



Assessing the accuracy of RC design code predictions through the use of artificial neural networks

Afaq Ahmad¹ · Gregoria Kotsovou¹ · Demitrios M. Cotsovos¹ · Nikos D. Lagaros²

Received: 21 July 2018 / Accepted: 19 September 2018
© The Author(s) 2018

Abstract

In light of recently published work highlighting the incompatibility between the concepts underlying current code specifications and fundamental concrete properties, the work presented herein focuses on assessing the ability of the methods adopted by some of the most widely used codes of practice for the design of reinforced concrete structures to provide predictions concerning load-carrying capacity in agreement with their experimentally established counterparts. A comparative study is carried out between the available experimental data and the predictions obtained from (1) the design codes considered, (2) a published alternative method (the compressive force path method), the development of which is based on assumptions different (if not contradictory) to those adopted by the available design codes, as well as (3) artificial neural networks that have been calibrated based on the available test data (the later data are presented herein in the form of a database). The comparative study reveals that the predictions of the artificial neural networks provide a close fit to the available experimental data. In addition, the predictions of the alternative assessment method are often closer to the available test data compared to their counterparts provided by the design codes considered. This highlights the urgent need to re-assess the assumptions upon which the development of the design codes is based and identify the reasons that trigger the observed divergence between their predictions and the experimentally established values. Finally, it is demonstrated that reducing the incompatibility between the concepts underlying the development of the design methods and the fundamental material properties of concrete improves the effectiveness of these methods to a degree that calibration may eventually become unnecessary.

Keywords Artificial neural network · Design codes · Ultimate limit state · RC beams · Compressive force path method · Physical models

Abbreviations

CFP Compressive force path
ANN Artificial neural network
ULS Ultimate limit state

List of symbols

a_v Shear span
 b Beam width
 d Effective depth
 x Depth of the compressive zone
 A_s Area of tensile reinforcement

A_s Area of compressive reinforcement
 A_{sw} Area of transverse reinforcement
 α_v/d Shear span-to-depth ratio
 f_c Uniaxial compressive strength of concrete
 f_{yl} Longitudinal reinforcement yield stress
 f_{yw} Transverse reinforcement yield stress
 s Spacing between shear links
 ρ_l Ratio of tensile reinforcement ($\rho_l = A_s/b \times d$)
 ρ_t Ratio of compressive reinforcement
($\rho_t = A_s'/b \times d$)
 ρ_w Ratio of transverse reinforcement ($\rho_w = A_{sw}/b \times s$)
 V_c Shear resistance of the RC beam without the contribution of the shear links
 V_s Shear resistance offered by of shear links
 V_u Shear developing along the span of the RC beam at failure
 M_u Bending developing along the span of the RC beam at failure
 M_f Flexural moment capacity of the cross section of the RC beam

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s40091-018-0202-4>) contains supplementary material, which is available to authorized users.

✉ Demitrios M. Cotsovos
d.cotsovos@hw.ac.uk

¹ Heriot-Watt University, Edinburgh, UK

² National Technical University of Athens, Athens, Greece



Introduction

It has recently been shown that the concepts underlying code specifications for the design of reinforced concrete structures (RC) is incompatible with fundamental concrete properties such as post-peak stress–strain characteristics, mechanism of crack extension, and sensitivity of strength and deformation characteristics to the presence of small confinement stresses (acting orthogonal to the direction considered) (Kotsovos 2015). Consequently, the need of calibrating these specifications with experimental data describing the behaviour of RC beam/column elements leads to an ever increasing complexity of the code-adopted design formulae. Moreover, it has been shown that compliance with the fundamental concrete properties can lead to the development of alternative design methods, such as the compressive force path (CFP) method, which may be, not only simpler and more efficient, but also rely on failure criteria derived from first principles (of mechanics), without the need of calibration through the use of experimental data (Kotsovos 2014).

To this end, the present work has been aimed at evaluating the shear design specifications of some of the most widely adopted codes of practice (ACI 445R-99 1999, ACI 318 2014, Eurocode 2 2004, JSCE 2007, CSA 2007, NZ 2006) and the CFP method (Kotsovos 2014) as regards, not only the closeness of the correlation of the calculated values of load-carrying capacity with their experimentally established counterparts, but also the validity of the concepts underlying the development of the method of calculation. The present study focuses on comparing the values of load-carrying capacity predicted by the design codes with those obtained through the use of a purely empirical (soft computing) method that employs artificial neural networks (ANNs) (Cladera and Mari 2004a, b; Mansour et al. 2004; Abdalla et al. 2007; Yang et al. 2007; Jung and Kim 2008) capable of providing predictions concerning load-carrying capacity that correlate closely to the available test data. In what follows the description of the ANN approach developed herein and the presentation and discussion of the results of the comparative study are preceded by a concise description of the methods of calculation employed and the presentation of the experimental information used for the training and verification of the ANNs.

Experimental information on RC structural response at ULS

To date, a large number of experiments have been conducted on simply supported RC beam specimens to study the mechanics underlying RC structural response throughout the loading process and establish the effect of a wide range of parameters [e.g., width (b), effective depth (d), shear

span-to-depth ratio (α_v/d), compressive strength of concrete (f_c), longitudinal (ρ_l) and transverse (ρ_w) steel reinforcement ratios, yield strength of longitudinal (f_{yl}) and transverse (f_{yw}) reinforcement bars, etc.] on the exhibited behaviour (Kani 1964; Panagiotakos and Fardis 2001; Reineck 2013; Reineck et al. 2014; Kotsovos 2014). During testing, attention is focused on measuring certain aspects of the exhibited behaviour such as the applied load and reaction forces, the displacement of certain points along the element span as well as the strain at specific critical locations (i.e., the compressive region or on the steel longitudinal and transverse reinforcement bars). In addition, the deformation profiles and crack patterns forming as well as the exhibited mode of failure are also recorded, as they provide important information concerning the internal stress state developing within the members considered. The test data published and the different parameters considered have been used to form an extensive database describing the effect of the above parameters on the exhibited behaviour mainly at ULS (see Tables A1 and A2 presented as supplementary material).

Design methods

Current design methods are intended to size the cross section of structural members and specify an amount and arrangement of reinforcement that will safeguard desired performance characteristics such as load-carrying capacity and ductility. This objective can only be achieved by ensuring that the load corresponding to shear types of failure is always higher than that corresponding to flexural capacity. To this end, the sizing of the cross section and the assessment of the longitudinal reinforcement are linked to flexural capacity, whereas the associated shear-force diagram underlies the method adopted for specifying an amount and arrangement of transverse reinforcement that will safeguard against shear failure occurring before flexural capacity is exhausted.

Code-adopted methods

The concepts underlying the methods adopted by the majority of the available RC design codes assume that load transfer at ULS is accomplished through the development of truss or various forms of strut-and-tie mechanisms (a simplified form of which is depicted in Fig. 1) in which the compressive zone and the tension flexural reinforcement form the longitudinal struts and ties, respectively, stirrups and inclined bars form the web ties and cracked concrete in the tensile zone (by means of aggregate interlock and dowel action) allows for the formation of inclined struts. On the basis of these concepts, shear failure is associated with failure of



either the web ties or the inclined struts, whereas the longitudinal struts and ties are designed such that flexural failure occurs when the concrete strut at the most critical location reaches its strength (in compression) after yielding of the longitudinal tie at the same cross section. To determine shear capacity each code employs its own empirical formulae (see Table 1), the derivation of which is largely based on regression analysis of the available test data, whereas the calculation of flexural capacity is based on assumptions such as plane cross sections remain plane during bending, full bond between concrete and steel, uniaxial stress–strain behaviour of concrete in the compressive zone, etc.

The CFP method

The CFP method is based on the incorporation of important mechanical characteristics of concrete such as brittle behaviour, cracking mechanism, sensitivity to triaxial stress conditions, etc. into beam theory which describes the mechanism of load transfer accomplished through the bending of the concrete cantilevers forming between consecutive flexural and/or inclined cracks (see Fig. 2). Failure is considered to be associated with the stress conditions in the compressive zone and reinforcement is specified to prevent non-flexural (commonly referred to as shear) types of failure before flexural capacity is exhausted. Unlike the code methods, the criteria for non-flexural types of failure have been derived from first principles without the need of calibration through the use of test data. The method is fully described elsewhere (Kotsovos 2014) and, therefore, is only briefly presented in what follows.

The CFP theory attributes what is commonly referred to as ‘shear’ failure to the development of tensile stresses at particular locations [dependent on the shear span-to-depth ratio

(a_v/d)] along the compressive zone. For an RC beam/column element with flexural reinforcement only, the values of the shear force corresponding to such types of failure are obtained from.

For $a_v/d > 2.5$, failure occurs when either the shear force at a distance equal to $2.5d$ from a simple support or a point of contra-flexure attains a critical value:

$$V_{II,1} = 0.5f_tbd, \tag{1}$$

where f_t is the tensile strength of concrete and b, d are the width and depth of the cross section, or after yielding of the flexural reinforcement in regions, where the value of the shear force attains a critical value:

$$V_{II,2} = F_c[1 - 1/(1 + 5|f_t|/f_c)], \tag{2}$$

where F_c is the value of the compressive force developing on account of bending when M_f is attained and f_c and f_t are the values of the compressive and tensile, respectively, strengths of concrete.

For $1 \leq a_v/d \leq 2.5$, failure is considered to occur when the bending moment at the cross section located at a distance equal to a_v from the closest simple support or point of contra-flexure exceeds a critical value:

$$M_{III} = M_{II} + (M_f - M_{II})(2.5 - a_v)/(1.5d), \tag{3}$$

obtained by linear interpolation between the values $M_{II} = V_{II}(2.5d)$ and $M_{IV} = M_f$ (flexural capacity) at cross sections at distances of $2.5d$ and d , respectively, from the closest simple support or point of contra-flexure.

For $a_v/d < 1$, failure occurs when the shear force exceeds a critical value provided by Eq. (4):

$$V_{IV} = M_f/a_v. \tag{4}$$

Fig. 1 Code design methods: **a** truss model of a beam-like RC element; **b** portion of truss in Fig. 1 between cuts 1–1 and 2–2 indicating the mechanism of load transfer

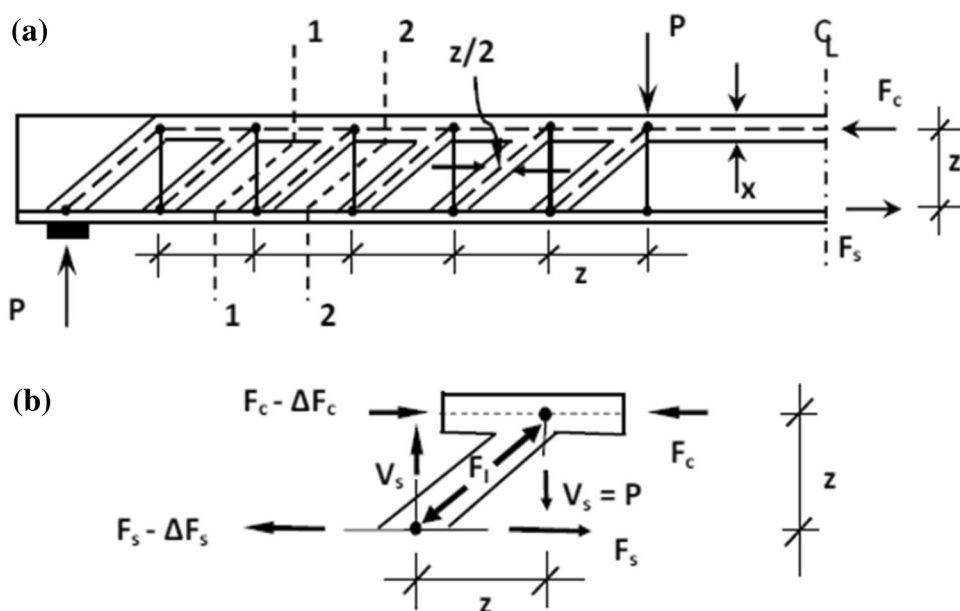


Table 1 Formulae adopted by the RC design codes for predicting shear capacity

Physical models	Contribution of the specimen cross section without stirrups V_c	Contribution of the stirrups V_s	Limitations	Factors used for multiplying shear capacity for the case of small shear spans
ACI-318-11 (SI units)	$V_c = \left(0.16f_c^{0.5} + 17 \frac{A_s \times V_u}{M_u}\right) bd$	$V_{s,45^\circ} = \frac{A_{sw} f_{yw} d}{s}$	$V_c \leq 0.29 \sqrt{f_c} bd$ $d \leq 500$ m $f_y \leq 420$ MPa $f_c \leq 70$ MPa	If $[3.5-2.5 (\alpha_v/d)] \geq 2.5$ then use 2.5 If $[3.5-2.5 (\alpha_v/d)] < 2.5$ and ≥ 1.0 then use 1.0 If $[3.5-2.5 (\alpha_v/d)] < 1.0$ then use $[3.5-2.5 (\alpha_v/d)]$
EC2 ^a (SI units)	$V_c = 0.18 k (100 \rho_v f_c)^{1/3} bd$	$V_s = \frac{A_{sw} f_{yw} 0.9d \cot \theta}{s}$, where $1 \leq \cot \theta \leq 2.5$	$V_c \leq v_{\min} b \times d$ $v_{\min} = 0.035 f_c^{1/2} k^{3/2}$ $\rho_1 = \frac{A_s}{bd} \leq 0.02$ $k = (1 + \sqrt{200/d}) \leq 2$ $f_{ck} < 90$ MPa	If $(\alpha_v/d) \leq 2.0$ then use $2d/\alpha_v$
JSCE (SI units)	$V_c = \beta_d \beta_p \beta_n \beta_{vcd} bd / \gamma_b$	$V_s = \frac{A_{sw} f_{yw} (\sin \alpha + \cos \alpha) z}{s \gamma_b}$ $z = d/1.15$	$\beta_p = (100 \rho_1)^{1/3} \leq 1.5$ $0 \leq \beta_n \leq 2.0$ $\beta_d = (1000/d)^{1/4} \leq 1.5$ $f_{vcd} = 0.20 \times f_c^{1/3} \leq 0.72$ $f_y \leq 400$ Mpa $f_c \leq 60$ MPa	No change required (see Section 9.2.2.2 of the relevant code)
KBCS-(SI units)	$V_c = \alpha_d \left(0.16f_c^{0.5} + 17.6 \rho_1 \frac{V_u d}{M_u}\right) bd$	$V_s = f_{yw} \rho_w bd$	$V_c \leq 0.29 \sqrt{f_c} bd$	$\alpha_d = [3.5 - 2.5(\alpha_v/d)] \leq 2.5$
CSA (SI units)	$V_c = \Phi_c \lambda \beta f_c^{0.5} bd$	$V_s = \frac{\Phi_s A_{sw} f_{yw} d \cot \theta}{s}$	$V_{r,max} \leq 0.25 \phi_c \times f_c bd$ $f_y \leq 400$ MPa $f_c \leq 60$ MPa	No change required (see Section 11.3.2 of the relevant code)
NZ (SI units)	$V_c = v_c A_{cv}, A_{cv} = bd$ $v_c = k_d k_a v_b, k_a = 1.0$ $v_b = \min[(0.07 + \rho_w) \sqrt{f_c}, 0.2 \sqrt{f_c}]$	$V_{s'} = \frac{A_{sw} f_{yw} d}{s}$	For $d \leq 400$ m $k_d = 1.0$ for $d > 400$ mm $k_d = (400/d)^{0.25}$ $f_y \leq 500$ Mpa $f_c \leq 50$ Mpa	No change required (see Section 9.3.9.3.1 of the relevant code)

^aFor design purposes, EC2 ignores the contribution of V_c , relying only in the contribution of V_s when assessing the shear capacity

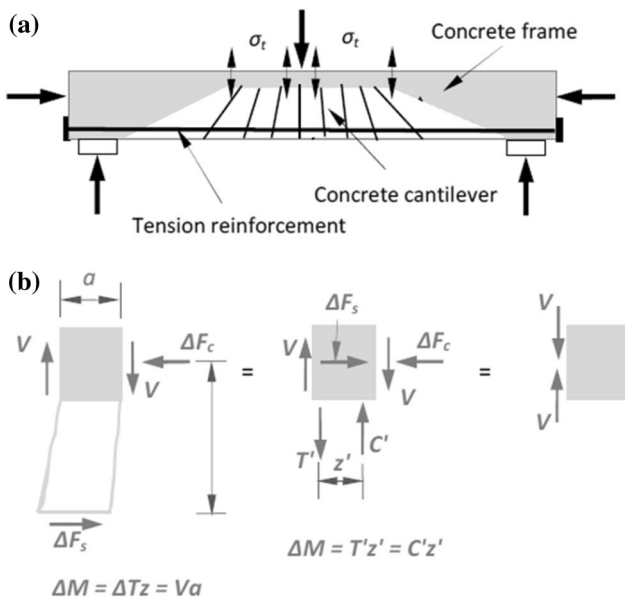


Fig. 2 CFP method: **a** schematic representation of the physical state of an RC beam at the ULS; **b** mechanism of load transfer

When the above types of failure occur before flexural capacity is attained, they can be prevented by specifying transverse reinforcement in the form of stirrups in an amount sufficient to undertake the total tension developing at the locations indicated earlier as critical. More specifically, for $\alpha_v/d > 2.5$, the transverse tension developing at a distance of $2.5d$ from the location of the simply support or the point of contra-flexure has been shown to be numerically equal to the shear force V_f developing at this location and, then, the amount of transverse reinforcement required to prevent failure is obtained from:

$$A_{sv} = V_f / f_{yv} \tag{5}$$

Such reinforcement is spread over a length of $2d$ extending symmetrically on either side of the location, where transverse tension is considered to develop (i.e., the location at a distance from a simple support or a point of contra-flexure).

This method is also used to calculate the amount of transverse reinforcement required to prevent failure in the region of points of contra-flexure.

When $V_f > V_{II,2}$ (where V_f is the shear force corresponding to M_f), the transverse tensile stresses σ_t exceed the tensile strength of concrete f_t in the compressive zone of the regions of beam/column elements, where yielding of the flexural

reinforcement occurs. Then, these tensile stresses can be calculated from Eq. (2) by replacing $V_{II,2}$ with V_f and f_t with σ_t , and solving for σ_t , that is

$$|\sigma_t| = f_c / [5(F_c / V_f - 1)], \tag{6}$$

and hence, the tensile force per unit length developing in this region is expressed by:

$$T_{II,2} = \sigma_t b / 2. \tag{7}$$

The second term of Eq. (7) is divided by 2 to allow for the non-uniform distribution of the transverse stresses within the compressive zone.

Therefore, the amount of transverse reinforcement required per unit length to sustain $T_{II,2}$ is equal to:

$$A_{sv,II2} = T_{II,2} / f_{yv}. \tag{8}$$

As regards the assessment of the transverse reinforcement to prevent failure for the case of $1 \leq a_v/d \leq 2.5$, this is obtained from:

$$A_{sv,III} = 2(M_f - M_{III}) / (a_v f_{yv}). \tag{9}$$

It should be noted that this reinforcement is uniformly distributed within the shear span and, in accordance with the CFP theory, is intended to increase flexural capacity from M_{III} to M_f rather than prevent shear failure.

It is also important to note that in accordance with the CFP theory, the specified transverse reinforce is effective when its spacing is smaller than $d/2$.

As regards the calculation of flexural capacity, this differs from the EC2 code method in that the normal stress distribution in the compressive zone is equivalent to a stress block covering the whole compressive zone; its intensity is obtained from:

$$\sigma_a = f_c + 5|f_t|. \tag{10}$$

The above expression has been derived by considering the triaxial stress conditions invariable developing in the compressive zone when ULS is approached.

Artificial neural networks

Artificial neural networks mimic their biological counterparts in the nervous system and the brains of animals and humans. They are used to estimate or approximate

functions that depend on a large number of parameters, the effect of which is not clearly established or quantified. Due to their adaptive nature and ability to remember information introduced to them during training (calibration), ANNs can learn, generalize, categorize, and empirically predict values.

A typical artificial neuron (AN) designed to perform specific operations is schematically depicted in Fig. 3. From the figure, it can be seen that the AN is fed with the values of parameters x_j (describing selected parameters associated with the design details and strength characteristics of the structural elements investigated), each of them being attributed an (initially assumed) weight w_{ji} and a bias θ_i . This input information is used to form the sum n_i of the products $x_j \times w_{ji}$ and θ_i , that is

$$n_i = \sum w_{ji} x_j + \theta_i. \tag{11}$$

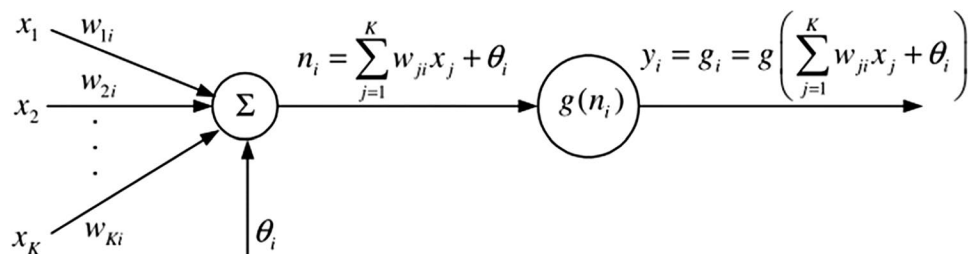
The sum n_i is inserted as argument into an adopted *activation* function g , that is

$$y_i = g_i = g(\sum w_{ji} x_j + \theta_i), \tag{12}$$

and the resulting value of g , $y_i = g(n_i)$ is the output O_i of the AN.

ANs, such as that of Fig. 3, have been combined to form the artificial neural networks (ANNs) adopted for the present work, as depicted in Fig. 4. The process through which the ANNs have been developed is presented in a flowchart form in Fig. 5. The architecture of the ANNs models, which is associated with the selection of the type of activation function employed for each layer, the number of hidden layers and the number of neurons in each hidden layer have been determined through a trial-and-error process due to the absence of relevant guidelines. This trial-and-error process indicated that the use of two hidden layers comprising nine neurons each produces a satisfactory solution for the purposes of the present work, whereas the types of activation functions adopted are provided in Table 2 which also shows the layers, where the functions were incorporated. The values of the weights (w_{kj}) and bias (θ_j) are initially randomly assigned, but obtain their final values through an iterative training (calibration) process

Fig. 3 Typical AN



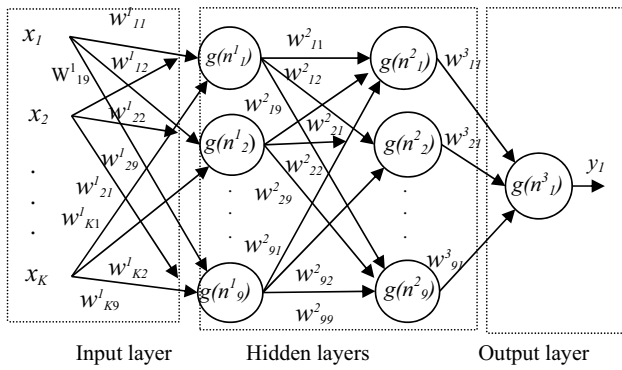


Fig. 4 Architecture of ANN adopted

(described later) which aims at reducing to an acceptable value the difference between the output values O_i (obtained from the ANN) and the target values T_i (provided in the

database) (Svozil et al. 1997; Basheer and Hajmeer 2000). A measure of the deviation of O_i from T_i is obtained from:

$$E = (1/2) \sum (T_i - O_i)^2. \tag{13}$$

Convergence is considered to be achieved when E becomes smaller than an acceptable value.

Experimental database

In the present study, two databases (DB) are formed which are presented as supplementary material at the end of this paper: DB-I, containing test data obtained from tests on 609 on RC beam specimens without stirrups (BWOS) (see Table A1 presented as supplementary material) and DB-II,

Fig. 5 Flowchart describing the process through which ANNs are developed

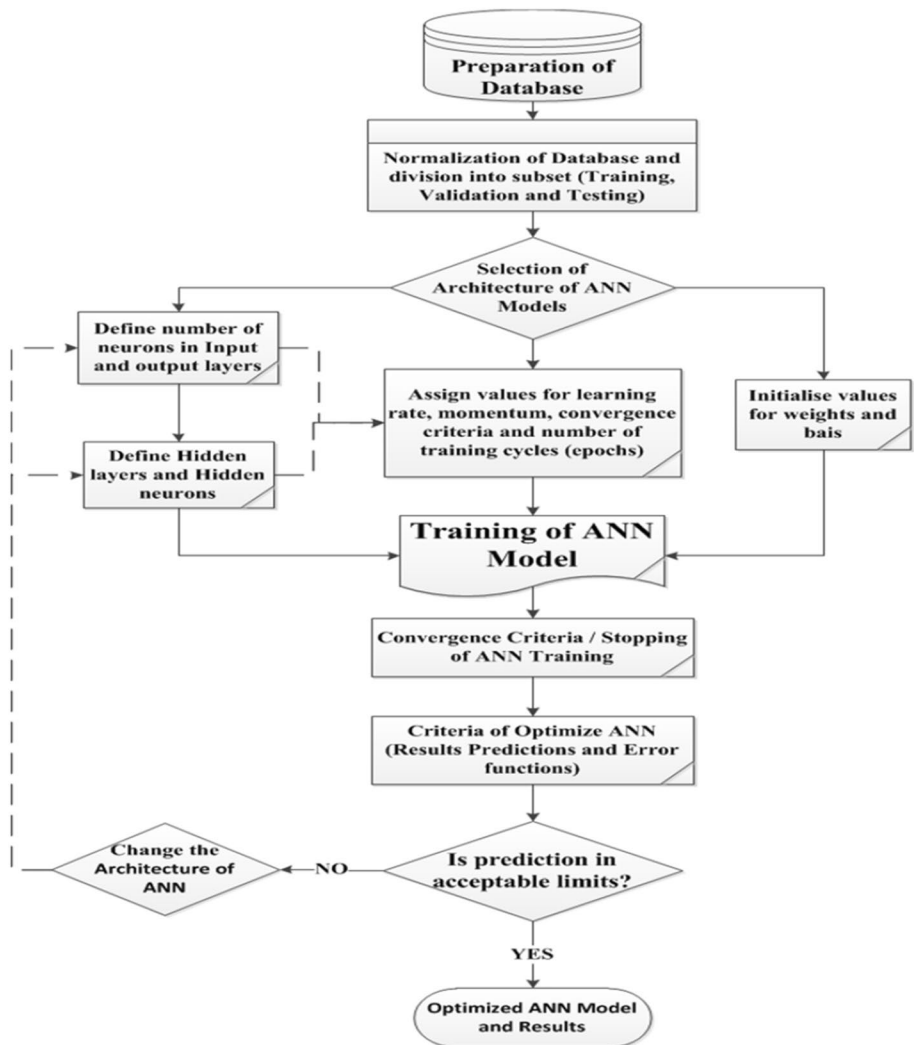
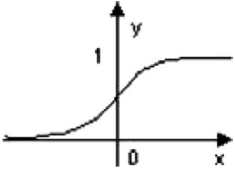
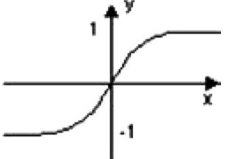


Table 2 Activation functions used in the ANN models

Activation functions	Layers	Mathematical expression	Range of values of x	Graphical expression
Sigmoid	B/w input layer and first hidden layer B/w first hidden layer and second hidden layer	$f(x) = \frac{1}{1+e^{-x}}$	(0, +1)	
Hyperbolic (tanh)	B/w second hidden layer and out put layer	$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$	(-1, +1)	

containing test data obtained from tests on 306 RC beam specimens with stirrups (BWS) (see Table A2 presented as supplementary material). Both databases include parameters associated with the design of the RC beams as well as aspects of RC structural response at ULS (load-bearing capacity and mode of failure). All specimens are simply supported and subjected to 3 or 4 points bending tests. Table 3 provides statistical information, essential for the development of the ANNs, on the variation of certain key parameters associated with the design details and load-carrying capacity of the specimens in DB-I and DB-II.

Selection of input parameters

The relationship between two parameters (in the databases) is expressed through the Pearson’s correlation coefficient (r) described by Eq. (14) (Mojtaba et al. (2013)):

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \sqrt{n(\sum y^2) - (\sum y)^2}}, \tag{14}$$

where n represents the total number of samples and x, y are the two variables considered.

Tables 4 and 5 describe the correlation between the parameters considered by the ANN employed. The linear

Table 3 Statistical information concerning the test data included in DB-I AND DB-II

		Min	Max	Diff	Avg.	St. dev	COV
For beams without stirrups (BWOS) DB-I							
b	(mm)	50	500	450	171.51	68.71	0.4
d	(mm)	65	483	418	255.15	73.6	0.29
$\alpha\sqrt{d}$	0	0.98	9.73	8.75	3.53	1.4	0.4
ρ_1	(%)	0.250261	7.462626	7.212365	2.14	1.09	0.51
f_{yl}	(MPa)	283	1779	1496	444.83	189.15	0.43
f_c	(MPa)	12.2	110.9	98.7	39.07	21.43	0.55
M_f	(kN mm)	2000	844,000	842,000	101,307	109,891.2	1.08
V_u	(kN)	7	360	353	69.09	47.07	0.68
For beams with stirrups (BWS) DB-II							
b	(mm)	100	510	410	207.27	67.64	0.33
d	(mm)	113	975	862	342.87	157.03	0.46
$a\sqrt{d}$		1	7.002188	6.002188	3.43	1.35	0.39
ρ_1	(%)	0.18	5.57	5.39	2.35	1.05	0.45
f_{yl}	(MPa)	250	888.48	638.48	411.33	77.88	0.19
f_c	(MPa)	13.8	126.2	112.4	44.51	22.41	0.5
ρ_w	(%)	0.08	2.24	2.16	0.57	0.5	0.88
f_{yw}	(MPa)	224.08	875.5	651.42	399.4	112.62	0.28
M_f	(kN mm)	2648	1,738,423	1,735,775	283,484.8	378,729.4	1.34
V_u	(kN)	4.0967	760.1942	756.0975	185.54	134.25	0.72

correlation coefficient (r) can obtain values between 1 and -1 ($1 > r > -1$). Negative values of r describe an inverse relation between the two associated parameters (i.e., an increase in the value of one parameter results in the decrease in the value of the other). Positive values of r have a similar effect on the values of the associated parameters (i.e., an increase in the value of one parameter results in an increase of the other). Higher absolute values of $|r|$ indicate a more pronounced relation between the two parameters considered. Depending on r and on the basis of current knowledge on RC structural response at ULS, different combinations of input parameters are selected for the case of the beams without and with stirrups, as shown in Tables 6 and 7.

Normalization of database

Expressing the parameters considered in the database in a non-dimensional form is of particular significance for the calibration and overall function of the ANNs, since this improves their ability to provide accurate predictions. To avoid problems associated with low learning rates of the ANNs, it is more effective to further normalize the values of the parameters between appropriate upper and lower limiting values. Furthermore, to compensate for the fact that purely experimental databases are characterized by regions (often located at their boundaries) which may not be as densely populated with data as other regions, it is considered more

effective to normalize the values of the parameters considered between (0.1, 0.9) instead of (0, 1), through the use of the following expression (Krogh and Vedelsby 1995; Utans et al. 1995; Castellano and Fanelli 2000):

$$X = \left(\frac{0.8}{x_{\max} - x_{\min}} \right) x + \left[0.9 - \left(\frac{0.8}{x_{\max} - x_{\min}} \right) x_{\max} \right], \tag{15}$$

where x_{\max} and x_{\min} represent the minimum and maximum values of a specific parameter included in the database considered.

Training process

The training of the ANNs is achieved through the multi-layer free-forward back-propagation (MLFFBP) process (Beale et al. 2015) (a schematic representation of which is provided in Fig. 6) and the provided databases DB-I and DB-II. The MLFFBP process consists of two phases: (1) the free-forward phase (input signal) and (2) the back-propagation phase (error signal). In the case of the free-forward phase, the information provided is processed from the input layer towards the output layer, where the predicted solution of the problem considered is obtained. If convergence is not achieved [i.e., the difference between the output (predicted) and target (database) values is expressed by E as defined by Eq. (13)] is larger than a pre-defined acceptable value), then the values of the weights and

Table 4 Values of r expressing the effect between parameters considered in the database for the case of beams without stirrups

	b	d	$a\sqrt{d}$	ρ_l	f_{yl}	f_c	M_f	V_u
b	1.00							
d	0.39	1.00						
$a\sqrt{d}$	-0.02	0.04	1.00					
ρ_l	-0.14	-0.16	0.23	1.00				
f_{yl}	0.09	0.03	0.02	-0.20	1.00			
f_c	0.20	-0.08	-0.09	0.25	0.09	1.00		
M_f	0.55	0.68	0.11	0.23	0.09	0.22	1.00	
V_u	0.56	0.48	-0.30	0.17	-0.04	0.30	0.67	1.00

Table 5 Values of r expressing the effect between parameters considered in the database for the case of beams with stirrups

	b	d	$a\sqrt{d}$	ρ_l	f_{yl}	f_c	ρ_w	f_{yw}	M_f	V_u
b	1									
d	0.71	1								
$a\sqrt{d}$	0.12	-0.10	1							
ρ_l	-0.17	0.08	-0.11	1						
f_{yl}	0.07	-0.24	0.57	-0.45	1					
f_c	-0.07	-0.33	0.36	0.31	0.28	1				
ρ_w	-0.23	-0.36	-0.37	0.10	-0.07	0.17	1			
f_{yw}	0.09	-0.36	0.21	-0.25	0.36	0.29	0.20	1		
M_f	0.71	0.74	0.15	0.33	-0.03	0.05	-0.14	-0.10	1	
V_u	0.26	0.46	-0.62	0.29	-0.60	-0.23	0.25	-0.26	0.33	1

Table 6 Architecture of ANN models

Sr. no	Model	Input	No of input neurons	Act. function b/w first two layers	Act. function in last layers	No of neurons in each hidden layer
For beams without stirrups (BWOS) DB-I						
1	BWOS-1	$b, d, a/d, \rho_l, f_{yl}, f_c$	6	Sigmoid	tanghn	12–12
2	BWOS-2	$b, d, a/d, M_f, f_c$	5	Sigmoid	tanghn	10–10
3	BWOS-3	$b/d, a/d, M_f/f_c b d^2$	3	Sigmoid	tanghn	6–6
4	BWOS-4	$d, a/d, M_f/f_c b d^2, f_c$	4	Sigmoid	tanghn	8–8
5	BWOS-5	$d, b/d, a/d, M_f/f_c b d^2$	4	Sigmoid	tanghn	8–8
6	BWOS-6	$d, b/d, a/d, M_f/f_c b d^2, f_c$	5	Sigmoid	tanghn	10–10
For beams with stirrups (BWS) DB-II						
1	BWS-1	$b, d, a/d, \rho_l, f_{yl}, f_c, \rho_w, f_{yw}$	8	Sigmoid	tanghn	16–16
2	BWS-2	$b, d, a/d, M_f, f_c, \rho_w, f_{yw}$	7	Sigmoid	tanghn	14–14
3	BWS-3	$b/d, a/d, M_f/f_c b d^2, \rho_w, f_{yw}$	5	Sigmoid	tanghn	10–10
4	BWS-4	$b/d, a/d, \rho_w/\rho_l, f_c/f_{yw}$	4	Sigmoid	tanghn	8–8
5	BWS-5	$d, a/d, M_f/b d^2, f_c, \rho_w f_{yw}$	5	Sigmoid	tanghn	10–10
6	BWS-6	$d, b/d, a/d, M_f/f_c b d^2, \rho_w, f_{yw}$	6	Sigmoid	tanghn	12–12
7	BWS-7	$d, b/d, a/d, M_f/f_c b d^2, f_c, \rho_w, f_{yw}$	7	Sigmoid	tanghn	14–14
8	BWS-8	$d, b/d, a/d, M_f/f_c b d^2, f_c, \rho_w f_{yw}$	6	Sigmoid	tanghn	16–16

Table 7 Results of different ANN models

V_u		Min	Max	Diff	Avg.	St. dev	COV
For beams without stirrups (BWOS) DB-I							
Exp	(kN)	7	360	353	69.09	47.07	0.68
BWOS-1	(kN)	5	307	302	69.53	46.61	0.67
BWOS-2	(kN)	0	300	300	69.57	45.75	0.66
BWOS-3	(kN)	6	394	388	78.95	63.04	0.8
BWOS-4	(kN)	7	371	364	69.29	46.14	0.67
BWOS-5	(kN)	7	319	312	70.72	50.48	0.71
BWOS-6	(kN)	7	306	299	68.38	44.46	0.65
For beams with stirrups (BWS) DB-II							
Exp	(kN)	4	761	757	185.54	134.25	0.72
BWS-1	(kN)	3	757	754	186.41	134.36	0.72
BWS-2	(kN)	13	780	767	184.99	130.65	0.71
BWS-3	(kN)	0	767	767	185.34	137.65	0.74
BWS-4	(kN)	4	656	652	184.5	133.08	0.72
BWS-5	(kN)	2	763	761	186.74	137.71	0.74
BWS-6	(kN)	7	846	839	187.31	137.19	0.73
BWS-7	(kN)	4	863	859	186.9	137.83	0.74
BWS-8	(kN)	3	834	831	187.31	141.15	0.75

bias are re-adjusted through the “gradient descent” method in which the change in weights (Δw_{kj}) is expressed in the form:

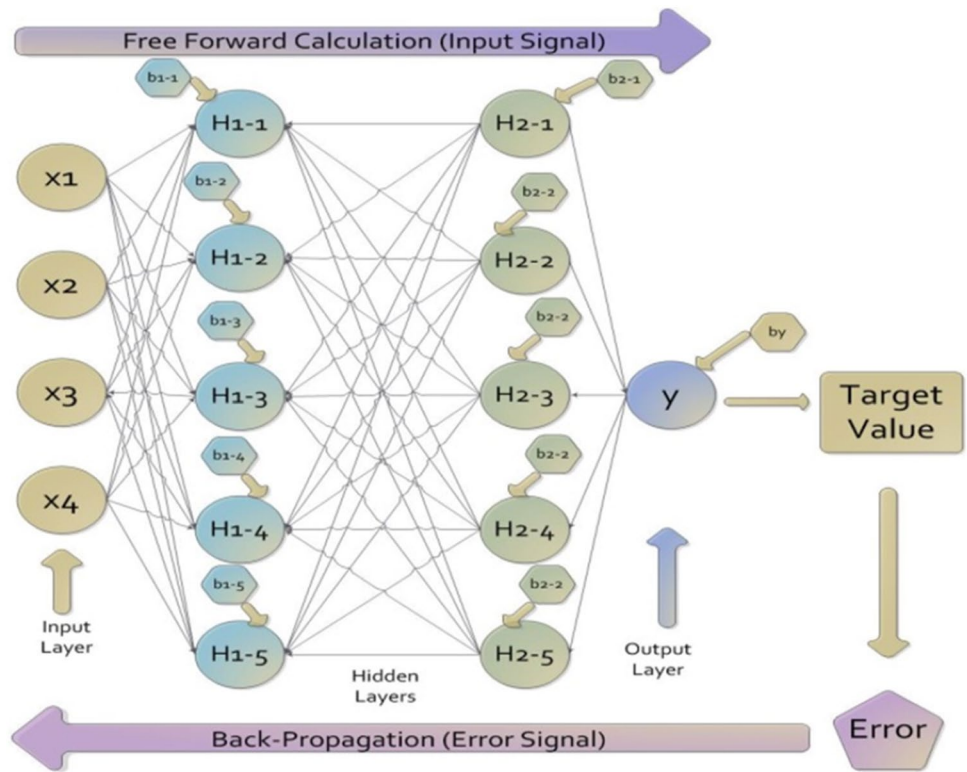
$$\Delta w_{kj} = -\alpha \frac{\partial E}{\partial w_{kj}}, \tag{16}$$

where a is the learning rate.

High values of learning rate (α) result in the values of the weights (w_{jk}) changing more drastically (large values of Δw_{jk})

during each iteration. This can result in the training process not converging to the most optimum combination of values of weights for which the error attains its minimum value for the whole database considered (global minima). Small values of α result in small changes to the weights (w_{jk}) (i.e., small values of Δw_{jk}) for each iteration. In the latter case, although more training time is required (due to the larger number of iterations that have to be carried out to satisfy the convergence criteria), the iterative process is able to identify more effectively the

Fig. 6 Representation of a multi-layer feed forward NN (MLFNN)



optimum combination of values w_{jk} for which the error function obtains its minimum value for the whole database considered allowing the ANN to provide more accurate predictions. To facilitate this process, the values of Δw_{kj} calculated for each iteration ($n > 1$) are provided as a function of the corresponding adjustment carried out in the previous iteration ($n - 1$) and the momentum factor ($0 < \eta < 1$) as follows:

$$\Delta w_{kj}^n = \alpha \times y_k + \eta \times \Delta w_{kj}^{n-1}. \tag{17}$$

The selection of the most accurate of the ANN models in Table 7 is based on assessing the correlation factor (R), the mean squared error (MSE), and the mean absolute error (MAE) which are analytically expressed by Krogh and Vedelsby (1995), Utans et al. (1995) and Castellano and Fanelli (2000):

$$R = \frac{\sum_{i=1}^n (T_i - \bar{T})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^n (T_i - \bar{T})^2 \sum_{i=1}^n (O_i - \bar{O})^2}}, \tag{18a}$$

$$MSE = \frac{\sum_{i=1}^n (T_i^2 - O_i^2)}{n}, \tag{18b}$$

$$MAE = \frac{\sum_{i=1}^n |T_i - O_i|}{n}, \tag{18c}$$

where $\bar{T} = \sum_1^n T_i/n$ and $\bar{O} = \sum_1^n O_i/n$ are the mean T_i and O_i values, respectively, and n is the total number of samples in databases, i.e., DB-I and DB-II (see Tables A1 and A2 which provided in the form of supplementary material).

During the calibration process, the normalized data are divided into three sub-sets. Each of these sub-sets is employed for training, validation, and testing purposes. Matlab (Beale et al. 2015) is used to develop the ANN models and to randomly divide the database into the three sub-sets: 60% of the data for training, 20% for validation, and 20% for testing (Krogh and Vedelsby 1995; Utans et al. 1995; Castellano and Fanelli 2000). Each ANN model is trained until one of the following conditions is met as proposed by the Levenberg–Marquardt back-propagation method (Beale et al. 2015):

1. For 100 iterations (epochs) including the full MLFFBP process.
2. Maximum of 6 validation failures are exhibited. Validation failure occurs when the performance of the ANN (assessed during each iteration) fails to improve or remains constant.
3. The values of performance goal indicators, expressing the difference between the target and output values (see Eqs. 18a–18c) attain small values (e.g., 10^{-6}), indicating that convergence has been achieved.

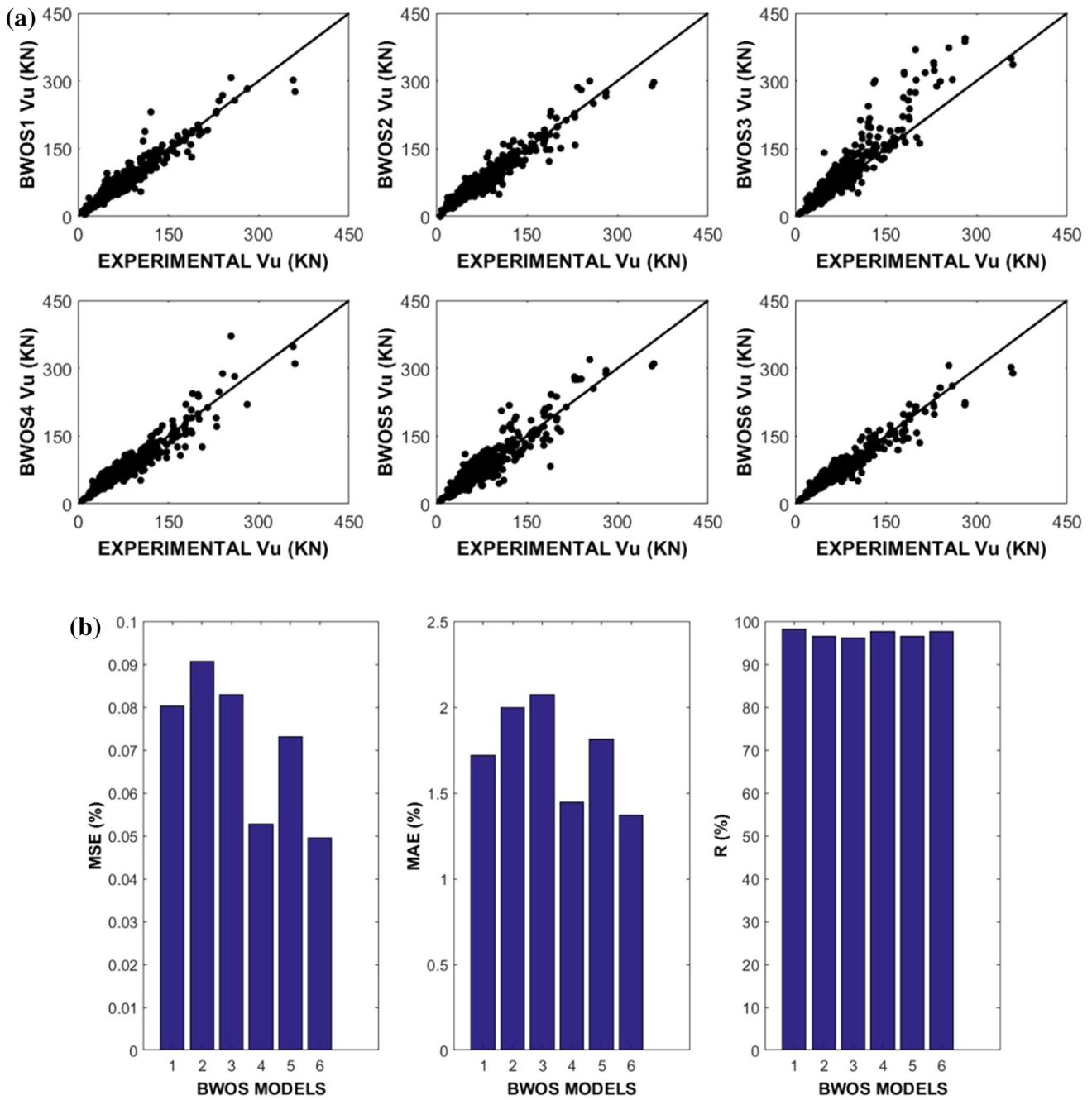


Fig. 7 a Predictions obtained from different ANNs for the case of RC beam specimens without stirrups (BWOS). b Performance exhibited by ANNs models for the RC beams without stirrups (BWOS)

4. Minimum performance gradient (related to the rate at which the weights are adjusted through the MLFFBP process) becomes 10^{-10} .

A number of ANN models have been presently considered employing different parameters, activation functions, and number of hidden neurons and layers (see Table 6) to study their effect on the ability of the subject models to

provide accurate predictions. The values of the input and output parameters are normalized between 0.1 and 0.9 at the start of the training process, whereas the values of the weight and biases vary between -1.0 and 1.0 , at the end of the training process. The previous work (Ahmad et al. 2016) has revealed that in such cases, the use of (1) the sigmoid activation function between the first three layers of the ANN (e.g., the input and the two hidden layers) and (2)

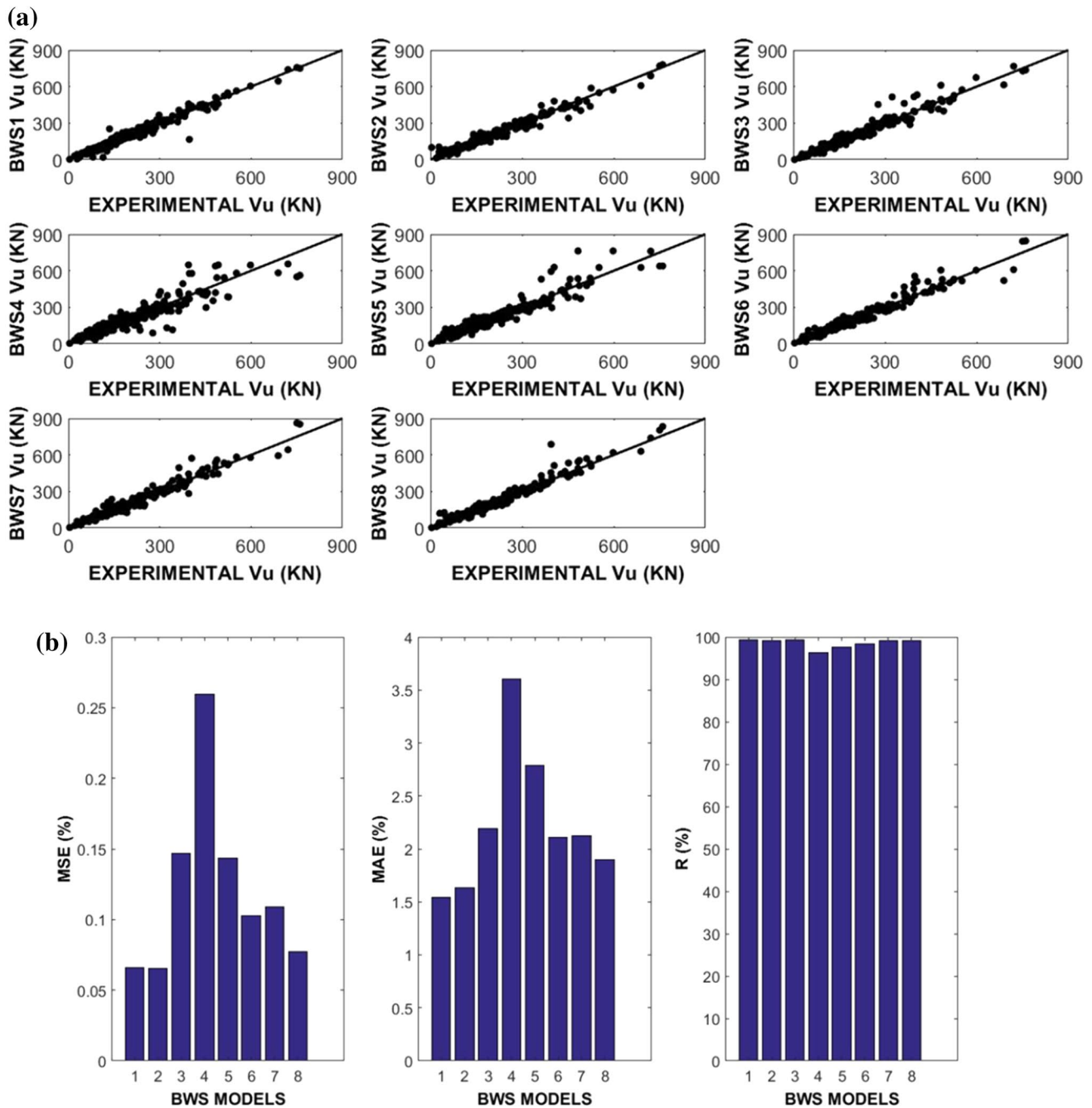


Fig. 8 a Predictions obtained from different ANNs for the case of RC beam specimens with stirrups (BWS). b Performance exhibited by ANNs models for the RC beams with stirrups (BWS)

the hyperbolic (\tanh) activation function between the second hidden layer and the output layer allows the ANN to provide accurate predictions (see Tables 2 and 6).

A total of six different ANN models are developed for the case of the RC beams without stirrups (BWOS) and eight ANN models for the case of the RC beams with stirrup (BWS) (as described in Table 6). Table 7 presents statistical information concerning the level of correlation

achieved between the predicted values of load-carrying capacity obtained from the ANN models and their experimentally established counterparts. Figures 7 and 8 show the same statistical information in the form of graphs. The ANN model considered to provide the most accurate predictions is that exhibiting the highest value of R (correlation function) in combination with the lowest values of MSE and MAE (Basheer and Hajmeer 2000). Figure 7b provides



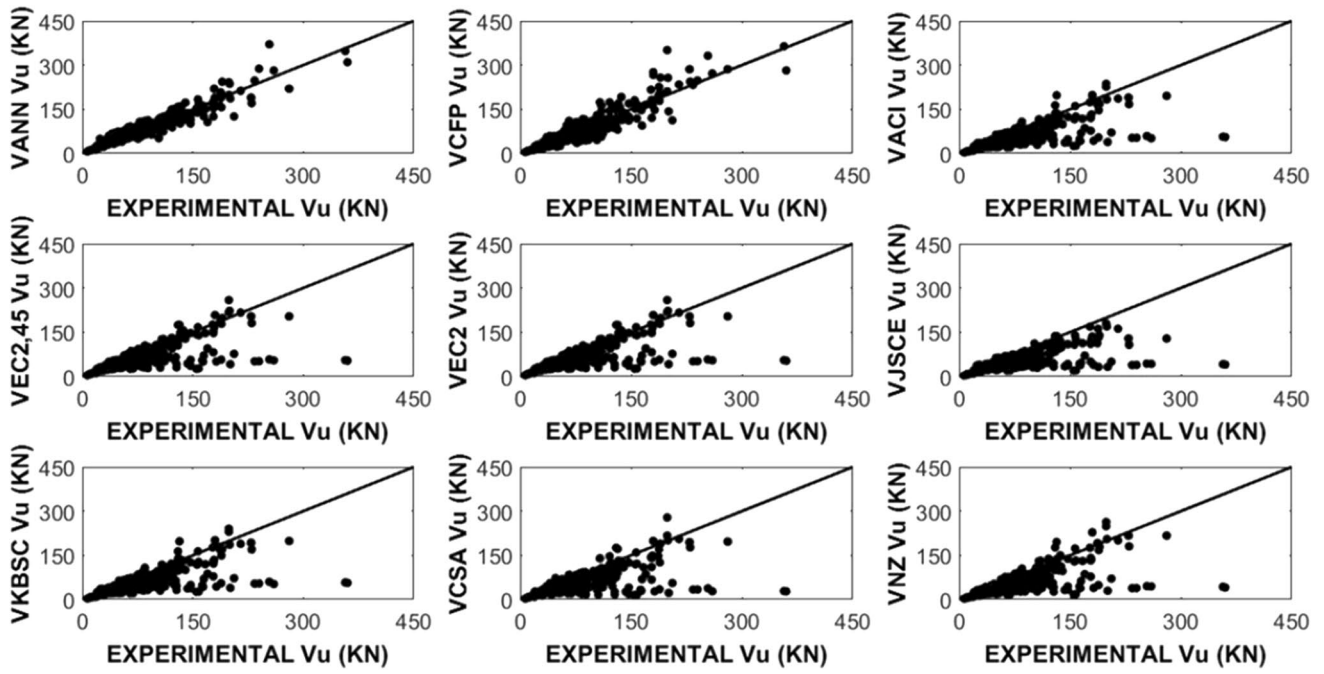
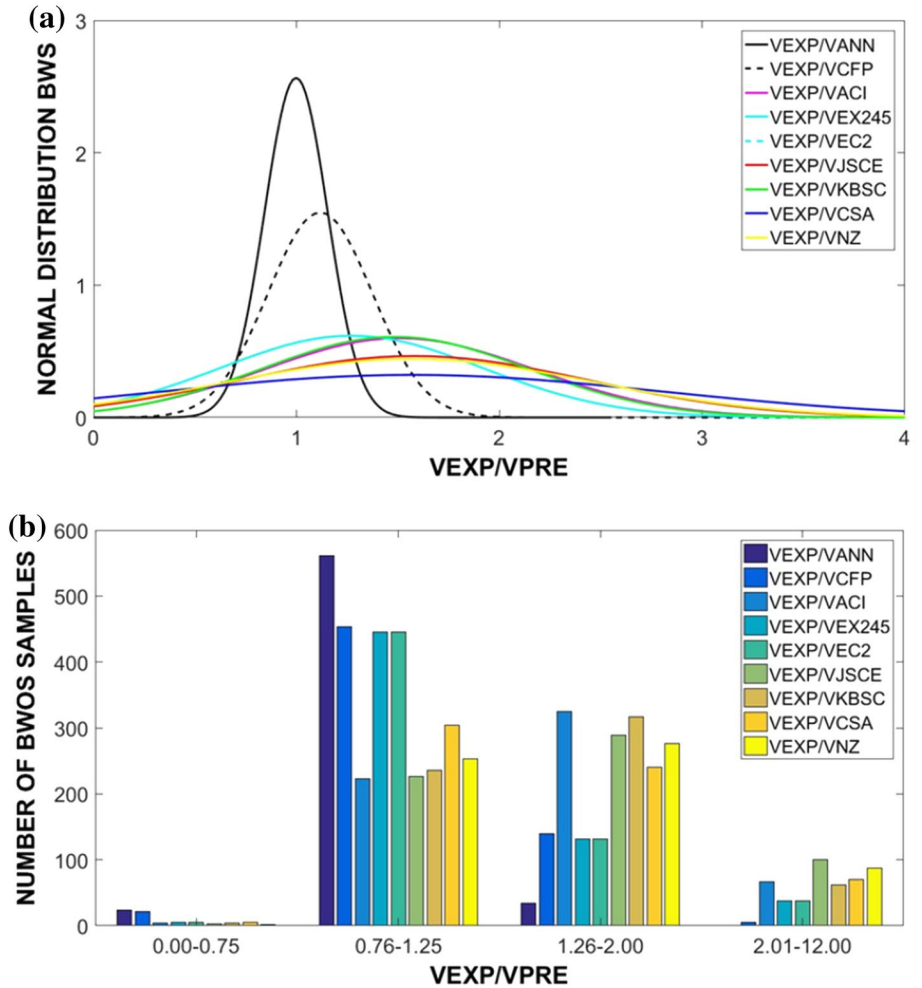


Fig. 9 Comparison between the predictions obtained from the various assessment methods employed and the ANN with their experimental counterparts for the case of beams without stirrups (BWOS)

Fig. 10 a Curves describing the normal distribution of V_{EXP}/V_{PRED} for beams without stirrups (BWOS) based on the predictions obtained from the various assessment methods employed and the use of ANNs. **b** Distribution V_{EXP}/V_{PRED} for beams without stirrups (BWOS) based on the predictions obtained from the various assessment methods employed and the use of ANNs



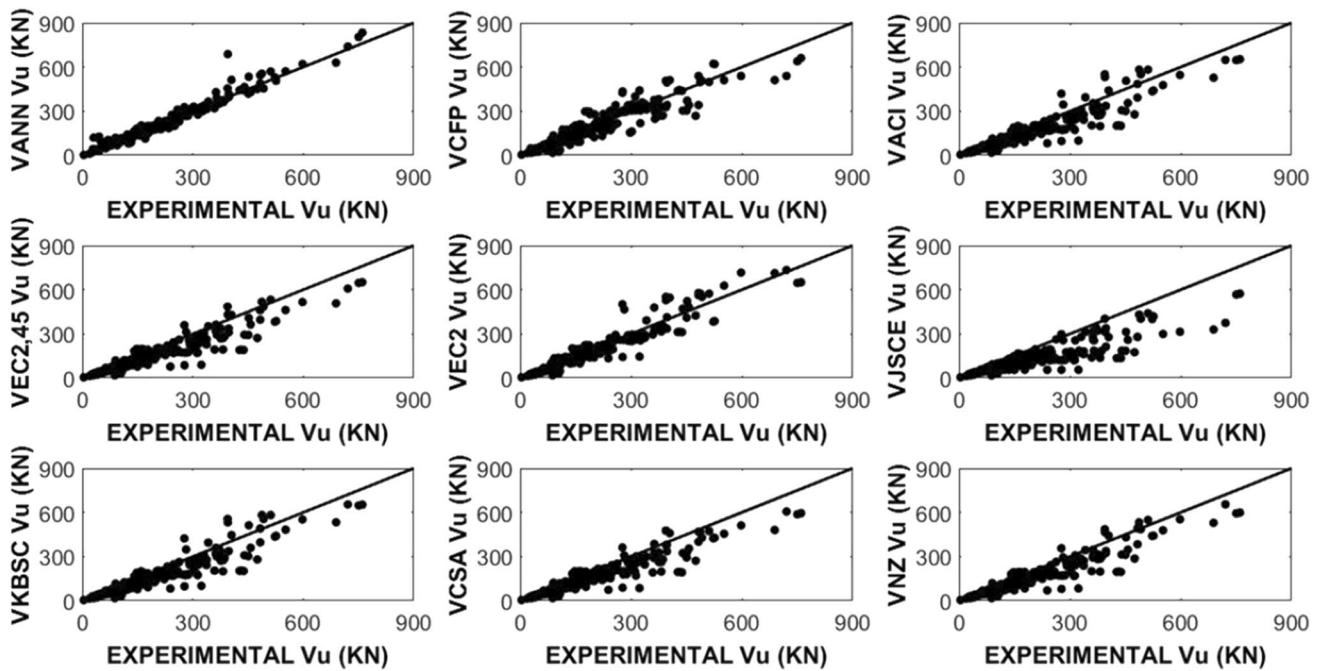


Fig. 11 Comparison between the predictions obtained from the various assessment methods employed and the ANN with their experimental counterparts for the case of beams with stirrups (BWS)

the values of R , MSE, and MAE calculated for the different ANN models developed for database DB-I (including beams without stirrups). From Table 7 and Fig. 8, it is observed the BWOS-4 exhibited the highest value of R (99.2%) in combination with the smallest values of MSE (0.05%) and MAE (1.5%), while, at the same time, employing the smallest number of parameters (i.e., d , a/d , $M_f/f_c b d^2$, and f_c). Similarly, based on the information provided in Table 7 and Fig. 8 for the case of database DB-II (including beams with stirrups), BWS-8 is the most efficient ANN model (with $R=99.2\%$, $MSE=0.08\%$, and $MAE=1.9\%$) and employs only seven parameters (i.e., d , b/d , a/d , $M_f/f_c b d^2$, f_c , and $\rho_w \times f_{yw}$).

In view of the above, ANN models BWOS-4 and DWS-8 have been selected for use in what follows.

Comparative study

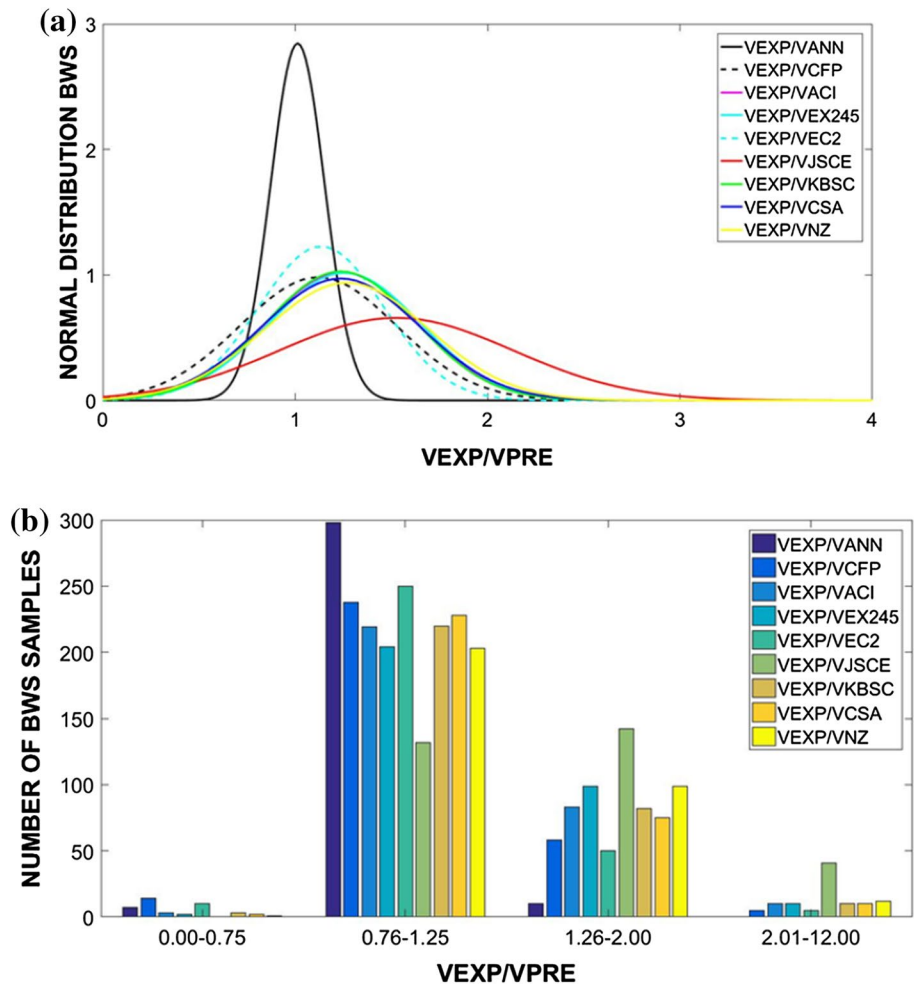
A comparative study between the predicted values of load-carrying capacity obtained from the ANNs, the design codes, and their experimentally established counterparts (forming the databases) is presented in Figs. 9, 10, 11 and 12 for the case of RC beams without (BWOS) and with (BWS) stirrups, respectively, and Table 8. From Table 8, it can be seen that the ANN technique produced the closest fit to the experimental values with the smallest standard deviation, with the deviation of the mean normalized values of load-carrying capacity from unity being 0% and

1% for the cases of RC beams without and with transverse reinforcement, respectively. The code predicted mean normalized values have been found to be conservative for both types of beams in all but one case, that of the NZ code for beams with stirrups, for which the predicted value overestimates its experimental counterpart by 8%. For the case of beams without transverse reinforcement, the code predicted mean normalized values deviate from unity by an amount ranging between approximately 27% and 60%, with the EC2 and NZ code values exhibiting the smaller and larger deviations, respectively. For the case of beams with transverse reinforcement, the corresponding deviations range between approximately 14% and 53%, with EC2 and JSCE values exhibiting the smaller and larger deviations, respectively. As regards the CFP method, this produces values which deviate from unity by an amount of the order of 12% for both types of beams considered (with and without stirrups). The trends of behaviour discussed above are presented in a graphical form in Figs. 9, 10, 11 and 12.

It should be noted at this stage that, unlike the other methods investigated, the CFP method is, by nature, capable of providing close predictions of not only the values of load-carrying capacities, but also of the locations and the causes of failure. In fact, the predictions of the latter have been successful in all cases investigated in the present work. Naturally, similar predictions cannot be expected to be obtained from a purely empirical method such as



Fig. 12 a Curves describing the normal distribution of V_{EXP}/V_{PRED} for beams with stirrups (BWS) based on the predictions obtained from the various assessment methods employed and the use of ANNs. **b** Distribution V_{EXP}/V_{PRED} for beams with stirrups (BWS) based on the predictions obtained from the various assessment methods employed and the use of ANNs



the ANN technique employed in the present work, which, however, has been found capable of producing the best possible numerical predictions of load-carrying capacity.

It is also important to note that, as already discussed, the code methods are also heavily dependent on the use of experimental data for the calibration of the formulae underlying the assumed mechanisms of load transfer. The use of the code-adopted formulae, when compared with the use of the ANN technique employed in the present work, appears to reduce the effectiveness of the calibration process and this may be considered to reflect shortcomings of the underlying theory in describing the actual mechanism of load transfer. On the other hand, the ability of the CFP method to produce numerical predictions exhibiting a relatively small deviation from those resulting from the ANN procedure, without the need of prior calibration, may be considered as an indication of the validity of the concepts underlying the method.

Conclusions

The accuracy of the predictions of the RC design methods considered is assessed primarily against the available test data. The calibrated ANNs developed were successful in providing predictions very close to their experimentally established counterparts. The good correlation between the available test data and the predictions of the ANNs demonstrate that the latter soft computing method provides an effective tool capable of objectively analyzing the available test data and quantifying the effect of specific parameters on certain important characteristics of RC structural behaviour (e.g., the load-carrying capacity and mode of failure).

The ANN procedure, which is purely empirical in nature and follows a detailed mathematical approach within a heuristic scheme allowing for all parameters widely assumed to affect load-carrying capacity, is found to produce a nearly perfect fit to the test data available in the literature on shear capacity of simply supported beam/column elements.

Current code methods for calculating shear capacity are semi-empirical in that, unlike the ANN procedure, the test



Table 8 Results of V_{Exp}/V_{Pred} for RC beams

	Min	Max	Mean	St. dev	CV%
For beams without stirrups (BWOS) DB-I					
V_{Exp}/V_{ANN}	0.48	2.00	1.00	0.16	0.16
V_{Exp}/V_{CFP}	0.57	2.26	1.12	0.26	0.23
V_{Exp}/V_{ACI}	0.67	6.47	1.51	0.67	0.44
V_{Exp}/V_{EC2-45}	0.72	6.62	1.27	0.65	0.51
V_{Exp}/V_{EC2}	0.72	6.62	1.27	0.65	0.51
V_{Exp}/V_{JSCE}	0.63	8.73	1.58	0.86	0.54
V_{Exp}/V_{KBSC}	0.67	6.36	1.49	0.65	0.44
V_{Exp}/V_{CSA}	0.66	12.93	1.57	1.24	0.79
V_{Exp}/V_{NZ}	0.56	11.53	1.60	1.08	0.67
For beams with stirrups (BWS) DB-II					
V_{Exp}/V_{ANN}	0.26	1.79	1.01	0.14	0.14
V_{Exp}/V_{CFP}	0.59	5.15	1.12	0.41	0.36
V_{Exp}/V_{ACI}	0.66	4.97	1.24	0.39	0.31
V_{Exp}/V_{EC2-45}	0.75	4.69	1.26	0.39	0.31
V_{Exp}/V_{EC2}	0.55	4.69	1.14	0.33	0.29
V_{Exp}/V_{JSCE}	0.82	5.79	1.53	0.61	0.40
V_{Exp}/V_{KBSC}	0.65	4.97	1.24	0.39	0.31
V_{Exp}/V_{CSA}	0.74	4.97	1.24	0.41	0.33
V_{Exp}/V_{NZ}	0.04	4.15	0.92	0.51	0.56

data are used for calibrating the formulation of the theoretical basis of the methods. The resulting formulae are found conservative: they underestimate shear capacity by an amount ranging between approximately 15% and 60%.

The inability of the code methods to provide a fit to the test data as close as that of the ANN procedure is considered to reflect the lack of effectiveness of the code formulae to describe the trends of behaviour exhibited by the test data and this may be viewed as an indication of shortcomings of the theoretical basis of the code-design methods.

In contrast with code methods, the formulation of the theoretical basis of the CFP method does not require calibration through the use of test data. The failure criteria have been derived from first principles and are functions of the strength of concrete of uniaxial compression and tension. Moreover, unlike the code methods, the CFP is capable of identifying both the location and the causes of failure.

An indication of the validity of the method is provided not only by the fit to the test data which is nearly as close as that of the ANN procedure for the case of simply supported beams, but also by the realistic predictions of the load-carrying capacity, location, and causes of failure of the structural elements exhibiting points of contra-flexure. However, since test data on the behaviour of the latter structural elements are rather sparse, additional work is required for drawing definitive conclusions on the method.

Acknowledgements This research was supported by the HORIZON 2020 Marie Skłodowska-Curie Research Fellowship Programme H2020-660545, titled: Analysis of RC Structures Employing Neural Networks (ARCSENN).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Abdalla JA, Elsanosi A, Abdelwahab A (2007) Modeling and simulation of shear resistance of R/C beams using artificial neural network. *J Franklin Inst* 344(5):741–756
- ACI 318-14 (2014) Building code requirements for structural concrete (ACI 318-14) and commentary (ACI 318R-14), Reported by ACI Committee 318, American Concrete Institute, Farmington Hills, pp 1–519
- ACI 445R-99 (1999) Recent approaches to shear design of structural concrete, Joint ACI-ASCE Committee 445, American Concrete Institute, Farmington Hills, pp 1–55
- Ahmad A, Cotsovovs DM and Lagaros ND (2016) Assessing the reliability of RC code predictions through the use of artificial neural network. In: 1st international conference on structural safety under fire & blast. Glasgow, UK
- Basheer IA, Hajmeer M (2000) Artificial neural networks: fundamentals, computing, design, and application. *J Microbiol Methods* 43:3–31
- Beale MH, Hagan MT, Demuth HB (2015) Neural Network Toolbox™—user's guide. The MathWorks, Inc. 3 Apple Hill Drive Natick, MA 01760-2098: The MathWorks, Inc
- Castellano G, Fanelli AM (2000) Variable selection using neural-network models. *Neurocomputing* 31:1–13
- Cladera A, Mari AR (2004a) Shear design procedure for reinforced normal and high-strength concrete beams using artificial neural networks. Part I: Beams without stirrups. *Eng Struct* 26(7):917–926
- Cladera A, Mari AR (2004b) Shear design procedure for reinforced normal and high-strength concrete beams using artificial neural networks. Part II: beams with stirrups. *Eng Struct* 26(7):927–936
- CSA-A23.3-04 (2007) Design of concrete structures. Canadian Standard Association, 5060 Spectrum Way, Suite 100, Mississauga, Ontario, Canada
- Eurocode 2 (2004) Design of concrete structures—Part 1-1: General rules and rules for buildings. In: EN 1992-1-1 2004: Management Centre: Avenue Marnix 17, B-1000 Brussels
- JSCE (2007) Standard specifications for concrete structures. In JSCE guideline for concrete no. 15. 2007: Yotsuya 1-chome, Shinjuku-ku, Tokyo 160-0004, Japan, pp 1–503
- Jung S, Kim KS (2008) Knowledge-based prediction of shear strength of concrete beams without shear reinforcement. *Eng Struct* 30(6):1515–1525
- Kotsovovs MD (2014) Compressive force-path method: unified ultimate limit-state design of concrete structures. Springer International Publishing, Cham
- Kotsovovs MD (2015) Finite-element modelling of structural concrete: short-term static and dynamic loading conditions. CRC Press/Taylor & Francis, Boca Raton, pp 1–355



- Krogh A, Vedelsby J (1995) Neural network ensembles, cross validation, and active learning. In: Tesauro G, Touretzky DS, Leen TK (eds) *Advances in neural information processing systems*. MIT Press, Cambridge
- Mansour MY, Dicleli M, Lee JY, Zhang J (2004) Predicting the shear strength of reinforced concrete beams using artificial neural networks. *Eng Struct* 26(6):781–799
- Mojtaba N, Bali M, Naeef MR, Amiri JV (2013) Prediction of lateral confinement coefficient in reinforced concrete columns using M5 machine learning method. *KSCE J Civ Eng* 17(7):1714–1719
- NZS 3101.1:2006 Concrete structures standard, Part-1: Design of concrete structures. New Zealand Standards
- Panagiotakos TB, Fardis MN (2001) Deformations of reinforced concrete members at yielding and ultimate. *ACI Struct J* 98(2):135–148
- Reineck KH (2013) ACI-Dafstb database of shear tests on slender reinforced concrete beams without stirrups. *ACI Struct J* 110(5):867–875
- Reineck KH, Bentz E, Fitik B, Kuchma DA, Bayrak O (2014) ACI-Dafstb databases for shear tests on slender reinforced concrete beams with stirrups. *ACI Struct J* 111(5):1147–1156
- Svozil D, Kvasnicka V, Pospichal J (1997) Introduction to multi-layer feed-forward neural networks. *Chemom Intell Lab Syst* 39:43–62
- Utans J, Moody J, Rehfuess S, Siegelmann H (1995) Input variable selection for neural networks: application to predicting the U.S. business cycle. *IEEE Trans Knowl Data Eng* 1995:118–122
- Yang KH, Ashour AF, Song JK (2007) Shear capacity of reinforced concrete beams using neural network. *Int J Concr Struct Mater* 1(1):66–73
- Kani GNJ (1964) The riddle of shear failure and its solution. *ACI Proc* 61(28):441–467

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

