СТРОИТЕЛЬНАЯ МЕХАНИКА ИНЖЕНЕРНЫХ КОНСТРУКЦИЙ И СООРУЖЕНИЙ
STRUCTURAL MECHANICS OF ENGINEERING CONSTRUCTIONS AND BUILDINGS

2025. 21(3). 231–241

ISSN 1815-5235 (Print), 2587-8700 (Online)

HTTP://JOURNALS.RUDN.RU/STRUCTURAL-MECHANICS

РАСЧЕТ И ПРОЕКТИРОВАНИЕ СТРОИТЕЛЬНЫХ КОНСТРУКЦИЙ
ANALYSIS AND DESIGN OF BUILDING STRUCTURES

DOI: 10.22363/1815-5235-2025-21-3-231-241
EDN: TJJGKF

Research article / Научная статья

Predicting the Strength of Eccentrically Compressed Short Circular Concrete Filled Steel Tube Columns

Tatiana N. Kondratieva✉, **Anton S. Chepurnenko**, **Batyr M. Yazyev**

Don State Technical University, *Rostov-on-Don, Russian Federation*
✉ ktn618@yndex.ru

Received: March 7, 2025.
Revised: May 28, 2025.
Accepted: June 2, 2025.

Abstract. The process of predicting the load-bearing capacity of eccentrically compressed circular concrete filled steel tube (CFST) columns using machine learning algorithms is investigated. The relevance of the work is established by the need to improve the accuracy of engineering calculations in the context of increasingly complex architectural solutions. The purpose of the study is to develop and evaluate the effectiveness of intelligent models for reliable prediction of CFST column strength based on key parameters of the structure and materials. The object of the study was short, eccentrically compressed CFST columns of circular cross-section. The input parameters of the machine learning models were the outer diameter of the column section, tube wall thickness, concrete strength, yield strength of steel and relative eccentricity. The load-bearing capacity of the column was taken as the output parameter. CatBoost and Random Forest Regressor (RFR) algorithms with hyperparameter optimization using the Optuna library were used for forecasting. The quality of the models was assessed using the MAE, MSE, and MAPE metrics. As a result of the study, intelligent models were developed. The CatBoost model demonstrated better accuracy rates (MAE = 67.1; MSE = 86.2; MAPE = 0.07%) compared to RFR (MAE = 72.6; MSE = 89.7; MAPE = 0.15%). The feature importance analysis showed that the outer diameter of the column and the relative eccentricity have the greatest influence on the bearing capacity. Correlation analysis confirmed the high dependence of the output parameter on these factors. The obtained results are recommended for use in calculation modules and supporting engineering systems for design solutions of load-bearing structures.

Keywords: CFST columns, machine learning, CatBoost, Random Forest, bearing capacity, strength prediction, intelligent models

Conflicts of interest. The authors declare that there is no conflict of interest.

Authors' contribution: *Kondratieva T.N.* — conceptualization, goals and objectives of the study, calculations, analysis of results, writing; *Chepurnenko A.S.* — supervision, review and editing, correction of conclusions; *Yazyev B.M.* — review and editing, final conclusions.

For citation: Kondratieva T.N., Chepurnenko A.S., Yazyev B.M. Predicting the strength of eccentrically compressed short circular concrete filled steel tube columns. *Structural Mechanics of Engineering Constructions and Buildings.* 2025;21(3):231–241. http://doi.org/10.22363/1815-5235-2025-21-3-231-241

Tatiana N. Kondratieva, Candidate of Technical Sciences, Associate Professor of the Department of Mathematics and Informatics of the Don State Technical University (DSTU), 1 Gagarin Sq., Rostov-on-Don, 344003, Russian Federation; eLIBRARY SPIN-code: 7794-2841, ORCID: 0000-0002-3518-8942; e-mail: ktn618@yndex.ru

Anton S. Chepurnenko, Doctor of Technical Sciences, Professor of the Department of Structural Mechanics and Theory of Structures, Don State Technical University, 1 Gagarin Sq., Rostov-on-Don, 344003, Russian Federation; ORCID: 0000-0002-9133-8546, eLIBRARY SPIN-code: 7149-7981; e-mail: anton_chepurnenk@mail.ru

Batyr M. Yazyev, Doctor of Technical Sciences, Professor of the Department of Structural Mechanics and Theory of Structures, Don State Technical University, 1 Gagarin Sq., Rostov-on-Don, 344003, Russian Federation; ORCID: 0000-0002-5205-1446; eLIBRARY SPIN-code: 5970-5350; e-mail: ps62@yandex.ru

© Kondratieva T.N., Chepurnenko A.S., Yazyev B.M., 2025

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

Прогнозирование прочности коротких внецентренно сжатых круглых трубобетонных колонн

Т.Н. Кондратьева✉, **А.С. Чепурненко**, **Б.М. Языев**

Донской государственный технический университет, *Ростов-на-Дону, Российская Федерация*
✉ ktn618@yndex.ru

Поступила в редакцию: 7 марта 2025 г.
Доработана: 28 мая 2025 г.
Принята к публикации: 2 июня 2025 г.

Аннотация. Исследован процесс прогнозирования несущей способности внецентренно сжатых круглых трубобетонных колонн (ТБК) с использованием алгоритмов машинного обучения. Актуальность работы обусловлена необходимостью повышения точности инженерных расчетов в условиях усложняющихся архитектурных решений. Цель исследования – разработка и оценка эффективности интеллектуальных моделей для надежного прогнозирования прочности ТБК на основе ключевых параметров конструкции и материалов. Объектом исследования выступили короткие внецентренно сжатые трубобетонные колонны круглого сечения. Входными параметрами моделей машинного обучения являлись наружный диаметр сечения колонны, толщина стенки трубы, прочность бетона, предел текучести стали и относительный эксцентриситет. В качестве выходного параметра принималась несущая способность колонны. Для прогнозирования использовались алгоритмы CatBoost и Random Forest Regressor (RFR) с оптимизацией гиперпараметров посредством библиотеки Optuna. Оценка качества моделей проводилась по метрикам MAE, MSE и MAPE. В результате исследования разработаны интеллектуальные модели. Модель CatBoost продемонстрировала лучшие показатели точности (MAE = 67,1; MSE = 86,2; MAPE = 0,07 %) по сравнению с RFR (MAE = 72,6; MSE = 89,7; MAPE = 0,15 %). Анализ важности признаков показал, что наибольшее влияние на несущую способность оказывают наружный диаметр колонны и относительный эксцентриситет. Корреляционный анализ подтвердил высокую зависимость выходного параметра от этих факторов. Полученные результаты рекомендуются к использованию в расчетных модулях и инженерных системах поддержки принятия решений при проектировании несущих конструкций зданий и сооружений.

Ключевые слова: модели машинного обучения, CatBoost, Random Forest, несущая способность, интеллектуальные модели

Заявление о конфликте интересов. Авторы заявляют об отсутствии конфликта интересов.

Вклад авторов: *Кондратьева Т.Н.* — формирование основной концепции, цели и задачи исследования, проведение расчетов, анализ результатов исследований, написание текста; *Чепурненко А.С.* — научное руководство, доработка текста, корректировка выводов; *Языев Б.М.* — доработка текста, итоговые выводы.

Для цитирования: *Kondratieva T.N., Chepurnenko A.S., Yazyev B.M.* Predicting the strength of eccentrically compressed short circular concrete filled steel tube columns // Строительная механика инженерных конструкций и сооружений. 2025. Т. 21. № 3. С. 231–241. http://doi.org/10.22363/1815-5235-2025-21-3-231-241

1. Introduction

Concrete filled steel tube (CFST) structures are widely used in construction due to their high strength, stiffness and cost-effectiveness, as well as good seismic stability and resistance to external factors. In [1] a modern review of publications on concrete filled steel tube columns, experimental and analytical basis of research under static and dynamic loads are presented. CFST elements can also be used in truss structures that can withstand high loads [2]. Concrete filling of steel structures increases the compressive strength of the upper chord of steel tube trusses and prevents local buckling of the tube wall.

Кондратьева Татьяна Николаевна, кандидат технических наук, доцент кафедры математики и информатики, Донской государственный технический университет (ДГТУ), Российская Федерация, 344003, г. Ростов-на-Дону, пл. Гагарина, 1; eLIBRARY SPIN-код: 7794-2841, ORCID: 0000-0002-3518-8942; e-mail: ktn618@yndex.ru

Чепурненко Антон Сергеевич, доктор технических наук, профессор кафедры строительной механики и теории сооружений, Донской государственный технический университет, Российская Федерация, 344003, г. Ростов-на-Дону, пл. Гагарина, 1; eLIBRARY SPIN-код: 7149-7981, ORCID: 0000-0002-9133-8546; e-mail: anton_chepurnenk@mail.ru

Языев Батыр Меретович, доктор технических наук, профессор кафедры строительной механики и теории сооружений, Донской государственный технический университет, Российская Федерация, 344003, Ростов-на-Дону, пл. Гагарина, 1; eLIBRARY SPIN-код: 5970-5350, ORCID: 0000-0002-5205-1446; e-mail: ps62@yandex.ru

A comprehensive study of the compressive load-bearing capacity of concrete filled steel tube columns is presented in [3]. In [4], a study of the efficiency of circular short CFST columns under axial load is carried out using thirty specimens of different types. In [5], the test results of twelve short CFST columns with circular steel casing and twelve square short columns subjected to axial compressive loading are presented. In addition to the experimental study of mechanical properties of concrete CFST column structures [6–8], it is necessary to create systems of practical recommendations for optimal design, analysis [9; 10] and application of such columns in various construction conditions [11–13]. The solution of problems of similar type has already attracted the attention of the authors of [14–16], where machine learning methods were used to obtain empirical formulas and statistical models for predicting the strength of CFST columns of square and circular section. The developed models show noticeable improvements in the prediction accuracy of model features [17], but their effectiveness depends largely on the quality of the initial data, the choice of algorithms, and the optimization of hyperparameters [18; 19]. In addition, it is necessary to consider the interpretability of models [20–22], their ability to generalize to new data, and their robustness to changes in input parameters [23–25].

Most of the existing machine learning models for predicting the load-bearing capacity of CFST columns are based on experimental data [26–28]. At the same time, experiments are usually conducted on specimens of relatively small size compared to real structures, which reduces the accuracy and limits the applicability of the models in engineering practice. Thus, there is a gap in the scientific literature associated with the lack of approaches allowing to create ML-models based on more generalizable and parametrically flexible data sources.

This study proposes construction of a predictive model based on the results of finite element modeling using a simplified methodology that has been previously tested on experimental data. This approach avoids the limitations inherent in laboratory testing and allows for a wide variation of parameters, which is critical for building a universal and engineerable model.

The purpose of this study is to develop reliable machine learning models for predicting the load-bearing capacity of eccentrically compressed circular concrete filled steel tube columns. The study is based on a hybrid approach that combines finite element modeling to form a training data set and further training of artificial intelligence models using the results of numerical experiments. This approach was previously successfully applied in [29] for predicting the bearing capacity of centrally compressed columns of square cross-section. The advantage over direct finite element analysis is the very high speed of the trained models: results for several sets of input parameters can be obtained in a fraction of a second by simple matrix multiplication. Another important advantage is that machine learning models allow to estimate the degree of significance of each input parameter in determining the load bearing capacity, which makes them particularly useful in engineering analysis and design.

2. Methods

To build the machine learning model, a database of 374975 numerical experiments using the finite element method according to the methodology presented in [30] was prepared for circular CFST columns. The methodology presented in [30] was previously tested on experimental data for 265 centrally compressed columns and 93 eccentrically compressed columns [31].

The following values were taken as the input parameters of the models: outer diameter of cross-section D_p (mm); wall thickness of circular steel tube t_p (mm); yield strength of steel R_y (MPa); compressive strength of concrete R_b (MPa); relative eccentricity e_o / D_p (dimensionless parameter). The output parameter of the model is the critical load N_{ult} (kN).

Such parameters as initial modulus of elasticity of concrete and tensile strength of concrete were not included in the number of input parameters because they are correlated with the compressive strength grade of concrete. The empirical formulas given in [30] were used in the finite element analysis to determine these parameters. In addition, the modulus of elasticity of steel was not included in the number of input parameters, because it has a weak influence on the compressive strength, and also has a small range of values, varying in the range of 195–210 GPa. The concrete tensile strength parameter, according to engineering standards, also has no significant effect on the column behavior in compression and was excluded from the final set of features. Preliminary analysis using a trained CatBoost model showed that when these parameters were varied by ±10%, the change in the predicted load bearing capacity was less than 1%, which is much lower than the contribution of the main geometric and strength characteristics.

In addition, only short columns, for which deflection does not significantly increase the eccentricity of the axial force were considered, and the design length of the element was also not included as an input parameter.

For building machine learning models, algorithms based on ensemble principles were chosen: CatBoost as a representative of gradient boosting and Random Forest Regressor as a representative of bagging. These algorithms demonstrate high accuracy and stability when working with tabular data, as well as have built-in mechanisms of model interpretation through the assessment of feature importance. The choice of CatBoost was additionally motivated by its high performance on sparse and categorical data, as well as low sensitivity to hyperparameter tuning. To increase the validity of the choice, other algorithms were also tested: XGBoost, LightGBM and the support vector regression (SVR). According to the test results, CatBoost showed the best values of MAE, MSE and MAPE metrics among the considered algorithms, which became the basis for its selection as the main method of boosting in this work. Random Forest was included as a benchmark algorithm for bagging-based ensembling, which is robust to overtraining and does not require feature scaling.

Table 1 partially summarizes the dataset used to train the artificial intelligence models.

Table 1

Fragment of the training data array

No.	D_p, мм / mm	t_p, мм / mm	R_y, МПа / MPa	R_b, МПа / MPa	e_o / D_p	N_{ult}, кН / kN
1	102	1.8	240	10	0	229.3777
2	102	1.8	240	10	0.04	208.9631
3	102	1.8	240	10	0.08	186.9428
4	102	1.8	240	10	0.12	168.3632
5	102	1.8	240	10	0.16	152.7655
6	102	1.8	240	10	0.2	139.9204
7	102	1.8	240	10	0.24	128.6809
8	102	1.8	240	10	0.28	118.8176
9	102	1.8	240	10	0.32	110.56
10	102	1.8	240	10	0.36	103.22
11	102	1.8	240	10	0.4	96.568
…	…	…	…	…	…	…
374965	1420	32	800	66	0.44	90861.91
374967	1420	32	800	66	0.48	85727.21
374968	1420	32	800	66	0.52	81039
374969	1420	32	800	66	0.56	76797.29
374970	1420	32	800	66	0.6	72778.83
374971	1420	32	800	66	0.64	68983.61
374972	1420	32	800	66	0.68	65634.89
374973	1420	32	800	66	0.72	62509.42
374974	1420	32	800	66	0.76	59607.2
374975	1420	32	800	66	0.8	56928.23

S o u r c e: made by T.N. Kondratieva

To improve the quality of the models, the correlations between the variables were analyzed. Stratified sampling methods were used to eliminate imbalances in the data, which allowed to distribute the values of the output parameter evenly. Additionally, a multicollinearity check was performed to eliminate redundant features and increase the robustness of the models. The resulting cleaned and balanced dataset became the basis for training and testing of the machine learning algorithms.

Gradient boosting algorithms (CatBoost, XGBoost) and Random Forest Regressor (RFR) bagging model were used to predict the bearing capacity of CFST columns. The quality of the models was assessed using MAE, MSE and MAPE metrics. In addition, feature importance was analysed based on the contributions of the parameters to predictions.

To improve model accuracy and prevent overtraining, the hyperparameters of the CatBoost and Random Forest Regressor algorithms were optimized. Optimization was performed using Grid search combined with 5-fold cross-validation, which ensured model robustness to random data fluctuations. Ranges of variation of hyperparameters were determined on the basis of preliminary analysis of model behavior on small samples and recommendations from the official documentation of CatBoost and Scikit-learn libraries.

For CatBoost, the parameters varied as: number of iterations — 100–5000; depth — 4–12; learning rate — 0.01–0.3; coefficient of the L2 regularization term in the loss function — 1.9–4.9.

For Random Forest Regressor: number of trees — 50–500; maximum depth — 3–20; minimum number of samples to be in the final node — 1–3.

The best combinations of hyperparameters were selected based on minimizing the MAE and MAPE errors obtained as a result of cross-validation.

For the trained models, feature importance was also analyzed by evaluating the degree of influence of each input parameter on the predictions. This approach allowed to determine which characteristics have the most significant effect on the final value of the bearing capacity.

3. Results and Discussion

The statistical characteristics of the original data set are presented in Table 2, showing the ranges of each characteristic. The main indicators are: sample size, sample mean, deviation, extremes of the variables. The set of these indicators helps to conduct statistical analysis of variables, to determine their scatter relative to their center, to show the asymmetry of distribution, to derive the laws of distribution of ordered series.

Table 2

Table of statistical characteristics

Index	D_p, mm	t_p, mm	R_y, MPa	R_b, MPa	e_o / D_p	N_{ult}, kN
Count	374976.00	374976.00	374976.0	374976.00	374976.00	374976.00
Mean	444.10	10.61	520.0	38.00	0.40	10009.60
Standard Deviation	361.76	7.50	183.3	18.33	0.24	16539.69
Min	102.00	1.80	240.0	10.00	0.00	59.41
Max	1420.00	32.00	800.0	66.00	0.80	223247.94

Source: made by T.N. Kondratieva

Since machine learning is essentially a multivariate interpolation, reliable model performance can be guaranteed only within the range of input parameter variation in the training data set. The minimum and maximum values of input parameters specified in Table 2 almost completely cover the possible range of variation of the characteristics of CFST columns.

Figure 1 shows the correlation between the input and output parameters of the model. There is almost no strong correlation between the input parameters except for the relationship between the wall thickness of the circular steel tube and the outer diameter of the cross-section $\rho_{t_p/D_p} = 0.8$. This is due to the fact that the geometric characteristics of steel tubes presented in GOST 10704-91[1] were used in the formation of the training data set. In this range, there is a tendency to increase the minimum and maximum wall thickness of the tube with increasing diameter to ensure the necessary stiffness and stability of the structure.

A correlation relationship between the input and output parameters of the model is observed: the wall thickness of the circular steel tube and the load-bearing capacity of CFST columns $\rho_{t_p/N_{ult}} = 0.6$; the outer diameter of the cross-section and the load-bearing capacity of CFST columns $\rho_{D_p/N_{ult}} = 0.8$. This result is expected from the standpoint of structural mechanics, since an increase in the wall thickness and outer diameter of the tube leads to an increase in the geometric characteristics of the equivalent cross-section, affecting the performance of the structure under eccentric compression: moment of inertia and area.

As a result of the correlation analysis, it was found that, although most of the parameters are relatively independent of each other, the main geometric characteristics of the column (outer diameter and wall thickness of the tube) have the greatest influence on its load-bearing capacity. Thus, it is important to consider the primary significance of these parameters when developing predictive models and designing CFST structures.

The absence of strong correlation between most of the input parameters of the tested model for predicting the load-bearing capacity of eccentrically compressed circular CFST columns indicates that each of these factors makes an independent contribution to the formation of the load-bearing capacity of the structure. This indicates a complex interaction of variables, in which the strength of columns is determined not by one dominant parameter, but by their combined influence.

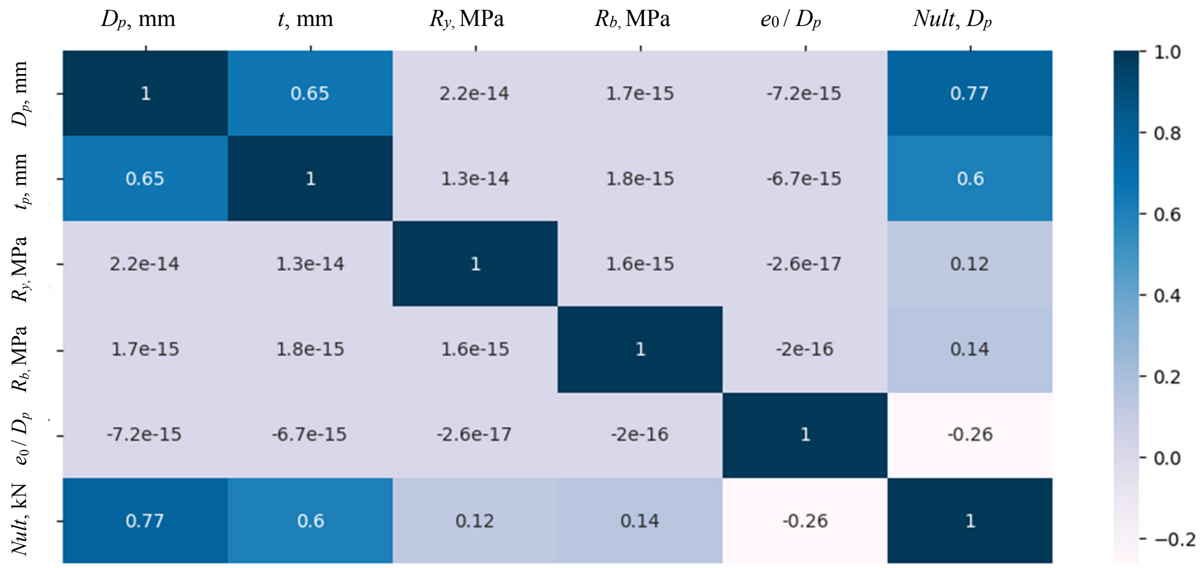

Figure 1. Correlation matrix
S o u r c e: made by T.N. Kondratieva

The results of the feature importance analysis obtained using CatBoost and Random Forest Regressor (RFR) machine learning models are shown in Figures 2 and 3. They allow to identify differences in the approaches of the algorithms to data analysis and to determine the key parameters affecting the load-bearing capacity of eccentrically compressed circular CFST columns.

The CatBoost model demonstrated a more balanced and engineering-based assessment of the influence of factors.

[1] GOST 10704-91. *Electrically welded steel line-weld tubes.* Moscow: Standardinform Publ.; 2007. (In Russ.)

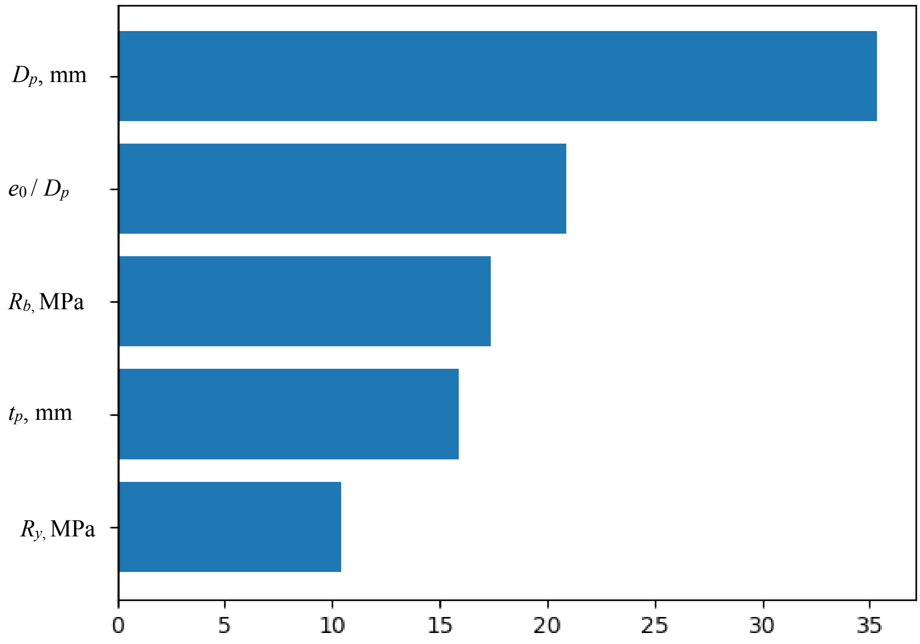

Figure 2. Feature Importance Assessment for CatBoost

S o u r c e: made by T.N. Kondratieva

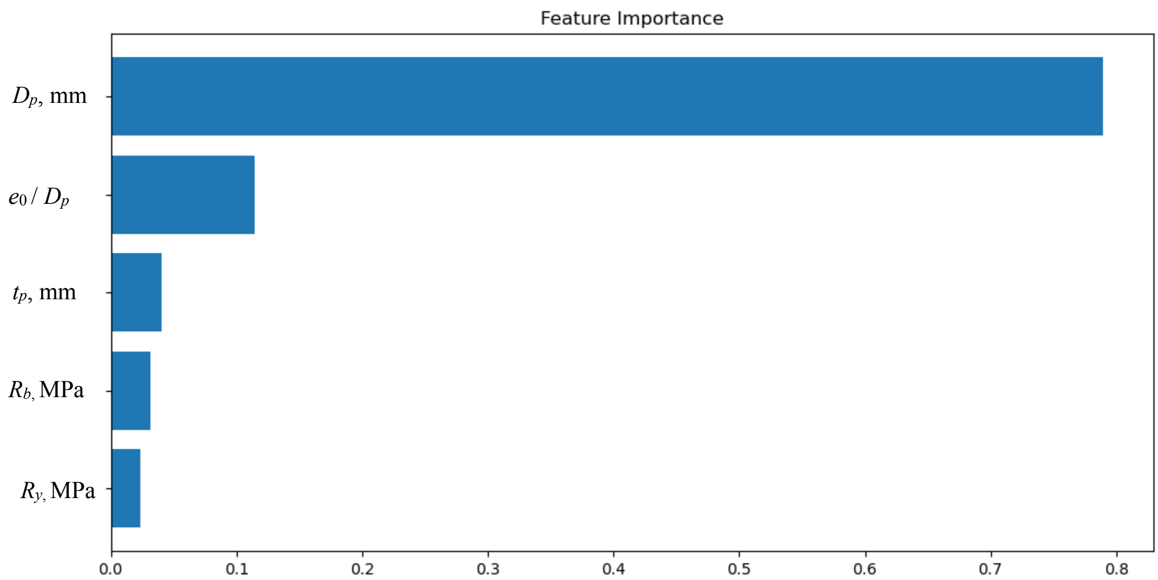

Figure 3. Feature Importance Assessment for RFR

S o u r c e: made by T.N. Kondratieva

The greatest influence is exerted by the outer diameter of the column cross-section (97%), which is reasonable, since an increase in diameter leads to an increase in the cross-sectional area and, consequently, increases the bearing capacity.

The relative eccentricity (60%) and tube wall thickness (40%) contribute significantly. The first parameter indicates the sensitivity of the column to eccentric loading, which is particularly important for complex loading conditions. The second parameter affects the resistance to local buckling and the overall interaction of the steel tube with the concrete.

Other significant parameters are the compressive strength of concrete (37%) and yield strength of steel (22%), which are the key material parameters that determine the behavior of the loaded column.

The Random Forest Regressor (RFR) model showed a strong dependence on a single parameter, the section outer diameter (97%), while other attributes such as wall thickness (0.7%), concrete strength (0.5%) and steel yield strength (0.2%) were rated as practically insignificant. This may indicate the limited ability of RFR to detect complex nonlinear relationships and interactions between attributes, which reduces its applicability for engineering analysis and prediction in conditions of high variability of design parameters.

The optimal parameter values obtained in the process of model training are presented in Table 3.

The quality assessment of the models is presented in Table 4.

Table 3

Optimal values of model parameters

Model	Parameter	Value
CatBoost	Iterations	1487
	Depth	8
	Learning rate	0,4
	l2 leaf reg	2,17
RFR	*N* estimators	150
	Max depth	16
	Min samples leaf	2

S o u r c e: made by T.N. Kondratieva

Table 4

Model quality metrics

Metric	CatBoost	RFR
MAE	67.1	72.6
MSE	86.2	89.7
MAPE (%)	0.07	0.15
R^2	0.99	0.98

S o u r c e: made by T.N. Kondratieva

The CatBoost and Random Forest Regressor algorithms are trained in a search for optimal hyperparameters, where each combination of parameters is tested for the smallest error. Higher mean average errors (MAE) for structural engineering problems are acceptable in both cases and are associated with the large range of the N_{ult} load values and their order of thousands of kN. High mean square errors (MSE) indicate the presence of outliers and indicate single, isolated large errors. MSE by itself is less informative in this case, but in combination with MAPE it indicates reasonable stability of the models. The MAPE values, which are less than 1%, indicate that the models are very accurate in predicting the bearing capacity values.

To increase the rigor of model comparison, the statistical significance of differences in MAE, MAPE, and R^2 between CatBoost and Random Forest Regressor (RFR) was analyzed. For this purpose, bootstrapping (1000 iterations) and nonparametric Wilcoxon test for paired samples on 5-fold cross-validation data were used.

Statistical significance of differences was confirmed by nonparametric Wilcoxon test with confidence intervals calculated by bootstrapping method is presented in Table 5.

Table 5

Comparison of models with confidence intervals and p-values

Metric	CatBoost (95% CI)	RFR (95% CI)	Difference (CatBoost — RFR)	95% CI of difference	p-value (Wilcoxon)
MAE	67.1 [64.3–70.5]	72.6 [69.8–76.1]	−5.5	[−6.8; −3.9]	0.018 < 0.01
MSE	86.2 [81.0–91.7]	89.7 [85.3–94.6]	−3.5	[−5.2; −1.7]	0.041 < 0.05
MAPE (%)	23.1 [21.2–25.4]	31.5 [29.4–34.2]	−0.08	[−0.10; −0.05]	0.003 < 0.01
R^2	0.94 [0.91–0.96]	0.86 [0.83–0.89]	+0.01	[+0.005; +0.02]	0.002 < 0.05

S o u r c e: made by T.N. Kondratieva

Table 5 shows the values of quality metrics of CatBoost and RFR models, as well as statistical analysis of differences between them based on bootstrapping and nonparametric Wilcoxon criterion. The obtained 95% confidence intervals of the differences of metrics between CatBoost and RFR confirmed the significant advantage of CatBoost, additionally, CatBoost demonstrates statistically better results in terms of MAE, MAPE and R^2 (p-value < 0.05), which confirms its advantage as a more accurate and stable model in this task.

Figures 4 and 5 show histograms of errors: actual values along the vertical axis, predicted values along the horizontal axis. The use of CatBoost algorithm allowed to achieve almost perfect coincidence between target and predicted values. Such high quality of prediction can be explained by the high smoothness of the data in the training array: the results of finite element analysis, unlike the experimental data, exclude random errors associated with measurement, scatter of material characteristics, etc. The results of the finite element analysis are not affected by random errors. In Figure 5, the deviation of points from the Actual = Predict line is larger compared to Figure 4. Thus, if choosing between CatBoost and Random Forest Regressor algorithms, preference should be given to the former. A similar result suggesting the better ability of CatBoost to determine complex nonlinear relationships in comparison with RFR was obtained earlier in [32] for another problem in the field of application of wireless sensor networks. This is because the CatBoost algorithm is based on the use of gradient boosting and has a built-in mechanism to deal with overtraining. Since CatBoost has shown the best quality of prediction, it is CatBoost that should be relied on when evaluating feature importance. CatBoost reveals the importance of even weak features without overfitting on them, RFR can ignore weak features because of averaging over trees.

Thus, the CatBoost model provides a more informative and interpretable approach to predicting the bearing capacity of CFST columns, taking into account a combination of factors, which makes it more preferable in the tasks of engineering design and analysis. CatBoost model also allows analyzing the influence of random factors. For this purpose, it is enough to pass the reference input values through the trained model, and then set random deviations for input parameters and see how the result will change.

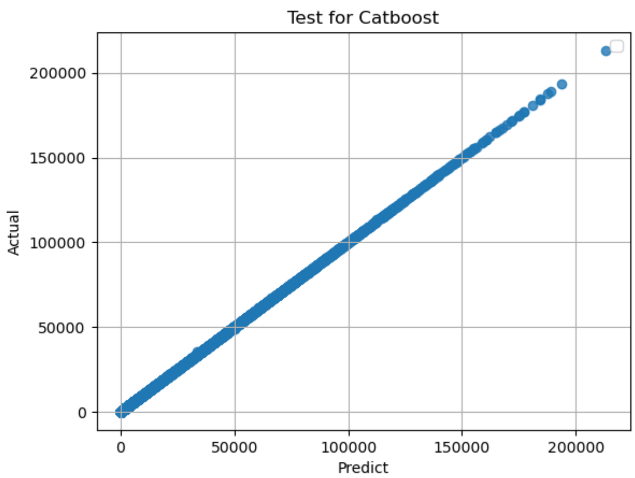

Figure 4. Error histogram for CatBoost

S o u r c e: made by T.N. Kondratieva

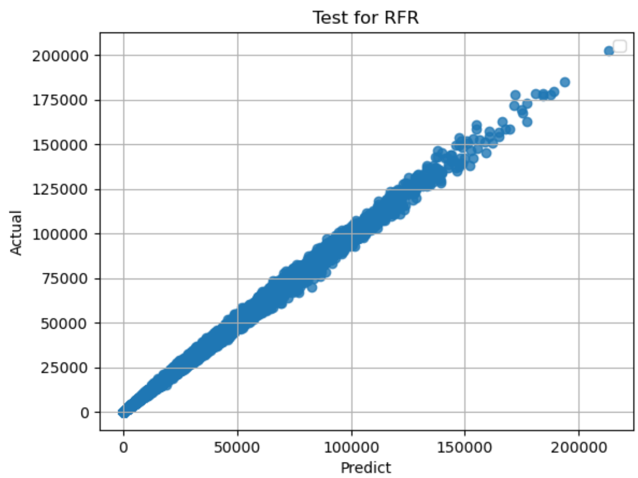

Figure 5. Error histogram for RFR

S o u r c e: made by T.N. Kondratieva

Thus, the comparison of CatBoost and Random Forest Regressor models has shown the advantage of the former both in accuracy and stability of prediction. This is due to the fact that CatBoost implements an advanced gradient boosting scheme with efficient processing of sparse and categorical data, and also has built-in mechanisms to combat overtraining. In contrast, Random Forest algorithm as a bagging method is prone to averaging and may lose sensitivity to weak but significant features. In spite of higher interpretability of Random Forest trees, CatBoost turned out to be the most appropriate for the task of estimation of load-bearing capacity of concrete filled steel tube columns, which is confirmed by low values of MAE, MSE and especially MAPE (less than 1%). The obtained results confirm the reasonableness of CatBoost as the main gradient boosting algorithm in this work. Random Forest, in turn, was used as a reference model to evaluate the relative efficiency of ensemble methods.

4. Conclusion

The main conclusions of the study are:

1. A machine learning model for predicting the load-bearing capacity of eccentrically compressed circular concrete filled steel tube (CFST) columns was built and validated.
2. The dominant influence of geometric characteristics (outer diameter and wall thickness) compared to the mechanical characteristics of materials was established.
3. CatBoost algorithm showed the best results among the tested models.
4. An accessible database suitable for further research and training of ML-models was formed and proposed.

In further studies it is planned to expand the use of machine learning methods in analyzing the characteristics of CFST structures, including their operation in extreme conditions (fire, dynamic effects). The solution of these issues will allow to expand the scope of application CFST columns and increase their reliability in modern construction projects.

Data Accessability Statement

The training dataset is available for download at: https://disk.yandex.ru/d/jVWmibREM10p3Q

References / Список литературы

1. Ilanthalir A., Regin J.J., Maheswaran J. Concrete-filled steel tube columns of different cross-sectional shapes under axial compression: A review. *IOP Conference Series: Materials Science and Engineering*; 2020 Sep 17–18; Tamil Nadu, India. Bristol: IOP Publ.; 2020:012007. https://doi.org/10.1088/1757-899X/983/1/012007 EDN: DMUHMU
2. Joseph J.R., Henderson J.H. Concrete-filled steel tube truss girders — a state-of-the-art review. *Journal of Engineering and Applied Science*. 2023;70:49. https://doi.org/10.1186/s44147-023-00220-w EDN: PWRPVD
3. Arokiaprakash A., Senthil Selvan S. Comprehensive study of compressive behavior of CFST columns with confinements and stiffeners. *Journal of Constructional Steel Research*. 2023;211:108127. https://doi.org/10.1016/j.jcsr.2023.108127 EDN: SKHBUQ
4. Wang C., Chan T.M. Machine learning (ML) based models for predicting the ultimate strength of rectangular concrete-filled steel tube (CFST) columns under eccentric loading. *Engineering Structures*. 2023;276:115392. https://doi.org/10.1016/j.engstruct.2022.115392 EDN: DIBHQH
5. Zhang S., Li X., Chen X., Chen J. Behavior of circular-steel-tube-confined square CFST short columns under axial compression. *Journal of Building Engineering*. 2022;51:104372. https://doi.org/10.1016/j.jobe.2022.104372 EDN: KHGDBW
6. Teng J.G., Wang J.J., Guan Lin, Zhang J., Feng P. Compressive behavior of concrete-filled steel tubular columns with internal high-strength steel spiral confinement. *Advances in Structural Engineering*. 2021;24(8):1687–708. https://doi.org/10.1177/1369433220981656 EDN: YFKXMK
7. Yuan F., Cao L., Li H. Axial compressive behaviour of high-strength steel spiral-confined square concrete-filled steel tubular columns. *Journal of Constructional Steel Research*. 2022;192:107245. https://doi.org/10.1016/j.jcsr.2022.107245 EDN: AJTUWQ
8. Hu H.S., Xu L., Guo Z.X., Shahrooz B.M. Behavior of eccentrically loaded square spiral-confined high-strength concrete-filled steel tube columns. *Engineering Structures*. 2020;216:110743. https://doi.org/10.1016/j.engstruct.2020.110743 EDN: RBSRZD
9. Kondratieva T.N., Chepurnenko A.S., Poliakova K.A., Rodionov K.A. CatBoost algorithms to predict the load-bearing capacity of centrally compressed short CFST columns of circular cross-section. *E3S Web of Conferences*. 2024; 583:06009. https://doi.org/10.1051/e3sconf/202458306009 EDN: YXISDV
10. Kondratieva T.N., Vysokovskiy D.V., Rusakova E.V., Khashkhozhev K.A., Poliakova K.A. Modeling the strength of the walls of I-shaped reinforced concrete beams. *International Conference on Recent Advances in Architecture and Construction*. Cham: Springer Nature Switzerland; 2024. p. 385–392. https://doi.org/10.1007/978-3-031-82938-3_42
11. Le T.T. Practical machine learning-based prediction model for axial capacity of square CFST columns. *Mechanics of Advanced Materials and Structures*. 2022;29(12):1782–1992. https://doi.org/10.1080/15376494.2020.1839608 EDN: LUHGDL

12. Zhou X.G., Hou C., Feng W.Q. Optimized data-driven machine learning models for axial strength prediction of rectangular CFST columns. *Structures*. 2023;47:760–780. https://doi.org/10.1016/j.istruc.2022.11.030 EDN: RJEFIJ 0

13. Cakiroglu C., Islam K., Bekdaş G., Isikdag U., Mangalathu S. Explainable machine learning models for predicting the axial compression capacity of concrete filled steel tubular columns. *Construction and Building Materials*. 2022;356:129227. https://doi.org/10.1016/j.conbuildmat.2022.129227 EDN: RDNJKT

14. Wang D., Ren Z., Kondo G. Interpretable domain knowledge enhanced machine learning framework on axial capacity prediction of circular CFST columns. *Journal Research Article: Computational Engineering, Finance, and Science*. 2024. https://doi.org/10.48550/arXiv.2402.04405

15. Megahed K., Mahmoud N.S., Abd-Rabou S.E.M. Prediction of the axial compression capacity of stub CFST columns using machine learning techniques. *Scientific Reports*. 2024;14:2885. https://doi.org/10.1038/s41598-024-53352-1 EDN: GANCNR

16. Faridmehr I., Nehdi M.L. Predicting axial load capacity of CFST columns using machine learning. *Structural Concrete*. 2022;23:1642–1658. https://doi.org/10.1002/suco.202100641 EDN: BOSJKP

17. Hakim S.J.S, Noorzaei J., Jaafar M.S., Jameel M., Mohammadhassani M. Application of artificial neural networks to predict compressive strength of high strength concrete. *International Journal of Physical Sciences*. 2011;6:975–981. https://doi.org/10.5897/IJPS11.023

18. Luat N.V., Han S.W., Lee K. Genetic algorithm hybridized with eXtreme gradient boosting to predict axial compressive capacity of CCFST columns. *Composite Structures*. 2021;278:114733. https://doi.org/10.1016/j.compstruct.2021.114733 EDN: KPUGZL

19. Zarringol M., Thai H., Thai S., Patel V. Application of ANN to the design of CFST columns. *Structures*. 2020;28:2203–2220. https://doi.org/10.1016/j.istruc.2020.10.048 EDN: CRVHBP

20. Lee S.C Prediction of concrete strength using artificial neural networks. *Engineering Structures*. 2003;25(7):849–857. https://doi.org/10.1016/S0141-0296(03)00004-X

21. Chepurnenko A.S., Kondratieva T.N., Deberdeev T.R., Akopyan V.F., Avakov A.A. Prediction of rheological parameters of polymers using the CatBoost gradient boosting algorithm. *Polymer Science, Series D*. 2024;17(1):121–128. https://doi.org/10.1134/S199542122370020X EDN: UFJNYY

22. Lee S., Vo T.P., Thai H.T, Lee J., Patel V. Strength prediction of concrete-filled steel tubular columns using Categorical Gradient Boosting algorithm. *Engineering Structures*. 2021;238:112109. https://doi.org/10.1016/j.engstruct.2021.112109 EDN: FTLKPF

23. Ma Lu., Zhou Ch., Lee D., Zhang Ji. Prediction of axial compressive capacity of CFRP-confined concrete-filled steel tubular short columns based on XGBoost algorithm. *Engineering Structures*. 2022;260:114239. https://doi.org/10.1016/j.engstruct.2022.114239 EDN: SEMQKN

24. Hou C., Zhou X. G. Strength prediction of circular CFST columns through advanced machine learning methods. *Journal of Building Engineering*. 2022;51:104289. https://doi.org/10.1016/j.jobe.2022.104289 EDN: CQXMDK

25. Vu Q.V., Truong V.H., Thai H.T. Machine learning-based prediction of CFST columns using gradient tree boosting algorithm. *Composite Structures*. 2021;259:113505. https://doi.org/10.1016/j.compstruct.2020.113505 EDN: HTWBNC

26. Tran V.L., Thai D.K., Kim S.E. Application of ANN in predicting ACC of SCFST column. *Composite* Structures. 2019;228:111332. https://doi.org/10.1016/j.compstruct.2019.111332

27. Du Y., Chen Z., Zhang C., Cao X. Research on axial bearing capacity of rectangular concrete-filled steel tubular columns based on artificial neural networks. *Frontiers of Computer Science.* 2017;11:863-873. https://doi.org/10.1007/s11704-016-5113-6 EDN: KGWXVD

28. Zarringol M., Thai H.T., Thai S., Patel V. Application of ANN to the design of CFST columns. *Structures.* 2020;28:2203-2220. https://doi.org/10.1016/j.istruc.2020.10.048 EDN: CRVHBP

29. Chepurnenko A.S., Yazyev B.M., Turina V.S., Akopyan V.F. Artificial intelligence models for determining the strength of centrally compressed pipe-concrete columns with square cross-section. *Magazine of Civil Engineering*. 2024;17(6). https://doi.org/10.34910/MCE.130.8 EDN: HBBADX

30. Chepurnenko A., Yazyev B., Meskhi B., Beskopylny A., Khashkhozhev K., Chepurnenko V. Simplified 2D finite element model for calculation of the bearing capacity of eccentrically compressed concrete-filled steel tubular columns. *Applied Sciences*. 2021;11:11645. https://doi.org/10.3390/app112411645 EDN: AKSSMM

31. Khashkhozhev K.N. *Improvement of the design of concrete-filled steel tube columns considering physical nonlinearity* [dissertation]. 2023. (In Russ.)

32. *Хашхожев К.Н.* Совершенствование расчета трубобетонных колонн с учетом физической нелинейности: дис. ... канд. техн. наук. 2023. 138 с.

33. Ileri K. Comparative analysis of CatBoost, LightGBM, XGBoost, RF, and DT methods optimised with PSO to estimate the number of k-barriers for intrusion detection in wireless sensor networks. *International Journal of Machine Learning and Cybernetics.* 2025:1–20. https://doi.org/1007/s13042-025-02654-5