Structures 56 (2023) 104904

Contents lists available at ScienceDirect

Structures

journal homepage: www.elsevier.com/locate/structures

Artificial neural network prediction of the initial stiffness of semi-rigid beam-to-column connections

J.M. Reinosa *, A. Loureiro, R. Gutierrez, M. Lopez

Laboratorio de Análisis Estructural, Campus Industrial de Ferrol, Centro de Investigacións Tecnolóxicas (CITENI), Universidade da Coruña, EPEF Ferrol, Mendizábal s/n Campus de Esteiro, 15471 Ferrol, Spain

ARTICLE INFO

Keywords:
Angle connections
Semi-rigid joints
Initial stiffness
Artificial neural networks

ABSTRACT

Joints are significant components in the design and construction of steel structures. The characteristic parameters of the connections must be reproduced in a reliable way to represent the actual behaviour of a structure. Accordingly, the study of semi-rigid joints is essential to better understand this issue. Among the different types of semi-rigid joints, angle connections stand out as a suitable solution in many cases. This paper presents a methodology using artificial neural networks for predicting the initial rotational stiffness of major axis symmetrical angle connections according to the Eurocode description. A consistent stiffness database was developed from the existing data in the Steel Connection Data Bank. Then, the database was cleansed to provide with a robust training set. Different network architectures were analysed until a topology that showed a good performance and generalisation features was obtained. The network was successfully checked with some saved tests from the database and with off-database tests; the network could be reliably used within the range of the training input parameters.

1. Introduction

The use of semi-rigid bolted angle joints in steel construction has several advantages. They reduce the execution times of the works and their cost [1], develop a better response to seismic events than other connection typologies due to their good ductility [2], and are suitable when a design is conceived for deconstruction [3]. In the European steel construction industry, angle connections are not as commonly used as other solutions such as end-plate connections, and the European standard (EC3) has not yet fully developed their capabilities. However, in recent years, different scientific contributions have provided mechanical or numerical models studying numerous essential aspects of the behaviour of angle connections, which could be included in EC3's component method. Several models have been developed to predict the rotational stiffness and resistance of these connections [4–6], the influence of bolt preloading [7], or how the presence of stiffeners affects the behaviour of the joint [8–10]. The difficulty in adequately representing the performance of the angle connections was evidenced in all these works, mainly due to the large number of components involved. Thus, the results of the different studies yielded complex mechanical models or numerical solutions that require a careful elaboration and subsequent validation and, therefore, are time-consuming and demanding [11,12]. One of the fundamental parameters according to EC3 is the initial rotational stiffness of connections. The Eurocode defines the initial rotational stiffness, $S_{j,ini}$, as *the slope of the elastic branch of the design moment-rotation characteristic* [13]. Different criteria are used in experiments provided in the literature to measure this initial stiffness. Furthermore, the initial stiffness is occasionally not explicitly provided by some authors, who only include the moment-rotation curves in their works. Therefore, one of the main tasks of this study was to obtain the initial stiffness from tests on semi-rigid joints with angles extracted from the Kishi and Chen Steel Connection Data Bank (SCDB) [14]. A stiffness database was obtained in a uniform and consistent way in order to validate the methodology developed in this work and that of any future works, whether analytical or numerical. In this study, the initial stiffness was predicted using artificial neural networks (ANN). A potential advantage of this specific ANN study is that it was based on experimental data which included real imperfections [15], so that the results presented here are probably more realistic than those obtained by other methods such as finite element models (FEM) [11,12], analytical approaches [14] or mechanical models [5]. A second advantage of this proposal is that the response was much faster than that

* Corresponding author.
E-mail address: jreinosa@udc.es (J.M. Reinosa).

https://doi.org/10.1016/j.istruc.2023.104904
Received 7 December 2022; Received in revised form 25 May 2023; Accepted 14 July 2023
Available online 19 July 2023

Table 1
Authors and number of tests: Top-and-Seat-Double-Web (TSDW) and Top-and-Seat (TS).

Author	N° of tests	Typology
Altman et al. (1982) [28]	20	TSDW
Rathbun (1936) [29]	2	TSDW
Roeder et al. (1996) [30]	1	TSDW
Elnashai et al. (1998) [31]	1	TSDW
Fu et al. (1998) [32]	4	TSDW
Calado et al. (2000) [33]	3	TSDW
Komuro et al. (2002) [34]	2	TSDW
Rathbun (1936) [29]	3	TS
Hetchman and Johnston (1947) [35]	12	TS
Marley et al. (1982) [36]	26	TS
Maxwell et al. (1981) [37]	12	TS
Davison et al. (1987) [38]	1	TS
Altman et al. (1982) [28]	2	TS
Harper (1990) [39]	1	TS
Mander et al. (1994) [40]	4	TS
Bernuzzi et al. (1996) [41]	1	TS
Kubo et al.(1999) [42]	5	TS
Komuro et al. (2002) [34]	1	TS
Sato et al. (2007) [43]	6	TS

derived from FEMs, which can be useful for research purposes or making decisions in experimental works without the need for expensive and complex numerical models. From the perspective of steel building design, ANNs are adequate tools for predicting the characteristic parameters of connections as long as the corresponding safety factors are adopted within the framework of some specified validity ranges [15]. In the literature, several authors have used ANNs to study semi-rigid joints. Abdalla and Stavroulakis [16] used ANNs to predict the moment-rotation curve of simple beam-to-column angle joints. Furthermore, Anderson et al. [17] used ANNs to study the behaviour of end-plate beam-to-column joints in the minor axis. Lima et al. [18] used ANNs to predict the stiffness and resistance of different types of beam-to-column joints. More recently, the moment-rotation behaviour of top and seat angle connections with double web angles was also investigated using ANNs [19]. In addition, Kim et al. [20] proposed a mechanical and an ANN informational model of steel beam-to-column connections, considering that some features of response are not represented by mechanical models.

The objective of this study was to generate a reliable ANN capable of predicting the initial rotational stiffness of semi-rigid beam-to-column connections. The first step was to produce a debugged database and to select the critical parameters for the determination of the initial stiffness. The second step consisted in finding the best network architecture. Once the network architecture was defined, a training process was recursively executed until the mean squared error (MSE) was minimised. The final step involved validating the network at the end of the optimisation process by verifying its generalisation power. Some of the tests from the debugged database were preserved to validate the network, which was additionally checked with some experimental tests not included in the database.

2. Experimental data

A databank of angle connection stiffness can help develop design methods and research without the need for costly experimental campaigns. In addition, some national regulations such as the American Institute of Steel Construction (AISC) standard [21] indicate that moment-rotation characteristics should be obtained from databanks to produce accurate results during the analysis of semi-rigid frames.

2.1. Steel connection data bank

Although other data references are available in the literature [22,23], this study will only use the SCDB database compiled by Kishi and Chen [14] to develop the ANN, since all the tests from this databank are deeply documented, including the points from the experimental moment-rotation curve. This databank originated from the studies conducted by Kishi and Chen [24] and by Chen and Kishi [25] which consisted of a comprehensive search on beam-to-column riveted, bolted, and welded connections. These studies provided the corresponding parameters and moment-rotation characteristics of semi-rigid connections related to steel construction published from 1936 to 1985, and all this information was compiled into a database. Then, Kishi revised the database, adding 93 experimental data for the period from 1986 to 1998 and Komuro and Kishi improved the connection database including the SCDB program. The connection data included in the databank currently consists of 486 experimental tests. The data were compared with some prediction equations: the analytical polynomial equation proposed by Frye and Morris [26], a three-parameter power model, and a modified

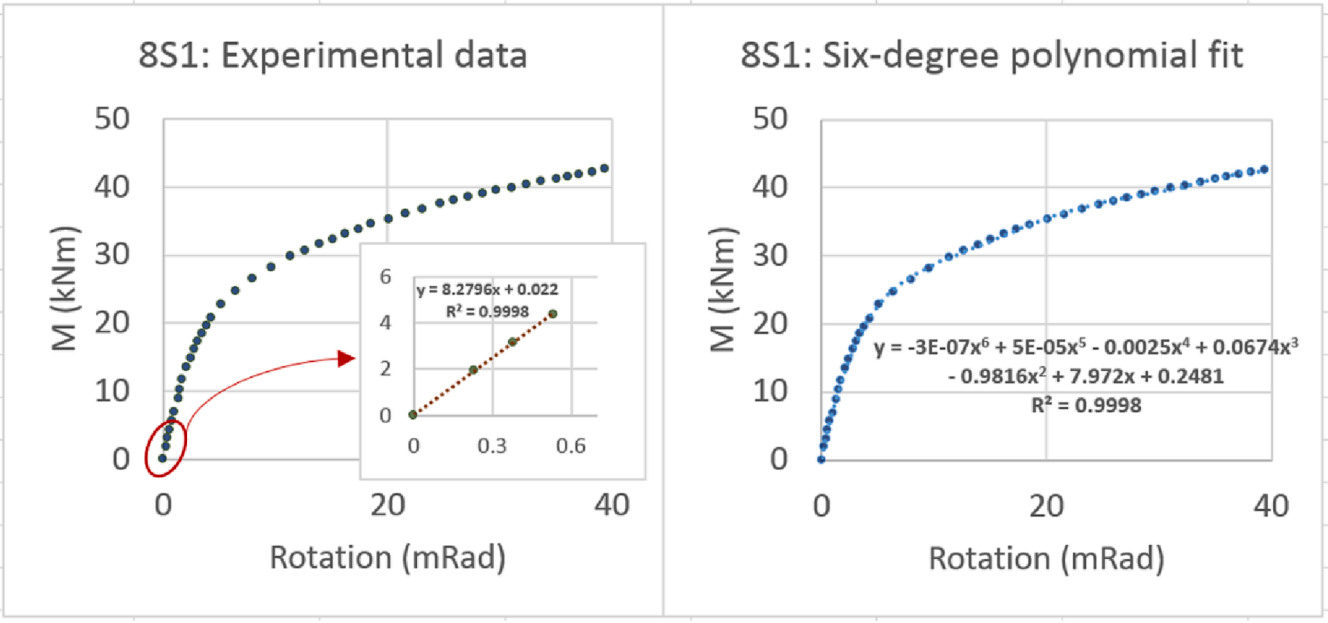

Fig. 1. Moment-Rotation curve of the 8S1 test and regression model for the first points (left) and polynomial fit (right).

Table 2
Obtained initial stiffness, coefficient of determination, and number of points used for the linear approximation.

N°	TEST	AUTHOR	Sjini (kN/mm)	R^2	N	N°	TEST	AUTHOR	Sjini (kN/mm)	R^2	N
1	8S1	AZIZINAMINI	8280	0.999	4	52	D2-1/4-1	MARLEY	773	1.000	4
2	8S2	AZIZINAMINI	15,593	0.990	4	53	D2-1/4-2	MARLEY	1755	1.000	3
3	8S3	AZIZINAMINI	12,813	0.999	4	54	D3-1/4-1	MARLEY	813	1.000	4
4	8S4	AZIZINAMINI	1875	1.000	4	55	D3-1/4-2	MARLEY	279	1.000	3
5	8S5	AZIZINAMINI	10,258	1.000	4	56	A-1/2-1	MARLEY	1948	1.000	4
6	8S6	AZIZINAMINI	5347	0.992	4	57	B-1/2-1	MARLEY	6930	1.000	4
7	8S7	AZIZINAMINI	5588	0.991	4	58	B-1/2-2	MARLEY	4679	0.995	4
8	8S8	AZIZINAMINI	6858	0.999	4	59	C1-1/2-1	MARLEY	3932	1.000	4
9	8S9	AZIZINAMINI	12,032	0.996	4	60	C1-1/2-2	MARLEY	3069	1.000	3
10	8S10	AZIZINAMINI	24,919	0.988	4	61	C2-1/2-1	MARLEY	6380	0.993	4
11	14S1	AZIZINAMINI	32,326	1.000	4	62	C2-1/2-2	MARLEY	2204	0.991	4
12	14S2	AZIZINAMINI	35,523	0.950	4	63	D1-1/2-1	MARLEY	3912	0.995	4
13	14S3	AZIZINAMINI	14,343	0.999	4	64	D1-1/2-2	MARLEY	6545	1.000	4
14	14S4	AZIZINAMINI	24,568	0.999	4	65	D2-1/2-1	MARLEY	10,976	1.000	3
15	14S5	AZIZINAMINI	22,310	0.999	4	66	D2-1/2-2	MARLEY	3431	1.000	4
16	14S6	AZIZINAMINI	27,427	0.999	4	67	D3-1/2-1	MARLEY	4784	0.985	4
17	14S8	AZIZINAMINI	55,014	1.000	4	68	D3-1/2-2	MARLEY	3087	1.000	4
18	14S9	AZIZINAMINI	28,172	0.999	4	69	A1	MAXWELL	32,504	0.999	4
19	B11	RATHBUN	24,613	0.992	5	70	A2	MAXWELL	24,446	1.000	4
20	B12	RATHBUN	28,483	0.992	4	71	A3	MAXWELL	18,789	1.000	4
21	L1	ROEDER	13,541	0.963	3	72	A4	MAXWELL	24,217	1.000	4
22	SBR01	ELNASHAI	14,237	1.000	4	73	B1	MAXWELL	20,616	1.000	4
23	LM-P	FU	9092	0.981	4	74	B2	MAXWELL	16,393	0.998	4
24	LM-T	FU	8891	0.992	4	75	B3	MAXWELL	17,375	0.998	4
25	LM-P15	FU	7658	0.961	4	76	B4	MAXWELL	18,651	1.000	4
26	LM-T15	FU	8946	0.998	4	77	C1	MAXWELL	15,851	1.000	4
27	BCC7-M	CALADO	23,557	0.995	4	78	C2	MAXWELL	17,063	0.999	4
28	BCC9-M	CALADO	18,819	0.996	4	79	C3	MAXWELL	22,390	1.000	4
29	BCC10-M	CALADO	19,975	0.999	4	80	C4	MAXWELL	24,389	1.000	4
30	W18-M	KOMURO	55,284	0.998	4	81	JT/08	DAVISON	15,089	0.998	4
31	W29-M	KOMURO	58,351	1.000	3	82	AZITS1	AZIZINAMINI	12,518	0.979	4
32	B8	RATHBUN	23,744	0.993	5	83	AZITS2	AZIZINAMINI	26,049	0.975	4
33	B9	RATHBUN	15,938	0.994	5	84	TEST3	HARPER	5831	0.990	3
34	B10	RATHBUN	23,968	0.990	5	85	R1-01	MANDER	7406	0.982	4
35	N02	HECHTMAN	10,166	0.999	6	86	R1-05	MANDER	15,010	0.999	4
36	N05	HECHTMAN	21,523	1.000	4	87	R1-06	MANDER	14,594	0.998	4
37	N09	HECHTMAN	39,321	0.993	4	88	R0-11	MANDER	11,975	1.000	4
38	N010	HECHTMAN	47,457	0.960	4	89	TSC/M	BERNUZZI	10,492	0.961	3
39	N016	HECHTMAN	17,837	1.000	4	90	N01	KUBO	3500	0.997	3
40	N020	HECHTMAN	28,246	1.000	4	91	N03	KUBO	4195	0.980	4
41	N022	HECHTMAN	26,592	1.000	4	92	N05	KUBO	4490	1.000	4
42	N024	HECHTMAN	38,767	0.991	4	93	N07	KUBO	3515	0.992	4
43	A-1/4-1	MARLEY	1333	1.000	3	94	N09	KUBO	5125	0.997	4
44	A-1/4-2	MARLEY	1078	1.000	4	95	W00-M	KOMURO	36,546	0.982	3
45	B-1/4-1	MARLEY	1229	1.000	3	96	G60	SATO	41,532	0.975	4
46	B-1/4-2	MARLEY	2589	0.999	5	97	G105	SATO	24,256	0.996	4
47	C1-1/4-2	MARLEY	2549	1.000	4	98	G150	SATO	11,465	0.969	4
48	C2-1/4-1	MARLEY	3398	0.999	4	99	GW60	SATO	43,407	1.000	3
49	C2-1/4-2	MARLEY	1438	1.000	4	100	A40	SATO	47,406	0.999	3
50	D1-1/4-1	MARLEY	594	0.995	5	101	A90	SATO	45,026	0.998	3
51	D1-1/4-2	MARLEY	557	1.000	4						

exponential model [27]. Each experimental datum included the moment-rotation characteristics and the parameters used in the prediction equations. The databank included 107 experimental data on semi-rigid angle connections: 74 experimental data for Top and Seat angle connections (TS) and 33 experimental data for Top and Seat angle connections with a Double Web angle (TSDW). Double Web angle connections were not included in this study because this typology is considered to be a simple connection. Table 1 summarises the compiled data used in this study. Tests *14SW1* and *14SW2* from Altman et al. [28] were disregarded because the angles were welded to the beam. In addition, tests *N011*, *N017*, *N018*, and *N023* from Hetchman [35] were discarded because they were related to minor axis joints. Thus, a total of 101 tests were considered in this research.

2.2. Determination of the rotational stiffness

According to experiments, beam-to-column angle connections usually display a moment-rotation behaviour that becomes non-linear during the first stages of the loading process. Therefore, a careless selection of the first points of the moment-rotation curve may occasionally give rise to the wrong values of initial stiffness. In this study, a univocal methodology based on a regression analysis was employed to compute the initial rotational stiffness. This linear regression was developed by considering the first points of the experimental moment-rotation curves, according to the data compiled in the SCDB database [14]. On this basis, the value of the initial stiffness could be obtained following the description of the Eurocode EC3 [13]. The general criteria to develop the linear regression were the following: 1) creating the regression model with at least four points whenever possible; 2) achieving a value of the coefficient of determination R^2 greater than 0.95; 3) and disregarding the points related to very low rotations. Regarding the range, the rotation is below 1.5 mRad in most tests, except in the series of tests conducted by Kubo et al. [42] where the range reached 4 mRad. Fig. 1 shows an example of the linear regression made for the 8S1 test. For this particular test, Altman et al. provided a value of 7540 kNm/rad obtained as the initial slope of the moment-rotation curve, measured as the derivative of a second degree polynomial fit through the first data points [28]. However, if a six-degree polynomial

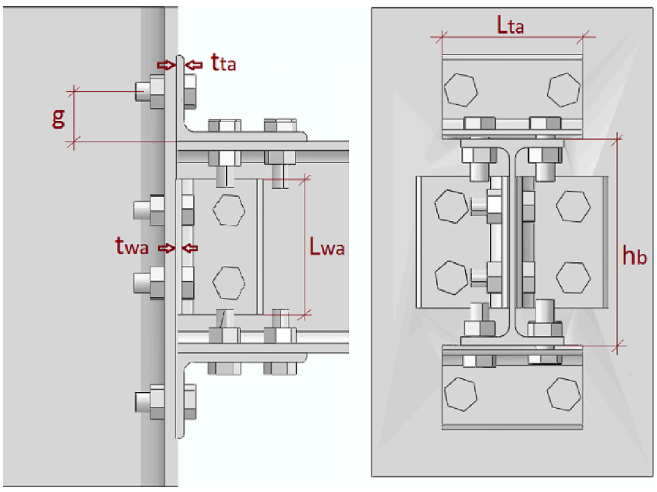

Fig. 2. Geometrical parameters considered in the study.

fit of the 8S1 test is considered, the tangent at the origin provides a stiffness of 7972 kNm/rad, a value that is closer to that obtained with the methodology proposed in this paper following EC3 prescriptions.

Table 2 lists the tests and the initial stiffness values obtained, the coefficients of determination R^2, and the number of points N used for the linear approximation (at least four points in 87 out of 101 specimens). The percentage of the resistance range covered in the regression analysis was between 10% and 50% of the plastic moment for TS connections, and between 20% and 80% of the plastic moment for TSDW connections. In almost all the tests, the linear approximation was satisfactory, with R^2 values close to 1, except for some which, despite presenting slightly lower values, still accounted for more than 95% of the variability in the response in all cases.

3. Parameter selection

If the number of points in the ANN training dataset is much larger than the number of parameters in the network, then there is no chance of overfitting. Since the supply of data is limited, and collecting more data under the same conditions as those provided by SCDB database is difficult, thus increasing the size of the training set, strategies should be considered to prevent overfitting [44]. The addition of parameters lead the model to memorise the data instead of finding generalisations. Therefore, an ANN trained on a dataset would probably reproduce the memorised data instead of learning. Therefore, numerous solutions can be obtained which fit the training data well but will not generalise at all. The optimal number of parameters in the ANN also depend on the data distribution, the quality of the obtained solution, the amount of noise, and the nature of the function being approximated [45]. Considering these premises, the rotational stiffness should be characterised through a reduced number of significant parameters. Azizinamini et al. [46] identified the most significant parameters affecting the initial stiffness of angle connections as the depth of the beam section, the thickness of the flange angles, and the gage in the leg of the flange angle attached to the column. They also detected three secondary parameters that also had an effect on the response of the connection: the length of the flange angles, that of the web angles, and the thickness of the latter. Since these parameters are also involved in the power model of Kishi and Chen [27], which is a simple and widely studied methodology, they were finally selected for this study. Therefore, the following predictors were considered in the ANN approach: the beam depth, h_b, the top angle thickness, t_{ta}, the top angle length, L_{ta}, the gage distance, g, the web angle thickness, t_{wa}, and the web angle length, L_{wa}. Fig. 2 shows these geometrical parameters in a typical configuration for a TSDW angle connection.

4. Statistical analysis for data cleansing

Data pre-processing was an important step in this study since the analysis of data that are not carefully examined can lead the ANN to produce misleading results. The development and assessment of experimental tests on semi-rigid joints, where a large number of components are involved, entails difficulties and possible execution and measurement inaccuracies; a prior data cleaning process can be used to detect these. A multilinear regression (ML) model may be used for the identification of anomalies in the data if the hypotheses of linearity and normality are confirmed. Therefore, a statistical analysis was carried out using the IBM *SPSS*® software package [47] concerning the data from the SCDB databank together with the previously obtained values of rotational stiffness. For this problem, the rotational stiffness was the dependent variable of the ML approximation, and it was related to several independent variables, the predictors described in the previous section, that are affected by regression coefficients. The method of least squares was used to estimate the regression coefficients in the ML model. Once these coefficients were obtained, the prediction for the stiffness was the output from SPSS®.

Table 3 provides a summary of the ML model results, including the multiple correlation coefficient R, the coefficient of determination R^2, the Fisher ratio F, the Durbin-Watson coefficient, and the regression coefficients. The value of the adjusted coefficient of determination R^2 indicates that the model accounts for just 71% of the variability in the initial stiffness response. However, the model can still be useful in detecting the presence of severe anomalies in the data. Table 4 shows that the test for significance of regression confirms the null hypothesis, with a Fisher value $F = 41.583 > F_{0.01,6,94} = 3.03$, so that the stiffness is linearly related to the predictors.

Fig. 3 shows a histogram of the regression standardised residuals vs frequency, and the normal probability P-P graphic of the regression standardised residuals. The histogram of the standardised residuals indicates a nearly symmetric, moderate tailed distribution. Therefore, the next step involved analysing the normal probability plot. Generally speaking, the "S" pattern, which could usually indicate non-linearity, is fairly mild. In addition, in the middle of the plot, the data exhibit a linear pattern and the first and last points do not show an increasing departure from the fitted line. Therefore, it can reasonably be concluded that the normal distribution provides an adequate fit for the residuals.

Table 5 shows the prediction from the ML model and the ratio r between the predicted rotational stiffness, S_{jini_MLR}, and the experimental values S_{jini_exp}.

In order to select a debugged set of input data to train the network, the tests that presented a very high or very low r ratio were removed. In particular, these selected ratio values were r greater than 1.6 or $r < 0.4$. The reason for choosing this relatively wide range was the need to obtain a database that was as clean as possible while maintaining a sufficient number of data for training the network, at least ten times the number of entry parameters. Thus, 33 experimental tests were removed from the original data, and therefore, 68 tests remained for developing the ANN approach. Table 6 shows the tests that remained in the same sorting order that they are presented in in the databank.

5. Artificial neural networks

5.1. Introduction

Artificial neural networks are computing systems that are inspired by biological ANNs and composed of simple elements operating in parallel. The connections between these elements determine the network function. An ANN can be trained to carry out a particular function by rearranging the values of the links between elements. These values are designated as weights. In an ANN, each neuron in a layer is connected via a weight to each neuron in the next activity. Each of these activities stores the weighted activities of previous layers. The bias is a parameter

Table 3
Model Summary.

R	R^2	Adjusted R^2	Std. Error	R^2 Change	F Change	Sig. F Change	Durbin-Watson
0,852	0,726	0,709	7532,7781	0,726	41,583	0,000	1,216
Predictor	*Constant*	h_b	t_{ta}	L_{ta}	g	t_{wa}	L_{wa}
Coefficient	−17036,48	46,93	1062,22	72,53	−89,32	−2944,30	143,25
Std. Error	4276,76	9,33	318,67	20,06	53,44	984,59	37,29

Table 4
Results of the analysis of variance.

	Sum of Squares	df	Mean Square	F	Sig.
Regression	14157231040,000	6	2359538506,000	41,583	0,000[b]
Residual	5333818058,000	94	56742745,300		
Total	19491049100,000	100			

of the ANN which is used to adjust the weighted sum of the inputs to the neuron together with the output. The output of the network is computed by multiplying the input by the weight and running the result through an activation function, for example, the Sigmoid. When the weight changes, the steepness of the sigmoid also changes, but changing the steepness is not enough for achieving the result. Instead, the entire curve must be shifted, and therefore a bias value is needed. Indeed, the use of biases in an ANN increases the capacity of the network to solve problems.

Artificial neural networks are adjusted so that a particular input leads to a specific target output, automatically generating identifying characteristics from the processed examples. Notably, ANNs have been trained to perform complex functions in various fields, including identification and pattern recognition, and to solve problems that are difficult for human beings or conventional computers.

Standard backpropagation is the common gradient descent algorithm used to train ANNs. After each forward pass-through, the algorithm performs a backward pass while adjusting the parameters of the model: i.e., the weights and biases. The weights move along the negative gradient of the performance function and learning is obtained through the backpropagation of errors.

The networks are properly trained if they are able to generalise. Generalisation means that good results can be obtained for data out of the representative input dataset where the network was trained. Network generalisation can be improved with several techniques. The use of a small-enough network is advisable since it will not have enough power to overfit the data. Nevertheless, it is relatively difficult to predict how large a network should be to avoid overfitting [44]. A commonly used method to improve generalisation is early stopping where the available data are divided into three sets: a training set, validation set, and test set. If the network overfits the data, the error of the validation set begins to increase and the training is stopped. Another method to improve generalisation is regularisation, which consists in modifying the performance function. This modification is done by adding the mean of the sum of squares of weights and biases to the basic performance function, which is the sum of squares of the network errors on the training set. Obtaining the optimal regularisation parameters can be challenging, and thus, a Bayesian approach can be used to ease and automate the process [45]. In the Bayesian framework, the weights and biases are random variables, and the regularisation parameters can be determined through statistical techniques. For small datasets and function approximation networks, the Bayesian approach provides a better performance than early stopping techniques since it does not require a validation dataset and uses all the data. However, the network must be trained until it reaches convergence. Finally, another strategy to improve generalisation, when facing difficulties associated with small and noisy datasets, involves training multiple ANNs and averaging their outputs using the Bayesian regularisation approach.

The data used in this research can be considered small and likely noisy, even after the cleansing process. Therefore, the Bayesian approach could be considered for improving generalisation, since overfitting becomes much more difficult to avoid in this small-data context. In any case, the training algorithm and the number of neurons in the hidden layer will be discussed in the following sections to ensure the best performance within the smallest network.

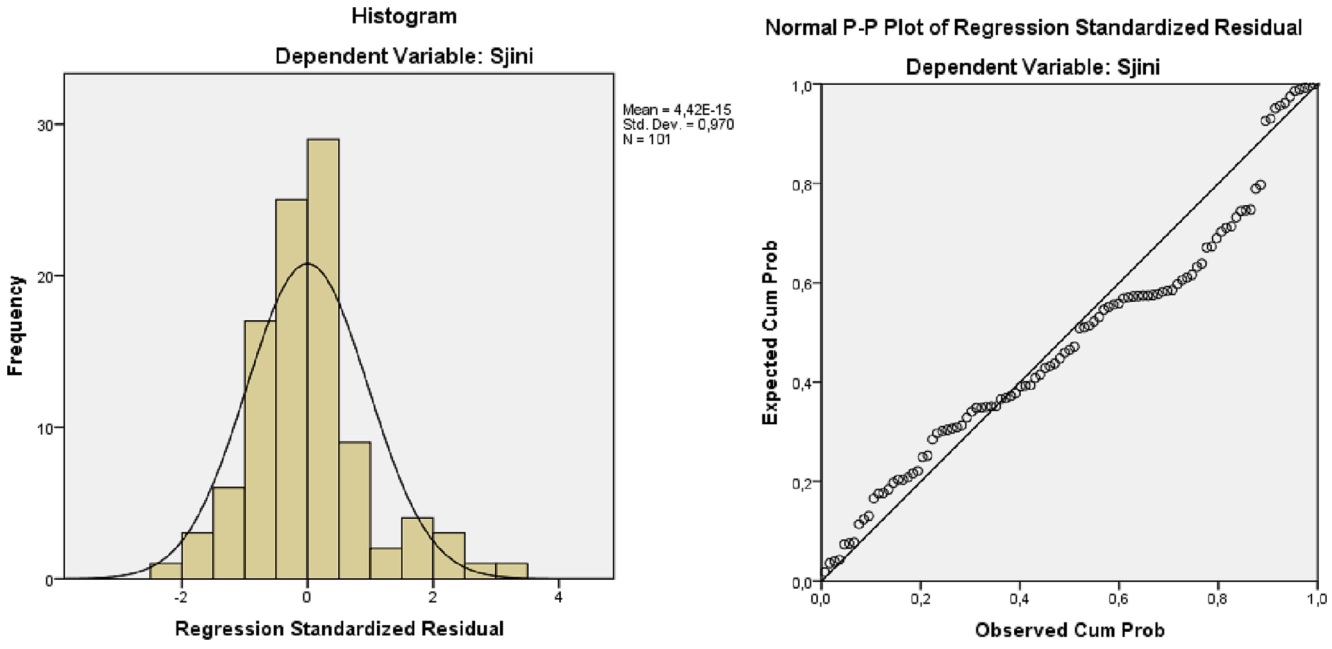

Fig. 3. Histogram of regression standardised residuals vs frequency and P-P plot.

Table 5
Prediction from the multilinear regression (ML) model and ratio r.

Test	S_{jini_exp}	S_{jini_MLR}	r	Test	S_{jini_exp}	S_{jini_MLR}	r
8S1	8280	8938	1.08	D2-1/4-1	773	−782	−1.01
8S2	15,593	10,638	0.68	D2-1/4-2	1755	−782	−0.45
8S3	12,813	12,623	0.99	D3-1/4-1	813	−782	−0.96
8S4	1875	4966	2.65	D3-1/4-2	279	−782	−2.80
8S5	10,258	13,188	1.29	A-1/2-1	1948	5963	3.06
8S6	5347	7804	1.46	B-1/2-1	6930	5963	0.86
8S7	5588	9503	1.70	B-1/2-2	4679	5963	1.27
8S8	6858	8938	1.30	C1-1/2-1	3932	5963	1.52
8S9	12,032	10,638	0.88	C1-1/2-2	3069	5963	1.94
8S10	24,919	14,037	0.56	C2-1/2-1	6380	5963	0.93
14S1	32,326	31,021	0.96	C2-1/2-2	2204	5963	2.71
14S2	35,523	34,420	0.97	D1-1/2-1	3912	5963	1.52
14S3	14,343	20,134	1.40	D1-1/2-2	6545	5963	0.91
14S4	24,568	21,894	0.89	D2-1/2-1	10,976	5963	0.54
14S5	22,310	31,021	1.39	D2-1/2-2	3431	5963	1.74
14S6	27,427	34,420	1.25	D3-1/2-1	4784	5963	1.25
14S8	55,014	37,819	0.69	D3-1/2-2	3087	5963	1.93
14S9	28,172	34,420	1.22	A1	32,504	21,367	0.66
B11	24,613	20,598	0.84	A2	24,446	24,993	1.02
B12	28,483	29,839	1.05	A3	18,789	21,367	1.14
L1	13,541	34,974	2.58	A4	24,217	24,993	1.03
SBR01	14,237	15,226	1.07	B1	20,616	23,491	1.14
LM-P	9092	12,758	1.40	B2	16,393	27,118	1.65
LM-T	8891	12,758	1.43	B3	17,375	23,491	1.35
LM-P15	7658	12,758	1.67	B4	18,651	27,118	1.45
LM-T15	8946	12,758	1.43	C1	15,851	26,678	1.68
BCC7-M	23,557	18,569	0.79	C2	17,063	30,304	1.78
BCC9-M	18,819	18,569	0.99	C3	22,390	26,678	1.19
BCC10-M	19,975	18,569	0.93	C4	24,389	30,304	1.24
W18-M	55,284	29,251	0.53	JT/08	15,089	9031	0.60
W29-M	58,351	45,009	0.77	AZITS1	12,518	18,922	1.51
B8	23,744	10,848	0.46	AZITS2	26,049	22,321	0.86
B9	15,938	14,533	0.91	TEST3	5831	21,625	3.71
B10	23,968	25,587	1.07	R1-01	7406	10,347	1.40
N02	10,166	21,071	2.07	R1-05	15,010	10,347	0.69
N05	21,523	34,496	1.60	R1-06	14,594	10,347	0.71
N09	39,321	37,895	0.96	R0-11	11,975	10,347	0.86
N010	47,457	41,188	0.87	TSC/M	10,492	17,273	1.65
N016	17,837	17,672	0.99	N01	3500	10,530	3.01
N020	28,246	33,278	1.18	N03	4195	10,440	2.49
N022	26,592	35,671	1.34	N05	4490	6842	1.52
N024	38,767	37,895	0.98	N07	3515	6851	1.95
A-1/4-1	1333	−782	−0.59	N09	5125	6860	1.34
A-1/4-2	1078	−782	−0.73	W00-M	36,546	24,076	0.66
B-1/4-1	1229	−782	−0.64	G60	41,532	26,816	0.65
B-1/4-2	2589	−782	−0.30	G105	24,256	22,797	0.94
C1-1/4-2	2549	−782	−0.31	G150	11,465	18,777	1.64
C2-1/4-1	3398	−782	−0.23	GW60	43,407	26,816	0.62
C2-1/4-2	1438	−782	−0.54	A40	47,406	26,816	0.57
D1-1/4-1	594	−782	−1.32	A90	45,026	26,816	0.60
D1-1/4-2	557	−782	−1.40				

5.2. Matlab neural network toolbox

The Matlab Neural Network (NN) Toolbox™ emphasises the use of ANNs in engineering and other practical applications. The graphical Matlab tools for training ANNs can be used to solve problems in function fitting since ANNs are good at fitting functions [44]. The NN Toolbox software can be used in several ways: through graphical user interfaces, using command-line operations, through customisation, and, finally, by modifying any of its functions. The customisation capability allows the creation of custom ANNs, while maintaining access to the full functionality of the toolbox. Networks with arbitrary connections can be created and trained using the existing toolbox training functions. The standard steps for designing ANNs to solve function fitting problems are shown in Fig. 4.

In this paper, the data was divided into two datasets: a training dataset and a testing dataset. The size of the training dataset was selected to guarantee a minimum number of samples regarding the number of input parameters, as will be described in the next section. In addition, the Matlab neural fitting tool was used to automatically generate a script that could be customised both to train and to validate the network. Otherwise, the validation of the network was developed not only with the preserved testing dataset, but also with some other test from the literature to further check the generalisability.

5.3. Datasets for the study

At the end of the network training process, the network performance must be checked with values that are not used to train it. Therefore, one out of eight of the remaining tests from Table 6 was extracted, leading the same order as sorted by the SCDB databank. Thus, eight tests were preserved to analyse the network at the end of the process, and, therefore, it was trained with the remaining 60 data. Tables 7 and 8 show the training and testing datasets, respectively. This training dataset ensured that there was an amount of data at least ten times greater than the number of parameters [48].

Table 6
Tests remaining after discarding.

Order	Test	Order	Test	Order	Test	Order	Test
1	8S1	18	B12	35	N022	52	C3
2	8S2	19	SBR01	36	N024	53	C4
3	8S3	20	LM-P	37	B-1/2-1	54	JT/08
4	8S5	21	LM-T	38	B-1/2-2	55	AZITS1
5	8S6	22	LM-T15	39	C1-1/2-1	56	AZITS2
6	8S8	23	BCC7-M	40	C2-1/2-1	57	R1-01
7	8S9	24	BCC9-M	41	D1-1/2-1	58	R1-05
8	8S10	25	BCC10-M	42	D1-1/2-2	59	R1-06
9	14S1	26	W18-M	43	D2-1/2-1	60	R0-11
10	14S2	27	W29-M	44	D3-1/2-1	61	N05
11	14S3	28	B8	45	A1	62	N09
12	14S4	29	B9	46	A2	63	W00-M
13	14S5	30	B10	47	A3	64	G60
14	14S6	31	N09	48	A4	65	G105
15	14S8	32	N010	49	B1	66	GW60
						67	
16	14S9	33	N016	50	B3	67	A40
17	B11	34	N020	51	B4	68	A90

Table 7
Training dataset.

Order	Test	Order	Test	Order	Test
1	8S1	21	BCC7-M	41	A2
2	8S2	22	BCC10-M	42	A3
3	8S3	23	W18-M	43	B1
4	8S5	24	W29-M	44	B3
5	8S6	25	B8	45	B4
6	8S8	26	B9	46	C3
7	8S9	27	B10	47	C4
8	14S1	28	N09	48	JT/08
9	14S2	29	N016	49	AZITS1
10	14S3	30	N020	50	R1-01
11	14S4	31	N022	51	R1-05
12	14S5	32	N024	52	R1-06
13	14S6	33	B-1/2-1	53	R0-11
14	14S8	34	B-1/2-2	54	N05
15	B11	35	C1-1/2-1	55	N09
16	B12	36	D1-1/2-1	56	W00-M
17	SBR01	37	D1-1/2-2	57	G105
18	LM-P	38	D2-1/2-1	58	GW60
19	LM-T	39	D3-1/2-1	59	A40
20	LM-T15	40	A1	60	A90

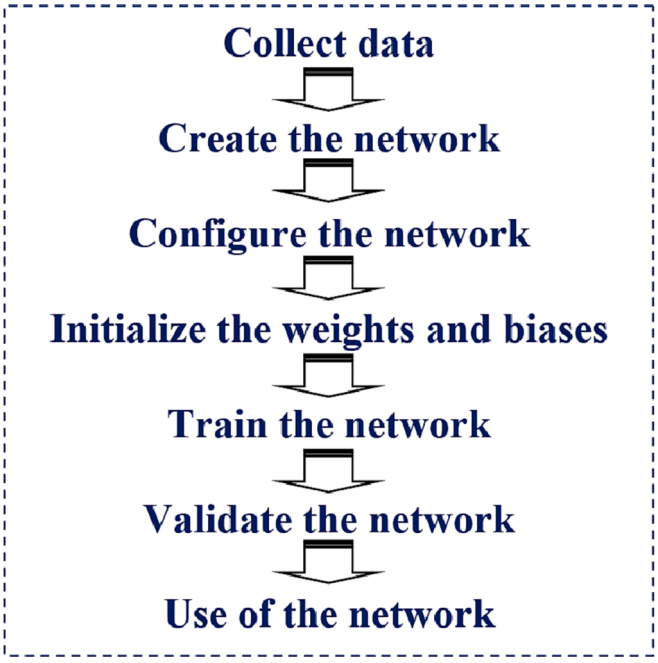

Fig. 4. Standard steps for designing an artificial neural network (ANN).

5.4. Selection of training algorithm and number of neurons

5.4.1. Selection of training algorithm

The NN Toolbox software contains a list of available training algorithms which compute either the gradient or the Jacobian by performing calculations backward through the network [44]. Two-layer feed-forward networks, with a sigmoid hidden layer and linear output layer, can fit multi-dimensional problems, given consistent data and enough neurons in the hidden layer. An ANN that is just large enough to obtain an adequate fit can improve the network generalisation. As the size of the network increases, the functions it can create become more complex. Therefore, small enough networks will not have enough power to produce overfitting. Among the different possibilities offered by the NN Toolbox, three of the most widely used algorithms were analysed: Levenberg-Marquardt (LM) backpropagation, scaled conjugated gradient (SCG) backpropagation, and Bayesian regularisation (BR) backpropagation.

Table 8
Testing dataset.

Order	Test
1	8S10
2	14S9
3	BCC9-M
4	N010
5	C2-1/2-1
6	A4
7	AZITS2
8	G60

5.4.1.1. Levenberg-Marquardt (LM) backpropagation: Trainlm. This training function updates the weight and bias values as stated by the LM optimisation. The *trainlm* routine is usually the fastest backpropagation algorithm in the toolbox, and is often recommended as the first-option algorithm, although it requires more memory than other algorithms. Particularly, in function approximation problems, *trainlm* will have the fastest convergence and can obtain lower mean square errors than other algorithms. However, the advantage of *trainlm* decreases with the increase of the number of weights in the network. Therefore, it is recommended for small and medium size networks, if there is enough memory available.

5.4.1.2. Scaled conjugate gradient (SCG) backpropagation: Trainscg. The SCG algorithm uses the scaled conjugate gradient method to update the weight and bias values, and performs well for networks with a large number of weights. The *trainscg* procedure is faster than the LM algorithm for function approximation problems where large networks and few memory demands are involved.

5.4.1.3. Bayesian regulation (BR) backpropagation: Trainbr. As previously discussed, the BR algorithm is particularly well-suited to cases involving a relatively small dataset available for network training, and it uses the whole of the available dataset for training purposes without requiring a validation set. Accordingly, the methodology prevents data discards and maximises the amount of training data. Furthermore, BR reduces previous works before training and preserves an optimal network size [44]. The *trainbr* routine updates the weight and bias values considering the LM optimisation. In the regularisation process, it minimizes a combination of squared errors and weights. Subsequently, it determines the right combination to produce a network that generalises adequately. The BR algorithm performs well on function approximation

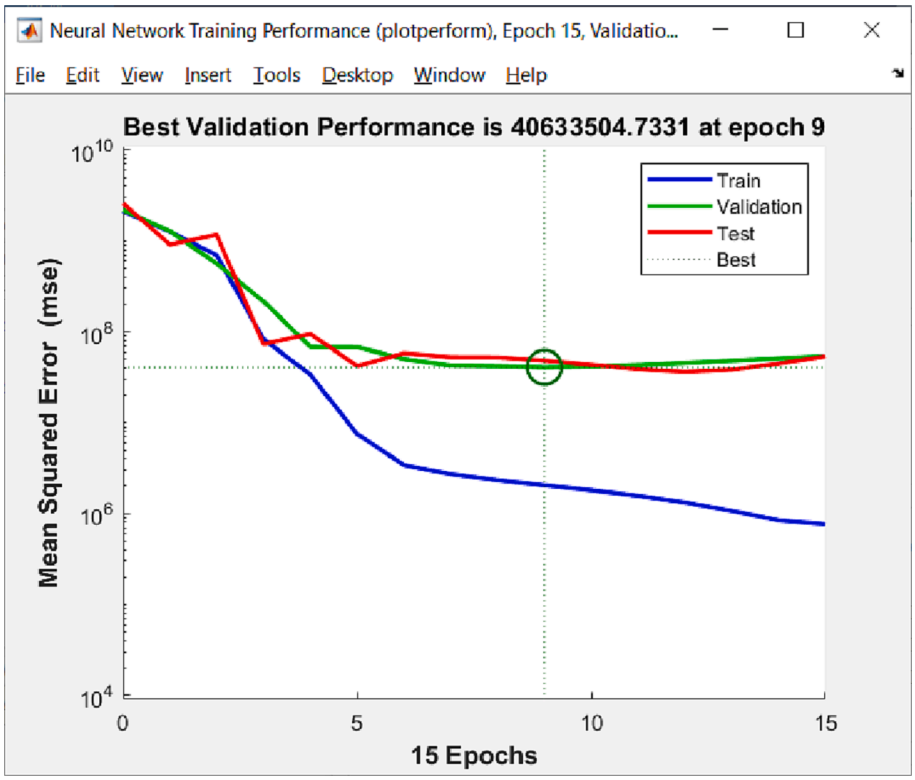

Fig. 5. Best performance found for the Levenberg-Marquardt (LM) training with ten neurons.

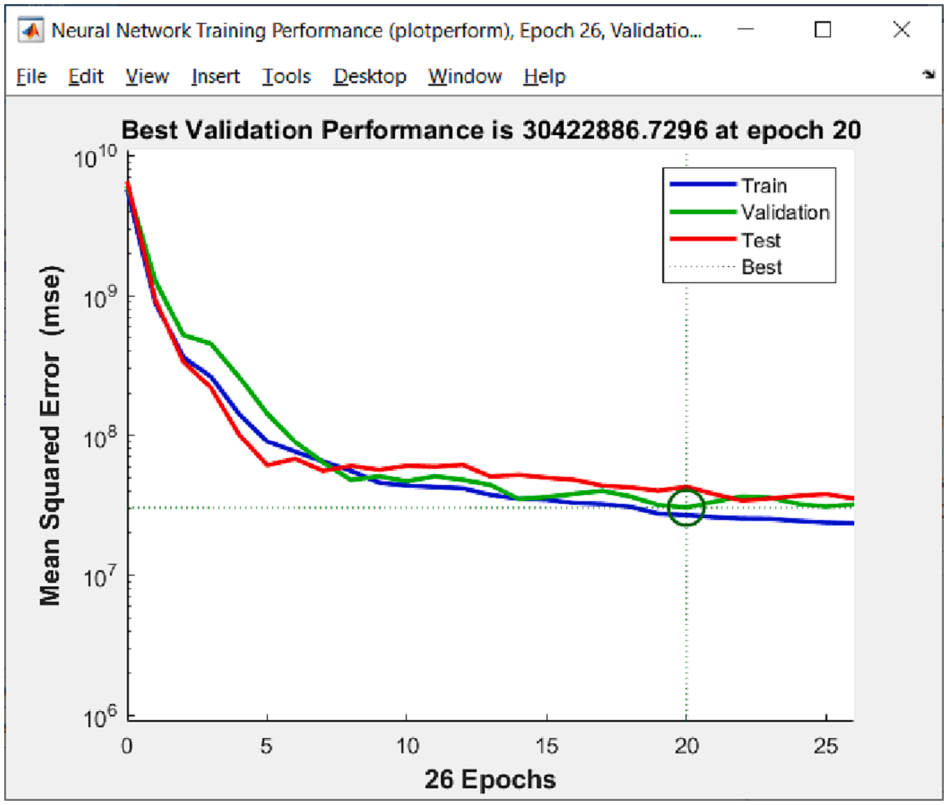

Fig. 6. Best performance found for the scaled conjugate gradient (SCG) training with ten neurons.

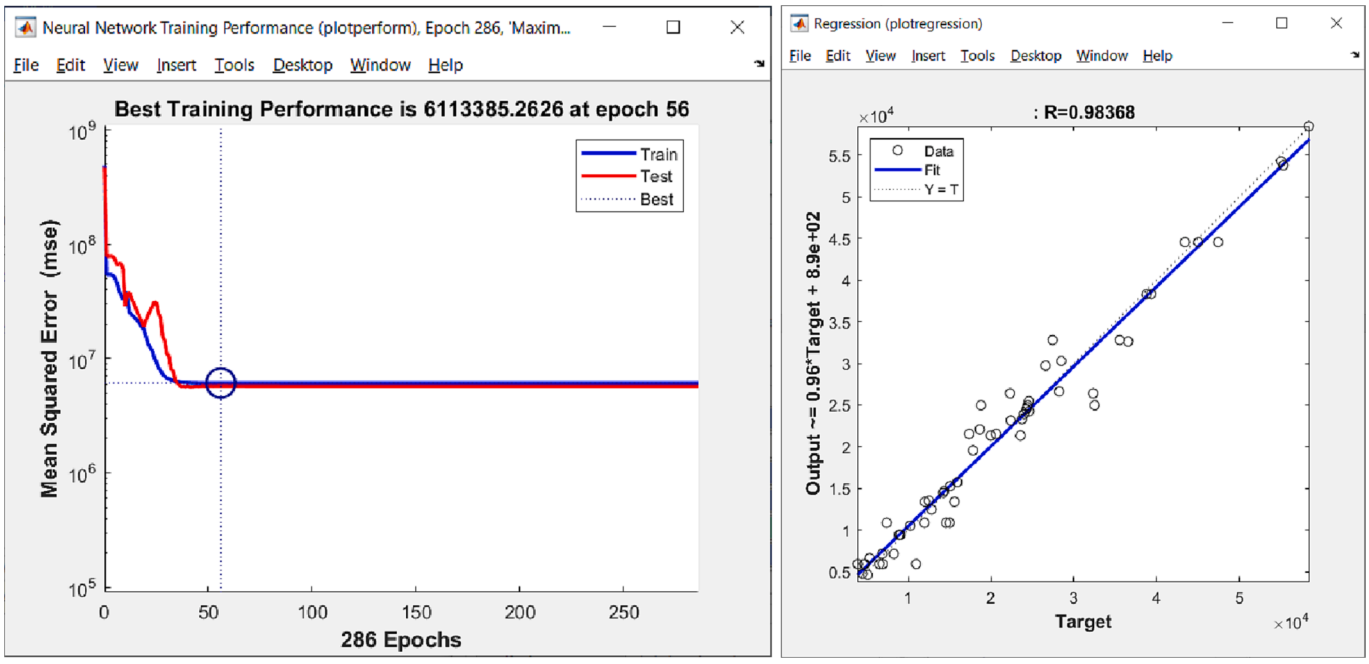

Fig. 7. Best performance found for the Bayesian regulation training with ten neurones.

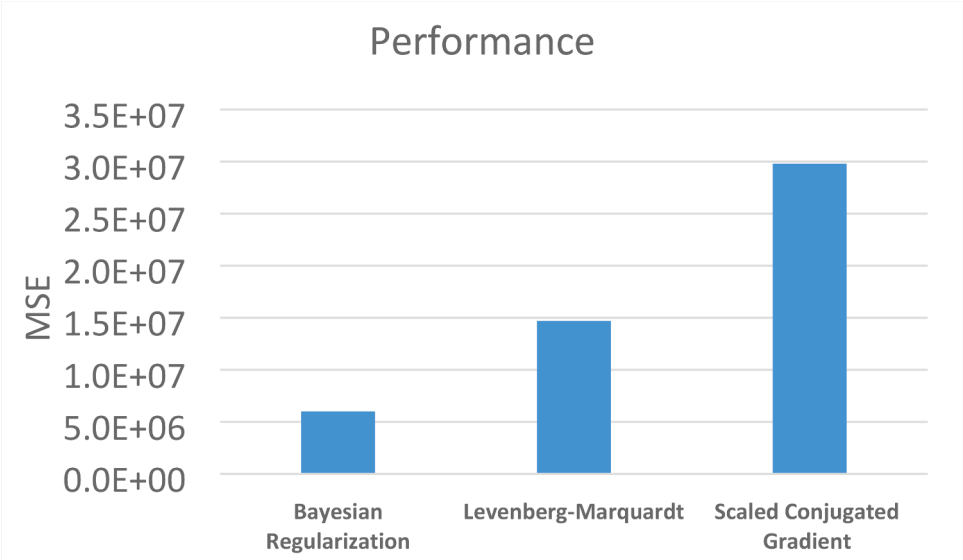

Fig. 8. Performance comparison for the studied training methods.

problems since the approximation to the Hessian used in the LM algorithm is accurate when the network output is not saturated. The main disadvantage of the BR method is that it takes longer to converge than early stopping.

A comparison between the three methods is here proposed for a network with ten neurons. For the LM and SCG training algorithms, 70% of the samples were set aside for training, 15% for validation, and 15% for testing. Validation allowed us to measure the network generalisation and stop the training before over-fitting. In addition, testing provided an independent measure of network performance during and after training. Several hundred networks were trained in each case until an optimal performance value, in terms of MSE, was reached. In the case of the LM and SCG methods, a reasonable good performance meant that the test set error and validation set error characterisations were similar and the test curve did not increase significantly before the validation curve increased. Figs. 5 and 6 show the best validation performance in terms of MSE regarding the number of epochs, where an epoch represent the training of the neural network with all the training data for one cycle.

Fig. 7 shows the results for the best solution found through BR training, leading to a correlation coefficient R equal to 0.98368 from the comparison between outputs and targets in the regression plot. Fig. 8 shows the performance comparison for the three studied training methods in terms of the MSE. As expected, the methodology that offered the best results was BR, which used all data for training in a small-data problem context. In the case of the other two methods, the LM training provided better results than the SCG training probably because the MSE stabilised early and the LM worked well when the number of weights in the neural network was not large. Regarding the speed, the fastest algorithm was LM, but for this network configuration, with just ten neurons in one layer, speed did not become a real problem.

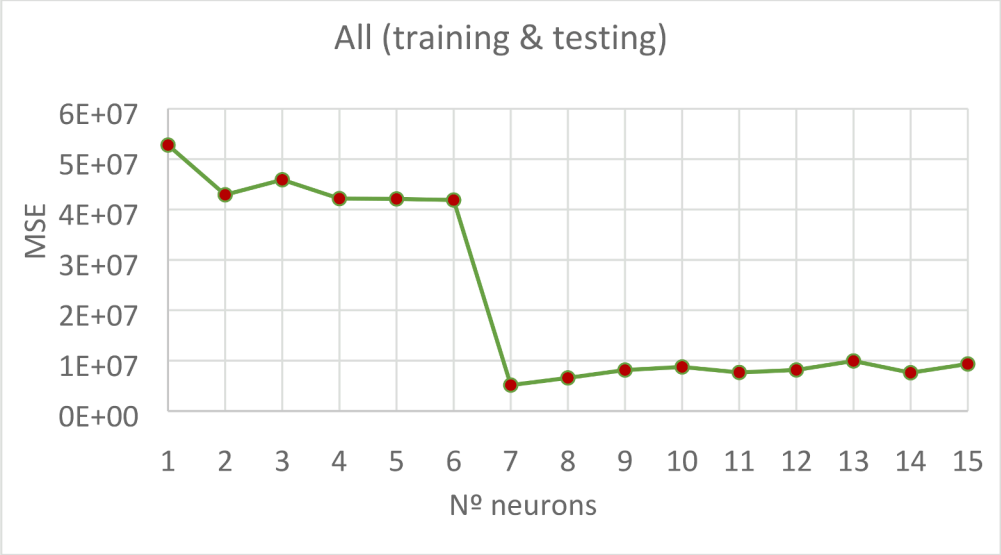

Fig. 9. Mean squared error regarding the number of neurons.

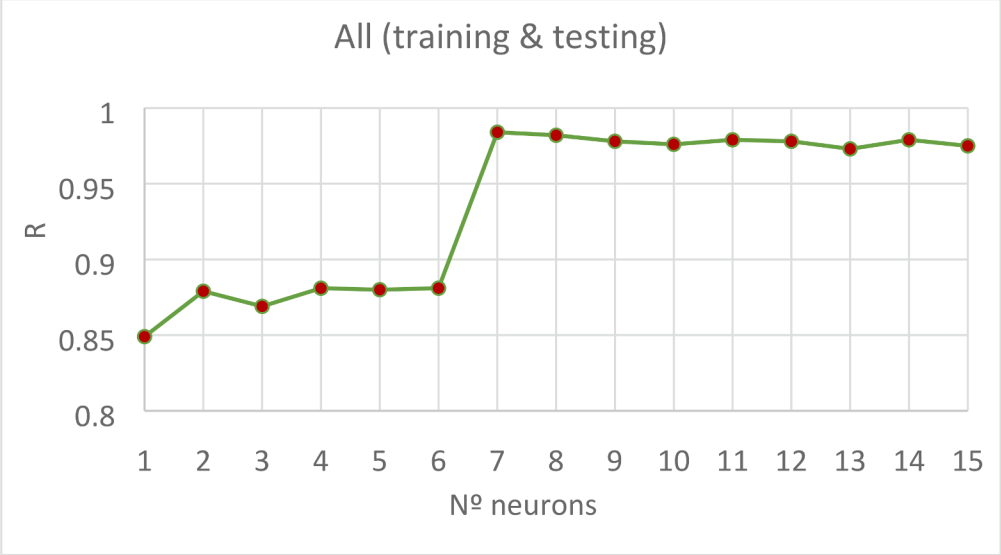

Fig. 10. Correlation coefficient regarding the number of neurons.

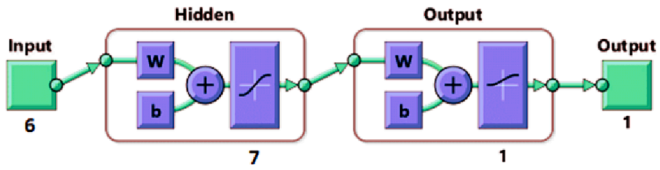

Fig. 11. Artificial neural network (ANN) architecture.

5.4.2. Number of neurons in the hidden layer

A large number of neurons in the hidden layer provides the network with more flexibility because it has more parameters to optimise. Nevertheless, if the hidden layer is too large, it might lead to an under-characterisation of the problem, since the network must optimise more parameters than existing data vectors to constrain these parameters. The training points could be adequately fitted, but the fitting curve would randomly oscillate between these points. In contrast, too few neurons can lead to underfitting.

There are still no strong mathematical methods for predicting the number of neurons in the hidden layers. However, several rules of thumb have been presented in the literature. Blum et al. [49] proposed that the size of this hidden layer must be between the size of the input layer size and that of the output layer. Furthermore, Berry et al. [50] proposed that the size should never be more than twice that of the input layer. Even so, these rules do not take into account the number of training samples. Other empirical rules, such as those of Ke et al. [51] and Shibata et al. [52], only work for particular situations. Huang and Babri [53] proved that a single hidden layer feedforward network with at most N_h hidden neurons can learn different N_s samples without error. This statement applies to any bounded, non-linear activation function with the limit at one infinity [53]. Thus, the upper bound for a single layer feedforward network is $N_h \leq N_s$. Hecht-Nielsen [54] stated that any continuous function f defined on an n-dimensional unit cube can be implemented exactly by a three-layer network with $2n + 1$ hidden nodes. According to Hecht-Nielsen, an upper bound for the number of hidden neurons using Kolmogorov's theorem could be $N_h \leq 2N_i + 1$ where N_h is the number of hidden neurons and N_i is the number of input

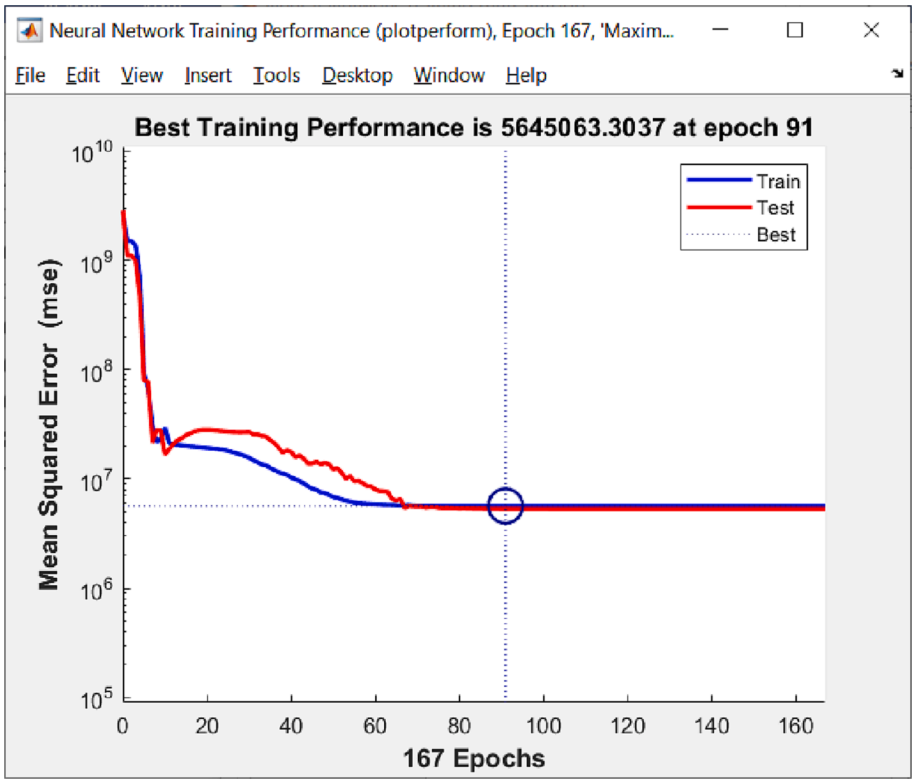

Fig. 12. Neural network training performance progress.

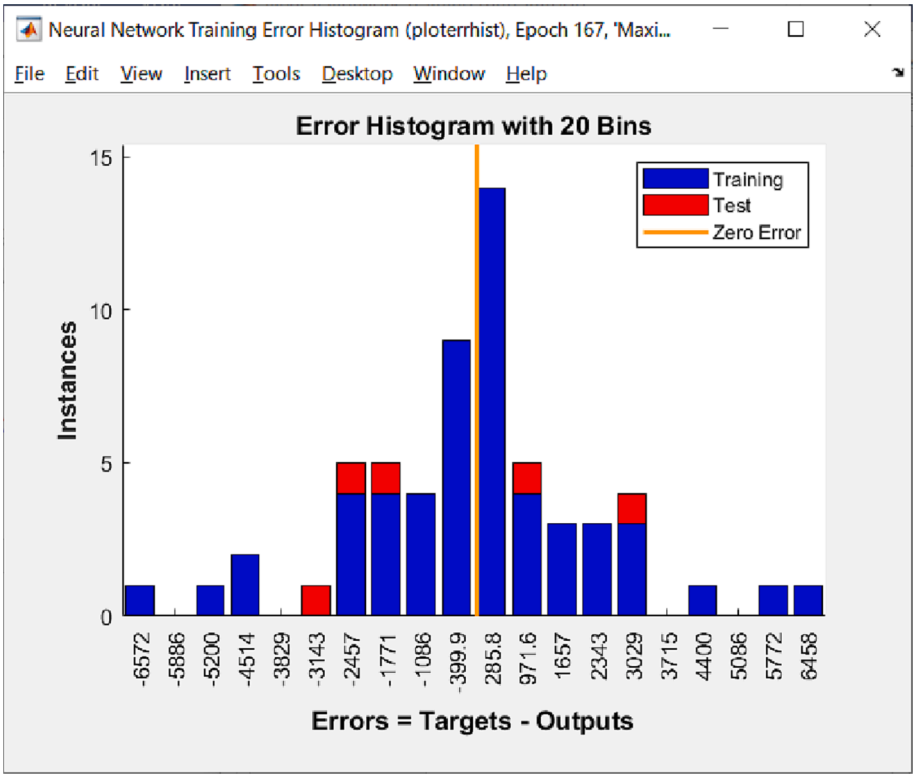

Fig. 13. Error histogram.

parameters. Nevertheless, there is another possibility that consist of searching through all the reasonable network topologies within the theoretical bounds and choosing the one with the least generalisation error. The main problem with this approach is that it can be very time-consuming but only for networks with more than one hidden layer. Since this study's network consisted of a single hidden layer, this methodology

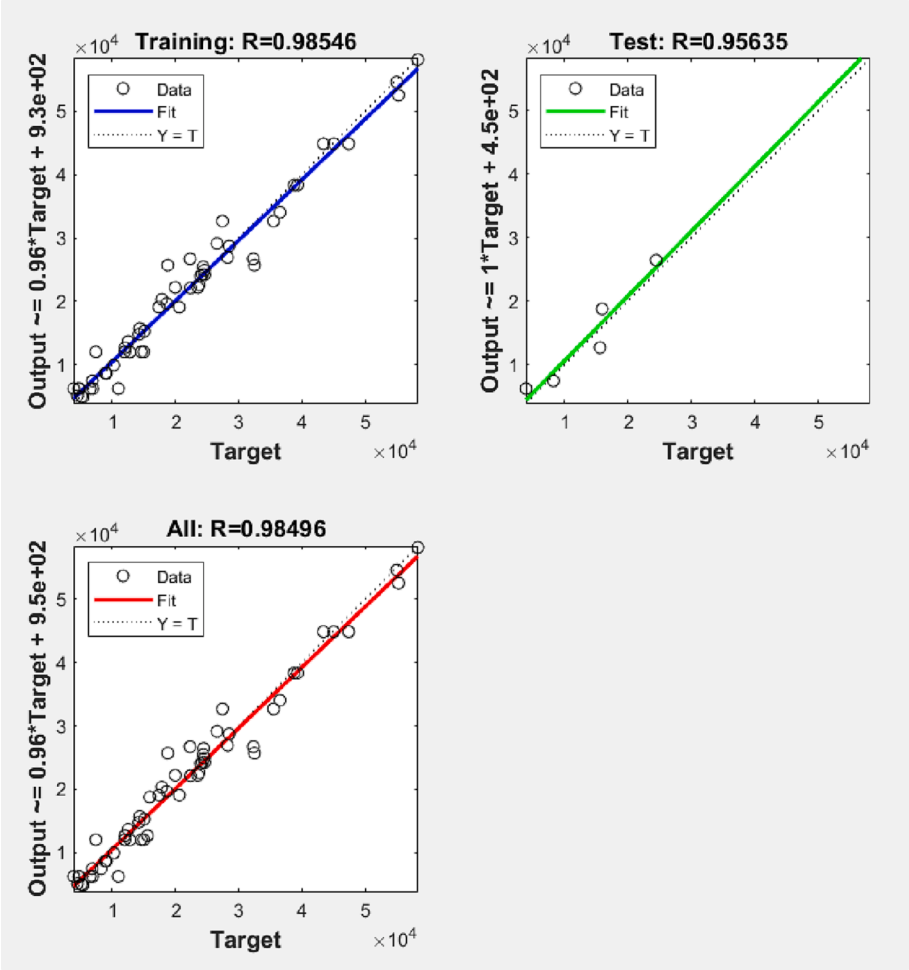

Fig. 14. NN Training Regression Plot.

Table 9
Validation of the neural network (NN) with the tests preserved from the databank.

	S_{j,ini_exp} (kNm/rad)	S_{j,ini_NN} (kNm/rad)	S_{j_NN}/S_{j_exp}	S_{j,ini_Kishi} (kNm/rad)	S_{j_Kishi}/S_{j_exp}	S_{j,ini_Faella} (kNm/rad)	S_{j_Faella}/S_{j_exp}
8S10	24,919	28,656	1.15	36,712	1,47	23,420	0,94
14S9	28,172	32,670	1.16	47,188	1,67	39,572	1,40
BCC9-M	18,819	22,204	1.18	41,878	2,23	31,575	1,68
N010	47,457	46,922	0.99	242,721	5,11	39,018	0.82
C2-1/2-1	6380	6244	0.98	3540	0,55	2646	0,41
A4	24,217	26,448	1.09	118,160	4,88	28,750	1,19
AZITS2	26,049	25,649	0.98	44,898	1,72	24,984	0,96
G60	41,532	44,881	1.08	105,589	2,54	32,128	0,77
		Mean	1.08	Mean	2,52	Mean	1,02
		StDev	0.08	StDev.	1,53	StDev.	0,37

could be used to find the optimum number of neurons [55]. Therefore, the network was trained several times, varying the number of neurons in the hidden layer between 1 and 16, until reiterative low values of MSE were obtained for each case. Figs. 9 and 10 show the performance results in terms of both the minimum MSE found and the correlation coefficient R. From $N_h = 7$, no significant improvements were obtained in either of the two indices. Besides, this number of neurons in the hidden layer was consistent with the recommendations of Berry et al. and also with respect to the suggestions of Huang-Babri and Hecht-Nielsen. Accordingly, the neural architecture described in Fig. 11 was chosen. It consisted of six input parameters, one layer of neurons with seven hidden neurons, and the rotational stiffness as the only output from the network.

5.5. Training and testing of ANN

The toolbox allowed the creation of custom functions leading to a great control over the network algorithms. Thus, a simple generated script was modified to check the network performance each time the training was completed. The possibility of using the multiple networks training technique was analysed, since it represented a good choice for problems linking small dataset with BR [44]. However, the results taking the average of the outputs from multiple networks did not improve the performance of the best network found after the iterative process.

Therefore, after hundreds of training cycles using an Intel® Core ™ i7 – 1065 G7 processor, several networks producing low MSE values were studied. The correlation coefficients and error histograms were

Table 10
Geometrical data of the tests out from SCDB databank (millimetres).

Test	Author	h_b	t_{ta}	L_{ta}	g	t_{wa}	L_{wa}
R1-04	Mander et al. [40]	210	9,5	165,1	50,8	0	0
R1-09	Mander et al. [40]	210	9,5	165,1	50,8	0	0
TSC-A	Bernuzzi et al. [41]	300	11,8	180	60.0	0	0
TSC-B	Bernuzzi et al. [41]	300	11,8	180	60.0	0	0
TSC-C	Bernuzzi et al. [41]	300	11,8	180	60.0	0	0
TSC-D	Bernuzzi et al. [41]	300	11,8	180	60.0	0	0
TP3A	Chasten et al. [63]	684	22,2	254	114,3	0	0
TP3B	Chasten et al. [63]	684	22,2	254	114,3	0	0
T1_Jobaer	Jobaer et al. [65]	354	12.0	202	64	10	214
TS18-i2	Zandonini et al. [64]	300	18.4	180	60	0	0
TS16-i2	Zandonini et al. [64]	300	16.4	180	60	0	0
TSW9-i2	Zandonini et al. [64]	300	16.4	180	60	8.4	260
TSW10-i2	Zandonini et al. [64]	300	18.4	180	60	10.4	260
TS18-i5	Zandonini et al. [64]	300	18.4	180	60	0	0
TS16-i5	Zandonini et al. [64]	300	16.4	180	60	0	0
TSW9-i5	Zandonini et al. [64]	300	16.4	180	60	8.4	260
TSW10-i5	Zandonini et al. [64]	300	18.4	180	60	10.4	260
101,003	Janss et al. [62]	200	15	150	50	0	0
101,012	Janss et al. [62]	300	15	150	50	0	0

Table 11
Neural network results for off-database tests.

	S_{j,ini_exp} (kNm/rad)	S_{j,ini_NN} (kNm/rad)	S_{j_NN}/S_{j_exp}
R1-04	13,876	12,043	0.87
R1-09	13,967	12,043	0.86
TSC-A	21,550	17,979	0.83
TSC-B	18,100	17,979	0.99
TSC-C	17,600	17,979	1.02
TSC-D	16,450	17,979	1.09
TP3A	96,032	93,916	0.98
TP3B	97,162	93,916	0.97
T1_Jobaer	29,244	26,368	0.90
TS18-i2	24,225	24,758	1.02
TS16-i2	38,130	22,903	0.60
TSW9-i2	44,595	39,534	0.89
TSW10-i2	50,150	55,111	1.10
TS18-i5	27,935	24,758	0.89
TS16-i5	35,010	22,903	0.65
TSW9-i5	32,830	39,534	1.20
TSW10-i5	38,875	55,111	1.42
101,003	17,650	14,201	0.80
101,012			
101,012	36,365	37,671	1.04
		Mean	0.95
		StDev.	0.18

checked and the best-trained network was selected. Fig. 12 shows the training performance progress for the best solution. This figure also indicates the iteration at which the performance reached a minimum. Fig. 13 shows the error histogram, which is the histogram of the errors between the target values and predicted values following the training of a feedforward neural network. The bins are the number of vertical bars on the graph and the Y-axis represents the number of samples from the dataset regarding a particular bin.

The neural network training regression plot is presented in Fig. 14, showing the relationship between the outputs of the network and the targets. The dashed line in each axis represents the ideal result where the outputs are equal to the targets. The solid line represents the best-fit line between the outputs and targets. The correlation value *R* indicates the relationship between the outputs and targets. If *R* is close to 1, this indicates that the relationship between the outputs and targets is quasi-linear. Indeed, a value of R between 0.9 and 1 indicates a very high positive correlation. In this case, the training and test correlations provided values of 0.98 and 0.95 respectively, indicating a relatively good fit, so that the network seemed suitable for predicting the rotational stiffness.

5.6. Validation of the ANN

Once the training was completed, the network performance was checked with the test set preserved from the databank. Additionally, the predictive capacity of the ANN was compared with two highly proven analytical methodologies: the power model approach of Kishi and Chen [14] and the methodology proposed by Faella et al. [56]. Table 9 shows the prediction results, where S_{j,ini_exp} is the experimental rotational stiffness, S_{j,ini_NN} is the prediction from the NN, and S_{j,ini_Kishi} and S_{j,ini_Faella} are the predictions from Kishi and Chen, and Faella et al., respectively.

The neural network proposal showed a good agreement with the experimental results, with a mean of 1.08 and standard deviation of 0.08.

To further assess the net generalisation capability, it could be helpful to check the ANN with tests not included in the SCDB database, some of them, if possible, including predictors with values outside the training ranges. Although experimental references are available in the literature, a few of them were not considered because they are not symmetrical tests [57], they are minor axis tests [58], or they are non-preloaded as in [59–61]; the TSC tests, from the second series of Bernuzzi et al. [41], and 101,006 from Janss et al. [62] were also excluded. Table 10 shows the geometries related to the off-databank tests and Table 11 shows the selected available tests with their respective references and geometrical data, where S_{j,ini_exp} is the experimental rotational stiffness and S_{j,ini_NN} is the neural network prediction for the rotational stiffness. The tests from Mander et al. [40], Bernuzzi et al. [41], and Chasten et al. [63] are tests carried out on the same geometry. The initial stiffness values were obtained from each reference, since the original data from the moment-rotation curves were not available. In the case of the experimental data from Zandonini et al. [64], the stiffness values were obtained with the *RSTAB*® structural analysis program, through an iterative procedure targeted to match the displacements measured in the experimental work. To carry out this task, the tests were simulated using the *RSTAB*® frame & truss analysis software as semi-rigid frame models where the connections were defined by means of a rotational spring. The starting rotational stiffness was iteratively modified until a maximum displacement equal to that referred to by the authors in each test was obtained. Fig. 15 shows the *RSTAB*® model for specimen TSW10-i2.

Table 11 shows the comparison between the experimental and stiffness prediction values for all off-database tests. The mean was 0.95 and standard deviation reached 0.18. This statistical results show a good performance of the ANN, even for cases where the stiffness was not evaluated in the same consistent way as for the training set.

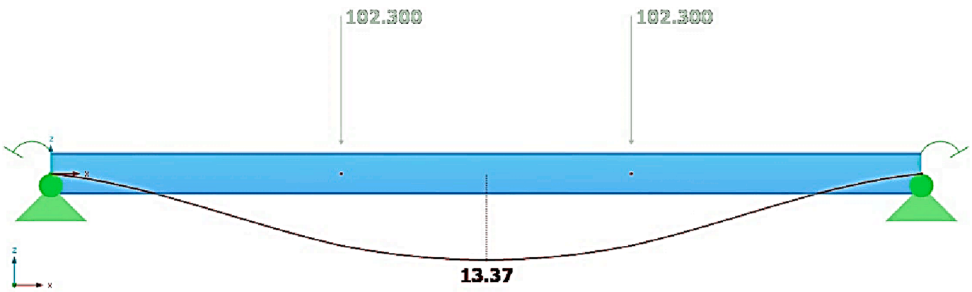

Fig. 15. RSTAB model for test TSW10-i2.

Table 12
Maximum and minimum values for the parameters in the training dataset.

Parameter	h_b (mm)	t_{ta} (mm)	L_{ta} (mm)	g (mm)	t_{wa} (mm)	L_{wa} (mm)
Max. Value	460,4	15,9	356	105	10	290
Min. Value	127,2	7,9	121	35	0	0

6. Discussion

The Kishi and Chen power model provided larger stiffness values for most of the tests, probably because it did not take into account the flexibility of the panel zone. Otherwise, the method of Faella et al. [56] yielded better results than the Kishi and Chen model for the analysed experimental tests, with a mean value close to 1, but the standard deviation was still high relative to the results from the neural network approach. However, considering that the results for test C2-1/2-1 showed a ratio of 0.42 between S_{j_Faella}/S_{j_exp}, if this result was removed from the test set, the standard deviation of the Faella's et al. methodology would be better. The prediction from the Kishi and Chen model for this test provided a unique S_{j_Kishi}/S_{j_exp} ratio that fell below 1. The C2-1/2-1 test was included in the experimental work of Marley et al. [36]. This series of tests presented low stiffness values, since the beam depth was just 127 mm, the minimum in the dataset. Specifically, this test had the same geometrical properties as the other 12 specimens (A1/2-1 to D3-1/2-2), but the experimental stiffness values were very different in some cases, ranging between 1948 kNm/rad and 10,976 kNm/rad. Nevertheless, the use of this specimen for the neural network testing was considered significant since its stiffness fell in the central range of results from the series. These low values of stiffness predictions for this particular case, in comparison with the other predictions, could indicate that both the power model and the Faella et al. method were sensitive to low values of beam depth, which is a parameter closely related to the lever arm in mechanical models.

The methodology of Faella et al. considered the panel components,

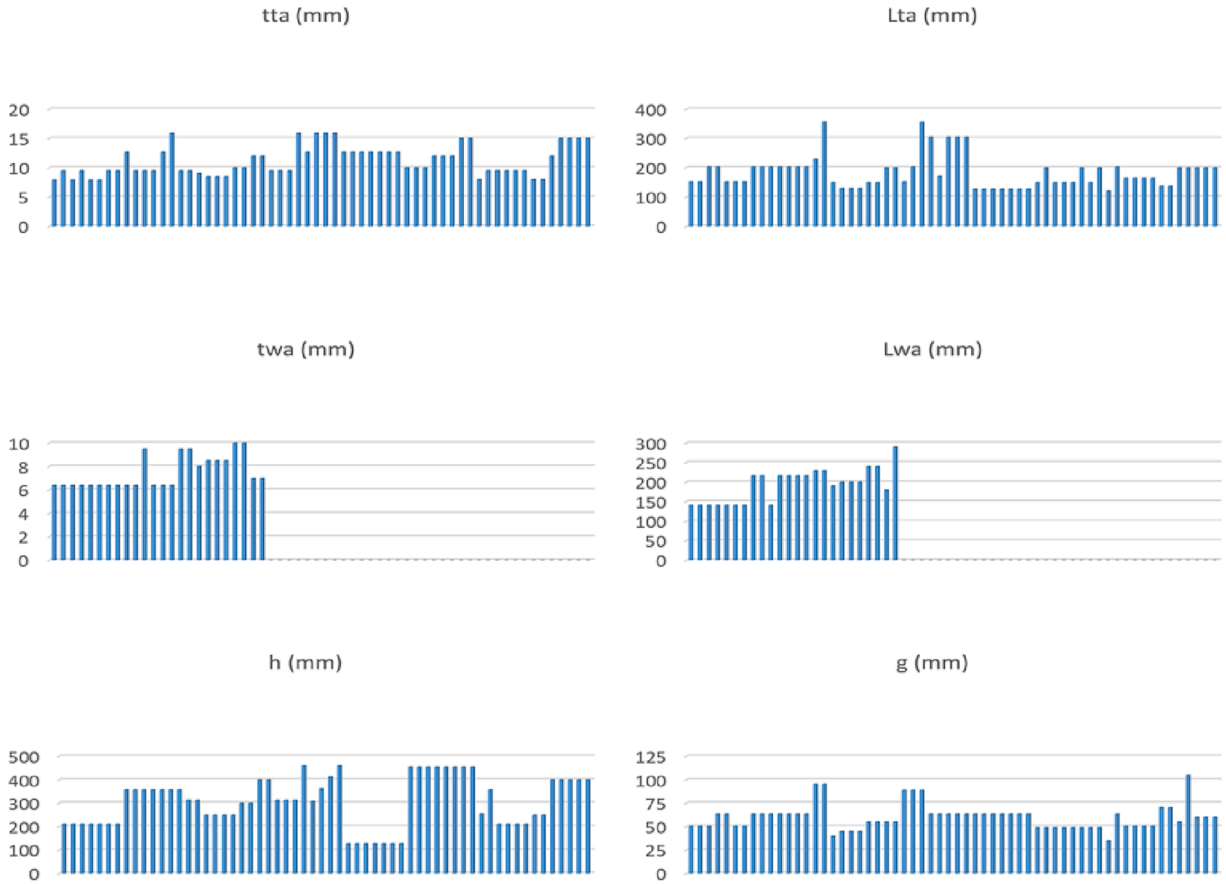

Fig. 16. Range of predictors in the training set.

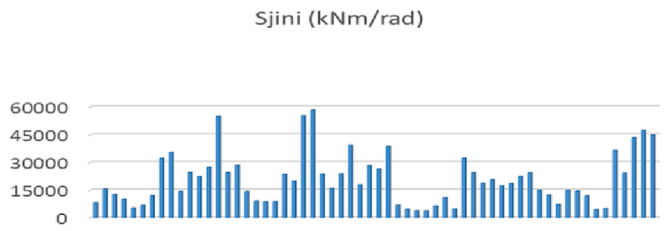

Fig. 17. Output range in the training set.

but required the introduction of a large number of input parameters. In the neural network model, the input process was quick and easy due to the small number of parameters involved.

The proposed ANN is just one of the possible solutions for the prediction of rotational stiffness: any other combination of tests in the training set will provide other valid but different solutions. However, generating a reliable ANN with such a limited amount of data is a challenging task. The problem required a detailed study of the different network architecture options, as well as the most appropriate training methodology. The number of data can be artificially increased, either with synthetic data or with data from a parametric study with the support of a finite element method. Methods such as the synthetic minority over-sampling technique (SMOTE) [66] could create new data points from the minority class data points, in this case, the TSDW samples. Nevertheless, an overgeneralisation problem can arise since the majority class is not taken into account. Moreover, there is a lack of flexibility related to the SMOTE technique because the number of synthetic samples is fixed in advance. Thus, the use of real data from available experimental tests has been preferred in order to develop a realistic model. These data inherently incorporate geometric imperfections, the actual material properties, and other issues regarding the practice of constructional steel design. However, the limited number of data require the setting of allowable predictor ranges and the further study of the ANN generalisation power. Table 12 shows the limit ranges of the predictors regarding the training dataset. Figs. 16 and 17 show the values of predictors and the output values, respectively, along the training set. The ranges of the beam depths and the top angle thicknesses were not too large, with relatively small upper limit values. The other predictors presented comparatively wide ranges but with the evident discontinuities with respect to the web angle parameters.

Regarding the off-database tests, the prediction for the specimens from Chasten et al. [63], which are characterised by high values of the top angle thickness and the beam depth, was fairly good. Moreover, the prediction for the tests presented in Zandonini et al. [64] concerning values of the top angle thickness outside the training range, was moderately good in most cases. Nevertheless, the results of the TS16-i2 and TS16-i5 tests were conservative while those of tests TSW10-i2 and TSW10-i5 were non-conservative. A possible explanation would be that, for some test geometries, the neural network does not work properly because these geometries, a particular geometrical parameter, or a combination of them are not well represented in the training set. The top and seat angle thicknesses were relatively large for these specimens and there were only five tests considering top angle thicknesses larger than 15 mm in the training set, as shown in Fig. 16. However, the prediction for tests TS18-i2 and TS18-i5, which considered TS angle connections with a top angle thickness of 18.4 mm, provided better results, even when this particular thickness was out of the training set range. This issue could indicate good generalisation characteristics, as well as a tendency to produce conservative results in the case of TS joints and non-conservative results in the case of TSDW joints when the top angle thickness is large.

7. Conclusions

A database of experimental rotational stiffnesses obtained for TS and TSDW angle connections was created through a linear regression of data points in the elastic range of the moment-rotation curves from the SCDB databank. Subsequently, a reduced number of significant input parameters was selected considering the size of the database and recommendations from well-known authors in the literature. Then, the data were analysed with a multiple linear regression model, once the normality hypothesis was confirmed, as a methodology to rule out outliers. For this, a discard criterion was established to obtain a database that was as clean as possible while maintaining a sufficient number of data for training the network. In this way, the ANN could be trained starting from a database that was, although small in size, consistent. Once the data were cleaned, an ANN was developed after having studied the network architecture in relation to the number of parameters to be considered and the number of neurons in the hidden layer. Among the different options studied, the best network architecture for this small-data problem was the one corresponding to a single hidden layer with seven neurons. The network was trained using a BR algorithm, which is particularly suitable for cases involving small datasets. The network was successfully tested with specimens from the SCDB databank and with other experimental tests not included in the databank. The results of the predictions made with these off-database tests were fairly good, even when some values of the top angle thickness and beam depth were outside the training range. The prediction accuracy in the TSDW and TS typologies, both for the in-database and off-database tests, was similar, with the exception of some off-database cases where the top angle thickness was large. The ANN proposed is valid for preloaded and symmetrical bolted angle connections in a major axis, and it can be reliably used within the range of the training input parameters. The presented methodology can be improved in the future by updating existing databases with the incorporation of new experimental results.

Declaration of Competing Interest

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

Acknowledgements

The financial support provided through grant PID2020-113895GB-C31 funded by MCIN/AEI/10.13039/501100011033 is gratefully acknowledged.

References

[1] Evers HGA, Maatje F. Cost based engineering and production of steel constructions. Strength and Design; 2002.
[2] J.D. Schippers, D.J. Ruffley, G.A. Rassati, J.A. Swanson. A design procedure for bolted top-and-seat angle joints for use in seismic applications, 7th International Workshop on Joints in Steel Structures (2012) Timisoara, Romania.
[3] Tingley DD, Davison B. Developing an LCA methodology to account for the environmental benefits of design for deconstruction. Build Environ 2012;57:387–95.
[4] Loureiro A, Gutiérrez R, Reinosa JM, Moreno A. Axial stiffness prediction of non-preloaded T-stubs: An analytical frame approach. J Constr Steel Res 2010;66(12):1516–22.
[5] Lemonis ME, Gantes CJ. Mechanical modeling of the nonlinear response of beam-to-column joints. J Constr Steel Res 2009;65(4):879–90.
[6] Reinosa JM, Loureiro A, Gutierrez R, Lopez M. Analytical frame approach for the rotational stiffness prediction of beam-to-column angle connections. J Constr Steel Res 2015;106:67–76.
[7] Reinosa JM, Loureiro A, Gutierrez R, Lopez M. Analytical frame approach for the axial stiffness prediction of preloaded T-stubs. J Constr Steel Res 2013;90:156–63.
[8] Skejic D, Dujmovic D, Beg D. Behaviour of stiffened flange cleat joints. J Constr Steel Res 2014;103(12):61–76.
[9] Reinosa JM, Loureiro A, Gutierrez R, Lopez M. Analytical plate approach for the axial stiffness prediction of stiffened angle cleats. J Constr Steel Res 2015;106:77–88.

[10] Reinosa JM, Loureiro A, Gutierrez R, Lopez M. Mechanical stiffness prediction of beam-to-column stiffened angle joints. J Constr Steel Res 2020;168:105875.
[11] Kishi N, Ahmed A, Yabuki N, Chen WF. Nonlinear finite element analyses of top and seat-angle with double web-angle connections. Int J Struct Eng Mech 2001;12:201–14.
[12] Citipitioglu A, Rami Hajali M, White DW. Refined 3D finite element modeling of partially-restrained connections including slip. J Constr Steel Res 2002;58(5–8):995–1013.
[13] Eurocode 3. Design of Steel Structures, Part 1-8: Design of Joints, European Committee for Standardization, EN 1993-1-8, Brussels; 2005.
[14] Chen WF, Kishi N, Komuro M. Semi-rigid connections handbook (Civil & environmental engineering). J Ross Publishing 2011.
[15] Tahir ZR, Mandal P. Artificial neural network prediction of buckling load of thin cylindrical shells under axial compression. Eng Struct 2017;152:843–55.
[16] Stavroulakis GE, Avdelas AV, Abdalla KM, Panagiotopoulos PD. A neural network approach to the modelling, calculation and identification of semi-rigid connections in steel structures. J Constr Steel Res 1997;44(1–2):91–105.
[17] Anderson D, et al. Application of artificial neural networks to the prediction of minor axis joints. Comput Struct 1997;63:685–92.
[18] de Lima LRO, Vellasco PCG da S, Andrade SAL de, Silva JGS da, Vellasco MMBR. Neural networks assessment of beam-to-column joints. J Brazil Soc Mech Sci Eng 2005;27(3):314–24.
[19] Salajegheh E, Gholizadeh S, Pirmoz A. Self-organizing parallel back propagation ANN for predicting the moment-rotation behavior of bolted connections. Asian J Civ Eng 2008;9(6):625–40.
[20] Kim J, Ghaboussi J, Elnashai AS. Mechanical and informational modeling of steel beam-to-column connection. Eng Struct 2010;32(2):449–58.
[21] Specification for Structural Steel Buildings. ANSI/AISC 360-16. AISC (2016).
[22] A.V. Goverdhan. A collection of experimental moment - rotation curves and evaluation of prediction equations for semi - rigid concertinos, 1984, Ph. D. Thesis, Vanderbilt University, Nashville, TN.
[23] Nethercot DA. Steel beams to column connections a review of test data and their applications to the evaluation of joint behaviour on the performance of steel frames. London: CIRIA Project Study; 1985. p. 338.
[24] N. Kishi, W.F. Chen. Database of steel beam–to-column connections, Vols I and II. Structural Engineering Report no. CE-STR-86-26, School of Civil Engineering, Purdue University, W. Lafayette, IN, 1986.
[25] Chen WF, Kishi N. Semirigid steel beam-to-column connections: Data base and modeling. J Struct Eng 1989;115:105–19.
[26] Frye MJ, Morris GA. Analysis of flexibly connected steel frames. Can J Civil Eng 1975;2(3):280–91.
[27] Chen WF, Lui EM. Stability design of steel frames. Boca Raton: CRC Press; 1991.
[28] Altman WG, Azizinamini A, Bradburn JH, Radziminski JB. Moment-rotation characteristics of semi-rigid steel beam-column connections structural research studies. Department of Civil Engineering: University of South Carolina, Columbia, SC; 1982.
[29] Rathbun JC. Elastic properties of riveted connections. T Am Soc Civ Eng 1936;101:524–63.
[30] Roeder CW, Knechtel B, Thomas E, Vaneaton A, Leon RT, Preece FR. Seismic behavior of older steel structures. J Struct Eng, ASCE 1996;122(4):365–73.
[31] Elnashai AS, Elghazouli AY, Denesh-Ashtiani FA. Response of semirigid steel frames to cyclic and earthquake loads. J Struct Eng, ASCE 1998;124(8):857–67.
[32] Fu Z, Ohi K, Takanashi K, Lin X. Seismic behavior of steel frames with semi-rigid connections and braces. J Constr Steel Res 1998;46(1):440–1.
[33] Calado L, De Matteis G, Landolfo R. Experimental response of top and seat angle semi-rigid steel frame connections. Mater Struct 2000;33(8):499–510.
[34] Komuro M, Kishi N, Matsuoka KG. Static loading tests for moment-rotation relation of top-and-seat-angle connections. J Constr Steel JSSC 2002;10:57–64 (Japanese).
[35] R.A. Hetchman, B.G. Johnston. Riveted semi-rigid beam-to-column building connections. Progress report No. 1, AISC Committee of Steel Structures Research, Lehigh University, Bethlehem, PA, 1947.
[36] Marley MJ, Gerstle KH. Analysis and tests of flexibly-connected steel frames. Chigago, IL: AISC; 1982. Report to AISC under Project 199,.
[37] Maxwell SM. A realistic approach to the performance and application of semi-rigid jointed structures. In: Howlett JH, Jenkins WM, Stainsby R, editors. Joints in Structural Steelwork. Plymouth: Pentech Press, UK; 1981. 2.71-2.98,.
[38] Davison JB, Kirby PA, Nethercot DA. Rotational stiffness characteristics of steel beam-to-column connections. J Constr Steel Res 1987;8:17–54.
[39] Harper WL. Dynamic response of steel frames with semi-rigid connections. Structural research studies. Department of Civil Engineering: University of South Carolina, Columbia, SC; 1990.
[40] Mander JB, Chen SS, Peckan G. Low-cycle fatigue behavior of semi-rigid top-and-seat angle connections. AISC Eng J 1994;31(3):111–22.
[41] Bernuzzi C, Zandonini R, Zanon P. Experimental analysis and modeling of semi-rigid steel joints under cyclic reversal loading. J Constr Steel Res 1996;38(2):95–123.
[42] Kubo N, Yoshida T, Hashimoto K, Tanuma Y. Column influence on the moment-rotation behaviour of semi-rigid angle connections. J Constr Steel, JSSC 1999;7:427–34 (Japanese).
[43] Sato Y, Komuro M, Kishi N. Experimental study on moment-rotation of top-and-seat angle connections. J Constr Steel, JSSC 2007;15:121–8 (Japanese).
[44] Beale MH, Hagan MT, Demuth HB. Neural Network Toolbox™, User's Guide R2013b. Natick, Mass, USA: The MathWorks; 2013.
[45] MacKay DJ. Bayesian interpolation. Neural Comput 1992;4(3):415–47.
[46] Azizinamini A, Radziminsky JB. Static and cyclic performance of semirigid steel beam-to-column connections. J Struct Eng 1989;115(12):2979–99.
[47] IBM SPSS Statistics Base V27. International Business Machines Corporation, 2022. https://www.ibm.com/docs/en/SSLVMB_27.0.0/pdf/en/IBM_SPSS_Statistics_Base.pdf.
[48] Harrell Jr FE, Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved prognostic prediction. Stat Med 1984;3(2):143–52.
[49] Blum A. ANN in C++: An object-oriented framework for building connectionist systems. New York, NY, USA: John Wiley & Sons, Inc.; 1992.
[50] Berry MJ, Linoff G. Data Mining Techniques: For Marketing, Sales, and Customer Support. New York, NY, USA: John Wiley & Sons, Inc.; 1997.
[51] J. Ke, X. Liu. Empirical Analysis of Optimal Hidden Neurons in Neural Network Modeling for Stock Prediction in Pacific-Asia. Workshop on Computational Intelligence and Industrial Application. PACIIA '08, 2008, vol. 2, pp. 828–832.
[52] Shibata K, Ikeda Y. Effect of number of hidden neurons on learning in large-scale layered ANN. ICCAS-SICE 2009;2009:5008–13.
[53] Huang G-B, Babri HA. Upper bounds on the number of hidden neurons in feedforward networks with arbitrary bounded nonlinear activation functions. IEEE Trans Neural Netw 1998;9(1):224–9.
[54] R. Hecht-Nielsen. Kolmogorov's mapping neural network existence theorem. Proceedings of the international conference on Neural Networks (Vol. 3, pp. 11-14). New York, NY, USA: IEEE Press.; 1987.
[55] Thomas AJ, Petridis M, Walters SD, Gheytassi SM, Morgan RE. On predicting the optimal number of hidden nodes. In: 2015 International Conference on Computational Science and Computational Intelligence (CSCI). IEEE; 2015. p. 565–70.
[56] C. Faella, V. Piluso, G. Rizzano, G. Structural steel semirigid connections: Theory, design and software. Boca Ratón, Florida. CRC Press, 1999.
[57] Yang J-G, Jeon S-S. Analytical models for the initial stiffness and plastic moment capacity of an unstiffened top and seat angle connection under a shear load. Int J Steel Struct 2009;9(3):195–205.
[58] Li LW. Behavior of semi-rigid beam-to-column minor axis connections in steel frames. Xi'an, China: Xi'an University of Architecture and Technology; 2007. Ph. D. Thesis,(in Chinese.
[59] Reinosa JM, Loureiro A, Gutiérrez R, López M. Experimental and numerical study of angle connections assembled with European profiles. Inf Constr 2014;66:1–6.
[60] Ghindea M, Catarig A, Ballok R. Behavior of beam-to-column joints with angles. Part 1- experimental investigations. J Appl Eng Sci 2015;18:21–8.
[61] Movaghati S, Abdelnaby AE. Experimental study on the nonlinear behavior of bearing-type semi-rigid connections. Eng Struct 2019;199:109609.
[62] Janss J, Jaspart JP, Maquoi R. Experimental study of the non-linear behaviour of beam-to-column bolted joints. In: State-of-the art workshop on connections and the behaviour, strength and design of steel structures. Elsevier Applied Science Publishers; 1988. p. 26–32.
[63] Chasten CP, Fleischman RB, Driscoll Jr GCLW. Top-and-seat-angle connections and end-plate connections: Behavior and strength under monotonic and cyclic loading. Lehigh University; 1989. ATLSS Report 89–05.
[64] Zandonini R, Zanon P. In: Advances in Steel Structures (ICASS '96). Elsevier; 1996. p. 359–64.
[65] Hasan MJ, Al-Deen S, Ashraf M. Behaviour of top-seat double web angle connection produced from austenitic stainless steel. J Constr Steel Res 2019;155:460–79.
[66] Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic minority over-sampling technique. J Artif Intel Res 2002;16:321–57.