potential overfitting and bias. The $R^2$ is the standard metric for evaluating the quality of fit in a regression model. The *RMSE* is a quadratic scoring rule which averages the magnitude of the error after squaring the errors, this gives rise to relatively high weight to large errors. This characteristic is desirable when large errors are unfavourable, particularly utilized in ANN models, where weights and biases impact how much a parameter influences the predicted output. The *RMSE* is also expressed in the same units of the predicted response parameter, making it easier to interpret the average error in each prediction. The $P_{15}$ metric provides a direct indication of how many predictions fall within an error range. The 15% value is chosen as it is generally regarded as an acceptable upper level of error in predictive models within the structural engineering practice.

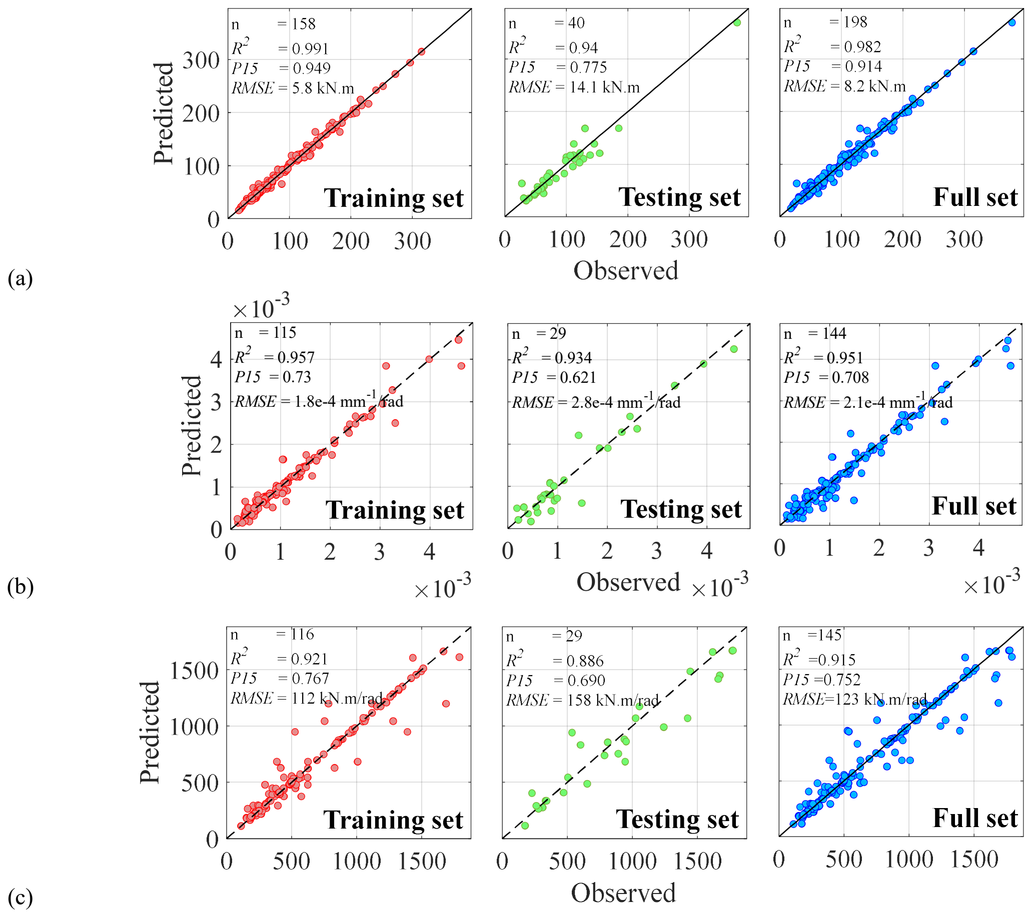



Figure 8. Predicted versus observed values: (a) $M_p$, (b) $\beta$, and (c) $K_s$

The quality-of-fit metrics are summarized in Table 5. Based on the full datasets, all models achieve an $R^2$ larger than 0.9. The mean error is 8.2 kN.m, 2.1x10$^{-4}$ mm$^{-1}$/rad, and 123.1 kN.m/rad for the $M_p$, $\beta$, and $K_s$ predictions, respectively. These mean errors are low as they constitute 8.5%, 17.5%, and 17.6%, respectively, with respect to the response parameters mean values (refer to Table 3). The high $R^2$ values are further collaborated with high $P_{15}$ values larger than 0.7, indicating no issue with overfitting. It should be noted that the quality-of-fit metrics for the stiffness parameters, $\beta$ and $K_s$, are not as high as that of $M_p$. This is expected, especially for $\beta$, since the stiffness is a very sensitive parameter that is affected by several connection details that are not accounted for herein, such as the fit between the different flat components (Mann and Morris 1981) and the bolt preload (Hellquist 1966; Faella et al. 1998). These limitations are further discussed at the end of this paper. Also, the elastic stiffness is sensitive to the method used to deduce it from test data, where typically, a variability of 10~15% is observed (Elkady 2022). Lastly, it is worth mentioning that maximum observed error in all parameters does not exceed 50% (i.e., $P_{50} \approx 1.0$).

Table 5. Summary of quality-of-fit metrics for the developed ANNs

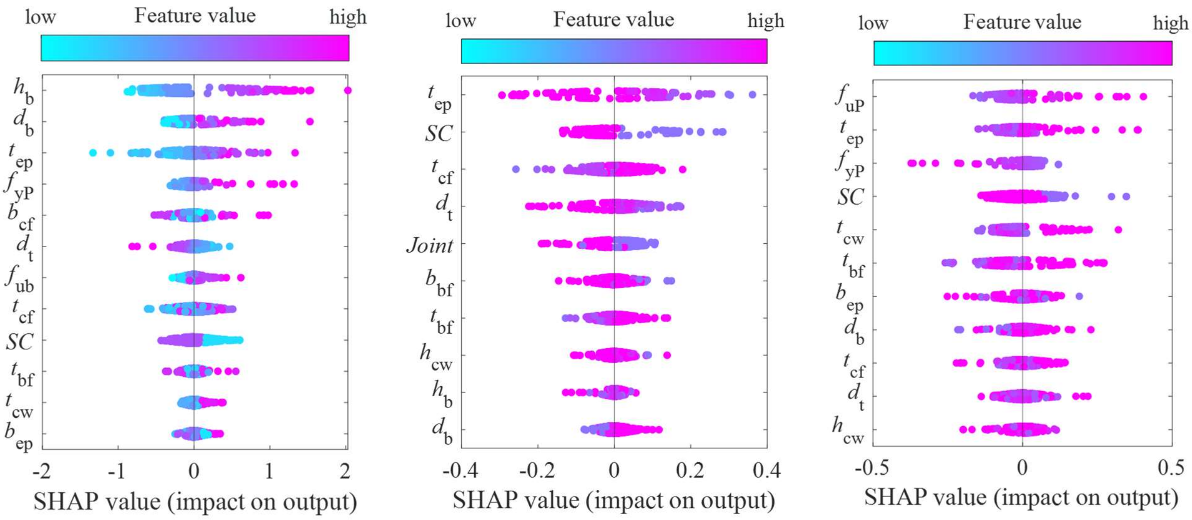
| Parameter | Training | | | Testing | | | Total | | |
|---|---|---|---|---|---|---|---|---|---|
| | $R^2$ | $P_{15}$ | RMSE | $R^2$ | $P_{15}$ | RMSE | $R^2$ | $P_{15}$ | RMSE |
| $M_p$ [kN.m] | 0.991 | 0.949 | 5.8 | 0.940 | 0.775 | 14.1 | 0.982 | 0.914 | 8.2 |
| $\beta$ [mm$^{-1}$/rad] | 0.957 | 0.730 | 1.8e-4 | 0.934 | 0.621 | 2.8e-4 | 0.951 | 0.708 | 2.1e-4 |
| $K_s$ [kN.m/rad] | 0.921 | 0.767 | 112.1 | 0.886 | 0.690 | 158.2 | 0.915 | 0.752 | 123.6 |

To shed light on the ANN model's inner workings, the model-agnostic ML interpreter known as SHapley Additive exPlanations (SHAP) (Lundberg and Lee 2017), is used here to explain how the input features affect the model output. Figure 9 shows the SHAP summary plots for the three networks; that is the distribution of SHAP values for each feature, which represents the magnitude of its positive/negative impact on the model output. Each point represents a single specimen, and the points are colour-coded with respect to the feature value.

With respect to $M_p$ (Figure 9(a)), the beam depth, the bolt diameter, the endplate thickness, and the endplate material's yield strength are ranked as the top features affecting the connections' plastic strength the most. All these features have a rational positive impact (correlation) on the value of $M_p$, as indicated by the colour code. This is expected and agrees with the mechanics of the problem (refer to Figure 3). The column flange width, and the distance between the tension bolt row and the beam flange follow in terms of significance. Column stiffening (*SC*) has a lower impact, where it is observed that stiffened joints (i.e., lower numeric *SC* value) tend to develop larger $M_p$, as expected. Lastly, the SHAP plot shows that the endplate width, the column web thickness, and the beam flange thickness have the least impact. Although $b_{ep}$ is chosen for its importance from a mechanical point of view, its apparent low impact can be attributed to its high correlation with $b_{cf}$ as highlighted earlier (see Figure 5). For $t_{cw}$ and $t_{bf}$, their low impact can be explained by the fact that plastification in most FEPC specimens was not controlled by beam flange buckling or column-web shear buckling. Designers commonly try to avoid these deformation modes to improve repairability and maintain structural stability.

For the elastic stiffness coefficient $\beta$ (Figure 9(b)), other significant features arise. The $t_{ep}$ and $d_t$ features are among the top predictors since these parameters control the rigidity of the endplate which is the primary deforming component in most tests. Similarly, the *SC* and $t_{cf}$ are among the top four predictors since these two control the rigidity of the column flange which is the second most probable component to deform in most tests, following the endplate. The *Joint* feature, which is unique for the $\beta$ ANN, is also important where cruciform/interior joints (encoded with the lower value of 1.0) result in a stiffer behaviour compared to cantilever/exterior joints (encoded with the larger value of 2.0). In the former, under symmetric loading conditions, the deformation of the column-web panel zone is limited. The bolt diameter ($d_b$) has the least impact on stiffness since FEPCs are mostly designed/tested based on a thin-plate strong-bolt approach. As such, the potential of bolt elongation, which can affect the rotational stiffness, is limited.

Similar observations are made with respect to $K_s$ (Figure 9(c)). Most notably, the yield and ultimate stresses of the endplate's material are among the top predictors. The ratio of these two mainly controls (with a positive correlation) the steepness of the post-yield slope. Therefore, the model should be valid for any FEPC fabricated from carbon, stainless, or high-strength steel. In summary, the SHAP plots confirm the validity of the ANN model from a mechanics point of view and that the observed model accuracy is not a result of blind overfitting.




(a) (b) (c)

Figure 9. SHAP value summary plots for the (a) $M_p$, (b) $\beta$, and (c) $K_s$ ANN models

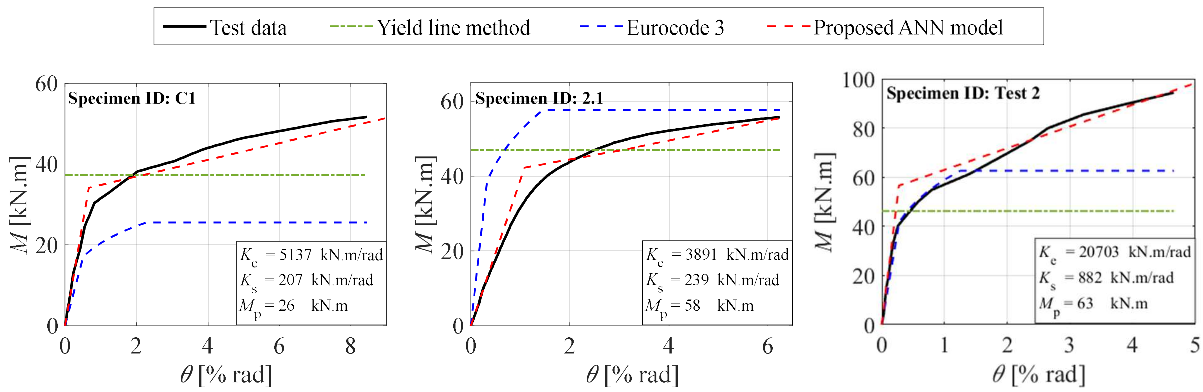
The predictions of the ANN model are further compared with moment-rotation test data. Figure 10 shows comparisons with nine tests on full-scale beam-to-column joints with FEPCs with different configurations (stiffened/unstiffened and exterior/interior) and different materials (carbon, stainless, and high-strength steel). In the same figure, predictions by the yield line method (Murray and Shoemaker 2002; AISC 2016; Eatherton et al. 2021) and the Eurocode 3 component method (CEN 2005), are superimposed for reference. The former predicts the plastic moment while the latter predicts both the plastic moment and the elastic rotational stiffness. Note that the yield line and the component methods were shown to be relatively better, in terms of accuracy and consistency of results, compared to other available models that are summarized in Table 1 (Ding and Elkady 2023b, a). For reference, the error in $M_p$ predicted by the yield line method ranges between ±27% and could reach up to ±60%. The error in $M_p$ predicted by the component method ranges between ±36% and could reach up to -70% and +177%. The error in $K_e$ predicted by the component method ranges between ±75% and could reach up to -94% and +286%.

In the cases shown in this figure, but also in others, the ANN model provides an accurate prediction of $M_p$ compared to the other models, by closely reproducing the transition point between the elastic and plastic slopes. This is already expected given the high quality-of-fit metrics for this response parameter.

Specifically, the error in predicting $M_p$ based on the ANN model falls mostly within the ±15% range and does not exceed 35%.

Concerning the elastic rotational stiffness, the model is notably good at capturing the true $K_e$ of the connection where the component method fails. The ANN model requires a limited number of input parameters and does not involve complex or lengthy computations as the component method. Furthermore, contrary to the component method, the ANN model can capture the strain-hardening branch. This is best demonstrated in the specimens tested by Bose et al. (1996) and Rölle (2013) with extreme values for the strain hardening slope (see Figure 10(e-f)). Capturing the post-yield slope, rather than assuming a perfectly plastic behaviour, is important as it is not a trivial value that can reach up to 20% $K_e$ in some cases (Elkady 2022). This makes it critical in numerical simulations concerned with the plastic behaviour of the joint under extreme events, such as column loss scenarios and collapse-level earthquakes.

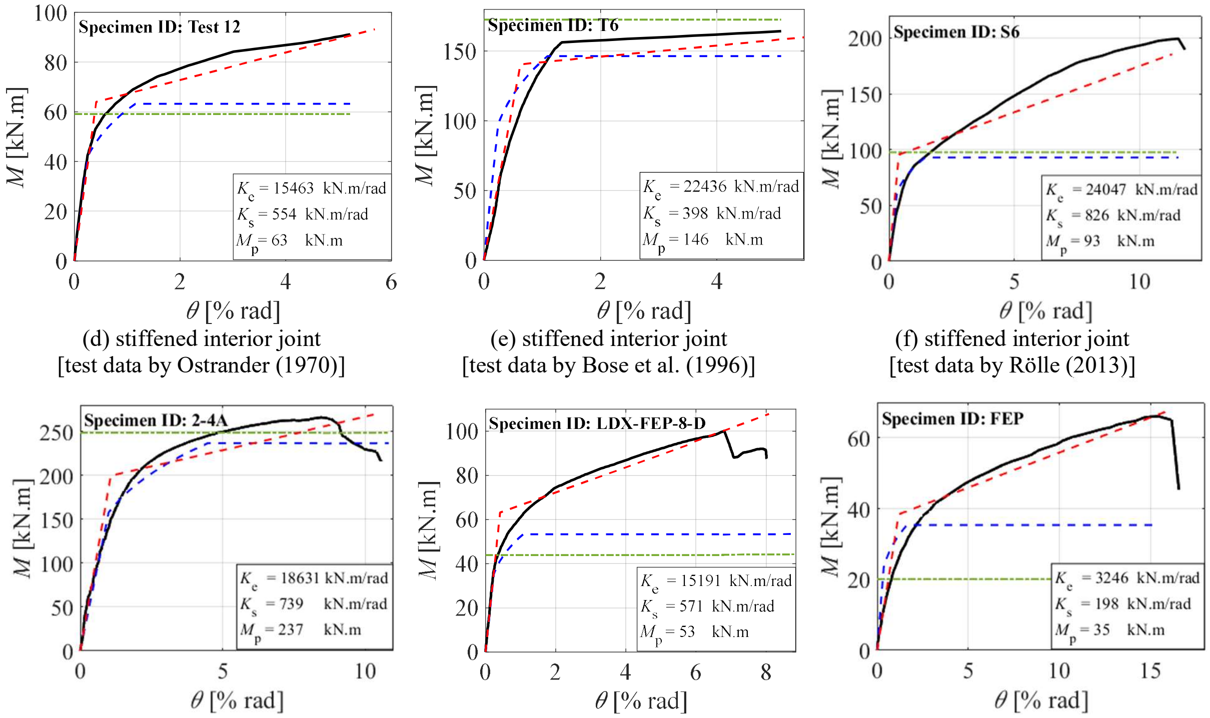
For FEPCs with endplates fabricated from stainless steel (Figure 10(h-i)) or high-strength steel (Figure 10(g)), the model performance is consistent with those fabricated from conventional carbon steel. In summary, these comparisons demonstrate the superior predictions of the ANN model across different connection configurations and materials, regardless of the developed damage mode(s).




(a) stiffened exterior joint [test data by Aggarwal (1994)]

(b) unstiffened exterior joint [test data by Thomson and Broderick (2002)]

(c) unstiffened exterior joint [test data by Zoetemeijer (1974)]




(d) stiffened interior joint [test data by Ostrander (1970)]
(e) stiffened interior joint [test data by Bose et al. (1996)]
(f) stiffened interior joint [test data by Rölle (2013)]
(g) unstiffened exterior high-strength joint [test data by Qiang et al. (2014)]
(h) stiffened exterior stainless steel joint [test data by Song et al. (2022)]
(i) unstiffened interior stainless steel joint [test data by Elflah et al. (2019)]

Figure 10. Sample comparisons of test data and predictions by the proposed ANN model, the yield line method, and the Eurocode 3 component method

## Model Utilization

To aid users in employing the proposed ANN model in design and numerical analysis, the mathematical procedure, to obtain the response parameters, is outlined in this section. For a given response parameter ($Y$), the steps to conduct a feed-forward pass through the ANN are as follows:

1) Create a column vector of the input features (*input*), relevant to the response parameter, of size $N_{\text{feature}}$, in the same order outlined in Table 2.
2) Normalize each of the input features ($X_{\text{norm}}$) using the appropriate normalization formula (Equation (1)) that corresponds to the target response parameter (refer to Table 4). The statistical parameters for each feature given a particular response parameter are outlined in Table 3.
3) Multiply the normalized input vector with the input-layer weight matrix ($W_{\text{hidden}}$), add the product with the input-layer bias vector ($B_{\text{hidden}}$), and then apply the hidden layer activation function (refer

to Table 4). For the $M_p$ ANN, this is the *Transig* function (called *tanh* in Matlab). For the $\beta$ and $K_s$ ANNs, this is the *Logsig* function (called *logsig* in Matlab). Tables 6 to 8 summarize the weight matrices and bias vectors of the hidden and output layers, for each response parameter.

$$Y_{\text{hidden}} = f_{\text{activation,hidden}}(W_{\text{hidden}} \cdot X_{\text{norm}} + B_{\text{hidden}}) \tag{2}$$

4) Multiply the output of the hidden layer ($Y_{\text{hidden}}$) by the output-layer weight matrix ($W_{\text{output}}$) then add the product with the output-layer bias vector ($B_{\text{output}}$), as shown in Equation (3).

$$Y_{\text{output}} = f_{\text{activation,output}}\left(W_{\text{output}} \cdot Y_{\text{hidden}} + B_{\text{ouput}}\right) \tag{3}$$

5) Lastly, de-normalize the output-layer output ($Y_{\text{output}}$) to obtain the response parameter ($Y$) using Equation (4), depending on the ANN normalization method (refer to Table 4). To obtain $K_e$, the $\beta$ value needs to be multiplied by the beam's rigidity $EI_x$.

$$Y = \begin{cases} Y_{\text{hidden}} \cdot \sigma_Y + \mu_Y & \text{for } z - score \\ Y_{\text{hidden}} \cdot (Y_{\min} - Y_{\min}) + Y_{\min} & \text{for } MinMax \end{cases} \tag{4}$$

The aforementioned procedure and the provided ANN data can be used as part of codified algorithms to automate the design or numerical modelling of FEPCs. The implementation of these mathematical procedures within a MATLAB script is made publicly available and downloadable from a GitHub repository (Elkady 2023). Furthermore, for quick and simple utilization of the ANN model, a computer tool with a friendly graphical user interface (GUI) is developed as shown in Figure 11(a). The GUI as well as the experimental database are available from the GitHub repository. The GUI also includes an optimization module to allow the users to optimize the connection design as shown in Figure 11(b). Specifically, this module requires the input of the connection's basic parameters such as the beam and column sections, the material properties, and the joint configuration. The user then needs to specify the target strength and stiffness parameters and the corresponding optimization weights. The module employs the ANN model described herein coupled with the particle swarm optimization algorithm (Kennedy and Eberhart 1995) to find the optimum endplate thickness, bolt diameter, gauge length, and tension row extension ($e_{rt}$).

Table 6. Weights and biases for the $M_p$ ANN

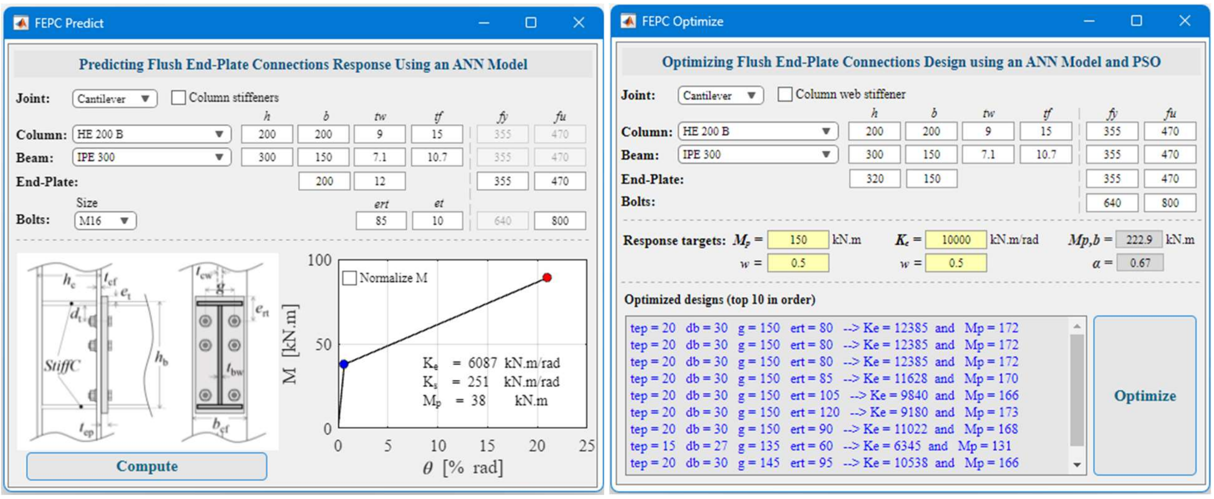
| $W_{hidden}$ [$N_{neuron}$ x $N_{feature}$] | | | | | | | | | | | | $B_{hidden}$ | $W_{output}$ | $B_{output}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| -0.1561 | -0.1052 | -0.5051 | 1.5925 | 0.2534 | -0.2811 | -0.1005 | 0.2101 | -1.2479 | -0.5133 | -0.4991 | -0.0778 | 2.3675 | 1.7993 | 0.4793 |
| 0.7769 | 0.1154 | -0.7984 | -0.0793 | 0.2784 | 0.4110 | -0.0957 | -0.1144 | -0.2526 | -0.0101 | 0.3160 | -0.1265 | -1.6290 | 0.8476 | |
| -0.0500 | 0.0293 | -0.3146 | -0.6680 | 0.6408 | 0.5310 | 0.0058 | -0.0551 | -0.7226 | 0.2243 | 0.5272 | 0.2861 | -1.0908 | -1.4234 | |
| 0.0114 | -0.5980 | -0.3546 | 0.1867 | 0.6470 | -0.5688 | -0.0934 | -0.4429 | -0.4474 | 0.0241 | 0.1160 | -0.1886 | -0.2539 | 0.7454 | |
| 0.3194 | 0.0201 | 0.4598 | -0.7007 | 0.1606 | 0.5038 | 0.1443 | -0.6837 | 0.6254 | 0.8730 | 0.2856 | 0.5911 | -1.4837 | 1.8515 | |
| 0.7984 | -0.0064 | -0.0299 | -0.3462 | -0.3462 | -0.0083 | -0.0431 | -0.4212 | 0.1476 | -0.2122 | 0.1855 | 0.6052 | -0.7678 | -0.4389 | |
| -0.0430 | -0.0889 | 0.1277 | -0.5202 | 1.2079 | 0.5101 | -0.5116 | -0.2164 | 0.2240 | -0.0235 | 0.6631 | 0.3020 | -0.3504 | 0.8081 | |
| -0.0459 | -0.3166 | -0.3118 | 0.3523 | -0.0074 | -0.2669 | -0.2262 | -0.4506 | -0.5188 | 0.0139 | -0.0449 | -0.4632 | -1.1475 | -0.9078 | |
| 0.7803 | 0.1481 | -0.0001 | -0.3586 | -0.3353 | 0.1846 | -0.3385 | -0.6374 | -0.1555 | -0.1041 | -0.3382 | -0.3220 | 2.2969 | -1.7263 | |
| 0.3352 | -0.0015 | 0.0989 | 0.3158 | -0.1477 | 0.0260 | 0.3469 | -0.2349 | -0.3160 | -0.1867 | 0.2090 | -0.2275 | 2.0693 | 0.5610 | |
| 0.6874 | 0.0948 | 0.5057 | -0.0202 | -0.2654 | -0.3087 | 0.0484 | 0.2451 | -0.0682 | -0.3652 | 0.0797 | 0.1407 | 1.3155 | -0.2270 | |

Table 7. Weights and biases for the $\beta$ ANN

| $W_{hidden}$ [$N_{neuron}$ x $N_{feature}$] | | | | | | | | | | $B_{hidden}$ | $W_{output}$ | $B_{output}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.6737 | -3.5636 | -5.9795 | -1.3544 | -2.3843 | -2.6358 | 4.3338 | 0.0051 | 1.4255 | -1.1912 | 2.5631 | 3.2652 | 0.5738 |
| 2.0886 | 1.6148 | 1.6487 | 2.4756 | 0.4893 | 1.0669 | -0.8688 | 0.5226 | 1.0163 | -0.6257 | -3.8645 | 1.4844 | |
| 2.1606 | 0.6211 | 2.3953 | 0.1015 | -1.1375 | 0.4203 | 1.6150 | -2.7943 | -2.5594 | -0.7846 | 0.4896 | -3.2233 | |
| -0.0491 | 1.1622 | 2.3309 | 0.3925 | 0.6881 | 1.4883 | -0.4225 | -1.5725 | 0.4301 | 1.3363 | -1.7660 | 4.3650 | |
| -1.8521 | 0.8547 | -2.0596 | 0.7335 | 2.5099 | 0.8620 | -1.4369 | -0.7955 | -1.8749 | -1.5590 | 1.7896 | 2.5703 | |
| 0.1169 | 0.3342 | 0.9147 | 1.9614 | 1.7312 | -0.1068 | 1.8641 | -0.1118 | 2.7481 | -1.5959 | -3.3999 | -2.3488 | |
| 1.5792 | -1.7161 | 1.4678 | 2.0036 | 0.3553 | -0.7988 | -1.5714 | -3.4534 | -2.0033 | 0.7603 | -0.0082 | 1.4426 | |
| -2.4398 | -0.4705 | -1.0927 | 0.6553 | 1.9301 | 1.7510 | 1.0317 | -1.4965 | 3.0194 | 2.9240 | -3.5538 | -2.4822 | |
| -0.9291 | 3.2119 | 1.2840 | -2.9120 | -0.8905 | -0.0193 | -2.3480 | -2.0539 | -2.8492 | -1.8051 | 1.9955 | -2.2881 | |
| 1.5825 | -0.5015 | 0.7783 | -0.7622 | 0.9383 | 1.2880 | 1.5855 | -1.6543 | 0.6486 | -0.4441 | 0.6443 | -0.2626 | |
| -0.5152 | -1.6942 | 0.6808 | 2.9432 | 0.7779 | 2.1363 | 1.0909 | 1.1213 | -1.8067 | 2.2180 | -0.0868 | -2.9505 | |

Table 8. Weights and biases for the $K_s$ ANN

| $W_{hidden}$ [$N_{neuron}$ x $N_{feature}$] | | | | | | | | | | | $B_{hidden}$ | $W_{output}$ | $B_{output}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.0964 | 3.6558 | 0.1004 | 4.4250 | -0.3974 | -1.4584 | 2.1737 | -0.8295 | -1.9409 | 0.1563 | 2.5873 | -3.9950 | 2.9940 | 1.1173 |
| 1.9564 | 3.0432 | -1.8556 | 3.5673 | 1.0240 | -1.4487 | 2.7917 | 4.0812 | 1.1277 | -1.6983 | -0.3383 | -5.8306 | 4.3030 | |
| 2.0783 | -0.6534 | 0.9335 | 4.7489 | 4.6275 | -3.2422 | 4.7006 | 2.3653 | -0.4612 | -0.7378 | -2.6238 | -5.3228 | -3.0635 | |
| -0.8262 | -3.3380 | -1.5086 | 3.6743 | -0.6628 | -1.5553 | -1.6243 | -2.2956 | 0.9421 | 2.7094 | -1.2926 | -0.0044 | 4.8968 | |
| -1.2614 | 2.3680 | 1.0986 | -2.3004 | 0.2709 | -0.6989 | -0.4501 | 2.1309 | 1.6630 | 3.1845 | -2.5515 | -1.4585 | -5.6332 | |
| -1.6697 | -1.7323 | 1.0516 | -0.9521 | 0.1916 | 1.7200 | -0.3316 | 0.0305 | -0.1203 | -0.8022 | 1.0066 | -2.0436 | -0.3030 | |
| 1.7964 | -0.8479 | -4.9731 | 0.4109 | -6.4301 | 1.1970 | -1.7062 | 0.1930 | -2.1441 | 0.8676 | 0.1087 | 3.8783 | -3.9594 | |
| -0.8394 | 1.5316 | 1.3979 | 0.0888 | -2.4408 | 0.8070 | -0.6903 | 1.9504 | -0.2030 | 1.8436 | 2.4091 | -2.9640 | 2.6221 | |
| 3.4067 | 3.3128 | 0.3067 | 3.6878 | 2.1387 | -3.1742 | 0.2058 | -0.6239 | -0.4970 | 0.3674 | 3.5597 | -1.7865 | -3.0713 | |
| -0.0299 | -0.4642 | 2.8773 | 4.1044 | -3.0830 | 3.0194 | 0.7973 | -3.3211 | 0.2638 | 0.6996 | -1.2593 | -2.7384 | -3.2549 | |
| -0.8880 | -1.3525 | -1.4061 | 1.4469 | 1.1228 | 0.4185 | -0.8428 | 1.5279 | -0.3123 | -1.5189 | 2.2137 | -0.7258 | -2.0232 | |
| 0.0123 | -0.6193 | 1.6863 | -0.2872 | 0.5602 | 4.2592 | 0.0963 | 3.3581 | 2.7626 | -2.7253 | -3.6252 | 0.8133 | 5.2351 | |




(a) (b)

Figure 11. A computer tool with a graphical user interface for FEPC (a) response prediction, and (b) design optimization

## Summary and Conclusions

Predicting the response of bolted semi-rigid steel connections can be challenging. It is demonstrated in the literature that existing and traditional analytical, empirical, and mechanical models do not yield consistent and accurate predictions even for the connection's fundamental parameters. This constituted a strong justification for utilizing available machine-learning methodologies to capture the complexity of such connections. In this paper, an artificial neural network (ANN) is developed for flush endplate connections (FEPCs) with an emphasis on bare steel beam-to-column connection types with I-shaped columns and major axis orientation. The availability of a large experimental database enabled the training of the ANN model. The proposed model employs a total of 13 geometric, material, and layout features to predict the elastic rotational stiffness, plastic strength, and post-yield stiffness of FEPCs. Therefore, the model provides a bilinear characterization of the connection while considering the strain-hardening slope, which is commonly either ignored or poorly characterized in past models.

The proposed ANN model provides robust predictions with a coefficient of determination $R^2$ larger than 0.9 for the three predicted response parameters. The error in predicting the response parameters for 70% of the specimens did not exceed 15%. The maximum observed error (mainly for the stiffness parameter) does not exceed 50%. The model constitutes an advancement in prediction accuracy over the yield line method and the Eurocode 3 component method while avoiding the computation complexity of the latter. This can help achieve efficient and optimized designs as well as more accurate numerical simulations.

To streamline the model implementation is research and industry practice, 1) a computer tool is developed and made publicly available, and 2) a full guide to the mathematical construction of the model is described in detail. Finally, the argument against the ability to understand Blackbox models is addressed by employing recent advancements in machine-learning models' explanatory algorithms to confirm the consistency of the model inner workings with the basics of joint mechanics.

## Model Limitations and Future Improvements

The developed model is limited to bare steel beam-to-column joints with major-axis orientation. The model applicability range covers a wide range of geometric and material design space. It is shown that this model achieves high accuracy in its predictions, particularly for the plastic strength. Some of the model's limitations are summarized below which are being addressed as part of future improvements to the model.

- Currently, there is not enough test data to establish a proper model for the post-peak response, i.e., the degrading post-capping plastic rotation and the residual moment. However, for four-bolt FEPCs, based on available research and from a conservative point of view, it is reasonable to assume that complete failure is coincident with the connection capacity.
- The model does not account for the column axial load and the beam axial load. Both of these can have a noticeable effect on both the plastic strength and elastic stiffness. The former affects the column-web panel zone stiffness and strength (Skiadopoulos et al. 2021). The latter can be critical in joints undergoing catenary action under column-loss scenarios, particularly in the posy-yield phase at large rotations (Izzuddin et al. 2007; Kukla and Kozlowski 2021). There was not enough data in the collected database to establish a clear correlation concerning these two parameters.

- The developed model does not predict the connection ductility, i.e., the rotation at failure. This is an important response parameter that is missing from most available models for semi-rigid bolted connections. This is attributed, in part, to the limited number of available experimental data, considering the multiple possible failure modes in FEPCs, such as bolt rupture, bolt stripping, weld failure, and plate tearing as well as the fact that most tests did not reach failure. Only two models were found by the authors that predict the ultimate/failure rotation (Ostrowski and Kozłowski 2017; Eladly and Schafer 2021). However, these models were shown to be inconsistent or inaccurate in that respect (Ding and Elkady 2023b, a). To be able to provide a more robust estimate of $\theta_f$, for failure modes controlled by bolt rupture, more data is needed to cover different bolt sizes and grades. Parametric finite element simulations, that account for the aleatoric uncertainty associated with the fracture phenomena, can be used for this purpose. For other failure modes, e.g., weld failure and plate tearing, a probabilistic model (e.g., a fragility model) would be more appropriate. This can achieved using the data summarized in Elkady (2022).
- The bolt preload ($P_t$) has a directly proportional impact on the connection stiffness. Past research showed that preloaded connections can develop up to double the elastic stiffness of un-preloaded ones (Hellquist 1966; Prescott 1987; Kline 1989; Jaspart and Maquoi 1995; Faella et al. 1998; Broderick and Thomson 2002). Nonetheless, this parameter is commonly ignored in existing predictive models. The challenge in utilizing the bolt preload as a feature in this study is attributed to the uncertainty associated with the values reported in the literature. Very limited researchers reported the actual $P_t$ value as measured through strain gauges or calibrated bolt-torquing methods. On the other hand, the majority reported either the presence/absence of bolt preload or the torque value used. Transforming the latter into a force induces notable uncertainty. The model performance for stiffness predictions can surely be further improved. This can be achieved by supplementing the experimental database with numerical data generated through thoroughly validated parametric finite element simulations.
- The model represents the moment-rotation response as a bilinear one. In the literature, several researches aimed to further capture the smooth power-shaped response using the popular Ramberg and

Osgood (1943) or Richard and Abbott (1975) models. Knowing the main response parameters of the connection (i.e., the pivot points), the parameters of these nonlinear models can simply be calibrated against available test data. This, however, will not affect structural design and is not expected to have major impact on overall system-level response. Similarly, the proposed model can be used to calibrate available hysteretic models with pinched response (e.g., Ibarra et al. (2005) and Kottari et al. (2014)).

## Data Availability

The experimental database upon which the current work is based, as well as the developed computer tool are publicly available for download through the second author's GitHub repository. Other data, models, or code that support the findings of this study can be made available from the corresponding author upon reasonable request.


## Acknowledgements

This work was conducted at the National Infrastructure Laboratory, University of Southampton (UoS). The authors gratefully acknowledge the financial support provided by UoS to the first author as part of his studentship.


## Notation

The following symbols are used in this paper:

| | |
|---|---|
| $b_{cf}$ | column flange width |
| $b_{ep}$ | endplate width |
| $B_{hidden}$ | bias vector for the hidden layer |
| $B_{output}$ | bias scalar for the output layer |
| $d_b$ | bolt diameter |
| $d_t$ | distance between the top bolt row and column flange center in tension |
| $E$ | modulus of elasticity |
| $e_{rt}$ | distance between the top bolt row and endplate edge in tension |
| $e_t$ | endplate extension at the tension side |
| $f_{y,i}$ | yield stress of component *i* [P: endplate, C: column, B: beam, b: bolt] |
| $f_{u,i}$ | ultimate stress of component *i* [P: endplate, C: column, B: beam, b: bolt] |
| $g$ | bolt gauge length |
| $h_b$ | beam depth |
| $h_c$ | column depth |

| | |
|---|---|
| $h_{cw}$ | column web depth |
| $I_x$ | beam second moment of inertia about the strong-axis |
| *Joint* | encoded categorical variable for the joint configuration |
| $K_e$ | initial elastic rotational stiffness |
| $K_s$ | post-yield hardening stiffness based on an equal-area bilinear fit |
| $L_b$ | beam length |
| $m$ | the bolt's inner horizontal distance to the nearest edge [see Eurocode 3, Part 1-8] |
| $m_2$ | the bolt's inner vertical distance to the nearest edge [see Eurocode 3, Part 1-8] |
| $M_c$ | capping moment |
| $M_p$ | the connection's equivalent yield (plastic) moment based on an equal-area bilinear fit |
| $M_{p,b}$ | the beam's plastic moment |
| $M_y$ | yield moment |
| $N$ | number of data points |
| $N_{feature}$ | number of features |
| $N_{neuron}$ | number of neurons in the hidden layer |
| $P_{15}$ | percentage of specimens with a prediction error equal to or less than 15% |
| $P_{y,b}$ | bolt yield load |
| $P_t$ | bolt preload |
| $p_f$ | distance from tension bolt centre to the centre line of beam tension flange |
| $p_t$ | bolt row pitch above and below the tension flange |
| $R^2$ | coefficient of determination |
| *RMSE* | root mean squared error |
| *SC* | encoded categorical variable for the column stiffener configuration |
| $t_{bf}$ | beam flange thickness |
| $t_{bw}$ | beam web thickness |
| $t_{cf}$ | column flange thickness |
| $t_{cw}$ | column web thickness |
| $t_{ep}$ | endplate thickness |
| $W_{hidden}$ | weight matrix for the hidden layer |
| $W_{output}$ | weight matrix for the output layer |
| $X$ | input parameter |
| $Y$ | output parameter |
| $z_{ex}$ | arm length between extreme bolt rows |
| $z_i$ | arm length between bolt row *i* and beam compression flange center |
| $\alpha$ | normalized moment |
| $\beta$ | normalized elastic stiffness |
| $\theta_f$ | failure rotation |
| $\mu$ | mean |
| $\sigma$ | standard deviation |

## References

Abdalla, K. M., and Stavroulakis, G. E. (1995). "A backpropagation neural network model for semi-rigid steel connections." *Computer-Aided Civil and Infrastructure Engineering*, 10(2), 77-87, DOI: 10.1111/j.1467-8667.1995.tb00271.x.

Abolmaali, A., Matthys, J. H., Farooqi, M., and Choi, Y. (2005). "Development of moment–rotation model equations for flush end-plate connections." *Journal of Constructional Steel Research*, 61(12), 1595-1612, DOI: 10.1016/j.jcsr.2005.05.004.

Aggarwal, A. K. (1994). "Comparative Tests on End Plate Beam-to-Column Connections." *Journal of Constructional Steel Research*, 30(2), 151-175, DOI: 10.1016/0143-974X(94)90048-5.

AISC (2016). "Prequalified connections for special and intermediate steel moment frames for seismic applications." *ANSI/AISC 358-16*, Chicago, IL.

Benterkia, Z. (1991). "End-plate connections and analysis of semi-rigid steel frames." Ph.D. Thesis, University of Warwick, Coventry, UK.

Bose, B., Youngson, J. K., and Wang, Z. M. (1996). "An appraisal of the design rules in Eurocode 3 for bolted end-plate joints by comparison with experimental results." *Proceedings of the Institution of Civil Engineers - Structures and Buildings*, 116(2), 221-234, DOI: 10.1680/istbu.1996.28289.

Broderick, B. M., and Thomson, A. W. (2002). "The response of flush end-plate joints under earthquake loading." *Journal of Constructional Steel Research*, 58(9), 1161-1175, DOI: 10.1016/S0143-974X(01)00073-6.

Brown, N. D., and Anderson, D. (2001). "Structural properties of composite major axis end plate connections." *Journal of Constructional Steel Research*, 57(3), 327-349, DOI: 10.1016/S0143-974X(00)00034-1.

Brown, N. D., Hughes, A. F., and Anderson, D. (2001). "Prediction of the initial stiffness of ductile end-plate steel connections." *Proceedings of the Institution of Civil Engineers - Structures and Buildings*, 146(1), 17-29, DOI: 10.1680/stbu.2001.146.1.17.

CEN (2005). "Eurocode 3 - Design of Steel Structures, Part 1-8: Design of Joints." *BS-EN 1993-1-8-2006*, Brussels, Belgium.

De Lima, L. R. O., Vellasco, P. C. G. d. S., De Andrade, S. A. L., Da Silva, J. G. S., and Vellasco, M. M. B. R. (2005). "Neural networks assessment of beam-to-column joints." *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 27(3), 314-324, DOI: 10.1590/S1678-58782005000300015.

Ding, Z., and Elkady, A. (2023a). "Accuracy assessment of predictive models for semi-rigid extended end-plate connections." *Proc., 10th Eurosteel Conference*, Amsterdam, Netherlands.

Ding, Z., and Elkady, A. (2023b). "Semirigid bolted end-plate moment connections: Review and experimental-based assessment of available predictive models." *Journal of Structural Engineering*, 149(9), DOI: 10.1061/JSENDH.STENG-11797.

Eatherton, M. R., Nguyen, T. N., and Murray, T. M. (2021). "Yield Line Patterns for End-Plate Moment Connections." *Report No. CE/VPI-ST-21/05*, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA.

Eladly, M. M., and Schafer, B. W. (2021). "Numerical and analytical study of stainless steel beam-to-column extended end-plate connections." *Engineering Structures*, 240, 112392, DOI: 10.1016/j.engstruct.2021.112392.

Elflah, M., Theofanous, M., Dirar, S., and Yuan, H. (2019). "Behaviour of stainless steel beam-to-column joints — Part 1: Experimental investigation." *Journal of Constructional Steel Research*, 152, 183-193, DOI: 10.1016/j.jcsr.2018.02.040.

Elkady, A. (2022). "Response characteristics of flush end-plate connections." *Engineering Structures*, 269, DOI: 10.1016/j.engstruct.2022.114856.

Elkady, A. (2023). "Semi-rigid connections experimental database (SRConED)." GitHub, <https://github.com/amaelkady/SRConED>.

Elkady, A., and Mak, L. (2022). "Data driven evaluation of existing numerical modelling guidelines for semi-rigid connections." *Proc., 10th International Conference on Behaviour of Steel Structures in Seismic Areas (STESSA)*, Timisoara, Romania, 244-251.

Faella, C., Piluso, V., and Rizzano, G. (1998). "Experimental analysis of bolted connections: Snug versus preloaded bolts." *Journal of Structural Engineering*, 124(7), 765-774, DOI: 10.1061/(ASCE)0733-9445(1998)124:7(765).

Faridmehr, I., Nikoo, M., Pucinotti, R., and Bedon, C. (2021). "Application of component-based mechanical models and artificial intelligence to bolted beam-to-column connections." *Applied Sciences*, 11(5), 2297, DOI: 10.3390/app11052297.

Frye, M. J., and Morris, G. A. (1975). "Analysis of flexibly connected steel frames." *Canadian Journal of Civil Engineering*, 2(3), 280-291, DOI: 10.1139/l75-026.

Ghassemieh, M., and Nasseri, M. (2012). "Evaluation of stiffened end-plate moment connection through optimized artificial neural network." *Journal of Software Engineering and Applications*, 5(3), DOI: 10.4236/jsea.2012.53023.

Hellquist, T. I. (1966). "The Behaviour of End Plate Connections." M.Sc. Thesis, University of Saskatchewan, Saskatchewan, Canada.

Ibarra, L. F., Medina, R. A., and Krawinkler, H. (2005). "Hysteretic Models that Incorporate Strength and Stiffness Deterioration." *Earthquake Engineering & Structural Dynamics*, 34(12), 1489-1511, DOI: 10.1002/eqe.495.

Izzuddin, B. A., Vlassis, A. G., Elghazouli, A. Y., and Nethercot, D. A. (2007). "Assessment of progressive collapse in multi-storey buildings." *Proceedings of the Institution of Civil Engineers - Structures and Buildings*, 160(4), 197-205, DOI: 10.1680/stbu.2007.160.4.197.

Jaspart, J.-P., and Maquoi, R. (1995). "Effect of bolt preloading on joint behaviour." *Proc., $1^{st}$ European Conference on Steel Structures*, Athens, Greece, 219-226.

Kennedy, J., and Eberhart, R. (1995). "Particle swarm optimization." *Proc., International conference on neural networks (ICNN'95)*, Perth, WA, Australia, 1942-1948.

Kline, D. P. (1989). "Performance of snug tight bolts in moment end-plate connections." M.Sc. Thesis, Virginia Polytechnic Institute and State University, Virginia, USA.

Kong, Z., Hong, S., Vu, Q.-V., Cao, X., Kim, S.-E., and Yu, B. (2020). "New equations for predicting initial stiffness and ultimate moment of flush end-plate connections." *Journal of Constructional Steel Research*, 175, 106336, DOI: 10.1016/j.jcsr.2020.106336.

Kottari, A. K., Charalampakis, A. E., and Koumousis, V. K. (2014). "A consistent degrading Bouc–Wen model." *Engineering Structures*, 60, 235-240, DOI: 10.1016/j.engstruct.2013.12.025.

Kozlowski, A., Kowalczyk, R., and Gizejowski, M. (2008). "Estimation of the Initial Stiffness and Moment Resistance of Steel and Composite Joints." *Proc., $8^{th}$ World Congress, Council on Tall Buildings and Urban Habitat (CTBUH)*, Dubai, United Arab Emirates.

Kueh, A. B. H. (2021). "Artificial neural network and regressed beam-column connection explicit mathematical moment-rotation expressions." *Journal of Building Engineering*, 43, 103195, DOI: 10.1016/j.jobe.2021.103195.

Kukla, D., and Kozlowski, A. (2021). "Parametric study of steel flush and extended end-plate joints under column loss scenario." *Engineering Structures*, 237, 112204, DOI: 10.1016/j.engstruct.2021.112204.

Kukreti, A. R., Murray, T. M., and Abolmaali, A. (1987). "End-plate connection moment-rotation relationship." *Journal of Constructional Steel Research*, 8(1), 137-157, DOI: 10.1016/0143-974X(87)90057-5.

Levenberg, K. (1944). "A method for the solution of certain non-linear problems in least squares." *Quarterly of applied mathematics*, 2(2), 164-168, DOI:

Lignos, D. G., Hartloper, A., Elkady, A., Deierlein, G. G., and Hamburger, R. (2019). "Proposed Updates to the ASCE 41 Nonlinear Modeling Parameters for Wide-Flange Steel Columns in Support of Performance-based Seismic Engineering." *Journal of Structural Engineering*, 145(9), 04019083, DOI: 10.1061/(ASCE)ST.1943-541X.0002353.

Lundberg, S. M., and Lee, S.-I. (2017). "A Unified Approach to Interpreting Model Predictions." *Proc., 31st Conference on advances in Neural Inforamtion Processing Systemd (NIPS 2017)*, Long Beach, CA, USA, 1–10.

Mak, L., and Elkady, A. (2021). "Experimental database for steel flush end-plate connections." *Journal of Structural Engineering*, 147(7), 04721006, DOI: 10.1061/(ASCE)ST.1943-541X.0003064.

Mann, A. P., and Morris, L. J. (1981). "Significance of Lack of Fit-Flush Beam-Column Connections." *Proc., Joints in structural steelwork*, 6.22-26.36.

Marquardt, D. W. (1963). "An algorithm for least-squares estimation of nonlinear parameters." *Journal of the society for Industrial and Applied Mathematics*, 11(2), 431-441, DOI:

Murray, T. M., and Shoemaker, W. L. (2002). "Flush and extended multiple-row moment end-plate connections." *Design Guide 16*, American Institute of Steel Construction, Chicago, Illinois, USA.

Ostrander, J. R. (1970). "An Experimental Investigation of End Plate Connections." M.Sc. Thesis, University of Saskatchewan, Saskatchewan, Canada.

Ostrowski, K., and Kozłowski, A. (2017). "Rotation Capacity of Bolted Flush End-Plate Stiffened Beam-to-Column Connection." *Civil and Environmental Engineering Reports*, 25(2), 173-184, DOI: 10.1515/ceer-2017-0028.

Prescott, A. T. (1987). "The Performance of End-Plate Connections in Steel Structures and their Influence on Overall Structural Behaviour." Ph.D. Thesis, Hatfield Polytechnic, Hatfield, UK.

Qiang, X., Bijlaard, F. S. K., Kolstein, H., and Jiang, X. (2014). "Behaviour of beam-to-column high strength steel endplate connections under fire conditions – Part 1: Experimental study." *Engineering Structures*, 64, 23-38, DOI: 10.1016/j.engstruct.2014.01.028.

Ramberg, W., and Osgood, W. R. (1943). "Description of Stress-Strain Curves by Three Parameters." *Report No. NACA-TN-902*, National Advisory Committee for Aeronautics, Washington, D.C, USA.

Richard, R. M., and Abbott, B. J. (1975). "Versatile Elastic-Plastic Stress-Strain Formula." *Journal of the Engineering Mechanics Division*, 101(4), 511-515, DOI: 10.1061/JMCEA3.0002047.

Rölle, L. (2013). "The load-bearing and deformation behavior of bolted steel and composite joints in fully plastic design and in extraordinary design situations (in german)." Ph.D. Thesis, University of Stuttgart, Stuttgart, Germany.

Skiadopoulos, A., Elkady, A., and Lignos, D. G. (2021). "Proposed Panel Zone Model for Seismic Design of Steel Moment-Resisting Frames." *Journal of Structural Engineering*, 147(4), 04021006, DOI: 10.1061/(ASCE)ST.1943-541X.0002935.

Song, Y., Uy, B., Li, D., and Wang, J. (2022). "Ultimate behaviour and rotation capacity of stainless steel end-plate connections." *Steel and Composite Structures*, 42(4), 569-590, DOI: 10.12989/scs.2022.42.4.569.

Statistics and Machine Learning Toolbox (2022). *version version: 12.4 (R2022b)*, The MathWorks Inc. https://www.mathworks.com, Natick, Massachusetts.

Thomson, A. W., and Broderick, B. M. (2002). "Earthquake resistance of flush end-plate steel joints for moment frames." *Proceedings of the Institution of Civil Engineers, Structures and Buildings*, 152(2), 157-165, DOI: 10.1680/stbu.2002.152.2.157.

Zoetemeijer, P. (1974). "A Design Method for the Tension Side of Statically Loaded, Bolted Beam-to-Column Connections." *HERON*, 20(1), DOI: