e-ISSN: 2689-4934

Available online at
http://uppp.uitm.edu.my/online-issues.html

ESTEEM Academic Journal Vol 22, March 2026, 1–16

ESTEEM
Academic Journal
(EAJ)

Shear capacity estimation of reinforced concrete deep beam with vertical bar using artificial neural network and small datasets

Syahrul Fithry Senin[1*], Nureen Natasya Mohamad Zamri[2], Rohamezan Rohim[3], Amer Yusuff[4], Chan Hun Beng[5], Nur Ashikin Marzuki[6]

[123456]*Civil Engineering Studies, Universiti Teknologi MARA Cawangan Pulau Pinang, Permatang Pauh Campus, 13500 Permatang Pauh, Penang, Malaysia*

ARTICLE INFO

Article history:
Received 14 December 2025
Revised 1 March 2026
Accepted 16 March 2026
Online first
Published 1 April 2026

Keywords:
Machine learning
Artificial neural network
Reinforced concrete
Deep beam
Shear capacity
Best network

DOI:
10.24191/esteem.v22iMarch.9841

ABSTRACT

In this study, we employ machine learning by applying an artificial neural network (ANN) to predict the shear capacity of simply supported reinforced concrete deep beams from a small dataset. A database of 76 experiments, comprising 13 key parameters, was prepared and used to train and tune various ANN configurations. The Levenberg−Marquardt algorithm converged fastest and most accurately after systematic trials and introducing a second hidden layer significantly enhanced the nonlinear mapping. An optimal network of 11-12 neurons with radial basis activation achieved a training root mean square error (RMSE) of 0.2345. Data validation revealed that correlation coefficients for training (0.999) and testing (0.992) were found, with over 95% of predictions within 5% of measured strengths. The model developed was shown to be overfitting as the number of datasets in this experiment is limited. Future studies need to be done to include more datasets to prevent overfitting.

1. INTRODUCTION

Reinforced concrete deep beams (RCDB) are among the most important components of present-day structures in multi-storey buildings, bridges, and offshore installations. Their unique geometric design makes distributing large loads from the application site to the support systems via a strut-based load transfer system very effective [1]. Due to this behaviour, RCDB applications are prevalent in critical structural elements such as pile caps, transfer girders, and shear walls. In tall buildings, these components are vital

[1*] Corresponding author. *E-mail address*: syahrul573@uitm.edu.my
https://doi.org/10.24191/esteem.v22iMarch.9841

for resisting lateral forces and will take precedence over changes in column alignment and the omission of columns.

When RCDB spans have a span-to-depth ratio (SDR) not exceeding 3.0 [2], Eurocode 2 recognises it as a deep geometry type. As highlighted in [3], the strain distribution in deep beams is markedly different from that in slender beams. They produce different shear deformations, and the classical flexural theory is underappreciated for these kinds of cases. Hence, RCDBs are more susceptible to shear failure; therefore, the accuracy of estimating their shear strength is a key factor to guarantee structural safety and achieve cost-effective design [4].

Many prediction equations estimate shear capacity for SDRs of less than 3, such as probabilistic, semi-empirical, and genetics-based models. Another promising prediction of the shear capacity of RCDB structures has come from Artificial Neural Network (ANN) techniques. To address these challenges in determining the shear capacity of RCDB, we set out to tackle two main fronts. The first, the limited case of the shear response of the supported RCDB, undergoing arch action (SDR smaller than 3), will be examined using ANN with vertical bars. The second aim is to investigate ANN architecture performance by varying the number of neurons in the hidden layer to improve predictive performance. The performance of the model is described by the Root Mean Square Error (RMSE and the correlation coefficient (R).

1.1 Existing methods and research gaps in predicting the shear capacity of RCDB

Predicting the shear capacity of RCDB is a problem with no exact solution [5]. To date, many theoretical and experimental investigations have been conducted to predict the strength of RCDB. Most experiments have aimed to develop basic predictive equations based on various shear mechanisms to facilitate deep beam design in concrete structures. The precision of such predictive equations appears to be compromised.

Since the 1950s, many investigations have been conducted on various formulations to estimate the shear strength of these members. The underlying mechanisms can usually be classified of: probabilistic models [6], semi-empirical models, and models created with genetic algorithms [7]. Probabilistic models are subject to variables predicted by a given distribution, which causes the behaviour of variables influencing deep beam to often be neither normal nor linear [8]. Of all these methods, semi-empirical models are most commonly used to predict shear capacity, and the majority of design code provisions are semi-empirical in nature [9]. For these models, the coefficients of equations derived from simplified mechanical mechanisms are typically determined by regression analysis of experimental shear test data. While the above-mentioned methods are powerful enough for prediction, it is impossible to implement them in practice in some cases, owing to the complexity of the underlying mathematics.

An alternative for estimating shear capacity is the Strut and Tie Method (STM). The STM has been formally adopted by major structural codes as a basic design methodology for deep beams. STM reduces the internal force transfer to a trussed state composed of struts, ties, and nodes, which is fundamental to STM's ability to provide a simplified account of the discontinuity regions. Regardless, STM depends almost entirely on the designer to specify where the load paths and fields of stress will be [10], which can result in disparate outcomes from different engineers. Not only that, but STM tends to deal with irregular geometries and difficult load redistribution, and it also lacks standardised procedures, which limits the reliability of STM's application. It also lacks standardised procedures, which limits the reliability of STM's application in some automated or complex design cases [11]. The presence of regions, namely disturbed (D-regions), especially where the SDR ratio is less than 3, is inconsistent with the basic assumptions of Euler–Bernoulli beam theory. Shear deformation significantly influences the structural response of RCDB; neglecting it may lead to substantial errors in shear capacity estimation [12].

Shear deformation has been incorporated into Timoshenko beam theory to more realistically represent deep-beam structural behaviour. In contrast, few reliable findings are provided regarding the determination

https://doi.org/10.24191/esteem.v22iMarch.9841

of shear resistance under load and strain between the concrete and the vertical bar, despite many researchers' efforts in this study. The vertical bar provides an intersecting critical diagonal crack, which results in peak shear resistance [13]. From an experimental report [14], the vertical bar reinforcement in a deep reinforced beam, when introduced at 0.38%, increases the shear capacity by 16.46 per cent. [15] reported a study of RCDB shear capacity without vertical bar using ANN and observed that the model outperformed the baseline models, especially under the trained condition on limited data.

In the year 2025, [16], the ultimate shear capacity of RCDB from deep reinforcement beams without a vertical bar was studied by ANN. The two data-driven dependent model is developed and is feasible, practical, and accurate. Later [17] used ANN technology to predict shear resistance without vertical bars and obtained a highly accurate coefficient of determination of 0.974 for its experimental data. However, just like before, the input of the vertical bar took no attention when calculating the shear capacity of the deep RC beam. To address these problems, some recent works have delved into machine learning techniques. ANN as a data-driven alternative for the prediction of RCDB shear capacity, where SDR is less than 3 in the presence of a vertical bar. Machine learning techniques (such as ANN) can discover complex, nonlinear relationships directly from experimental data and do not impose mathematical formulations.

1.2 ANN architectures and comparison between ANN and STM for predicting shear capacity

ANNs are available in various architectures, which can either reflect the nature of the problem or the structure of the data to be implemented. One well-established architecture we can use is a feedforward backpropagation (BP) network. Most commonly, this architecture is applied with the BP algorithm, a popular supervised learning technique [18]. Two main stages, namely forward propagation and backward propagation, were conducted during training. [19]

In the forward phase, the input data is transmitted through the network layers and consequently returns the output. Using a loss function, the predicted output is compared with the target, where the prediction error is calculated. During the backward phase, this error is propagated back through the network to check whether each weight and bias contributes directly to the overall error [20], as shown in Figure 1.

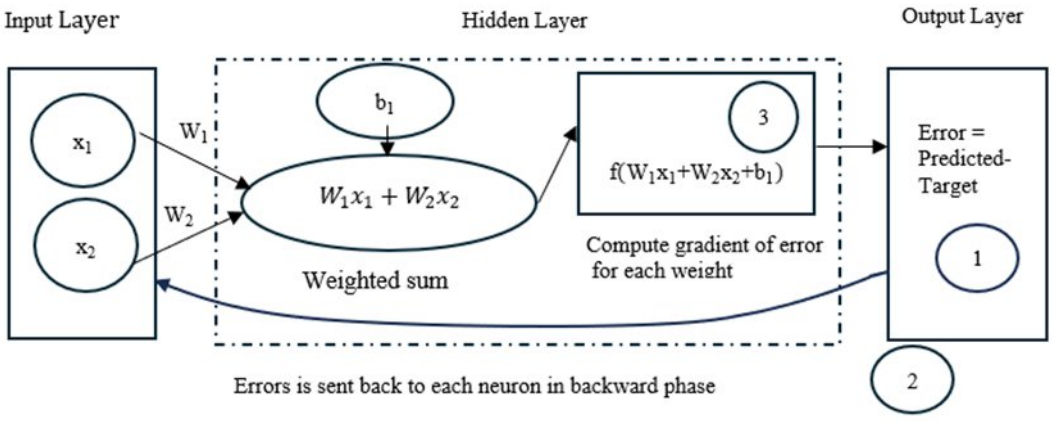

Fig. 1. Backpropagation phase of ANN (The numbers in circles show the order of BP)

https://doi.org/10.24191/esteem.v22iMarch.9841

This feedback is then used to compute a gradient of the network parameters and update them. These parameters are optimised using techniques such as gradient descent and its variants. With subsequent training in iterations, the error decreases progressively and the prediction accuracy improves [21]. The ANN algorithm was proposed as it outperforms STM for predicting the shear capacity of RCDB. Accordingly, [22] investigates the shear strength predicted by the ANN model against the other two important techniques in the standard practices code, ACI318 and EC2, which both utilise STM. ANN has well-defined shear strength prediction results to only 10% error superior to the STM technique. This STM methodology makes estimating shear-induced load transfer very difficult because it is conditioned by the span-to-depth ratio and the spacing of vertical bars [23]. The [5] article explored STM to estimate deep RC beam shear strength, in an alternative manner compared with ANN. They found that the ANN technique is more sensitive for predicting shear strength than STM. Based on this information, the ANN technique will be employed as the established method to accurately predict the shear capacity of RCDB.

2. METHODOLOGY

This work employs a quantitative modelling methodology to predict the structural shear capacity of supported RCDBs using an ANN. The method, performance evaluation, and comparison of predicted and actual datasets from the experiment. All computation was carried out in MATLAB R2020a, and early-stage processes, including data cleaning and normalisation, were undertaken in RapidMiner. The study flow diagram is shown in Figure 2, and we present the details of each phase in this paper.

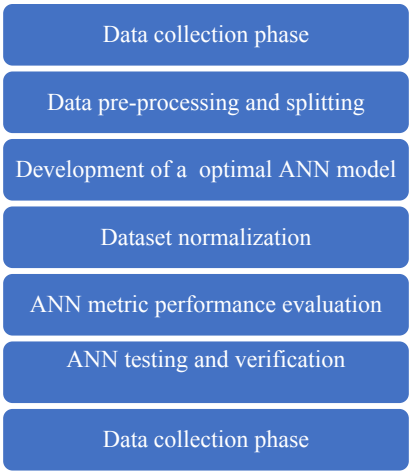

Fig. 2. Research flow diagram summary

2.1 Data collection phase

In general, the quality and suitability of the dataset influence the ANN model's performance during training. As mentioned earlier, the dataset used in this research is derived from reference [24] and comprises 76 experimental results for supported RCDB. An important reason cited for the limited availability of data in civil engineering is the high cost and time associated with data collection, sample preparation, sample setup, and so on. Financial resources and specialised equipment are also required for field observations, mapping, measurements, and long-term monitoring systems. Moreover, large-scale and physical tests require strict adherence to a schedule or process that typically involves two or more test repetitions in

https://doi.org/10.24191/esteem.v22iMarch.9841

response to different materials or situations, human or equipment errors, and inadequate testing setups. And this makes it challenging to lay the groundwork for large, diverse datasets for ML model development and validation. The experimental records contained 13 input parameters for geometric dimensions, material properties, and reinforcement characteristics, and 1 output variable: the experimentally obtained structural shear capacity. Table 1 describes these variables as simple input parameters and provides minimum, maximum and mean values.

2.2 Data pre-processing and data splitting determination

The dataset was pre-processed on RapidMiner to ensure its readiness for ANN modelling. It is pertinent to note that the pre-processing includes recognising and removing duplicate records, and addressing missing entries via removal or statistical substitution. We imported the original Excel-based data file directly into RapidMiner to ensure its format was compatible. The cleaned dataset was then exported back to Excel, after confirming the completeness, consistency, and proper labelling of all entries. To enable detailed analysis and model training, this cleaned dataset (with no incomplete or repetitive entries) was imported into MATLAB. At this stage, we gained foundational knowledge to enhance the reliability and predictive power of the ANN model. Following the approach by Bichri [25], we assigned 80 per cent of the data set for training and 20 per cent for testing, as recommended [26]. For the case of 13 features, the 80–20 distribution is consistent with Equation 1 [27] to get a good balance between learning, reducing overfitting and allowing for objective model evaluation:

$$\frac{\text{Optimal training}}{\text{Testing Ratio}} = \frac{\sqrt{p}}{1} = \frac{\sqrt{13}}{1} = \frac{3.6}{1} \approx 4{:}1 \tag{1}$$

2.3 ANN Model development

The ANN framework developed in this work is designed to estimate the shear capacity of RCDB. All the ANN models were developed, trained, and verified using MATLAB through an in-house script that enabled the construction of flexible network architectures, training as appropriate, and evaluation of various models.

Four basic hyperparameters, i.e., (i) hidden layer numbers, (ii) neuron numbers per layer, (iii) type of training algorithm (six options), and (iv) activation function (eight options) were systematically varied to generate several ANN architectures. From this training dataset, 13 relevant input features affecting shear capacity were selected. Different configurations were tested to determine how each affected predictive accuracy. We have analysed both shallow (one hidden layer) and moderately deep (two hidden layers) feedforward networks. Two hidden layers are widely reported to capture complex nonlinear relationships [28], and Heaton [29] shows that one hidden layer can approximate any function with good design and training.

The value of neurons was computed using a common heuristic, with values between input and output neurons [30]. As a consequence, hidden layers are assigned to 1 to 13 neurons. This practice helps to balance the risks of overfitting and underfitting [31]. Other hyperparameters were kept unchanged to ensure a fair comparison between networks. Epochs were set to 1000 (after [26]) to achieve sufficient training but less overfitting. More importantly, the performance threshold of 0.95 accuracy was selected to provide meaningful predictions and strong generalisation [32]. All the ANN architectures adopted a multilayer feedforward structure trained using the backpropagation (BP) algorithm.

https://doi.org/10.24191/esteem.v22iMarch.9841

2.4 Dataset normalisation phase

Normalising the dataset before its training was required to ensure homogeneity in dataset scaling (input and output). Normalisation speeds up convergence during training and increases overall network stability [33]. This was especially important in view of the variable features with different units and ranges of measurement. The min–max normalization technique was used for scaling, with all the inputs scaled between 0 and 1 based on Equation 2. This method prevents higher-magnitude variables from dominating the learning algorithm and promotes an equitable representation of features [34].

$$k_{i,normal} = \frac{k_i - \min(k_i)}{\max(k_i) - \min(k_i)} \tag{2}$$

where

$k_{i,\ normal}$ = the i-th k normalised input features value
k_i = the actual i-th k input features value from the dataset
$\min(k_i)$ = the minimum i-th k input feature values from the whole dataset
$\max(k_i)$ = the maximum i-th k input feature values from the whole dataset

2.5 ANN performance metric evaluation phase

Using the validation dataset, we assessed the model's baseline performance. At this time, hyperparameters were tuned to improve model accuracy and select the best network structure. The model evaluation was primarily based on the Root Mean Square Error (RMSE), which is highly sensitive to large deviations and effectively captures average prediction errors. This smaller measure, RMSE, reflects a stronger prediction ability. In addition, the correlation coefficient (R) was computed to examine the linear correlation between the predicted and observed values. Equations 3 and 4 are the formulations for RMSE and R.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \left(Shear_{act,i} - Shear_{pred,i}\right)^2}{N}} \tag{3}$$

$$R = 1 - \frac{\sum_{i=1}^{N} \left(Shear_{act,i} - Shear_{pred,i}\right)^2}{\sum_{i=1}^{N} \left(Shear_{act,i} - \overline{Shear}\right)^2} \tag{4}$$

where

$Shear_{act,i}$ = the value of actual structural shear capacity for the i-th from the experimental work
$Shear_{pred,i}$ = the value of predicted structural shear capacity for the i-th from the ANN model
$\overline{Shear}$ = the mean of the actual structural shear capacity of N samples from the experimental work

https://doi.org/10.24191/esteem.v22iMarch.9841

N = number of datasets used

2.6 ANN testing and performance verification phase

Once the ANN model was selected through validation, its predictive performance was further evaluated on the test dataset. The predicted shear capacities from the ANN model were compared with experimental measurements to assess accuracy and generalisation.

3. RESULTS AND DISCUSSIONS

This part discusses the performance outcomes of various ANN architectures developed in MATLAB for predicting the shear capacities of supported RCDB. The dataset described earlier formed the basis for model training. A comparative analysis of single- and two-hidden-layer networks revealed notable patterns related to network structure, activation functions, and training algorithms. The performance metrics, Root Mean Square Error and correlation coefficient, R, were monitored during training and testing phases.

3.1 Data descriptions

The dataset contains 76 experimental records with 13 input features (x_1 to x_{13}) and one output. It is derived from [24], a well-recognised reference in the field of civil engineering. Descriptive statistics of the dataset are depicted in Table 1.

Table 1. Dataset statistical description of the study

Features	Units	Minimum	Maximum	Mean	Standard deviation	Type
Beam's width (x_1)	mm	80	533	238.76	157.13	Input
Beam's height (x_2)	mm	250	1905	751.91	455.30	Input
Effective depth (x_3)	mm	207	1750	673.16	418.90	Input
Span-to-depth ratio (x_4)	-	0.56	2.34	1.34	0.49	Input
Load position (x_5)	mm	50	417	194.84	121.98	Input
Support bearing (x_6)	mm	53	406	218.54	117.62	Input
Tension bar ratio (x_7)	%	1.07	3.24	2.08	0.74	Input
Vertical bar ratio (x_8)	%	0.12	1.25	0.30	0.18	Input
Horizontal bar ratio (x_9)	%	0.12	0.68	0.23	0.16	Input
Yield stress tension bar (x_{10})	MPa	420	555	482.77	44.13	Input
Yield stress horizontal;	MPa	400	558	462	48.86	Input

https://doi.org/10.24191/esteem.v22iMarch.9841

bar (x_{11})						
Yield stress vertical bar (x_{12})	MPa	414	558	488.22	43.16	Input
Concrete compressive strength (x_{13})	MPa	16.1	97.7	30.31	14.98	Input
Shear strength capacity	kN	85	5747	894.70	1147.95	Output

3.2 ANN model in one hidden layer

The performance of an ANN relies heavily on the proper selection and coordination of several critical factors. An extensive evaluation of various single-hidden-layer structures was done; the results are depicted in Table 2.

Table 2. Neuron number and corresponding RMSE for selected training and activation functions (one hidden layer)

Training algorithm	Activation Function							
	tansig	logsig	poslin	purelin	softmax	satlin	hardlim	radbas
trainlm	0.80	0.95	30.25	305.54	2.73	46.70	685.67	0.82
trainscg	24	29	35	306	23	29	538	19
trainbr	29.1	30.5	35.2	309	33	45	539	28
trainrp	30	31	33	306	30	47	555	30
traincgb	24	25	35	305	16	30	686	13
traingcf	30	32	42	306	26	29	678	18

Regarding the single-hidden-layer models, a setup that leveraged the Levenberg–Marquardt (trainlm) training algorithm along with the activation function (tansig) and 13 neuron numbers, achieved the highest performance. The best setup had the lowest training RMSE = 0.80, as presented in Table 2. A second-best configuration achieved an RMSE of 0.82 by coupling trainlm with the radial basis function (radbas) and using 11 neurons. A third prominent model, using trainlm with the logsig activation function and 12 neurons, achieved an RMSE of 0.95. These results were clear for single-hidden-layer combinations, which also showed which combinations performed best overall. Several previous works [35, 36] also found that trainlm works well together with the nonlinear activation functions for nonlinear features on predictive tasks. Thus, the setup with tansig and 13 neurons in one trainlm was used as the baseline for the latter 2 hidden-layer experiments. On the other hand, activation functions (i.e. poslin, satlin, purelin and hardlim) had higher RMSE values than any other of the training algorithms, which means that they could not efficiently characterise the complex behaviour in the nonlinear aspect caused by the shear capacity due to the RCDB.

https://doi.org/10.24191/esteem.v22iMarch.9841

3.3 ANN model from the data in two hidden layers

In analysing 1-hidden-layer networks, it was found that trainlm (Levenberg–Marquardt) yielded the most outstanding output of the six algorithms assessed. Thus, the second hidden layer of the model was addedand its neuron number , and its neuron count was systematically adjusted from 1 to 13, repeating the same trial-and-error algorithm. The first hidden layer was initialised at 13 neurons (the previous best number) to provide the steady-state foundation of error minimisation in the second hidden layer. This method is consistent with the progressive optimisation procedures, which are particularly appropriate for deep neural networks [37,38]. The results of the most advanced ANN model with tansig activation function with trainlm are shown in Table 3.

The relationship between the number of neurons in the second hidden layer and RMSE was clearly a kind of U-curve. The minimum possible RMSE was 0.2383 at 6 neurons in the first hidden layer, with the second layer having 1 to 12 neurons. This nonlinear relationship indicates the complicated effect of network complexity on the generalisation error. The U-shaped behaviour corresponds to the bias–variance trade-off in statistical learning theory [39]. In the case of low neuron counts (1–5), the bias is high, and the model cannot model complex relationships [40]. Growing the number of neurons to 6 or beyond results in a lot of variance — the model starts fitting noise rather than simple patterns, which is aggravated by trainlm's sensitivity to complex error surfaces [41]. Moreover, tansig's saturating nonlinearity can render the gradients of deeper structures vanishing, while this effect in the two-layer design is mitigated. Too many neurons, on the other hand, cause gradient fluctuation and also interfere with handling stability and weight updates during backpropagation, which accounts for the performance drop beyond six neurons [42].

Table 3. Two-hidden-layer ANN training performance (first-ranked model)

Activation function	Training algorithm	Neuron number in the first hidden layer*	Neuron number in the second hidden layer	RMSE (training)
tansig	trainlm	13	1	132.41
			2	2.6055
			3	3.0657
			4	0.2389
			5	1.3623
			6	0.2383
			7	0.2444
			8	0.2439
			9	0.2441
			10	0.2428
			11	1.1682
			12	0.2414

* The number of neurons is fixed in this hidden layer

A different pattern was illustrated for the radbas activation function with trainlm, and 11 neurons in the first hidden layer, as shown in Table 4. RMSE values showed significant differences as the number of neurons in the second layer increased from 1 to 13, with values varying from 1147.35 to 0.2345. This difference reflects the high effect of the second-layer neuron count on performance. A similar trend was shown by Yu et al [43], who attributed the performance of radbas in low neuron numbers to its localised approximation properties, where each neuron responds to inputs within a specific range.

https://doi.org/10.24191/esteem.v22iMarch.9841

Table 4. Two-hidden-layer ANN training performance (second-ranked model)

Activation function	Training algorithm	Neuron number in the first hidden layer*	Neuron number in the second hidden layer	RMSE (training)
radbas	trainlm	11	1	1147.35
			2	128.37
			3	0.241
			4	1.3686
			5	2.2104
			6	1.0195
			7	1.7056
			8	0.2438
			9	1.072
			10	0.6509
			11	0.902
			12	0.2425
			13	0.238

* The number of neurons is fixed in this hidden layer

Table 5 shows the performance of the logsig activation function with trainlm and 12 neurons in the first hidden layer. As the second-layer neurons changed, RMSE values also varied from 1064.64 at one neuron, 0.241 at three neurons, to 0.2385–2.2101 for larger neuron counts. The lowest RMSE, 0.2385 (13 neurons), was achieved. These results demonstrate the architectural dynamic impact on logsig networks. Such behaviour is associated with vanishing gradients in asymmetric functions, where weight updates weaken as training proceeds [44]. The logit function's "logistic S-curve," which compresses outputs between 0 and 1, affects learning dynamics depending on the distribution of neurons.

Table 5. Two-hidden-layer ANN training performance (third-ranked model)

Activation function	Training algorithm	Neuron number in the first hidden layer*	Neuron number in the second hidden layer	RMSE (training)
logsig	trainlm	11	1	1064.64
			2	2.2322
			3	1147.35
			4	2.4803
			5	3.7533
			6	0.2407
			7	0.2417
			8	0.2439
			9	0.2424
			10	0.2465
			11	1.242
			12	0.2345
			13	0.2384

* The number of neurons is fixed in this hidden layer

https://doi.org/10.24191/esteem.v22iMarch.9841

According to the results, the configuration using radbas with 12 neurons in the second hidden layer had the lowest overall training RMSE, at 0.2345 (Table 3). It was better than the tansig config (RMSE = 0.2383 with 6 neurons in the second layer) and the logsig config (RMSE = 0.2385 with 13 neurons), respectively, suggesting a hierarchy of performance: radbas, followed by tansig, and ending with logsig.

3.4 Comparison between the predicted and actual shear capacities of the deep beam

A comparison is made between the predicted shear capacities and the actual shear capacities of the deep beam. To validate the predictive ability of the chosen ANN model, a direct comparison of the predicted shear capacities to the measured ones was performed. Radbas-trainlm (structured as 11–12 neurons) was chosen as the best model for all data points. Figure 3 shows a bar graph comparing the predicted shear capacities with the model's actual capacities for 15 of the 76 representative testing datasets. It appears that the predicted and actual values coincide closely in most cases, indicating the model's high generalisation. The ANN performed particularly well on moderate-sized shear-capacity datasets. There were a few departures with high-capacity values, especially in Dataset 13, which may be due to the limited representation of extreme cases in training. At most, prediction errors stayed within 5% for most of the data; Dataset 1 had the largest error, at 11.15%. Even so, the network demonstrated strong predictive performance on the dataset, underscoring its stability.

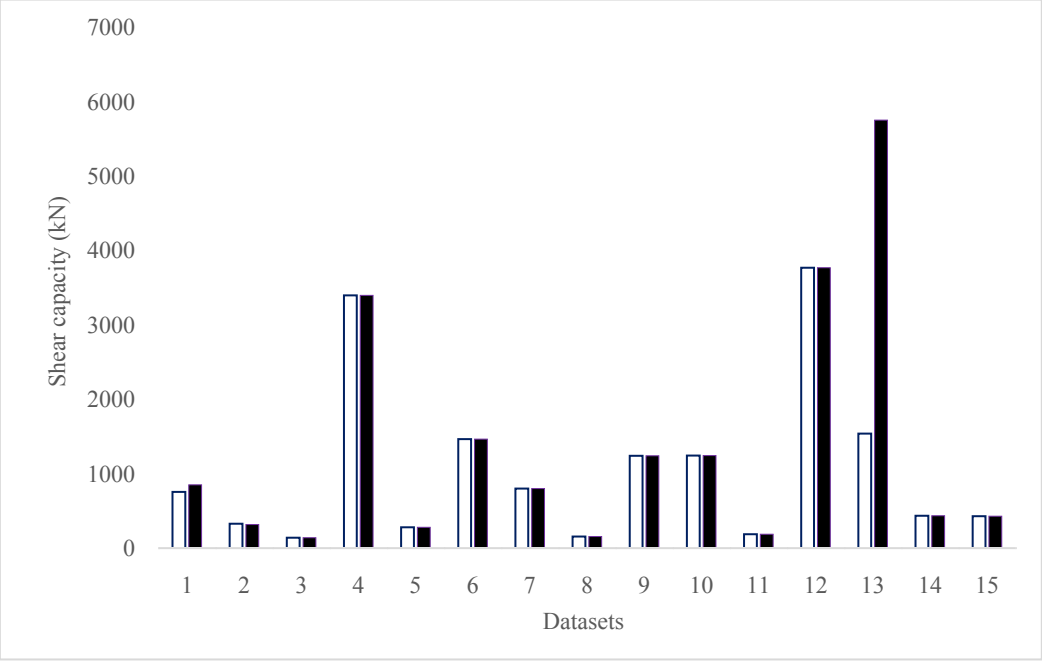

Fig. 3. Actual and predicted structural shear capacity based on the best with 2 hidden layer ANN Model (on the testing dataset)

https://doi.org/10.24191/esteem.v22iMarch.9841

3.5 Correlation coefficients analysis for the ANN of the two hidden layers

The correlation coefficients for the ANN's two hidden layers. Regression plots of actual shear capacities and R values for the training and testing datasets are presented in Figure 4. In the correlation between the predicted and actual shear capacities and R values, very high correlation coefficients of R = 0.999 for training and R = 0.992 for validation were observed for the linear fit between predictions and reality. In the case of test data, the model achieved R = 0.992, indicating a significant increase in correlation on unseen data, with some signs of overfitting. A small dataset (76 samples) may not adequately represent the true variability of a population. In machine learning, models try to capture how input features relate to output responses. In the limited sample sizes, the dataset might contain an incomplete or biased part of the potential patterns. works. It was concluded that a larger dataset should be used in future work to obtain a more accurate machine learning model.

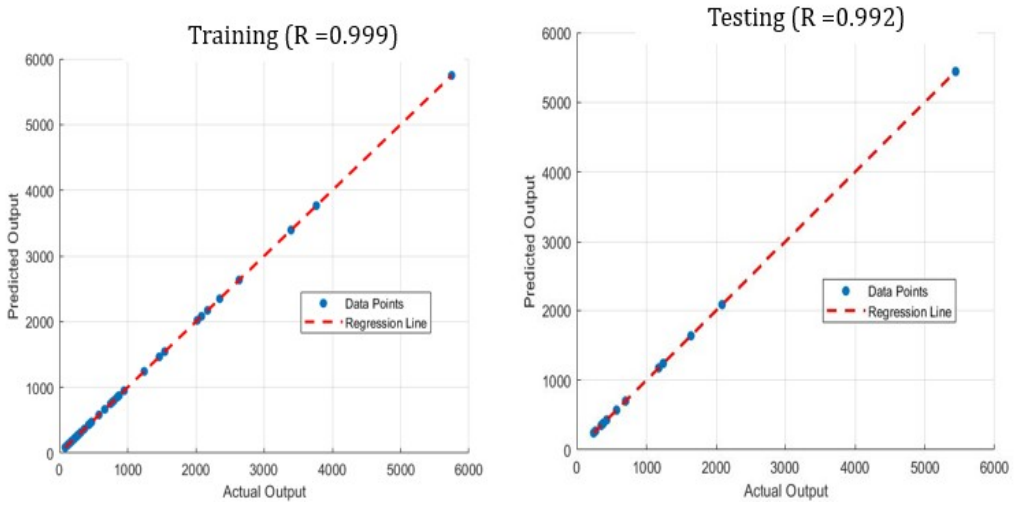

Fig. 4. The coefficient of determination, R, for the training datasets for testing and training dataset of the best 2-hidden layer ANN model

4. CONCLUSIONS

The present study has successfully used an ANN model to predict the shear capacity of simply supported RCDBs with vertical bars and span-to-depth ratios less than 3. Among the tested algorithms, Levenberg–Marquardt (trainlm) consistently produced the best performance across diverse network models and activation functions, making it the best training method. Architectural optimisation indicated that a second hidden layer significantly increased prediction performance compared to single-layer networks. Here are the key findings:

a) A second hidden layer added significantly better prediction than the single-layer architectures
b) For the best-performing architecture, the radbas activation function was employed.
c) The best structure consisted of 11 neurons in the first hidden layer and 12 neurons in the second hidden layer
d) This configuration produced the lowest RMSE of 0.2345.
e) The performance was better than the tansig configuration (RMSE = 0.2383).

https://doi.org/10.24191/esteem.v22iMarch.9841

f) It also outperformed the logsig configuration (RMSE = 0.2385).

A U-shaped RMSE curve for both bias and variance was obtained from the neuron counts in the second layer. This also stresses the need for proper architectural selection to counter underfitting and overfitting. Experimental results verified that most of the predictions had percentage errors below 5%, with R = 0.999 in the training/validation phase and R = 0.992 in the testing phase, which indicates the model was highly overfitting. To prevent possible overfitting in the ANN model in future work, sufficient datasets with more than 76 samples shall be used in future research studies.

5. CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this paper. All affiliations supporting this study are transparently acknowledged, and no competing financial or non-financial interests exist that could have influenced the outcome of this research.

6. AUTHORS' CONTRIBUTIONS

Syahrul Fithry Senin: Conceptualisation, supervision, methodology, and writing (review and editing); **Nureen Natasya Mohamad Zamri**: Formal analysis, writing the original draft; **Rohamezan Rohim**: Software; **Amer Yusuff**: Visualisation; **Chan Hung Beng**: Validation; **Nur Ashikin Marzuki**: Methodology

7. REFERENCES

[1] K. S. Abdul-Razzaq, A. M. Jalil, and S. F. Jebur, "Behaviour of reinforced concrete deep beams in previous studies," IOP Conference Series: Materials Science and Engineering, vol. 518, no. 2, p. 022065, May 2019, doi: https://doi.org/10.1088/1757-899x/518/2/022065.

[2] O. M. Makki and Hayder M K Al-Mutairee, "Continuous Deep Beams Behavior Under Static Loads: A Review Study," IOP Conference Series Earth and Environmental Science, vol. 961, no. 1, pp. 012034–012034, Jan. 2022, doi: https://doi.org/10.1088/1755-1315/961/1/012034.

[3] H. A. Jabir, J. M. Mhalhal, T. S. Al-Gasham, and S. R. Abid, "Mechanical characteristics of deep beams considering variable a/d ratios: an experimental investigation," IOP Conference Series: Materials Science and Engineering, vol. 988, no. 1, p. 012030, Dec. 2020, doi: https://doi.org/10.1088/1757-899x/988/1/012030.

[4] Tadesse Wakjira, M. Ibrahim, B. Sajjad, and Usama Ebead, "Shear capacity of reinforced concrete deep beams using genetic algorithm," IOP Conference Series Materials Science and Engineering, vol. 910, no. 1, pp. 012002–012002, Aug. 2020, doi: https://doi.org/10.1088/1757-899x/910/1/012002.

[5] G. S. Harsha and P. Poluraju, "Modified Strut and Tie Method and Truss Reinforcement for Shear Strengthening of Reinforced Concrete Deep Beams," International Journal of Technology, vol. 12, no. 1, p. 65, Jan. 2021, doi: https://doi.org/10.14716/ijtech.v12i1.3920.

[6] X. Ni, Q. Zhang, and G. Xu, "Probabilistic prediction method for shear strength capacity of RC deep beams based on the fusion of multiple machine learning models," Structures, vol. 76, p. 108864, Jun. 2025, doi: https://doi.org/10.1016/j.istruc.2025.108864.

https://doi.org/10.24191/esteem.v22iMarch.9841

[7] M. Shahnewaz, A. Rteil, and M. S. Alam, "Shear strength of reinforced concrete deep beams – A review with improved model by genetic algorithm and reliability analysis," Structures, vol. 23, pp. 494–508, Feb. 2020, doi: https://doi.org/10.1016/j.istruc.2019.09.006.

[8] Z. Li, X. Liu, D. Kou, Y. Hu, Q. Zhang, and Q. Yuan, "Probabilistic Models for the Shear Strength of RC Deep Beams," Applied Sciences, vol. 13, no. 8, p. 4853, Apr. 2023, doi: https://doi.org/10.3390/app13084853.

[9] J.-H. Song, W.-H. Kang, K.-S. Kim, and S.-M. Jung, "Probabilistic shear strength models for reinforced concrete beams without shear reinforcement," Structural Engineering and Mechanics, vol. 34, no. 1, pp. 15–38, Jan. 2010, doi: https://doi.org/10.12989/sem.2010.34.1.015.

[10] J. S. Jyothis and J. Jayasree, "Beyond the Standards: Analyzing RC Deep Beam Design Codes," International Research Journal of Engineering and Technology (IRJET), vol. 11, no. 6, pp. 133–139, Jun. 2024. [Online]. Available: https://www.irjet.net/archives/V11/i6/IRJET-V11I655.pdf.

[11] S. Fan, Y. Zhang, Y.-X. Ma, and K. H. Tan, "Strut-and-tie and finite element modelling of unsymmetrically-loaded deep beams," *Structures*, vol. 36, pp. 805–821, Dec. 2021, doi: https://doi.org/10.1016/j.istruc.2021.12.037.

[12] S. S. Hashemi, K. Sadeghi, Saeid Javidi, and Mahmoud Malakouti, "Analysis of RC Deep Beams Considering the Shear Deformations and Bar-concrete Interaction," *Periodica Polytechnica Civil Engineering*, Oct. 2020, doi: https://doi.org/10.3311/ppci.16192.

[13] M. Zhang, M. Deng, J. Yang, and Y. Zhang, "Experimental Study on the Shear Behavior of Reinforced Highly Ductile Fiber-Reinforced Concrete Beams with Stirrups," *Buildings*, vol. 12, no. 8, pp. 1264–1264, Aug. 2022, doi: https://doi.org/10.3390/buildings12081264.

[14] H. Mirzaaghabeik, S. K. Shukla, and N. S. Mashaan, "Effects of vertical reinforcement on the shear performance of UHPC deep beams with synthetic and steel fibres," Structures, vol. 76, p. 109038, Jun. 2025, doi: https://doi.org/10.1016/j.istruc.2025.109038.

[15] H. Pak, S. Leach, S. H. Yoon, and S. G. Paal, "A knowledge transfer enhanced ensemble approach to predict the shear capacity of reinforced concrete deep beams without stirrups," *Computer-Aided Civil and Infrastructure Engineering*, vol. 38, no. 11, pp. 1520–1535, Jul. 2023, doi: https://doi.org/10.1111/mice.12965.

[16] P. D. Nguyen and V. H. Dang, "Shear strength of FRP - reinforced concrete deep beams: Extension of beam and arch action model based on data-driven analysis," *Structures*, vol. 74, pp. 108553–108553, Mar. 2025, doi: https://doi.org/10.1016/j.istruc.2025.108553.

[17] A. B. David, O. B. Olalusi, P. O. Awoyera, and L. Simwanda, "Suitability of Mechanics-Based and Optimized Machine Learning-Based Models in the Shear Strength Prediction of Slender Beams Without Stirrups," *Buildings*, vol. 14, no. 12, p. 3946, Dec. 2024, doi: https://doi.org/10.3390/buildings14123946.

[18] I. U. Haq, "Activation Functions in Machine Learning: Theory and Applications," *ResearchGate*, Jun. 14, 2025. [Online]. Available: https://www.researchgate.net/publication/392696340_Activation_Functions_in_Machine_Learning_Theory_and_Applications

[19] Y. Zhang, S. Wang, and J. Liu, "Improved Backpropagation Algorithm for Feedforward Neural Networks," *Neural Computing and Applications*, vol. 33, no. 14, pp. 8535–8545, 2021.

[20] H. Li, X. Chen, and Q. Zhao, "Gradient-Based Optimization Techniques in Deep Learning: A Review," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 6, pp. 2456–2470, 2022.

[21] L. Chen, M. Wang, and T. Zhang, "Recent Advances in Neural Network Training and Generalization," *Artificial Intelligence Review*, vol. 57, no. 2, pp. 1203–1225, 2024.

https://doi.org/10.24191/esteem.v22iMarch.9841

[22] K. Megahed, "Prediction and reliability analysis of shear strength of RC deep beams," Scientific Reports, vol. 14, no. 1, Jun. 2024, doi: https://doi.org/10.1038/s41598-024-64386-w.

[23] K. S. Ismail, M. Guadagnini, and K. Pilakoutas, "Strut-and-Tie Modeling of Reinforced Concrete Deep Beams," *Journal of Structural Engineering*, vol. 144, no. 2, p. 04017216, Feb. 2018, doi: https://doi.org/10.1061/(asce)st.1943-541x.0001974.

[24] H. Chen, W.-J. Yi, and Z. J. Ma, "Shear size effect in simply supported RC deep beams," *Engineering Structures*, vol. 182, pp. 268–278, Mar. 2019, doi: https://doi.org/10.1016/j.engstruct.2018.12.062.

[25] H. Bichri, A. Chergui, and M. Hain, "Investigating the Impact of Train/Test Split Ratio on the Performance of Pre-Trained Models with Custom Datasets," *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 2, 2024, doi: 10.14569/ijacsa.2024.0150235

[26] S. F. Senin, R. Rohim, and A. Yusuff, "Concrete Mix Design by Using the Artificial Neural Network Approach with Limited Datasets," *International Journal of Engineering Advanced Research*, vol. 4, no. 3, pp. 35–46, 2022

[27] V. R. Joseph, "Optimal ratio for data splitting," *Statistical Analysis and Data Mining: The ASA Data Science Journal*, vol. 15, no. 4, pp. 531–538, Apr. 2022, doi: https://doi.org/10.1002/sam.11583.

[28] I. Lishner and A. Shtub, "Using an Artificial Neural Network for Improving the Prediction of Project Duration," *Mathematics*, vol. 10, no. 22, p. 4189, Nov. 2022, doi: https://doi.org/10.3390/math10224189.

[29] J. Heaton, "Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning," *Genetic Programming and Evolvable Machines*, vol. 19, no. 1–2, pp. 305–307, Oct. 2017, doi: https://doi.org/10.1007/s10710-017-9314-z.

[30] K. G. Sheela and S. N. Deepa, "Review on Methods to Fix Number of Hidden Neurons in Neural Networks," *Mathematical Problems in Engineering*, vol. 2013, pp. 1–11, 2013, doi: https://doi.org/10.1155/2013/425740.

[31] M Priyadharshini *et al.*, "A population based optimization of convolutional neural networks for chronic kidney disease prediction," *Scientific Reports*, vol. 15, no. 1, Apr. 2025, doi: https://doi.org/10.1038/s41598-025-99270-8.

[32] Md. Sharafat Chowdhury, "Comparison of Accuracy and Reliability of Random Forest, Support Vector Machine, Artificial Neural Network and Maximum Likelihood method in Land use/cover Classification of Urban Setting," *Environmental Challenges*, pp. 100800–100800, Nov. 2023, doi: https://doi.org/10.1016/j.envc.2023.100800.

[33] C. B. Nayak *et al.*, "Experimental, analytical and numerical performance of RC beams with V-shaped reinforcement," *Innovative Infrastructure Solutions*, vol. 6, no. 1, Oct. 2020, doi: https://doi.org/10.1007/s41062-020-00363-2.

[34] J. Brownlee, *Deep Learning with Python: Develop Deep Learning Models on Theano and TensorFlow Using Keras*. Machine Learning Mastery, 2016.

[35] Salim Almaliki, Reza Alimardani, and Mahmoud Omid, "Artificial Neural Network Based Modeling of Tractor Performance at Different Field Conditions," *Agricultural Engineering International: CIGR Journal*, vol. 18, no. 4, pp. 262–274, 2016, Accessed: Apr. 07, 2026. [Online]. Available: https://cigrjournal.org/index.php/Ejounral/article/view/3880.

[36] S. Syaharuddin, F. Fatmawati, and H. Suprajitno, "Best Architecture Recommendations of ANN Backpropagation Based on Combination of Learning Rate, Momentum, and Number of Hidden Layers," *JTAM (Jurnal Teori dan Aplikasi Matematika)*, vol. 6, no. 3, p. 629, Jul. 2022, doi: https://doi.org/10.31764/jtam.v6i3.8524.

[37] H. Cai, T. Chen, W. Zhang, Y. Yu, and J. Wang, "Efficient Architecture Search by Network Transformation," *UCL Discovery (University College London)*, vol. 32, no. 1, Apr. 2018, doi: https://doi.org/10.1609/aaai.v32i1.11709.

https://doi.org/10.24191/esteem.v22iMarch.9841

[38] Y. Bengio, "Practical Recommendations for Gradient-Based Training of Deep Architectures," *Lecture Notes in Computer Science*, vol. 7700, pp. 437–478, 2012, doi: https://doi.org/10.1007/978-3-642-35289-8_26.

[39] S. Geman, E. Bienenstock, and R. Doursat, "Neural Networks and the Bias/Variance Dilemma," *Neural Computation*, vol. 4, no. 1, pp. 1–58, Jan. 1992, doi: https://doi.org/10.1162/neco.1992.4.1.1.

[40] O. Bournez, J. Cohen, and A. Wurm, "A universal uniform approximation theorem for neural networks," in *Proc. 50th Int. Symp. Mathematical Foundations of Computer Science (MFCS)*, Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2025, pp. 29:1–29:20, doi: 10.4230/LIPIcs.MFCS.2025.29.

[41] Y. Zhang, Q. Liu, S. Song, and W. Zhang, "Theoretical analysis of overfitting phenomenon in Levenberg-Marquardt trained neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 7, pp. 2894–2906, 2022.

[42] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Transactions on Neural Networks*, vol. 5, no. 2, pp. 157–166, Mar. 1994, doi: https://doi.org/10.1109/72.279181.

[43] Z. Hu, J. Zhang, and Y. Ge, "Handling Vanishing Gradient Problem Using Artificial Derivative," *IEEE Access*, vol. 9, pp. 22371–22377, 2021, doi: https://doi.org/10.1109/access.2021.3054915.

[44] H. Yu, T. Xie, S. Paszczynski, and B. M. Wilamowski, "Advantages of Radial Basis Function Networks for Dynamic System Design," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 12, pp. 5438–5450, Dec. 2011, doi: https://doi.org/10.1109/tie.2011.21647

https://doi.org/10.24191/esteem.v22iMarch.9841