applied sciences

MDPI

Article

Prediction of the Yield Strength of RC Columns Using a PSO-LSSVM Model

Bochen Wang [1,2], Weiming Gong [1,2,*], Yang Wang [1,2], Zele Li [1,2] and Hongyuan Liu [3]

[1] Key of Laboratory for Concrete and Prestressed Concrete Structures of Ministry of Education, Southeast University, Nanjing 211100, China
[2] School of Civil Engineering, Southeast University, Nanjing 211100, China
[3] School of Civil Engineering and Architecture, Jiangsu University of Science and Technology, Zhenjiang 212110, China
* Correspondence: wmgong@seu.edu.cn

Abstract: Accuracy prediction of the yield strength and displacement of reinforced concrete (RC) columns for evaluating the seismic performance of structure plays an important role in engineering the structural design of RC columns. A new hybrid machine learning technique based on the least squares support vector machine (LSSVM) and the particle swarm optimization (PSO) algorithm is proposed to predict the yield strength and displacement of RC columns. In this PSO-LSSVM model, the LSSVM is applied to discover the mapping between the influencing factors and the yield strength and displacement, and the PSO algorithm is utilized to select the optimal parameters of LSSVM to facilitate the prediction performance of the proposed model. A dataset covering the PEER database and the available experimental data in relevant literature is established for model training and testing. The PSO algorithm is then evaluated and compared with other metaheuristic algorithms based on the experiment's database. The results indicate the effectiveness of the PSO employed for improving the prediction performance of the LSSVM model according to the evaluation criteria such as the root mean square error (RMSE), mean absolute error (MAE) and coefficient of determination (R^2). Overall, the simulation demonstrates that the developed PSO-LSSVM model has ideal prediction accuracy in the yield properties of RC columns.

Keywords: RC column; skeleton curve; yield strength; least squares support vector machine; particle swarm optimization

Citation: Wang, B.; Gong, W.; Wang, Y.; Li, Z.; Liu, H. Prediction of the Yield Strength of RC Columns Using a PSO-LSSVM Model. *Appl. Sci.* **2022**, *12*, 10911. https://doi.org/10.3390/app122110911

Academic Editor: Sang Whan Han

Received: 9 October 2022
Accepted: 26 October 2022
Published: 27 October 2022

1. Introduction

Reinforced concrete (RC) structures are the main parts of modern buildings, and RC columns are one of the crucial weight-bearing components in RC structures, which is closely related to the seismic performances of buildings [1,2]. The post-earthquake reconnaissance [3] demonstrates that the collapse of buildings in the earthquake was mostly caused by the destruction and failure of RC columns, which play an important role in resisting lateral load. It is thus necessary to evaluate the existing structural performance and load-carrying capacity of post-earthquake RC columns, and quickly identify the most collapse-prone RC frames across the region, providing insights for repair strategies and useful information to emergency teams for making plans for safety rescues. Seismic resistance assessment of post-earthquake RC structures is an important issue that has received widespread attention in the structural engineering field. Several existing modeling approaches were proposed for accurately estimating the residual collapse capacity of buildings, but they are time-consuming and have low prediction accuracy and lack general applicability [4–6]. Hence, efficient and accurate prediction of the yield strength and displacement of RC columns has great practical significance for evaluating the seismic performance and engineering design of RC structures [7].

Several numerical and experimental [8] methods have been extensively employed to investigate the nonlinear behaviors of RC structures. The numerical models, based on the finite element method, are usually used for nonlinear static or dynamic analyses. However, many FE-based models are incapable of capturing the hysteretic behaviors accurately, something which depends on the selection of model parameters and boundary conditions [4–6]. It is, therefore necessary to further improve the assessment accuracy of the behaviors of RC columns by the numerical simulation method. Additionally, the calculation of tedious iterative processes is time-consuming. In contrast, the experimental method is more effective and capable of identifying the real responses of a structure [9,10], but it is quite costly, time-consuming and labor-intensive.

Over the past few decades, artificial intelligence (AI) technologies, with the merits of their strong nonlinear learning capacities, have been widely utilized in the earthquake and civil engineering fields [11–16]. The technologies have gradually developed into an effective research method to study the abovementioned problems [17–24], one which enables civil engineers to predict the structural performance of concrete members. For instance, Quaranta et al. [24] performed innovative work in which a machine-learning-aided approach was proposed to improve the accuracy of the mechanics-based shear capacity equation for RC beams and columns. On the other hand, many models based on hybrid machine learning (ML) techniques were developed to predict structural performance straightforwardly without explicit formulae. Typical of these predictive models, the support vector machine (SVM) proposed by Vapnik [25] is a widely used machine learning technique. Thereafter, the least squares support vector machine (LSSVM) [26], as a modified version of SVM, is reported to be able to conduct a faster training process for tackling many complex and non-linear problems in engineering compared to the standard SVMs, i.e., with faster computational and generalized capacity [27,28]. Some work on the application of machine learning (ML) to the study of RC structures is listed in Table 1, as follows:

Table 1. Application of machine learning in the study of RC structures.

Study	Model	Issues or Objectives
Xu et al. [18]	LSSVM	Predict the strength of radiation-shielding concrete
Vu and Hoang [19]	LSSVM	Predict the ultimate punching shear capacity of FRP-RC slabs.
Luo and Paal [20]	ML-BCV	Predict backbone curves of RC columns
Mangalathu and Jeon [21]	MLs	Predict shear strength for RC beam-column joints
Ning et al. [22]	DE-ANN	Predict the hysteresis loop of RC columns

It is worth noting that Luo and Paal [20] developed a novel machine learning–based backbone curve (ML-BCV) model to predict backbone curves of RC columns subjected to three different failure modes. This model consists of a modified LSSVM to address the multioutput case, and a grid search algorithm (GSA) to facilitate the training process. The essence of the training process is to select the optimal solution for two key parameters that significantly affect the accuracy level of predicted results. The selection process of these two parameters can be regarded as an optimization problem [20]. Thus, an appropriate swarm intelligence (SI) algorithm is needed to solve the optimization problem. Several SI algorithms have been widely applied for solving various optimization problems owing to their powerful global search capability [29], typically genetic algorithm (GA) [30], differential evolution (DE) [31], artificial bee colony (ABC) [32] and particle swarm optimization (PSO) [33] algorithms. The PSO algorithm developed by Kennedy and Eberhart [33], is regarded as a reliable tool when combined with AI techniques due to its outstanding features of high calculation efficiency, high accuracy and powerful global optimization and search capability in searching for the optimal solution [34–37]. The PSO has been used to enhance the training process of LSSVM or other models, and further improve the predictive

accuracy of forecasting models in the fields of engineering [38–42]. A detailed introduction to the PSO algorithm used in the present work can be found in Section 4.2.

To the best of our knowledge, there are still few computational models that can rapidly calculate the strength-related properties of RC columns with high accuracy [7,43,44]. Moreover, the predictive models or formulae that enable the comprehensive covering of most of the influencing factors are still scarce due to the complexity of multi-factor issues. Based on the challenge, machine learning (ML) techniques are widely adopted by virtue of their data-driven nature and their expertise in dealing with nonlinear correlations between multiple factors.

The present work focuses on developing a new prediction model based on machine learning techniques for achieving the goal of efficiently and accurately predicting yield strength and displacement of RC columns with more influence factors considered. In this model, the LSSVM is applied to establish a prediction model to estimate the nonlinear mapping between influence factors and yield properties. The PSO algorism is adopted to facilitate the training process of the LSSVM model for the purpose of improving the predictive accuracy of this model through selecting the optimal solution of key parameters i.e., regularization parameter (γ) and RBF kernel parameter (σ). For training and testing the performance of the present model, a test database of the proposed model is built up by an efficient program script compiled to collect and collate the existing experimental data obtained by pseudo-static cyclic testing of RC columns. The performance of the trained LSSVM model can be evaluated by the root mean square error (RMSE), mean absolute error (MAE), explained variance (EV) and coefficient of determination (R^2). Meanwhile, the effectiveness of the PSO in enhancing the prediction accuracy of the LSSVM model is evaluated compared to other metaheuristics.

2. Pseudo-Static Cyclic Test and Yield Point

2.1. Pseudo-Static Cyclic Test

The pseudo-static cyclic test, pseudo-dynamic test and simulated seismic shaking table test are three test methods which are commonly used for testing the seismic performance of building structures and their components. Therein, the pseudo-static cyclic test is an approach which records the whole process from the elastic stage to the failure or damage by controlling the load or displacement for low circumferential reciprocal cyclic loading [20,45]. The pseudo-static test is widely adopted by researchers in the field of study of the seismic performance of structures or components relying on its advantages, such as the better adaptability of the site and experimental equipment, which is convenient for conducting full-scale experiments, and low loading rate that allows detailed observation of the whole process of specimen failure.

Under the action of low circumferential cyclic load, a hysteresis loop denoting the relationship between the load and the corresponding deformation of the structure is formed during a loading and unloading process, and the series of hysteresis loops formed after multiple cycles of loading constitute the hysteretic curves of the structure, namely the load-displacement curves. Then, the skeleton curve of the specimen shown in Figure 1 can be obtained by connecting the peak points of each hysteretic curve of the specimen, which reflects the relationship between the force and deformation of the experimental model. The skeleton curve is a comprehensive reflection of the seismic performance of the structure; it thus can be an important basis for analyzing the elastic-plastic dynamic response of the specific structure. For the concrete specimen, its skeleton curve is taken as the envelope connected by the peak points of the first cycle at each loading level of each load-displacement curve, i.e., the envelope of the hysteretic curve.

In recent years, many experimental studies on the seismic performance of RC columns have been conducted by using the pseudo-static cyclic test [46,47]. The influences of relevant factors on their seismic performance are investigated. The load-displacement curve can be obtained through the experimental test, and it reflects the nonlinear performance, energy dissipation characteristics and ultimate damage mechanism of the structure or component.

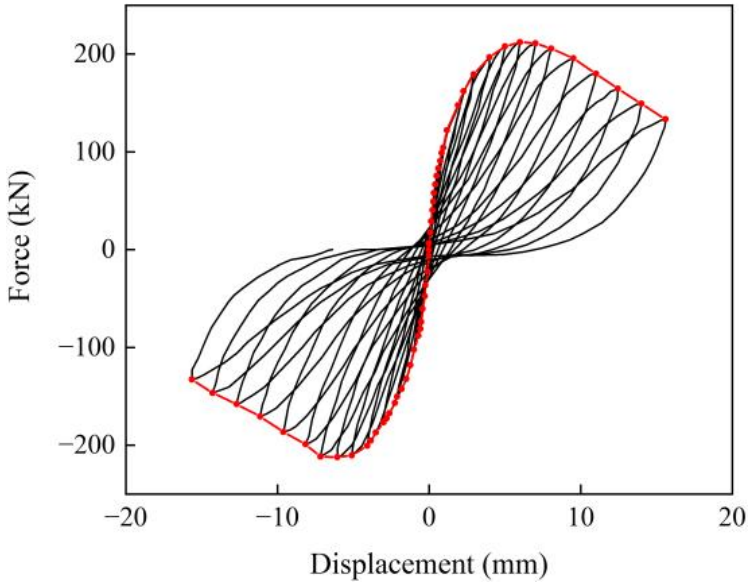


Figure 1. The hysteretic curve and skeleton curve.

2.2. Yield Point

It is well known that how to judge the yield point on the specific skeleton curve is still a complicated problem, and there is no unified standard in the related research field [7,43,48]. Here are three judgment methods commonly used to determine the yield point as follows:

1. For the Geometric Graphic Method shown in Figure 2a, the yield displacement and yield point are defined as follows: Draw the tangent line to the curve at 'O'. The line is extended to the intersection with a horizontal line through 'D' at 'A', where the 'D' corresponds to the maximum applied shear P_{max} shown in Figure 2a. The perpendicular of line 'OA' intersects the curve at 'B'. Connect 'O' and 'B' and extend line 'OB' to meet the horizontal line 'DA' at 'C', and then project onto the horizontal axis to obtain the yield displacement Δ_y and the yield point 'E' on the curve, where it corresponds to the yield applied shear P_y.
2. For R. Park Method shown in Figure 2b [48], the yield displacement and yield point are defined as follows: A secant 'OB' is drawn to intersect the lateral load-displacement relation at a certain proportion of the maximum applied shear, i.e., the 'B' on the curve corresponding to αP_{max} shown in Figure 2b. Similarly, this extension of line 'OB' and the horizontal line corresponding to the maximum applied shear P_{max} intersect at 'A', and then projects onto the horizontal axis to obtain the yield displacement Δ_y. The intersection point 'C' of the vertical and curve is defined as the yield point, which corresponds to the yield applied shear P_y.
3. For Equivalent Elasto-Plastic Energy Method shown in Figure 2c, the yield displacement and yield point are defined as follows: Determine a point 'B' on the curve and draw secant 'OB' to intersect the curve for satisfying the principle that the energy absorbed by the ideal elastoplastic structure is equal to that absorbed by the actual structure, i.e., the areas of shaded area 'OAB' and 'BCD' are equal shown in Figure 2c. Similar to the R. Park Method in Figure 2b, this extension of line 'OB' and the horizontal line corresponding to the maximum applied shear P_{max} intersect at 'C', and is then projected onto the horizontal axis to obtain the yield displacement Δ_y. The intersection point 'E' of the vertical and curve is defined as the yield point, which corresponds to the yield applied shear P_y.

In this paper, the yield point of the skeleton curve of the load-displacement curve is calculated by using the Equivalent Elasto-Plastic Energy Method.

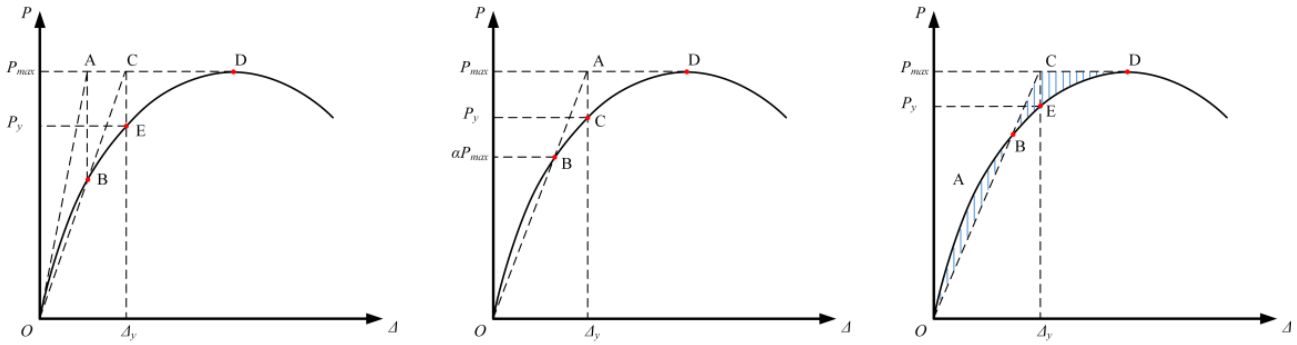


Figure 2. The yield point judgment methods: (**a**) Geometric Graphic Method; (**b**) R. Park Method; (**c**) Equivalent Elasto-Plastic Energy Method.

3. Data Collection and Pre-Processing

In this work, a quantity of data adopted for the following calculations is collected from the PEER database [49,50] and the relevant literature [22,44]. The PEER database covers extensive hysteretic curves of the RC columns with various material properties obtained by pseudo-static cyclic test. As for the collected original experimental data, relevant pre-processing should be conducted to obtain the corresponding yield point whose values are used to compare with predictive values (see Section 4 for details). The steps of data pre-processing are as follows: Firstly, determine the skeleton curve of each hysteretic curve according to the rule seen in Section 2.1; secondly, determine the yield point on the obtained skeleton curve according to the method seen in Section 2.2, and the yield strength and displacement are obtained. However, the amount of data involved in the data collection and pre-processing is enormous. To improve the data-processing efficiency, the above two steps are implemented via programming calculation. Thus, a program script was written for plotting the skeleton curves, and then the yield strength and yield displacement are both outputted. Thereafter, a total of 382 test results were selected and organized in bulk according to the parameter types, and then all the types of the parameter were categorized into two parts: input parameters and output properties.

The input parameters are the influencing factors including the basic parameters of the RC column. The output properties denote performances of the RC column, e.g., yield strength and displacement, which are codetermined by the influence factors. In the present model, nine representative influence factors are identified as the input parameters of the proposed model for the prediction of output properties. The set of influence factors can be expressed as $\{D_c, \rho_r, SecT, f_c, f_{yl}, f_{yv}, \rho_l, \rho_v, \rho_n\}$, where D_c is the section size; ρ_r is the span-to-depth ratio defined as L/D_c, in which L is length; $SecT$ is the section type, $SecT$ = 1 or $SecT$ = 2 denotes the square or circle section type; f_c is the strength of concrete; f_{yl} is the yield strength of longitudinal bars; f_{yv} is the yield strength of transverse bars; ρ_l is the longitudinal reinforcement ratio; ρ_v is the transverse reinforcement ratio; ρ_n is the axial load ratio defined by $\rho_n = P/(A_g \cdot f_c)$, in which P is the constant axial load and A_g is the gross section area. These input parameters, i.e., influence factors mentioned above, are summarized in Table 2.

The focus of the developed model is a prediction of the yield strength (F_y) and yield displacement (Δ_y) of RC columns, which are considered as the output properties in the present study. The ranges of the input parameters and output properties are presented in Table 3 in detail. It can be seen that the input parameters and output properties are well-distributed over a wide range. Thus, each specimen in the test database can be written as $\{\{D_c, \rho_r, SecT, f_c, f_{yl}, f_{yv}, \rho_l, \rho_v, \rho_n\}, \{F_y, \Delta_y\}\}$, which involves the input parameters and output properties. Figure 3 displays the histograms of the input parameters and output properties.

Table 2. Input parameters for LSSVM training.

Influencing Factors	Parameters	Description
Geometries	Dc	Section size (mm)
	$\rho_r = L/D_c$	Span-to-depth ratio (%)
	$SecT$	Section type
Materials	f_c	Strength of concrete (MPa)
	f_{yl}	Yield strength of longitudinal bars (MPa)
	f_{yv}	Yield strength of transverse bars (MPa)
Reinforcement	ρ_l	Longitudinal reinforcement ratio (%)
	ρ_v	Transverse reinforcement ratio (%)
Loading	$\rho_n = P/(A_g \cdot f_c)$	Axial load ratio

Table 3. Statistical description of the input and output variables.

Variables	Min.	Max.	Mean	Std.
Dc	80	1520	345.85	149.98
ρ_r	1	10	3.44	1.65
f_c	16	118	46.54	25.98
f_{yl}	0	587.1	422.84	72.41
f_{yv}	0	1424	454.52	206.60
ρ_l	0.0046	0.0603	0.03	0.01
ρ_v	0	4.27	0.40	0.69
ρ_n	−0.099	0.9	0.22	0.18
F_y	16.27	2654.11	202.49	204.25
Δ_y	0.54	114.70	12.64	14.55

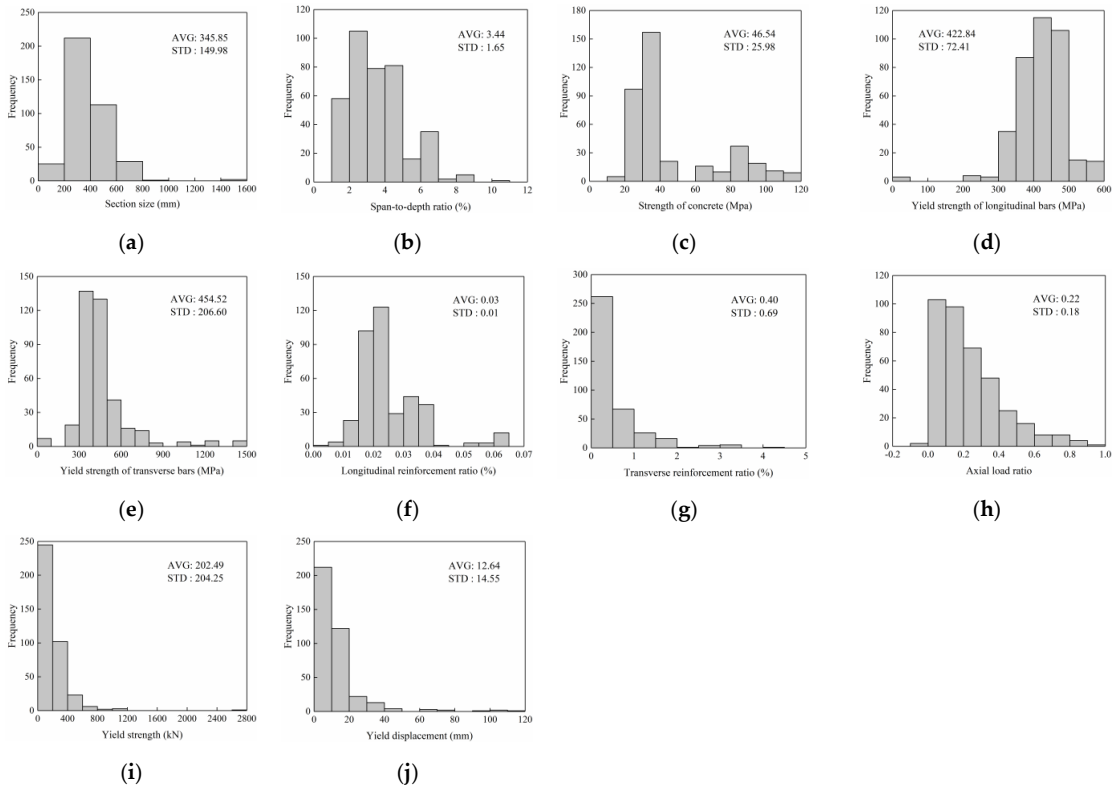


Figure 3. Statistical distribution of the input and output variables: (**a**) Section size; (**b**) Span-to-depth ratio; (**c**) Strength of concrete; (**d**) Yield strength of longitudinal bars; (**e**) Yield strength of transverse bars; (**f**) Longitudinal reinforcement ratio; (**g**) Transverse reinforcement ratio; (**h**) Axial load ratio; (**i**) Yield strength; (**j**) Yield displacement.

4. Model Development and Verification

4.1. Least Squares Support Vector Machine

The LSSVM, first introduced by Suykens et al. [26], is a supervised machine learning method based on statistical learning theory. It is well-known that the LSSVM methodology possesses an outstanding feature in learning nonlinear functions, which simply solves a set of linear equations. Therefore, LSSVM, as an extension of SVM equipped with the advantage of fast calculation and generalized capacity, has been widely applied for practical problems such as nonlinearly, high-dimensional input space and small training samples. These problems are generally divided into two categories, i.e., regression problems and classification problems. The focus of the present study is predicting the yield strength and displacement with the LSSVM adopted in this prediction model, which is thus considered as a regression problem. The governing equations and numerical algorithms are described in detail as follows.

For a set of given nonlinear training samples $\{x_i, y_i\}$, $i = 1,2, \ldots ,l$, where l is the number of training samples. $x_i \in R^n$ is the input vector, and $y_i \in R$ is its corresponding output vector. The regression function is obtained by mapping the input space to high-dimensional feature space using the nonlinear mapping function $\phi(\cdot)$, the equation of the regression function can be written as

$$f(x) = w^T \phi(x) + b \tag{1}$$

where w is the weight vector; b is the bias term (also called the partial vector).

The objective of LSSVM regression can be transformed into an optimization problem that is expressed as

$$\begin{gathered} \min_{w,b,\xi} J(w,\xi) = \frac{1}{2}\|w\|^2 + \frac{\gamma}{2}\sum_{i=1}^{l}\xi_i^2 \\ \text{s.t. } y_i = w^T f(x_i) + b + \xi_i,\ i = 1,2,\ldots,l \end{gathered} \tag{2}$$

where γ is the regularization parameter (also called the penalty parameter); $\xi \in R$ are the error variables.

The Lagrange function of the optimization problem, i.e., Equation (2) is formulated as follows:

$$L = J - \sum_{i=1}^{l} \alpha_i \left[w^T \cdot \phi(x_i) + b + \xi_i - y_i \right] \tag{3}$$

where α_i is the Lagrange multiplier, and sample ($\alpha_i \neq 0$) is the support vector.

The Karush-Kuhn-Tucker (KKT) conditions are employed for the optimal solution of the object function in a nonlinear optimization problem [51,52]. The resulting regression model of LSSVM ultimately is obtained and can be evaluated by

$$y(x) = \sum_{i=1}^{l} \alpha_i K(x, x_i) + b \tag{4}$$

where $K(x, x_i) = \phi(x)^T \phi(x_i)$ is the kernel function, which is used for classification by mapping the input data from the featured space into a high-dimensional space. In the present work, the Radial Basis Function (RBF) kernel function is adopted in the implementation of the LSSVM and given by

$$K(x, x_i) = e^{-\frac{\|x - x_i\|^2}{2\sigma^2}} \tag{5}$$

where σ is the RBF kernel function parameter that needs to be determined, which is optimized via an external optimization technique during the training process. The pseudo-code of LSSVM algorithm is presented in Algorithm 1 as follows.

Algorithm 1: Pseudo-code of LSSVM algorithm

```
Initialize LSSVM parameters
Normalize data using Equation (8)
while stopping condition is not met do
    Train LSSVM model with parameters (σ and γ) using each training data point
end while
Predict testing data using trained LSSVM model
return accuracy
```

In the current study, the learning performance of the LSSVM is mainly determined by two tuning parameters, i.e., the RBF kernel function parameter (σ) and the regularization parameter (γ). The RBF kernel function parameter (σ) influences the smoothness of the approximated nonlinear function, and the regularization parameter (γ) controls the penalty imposed on data point that deviates from the regression function. Particle swarm optimization (PSO) is a popular approach that is employed to search for the optimal solution of parameters (σ, γ). The minimum computational error can be obtained by selecting the optimal solution (σ, γ) during the training process.

4.2. Particle Swarm Optimization

In this section, the basic algorithm of particle swarm optimization (PSO) [33] is introduced in detail. It is well-known that the PSO is a heuristic randomized algorithm and developed based on animal social behavior, e.g., food searching of bird flocks and fish schools. As the LSSVM starts working for regression fitting, the solution of tuning parameters (σ, γ) has a direct impact on the feature space and solution accuracy of the prediction model. Compared with the traditional manual selection method, the PSO algorithm with the advantages of short time consumption and high accuracy in searching for the optimal solution is adopted in this paper. The PSO algorithm is widely proposed for simulating population intelligence, in which swarm the individual is treated as a particle without volume. The PSO begins with a group of particles migrating at various initial speeds and directions. The assignment of initial speed and direction proceeds randomly in a certain range. Thereafter, each particle iteratively updates its speed and direction according to individual and group behavior. Finally, all the particles in the swarm migrate to the optimal location, which is considered as the optimal solution to the corresponding optimization problem. The PSO algorithm possesses the strong power of global optimization property and search capability.

As mentioned above, the principle of PSO is that particles iteratively search for the optimal solution from a random solution, and find the global optimum by following the currently searched optimal solution. Assuming that the initial location and velocity of the ith particle in N-dimensional space are $U_i = (u_{i1}, u_{i2}, \ldots, u_{ij})$ and $V_i = (v_{i1}, v_{i2}, \ldots, v_{ij})$, $j = 1,2, \ldots, N$; $i = 1,2, \ldots, M$, where M is the swarm size. In PSO, the optimal values of location and velocity are found by tracking the extreme values of $p_{\text{best}i}$ and g_{best} in each iteratively updated process of the ith particle, where the $p_{\text{best}i}$ is the individual optimal experience of the ith particle denoted by $P_i = [p_{i1}, p_{i2}, \ldots, p_{ij}]$, and g_{best} is the global optimal experience of the swarm denoted by $P_g = [p_{g1}, p_{g2}, \ldots, p_{gj}]$. According to the fitness value of particles, the location and velocity of each particle can be updated using the equations below:

$$v_{id}^{k+1} = \chi v_{id}^{k} + c_1 r_1 (p_{id} - u_{id}^{k}) + c_2 r_2 (p_{gd} - u_{id}^{k}) \quad (6)$$

$$u_{id}^{k+1} = u_{id}^{k} + v_{id}^{k+1} \quad (7)$$

where v_{id}^{k+1} and v_{id}^{k} are the velocities of the ith particle at kth and $(k + 1)$th iterations, respectively; u_{id}^{k+1} and u_{id}^{k} are the locations of the ith particle at kth and $(k + 1)$th iterations, respectively; χ is the inertia weight coefficient; c_1 and c_2 are the learning coefficients, which are two prescribed constants in PSO; r_1 and r_2 are two random constants in the range of [0, 1].

In the present work, two tuning parameters (σ, γ) are required to be optimized, thus the value of N is selected as $N = 2$. Considering that the assignment of M is relevant to the specific problems, whose value is 10~50 in general. For the present problem of optimization and prediction, the size of the solution M is set to be $M = 30$. The empirical values of the learning coefficients $c_1 = c_2 = 2$ is adopted [38,42]. Since a higher value of the inertia weight coefficient (χ) is conducive to jumping out of local convergence, while a lower value of χ has a stronger local search capability, the value of χ is taken as 0.8 [38]. And the maximum number of swarm evolution (i.e., iteration number) $k_{\max} = 200$ is pre-set in this work. The pseudo-code of PSO algorithm is presented in Algorithm 2 as follows.

Algorithm 2: Pseudo-code of LSSVM algorithm

```
Initialize population of particles and derive local and global best particles (p_besti and g_best)
for k = 1 to maximum number of iterations do
    for i = 1 to population size do
        Update the velocity of the ith particle (v_i) using Equation (6)
        Update the location of the ith particle (u_i) using Equation (7)
        if F(u_i) < F(p_besti) then
            p_besti = u_i.
            if F(p_besti) < F(g_best) then
                g_best = p_besti
            end if
        end if
    end for
end for
return g_best
```

When the RMSE of the LSSVM is minimal, the corresponding σ and γ can be regarded as the optimal parameters. The detailed optimization steps are stated as follows:

(i) Initialize the parameters in PSO algorithm.
(ii) Calculate the fitness value of each particle, and evaluate their fitness with Equation (8), i.e., $F(u) = F(\sigma, \gamma) = \text{RMSE}$.
(iii) Update the location and velocity of the ith particle with Equations (6) and (7) and compare the current fitness value of each particle $F(u_i)$ with the individual best fitness value $F(p_{\text{best}i})$, if satisfying $F(u_i) < F(p_{\text{best}i})$, $p_{\text{best}i} = u_{i.}$
(iv) Compare the current fitness values of all particles in the swarm $F(u_i)$ with the fitness value of the best location of the swarm $F(g_{\text{best}})$, if satisfying $F(u_i) < F(g_{\text{best}})$, the global optimal solution $g_{\text{best}} = u_{i.}$
(v) Check whether the termination condition is met. If the error accuracy is satisfied or the maximum number of swarm evolution is reached, then the process of searching for the optimal solution (σ, γ) ends and the optimal values of σ and γ are outputted. Otherwise, proceed to the next step from step (ii) to continue the process of parameter optimization.
(vi) Substitute the optimal parameters (σ, γ) into Equation (1) for predicting the yield strength and displacement of RC columns.

4.3. Implementation of the Yield Strength and Displacement Prediction

This section introduces the implementation of the proposed PSO-LSSVM model for the yield strength and displacement prediction in detail, as shown in Figure 4. At the beginning of the PSO-LSSVM, to prevent the formation of larger range numerical attributes dominating smaller range attributes, data initialization is required. The normalized data are in the range of [0, 1] through Equation (8), which is defined by

$$x_i^n = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{8}$$

where x_i is any data point, $x_{\min}$ and $x_{\max}$ are the minimum and maximum values of the entire dataset; x_i^n is the normalized value of the data point.

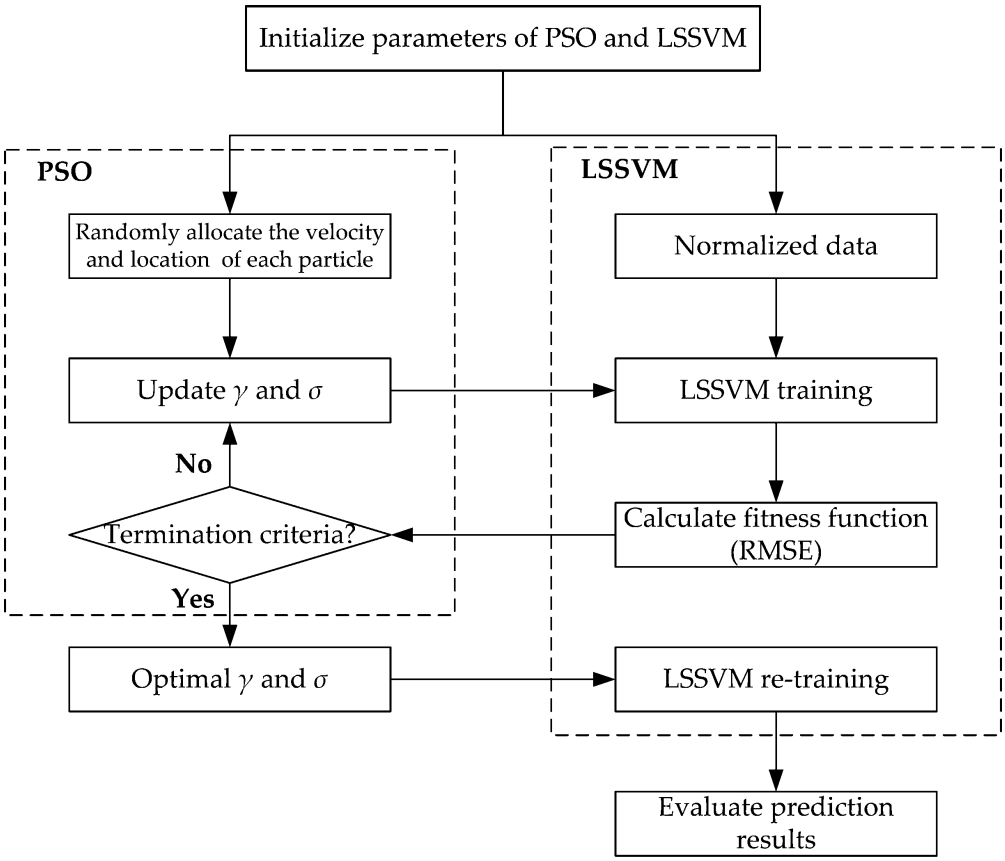


Figure 4. Flowchart of PSO-LSSVM scheme.

4.4. Evaluation of the PSO-LSSVM Prediction Performance

4.4.1. Performance Evaluation Indicators

The prediction performance of PSO-LSSVM is evaluated by the coefficient of determination (R^2), mean absolute error (MAE) and root mean square error (RMSE). The coefficient R^2 represents the level of explained variability between the actual values measured in the experiment and predicted values computed by the PSO-LSSVM model. The closer the coefficient R^2 varies toward 1, the more similar the actual values and predicted values are. Likewise, the lower values for MAE and RMSE, the higher accuracy in the predicted values indicate. Equations (9)–(12) are the mathematical formulation of the adopted performance evaluation criteria RMSE, MAE and R^2, respectively, as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{9}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{10}$$

$$R^2 = \frac{(n\sum y_i \times \hat{y}_i - \sum y_i \sum \hat{y}_i)^2}{\left(n\sum y_i^2 - (\sum y_i)^2\right)\left(n\sum \hat{y}_i^2 - (\sum \hat{y}_i)^2\right)} \tag{11}$$

$$\text{Explained Variance} = \left(1 - \frac{\text{Var}(y_i - \hat{y}_i)}{\text{Var}(y_i)}\right) \times 100\% \tag{12}$$

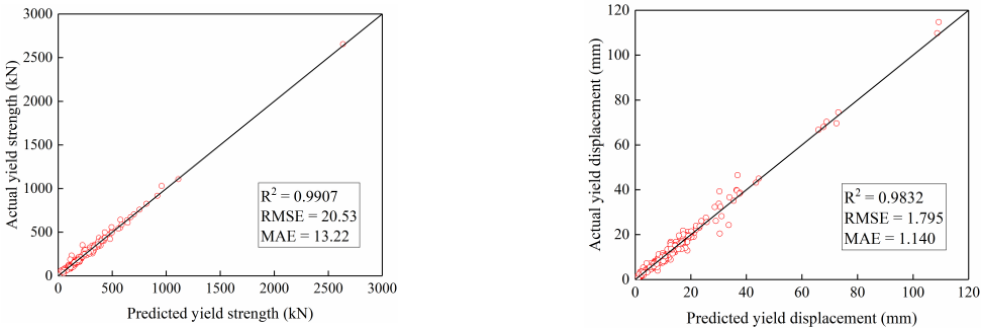
where y_i is the experimental value of the ith sample, $\hat{y}_i$ is the predicted value of the ith sample, and n is the total number of samples, and Var() is the variance.

4.4.2. Evaluation of the PSO-LSSVM Prediction Results

Before evaluating the proposed model, the dataset composed of 382 test results is randomly divided into two sets: the training set (305 data samples) and the testing set (77 data samples). With the optimal parameters in Table 4 adopted, the training performance and testing results of the PSO-LSSVM are displayed in Figures 5 and 6, respectively. The scatter-plots in the two figures show that the proposed model has yielded prediction results that present quite excellent agreement with the experimental test results in both training and testing processes. Moreover, it also can be seen that the PSO-LSSVM earned relatively high values of R^2 for testing, the corresponding values of the coefficient of determination (R^2) are 0.99 and 0.96 for the training and testing cases of yield strength, respectively (Figures 5a and 6a), and the corresponding values of R^2 are 0.98 and 0.97 for the training and testing cases of yield displacement, respectively (Figures 5b and 6b). The high values of R^2 imply that a strong correlation exists between the predicted and measured yield point, which indicates that the PSO-LSSVM is proficient at capturing the underlying function of yield strength and yield displacement of RC columns.

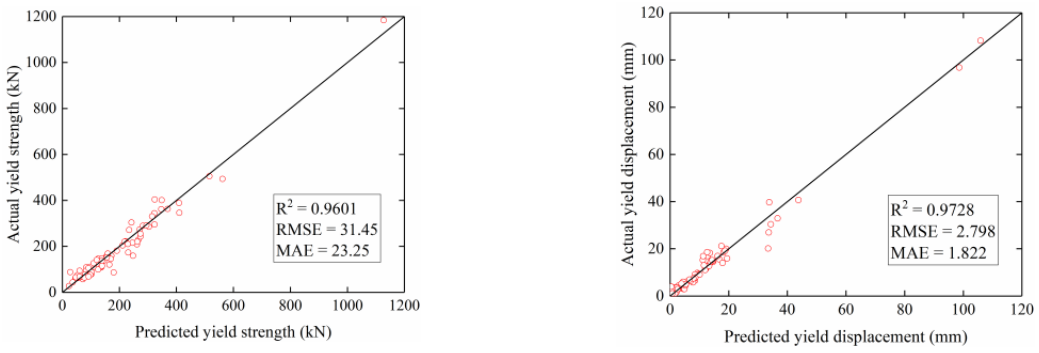
Table 4. Obtained optimal parameters of LSSVM by PSO.

	σ	γ
F_y	0.47	195.45
Δy	0.43	49.93



(**a**) (**b**)

Figure 5. Training performance and result of PSO-LSSVM (number of samples = 305): (**a**) Yield strength; (**b**) Yield displacement.



(**a**) (**b**)

Figure 6. Testing result of PSO-LSSVM (number of samples = 77): (**a**) Yield strength; (**b**) Yield displacement.

In addition, the mean absolute error (MAE) and root mean square error (RMSE) for evaluating the predictive accuracy of the PSO-LSSVM model, are also presented in Figures 5 and 6. The PSO-LSSVM model attains good perdition performances in both yield strength and yield displacement reflected in the low values of MAE and RMSE. Nevertheless, the values of MAE and RMSE in the testing process of the PSO-LSSVM are low enough already, which demonstrates that the new model has a good capability of predicting yield strength and displacement accurately. To sum up, the results indicate that the estimated yield strength and displacement are trustworthy and reliable due to the high R^2 and low RMSE and MAE. Thus, the proposed PSO-LSSVM model is deemed suited for predicting yield strength and displacement of RC columns as a new and effective method. Meanwhile, the results shown in Figures 5 and 6 also indicate that the PSO algorithm adopted for searching optimal solutions to the LSSVM model is quite effective and meets the calculation accuracy requirement in both training and testing results.

4.4.3. Comparative Evaluation of Model Prediction Performance

For further elaborating the effectiveness of the PSO algorithm in terms of parameter (σ, γ) optimization, other metaheuristic algorithms such as GA [30], DE [31] and ABC [32] are also applied to optimize the parameters of the LSSVM model. The identical training and testing samples are employed to make a performance comparison among standard LSSVM models without the optimized algorithm, and LSSVM optimized by GA, ABC, DE and PSO algorithms, respectively. The configuration of the standard LSSVM model is the same as that of the proposed model with the parameters of standard LSSVM and kernel set as σ^2 = 0.1, γ = 10 [39]. The parameter settings of the GA, ABC and DE algorithms are identical to that used in the PSO, i.e., the size of the solution M = 30 and the maximum iteration number $k_{\max}$ = 200. Table 5 is the evaluation results of the LSSVM, GA-LSSVM, ABC-LSSVM, DE-LSSVM and PSO-LSSVM models. As shown in Table 5, it is obvious that the prediction performance of the standard LSSVM is improved by adopting the optimized algorithms due to the increase in R^2 and EV, and the decrease in RMSE and MAE. The higher R^2 and EV, and lower RMSE and MAE imply the higher accuracy of models. Meanwhile, comparing the evaluation indices (R^2, EV, RMSE and MAE) of testing data between the LSSVM optimized by other metaheuristic algorithms and the proposed PSO-LSSVM model, it can be found that the R^2 of the PSO-LSSVM model in predicting yield strength (F_y) and yield displacement (Δ_y) can be up to 0.96 and 0.97, respectively. The RMSE and MAE of the PSO-LSSVM model are both sufficiently minimal. For the prediction of F_y, there is little difference in the prediction accuracy between GA-LSSVM, ABC-LSSVM, DE-LSSVM and PSO-LSSVM models. While for the prediction of Δ_y, the prediction accuracy of the PSO-LSSVM model is higher than other models. This demonstrates that the PSO-LSSVM model performs better than GA-LSSVM, ABC-LSSVM and DE-LSSVM models in terms of R^2, EV, RMSE and MAE, indicating the proposed PSO algorithm has an outstanding advantage over the GA, ABC and DE algorithms in establishing the LSSVM model used for effectively predicting the yield properties of RC columns.

Table 5. Evaluation results of standard LSSVM, GA-LSSVM, ABC-LSSVM, DE-LSSVM, PSO-LSSVM.

	Model	Training				Testing			
		RMSE	MAE	R^2	EV	RMSE	MAE	R^2	EV
F_y	LSSVM	25.41	15.65	0.9859	98.5933	43.29	26.58	0.9248	92.6090
	GA-LSSVM	20.61	13.41	0.9907	99.0746	31.92	23.93	0.9591	95.9857
	ABC-LSSVM	20.54	13.25	0.9908	99.0802	31.60	23.38	0.9599	96.0527
	DE-LSSVM	19.90	12.94	0.9913	99.1371	31.42	23.17	0.9603	96.1449
	PSO-LSSVM	20.53	13.22	0.9907	99.0821	31.45	23.25	0.9601	96.0898
Δ_y	LSSVM	2.067	1.282	0.9775	99.7536	7.870	2.895	0.7848	78.5773
	GA-LSSVM	1.800	1.108	0.9830	98.2953	4.647	2.87	0.9249	92.5010
	ABC-LSSVM	1.876	1.168	0.9815	98.1484	4.943	2.674	0.9151	91.5142
	DE-LSSVM	1.831	1.139	0.9823	98.2360	4.975	2.686	0.9140	91.4036
	PSO-LSSVM	1.795	1.140	0.9832	98.3245	2.798	1.822	0.9728	97.2999

5. Conclusions

A new ML-based method incorporating algorithms of the LSSVM and PSO is proposed for predicting the yield strength and displacement of RC columns. The present PSO-LSSVM model is capable of learning the nonlinear regression function that controls the mapping between the influencing factors and the yield strength and displacement with the PSO utilized for assisting to adaptively and quickly determine two optimal tuning parameters (σ, γ). A test database of 382 pseudo-static cyclic test of RC columns is built, then the effects of the column featured parameters on the yield-related properties are investigated concerning the related research results. The proposed PSO-LSSVM model is then compared with the standard LSSVM and the LSSVM optimized by other metaheuristic algorithms, such as GA, ABC and DE algorithms, with the identical experimental database used for model training and testing. The comparative results show that the PSO empowers the LSSVM model to efficiently and accurately predict the target data with a self-optimized machine learning algorithm based on the collected data samples. Based on the analysis results about the prediction performance of the PSO-LSSVM for the yield strength and displacement of RC columns, the following conclusions are drawn:

(1) The determination coefficient R^2 is 0.96, if 80% of the whole dataset is used for training the PSO-LSSVM model, which means a low prediction error. Meanwhile, the RMSE and MAE are 31.45 and 23.25, respectively, which indicates that the prediction model has a low prediction deviation.
(2) The PSO adopted for parameter optimization of σ and γ that are embedded in the LSSVM model, can quickly find the optimal parameters to effectively assist the LSSVM model in prediction work.
(3) The proposed PSO-LSSVM model can predict the yield strength and displacement of RC columns with high efficiency and accuracy by comparing them with the LSSVM models optimized by other metaheuristic algorithms.

Hence, the developed PSO-LSSVM model can be a useful tool to provide effective guidance for the structural design of RC columns in practical engineering applications. Essentially, the proposed model based on the ML technique is a data-driven approach that focuses on correlations between research objects, different from the traditional research method that is more concerned with cause-effect relationships. Moreover, ML methods characterized by data-driven approaches are very powerful at solving regression and optimization problems, as long as sufficient data can be provided. In our future work, the model will be extended to study the deformation properties of other key component units of building structures, such as RC beams and shear walls, and, the data collection and processing of the relevant experimental database will be conducted.

Author Contributions: Conceptualization, B.W., W.G. and Z.L.; methodology, B.W. and W.G.; validation, B.W., W.G. and Y.W.; formal analysis, Y.W. and Z.L.; investigation, B.W. and H.L.; writing—review and editing, B.W and W.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China (51808112, 51878160, 52078128), Natural Science Foundation of Jiangsu Province (BK20180155) and Postgraduate Research & Practice Innovation Program of Jiangsu Province (SJCX21_1794).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Lehman, D.; Moehle, J.; Mahin, S. Experimental evaluation of the seismic performance of reinforced concrete bridge columns. *J. Struct. Eng.* **2004**, *130*, 869–879. [CrossRef]
2. Du, X.L.; Chen, M.Q.; Han, Q. Experimental evaluation of seismic performance of reinforced concrete hollow bridge columns. *J. Vib. Shock* **2011**, *30*, 254–259.
3. Basöz, N.I.; Kiremidjian, A.S.; King, S.A.; Law, K.H. Statistical analysis of bridge damage data from the 1994 Northridge, CA, earthquake. *Earthq. Spectra* **1999**, *15*, 25–54. [CrossRef]
4. Park, H.G.; Yu, E.J.; Choi, K.K. Shear-strength degradation model for RC columns subjected to cyclic loading. *Eng. Struct.* **2012**, *34*, 187–197. [CrossRef]
5. Shrestha, R.; Smith, S.T.; Samali, B. Finite element modelling of FRP-strengthened RC beam-column connections with ANSYS. *Comput. Concr.* **2013**, *11*, 1–20. [CrossRef]
6. Wang, Q.; Song, X.; Hao, Z.; Wang, Z. Numerical simulation of failure process of RC columns based on fiber beam element in ABAQUS. *China Civ. Eng. J.* **2014**, *47*, 16–26.
7. Panagiotakos, T.B.; Fardis, M.N. Deformations of reinforced concrete members at yielding and ultimate. *ACI Struct. J.* **2001**, *98*, 135–148.
8. Frederic, L.; Patrick, P. Behavior of high-strength concrete columns under cyclic flexure and constant axial load. *ACI Struct. J.* **2000**, *97*, 591–601.
9. Lu, X.Z.; Ye, L.P.; Pan, P. Pseudo-static collapse experiments and numerical prediction competition of RC frame structure II Key elements experiment. *Build. Struct.* **2012**, *42*, 23–26.
10. Li, X.; Wang, J.; Luo, M. Pseudo-static experiment and analysis on seismic behavior of the RC columns strengthened by GHPFRCC. *Nephron Clin. Pract.* **2015**, *22*, 56–60. [CrossRef]
11. Lee, S.C. Prediction of concrete strength using artificial neural networks. *Eng. Struct.* **2003**, *25*, 849–857. [CrossRef]
12. Gencel, O. The application of artificial neural networks technique to estimate mass attenuation coefficient of shielding barrier. *Int. J. Phys. Sci.* **2009**, *4*, 743–751.
13. Chou, J.S.; Pham, A.D. Enhanced artificial intelligence for ensemble approach to predicting high performance concrete compressive strength. *Constr. Build. Mater.* **2013**, *49*, 554–563. [CrossRef]
14. Pham, A.D.; Hoang, N.D.; Nguyen, Q.T. Predicting compressive strength of high-performance concrete using metaheuristic-optimized least squares support vector regression. *J. Comput. Civ. Eng.* **2016**, *30*, 06015002. [CrossRef]
15. Sharma, L.K.; Singh, T.N. Regression-based models for the prediction of unconfined compressive strength of artificially structured soil. *Eng. Comput.* **2017**, *34*, 175–186. [CrossRef]
16. Yue, Z.; Ding, Y.; Zhao, H.; Wang, Z. Case study of deep learning model of temperature-induced deflection of a cable-stayed bridge driven by data knowledge. *Symmetry* **2021**, *13*, 2293. [CrossRef]
17. Jeon, J.S.; Shafieezadeh, A.; Desroches, R. Statistical models for shear strength of RC beam-column joints using machine-learning techniques. *Earthq. Eng. Struct. Dyn.* **2014**, *43*, 2075–2095. [CrossRef]
18. Xu, J.; Ren, Q.; Shen, Z. Prediction of the strength of concrete radiation shielding based on LS-SVM. *Ann. Nucl. Energy* **2015**, *85*, 296–300.
19. Vu, D.T.; Hoang, N.D. Punching shear capacity estimation of FRP-reinforced concrete slabs using a hybrid machine learning approach. *Struct. Infrastruct. Eng.* **2016**, *12*, 1153–1161. [CrossRef]
20. Luo, H.; Paal, S.G. Machine Learning-Based Backbone Curve Model of Reinforced Concrete Columns Subjected to Cyclic Loading Reversals. *J. Comput. Civ. Eng.* **2018**, *32*, 04018042. [CrossRef]
21. Mangalathu, S.; Jeon, J.S. Classification of failure mode and prediction of shear strength for reinforced concrete beam-column joints using machine learning techniques. *Eng. Struct.* **2018**, *160*, 85–94. [CrossRef]
22. Ning, C.; Wang, L.; Du, W. A practical approach to predict the hysteresis loop of reinforced concrete columns failing in different modes. *Constr. Build. Mater.* **2019**, *218*, 644–656. [CrossRef]
23. Jiang, C.S.; Liang, G.Q. Modeling shear strength of medium- to ultra-high-strength concrete beams with stirrups using SVR and genetic algorithm. *Soft Comput.* **2021**, *25*, 10661–10675. [CrossRef]
24. Quaranta, G.; De Domenico, D.; Monti, G. Machine-learning-aided improvement of mechanics-based code-conforming shear capacity equation for RC elements with stirrups. *Eng. Struct.* **2022**, *267*, 114665. [CrossRef]
25. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
26. Suykens, J.A.; Vandewalle, J. Least squares support vector machine classifiers. *Neural Process. Lett.* **1999**, *9*, 293–300. [CrossRef]
27. Suykens, J.A.; Vandewalle, J.; Moor, B.D. Optimal control by least squares support vector machines. *Neural Netw.* **2001**, *14*, 23–35. [CrossRef]
28. Suykens, J.A.; De Brabanter, J.; Lukas, L.; Vandewalle, J. Weighted least squares support vector machines: Robustness and sparse approximation. *Neurocomputing* **2002**, *48*, 85–105. [CrossRef]
29. Zheng, C.; Kasihmuddin, M.S.M.; Mansor, M.A.; Chen, J.; Guo, Y. Intelligent Multi-Strategy Hybrid Fuzzy K-Nearest Neighbor Using Improved Hybrid Sine Cosine Algorithm. *Mathematics* **2022**, *10*, 3368. [CrossRef]
30. Holland, J.H. Genetic algorithms and the optimal allocation of trials. *SIAM J. Comput.* **1973**, *2*, 88–105. [CrossRef]
31. Storn, R.; Price, K. Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces. *J. Glob. Optim.* **1997**, *11*, 341–359. [CrossRef]

32. Karaboga, D.; Basturk, B. A powerful and efficient algorithm for numerical function optimization: Artificial bee colony (ABC) algorithm. *J. Glob. Optim.* **2007**, *39*, 459–471. [CrossRef]
33. Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceedings of the ICNN'95-International Conference on Neural Networks, Perth, Australia, 27 November–1 December 1995; Volume 4, pp. 1942–1948.
34. Ma, W.T. Evaluation of rock slope stability based on PSO and LSSVM. *Rock Soil Mech.* **2009**, *30*, 845–848.
35. Chatterjee, S.; Sarkar, S.; Hore, S.; Dey, N.; Ashour, A.S.; Shi, F.; Le, D.N. Structural failure classification for reinforced concrete buildings using trained neural network based multi-objective genetic algorithm. *Struct. Eng. Mech.* **2017**, *63*, 429–438.
36. Xue, X. Prediction of slope stability based on hybrid PSO and LSSVM. *J. Comput. Civ. Eng.* **2017**, *31*, 04016041. [CrossRef]
37. Yang, J.; Zhu, H.; Wang, Y. An orthogonal multi-swarm cooperative PSO algorithm with a particle trajectory knowledge base. *Symmetry* **2017**, *9*, 15. [CrossRef]
38. Cao, J.; Li, W.Y.; Zhao, D.S.; Song, Z.G.; Ding, W.Y. Forecast of building inclination around foundation pit based on PSO-LSSVM model. *Comput. Eng. Appl.* **2016**, *52*, 254–259.
39. Yu, Y.; Zhang, C.; Gu, X.; Cui, Y. Expansion prediction of alkali aggregate reactivity-affected concrete structures using a hybrid soft computing method. *Neural Comput. Appl.* **2019**, *31*, 8641–8660. [CrossRef]
40. Mo, P.Q.; Ma, D.Y.; Zhu, Q.Y.; Hu, Y.C. Interpretation of heating and cooling data from thermal cone penetration test using a 1D numerical model and a PSO algorithm. *Comput. Geotech.* **2020**, *130*, 103908. [CrossRef]
41. He, S.; Chen, W.; Mu, X.; Cui, W. Constrained optimization model of the volume of initial rainwater storage tank based on ANN and PSO. *Environ. Sci. Pollut. R.* **2020**, *27*, 21057–21070. [CrossRef]
42. Khari, M.; Danial, J.A.; Dehghanbanadaki, A. Prediction of lateral deflection of small-scale piles using hybrid PSO–ANN model. *Arab. J. Sci. Eng.* **2020**, *45*, 3499–3509. [CrossRef]
43. Haselton, C.B.; Liel, A.B.; Taylor-Lange, S.C.; Deierlein, G.G. Calibration of model to simulate response of reinforced concrete beam-columns to collapse. *ACI Struct. J.* **2016**, *113*, 1141–1152. [CrossRef]
44. Liu, Z.; Guo, A. Empirical-based support vector machine method for seismic assessment and simulation of reinforced concrete columns using historical cyclic tests. *Eng. Struct.* **2021**, *237*, 112141. [CrossRef]
45. Zhang, J.; Li, X.; Cao, W.; Yu, C. Cyclic behavior of steel tube-reinforced high-strength concrete composite columns with high-strength steel bars. *Eng. Struct.* **2019**, *189*, 565–579. [CrossRef]
46. Hassan, W.M.; Farag, M. Seismic performance of steel-reinforced concrete composite columns in existing and modern construction. *Soil Dyn. Earthq. Eng.* **2021**, *151*, 106945. [CrossRef]
47. Liu, Z.; Wang, X.; Zhou, Z.; Xue, J.; Lai, B.; Xu, J. Experimental and numerical study on seismic performance of steel reinforced concrete spatial frame with irregular section columns. *Eng. Struct.* **2021**, *242*, 112507. [CrossRef]
48. Park, R. State of the art report-ductility evaluation from laboratory and analytical testing. In Proceedings of the 9th World Conference on Earthquake Engineering, Tokyo-Kyoto, Japan, 2–9 August 1988; Volume 8, pp. 605–616.
49. Berry, M.; Parrish, M.; Eberhard, M. *PEER Structural Performance Database User's Manual Version 1.0*; University of California: Berkeley, CA, USA, 2004.
50. Ghannoum, W.; Sivaramakrishnan, B.; Pujol, S.; Catlin, A.C.; Fernando, S.; Yoosuf, N.; Wang, Y. NEES: ACI 369 Rectangular Column Database. 2015. Available online: https://datacenterhub.org/resources/255 (accessed on 25 October 2022).
51. Platt, J.C. A fast algorithm for training support vector machines. *J. Inf. Technol.* **1998**, *2*, 1–28.
52. Shevade, S.K.; Keerthi, S.S.; Bhattacharyya, C.; Murthy, K. Improvements to the SMO algorithm for SVM regression. *IEEE Trans. Neural Networ.* **2000**, *11*, 1188–1193. [CrossRef]