

## Article

# Optimal Dimensions of Post-Tensioned Concrete Cylindrical Walls Using Harmony Search and Ensemble Learning with SHAP

Gebrail Bekdaş<sup>1,\*</sup>, Celal Cakiroglu<sup>2</sup>, Sanghun Kim<sup>3</sup> and Zong Woo Geem<sup>4,\*</sup><sup>1</sup> Department of Civil Engineering, Istanbul University-Cerrahpasa, 34320 Istanbul, Turkey<sup>2</sup> Department of Civil Engineering, Turkish-German University, 34820 Istanbul, Turkey; cakiroglu@tau.edu.tr<sup>3</sup> Department of Civil and Environmental Engineering, Temple University, Philadelphia, PA 19122, USA; sanghun.kim@temple.edu<sup>4</sup> Department of Smart City & Energy, Gachon University, Seongnam 13120, Republic of Korea

\* Correspondence: bekdas@iuc.edu.tr (G.B.); geem@gachon.ac.kr (Z.W.G.)

**Abstract:** The optimal design of prestressed concrete cylindrical walls is greatly beneficial for economic and environmental impact. However, the lack of the available big enough datasets for the training of robust machine learning models is one of the factors that prevents wide adoption of machine learning techniques in structural design. The current study demonstrates the application of the well-established harmony search methodology to create a large database of optimal design configurations. The unit costs of concrete and steel used in the construction, the specific weight of the stored fluid, and the height of the cylindrical wall are the input variables whereas the optimum thicknesses of the wall with and without post-tensioning are the output variables. Based on this database, some of the most efficient ensemble learning techniques like the Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), Categorical Gradient Boosting (CatBoost) and Random Forest algorithms have been trained. An  $R^2$  score greater than 0.98 could be achieved by all of the ensemble learning models. Furthermore, the impacts of different input features on the predictions of different machine learning models have been analyzed using the SHapley Additive exPlanations (SHAP) methodology. The height of the cylindrical wall was found to have the greatest impact on the optimal wall thickness, followed by the specific weight of the stored fluid. Also, with the help of individual conditional expectation (ICE) plots the variations of predictive model outputs with respect to each input feature have been visualized. By using the genetic programming methodology, predictive equations have been obtained for the optimal wall thickness.

**Keywords:** optimization; machine learning; XGBoost; SHAP; prestressed concrete; post-tensioning; genetic programming



check for updates

**Citation:** Bekdaş, G.; Cakiroglu, C.; Kim, S.; Geem, Z.W. Optimal Dimensions of Post-Tensioned Concrete Cylindrical Walls Using Harmony Search and Ensemble Learning with SHAP. *Sustainability* **2023**, *15*, 7890. <https://doi.org/10.3390/su15107890>

Academic Editors: Uroš Klanšek and Tomaž Žula

Received: 10 April 2023

Revised: 30 April 2023

Accepted: 9 May 2023

Published: 11 May 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Prestressing has well-known benefits in terms of increasing the load-carrying capacity and durability of concrete structures like liquid storage tanks, silos, and nuclear facilities [1]. It significantly reduces concrete cracks at the tension side by pre-compressing those parts of the concrete structures under tensile strain before the service loads are applied thereby countering the effects of tension caused by bending. Post-tensioning is a type of prestressing where the tendons are stressed after the concrete has been poured and hardened. Post-tensioning is a particularly favorable technique when the serviceability requirements on a large-scale structure cannot be met using conventional reinforcement techniques while keeping the dimensions of the structure in an economically feasible range [2].

Previous studies of the reinforced concrete structures optimization with and without prestressing have demonstrated the applicability of metaheuristic optimization techniques like the harmony search (HS) methodology to reduce the structural cost and environmental

impact [2–6]. Metaheuristic optimization techniques have been applied to a variety of problems in civil and structural engineering including the dimensioning of trusses [7–13], retaining walls [14–22], laminated composite plates [23–30], and steel plate girders [31]. On the other hand, post-tensioning in structures is an area where metaheuristic optimization techniques found relatively few applications. The research in the area of structural post-tensioning can be grouped into studies related to beams [32], slabs [33–35], experimental studies [33,34,36], and numerical studies [32,34–36]. Elbelbisi et al. [32] performed a parametric study of post-tensioned fiber reinforced polymer (FRP) systems. The flexural behavior of beams externally post-tensioned with FRP tendons was investigated using finite element analysis. The application of external post-tensioning significantly increased the load-carrying capacity of the beams. Elsheshtawy et al. [33] investigated the effect of prestressing force and layout of strands on the punching shear strength of slab-column connections and found that the banded layout of the post-tensioning strands was most favorable to increase the punching shear strength of the post-tensioned flat slabs. Attia et al. [34] examined the load-carrying capacity of two-way flat slabs under post-tensioning with external FRP laminates using finite element analysis. A parametric study was carried out and it was shown that post-tensioning with FRP laminates increase the ductility. Furthermore, strengthening near the supports was shown to be more effective than in the middle of the slab. Tahmasebinia et al. [35] investigated post-tensioned concrete flat slabs under dynamic loading. According to finite element analysis, increasing the slab thickness and damping ratio is favorable towards better vibration serviceability. Vavrus and Kralovanec [36] investigated the application of steel fiber reinforced concrete in the anchorage zones of post-tensioning tendons to increase the load-carrying capacity. Based on numerical analysis, increasing the fiber content near the anchorage plate is favorable for the increase of the load carrying capacity. In some of the more recent studies in the area of post-tensioned structural members, bearing capacity of anchorage zones [37], precast post-tensioned girders [38], and post-tensioned self-compacting concrete beams with recycled coarse aggregate [39] have been experimentally and numerically investigated. Lei et al. [37] proposed a new formula for the bearing capacity of anchorage zones using the stress field approach. Joyklad et al. [38] carried out an experimental program with full-scale pre-cast post-tensioned (PCPT) girders accompanied with numerical simulations. Two PCPT girders of 29.95 m length, 1.8 m depth and 0.69 m width were manufactured. The upward deflection response of these girders, post-tensioning tendon strains, and reinforcement strains were measured with displacement transducers and strain gauges for 120 days. Furthermore, the thermal response of fresh concrete was measured using thermocouple wires. All measured parameters were observed to increase during post-tensioning. Yu et al. [39] carried out four-point bending tests on post-tensioned beams with and without recycled coarse aggregate. The specimens made with recycled coarse aggregate and self-compacting concrete were demonstrated to have similar flexural capacity as the specimens with natural coarse aggregate. To the best of the authors' knowledge the only research studies involving the application of metaheuristic optimization techniques to post-tensioned reinforced concrete cylindrical containers was carried out by Bekdaş et al. [40]. It was found that the addition of multiple layers of post-tensioning cables reduces the amount of carbon emission associated with the manufacturing process. Harmony search (HS), teaching-learning based optimization (TLBO), and flower pollination algorithm (FPA) were utilized for minimizing the carbon footprint.

The current study aims to develop predictive models using ensemble machine learning techniques and genetic programming in order to predict the optimal value of wall thickness for a post-tensioned reinforced concrete cylindrical wall. In order to train these predictive models a large database of optimal design configurations has been generated using the harmony search algorithm. The statistical analysis of this data set as well as the harmony search methodology have been presented in Section 2 which also contains the details of the genetic programming and ensemble learning algorithms. Using the genetic programming algorithm, closed-form equations have been proposed for the prediction of the optimal

wall thickness for cylindrical walls with and without post-tensioning. The performances of the ensemble learning and genetic programming algorithms in predicting the optimal wall thickness have been presented in Section 3. In the same section also, the input feature impact analysis has been carried out using SHAP and individual conditional expectation plots. The major novelty of the current paper is the proposal of a new methodology to overcome the data scarcity for the training of reliable predictive models using machine learning algorithms. Since reliable machine learning models necessitate large data sets which are not readily available in case of post-tensioned cylindrical walls, the current proposed technique can make a valuable contribution to this area.

## 2. Machine Learning and Optimization Methods

In this section, the procedures of building data sets of optimum post-tensioned cylindrical walls and developing machine learning (ML) models based on these data sets are explained. The database for the training of the ML models is generated using the harmony search optimization algorithm. In this process the serviceability requirements on the cylindrical wall such as the limitation on the crack size according to ACI 318 code [41] are considered as the constraints of the optimization. The structural analysis of the cylindrical walls has been carried out using the superposition method (SPM), which is based on St. Venant's principle according to which stresses sufficiently far from an applied load are not significantly altered if this load is changed to a statically equivalent load combination [42].

$$M_A + \frac{P_{0A}}{4\beta} + \frac{P_{0B}}{4\beta}C_{\beta 1} + \frac{M_{0A}}{2} + \frac{M_{0B}}{2}D_{\beta 1} = 0, \quad (1)$$

$$Q_A - \frac{P_{0A}}{2} + \frac{P_{0B}}{2}D_{\beta 1} - \frac{\beta M_{0A}}{2} + \frac{\beta M_{0B}}{2}A_{\beta 1} = 0, \quad (2)$$

$$M_B + \frac{P_{0A}}{4\beta}C_{\beta 1} + \frac{P_{0B}}{4\beta} + \frac{M_{0A}}{2}D_{\beta 1} + \frac{M_{0B}}{2} = 0, \quad (3)$$

$$Q_B - \frac{P_{0A}}{2}D_{\beta 1} + \frac{P_{0B}}{2} - \frac{\beta M_{0A}}{2}A_{\beta 1} - \frac{\beta M_{0B}}{2} = 0, \quad (4)$$

$$A_{\beta 1} = e^{-\beta 1}(\cos \beta 1 + \sin \beta 1), \quad B_{\beta 1} = e^{-\beta 1}\sin \beta 1 \quad (5)$$

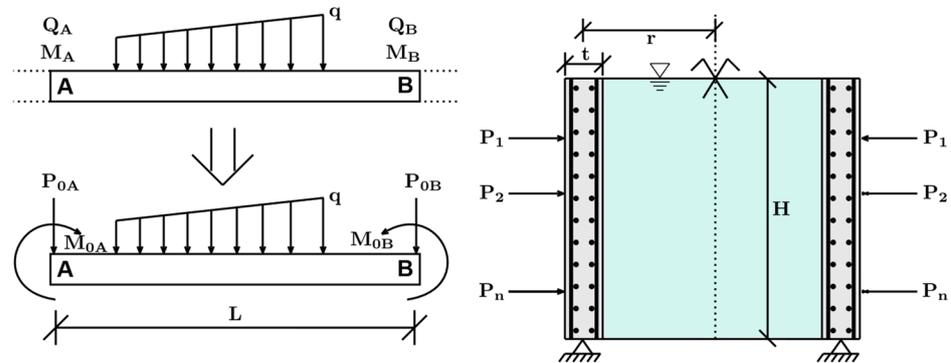
$$C_{\beta 1} = e^{-\beta 1}(\cos \beta 1 - \sin \beta 1), \quad D_{\beta 1} = e^{-\beta 1}\cos \beta 1 \quad (6)$$

The process of SPM has been visually depicted in Figure 1 where the beam in the top left shows a section of length  $L$  from an infinite wall,  $q$  is a distributed load representing the forces acting on the wall,  $M_A, M_B, Q_A, Q_B$  are the moments and shear forces at the positions A and B, respectively and  $P_0, M_0$  at the bottom left portion of Figure 1 are the end conditioning forces and moments, respectively which should be statically equivalent and in opposite direction to  $M_A, M_B, Q_A, Q_B$  such that these moments and shear forces at the positions A,B are cancelled and the infinite wall is equivalent to a finite wall with free ends. The end-conditioning forces can be found by solving Equations (1) to (4). Further details of the SPM can be found in Hetenyi [43] and Bekdaş [2].

The post-tensioning forces are denoted with  $P_1$  to  $P_n$  in the right side of Figure 1. The optimal distances of these forces from the ground as well as their optimal magnitudes were determined using the harmony search technique. The wall thickness, radius of the cylindrical wall and the height of the wall are denoted with  $t, r,$  and  $H,$  respectively. Using the methodology described in the previous paragraph, data sets of 1925 samples were generated for cylindrical walls with different numbers of post-tensioning loads. The

data samples were generated by considering the total cost of the resulting structures and minimizing the objective function given in Equation (7).

$$f(\mathbf{x}) = C_c V_c + C_s W_s + C_{pt} W_{pt} + C_{fw} A_{fw} \quad (7)$$



**Figure 1.** SPM for cylindrical walls.

In Equation (7),  $f(\mathbf{x})$  denotes the function that outputs the total cost of the resulting structure,  $\mathbf{x}$  denotes a vector containing the values for the wall dimensions and unit costs, and  $C_c, C_s, C_{pt}, C_{fw}$  are the cost of concrete per unit volume, cost of steel rebars per unit weight, cost of post-tensioning per unit weight and cost of formwork per unit area, respectively. The details of the objective function can be found in [2]. The database consisting of the optimal design configurations has been taken for the prediction of the optimal wall thickness. The height of the wall ( $H$ ), the unit cost of concrete ( $C_c$ ), specific weight of the liquid ( $\gamma$ ), and unit cost of steel ( $C_s$ ), have been used as the input features defining the load cases. The following section shows the details of each data set used in the development of the predictive models.

### 2.1. Analysis of the Data Set

The data sets used in the prediction of the optimal wall thickness for the case with and without post-tensioning consist of 1908 and 1925 data samples, respectively. The data sets were generated using the harmony search algorithm to find the optimal wall thickness value corresponding to certain combinations of wall height, fluid specific weight and unit costs of concrete and steel. These data sets are split into a training set and a test set in 70% to 30% ratio. The ensemble learning models have been trained on the training set using the 10-fold cross-validation process. Afterward, the performances of these models have been measured on the test set which consists of data samples yet unseen by the models. In the 10-fold cross-validation process the training set is split into 10 disjoint subsets. The models are trained using 9 of these subsets while the tenth subset is being used for model validation. After 10 passes of the training set the best performing model is being selected which is used for performance evaluation on the test set.

The correlation plots in Figures 2 and 3 show the Pearson correlation coefficients between different features in the upper right triangular parts. It can be seen that the wall height is the most correlated input feature with the wall thickness with correlation coefficients of 0.89 and 0.91 followed by the fluid specific weight ( $\gamma$ ) with correlation coefficients of 0.34 and 0.37. The strength of the correlation is also denoted by stars in Figures 2 and 3. A Pearson correlation coefficient close to 1 would indicate a highly linear relationship between two variables. The computation formula for the Pearson correlation coefficient is given in Equation (8) where  $x_i, y_i$  denote two data series of equal length,  $n$  is

the length of these data series, and  $r_{xy}$  is the Pearson correlation coefficient between these two data series.

$$r_{xy} = \frac{n\sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{n\sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \sqrt{n\sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2}} \quad (8)$$

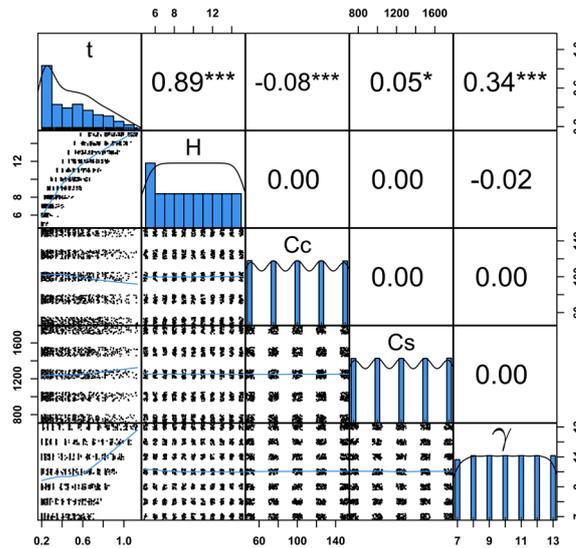


Figure 2. Correlation plot of the data set with one layer of post-tensioning cables.

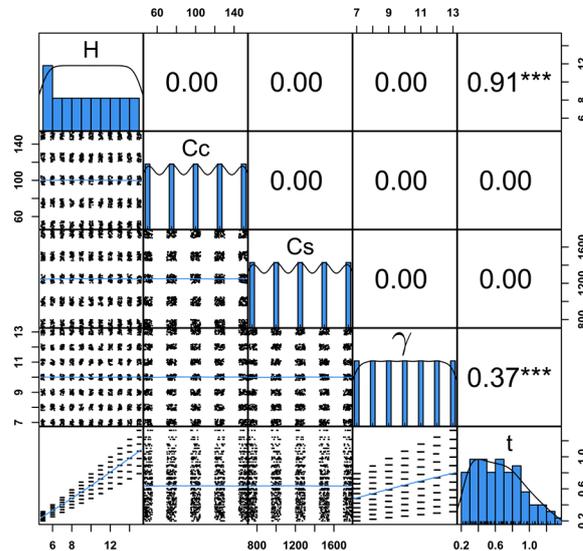


Figure 3. Correlation plot of the data set without post-tensioning cables.

In addition to the Pearson coefficient, Figures 2 and 3 also display the distributions of different features. The diagonals of each plot contain frequency distributions of the features whereas the lower left triangular area contains bivariate scatter plots with regression lines. For each feature denoted on a diagonal tile, the value range of this feature is shown in both the horizontal and the vertical axis.

### 2.2. Harmony Search Algorithm

The harmony search (HS) algorithm was developed by Geem et al. [44] and has been used for solving numerous optimization problems in broad areas including the numerical solution of differential equations [45], project scheduling [46], weapon target

assignment [47], law enforcement [48], internet of things [49], robotics [50], soil stability analysis [51], and structural design [52–54]. The HS method is based on the evolutionary improvement of an initially randomly generated population of optimal solution candidates denoted with  $x$  in Equations (10) and (11). The size of this population is denoted with HMS in Equation (9) which stands for harmony memory size. As the members of the population go through the HS iterations given in Equations (9) to (12), new and better-performing population members replace the members of the previous generations.

$$k = \text{int}(\text{rand} \cdot \text{HMS}), \text{rand} \in (0, 1) \quad (9)$$

$$x_{i,\text{new}} = x_{i,\text{min}} + \text{rand} \cdot (x_{i,\text{max}} - x_{i,\text{min}}), \text{if HMCR} > \text{rand} \quad (10)$$

$$x_{i,\text{new}} = x_{i,k} + \text{rand} \cdot \text{PAR} \cdot (x_{i,\text{max}} - x_{i,\text{min}}), \text{if HMCR} \leq \text{rand} \quad (11)$$

$$\text{HMCR} = 0.5 \left( 1 - \frac{i}{\max(i)} \right), \text{PAR} = 0.05 \left( 1 - \frac{i}{\max(i)} \right) \quad (12)$$

### 2.3. Ensemble Learning Algorithms

Ensemble learning methods are based on the idea of combining the predictions of multiple predictive models to obtain a strong learning algorithm. In this study XGBoost, Random Forest, LightGBM, and Catboost ensemble learning algorithms have been applied. The ensemble learning models have been trained using the scikit-learn library available for the Python programming language. The XGBoost algorithm iteratively generates decision trees while each newly generated tree corrects the errors of the previous trees. In these iterations the XGBoost algorithm aims at minimizing the objective function given in Equation (13) where  $M$  is the size of the training set,  $N$  is the number of decision trees,  $l$  is a loss function and  $\Omega$  is a penalty function. The XGBoost algorithm is summarized in Equation (14) where  $g$  denotes the strong learner model which is a linear combination of  $N$  weak learners denoted with  $g_k$  and  $\hat{y}$  is the model prediction [55].

$$L = \sum_{i=1}^M l(y_i, \hat{y}_i) + \sum_{k=1}^N \Omega(f_k) \quad (13)$$

$$g(x) = \sum_{k=1}^N g_k(x) = \hat{y} \quad (14)$$

The LightGBM, Random Forest, and CatBoost algorithms function on similar principals as the XGBoost algorithm by iteratively adding weak learner trees to the model and combining their output to build a strong learner model. The distinguishing feature of the LightGBM algorithm is its computational speed. The LightGBM algorithm implements techniques such as histogram-based algorithm, and Gradient-based One-Side Sampling (GOSS) to achieve improved performance [56]. On the other hand, the CatBoost algorithm stands out by its ability to work with categorical features more efficiently. The CatBoost algorithm incorporates ordered boosting, greedy method, and L2-regularization to improve model performance [57].

### 2.4. Genetic Programming

Genetic programming (GP) is a population based evolutionary algorithm. The algorithm generates a population of programs where each program is represented by a tree structure. Each node in these programs represents an operation (such as addition or multiplication) or a numerical value. Figure 4 shows an example of representing a program as a tree structure where nodes containing numerical values, unary operators and binary operators are shown in different colors. The fitness of each program is evaluated by

running on a test set, and the programs that perform the best are selected for producing the next generation of programs through crossover and mutation [58–60]. The process of producing new programs from the existing ones using crossover and mutation operations is schematically explained in Figure 5. In the crossover part of Figure 5 the identical parts of the programs are shown inside rectangles of the same color whereas in the mutation part the mutated operator is shown inside a red rectangle.

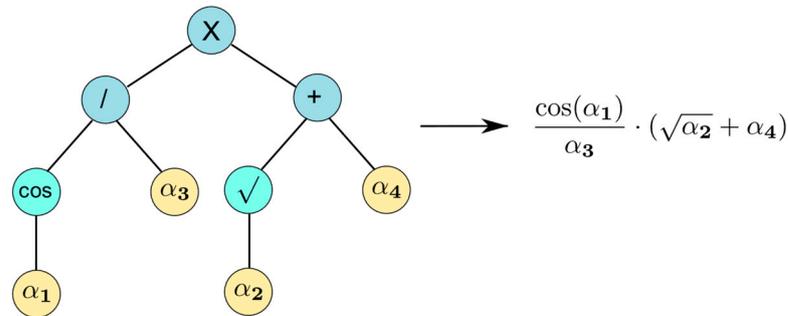


Figure 4. Tree representation of a program consisting of 5 operators and 4 parameters.

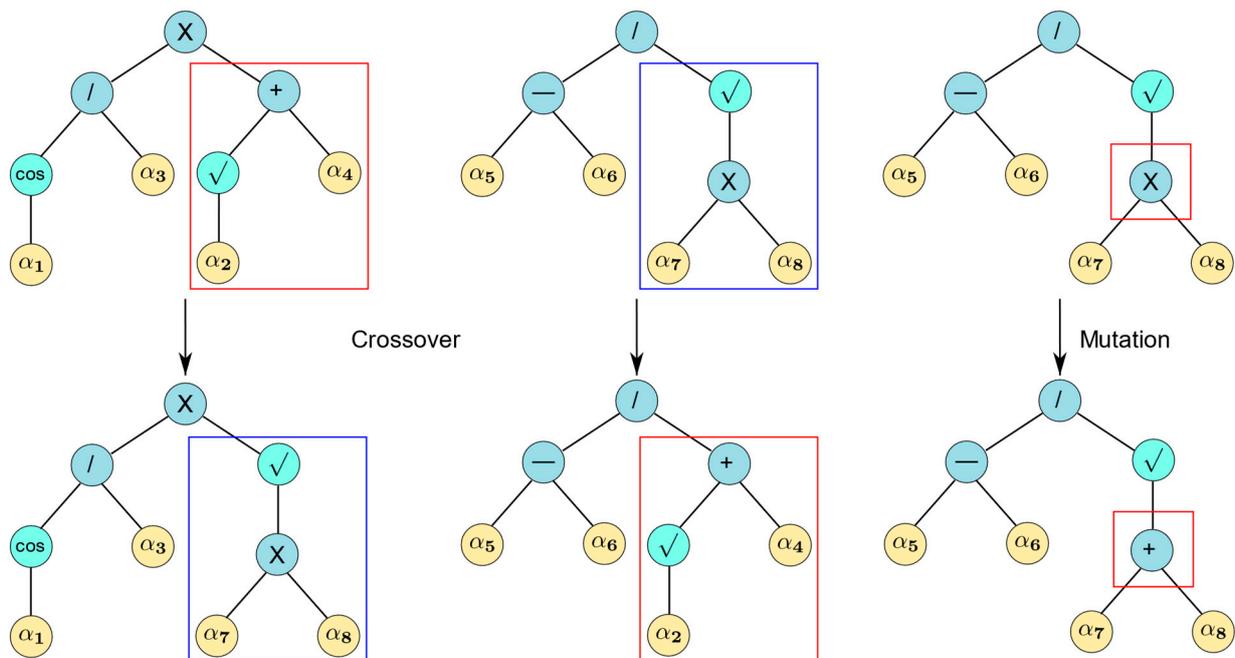
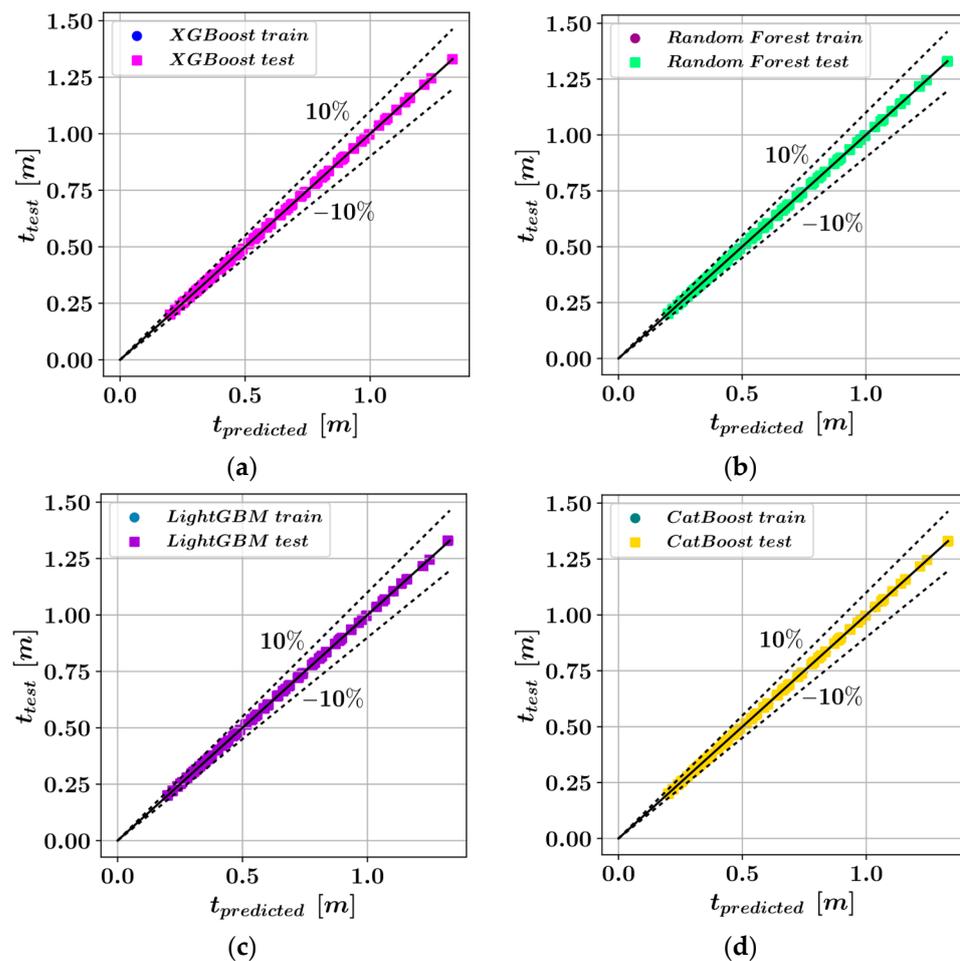


Figure 5. Visualizations of program updates by crossover and mutation.

### 3. Results

This section gives a detailed presentation of the predictive performances of the machine learning models. The predicted and actual optimal wall thickness values are plotted against each other and the percentage deviations of the predicted values from the actual thickness values have been shown. The prediction accuracies of ML models have been quantified using root mean squared error, mean absolute error, and the coefficient of determination ( $R^2$ ). ML models have been trained on datasets of wall geometries with and without post-tensioning cables. Figure 6 shows the prediction of the optimal wall thickness as a function of the concrete and steel unit costs, liquid specific weight, and wall height for a cylindrical wall without post-tensioning. The wall radius has been fixed at a constant value in all the databases in this study since the liquid pressure on the wall is not affected by this quantity. It can be observed from Figure 6 that the predicted and actual optimal values for the thickness perfectly overlap.



**Figure 6.** Comparison of the predicted and actual wall thickness values without post-tensioning cables using (a) XGBoost, (b) Random Forest, (c) LightGBM, (d) CatBoost.

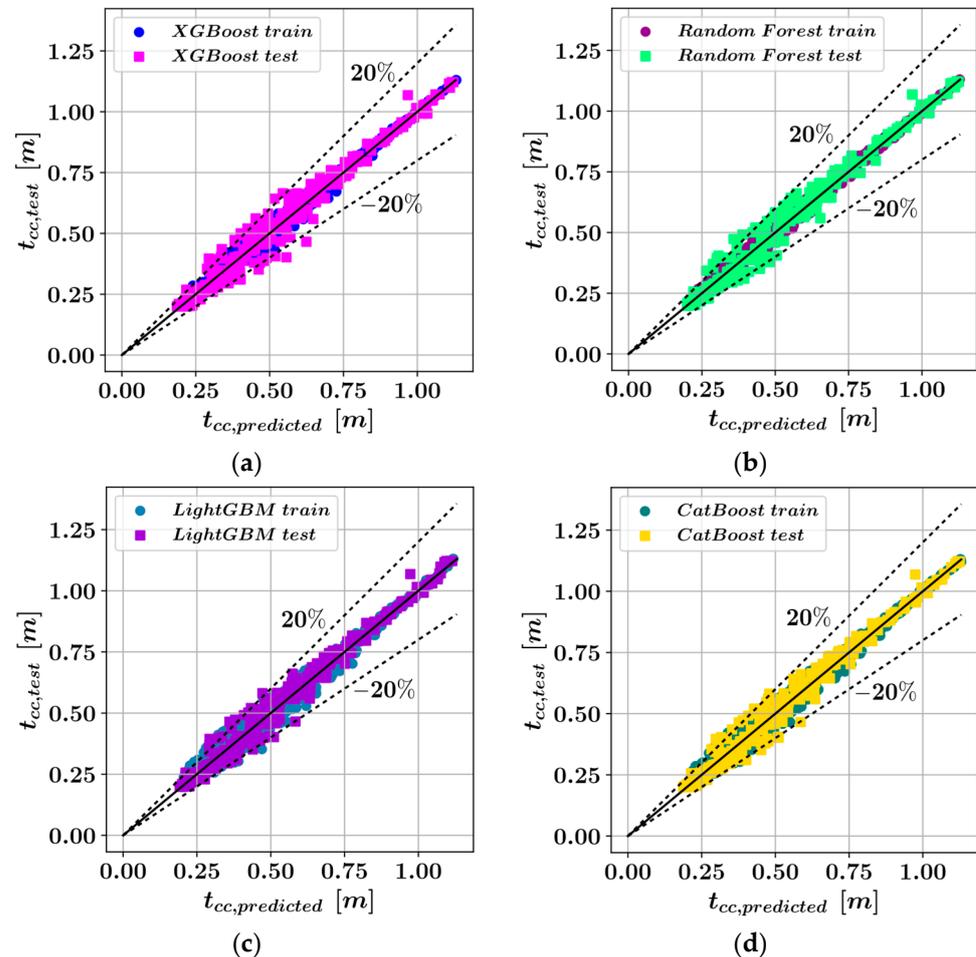
The performances of four different ML models have been quantified and listed in Table 1. It can be observed that all of the ML models were able to predict the optimal wall thickness values with near-perfect accuracy since the  $R^2$  scores are close to 1 for all ML models. Table 1 also lists the performance values for the genetic programming model (GP). Clearly, in terms of both accuracy and computational efficiency the GP model performed worse than the ensemble learning models. However, the GP algorithm has been included in this study, because this algorithm delivers closed-form equations that could be easily incorporated into practical engineering applications for the prediction of the optimal wall thickness.

**Table 1.** Model accuracies in predicting the optimal wall thickness without post-tensioning.

Algorithm	$R^2$		MAE		RMSE		Duration [s]
	Train	Test	Train	Test	Train	Test	
XGBoost	0.9999	0.9999	0.0001	0.0002	0.0003	0.0003	4.81
Random Forest	0.9999	0.9999	$10^{-5}$	$3 \times 10^{-5}$	$9 \times 10^{-5}$	0.0002	3.87
LightGBM	0.9999	0.9999	0.0005	0.0006	0.0012	0.0014	4.52
CatBoost	0.9999	0.9999	0.0002	0.0003	0.0002	0.0004	32.82
GP	0.9584	0.9573	0.0465	0.0460	0.0573	0.0570	359

The predicted and actual optimal wall thickness values in the case of a single layer of post-tensioning have been plotted in Figure 7. It can be observed that the introduction of

post-tensioning into the models caused a slight reduction in the accuracy of the ML models. However, Table 2 shows that all ensemble models were capable of predicting the optimal wall thickness with an  $R^2$  score greater than 0.98 on the test set.



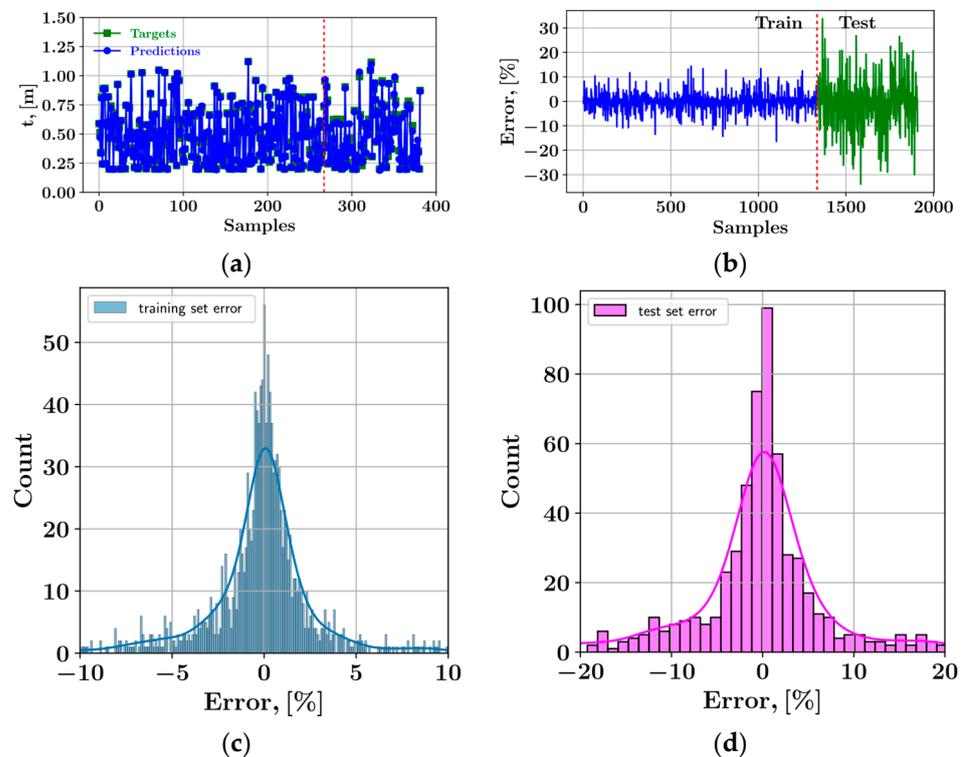
**Figure 7.** Comparison of the predicted and actual optimal wall thicknesses with post-tensioning for (a) XGBoost, (b) Random forest, (c) LightGBM, (d) CatBoost.

**Table 2.** Model accuracies in predicting the optimal wall thickness with post-tensioning.

Algorithm	$R^2$		MAE		RMSE		Duration [s]
	Train	Test	Train	Test	Train	Test	
XGBoost	0.9967	0.9825	0.0087	0.021	0.0139	0.0346	5.66
Random Forest	0.9973	0.9850	0.0077	0.019	0.0126	0.0320	5.14
LightGBM	0.9908	0.9864	0.0153	0.019	0.0234	0.0305	4.39
CatBoost	0.9931	0.9863	0.0129	0.018	0.0203	0.0306	25.15
GP	0.9588	0.9581	0.0422	0.045	0.0501	0.0531	150

To have a clear visualization of the model performance, the percentage errors on the training and test sets have been plotted for the CatBoost model in Figure 8. Figure 8a shows the overlap between the actual and predicted optimal wall thickness values. According to Figure 8b, the error percentages fluctuate in the  $\pm 10\%$  and  $\pm 30\%$  ranges for the training and test sets respectively. Figure 8c,d show the distributions of the error percentages for the training and test sets respectively. Figure 8d, shows that most of the error percentage fluctuations on the test set are accumulated in the  $\pm 20\%$  range whereas smaller error percentages are observed in Figure 8c in the  $\pm 10\%$  range. The entire data set was split into

a training set and a test set in 70% to 30% ratio. In terms of mean absolute error (MAE), the CatBoost model performed best on the test set, whereas the LightGBM model performed best in terms of RMSE and  $R^2$ . It should be noted that all of the four ensemble learning models tested in this study demonstrated similar performances and the differences in performance are deemed negligible. On the other hand, a significant difference in terms of the computational speed could be observed for the CatBoost model. It was observed that the duration of training and testing the models is significantly longer in the case of CatBoost. The LightGBM model was observed to be the most efficient model in terms of computational speed, although the XGBoost and Random Forest models also have similar performances to LightGBM in terms of computational speed. Finally, the closed-form equation obtained from the genetic programming (GP) algorithm performed less accurately than the ensemble learning models.



**Figure 8.** Error percentages of the CatBoost model.

Figure 9 shows the first of the decision trees that constitute the XGBoost model developed for the case with one layer of post-tensioning cables. The XGBoost model consists of a total of 100 iteratively added decision trees such that each new decision tree corrects the errors of the trees before itself. The final prediction of the XGBoost model is the sum of all the predictions made by the decision trees. The decision tree in Figure 9 consists of a root node which is split according to the level of wall height, 41 internal nodes, and 43 leaf nodes. The internal nodes are split at different levels of the fluid specific weight, and the unit costs of steel and concrete. The green colored leaf nodes contain the possible contributions of the decision tree to the final model prediction.

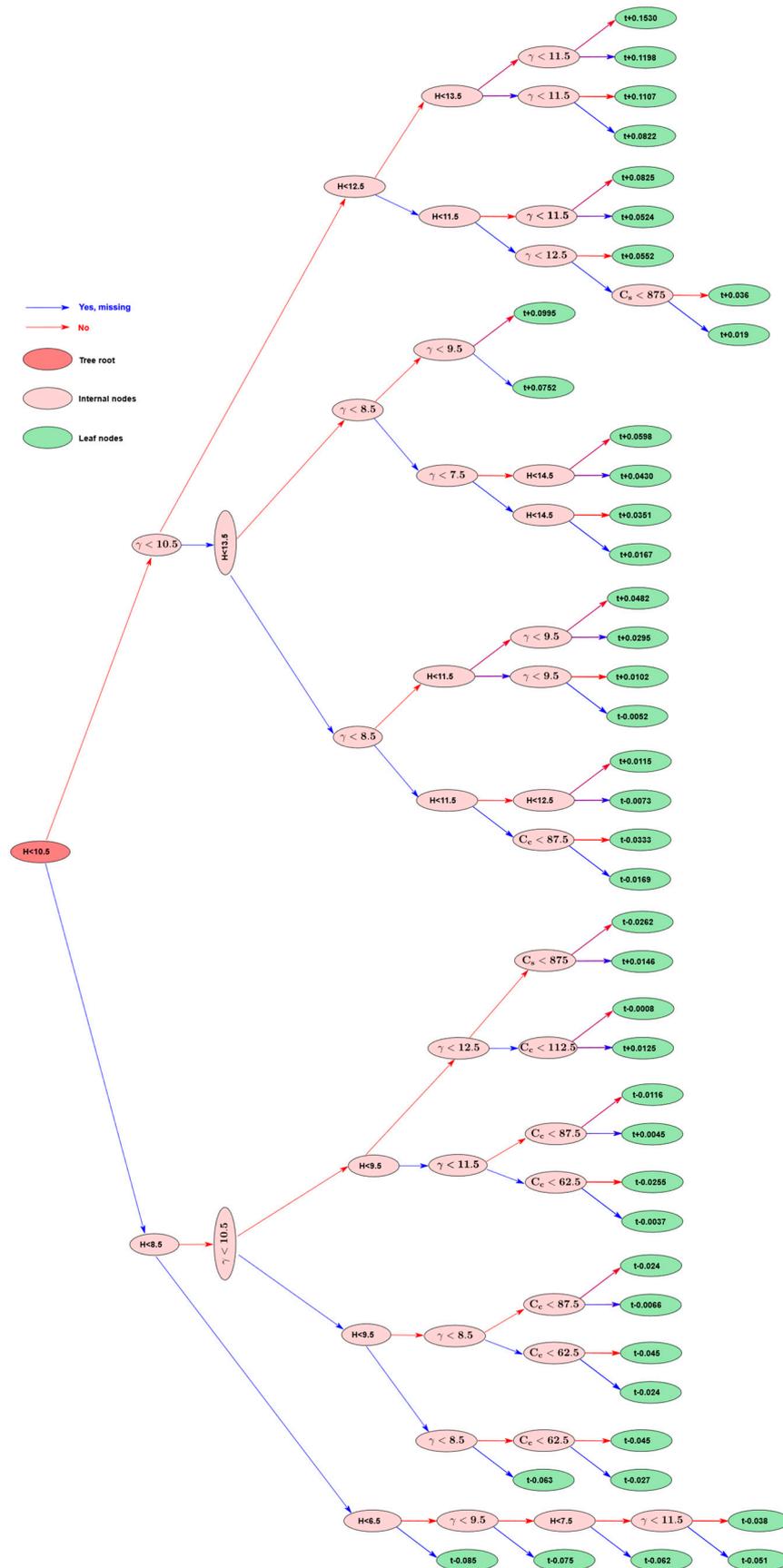


Figure 9. First tree of the XGBoost model.

### 3.1. Interpretation of the Ensemble Learning Models Using SHAP Approach

The SHAP summary plots in Figures 10 and 11 visualize the impact of each input feature on the CatBoost model predictions. The load cases with and without post-tensioning are visualized in Figures 10 and 11, respectively. In the SHAP summary plots each dot represents one of the data points in the data set. The SHAP value corresponding to an input feature in a data point is represented on the horizontal axis, whereas the value of the input feature is color coded as shown with a color bar on the right-hand side of the plot. The higher values of an input feature are displayed with shades of red, while blue colors represent the lower values of an input feature. The input features are sorted by the magnitude of their impact, in decreasing order from top to bottom. The SHAP value measures the contribution of a feature to the model prediction such that positive SHAP values indicate an increasing effect of an input feature on the model output and negative SHAP values indicate a decreasing effect on the model output. The computation of the SHAP values can be described as in Equation (15) where  $F$  is the set of all input features and  $S$  is a subset of  $F$  where the feature with the index  $i$  has been withheld. In Equation (15),  $x$  and  $\phi_i$  represent a vector of input feature values and the corresponding SHAP value respectively [61].

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} [f_{S \cup \{i\}}(x_{S \cup \{i\}}) - f_S(x_S)] \quad (15)$$

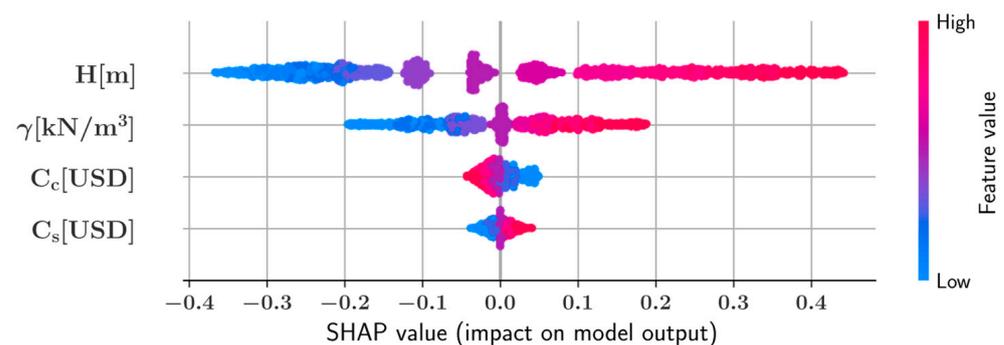


Figure 10. SHAP summary plot with one layer of post-tensioning.

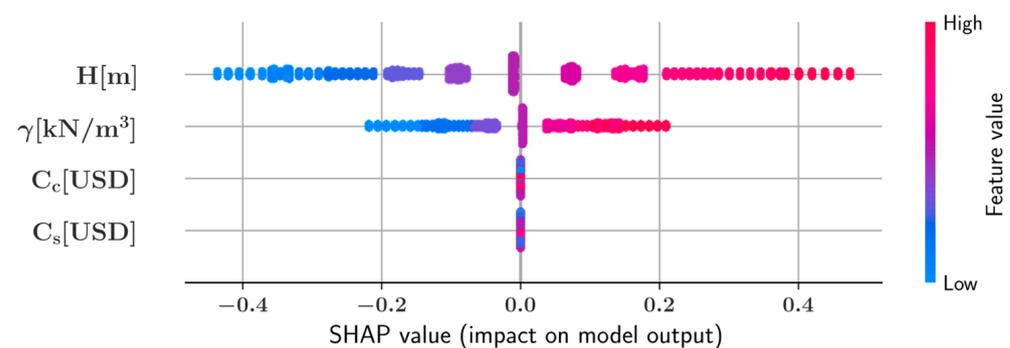


Figure 11. SHAP summary plot without post-tensioning.

According to Figures 10 and 11 the height of the cylindrical wall has the greatest impact on the predicted optimal wall thickness for both load cases. The fluid specific weight is the second most impactful input feature whereas the concrete and steel unit costs have a relatively minor impact on the model predictions. Particularly, in the load case without post-tensioning, the impacts of the concrete and steel unit costs are significantly less than the fluid specific weight and the wall height.

Figures 12 and 13 show feature dependence plots where the y-axis represents the SHAP value for the feature of interest, and the x-axis represents the actual value of the feature for each point in the dataset for the cases with and without post-tensioning respectively. The color of each dot on the feature dependence plots represents the value of a second feature that is most correlated with the main feature represented in the horizontal axis. The feature dependence plots convey information about how the predictions of the model are affected as the value of an input feature varies. A positive relationship between a feature's value and its corresponding SHAP value indicates that an increase in that feature's value will result in an increase in the model predictions. In both Figures 12 and 13, there is a positive relationship between the height of the wall and the corresponding SHAP values which indicates that increased wall height leads to increased predictions of the wall thickness. The coloring of the dots in Figures 12 and 13 indicate that for any given value of the wall height, an increase in the fluid specific weight is associated with a greater increasing impact on the model output wall heights greater than 10 m. On the other hand, for wall heights less than 10 m, in both load cases with and without post-tensioning, the fluid specific weight has the opposite effect on the predicted wall thickness. A similar relationship is also observed between the fluid specific weight and its SHAP value. For values greater than  $10 \text{ kN/m}^3$  the addition of this variable into the models has an increasing impact on the model output whereas specific weight values less than  $10 \text{ kN/m}^3$  are associated with a decrease in the predicted wall thickness values. Furthermore, for any given value of the fluid specific weight, increased wall height also increases the impact of  $\gamma$  on the model output as the coloring of the data points shows.

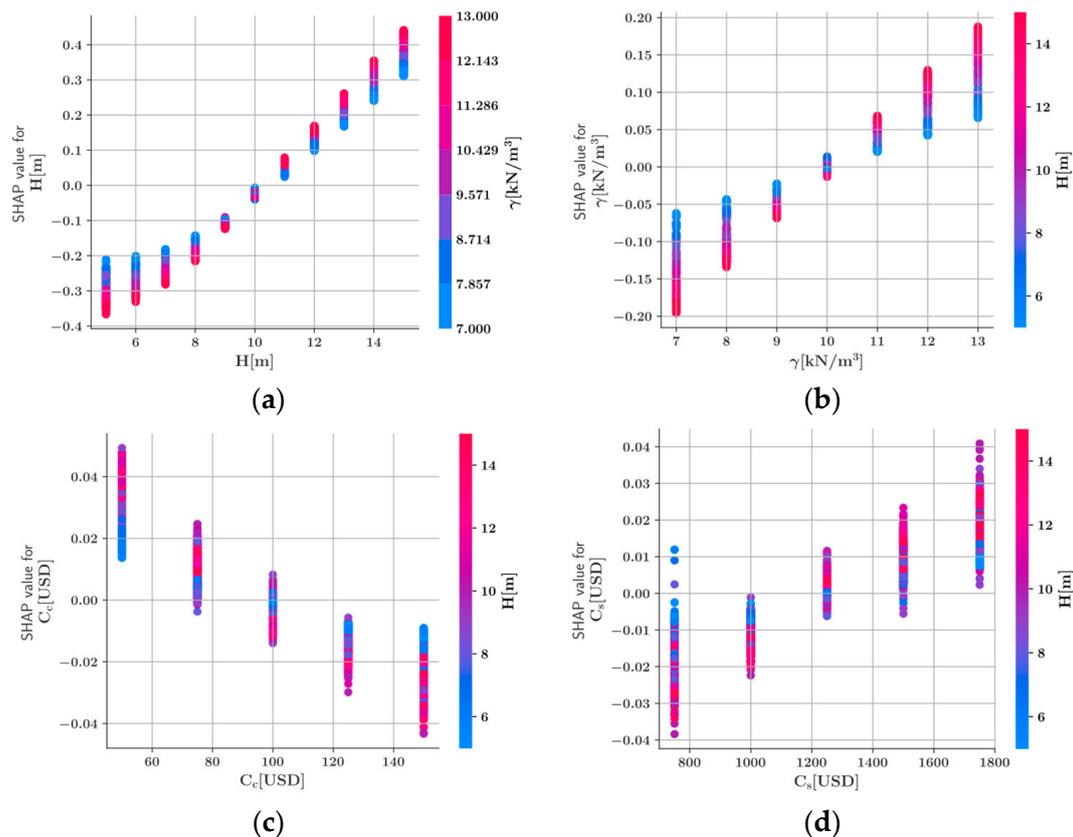


Figure 12. Feature dependence plots (CatBoost) with post-tensioning for (a) H (b)  $\gamma$  (c) C<sub>c</sub> (d) C<sub>s</sub>.

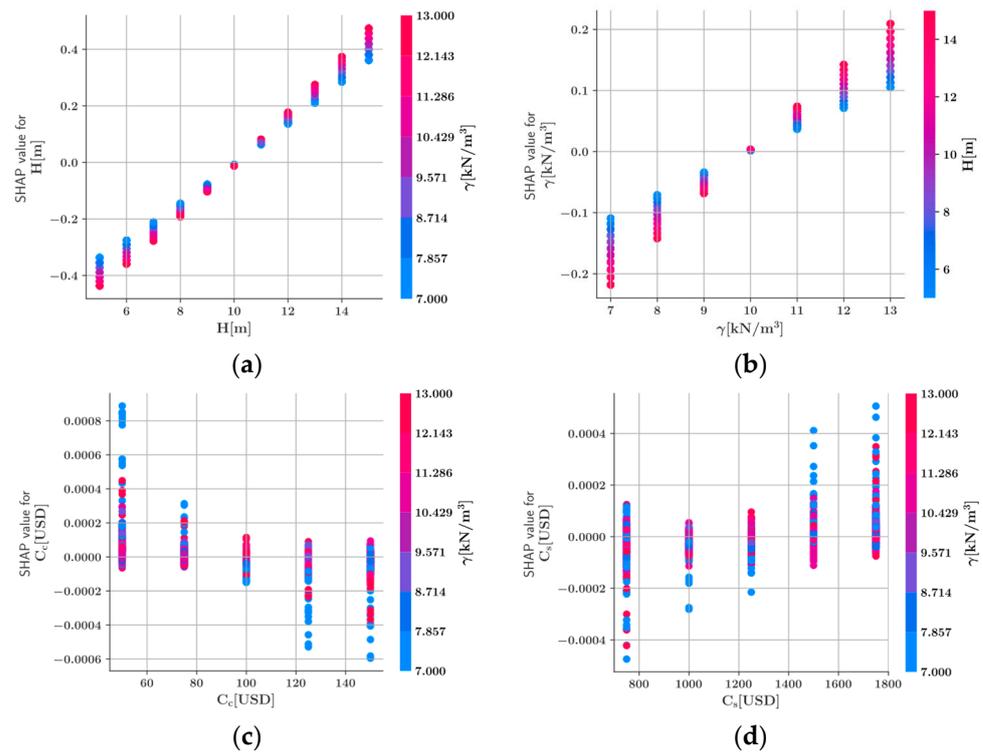


Figure 13. Feature dependence plots (CatBoost) without post-tensioning for (a) H (b)  $\gamma$  (c)  $C_c$  (d)  $C_s$ .

Individual conditional expectation (ICE) plots are helpful for gaining a better understanding of how a single feature affects the predictions made by a machine learning model. The ICE plots in Figures 14 and 15 show the variation of the model predictions concerning a single input feature while the values of the remaining input features are kept constant. Each line in the ICE plot represents one of the data points in the data set and the average value of the predictions for all the data points is plotted with a thick blue line. Figures 14a and 15a clearly show that the predicted optimal wall thickness values increase with the wall height while the range of optimal wall thickness values becomes wider as the wall height increases. A less steep increase in the optimal wall thickness values can be observed with respect to the  $\gamma$  values in Figures 14b and 15b. Also, for any given  $\gamma$  value, the optimal wall thicknesses take values on a much wider range compared to the wall height. Finally, the ICE plots of  $C_c$  and  $C_s$  show that these features have significantly less impact on the model output compared to H and  $\gamma$ .

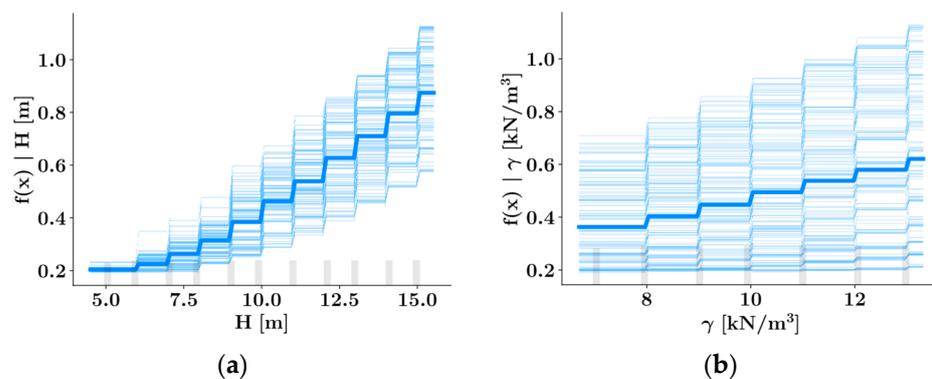


Figure 14. Cont.

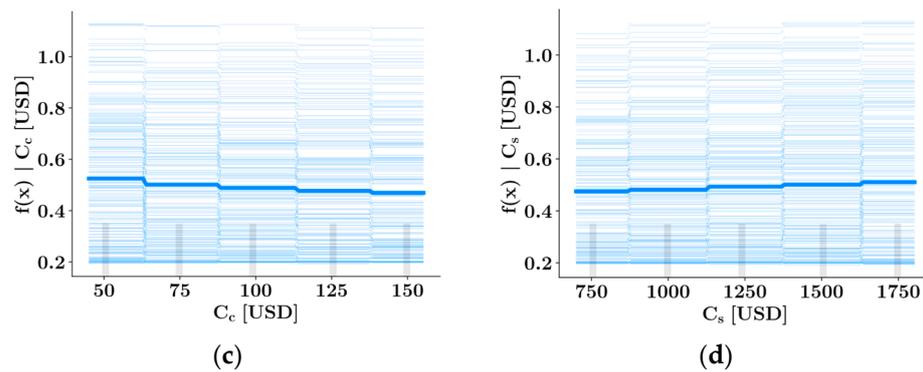


Figure 14. ICE plots (XGBoost) with post-tensioning for (a) H (b)  $\gamma$  (c)  $C_c$  (d)  $C_s$ .

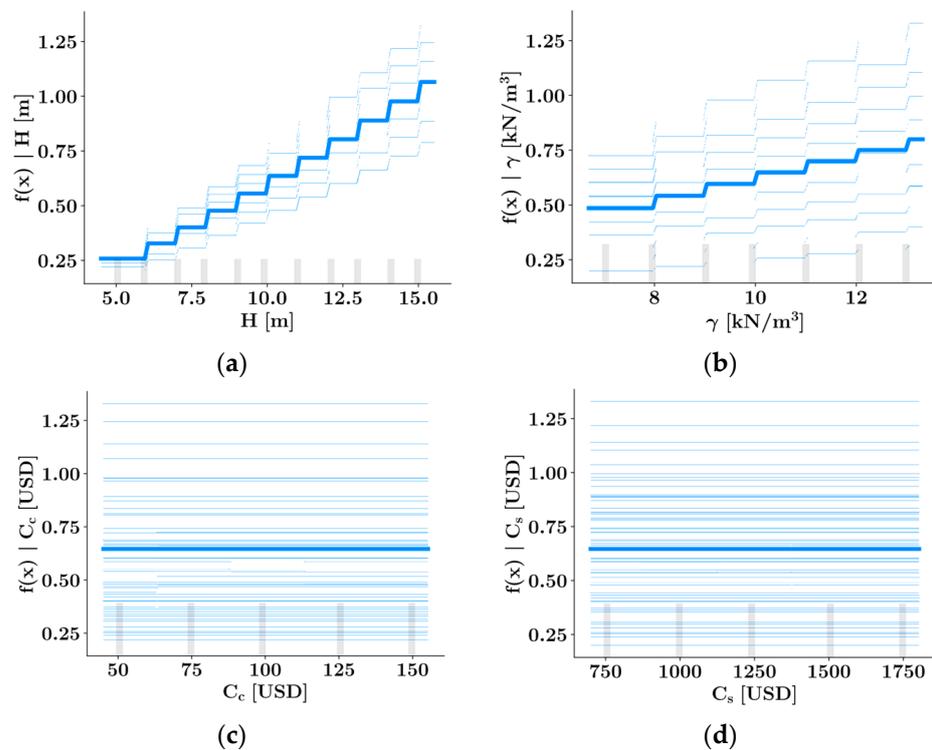


Figure 15. ICE plots (XGBoost) without post-tensioning for (a) H (b)  $\gamma$  (c)  $C_c$  (d)  $C_s$ .

### 3.2. Genetic Programming

Equations (16) and (17) have been obtained by the genetic programming algorithm for the prediction of the optimal wall thickness as a function of the wall height and liquid specific weight for the case of no post-tensioning and one post-tensioning cable, respectively. The tree representations of Equations (16) and (17) are shown in Figure 16. The tree population size was 5000 and mean absolute error (MAE) was used as the error metric. MAE values of 0.041 and 0.046 could be achieved in less than 50 iterations in cases with and without post-tensioning, respectively.

$$t(H, \gamma) = 0.095 \cdot H \cdot \log(0.878 \cdot \log(\gamma)) \tag{16}$$

$$t(H, \gamma) = \tan\left(\tan\left(0.093 \cdot H \cdot \tan\left(\log\left(\sqrt{\log(\gamma)}\right)\right)\right)\right) \tag{17}$$

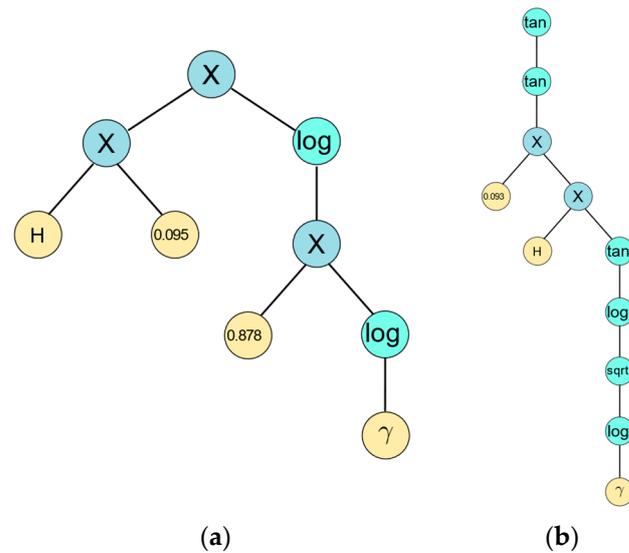


Figure 16. Tree representations of (a) Equation (16), (b) Equation (17).

Figures 17 and 18 show the development of the equation length and accuracy throughout the generations. It can be seen from Figure 17 that after 50 iterations an MAE value of 0.046 could be achieved with a predictive equation that consists of 9 components in the load case without post-tensioning. These 9 components consist of the log function which appears 2 times, the multiplication operator which appears 3 times, the coefficients 0.095 and 0.878, and the variables H and  $\gamma$ . Figure 18 shows that after 50 iterations an MAE value of 0.041 is achieved with a predictive equation consisting of 11 components. These components are the tangent function which appears 3 times, the log function which appears 2 times, the square root function which appears 1 time, the multiplication operator which appears 2 times, the coefficient 0.093, and the variables H and  $\gamma$ . The larger MAE value associated with the case without post-tensioning is due to the larger wall thicknesses necessary in the absence of post-tensioning. A complete list of the hyperparameters modified in obtaining Equations (16) and (17) is given in Table 3. For the remaining hyperparameters, the default values have been used.

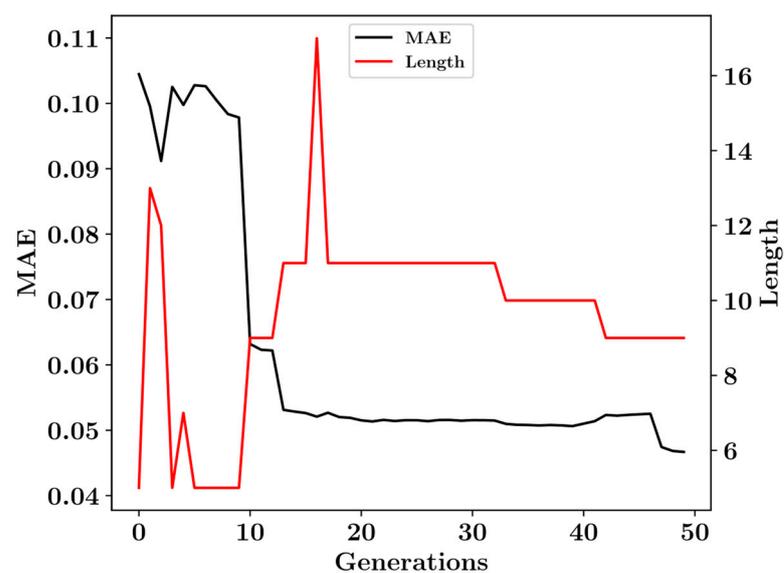


Figure 17. Development of the genetic programs without post-tensioning cables.

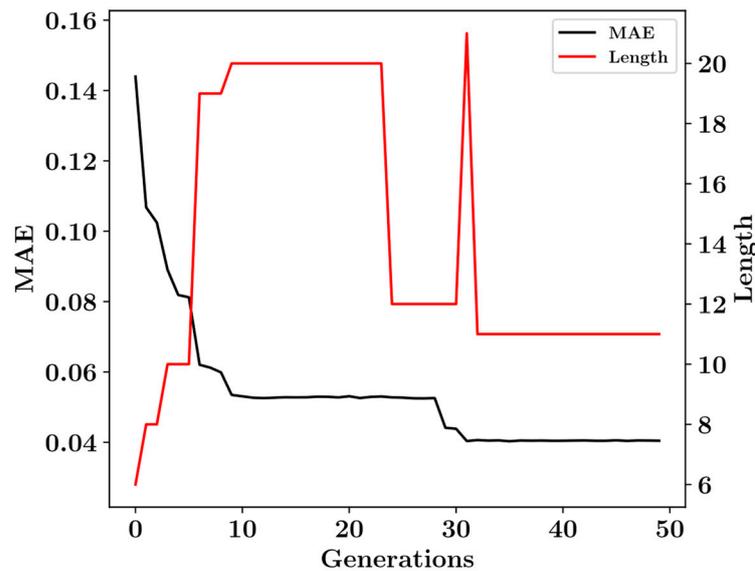


Figure 18. Development of the genetic programs for a single level of post-tensioning cables.

Table 3. Hyperparameters for the Genetic Programmin (GP) Model of the wall thickness.

Model	Parameter Name	Value
GP	population_size	5000
	p_crossover	0.7
	p_subtree_mutation	0.1
	p_hoist_mutation	0.05
	p_point_mutation	0.1
	tournament_size	150
	function_set	('add', 'sin', 'cos', 'tan', 'log', 'sub', 'mul', 'div', 'sqrt')

#### 4. Discussion

The current study demonstrates the application of the harmony search optimization algorithm in generating large data sets for training state-of-the-art machine learning models. The problem of predicting the optimal wall thickness of a cylindrical post-tensioned reinforced concrete wall of a liquid container has been investigated. In addition to ensemble learning models such as XGBoost and CatBoost also, the genetic programming methodology has been utilized to obtain closed-form predictive equations. The ensemble learning models have been trained with the height of the wall, specific weight of the liquid, and unit costs of steel and concrete as the input features. It was shown that the inclusion of post-tensioning cables leads to a decrease in the optimal wall thickness. The output of the ensemble learning models has been further analyzed using the SHAP approach and individual conditional expectation plots. The SHAP analysis showed that the wall height is the most impactful input feature affecting the model predictions of the optimal wall thickness followed by the liquid specific weight. An increase in the wall height and liquid specific weight was shown to also increase the optimal wall thickness. Furthermore, the genetic programming methodology provided predictive equations as functions of wall height and liquid specific weight for the load cases with and without post-tensioning. All predictive models were able to provide highly accurate predictions of the optimal wall thickness, with of the ensemble learning models achieving  $R^2$  scores greater than 0.98 on the test sets. The LightGBM model delivered the fastest and most accurate predictions in the case of walls with post-tensioning whereas the CatBoost model was the slowest.

## 5. Conclusions

The application of optimization techniques to generate big data sets is a novel approach to overcome the lack of experimental data points for the training of ML models related to structural engineering. Furthermore, the availability of closed-form equations for the prediction of optimal design configurations has significant practical benefits. The current study demonstrated the application of the genetic programming methodology in obtaining closed-form equations for the prediction of the optimal wall thickness for a liquid containing post-tensioned cylindrical wall. The data set necessary for this procedure was generated using the harmony search optimization technique. The obtained predictive equations could be incorporated into the engineering design process to facilitate structural optimization. However, further research needs to be done with larger data sets for the performance validation of the presented equations. On the other hand, higher  $R^2$  scores could be obtained from the ensemble learning algorithms. A limitation of the current study is the range of design variables included in the data set generation process. Future research in this area can incorporate additional input features into the data set such as the mechanical properties of steel and concrete or the positions of post-tensioning cables. Data sets can also be further enhanced with the results of numerical studies.

**Author Contributions:** Writing—original draft preparation, G.B. and C.C.; Conceptualization, G.B. and C.C.; data curation, C.C.; visualization, C.C.; writing—review & editing, G.B., S.K. and Z.W.G.; supervision G.B. and Z.W.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** Data will be made available upon reasonable request.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Nomenclature

$C_c$	Unit cost of concrete	$C_c$	Unit cost of steel
$C_{pt}$	Unit cost of post-tensioning	$C_{fw}$	Unit cost of formwork
$F$	Set of all input features	FPA	Flower pollination algorithm
FRP	Fiber reinforced polymer	HMCR	Harmony memory consideration rate
HMS	Harmony memory size	$i$	Index of a design variable
$\gamma$	Specific weight of the liquid	GOSS	Gradient-based One-Side Sampling
GP	Genetic programming	$H$	Height of the wall
HS	Harmony search	ICE	Individual conditional expectation
$k$	Index of a population member	$M$	Size of the training set
MAE	Mean absolute error	ML	Machine learning
$N$	Number of decision trees	$\Omega$	Penalty function
$P_1 \dots P_n$	Post-tensioning forces	PAR	Pitch adjustment rate
$\phi_i$	SHAP value of the $i$ -th feature	PCPT	Pre-cast post-tensioned
$r$	Radius of the cylindrical wall	$R^2$	Coefficient of determination
RMSE	Root mean squared error	$r_{xy}$	Pearson correlation coefficient
SHAP	SHapley Additive exPlanations	$t$	Wall thickness
TLBO	Teaching learning based optimization	SPM	Superposition method
$V_c$	Volume of concrete	$W_{pt}$	Weight of post-tensioning cables
$W_s$	Weight of steel	$\hat{y}$	Model prediction
$A_{fw}$	Area of formwork		

## References

1. *Post-Tensioning Manual*, 6th ed.; Post-Tensioning Institute: Phoenix, AZ, USA, 2006; ISBN 0977875202.
2. Bekdaş, G. Harmony Search Algorithm Approach for Optimum Design of Post-Tensioned Axially Symmetric Cylindrical Reinforced Concrete Walls. *J. Optim. Theory Appl.* **2014**, *164*, 342–358. [[CrossRef](#)]
3. Bekdas, G. Optimum design of axially symmetric cylindrical reinforced concrete walls. *Struct. Eng. Mech.* **2014**, *51*, 361–375. [[CrossRef](#)]

4. Bekdaş, G. New improved metaheuristic approaches for optimum design of posttensioned axially symmetric cylindrical reinforced concrete walls. *Struct. Des. Tall Spec. Build.* **2018**, *27*, e1461. [[CrossRef](#)]
5. Bekdaş, G.; Nigdeli, S.M. Optimum Reduction of Flexural Effect of Axially Symmetric Cylindrical Walls with Post-tensioning Forces. *KSCE J. Civ. Eng.* **2017**, *22*, 2425–2432. [[CrossRef](#)]
6. Bekdaş, G.; Cakiroglu, C.; Islam, K.; Kim, S.; Geem, Z.W. Optimum Design of Cylindrical Walls Using Ensemble Learning Methods. *Appl. Sci.* **2022**, *12*, 2165. [[CrossRef](#)]
7. Kaveh, A.; Zolghadr, A. Meta-heuristic methods for optimization of truss structures with vibration frequency constraints. *Acta Mech.* **2018**, *229*, 3971–3992. [[CrossRef](#)]
8. Renkavieski, C.; Parpinelli, R.S. Meta-heuristic algorithms to truss optimization: Literature mapping and application. *Expert Syst. Appl.* **2021**, *182*, 115197. [[CrossRef](#)]
9. Mortazavi, A.; Togan, V. Metaheuristic Algorithms for Optimal Design of Truss Structures. In *Advances in Structural Engineering—Optimization. Studies in Systems, Decision and Control*; Nigdeli, S.M., Bekdaş, G., Kayabekir, A.E., Yucel, M., Eds.; Springer: Cham, Switzerland, 2021; Volume 326. [[CrossRef](#)]
10. Mortazavi, A. The Performance Comparison of Three Metaheuristic Algorithms on the Size, Layout and Topology Optimization of Truss Structures. *Mugla J. Sci. Technol.* **2019**, *5*, 28–41. [[CrossRef](#)]
11. Bekdaş, G.; Yucel, M.; Nigdeli, S.M. Evaluation of Metaheuristic-Based Methods for Optimization of Truss Structures via Various Algorithms and Lévy Flight Modification. *Buildings* **2021**, *11*, 49. [[CrossRef](#)]
12. Grzywinski, M.; Selejdak, J.; Dede, T. Shape and size optimization of trusses with dynamic constraints using a metaheuristic algorithm. *Steel Compos. Struct.* **2019**, *33*, 747–753. [[CrossRef](#)]
13. Yousefpoor, H.; Kaveh, A. Chaos Embedded Meta-heuristic Algorithms for Optimal Design of Truss Structures. *Sci. Iran.* **2022**, *29*, 2868–2885. [[CrossRef](#)]
14. Uray, E.; Tan, Ö.; Çarbaş, S.; Erkan, H. Metaheuristics-based Pre-Design Guide for Cantilever Retaining Walls. *Tek. Dergi* **2021**, *32*, 10967–10993. [[CrossRef](#)]
15. Sharma, S.; Saha, A.K.; Lohar, G. Optimization of weight and cost of cantilever retaining wall by a hybrid metaheuristic algorithm. *Eng. Comput.* **2021**, *38*, 2897–2923. [[CrossRef](#)]
16. Aydoğdu, I. Comparison of metaheuristics on multi objective (Cost-C02) optimization of RC cantilever retaining walls. *Pamukkale Univ. J. Eng. Sci.* **2017**, *23*, 221–231. [[CrossRef](#)]
17. Bekdaş, G.; Cakiroglu, C.; Kim, S.; Geem, Z.W. Optimal Dimensioning of Retaining Walls Using Explainable Ensemble Learning Algorithms. *Materials* **2022**, *15*, 4993. [[CrossRef](#)]
18. Kaveh, A.; Hamedani, K.B.; Bakhshpoori, T. Optimal Design of Reinforced Concrete Cantilever Retaining Walls Utilizing Eleven Meta-Heuristic Algorithms: A Comparative Study. *Period. Polytech. Civ. Eng.* **2020**, *64*, 156–158. [[CrossRef](#)]
19. Kalemci, E.N.; Ikizler, S.B.; Dede, T.; Angın, Z. Design of reinforced concrete cantilever retaining wall using Grey wolf optimization algorithm. *Structures* **2019**, *23*, 245–253. [[CrossRef](#)]
20. Aral, S.; Yılmaz, N.; Bekdaş, G.; Nigdeli, S.M. Jaya Optimization for the Design of Cantilever Retaining Walls with Toe Projection Restriction. In *Proceedings of the 6th International Conference on Harmony Search, Soft Computing and Applications (ICHSA 2020)*, Istanbul, Turkey, 22–24 April 2020; Nigdeli, S.M., Kim, J.H., Bekdaş, G., Yadav, A., Eds.; *Advances in Intelligent Systems and Computing*. Springer: Singapore, 2021; Volume 1275. [[CrossRef](#)]
21. Kalemci, E.N. Rao-3 algorithm for the weight optimization of reinforced concrete cantilever retaining wall. *Geomech. Eng.* **2020**, *20*, 527–536. [[CrossRef](#)]
22. Cakiroglu, C.; Islam, K.; Bekdaş, G.; Nehdi, M.L. Data-driven ensemble learning approach for optimal design of cantilever soldier pile retaining walls. *Structures* **2023**, *51*, 1268–1280. [[CrossRef](#)]
23. Jalili, S.; Khani, R.; Maheri, A.; Hosseinzadeh, Y. Performance assessment of meta-heuristics for composite layup optimisation. *Neural Comput. Appl.* **2021**, *34*, 2031–2054. [[CrossRef](#)]
24. Nicholas, P.E.; Dharmaraja, C.; Sofia, A.S.; Vasudevan, D. Optimization of laminated composite plates subjected to nonuniform thermal loads. *Polym. Polym. Compos.* **2019**, *27*, 314–322. [[CrossRef](#)]
25. Cakiroglu, C.; Islam, K.; Bekdaş, G.; Kim, S.; Geem, Z.W. Metaheuristic Optimization of Laminated Composite Plates with Cut-Outs. *Coatings* **2021**, *11*, 1235. [[CrossRef](#)]
26. Nicholas, P.E.; Padmanaban, K.; Vasudevan, D. Buckling optimization of laminated composite plate with elliptical cutout using ANN and GA. *Struct. Eng. Mech.* **2014**, *52*, 815–827. [[CrossRef](#)]
27. Javidrad, F.; Nazari, M.; Javidrad, H.R. An Innovative Optimized Design for Laminated Composites in terms of a Proposed Bi-Objective Technique. *J. Soft Comput. Civ. Eng.* **2020**, *4*, 1–28. [[CrossRef](#)]
28. Topal, U.; Vo-Duy, T.; Dede, T.; Nazarimofrad, E. Buckling load optimization of laminated plates resting on Pasternak foundation using TLBO. *Struct. Eng. Mech. Int. J.* **2018**, *67*, 617–628. [[CrossRef](#)]
29. Jafari, M.; Chaleshtari, M.H.B.; Khoramishad, H.; Altenbach, H. Minimization of thermal stress in perforated composite plate using metaheuristic algorithms WOA, SCA and GA. *Compos. Struct.* **2023**, *304*, 116403. [[CrossRef](#)]
30. Cakiroglu, C.; Bekdaş, G.; Geem, Z.W. Harmony Search Optimisation of Dispersed Laminated Composite Plates. *Materials* **2020**, *13*, 2862. [[CrossRef](#)]
31. Cakiroglu, C.; Bekdaş, G.; Kim, S.; Geem, Z.W. Optimisation of Shear and Lateral-Torsional Buckling of Steel Plate Girders Using Meta-Heuristic Algorithms. *Appl. Sci.* **2020**, *10*, 3639. [[CrossRef](#)]

32. Elbelbisi, A.H.; El-Sisi, A.A.; Hassan, H.A.; Salim, H.A.; Shabaan, H.F. Parametric Study on Steel–Concrete Composite Beams Strengthened with Post-Tensioned CFRP Tendons. *Sustainability* **2022**, *14*, 15792. [CrossRef]
33. Elsheshtawy, S.S.; Shoeib, A.K.; Hassanin, A.; Ors, D.M. Influence of the Distribution and Level of Post-Tensioning Force on the Punching Shear of Flat Slabs. *Designs* **2022**, *7*, 1. [CrossRef]
34. Attia, M.M.; Khalil, A.H.H.; Mohamed, G.N.; Samaan, M.F.; Katunský, D. Nonlinear Behavior of Bonded and Unbonded Two-Way Post-Tensioned Slabs Pre-Strengthened with CFRP Laminates. *Buildings* **2022**, *13*, 35. [CrossRef]
35. Tahmasebinia, F.; Hu, Z.; Wei, Q.; Ma, W. Numerically Evaluation of Dynamic Behavior of Post-Tensioned Concrete Flat Slabs under Free Vibration. *Sustainability* **2023**, *15*, 845. [CrossRef]
36. Vavrus, M.; Kralovanec, J. Study of Application of Fiber Reinforced Concrete in Anchorage Zone. *Buildings* **2023**, *13*, 524. [CrossRef]
37. Lei, H.; Liu, Z.; Alsomiri, M. Bearing capacity of concentric anchorage zones in post-tensioned members: A stress field solution. *Structures* **2023**, *50*, 1368–1375. [CrossRef]
38. Joyklad, P.; Ali, N.; Chaiyasarn, K.; Suparp, S.; Hussain, Q. Time-dependent behavior of full-scale precast post-tensioned (PCPT) girders: Experimental and finite element analysis. *Case Stud. Constr. Mater.* **2022**, *17*, e01310. [CrossRef]
39. Yu, F.; Wang, M.; Yao, D.; Liu, Y. Experimental research on flexural behavior of post-tensioned self-compacting concrete beams with recycled coarse aggregate. *Constr. Build. Mater.* **2023**, *377*, 131098. [CrossRef]
40. Bekdaş, G.; Yucel, M.; Nigdeli, S.M. Generation of eco-friendly design for post-tensioned axially symmetric reinforced concrete cylindrical walls by minimizing of CO<sub>2</sub> emission. *Struct. Des. Tall Spéc. Build.* **2022**, *31*, e1948. [CrossRef]
41. ACI 318M–05; Building Code Requirements for Structural Concrete and Commentary. American Concrete Institute: Farmington Hills, MI, USA, 2005.
42. Love, A. A Treatise on the Mathematical Theory of Elasticity. 1, 1892. (hal-01307751). Available online: <https://hal.science/hal-01307751> (accessed on 23 February 2023).
43. Hetenyi, M. *Beams on Elastic Foundation*; The University of Michigan Press: Ann Arbor, MI, USA, 1946.
44. Geem, Z.W.; Kim, J.H.; Loganathan, G.V. A new heuristic optimization algorithm: Harmony search. *Simulation* **2001**, *76*, 60–68. [CrossRef]
45. Yadav, N.; Ngo, T.T.; Kim, J.H. An algorithm for numerical solution of differential equations using harmony search and neural networks. *J. Appl. Anal. Comput.* **2022**, *12*, 1277–1293. [CrossRef]
46. Geem, Z.W. Multiobjective Optimization of Time-Cost Trade-Off Using Harmony Search. *J. Constr. Eng. Manag.* **2010**, *136*, 711–716. [CrossRef]
47. Chang, Y.-Z.; Li, Z.-W.; Kou, Y.-X.; Sun, Q.-P.; Yang, H.-Y.; Zhao, Z.-Y. A New Approach to Weapon-Target Assignment in Cooperative Air Combat. *Math. Probl. Eng.* **2017**, *2017*, 2936279. [CrossRef]
48. Shih, P.-C.; Chiu, C.-Y.; Chou, C.-H. Using Dynamic Adjusting NGHS-ANN for Predicting the Recidivism Rate of Commuted Prisoners. *Mathematics* **2019**, *7*, 1187. [CrossRef]
49. Xu, C.; Wang, X. Transient content caching and updating with modified harmony search for Internet of Things. *Digit. Commun. Netw.* **2018**, *5*, 24–33. [CrossRef]
50. Gonzalez, P.; Mora, A.; Garrido, S.; Barber, R.; Moreno, L. Multi-LiDAR Mapping for Scene Segmentation in Indoor Environments for Mobile Robots. *Sensors* **2022**, *22*, 3690. [CrossRef]
51. Cheng, Y.; Li, L.; Lansivaara, T.; Chi, S.; Sun, Y. An improved harmony search minimization algorithm using different slip surface generation methods for slope stability analysis. *Eng. Optim.* **2008**, *40*, 95–115. [CrossRef]
52. Lee, K.S.; Geem, Z.W. A new structural optimization method based on the harmony search algorithm. *Comput. Struct.* **2004**, *82*, 781–798. [CrossRef]
53. Lee, K.S.; Geem, Z.W.; Lee, S.-H.; Bae, K.-W. The harmony search heuristic algorithm for discrete structural optimization. *Eng. Optim.* **2005**, *37*, 663–684. [CrossRef]
54. García-Segura, T.; Yepes, V.; Alcalá, J. Computer-support tool to optimize bridges automatically. *Int. J. Comput. Methods Exp. Meas.* **2017**, *5*, 171–178. [CrossRef]
55. Aladsani, M.A.; Burton, H.; Abdullah, S.A.; Wallace, J.W. Explainable Machine Learning Model for Predicting Drift Capacity of Reinforced Concrete Walls. *ACI Struct. J.* **2022**, *119*, 191–204. [CrossRef]
56. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.Y. Lightgbm: A highly efficient gradient boosting decision tree. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 3146–3154.
57. Dorogush, A.V.; Ershov, V.; Gulin, A. CatBoost: Gradient boosting with categorical features support. *arXiv* **2018**, arXiv:1810.11363.
58. Gandomi, A.H.; Roke, D.A. Assessment of artificial neural network and genetic programming as predictive tools. *Adv. Eng. Softw.* **2015**, *88*, 63–72. [CrossRef]
59. Madár, J.; Abonyi, J.; Szeifert, F. Genetic Programming for the Identification of Nonlinear Input–Output Models. *Ind. Eng. Chem. Res.* **2005**, *44*, 3178–3186. [CrossRef]

60. Gondia, A.; Ezzeldin, M.; El-Dakhakhni, W. Mechanics-Guided Genetic Programming Expression for Shear-Strength Prediction of Squat Reinforced Concrete Walls with Boundary Elements. *J. Struct. Eng.* **2020**, *146*, 04020223. [[CrossRef](#)]
61. Mangalathu, S.; Hwang, S.-H.; Jeon, J.-S. Failure mode and effects analysis of RC members based on machine-learning-based SHapley Additive exPlanations (SHAP) approach. *Eng. Struct.* **2020**, *219*, 110927. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.