Connection Design of Precast Concrete Structures Using Machine Learning Techniques

Nitin Dahiya\textsuperscript{1*}, Babita Saini\textsuperscript{2}, H.D. Chalak\textsuperscript{3}

1 Ph.D. Student, Faculty of Civil Engineering, National Institute of Technology Kurukshetra, Haryana, India
2. Associate Professor, Faculty of Civil Engineering, National Institute of Technology Kurukshetra, Haryana, India
3. Assistant Professor, Faculty of Civil Engineering, National Institute of Technology Kurukshetra, Haryana, India

Corresponding author: nitindahiya17@gmail.com
https://doi.org/10.22115/SCCE.2023.356547.1506

ARTICLE INFO

Article history:
Received: 17 August 2022
Revised: 25 January 2023
Accepted: 04 April 2023

Keywords:
Machine learning;
Gradient boosting;
Support vector machines;
Precast concrete structures;
Computer programming.

ABSTRACT

In this research, the number of dowels (horizontal connection) has been determined using support vector machines (SVM), gradient boosting and artificial neural networks (ANN-Multilayer perceptron). Building height, length and thickness of the wall, maximum shear, maximum compressive force and maximum tension were the input parameters while the output parameter was the number of dowels. 1140 machine learning models were used, out of which 814 were used as training datasets and 326 as test datasets. A coefficient of correlation of 0.9264, root mean square error of 0.3677 and scattering Index of 4.75 % was achieved by SVM radial basis kernel function (SVM-RBF) as compared to a coefficient of correlation of 0.9232, root mean square error of 0.3743 and scattering Index of 4.83 % by resilient ANN-Multilayer perceptron, suggesting that SVM-RBF is more accurate in estimating the number of dowels. The study's encouraging findings highlight the need for additional research into the use of machine learning in civil engineering.
1. Introduction

The precast structure includes the analysis and design of members, the method of fabrication, and the various aspects including the members lifting, transporting and erecting the members on site. The analysis and design require an examination of the strength and stability of the structure. Apart from design, the casting of members, transportation, and temporary stability of the structure during construction has been also considered.

The following aspects have been considered during the design stage for achieving good-quality precast concrete elements:

- Dimension and profile of precast members.
- Member joints and connections.

1.1. Dimensions and profile of precast members

The size of precast members mainly depends on the lifting capability of the lifting crane at the casting area and site. The designer attempts to provide the largest size of elements to minimize joints and handling at the site. Casting and construction quality has been achieved by the combination of different members, such as shear wall system, 3D units and beam-column system. The design has been regulated in such a way allowing for maximum replication in the dimension and profile of the members. This is an advantageous factor for achieving economic and quality assurance as fewer steel mould changes will make possible the construction agenda, thereby reducing production time.

1.2. Joints & connections

The design of connections is to effectively transfer the loads from members to the structure or a neighbouring member. Strength, Stability, ductility, durability and fire resistance are important considerations considered in the design of joints. Design criteria not only consider the production of elements but also consider the erection process of the element as well. The joints and connections are designed considering the standardization of joints, avoiding congestion, providing erection tolerances and easy assembly of joints. The joint and connection provided should be practical, feasible and serviceable. The joint and connection provided should be at the logical location, with no local stresses and it can easily transfer all loads considering the factor of safety. A typical precast connection is shown in Figure 1 and a connection designed as a dowel for shear, tension & compression is shown in Figure 2.

2. Literature

The study of precast concrete structures and their connections has been the focus of research. The specifications for precast buildings are not covered by the current IS codes. The available related guidelines also don't reflect the most recent theories and methods used by the global precast concrete industry, which makes the engineering community wary when examining the design strategies used by the developing precast sector in India. Benjamin [1] focused on the variability analysis of shear wall structures, which are highly variable in terms of rigidity and strength.
The transfer matrix method was used by Bozdogan et al. [2] to analyse vibrations of asymmetric shear wall structures and the governing differential equations for the equivalent bending-warping torsion beam are developed using the continuum approach. Using openSEES, a numerical analysis of cyclic loading tests on shear walls was done by Xiaolei et al. [3]. Carpinteri et al. [4] investigated shear wall structures and the effects of lateral loads on the structure of various heights. The comparison of finite element solutions was evaluated, in which the bracings were modelled as 3-D structures. Biswas et al. [5] conducted a three-dimensional analysis of a multi-story shear wall building and investigated the role of torsion in multi-story buildings with shear wall layouts. Fahjan et al. [6] investigated the non-linear analysis for reinforced concrete buildings with shear walls. R. Vidjeapriya et al. [7] conducted an experimental study on the
performance of two dry connections, with double stiffener and single stiffener. These models were compared to monolithic connections and tested under reverse cyclic loading, with energy dissipation, load-carrying capacity, and ductility. Maya et al. [8] proposed a connection for precast construction of beam-column to attain a shorter splice length using ultra-high-performance fibre-reinforced concrete (UHPFRC). Shariatmadar et al. [9] tested precast beam-to-column connections with different detailing of three full-scale buildings, i.e., U-shaped spliced with steel plates, U-shaped spliced, straight spliced, under reverse cyclic loading. Blaz Zoubek et al. [10] investigated the capacity of dowel connections in beam-column joints, the failure types observed were a local failure with crushing of the surrounding concrete and dowel bar yielding, and a global failure with spalling of concrete between the edge and dowel. In the column-to-foundation joint, bolted precast connections under cyclic loading were studied by Elena Camnasio et al. [11]. The proposed anchor bolt connection showed acceptable ductility and stiffness, and it was recommended that more research be done on the cyclic performance of beam-column connections. Douglas J. Provost-Smith et al. [12] tested connections with the grouted dowel to determine bond strength between dowel connections used in precast wall construction. To determine the bond strength at a grouted dowel connection, pull-out tests were performed. The steel duct created a greater confinement effect, resulting in shear pull-out failure rather than split tensile failure. The absence of ducts can reduce connection strength by 30%, and the effect of eccentric bar placement was also investigated. Ibrahim M.H. et al. [13] studied the most recent developments in strengthened beam-column and beam-column-slab connections made of precast concrete. The goal of the study is to offer insightful information to researchers, designers, and specialists who are interested in researching the progressive collapse of precast concrete (PC) structures.

The various approaches for linear and nonlinear modelling of shear walls in structural analyses of buildings are investigated and applied to RCC buildings with shear walls.

The research work done by many researchers has made it possible to develop analysis design criteria for precast concrete structures and their connections. In India, the design of precast concrete structural systems is in the developing stage. The precast structure has to emulate the performance and behaviour of cast-in-place reinforced concrete structure under seismic loading and the vertical and horizontal connections are designed. Machine learning techniques such as artificial neural networks, support vector machines and gradient boosting can play an important role in predicting the number of dowels required in connection with the precast member. Pal and deswal [14], and Asteris et al. [15–19] used various machine-learning techniques in civil engineering problems.

3. Research significance

In this research, models are first made in Etabs, data were extracted from all the models including axial load, shear, torsion and moment. An excel sheet was prepared for the calculation of the number of dowels required based on compressive strength and tensile strength capacity. The input parameters including the building height, length and thickness of the wall, maximum shear, compressive force and tension in the wall were fed into the machine learning models to predict the number of dowels using computer programming based on Rstudio [20]. The goal of
the study is to offer insightful information to researchers, designers, and specialists who are interested in researching the connection design of precast concrete structures using machine learning techniques.

4. Machine learning methods

Support vector machines are supervised learning techniques used in regression and classification. SVM was developed by Vapnik [21] in 1995 in the UK. SVM uses supervised learning models and learning algorithms to analyse data for classification and regression.

An SVM model is an abstraction of the examples as a point in space, traced in such a way that the examples of the various categories are separated by as wide of a gap as possible. A variation of the support vector called SVM employs kernels to generate non-linear boundaries. A kernel is a similarity function between two observations. There are major three kernel functions- linear, polynomial and radial. The linear kernel takes the inner product of two observations, the polynomial kernel used the power function to create non-linear boundaries, and the radial kernel uses the radial function to create radial boundaries. In the support vector machine model following hyper tuning parameters are required: Cost, degree and gamma. When the cost is low, margins will be large, there will be a large number of support vectors, and a large number of observations will be incorrectly classified. A high cost will result in narrow margins, fewer support vectors, and fewer misclassified values. Low-cost value, however, prevents overfitting and might result in better test performance. The degree will determine the flexibility of the polynomial boundary and gamma is defined as how much influence a single training example has at any given point. The flow chart of the SVM algorithm is shown in figure 3. A reference is made for an extensive analysis of SVM in Nitin Dahiya et al. [22].

Gradient boosting is an ensemble technique for regression and classification that was proposed by Friedman (1999) [23]. Gradient boosting is a slow learning procedure in which a tree is fit using current residuals rather than the outcome as the response. Each iteration of the randomly selected training set is validated from the base model in the gradient boosting process. Randomly, a sub-sample of training data may enhance gradient-boosting execution speed and precision and aid in avoiding overfitting. A reference is made for an extensive analysis of gradient boosting in Nitin Dahiya et al. [24].

ANN-Multilayer perceptron is called a neural network with multiple hidden layers. A perceptron is a single neuron model that is a precursor to a larger neural network that uses backpropagation to learn a multi-layer perceptron. It can be developed manually or by using a simple heuristic setup therefore, it is possible to monitor and modify the parameters in the training period. There are only sigmoid nodes in this network. Batch size, momentum, learning rate and hidden layer, are the parameters required in multilayer perceptron regression. A reference is made for an extensive analysis of neural networks in Nitin Dahiya et al. [25].
The current study examined 5 bay models of precast buildings with varying building heights (Table 1). The input parameters including the building height, length and thickness of the wall, maximum shear, compressive force and tension in the wall were fed into the machine learning models. To calculate the outcome of input parameters on the number of dowels in the horizontal connection of structures, 1140 data were prepared. ETABS [26] 2018 is used to model and analyse the structure. The PEER database was used to get nonlinear ground motion data. For the analytical research, an ensemble of 100 distinct ground motion data was used. Bispec software is used to combine these ground motions. The adopted ground motion records have only used near-fault data to obtain effective variance in the seismic response parameters. The three-dimension model of the structure is shown in Figure 4.
Rstudio was used on 70 per cent of the 1140 models to acquire the split data needed for training. In this research, 814 and 326 samples were chosen at random to obtain split data for training and testing, respectively. The input variable was the height of the building, length and thickness of the wall, maximum shear, compressive force and tension in the wall. Table 2 presents a summary of the training and test data. Scattering index (SI), root mean square error (RMSE), mean absolute percentage error (MAPE), mean absolute error (MAE) and correlation coefficient (CC) were used to compare the efficacy of different techniques using test data. The tuning parameters regulate the efficacy of ANN-Multilayer perceptron, SVM and gradient boosting. The ideal value of the tuning parameters was determined through trail runs by comparing the SI, MAE, MAPE, RMSE and CC. The ideal value of the tuning parameters employed in this investigation is shown in Table 3. The performance of support vector machines (SVM), gradient boosting and multilayer neural network algorithms were compared and validated using the ANOVA single-factor test.

A single-factor test also known as ANOVA is a one-way tool which is employed to obtain a number of mean values for various equivalent clusters. It is a statistical testing method that uses hypothetical data. F-critical, p-value and F-value were obtained with this technique, which is significant for the result. F-statics are used to derive F-critical and F-value. The difference between the groups is statistically significant if the F-value is greater than the F-critical.

**Table 1**

<table>
<thead>
<tr>
<th>Building parameters</th>
<th>Building Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bays (Nos.)</td>
<td>5</td>
</tr>
<tr>
<td>Type of Frame</td>
<td>Shear wall with the opening</td>
</tr>
<tr>
<td>Length (in m)</td>
<td>8, 12, 16, 20</td>
</tr>
<tr>
<td>Breadth (in m)</td>
<td>3, 6, 9, 12</td>
</tr>
<tr>
<td>Height (in m)</td>
<td>9, 18, 27, 36, 45, 54</td>
</tr>
<tr>
<td>Column (in mm)</td>
<td>300 x 600</td>
</tr>
<tr>
<td>Beam (in mm)</td>
<td>300 x 450</td>
</tr>
<tr>
<td>Shear Wall thickness (in mm)</td>
<td>200 and 150</td>
</tr>
<tr>
<td>Grade of Concrete</td>
<td>M 30</td>
</tr>
</tbody>
</table>
Table 2
Training and testing data set summary.

<table>
<thead>
<tr>
<th>Input</th>
<th>Training set</th>
<th></th>
<th></th>
<th>Testing set</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min</td>
<td>Max</td>
<td>Mean</td>
<td>Std. dev.</td>
<td>Min</td>
<td>Max</td>
</tr>
<tr>
<td>Length (L) in m</td>
<td>2.8</td>
<td>3.8</td>
<td>3.301</td>
<td>0.49</td>
<td>2.8</td>
<td>3.8</td>
</tr>
<tr>
<td>Thickness (T) in m</td>
<td>0.15</td>
<td>0.2</td>
<td>0.1749</td>
<td>0.025</td>
<td>0.15</td>
<td>0.2</td>
</tr>
<tr>
<td>Height (H) in m</td>
<td>9</td>
<td>54</td>
<td>22.9</td>
<td>2</td>
<td>12.95</td>
<td>9</td>
</tr>
<tr>
<td>Max. Shear (MS) in kN</td>
<td>10.99</td>
<td>530.70</td>
<td>147.18</td>
<td>103.12</td>
<td>13.02</td>
<td>515</td>
</tr>
<tr>
<td>Max. Compressive Force (MC) in kN</td>
<td>0.00</td>
<td>851.66</td>
<td>130.90</td>
<td>169.81</td>
<td>0.00</td>
<td>851.33</td>
</tr>
<tr>
<td>Max. Tension (MT) in kN</td>
<td>-848.64</td>
<td>-2.647</td>
<td>-137.389</td>
<td>167.93</td>
<td>-849</td>
<td>-3.342</td>
</tr>
</tbody>
</table>

Table 3
Optimum value of tuning parameters.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Tuning parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN-Multilayer perceptron</td>
<td>Layer dense- 2 (8, 5 nodes), epochs-1500, learning rate-0.6, momentum-0.2 Batch size-20.</td>
</tr>
<tr>
<td>SVM-RBF</td>
<td>Cost – 20, Gamma- 0.8, Epsilon- 0.1</td>
</tr>
<tr>
<td>SVM-Poly</td>
<td>Cost – 10, Degree- 3, Gamma- 0.16, Epsilon- 0.1</td>
</tr>
<tr>
<td>Gradient Boosting</td>
<td>Interaction Depth = 7, Shrinkage rate = 0.06, n.trees = 216</td>
</tr>
</tbody>
</table>

6. Results

The scattering index (SI), root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE) and correlation coefficient (CC) were obtained with test data using support vector machines, gradient boosting, and an ANN multilayer perceptron are shown in Table 4. Table 4 indicates that SVM-RBF has improved accuracy and efficiency than gradient boosting, SVM-Poly and ANN multilayer perceptron for SI, MAE, MAPE, RMSE and CC. The optimum value of Cost- 20, Gamma- 0.8, Epsilon- 0.1, and the hyper-tuning parameters were obtained by the trial-and-error method. In Figure 5, the performance of SVM-RBF has been presented in gamma and cost. Table 4 shows that the SVM-RBF outperforms the gradient boosting, SVM-Poly, and ANN multilayer perceptron in terms of precision and efficacy in predicting the number of dowels. Figures 6, 7, 8 and 9 provide the plots between the actual numbers of dowels vs predicted numbers of dowels using test data of SVM-RBF, gradient boosting, SVM-Poly and ANN multilayer perceptron algorithms. Figures 6, 7, 8 and 9 show SVM-RBF, gradient boosting, SVM-Poly and ANN multilayer perceptron are good performers for numbers of dowels values in the lower and middle ranges.
Table 4
SI, RMSE, MAE, MAPE, and CC values with test data.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Test Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CC</td>
</tr>
<tr>
<td>ANN-Multilayer perceptron</td>
<td>0.9232</td>
</tr>
<tr>
<td>SVM-RBF</td>
<td>0.9264</td>
</tr>
<tr>
<td>SVM-Poly</td>
<td>0.9118</td>
</tr>
<tr>
<td>Gradient Boosting</td>
<td>0.9087</td>
</tr>
</tbody>
</table>

Fig. 5. Performance of SVM-RBF.

Fig. 6. Actual versus predicted dowels using SVM-RBF.
Fig. 7. Actual versus predicted dowels using Gradient Boosting.

Fig. 8. Actual versus predicted dowels using SVM-Poly.

Fig. 9. Actual versus predicted dowels using ANN.
Table 5
Test result of ANOVA.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Modelling Technique</th>
<th>p-value</th>
<th>F-value</th>
<th>F-critical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Testing</td>
<td>ANN</td>
<td>0.5823</td>
<td>0.3027</td>
<td>3.855</td>
</tr>
<tr>
<td></td>
<td>SVM-RBF</td>
<td>0.6659</td>
<td>0.1865</td>
<td>3.855</td>
</tr>
<tr>
<td></td>
<td>SVM-Poly</td>
<td>0.9164</td>
<td>0.0110</td>
<td>3.855</td>
</tr>
<tr>
<td></td>
<td>Gradient Boosting</td>
<td>0.8799</td>
<td>0.0228</td>
<td>3.855</td>
</tr>
</tbody>
</table>

The test was conducted on the variance between the actual and predicted number of the dowels using ANOVA statistical test and the results are depicted in Table 5. According to Table 5, the entire machine learning model was validated because the p-value was higher than 0.05 and the F-critical was higher than the F-value. Table 4 suggests that the SVM-RBF regression approach performs better than any other approaches, SVM-RBF regression has a CC of 0.9264, RMSE of 0.3677, MAPE of 0.0197, MAE of 0.1631 and SI of 4.75 per cent.

The CC, RMSE, MAPE, MAE, and SI (%) on all data sets were calculated using the SVM-RBF model using the same tuning parameter. Table 6 demonstrates unequivocally that all data sets fit the data better than the test data set.

Table 6
Testing and all data sets.

<table>
<thead>
<tr>
<th>SVM-RBF</th>
<th>CC</th>
<th>RMSE</th>
<th>MAPE</th>
<th>MAE</th>
<th>SI (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test data</td>
<td>0.9264</td>
<td>0.3677</td>
<td>0.0197</td>
<td>0.1631</td>
<td>4.75</td>
</tr>
<tr>
<td>All data</td>
<td>0.9713</td>
<td>0.2380</td>
<td>0.0124</td>
<td>0.1023</td>
<td>3.07</td>
</tr>
</tbody>
</table>

7. Sensitivity analysis

To understand the significance of the input variable significance in predicting the dowels number requirement, sensitivity analysis was performed. The training data is combined with the kernel function for analysis. By removing each input parameter, SVM-RBF is used on a different set of training data:

- The first iteration used all the input parameters,
- In the second iteration, max. tension (MT) was removed,
- In the third iteration, max. compressive force (MC) is removed,
- In the fourth iteration, max. shear (MS) is removed,
- In the fifth iteration, height (H) is removed,
- In the sixth iteration, wall thickness (T) is removed,
- In the seventh iteration, the length of the wall (L) is removed,
- For the final iteration, the length of the wall (L) and wall thickness (T) is removed and results for all eight iterations were obtained and analysed in form of CC, RMSE, MAPE, MAE and SI (%) are shown in table 7.

Results in table 7 show that the length (L), thickness (T) and height (H) have the least impact in predicting the number of dowels whereas, max shear (MS), max compressive force (MC) and max tension (MT) has a major influence in predicting the number of dowels.
Table 7
Sensitivity Analysis.

<table>
<thead>
<tr>
<th>Input Parameter</th>
<th>CC</th>
<th>RMSE</th>
<th>MAPE</th>
<th>MAE</th>
<th>SI (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L, T, H, MS, MC, MT</td>
<td>0.9264</td>
<td>0.3677</td>
<td>0.019</td>
<td>0.1631</td>
<td>4.75</td>
</tr>
<tr>
<td>L, T, H, MS, MC</td>
<td>0.8549</td>
<td>0.5044</td>
<td>0.020</td>
<td>0.1833</td>
<td>6.58</td>
</tr>
<tr>
<td>L, T, H, MS, MC, MT</td>
<td>0.8678</td>
<td>0.5057</td>
<td>0.022</td>
<td>0.1931</td>
<td>6.52</td>
</tr>
<tr>
<td>L, T, H, MC, MT</td>
<td>0.7891</td>
<td>0.5950</td>
<td>0.025</td>
<td>0.2152</td>
<td>7.74</td>
</tr>
<tr>
<td>L, T, MS, MC, MT</td>
<td>0.9228</td>
<td>0.3743</td>
<td>0.018</td>
<td>0.1532</td>
<td>4.85</td>
</tr>
<tr>
<td>L, H, MS, MC, MT</td>
<td>0.9266</td>
<td>0.3673</td>
<td>0.019</td>
<td>0.1623</td>
<td>4.75</td>
</tr>
<tr>
<td>T, H, MS, MC, MT</td>
<td>0.9267</td>
<td>0.3671</td>
<td>0.019</td>
<td>0.1622</td>
<td>4.74</td>
</tr>
<tr>
<td>H, MS, MC, MT</td>
<td>0.8030</td>
<td>0.5978</td>
<td>0.052</td>
<td>0.4121</td>
<td>7.75</td>
</tr>
</tbody>
</table>

8. Conclusions

In the present study, computer programming based on Rstudio was developed for support vector machines (SVM), ANN-Multilayer perceptron and gradient boosting to predict the number of dowels required in a precast concrete structure.

- The obtained results show that SVM-RBF works well in comparison to other machine learning techniques for the data set, SVM-RBF provides the best outcomes- CC of 0.9264, RMSE of 0.3677, MAPE of 0.0197, MAE of 0.1631 and SI of 4.75 per cent for predicting the number of dowels required in precast concrete structures.
- From the sensitivity analysis, length (L), thickness (T) and height (H) have the least impact in predicting the number of dowels whereas, max shear (MS), max compressive force (MC) and max tension (MT) has a major influence in predicting the number of dowels.
- ANOVA test was performed on all machine learning techniques, the p-value was greater than 0.05, and the F-critical was greater than the F-value, which means the results are statistically significant.
- It can also be concluded from this study that machine learning techniques are robust and accurate modelling approaches and need to be used further in structural and civil engineering problems.

References


