Prediction of Concrete Breakout Strength of Single Anchors in Shear

Oladimeji B. Olalusi and Panagiotis Spyridis

Abstract This study proposes a machine learning algorithim—a Gaussian process regression (GPR)—for predicting the concrete breakout capacity of single anchors in shear. To this end, experimental strength of 366 tests on single anchors with concrete edge breakout failures were collected from literature to establish the experimental database to train and test the model. 70% of the data were used for the model training, and the rest were used for the model testing. Shear influence factors such as the concrete strength, the anchor diameter, the embedment depth (technically the influence length), and the concrete edge distance were taken as the model input variables. The generated predictive model yielded a determination coefficient $R^2 = 0.99$ for both the training and testing data sets. Predictions from the developed models were compared to that of the other existing models (Eurocode 2 and ACI 318) to validate its performance. The developed model provided a better prediction of the experimentally observed shear strength, compared to the existing models, yielding low mean absolute error, low bias and variability when tested.

Keywords Machine learning · Gaussian process regression · Breakout strength · Fasteners · Model uncertainty

1 Introduction

Recent advances in computing have given a boost in the use of advanced soft computing methods across all industries; the field of artificial intelligence has been developing since the 1950s. Knowledge-based methods or expert systems have been deployed to assist risk-related decisions under uncertainty. Applications in structural engineering are witnessed in the research field as early as the 1980s [1] since structural engineering problems in practice are governed by a wide range of uncertainties

O. B. Olalusi
University of KwaZulu-Natal, Durban, South Africa

P. Spyridis (✉)
Technical University of Dortmund, Dortmund, Germany
e-mail: panagiotis.spyridis@tu-dortmund.de

related for example to the applied actions, the material performance and homogeneity, or the models used to describe the problem itself. A benefit of such soft computing methods is that they can provide reliable solutions on multi-parametric problems and highly nonlinear correlations of the input. Simultaneously, various disciplines of structural engineering rely on empirical, semi-empirical, or numerical predictive models.

Design of anchorage to concrete, offers an example of semi-empirical predictive and design models, particularly when it comes to concrete-related failure modes. The assessment of anchor capacity using Artificial Neural Networks (ANN) has been previously investigated in [2, 3]. More recent studies present paradigms of machine learning algorithms such as the Gaussian process regression (GPR) with very high efficiency in structural engineering problems [4, 5]. The present paper investigates the feasibility of using GPR algorithms to predict the concrete breakout strength of single anchors loaded in shear. Towards developing the model, experimental results of 366 tests on single anchors with concrete edge breakout failures were collected based on an extensive literature search to establish the database to train and test the model. A parametric study and comparison of the proposed models with other existing predictive models were reported to assess the accuracy and efficiency of the shear capacity design methods for anchors subjected to shear loads.

2 Existing Strength Models for Concrete Breakout Capacity of Single Anchors in Shear

The current EN 1992-4 [6] and ACI 318 [7] design standards provide Eqs. 1 and 2, respectively, for the evaluation of a single anchor's resistance against concrete edge failure in non-cracked concrete. The European design is valid for a $d_{nom} \leq$ 60 mm and influence length $l_f \leq 12d_{nom}$ in case of $d_{nom} \leq 24$ mm and otherwise $\leq \max\{8 \cdot d_{nom}; 300$ mm$\}$. The design calculations covered by CEN are valid only up until $f_{ck} < 60$ N/mm^2. ACI 318 is valid for a concrete compressive cylinder strength of 10,000 psi (70 MPa) for cast in anchors and 8000 psi (55 MPa) for post-installed anchors and an anchor diameter up to 4in. (100 mm).

$$V_{EC2-k} = 2.4.d_{nom}^{\alpha}.l_f^b.\sqrt{f_{ck}}.c_1^{1.5}. \tag{1a}$$

The mean concrete breakout capacity of single shear anchor in non-cracked concrete for the EN 1992-4 standard is calculated according to Eq. 1b [8].

$$V_{EC2-m} = 3.d_{nom}^{\alpha}.l_f^{\beta}.\sqrt{f_{cm}}.c_1^{1.5}. \tag{1b}$$

where

$$\alpha = 0, 1 \cdot \left(\frac{l_f}{c_1}\right)^{0,5}$$

$$\beta = 0, 1 \cdot \left(\frac{d_{nom}}{c_1}\right)^{0,2}$$

The design and the mean concrete breakout capacity of single shear anchor in non-cracked concrete for ACI 318 are calculated according to Eq. 2a, b respectively [9, 10].

$$V_{ACI} = \min\left\{\left(0.6 \cdot \left(\frac{l_f}{d_{nom}}\right)^{0.2} \cdot \sqrt{d_{nom}}\right)\lambda_a\sqrt{f_c'}(c_1)^{1.5}, 3.7.\lambda_a.\sqrt{f_c'}(c_1)^{1.5}\right\} \quad (2a)$$

$$V_{ACI-m} = \min\left\{\left(\frac{l_f}{d_{nom}}\right)^{0.2} \cdot \sqrt{d_{nom}}\sqrt{f_{cm}}(c_1)^{1.5}, 7.1.\sqrt{f_{cm}}(c_1)^{1.5}\right\} \quad (2b)$$

where d_{nom} is the outside diameter of the anchor. f_c' is the concrete cylinder strength per the ACI acceptance standards. λ_a is the modification factor for applications in lightweight concrete. f_{cm} is the mean concrete cylinder compressive strength. l_f is the influence length of the anchor loaded in shear.

3 Gaussian Process Regression (GPR)

The Gaussian process model is a kernel machine type which can be used as a supervised learning technique for classification as well as regression. Gaussian processes can give a simple probabilistic representation of random processes and can be used for many different types of nonparametric estimation. This method is currently well adopted and applied in various areas in structural engineering [4, 5]. A summary of the GPR algorithm is presented in this section. More details on the GPR methodology can be found in [11].

Given a training set, $U = \{(x_i, y_i); i = 1, 2, \ldots, n\}$, where the input $x_i \in \mathbb{R}^{U.n}$ denotes the design matrix and $y_i \in \mathbb{R}^n$ denotes the vector of the desired output, drawn from an unknown distribution. A GPR model predicts the value of the output variable y_{new}, given a new input vector x_{new}, and training data. In the setting of classic linear regression, we model the output variable y by a function of an input variable x expressed in Eq. 3 [11].

$$y = x^T\beta + \varepsilon \quad (3)$$

where x is the input vector, and y is the observed target value. The random error term $\varepsilon \sim N\left(0, \sigma_n^2\right)$. The error variance σ_n^2 and the coefficients β are estimated from the data.

The multivariate Gaussian distribution, which has a mean vector μ and covariance matrix Σ have the joint probability density expressed as Eq. 4.

$$p(x|\mu) = (2\pi)^{-\frac{D}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right) \tag{4}$$

Unlike the Gaussian distribution, which is a distribution over vectors, the Gaussian process is a distribution over functions with a covariance function $k(x, x')$ and a mean function $m(x)$ (Eq. 5).

$$f(x) \sim GP\big(m(x), k(x, x')\big) \tag{5}$$

The indexes of the GP is x. Where the mean function and covariance function of a real process $f(x)$ is defined as Eqs. 6 and 7, respectively.

$$m(x) = E[f(x)] \tag{6}$$

$$k(x, x') = E\big[(f(x) - m(x))(f(x') - m(x'))\big] \tag{7}$$

The Gaussian Process is a multivariate Gaussian of infinite length. Following the GPR procedure, the n observations in an arbitrary data set, $y = \{y_1, \ldots, y_n\}$ can be taken as a sample from some multivariate Gaussian distribution. Hence, going from the process of distribution, we can get an understanding of a GP and then draw samples from it. The Gaussian process $f \sim GP(m, k)$ is defined with a mean function $m(x) = 0$ (Eq. 8) and covariance/kernel function $k(x, x')$ (Eq. 9). The goal of only working with finite quantities is simply achieved by requiring the values of f at a distinct number of n locations. Given the x-values we can evaluate the GP, which is now reduced to a multivariate Gaussian distribution [11].

$$\mu = m(x) = 0 \tag{8}$$

$$k(x, x') = \sigma_f^2 \exp\left(-\frac{1}{2l^2}(x - x')^2\right) + \sigma_n^2 \delta(x, x') \tag{9}$$

where l denotes the length parameter of the kernel function. $\delta(x, x')$ is denotes the Kronecker delta function.

Gaussian Process regression has different types of kernel functions, some of which includes the squared exponential kernel, Laplace kernel and Linear Kernel. Since different kernel functions are suitable for different type of data, several kernel functions need to be trialled to choose the most appropriate. In this study, the two most suitable kernel functions (non-linear kernel functions) obtained for the database is the Gaussian or Radial Basis Function (RBF) $k(x, x') = \exp\left(-\frac{1}{2\sigma^2}x - x'^2\right)$. Where σ is the width of the kernel which are user-defined parameters.

4 Development of the GPR Model

The experimental database considered in this investigation was compiled from different research published in technical literature. They are majorly experiments conducted by [10, 12–22]. The different experiments were aimed and designed to reflect the behaviour of single anchors in shear. The experimental database consists of the failure load of 366 single anchors in shear, failing due to concrete edge breakout in non-cracked concrete. The database covers a wide range of anchor configurations that can be used to assess any anchor design method against experimental results.

The development of an efficient model for predicting the concrete breakout strength of single anchors in shear requires the inclusion of the main factors affecting anchors in shear as inputs. The various parameters affecting the breakout strength of single anchors in shear are discussed in [9]. In this study, the input parameters considered for the implementation of the GPR model includes edge distance c_1, anchor diameters d_{nom}, influence length l_f and concrete strength f_c.

In order to implement the GPR model, the database of experimental anchor tests was split into two subsets, namely: the training data set, and testing data set. The training data set is used to develop both the GPR model, whereas the performance of the developed model is evaluated using the testing dataset. While splitting the database into subsets, it is essential to ensure data patterns that are statistically consistent in both the training and testing datasets. This was achieved by randomly sorting both the training and testing data set until an acceptable consistency is maintained among the input variables, in terms of statistical properties (such as mean and standard deviation) and range of data. This is summarised in Table 1. In this study, 70% of the data (256 out of 366 cases) were used for training, and the remainder (110 cases) were used for testing the models.

The nonlinear regression technique of the GPR models, implemented in a MATLAB environment, was used to predict the concrete breakout strength of anchors in shear. In order to map input data into feature space, nonlinear regression technique requires kernel functions. The optimum search method was used to obtain the optimum parameters. The performance of the developed model was detailed using statistical parameters, namely the coefficient of determination (R^2), mean absolute percentage error (MAPE), Mean square error (MSE) and root-mean-squared error (RMSE).

5 Results and Discussion

5.1 Performance of the GPR Model

The investigation of the performance of the developed model using the training and testing data set is discussed in this section. Using the experimental data presented in Table 1, the GPR model was adopted to successfully learn the interrelationships

Table 1 Statistical summary of the experimental dataset used for the model development

Statistical parameter	Datasets	d_{nom} (mm)	l_f (mm)	c_1 (mm)	f_c (MPa)	V_u (kN)
Range	Training	8–88.9	50–762	45–508	18.4–85.4	4.72–511.27
	Testing	8–88.9	50–762	50–508	18.4–85.4	5.17–518.72
Mean	Training	25.05	206.7	148.1	27.8	71.6
	Testing	25.1	234.7	160.3	28.2	89
SD	Training	15.2	150.3	107.6	9.9	97.4
	Testing	17.7	175.2	126.2	10.4	127.9

SD denotes the standard deviation, V_u is the experimental concrete breakout strength

Table 2 Statistical properties of the developed models

Parameters	GPR model	
Dataset	Training data set (256)	Testing data set (110)
R^2	0.99	0.99
RMSE	8.7	11.1
MSE	76.7	123.9
MAE	5.8	5.6

between concrete breakout strength and varied shear strength parameters (input variables). The model accuracy is assessed using statistical parameters such as R^2, RMSE, MAE and MSE, calculated between the experimental and predicted results, and the results presented in Table 2. The statistical result for the training database (256 experimental tests) is reported as follows: RMSE = 8.7; MAE = 5.8 and R^2 = 0.99. The R^2 and MAE value of the testing database is comparable to that of the training database (110 experimental tests), but with a slightly higher RMSE value. These results indicate that the GPR model is a good predictor of concrete breakout strength. The results demonstrate the generalization capability of the developed GPR model.

5.2 Comparative Study of the GPR with Existing Concrete Breakout Strength Models

The trained GRP model was compared to existing concrete breakout capacity predictive models, using the testing dataset, to examine the predictive performance of the models. Two existing concrete breakout capacity models in shear were considered, namely the predictive model of (1) EN 1992-4 (Eq. 1) (2) ACI 318 (Eq. 2). The mean shear resistance function/best estimate model of the EN 1992-4 and ACI 318 models were used in this comparative assessment.

The plot of the experimental breakout capacity against the predicted breakout capacity of the developed models and other existing models, using the testing

database, are presented in Fig. 1. The figure portrays the deviation of the predicted strength by the methods from the line of 'Perfect model', which is defined as the line position of all the points where the experimental breakout capacity is equal to predicted breakout capacity. As seen from the figures, the predicted values by the GPR model are much less scattered compared to the values from the other models.

A comparison of the GPR model with other existing models in terms of MAE and RMSE values is presented in Table 3. The table revealed that the MAE and RMSE values of the GPR model is less than that of the other models. Güneyisi et al. [23] developed a model to predict the concrete edge breakout capacity of single adhesive anchors using gene expression programming (GEP). Their developed model yielded R^2 of 0.92 (lower than what is obtained for GPR model in this study) and the value of RMSE $= 13$; MAPE $= 14.2$, MSE $= 168.7$ (higher than what is achieved for GPR model in this study), using the testing database of 34 anchor experiments. However, it should be highlighted that a larger database is utilized for training and testing in this study compared to [23]. A general assessment of the results presented in Table 3 suggests that the GPR model outperformed all the other models investigated in this study.

The summary of the statistical properties of the model uncertainty (obtained as the ratio of experimental breakout strength to predicted breakout strength) associated to the GPR and the other models are presented in Table 3, and the distributions are plotted in a box plot shown in Fig. 2. A box plot is a statistical tool that can be used to provide statistical summaries of the underlying distribution of a dataset. The box plot displays the maximum and minimum values in the dataset, the lower and upper quartiles, the mean and the median. A model uncertainty mean value $\mu_{ME} = 1$ is a condition for an ideal model. The GPR and the other models are assessed, using the criteria that an ideal model is expected to have in addition to model uncertainty mean value $\mu_{ME} = 1$; high precision (that is, small scatter of data) [24, 25].

Assessment of the box plot revealed that the GPR model has the smallest length of the interquartile range of all the models investigated, thereby suggesting less variability of the GPR model uncertainty. As presented in Table 3, the model uncertainty variable associated with the GPR model has a mean value of $\mu_{ME} = 0.99$ (closest

Table 3 Statistical properties of the proposed and existing models

Parameters	GPR	EC2	ACI 318
R^2	0.99	0.98	0.98
RMSE	11.1	41.78	22.2
MSE	123.9	1747.2	196
MAPE	5.6	20.7	14.0
Mean μ_{ME}	0.99	0.86	0.93
Standard deviation σ_{ME}	0.11	0.18	0.23

μ_{ME} and σ_{ME} are respectively, the mean and standard deviation of the model uncertainty (ratio of the experimentally observed shear capacity and the predicted shear capacity by the models). R^2 is the coefficient of determination
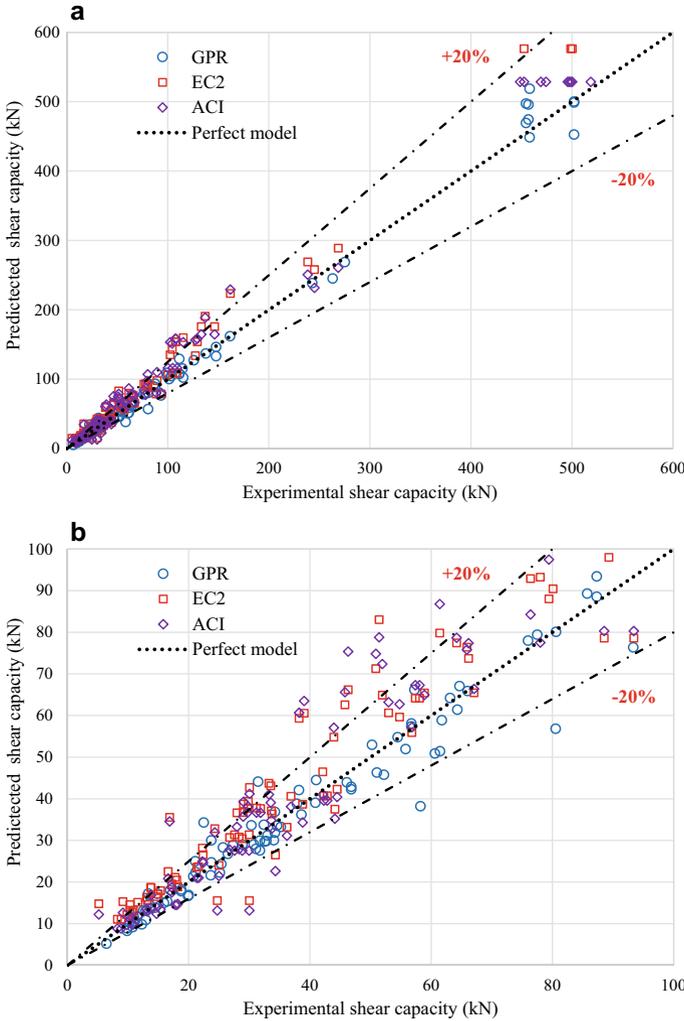
Fig. 1 **a** The full plot of experimental breakout capacity versus predicted breakout capacity. **b** Partial plot of experimental breakout capacity versus predicted breakout capacity (breakout strength up to 100 kN only)

to mean value of 1), suggesting that the model reasonably predicts the breakout strength. Regarding standard deviation, the GPR model yields the lowest dispersion of all the models investigated with $\sigma_{ME} = 0.11$. Table 3 also shows that the best estimate models of EN 1992-4 and ACI 318 overpredict the shear breakout strength, but predictive models in design standards may typically lie on the conservative side.
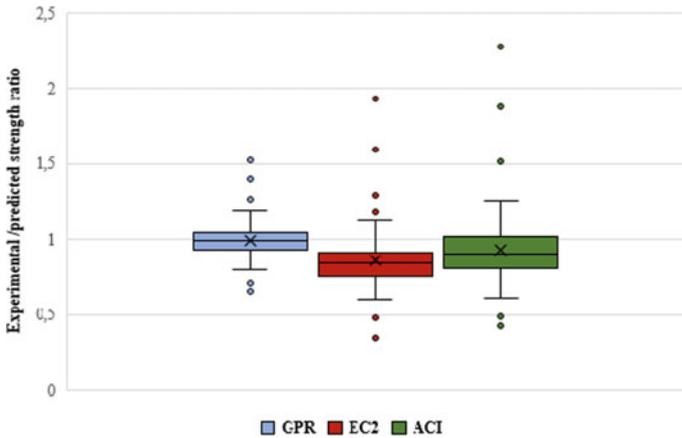
Fig. 2 Comparison of the prediction error of GPR to other models

6 Conclusions

A possible failure mode for shear loaded anchors is the concrete breakout failure. Concrete related failure mode poses a significant safety issue, since they may develop abruptly, without preceding signs of damage. Given this, accurate prediction of the concrete breakout resistance of anchors in shear is crucial. This contribution focuses on the feasibility of using the Gaussian Process Regression (GPR) machine learning algorithms to predict the shear breakout strength of single anchors and quantifies the model uncertainties in existing predictive models. The following general conclusions may be drawn from the present study:

- A reasonable accuracy was obtained for both the training and testing datasets in terms of low RMSE, MAE and high determination coefficients R^2, even though the database for testing were not utilized for training. This reflects the generalization capability of the developed GPR model.
- The prediction capability of the developed model was compared to that of the existing models proposed in EN 1992-4 and ACI 318. The statistical analysis revealed that the proposed GPR model had relatively lower errors and higher determination coefficient than the existing codified models investigated.
- The model uncertainty associated with the GPR model has the closest mean value to 1 ($\mu_{ME} = 0.99$) and the lowest standard deviation ($\sigma_{ME} = 0.11$). Therefore, the GPR model is described as the best performer of all the models analysed in this study.
- In the context of the reliability analyses, a limit state function should ideally be based on a good predictive model, with low bias (with a mean close to 1) and uncertainty coupled. Such a model can be used as a general probabilistic model in the reliability analysis of fastening to concrete design provisions.

References

1. Adeli, H. (1986). Artificial intelligence in structural engineering. *Engineering Analysis, 3*(3), 154–160.
2. Alqedra, M. A., & Ashour, A. F. (2005). Prediction of shear capacity of single anchors located near a concrete edge using neural networks. *Computers and structures, 83*(28–30), 2495–2502.
3. Sakla, S. S., & Ashour, A. F. (2005). Prediction of tensile capacity of single adhesive anchors using neural networks. *Computers and structures, 83*(21–22), 1792–1803.
4. Olalusi, O. B., & Spyridis, P. (2020). Machine learning-based models for the concrete breakout capacity prediction of single anchors in shear. *Advances in Engineering Software, 147,* 102832.
5. Hoang, N. D., Pham, A. D., Nguyen, Q. L., & Pham, Q. N. (2016). Estimating compressive strength of high performance concrete with Gaussian process regression model. *Advances in Civil Engineering.*
6. Technical Committee 250/European Committee for Standardisation (CEN/TC 250). EN 1992-4 Eurocode 2. (2018). *Design of concrete structures. Design of fastenings for use in concrete.*
7. ACI Committee and International Organization for Standardization. (2014). *Building code requirements for structural concrete (ACI 318-14) and commentary.* American Concrete Institute.
8. Hofmann, J. (2004). Tragverhalten und Bemessung von Befestigungen unter beliebiger Querbelastung in ungerissenem Beton. (*Load—Bearing behavior and design of fastenings under arbitrary shear loading in uncracked concrete—In German*). Dissertation, University of Stuttgart.
9. Fuchs, W., Eligehausen, R., & Breen, J. E. (1995). Concrete capacity design (CCD) approach for fastening to concrete. *Structural Journal, 92*(1), 73–94.
10. Lee, N. H., Park, K. R., & Suh, Y. P. (2010). Shear behavior of headed anchors with large diameters and deep embedments. *ACI Structural Journal, 107*(2).
11. Bentsen, F. H. (2019). *Model Construction with Support Vector Machines and Gaussian Processes through Kernel Search* (Master's thesis, The University of Bergen).
12. Ueda, T., Kitipornchai, S., & Ling, K. (1990). Experimental investigation of anchor bolts under shear. *Journal of Structural Engineering, 116*(4), 910–921.
13. Ueda, T., Stitmannaithum, B., & Matupayont, S. (1991). Experimental investigation on shear strength of bolt anchorage group. *Structural Journal, 88*(3), 292–300.
14. Dong-Hyun, K., Yong-Myung, P., Ho-Jung, J., & Moon-Ki, K. (2014). Concrete breakout capacity of cast-in-place anchor under shear loading. *Advanced Materials Research.*
15. Hofmann, J., Fuchs, W., & Eligehausen, R. (2004). Quertragfähigkeit randnaher Befestigungsmittel mit Belastung senkrecht zum Bauteilrand (Behavior of anchorages arranged close to the edge and loaded towards the edge). In: Beton- und Stahlbetonbau; 99 No. 10, pp. 806–807 (in German).
16. Shaikh, A. F., & Yi, W. (1985). In-place strength of welded headed studs. *PCI Journal, 30*(2), 56–81.
17. Toth, M., Bokor, B., & Sharma, A. (2019). Anchorage in steel fiber reinforced concrete–concept, experimental evidence and design recommendations for concrete cone and concrete edge breakout failure modes. *Engineering Structures, 181,* 60–75.
18. Unterweger, A. (2008). Randnahe Anker unter Querlast mechanische Modellierung und Bemessung (Anchors close to edge under shear load). Vienna, Austria: Universität für Bo-denkultur, Dissertation, (In German).
19. Klingner, R. E., Muratli, H., & Shirvani, M. (1999). A technical basis for revision to anchorage criteria. Division of Engineering Technology, Office of Nuclear Regulatory Research, US Nuclear Regulatory Commission.
20. Grosser, P. (2010). Single anchors loaded in shear close to the edge in high strength concrete. Test Report, Stuttgart, Germany: Institute of Construction Materials, University of Stuttgart. Report No. E 10/19 – Bft/17, not published
21. Spyridis, P. (2011). *Behavior of anchor groups under shear loads—Influence of assembly tolerances.* Dissertation. University of Natural resources and Life Sciences of Vienna.

22. Senftleben, S. (2010). Zum Tragverhalten von randnahen Reihenbefestigungen beansprucht durch eine Querlast parallel zum Bauteilrand. (*Behavior of multiple anchor connections close to the edge under shear loading parallel to the edge*). Stuttgart, Germany: Institute of Construction Materials, University of Stuttgart, Diploma Thesis, (in German).
23. Gesoğlu, M., Güneyisi, E. M., Güneyisi, E., Yılmaz, M. E., & Mermerdaş, K. (2014). Modeling and analysis of the shear capacity of adhesive anchors post-installed into uncracked concrete. *Composites Part B: Engineering, 60,* 716–724.
24. Olalusi, O. B., & Viljoen, C. (2020). Model uncertainties and bias in SHEAR strength predictions of slender stirrup reinforced concrete beams. *Structural Concrete, 21*(1), 316–332.
25. Olalusi, O. B., & Spyridis, P. (2020). Uncertainty modelling and analysis of the concrete edge breakout resistance of single anchors in shear. *Engineering Structures, 222,* 111112.