Article

Seismic Response Prediction of Porcelain Transformer Bushing Using Hybrid Metaheuristic and Machine Learning Techniques: A Comparative Study

Quan Zhou [1,2,3], Yongheng Mao [1], Fengqi Guo [1,*] and Yuxuan Liu [1]

[1]   School of Civil Engineering, Central South University, Changsha 410075, China; zhou-quan@cscec.com (Q.Z.);
      234802092@csu.edu.cn (Y.M.); 234812331@csu.edu.cn (Y.L.)
[2]   China Construction Fifth Engineering Bureau Co., Ltd., Changsha 410004, China
[3]   State Grid Hunan Electric Power Co., Ltd., Changsha 410004, China
[*]   Correspondence: fengqiguo@csu.edu.cn

Abstract: Although seismic response predictions are widely used for engineering structures, their applications in electrical equipment are rare. Overstressing at the bottom of the porcelain insulators during seismic events has made power transformer bushings in substations prone to failure. Thus, this paper proposed and compared six integrated machine learning (ML) models for seismic stress response predictions for porcelain transformer bushings using easily monitored acceleration responses. Metaheuristic algorithms such as particle swarm optimization were employed for architecture tuning. Prediction accuracies for stress response values and classifications were evaluated. Finally, shaking table tests and simulation analyses for a 1100 kV bushing were implemented to validate the accuracy of the six ML models. The results indicated that the proposed ML models can quickly forecast the maximum stress experienced by a porcelain bushing during earthquakes. Swarm intelligence evolutionary technologies could quickly and automatically aid in the retrofitting of architecture for the ML models. The K-nearest neighbor regression model had the best level of prediction accuracy among the six selected ML models for experimental and simulation validations. ML prediction models have clear benefits over frequently used seismic analytical techniques in terms of speed and accuracy for post-earthquake emergency relief in substations.

Keywords: porcelain transformer bushing; seismic response prediction; metaheuristic optimization; machine learning; simulation analysis; shaking table tests

MSC: 68W99

1. Introduction

Substations play a crucial role in power transmission since they are responsible for converting and regulating electricity [1]. Nevertheless, during past seismic events [2–4], including the 2022 Luding earthquake in China [5], electrical equipment in substations endured significant harm and displayed clear susceptibility. Damage to substation equipment can interrupt normal power delivery and result in high financial costs, further hindering post-earthquake relief work.

Electrical equipment used in substations typically includes a vertical or inclined cantilever insulator equipment body together with a supporting structure such as a steel frame or turret, as shown in Figure 1. Substations consist of a variety of equipment, such as post insulators, transformer bushings, disconnect switches, circuit breakers, and surge arresters. The equipment damaged in the 2008 Wenchuan earthquake exhibited numerous common failure mechanisms, including failure of the porcelain insulators, the cracking of connection flanges [5], and oil leakage from insulators [6]. Damage to equipment often occurs due to high stress in the bottom positions, surpassing the materials' maximum

strength. Thus, determining the peak stresses of the porcelain insulators during earthquakes is a crucial factor for maintaining structural integrity.
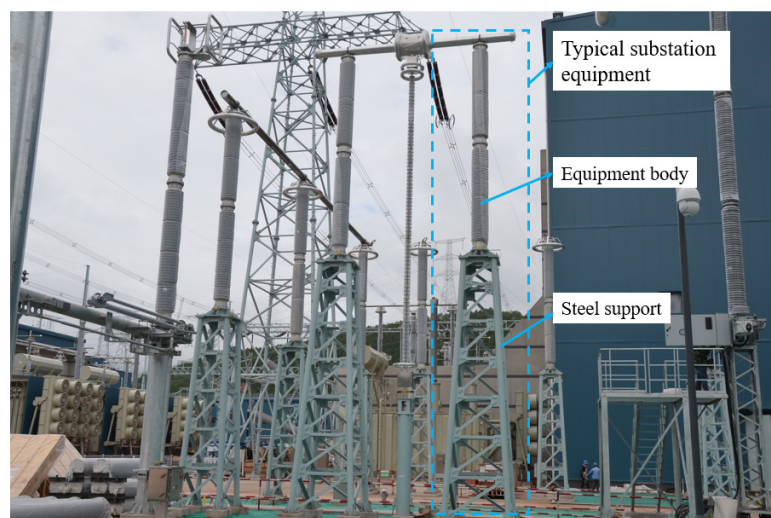

Figure 1. Typical electrical equipment found in substations.

In recent years, studies with a particular emphasis on substation equipment, including power transformers [7–11] and other equipment [12–14], have been conducted. These studies have involved seismic response analysis [15,16], seismic mitigation [17–19], and vulnerability and risk assessments [20–22]. Recently, research has been conducted on the evaluation of substation equipment after earthquakes [23,24]. Among them, power transformers are of high concern since they are the core equipment in substations. In 1998, Bellorini et al. conducted static calculations and vibrational experiments on high-voltage bushings to assess the stipulated amplification factor between the ground and the transformer bushing flange [25]. In 2019, He et al. conducted shaking table experiments to assess the seismic performance of five UHV transformer bushings mounted on support frames [26,27]. Their study indicated that the metal flange may be susceptible to collapsing during earthquakes. Based on this analysis, a revised flange design was suggested to enhance its ability to withstand seismic activity [28]. Additional research on the relationship between flanges and their stiffness has also been conducted [29]. Xie et al. conducted a numerical investigation of the seismic reactions of post-electrical equipment using linear-elastic analysis [30]. The validity of the theoretical model was confirmed using shaking table experiments. These studies contribute to improving the structural safety of electrical systems in substations.

The most recent research is relevant to the design or retrofitting of structures undergoing seismic activity, rather than for assessing damage after earthquakes. There is a significant paucity of research on post-earthquake efforts. The analytical procedures often used in seismic studies on electrical equipment, such as shaking table tests and simulation calculations, are too time-consuming to provide prompt post-earthquake identification due to the huge number of equipment involved and the damaged bushings or insulators are clearly a result of inadequate strength. In some instances, the structures may still look whole, but micro-cracks may have developed owing to excessive stress or cumulative damage. Porcelain insulators containing micro-cracks pose a risk to structural safety and the transmission of electricity. Therefore, it is necessary to identify and replace them. However, the task of identifying every individual piece of equipment in substations after earthquakes is difficult for maintenance staff due to the substantial quantity of equipment there. The peak root stress is crucial for making informed decisions regarding post-earthquake services, but the installation of strain gauges is challenging due to the presence of a silicone layer used for protection. In addition, they are also susceptible to interference from elec-

tromagnetic fields, resulting in inaccurate monitoring data. Given these constraints, it is imperative to use non-contact signal-gathering technologies and evaluation methods.

To assess the physical indexes that are difficult to measure, artificial intelligence (AI) technologies, especially machine learning (ML) and deep learning (DL), have been introduced in structural engineering [31,32]. Using AI algorithms, prediction models can be developed to estimate the failure probability or damage index of building structures. In 2015, Goulet et al. proposed an assessment framework that could learn the relationship between the frequency shift and damage state of buildings using data-driven methods, and this assessment method could estimate the safety state by monitoring data [33]. Ni et al. proposed a data-driven approach for post-earthquake reliability assessments of civil structures and the updated probability density functions can be used to assess reliability [34]. However, the predicted objectives of these studies are always probability or defined performance indexes, or response time history, and the hyper-parameter tuning is a time-consuming grid search or is artificial. Thus, several academics accurately developed non-linear structural seismic responses using long short-term memory (LSTM) networks [35] and attention-based LSTM networks [36]. The previous paragraph has stated that the primary focus is on the peak seismic stress response of substation equipment. Therefore, the network architecture of DL, which requires significant time for superfluous time history adjustments, can be disregarded. A rapid method for architecture tuning of ML models for predicting the peak response of equipment after earthquakes is desperately needed.

Based on the above-mentioned research background and literature review, this paper conducts a seismic assessment of electrical equipment in substations. In the current seismic study of substation equipment, the existing literature focuses mainly on pre-earthquake seismic performance analysis, mitigation analysis, and fragility assessment. Post-earthquake rapid prediction of seismic stress responses is beneficial to the rapid detection of substations with hundreds of equipment. However, investigations of seismic response predictions for substation equipment are lacking in the current literature, indicating a gap in current research. To solve this problem, ML techniques, which can estimate the response results in a short time and provide potential solutions, are introduced in this work. The approach of using artificial search or grid search to find suitable ML model hyperparameters is time-consuming. Instead, metaheuristic optimization algorithms are employed due to their automatic search performance. As the most commonly used metaheuristic optimization algorithm, particle swarm optimization (PSO) is adopted as a representation to illustrate the feasibility of using metaheuristic algorithms to optimize ML models of electrical equipment.

Therefore, the goal of this paper is to enrich the seismic research on substation equipment by using ML techniques to predict peak stress responses of equipment. It provides an approach for finding the peak response of one piece of equipment after earthquakes; this can be further used for all equipment, thus reducing the post-earthquake decision time. The specific methodology includes: (1) selecting a large number of natural ground motions to excite the numerical model of equipment to obtain a batch seismic response containing multiple peak accelerations and peak stress responses at vulnerable positions; (2) each ground motion produces a sample, thus there are lots of samples that are then split into a training set and a test set; (3) The training set is used to train ML models, taking peak accelerations as inputs and the peak stress as the output; (4) During the training process, the model structures can be adjusted using PSO to optimize the structural hyperparameters; (5) after adjusting the model structures, the final ML models can be employed as prediction models after earthquakes. This paper takes a typical 1100 kV porcelain transformer bushing as an example to conduct the case study. The remainder of this paper is organized as follows: In Section 2, six ML algorithms and PSO are introduced. Section 3 illustrates a typical 1100 kV porcelain transformer. The overall application framework and specific methodology for predicting the peak responses of the bushing are given in Section 4 and the results are presented in Section 5. Finally, the prediction models are validated using shaking table tests and simulation analyses in Section 6.

2. Methodology

2.1. Machine Learning Models

ML is a commonly used technology in the data-driven area, known for its robust learning and predictive capabilities. By using data learning, the prediction model can be constructed as a mapping between various physical variables. This research aims to estimate the peak stress at the bottom of the transformer bushing. Unlike a classified index, this quantity is a continuous value; therefore, regression techniques are used instead of classification algorithms. In the context of a certain prediction model, the optimum hyperparameter combination is considered as the solution that leads to the best performance. This implies that it maximizes the performance function or minimizes the error function. While ML models can enhance their prediction performance by artificial manipulation of hyperparameters, the excessive time required for this procedure is often unsatisfactory. Fortunately, biological metaheuristic algorithms provide possible replacements for artificial methods, allowing for the automated discovery of the best combinations of hyperparameters [37]. This work used the particle swarm optimization (PSO) method as the search strategy for optimal hyperparameters. PSO is a stochastic search algorithm that mimics the cooperative foraging behavior of birds. It should be pointed out that other heuristic optimization algorithms can also be used. The main role is to find the optimal hyperparameters for ML models, hence any optimization algorithm can be accepted as long as it can help the ML model achieve the required predictive ability. In this paper, PSO is used as a case study to illustrate the overall process. Its basic principle is introduced in Section 2.2.

2.1.1. Multi-Layer Perceptron (MLP)

The multi-layer perceptron (MLP) is a common neural network structure. Artificial neural networks, convolutional neural networks, and recurrent neural networks are very effective at making accurate predictions. MLP, as a basic network, also has a robust capacity to determine both linear and non-linear relationships. The MLP model is capable of converting a mapping from an m-dimensional input $X^m = [x_1, x_2, \ldots, x_m]$ to an n-dimensional output $Y^n = [y_1, y_2, \ldots, y_n]$. Figure 2 illustrates the overall structure of an MLP with two hidden layers and a one-dimensional output. Every neuron in the hidden layer is responsible for converting the input from the preceding layer into the output of that neuron. It is achieved using a weighted linear summation and a non-linear activation function. Once the numbers of hidden layers and neurons in each layer are determined, the learned objectives are the weighted coefficients and bias values. Following this, the output can be provided by the MLP model if the m-dimension vector is given.
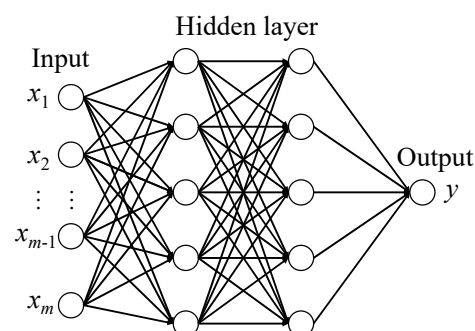

Figure 2. MLP with two hidden layers [38].

2.1.2. Support Vector Regression (SVR)

The goal of the support vector machine (SVM) technique is to identify a hyperplane in an n-dimensional space that can accurately classify input points. Support vector regression (SVR) employs the same underlying ideas as SVM, but instead of doing classification, SVR focuses on predicting numerical values. Although SVM is more well-known, SVR is known for its efficacy in estimating real-value functions. SVR is a supervised learning technique

that builds models using symmetric loss functions that may punish both overestimation and underestimation. The SVR model, using the ordinary least squares approach, establishes a threshold ε around the regression line (or hyperplane) to exclude any data points inside ε from being penalized for mistakes. The range of deviation is referred to as the ε-pipeline. The corresponding mathematical issue involves minimizing the norm of the weight vector, as represented by Equation (1), while satisfying the constraint condition described in Equation (2).

$$\min_{\text{w},b} \frac{1}{2}\|\text{w}\|_2^2 \tag{1}$$

$$\left|y_i - \left(\text{w}^\text{T}\text{x}_i + b\right)\right| \leq \varepsilon \quad i = 1, 2, ..., N \tag{2}$$

where w represents the weight and b represents the bias value. (x_i, y_i) is the i-th sample.

2.1.3. Kernel Ridge Regression (KRR)

Ridge regression is a modified version of the least squares approach that imposes a penalty on the size of the coefficients. Kernel ridge regression (KRR) is a combination of ridge regression and the kernel trick, as described by Murphy [39]. The space resulting from combining the kernel and the data is determined, and the KRR algorithm then estimates a linear relationship. A non-linear relationship may be obtained from the original data using non-linear kernels. Unlike the least squares approach, KRR sacrifices its unbiasedness in order to achieve great numerical stability, resulting in improved computational accuracy. The KRR model can offer mapping between the covariates x and the output variables y, both of which are continuous. The primary objective is to minimize the overall loss function, which can be expressed as:

$$\sum_i \left(w^T \phi(x_i) - y_i\right)^2 + \frac{\lambda}{2}\|w\|^2 \tag{3}$$

where w is the weight and λ is the ridge parameter. $\phi(\cdot)$ is a feature map and (x_i, y_i) is the i-th sample. In Equation (3), the leftmost items are the cumulative error and regularization. The goal of KRR is to determine the estimation model $f(x) = w^T\phi(x)$. In KRR, the kernel $k(x, y) = \phi(x)^\text{T}\phi(y)$ such that the feature map is not included in the estimation model f.

2.1.4. Stochastic Gradient Descent Regression (SGDR)

Gradient descent is a widely used approach for optimizing model parameters in machine learning architecture. The gradient descent approach may be used to iteratively get the minimal loss and related model hyperparameters while searching for the minimum value of the loss function. Machine learning encompasses two gradient descent algorithms that are derived from fundamental theory: stochastic gradient descent and batch gradient descent. The pseudocodes for stochastic gradient descent are shown in Algorithm 1. Only the former is examined in this research due to its efficiency and ease of implementation. Each training sample promptly changes settings to achieve a higher speed.

The loss function quantifies the discrepancy between the estimated and real values in the linear estimation Equation (4), as shown in Equation (5).

$$f(x_i) = \theta^T x_i \tag{4}$$

$$L(\theta) = \frac{1}{2}(y_i - f(x_i))^2 \tag{5}$$

The loss function is directly associated with the model parameter, where the model parameter is considered the independent variable and the loss is the dependent variable. In order to reduce the amount of loss, the model parameter θ is adjusted in the direction of the gradient (as shown in Equation (6)). Thus, the hyperparameter η determines the pace

at which the best solution, namely the learning rate, is obtained. Both excessively big and tiny values are unsuitable.

$$\theta := \theta - \eta * \nabla L = \theta - \eta \times \frac{\partial}{\partial \theta} L(\theta) \tag{6}$$

where x_i represents an input vector, y_i represents an output corresponding to x_i, and L represents the loss function.

Algorithm 1 Stochastic Gradient Descent.

1: **Require:** Learning rate η_k
2: **Require:** Initial parameter θ
3: **while** Stopping criterion not met **do**
4:     Sample a minibatch of m examples from the training set
5:     $\{x^1,\ldots,x^m\}$
6:     Set g = 0.
7:     **for** i = 1 to m **do**   **if m = 1 $\rightarrow$ single example SGD**
8:         Compute gradient estimate:
        $\hat{g} \leftarrow \hat{g} + \frac{1}{m}\nabla_\theta L(f(x^i;\theta),y^i)$
9:     **end for**
10:   Apply update: $\theta \leftarrow \theta - \eta_k\hat{g}$
11: **end while**

2.1.5. K-Nearest Neighbor Regression (KNR)

If the majority of the K-nearest samples in the feature space belong to a certain category, then the sample should be classified in that category. This technique, called the K-nearest neighbor algorithm, is used in the field of machine learning. K-nearest neighbor regression (KNR) is an extension that is used to address challenges associated with the prediction of continuous variables, as opposed to classification problems. KNR is a non-parametric learning approach, which implies that it does not need training a function as an estimating model. It only depends on the initial data. In Figure 3, the blue point is currently unknown and we anticipate acquiring its output. Given that K = 4, there are four data points that are in close proximity to this blue point. The proximity between any two points is contingent upon the definition of distance. Typically, Euclidean distance or Manhattan distance are used. After performing the distance computation, the outputs of these four data points are retrieved, and the average value is used as the estimated output of the blue point.
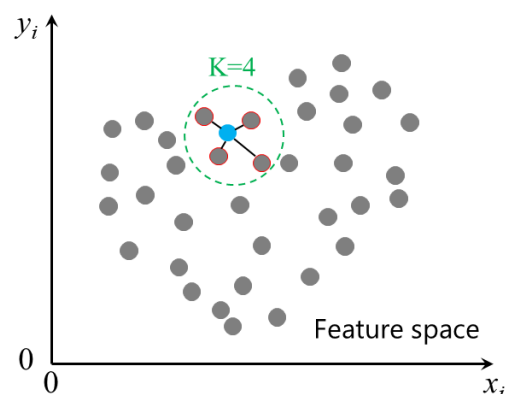

Figure 3. K-nearest neighbor regression with K = 4.

2.1.6. Decision Tree Regression (DTR)

A decision tree (DT) is a graphical representation in the form of a tree, as shown in Figure 4. A DT does not need to assume a relationship between input and output before training, unlike other models. This is because a DT relies on the original data

structure. This approach is applicable for both linear and non-linear relationship estimation. Algorithm 2 shows the pseudocode for the decision tree algorithm. Figure 4 illustrates a straightforward example of DTR in a 2D feature space. The tree originates from the root node, which encompasses all the data points. The first optimum division (x_1, a) is determined by calculating the error, resulting in the creation of two data subsets. The same computations are performed for each subset in order to partition the spaces into further subspaces. Following two divisions, a total of four-leaf nodes are identified. The decision tree regression operates on the premise of averaging the outputs of all data points inside a subspace to get the estimated output. If a newly acquired data point is verified as belonging to this specific subspace, its output is determined solely based on the output of this subspace.

Algorithm 2 Pseudocode for the decision tree algorithm.

1: **Input:** Training dataset D;
2: **Output:** Regression tree $f(x)$
3:　　In the input space where the training dataset is located, recursively divide each
4: region into two subregions and determine the output values on each subregion and
5: construct a binary decision tree:
6: (1) Select the optimal segmentation variable and segmentation point s, and solve it
7: $\min\limits_{j,s}\left[\min\limits_{c_1}\sum\limits_{x_i\in R_1(j,s)}(y_i-c_1)^2+\min\limits_{c_2}\sum\limits_{x_i\in R_2(j,s)}(y_i-c_2)^2\right]$
8:　　Traverse variable j, scan the segmentation point s for a fixed segmentation variable
9: j, and select the pair (j, s) that minimizes equation
10: (2) Divide the region using the selected pairs (j, s) and determine the corresponding
11: output values:
12: $R_1(j,s)=\left\{x\big|x^{(j)}\leq s\right\},R_2(j,s)=\left\{x\big|x^{(j)}\leq s\right\}$
13: $\hat{c}_m=\frac{1}{N_m}\sum\limits_{x_i\in R_m(j,s)}y_i,x\in R_m,m=1,2$
14: (3) Continue to call the above two steps on the two subregions until the stop
15: condition 1:is met.
16: (4) Divide the input space into **M** regions R_1, R_2, ..., R_M, and generate a decision tree:
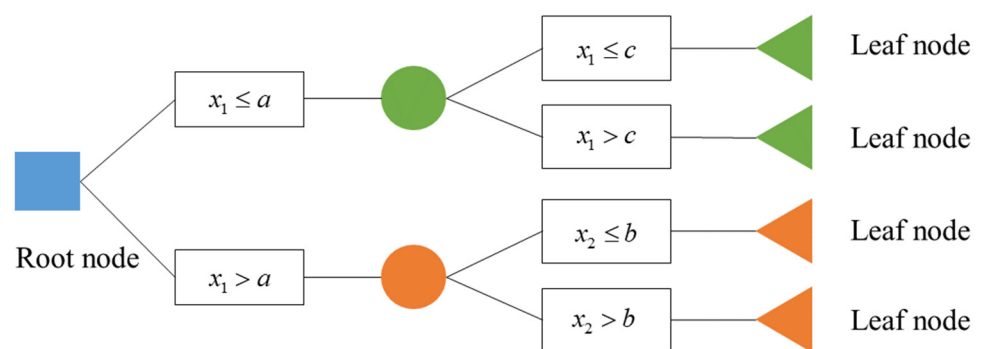17: $f(x)=\sum\limits_{m=1}^{M}\hat{c}_m I(x\in R_m)$

Figure 4. Typical decision tree analysis model.

2.2. Particle Swarm Optimization Technology

Particle swarm optimization (PSO) is a swarm intelligence algorithm [33] that mimics the predatory behavior of bird groups in nature. The flock's objective is to locate sustenance within a designated area. During the search, the birds communicate and exchange information, including their unique positions. In order to determine whether they have discovered the most effective answer via collaboration, the birds assess the situation and share their own ideal solution with the whole flock. Ultimately, the flock may converge on the food source, thus establishing the ideal solution.

Algorithm 3 displays the whole sequence of steps in the method. A group of particles with no mass is used to replicate the collective behavior of bird flocks, which symbolizes the random solutions. The function output represents the fitness value associated with a particle. Every particle has two distinct characteristics: velocity (*V*) and position (*X*). During each iteration, the particles undergo displacement from their present position *X* based on their velocity *V* and two extrema. There are two types of extrema: the individual extremum, known as *pbest*, and the global extremum, known as *gbest*. The term "*gbest*" refers to the highest fitness value among all fitness values associated with all particles, as shown in Equations (7) and (8).

$$V_j := \omega \times V_j + c_1 \times rand() \times (pbest_j - X_j) + c_2 * rand() \times (gbest_j - X_j) \tag{7}$$

$$X_j := X_j + V_j \tag{8}$$

where V_j and X_j represent the speed and location of particle j, respectively; ω represents the inertia weight; c_1 and c_2 represent the learning factors, generally $c_1 = c_2 = 2$. After enough iterations, the particles find the optimal solution for the best fitness value.

Algorithm 3 Basic algorithm flow for PSO.

1: (1) Determine the fitness function f according to the actual problem;
2: (2) Give the total number of particles and the maximum generation, and initialize
3: them randomly;
4: (3) Calculate global optimum *gbest*;
5: (4) Update the speed *V* and position *X* of all particles per Equations (7) and (8);
6: (5) Calculate the fitness value of each particle;
7: (6) Update local optimum *gbest* for each particle;
8: (7) Update global optimum *gbest*;
9: (8) Is the maximum generation met? Yes, end; No, repeat Steps 4~8.
10: Output the optimal fitness and solutions.

---

3. Structure of the 1100 kV Transformer Bushing

This section introduces the detailed physical structure of a 1100 kV oil-paper insulated capacitive transformer bushing. The ultra-high-voltage transformer bushing, working at 1100 kV, consists of two porcelain insulators (air-side and oil-side), a connection flange used for connecting them, four corona rings, and a dome, as shown in Figure 5. The total length of the bushing is 13,315 mm, with 2310 mm for the oil-side insulator and 11,005 mm for the air-side insulator. The oil-side porcelain insulator is consistently immersed in dielectric oil. In the same way, the air-side porcelain insulator is on the outside of the tank. The bottom cross-section of the bushing has an outer diameter of 640 mm and an internal diameter of 550 mm. The total mass of the bushing is 7135 kg. The connection flange serves to link the two porcelain insulators and securely fasten the UHV bushing to the transformer turret using bolts. Within the porcelain insulators, a central conductor with an aluminum core is installed to facilitate the transmission of electric current. Additionally, to compress the bushing before any external loads are applied to it, a seven-ton pretension is placed on the conductor. The connecting flange is constructed from cast aluminum, which is a commonly used material in high-voltage electrical apparatus, as shown in Figure 5b. It consists of the top plate, the bottom plate, the installation plate, the cylinder, and the stiffeners. The installation plate is installed to connect this flange and the transformer tank, making them whole.

The finite element (FE) model was generated using ABAQUS2023 software, as shown in Figure 5c. The corona ring, air-side porcelain insulator, flange, and oil-side porcelain insulator were all represented using C3D8R solid elements in the simulation. The S4R shell elements were used to imitate the stiffeners equipped on the flange. When considering the air-side insulator, the impact of umbrellas on the shape was disregarded and only the weight of the porcelain insulator was taken into account. The two horizontal directions and

the vertical direction of the bushing frame system were set to the X, Z, and Y directions, respectively. To accommodate the earthquake inputs in both the FE model and the experimental model, a steel frame measuring 2.7 m in height was used to provide support for the bushing. The connection between the frame and the mounting flange was established at the installation plate. In order to verify the accuracy of the FE model, the researchers conducted shaking table tests and compared the dynamic features, namely the vibration frequencies (2.48 Hz for the FE model and 2.38 Hz for the test) and mode shapes, as stated in reference [26]. Hence, it can be concluded that the FE simulation model is valid and may serve as a substitute for the physical bushing when it comes to seismic responses [26]. This simulation model is used to compute the dynamic responses, which are deemed to be consistent with the real structure.
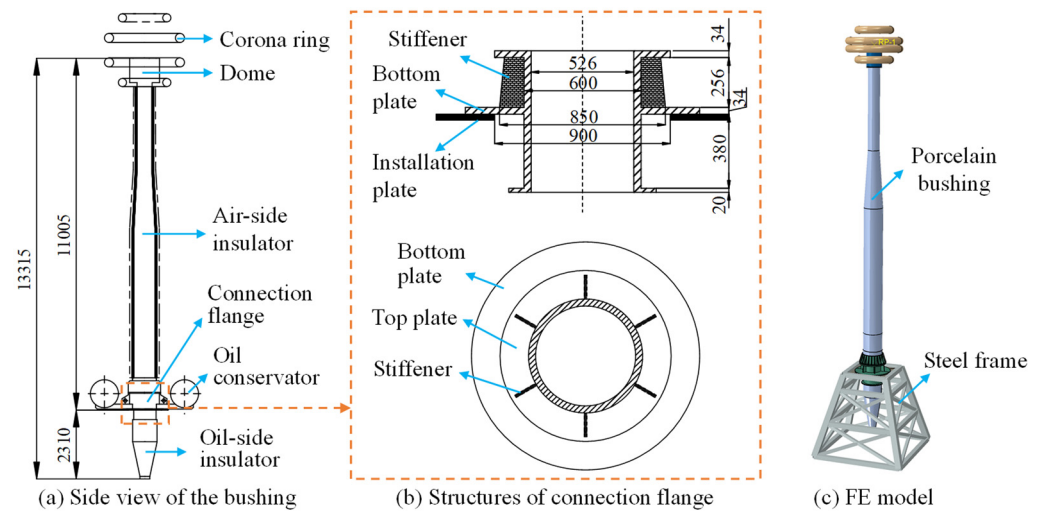

Figure 5. Structures and FE model of the bushing (unit: mm).

4. Framework and Methodology

4.1. Application Framework

To clearly understand the application of the ML models aided by PSO, a framework of rapid prediction of seismic responses of substation equipment is illustrated in Figure 6. It runs through the entire earthquake process, involving pre-earthquake sample collection and model training, during-earthquake monitoring, and post-earthquake prediction using the trained ML models. The optimal ML prediction model is the target of this study.
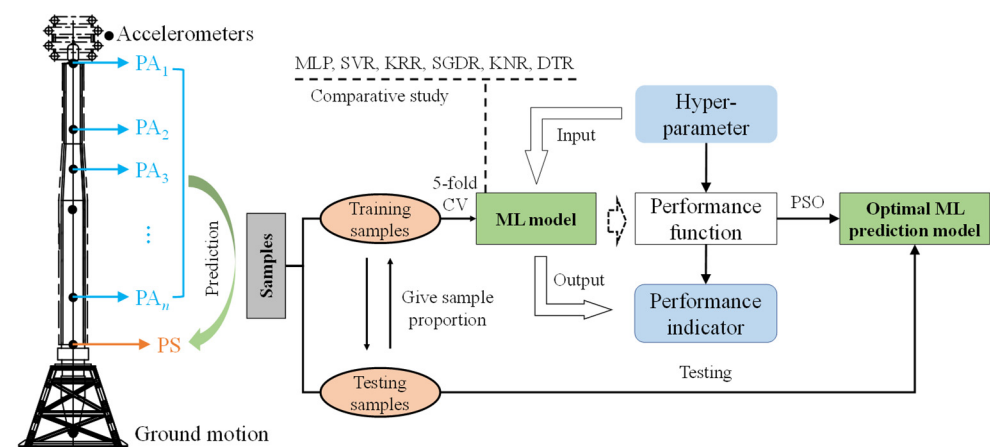

Figure 6. Overall application framework for predicting peak stress responses of substation equipment.

The main steps of this application framework can be summarized as follows:

Prior to the occurrence of an earthquake, certain locations (key points) are chosen based on their ease of monitoring acceleration, taking into account the circumstances for installation. High-speed cameras and infrared cameras may be used as methods for acquiring acceleration response in monitoring technologies.

A sophisticated FE simulation model is created to analyze the seismic responses, considering the precise dimensions of the substation equipment.

A variety of ground motions are selected based on the soil type as inputs for the FE model, resulting in a collection of response samples consisting of peak accelerations (PA_1, PA_2, ..., PA_n) and peak stress at the bottom of the air-side bushing (PS).

The samples are partitioned into one training set and one test set. The prediction performance of an ML model with predefined architectural parameters may be assessed using k-fold cross-validation (CV) [40–42].

The performance of the ML model is solely determined by its internal structures, while the training sets are fixed. There exists a mapping connection between the architectural hyperparameters and performance indicators, which is referred to as the performance function. The grid search and artificial search, which often use hyperparameter search methods, are known to be time-consuming. One may also use biological heuristic algorithms like PSO to evolve the performance function. The best architecture is found and may be utilized to construct a task-optimized ML model.

All training and testing samples are used to train and test the model, respectively. Additionally, the ML model is sufficiently advanced so that it can replace physical post-earthquake equipment for seismic response analysis.

Sensors or other monitoring technologies can be used to measure the accelerations at crucial places during an earthquake. Next, the maximum accelerations are evenly obtained over a brief period.

Once an earthquake occurs, the built ML model is promptly used to evaluate the maximum stress at the bottom of the air-side insulator according to the peak accelerations of key positions as inputs in the model. The predicted peak stress can assist with post-earthquake rapid decision-making of response-level estimations for the detection of multiple electrical equipment in a substation.

4.2. Data and Sample Collection

To obtain the seismic response samples to train the ML model, a large number of natural ground motions should be selected. Since the ground motions are used for training the ML model, which should be suitable for as many cases as possible, the selection only considers the soil type and as many ground motions as possible. In this paper, the bushing services in a specific location are designed to meet the eight-degree fortification requirement with soil type II. However, it is recommended that nine degrees of fortification be adopted due to the significance of the bushing [27]. The fundamental acceleration of the design is 0.4 g. Therefore, a total of 340 ground motions were chosen from the PEER ground motion database, focusing on the Vs30 range that corresponds to soil type II [43]. The characteristic information for the selected earthquake is shown in Table 1.

Table 1. Characteristics of the selected earthquake.

Characteristic	Magnitude	Epicenter Distance (km)	Fault Type Rrup (km)	PGA (g)
Range	4.37~7.36	1.06~247.04	3.21~222.41	0.005~0.84
Mean value	5.91	42.74	38.09	0.128
Standard deviation	0.67	42.26	38.58	0.124

After inputting the ground motions into the bushing frame system, a series of response samples can be obtained, each including the peak accelerations and the peak stress responses. According to the application framework, they are divided into training sets and testing sets. As shown in Figure 7, the training samples are divided into k parts (k = 5 in

this paper). For each training session, only four sections are used as training data, while the remaining sections are used as validation data. As a consequence, every training session yields a validation outcome. Each of the five parts is used as validation data in rotation, while the remaining four parts are utilized for training. Model performance is determined by averaging five validation results and using them for comparison. The testing samples, obtained by subtracting the training samples from the total samples, are considered prediction samples. The estimated findings derived from these samples serve as the foundation for error analysis.
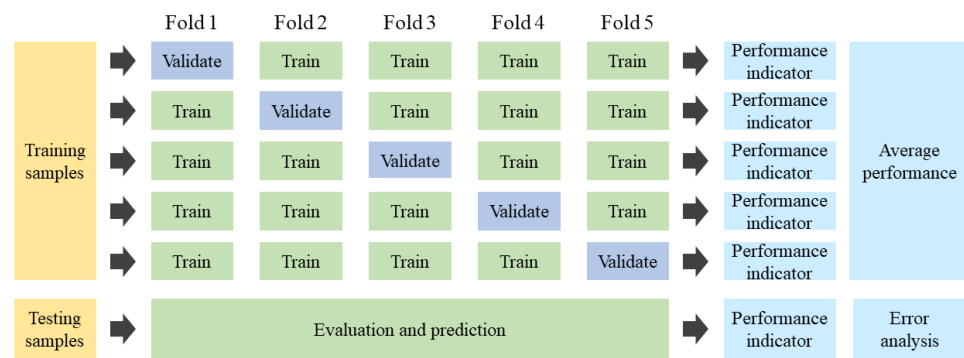

Figure 7. Sample partition and application.

4.3. Evaluation Indicators

It is essential to compare the predicted results and actual results of stresses obtained from the samples during the validation and testing processes. Thus, evaluation indicators must be supplied. Several indicators are often employed, including the mean absolute percentage error (MAPE) and the Willmott index of assessment. Four indicators are used in this study: MAPE, Pearson correlation coefficient (*R*), mean square error (MSE), and mean absolute error (MAE). MAPE is an error indicator that decreases as the estimated and real stresses get closer to each other. On the other hand, the Pearson correlation coefficient, *R*, is a consistency indicator that approaches 1 when the estimation performance is better. The formulas are represented by Equations (9)–(12).

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{\hat{y}_i - y_i}{y_i} \right| \tag{9}$$

$$R = \sum_{i=1}^{n} (y_i - \overline{y})(\hat{y}_i - \overline{\hat{y}}) \Big/ \left(\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2} \sqrt{\sum_{i=1}^{n}(\hat{y}_i - \overline{\hat{y}})^2} \right) \tag{10}$$

$$\text{MSE} = \frac{1}{m} \sum_{i=1}^{m} \left(y_i - \hat{y}_i \right)^2 \tag{11}$$

$$\text{MAE} = \frac{1}{m} \sum_{i=1}^{m} \left| y_i - \hat{y}_i \right| \tag{12}$$

where y_i, $\hat{y}_i$ represent the actual and predicted values of testing samples, respectively; $\overline{y}_i$, $\overline{\hat{y}}_i$ represent the means of the actual and predicted values of testing samples, respectively; n represents the number of test samples.

The practitioner also pays attention to the response level of substation equipment. The relevant standard [44] stipulates a factor of safety of 2.0 for the porcelain insulators used in the electrical equipment. Therefore, the workers are more concerned with the critical stress response (*CSR*), which is defined as the ratio of material strength (*MS*) to the factor of safety (*FOS*), as given in Equation (13).

$$CSR = MS/FOS \tag{13}$$

Therefore, the stress responses are divided into two classes: one contains responses <*CSR* while the other contains responses ≥*CSR*. The response prediction results can further output the classifications, thus there will be a comparison. For a classification model, the confusion matrix is the frequently used error evaluation variable. A typical confusion matrix is given in Table 2. TP and TN represent the numbers of correctly predicted stress <*CSR* and ≥*CSR*, respectively. FN and FP represent the numbers of incorrectly predicted stress <*CSR* and ≥*CSR*, respectively. The accuracies, TPR and TNR, are critical evaluation indicators; their definitions are provided in Table 2.

Table 2. Typical confusion matrix and evaluation indicators for classification accuracy.

Actual Stress	Predicted Stress		Accuracy = (TP + TN)/(TP + TN + FP + FN)	
	<*CSR*	≥*CSR*		
<*CSR*	TP	FN	TPR = TP/(TP + FN)	FNR = FN/(TP + FN)
≥*CSR*	FP	TN	FPR = FP/(FP + TN)	TNR = TN/(FP + TN)

4.4. Hyperparameter Tuning

The predictive capabilities of ML models are heavily influenced by the algorithm's design, particularly the hyperparameters. In order to enhance the estimating capabilities of ML models, a thorough search for appropriate hyperparameters is performed for each individual model. PSO is used as a replacement for the time-consuming grid search and artificial search to aid with the optimization of the hyperparameters. The overall process employing PSO to optimize the model's hyperparameters is shown in Figure 8. Firstly, for each ML model with a given training set, a group of hyperparameters will produce a performance indicator, such as MAPE or R. Therefore, the fitness function *f* should be determined based on the ML model. The function returns an evaluation indicator based on a combination of hyperparameters provided as input. The average Mean Absolute Percentage Error (MAPE) is the result obtained from each assessment using a 5-fold cross-validation (CV) technique. Secondly, several initial groups of hyperparameters are randomly generated. Then, the MAPE is calculated to obtain the minimum MAPE as the *gbest*. If the generation is less than the given value, the hyperparameters are updated and then the MAPEs for all groups are calculated so that there are minimum MAPE values for all groups of hyperparameters, *gbest,* and minimum MAPE for each group of hyperparameter, *pbest,* during the existing and previous generations. The *pbest* for each group and the *gbest* are continuously updated according to the PSO algorithm until the maximum number of generations is reached. Then, the ideal combination of hyperparameters is obtained. The ML model is trained using the whole set of training samples, with the optimal hyperparameter combination that was searched for. The trained model is then used to predict stress levels using the testing examples in order to assess its performance.
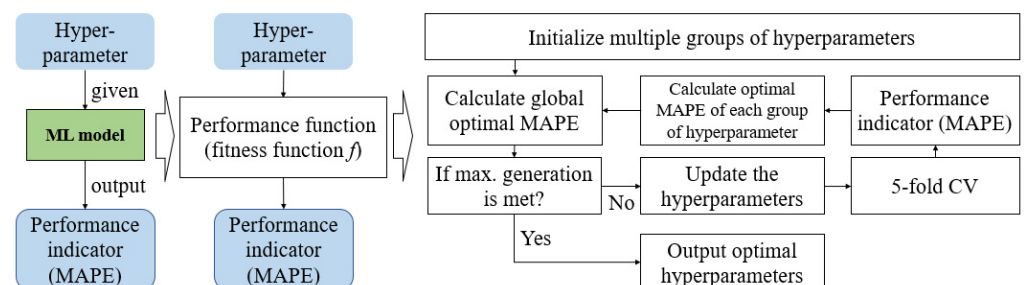

Figure 8. Approach employing PSO to tune the model structures.

5. Results and Discussion

5.1. Data Collection Results

In this paper, five key points were selected to gain the peak acceleration responses as the model inputs. The specific positions are the top of the air-side insulator, the two junctions with variable cross-sections, the bottom of the insulator, and the top of the steel frame. A total of 340 sets of ground motions were selected. The response spectra and average spectrum are shown in Figure 9a. Subsequently, the 340 ground motions were inputted into the bottom of the frame in order to get the seismic responses. After seismic response analyses, the acceleration time histories of these points were derived, and the peaks were extracted. Additionally, the peak stresses at the bottom of the porcelain insulator are gathered. During an earthquake, five peak accelerations and one peak stress make up a sample, with the 340 samples corresponding to the 340 ground motions. Their statistics are given in Figure 9b.
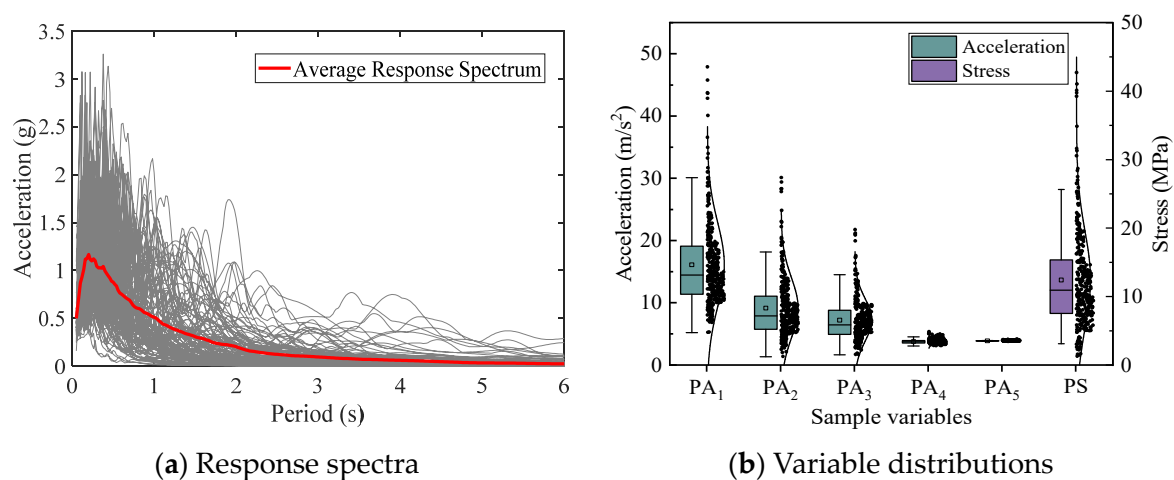

(**a**) Response spectra   (**b**) Variable distributions

Figure 9. Spectral information for the selected ground motions and distributions of the variables.

5.2. Hyperparameter Tuning Results

For the six ML models selected in this paper, the hyperparameters are searched using PSO. The hyperparameters and their search ranges are shown in Table 3. The search ranges are artificially determined because the hyperparameters should be accepted as long as the prediction of the ML models is accurate enough. Additionally, notice that only two layers are determined in MLP to simplify the tuning process because the more the layers, the more complex the model. Inertia weight ω plays a pivotal role in maintaining the balance between the exploration and exploitation processes. During the searches, the inertia weight of PSO is set to 0.4, which is determined based on the commonly used range, 0.4~0.9.

The optimal hyperparameter results are provided in the last column of Table 3. The change in the MAPE indicator is given in Table 4, where we can see that MAPE levels off when the number of generations reaches 300, which indicates search stability. Table 4 shows that SVR and KRR, which are the easiest ML models whose MAPEs level off, maintain stability from the 20th generation. MLP and KNR reach stability from the 60th generation and the 120th generation, while SGDR and DTR are the last two models whose MAPEs reach stability.

Table 3. Hyperparameter search ranges and results of the estimation models.

Ml Models	Hyperparameters	Search Range	Optimal Results
MLP	Number of neurons in the 1st layer	1~20, integer	19
	Number of neurons in the 2nd layer	1~20, integer	13
SVR	Regularization parameter	0.01~10	4.676
	Kernel	'linear', 'poly', 'rbf'	'linear'
KRR	Regularization strength	0.01~10	0.01
	Kernel	'linear', 'polynomial', 'rbf'	'linear'
KNR	Number of neighbors	1~50, integer	2
	Leaf size	1~100, integer	40
SGDR	Regularization term (RT)	0.1~10	0.1103
	The maximum iteration number (MI)	100~10,000, integer	6139
	Stopping criteria (SC)	10^{-5}~10^{-2}	0.00952
DTR	The max depth of the tree (MD)	1~20, integer	11
	The minimum samples required to split (MSS)	2~10, integer	7
	The minimum samples required for a leaf (MSL)	1~10, integer	2

Table 4. MAPE changes versus PSO generations.

Generation	20	60	100	120	160	180	220	240	280	300
MLP	0.081	0.078	0.078	0.078	0.078	0.078	0.078	0.078	0.078	0.078
SVR	0.077	0.077	0.077	0.077	0.077	0.077	0.077	0.077	0.077	0.077
KRR	0.079	0.079	0.079	0.079	0.079	0.079	0.079	0.079	0.079	0.079
KNR	0.147	0.116	0.116	0.112	0.112	0.112	0.112	0.112	0.112	0.112
SGDR	0.192	0.164	0.109	0.105	0.101	0.11	0.103	0.095	0.095	0.095
DTR	0.123	0.12	0.119	0.12	0.119	0.12	0.12	0.119	0.119	0.119

5.3. Regression Prediction Error Analysis

Comparisons were conducted to evaluate the prediction performance of the established ML models. The testing samples are used to make comparisons between the predicted results of the ML models and the simulated results of the simulation model of the 1100 kV bushing. The peak acceleration data are inputted into the six trained ML models and the outputted peak stresses are compared with the stress data in the testing samples, as shown in Figure 10. The six ML models have enough ability to predict the peak stress of the bushing. Most predicted points are close to the actual points and the similar variations in trends also illustrate the prediction accuracy. SGDR also shows relatively larger prediction deviations compared with the other five ML models.

Evaluation indicators for the six ML models are shown in Table 5. Analysis of the evaluation indicators shows that the *R* values of the ML models are all approximately 0.99, and in some cases considerably higher, approaching 1. Among these models, MLP, SVR, and KRR have the lowest MAPE values and the highest *R* values (0.994). These ML models are built using linear methods, indicating that there is a significant linear correlation between the inputs and outputs. However, Other ML models provide accurate estimation results. Therefore, the six ML models are considered suitable for estimating the peak stress of the transformer bushing.
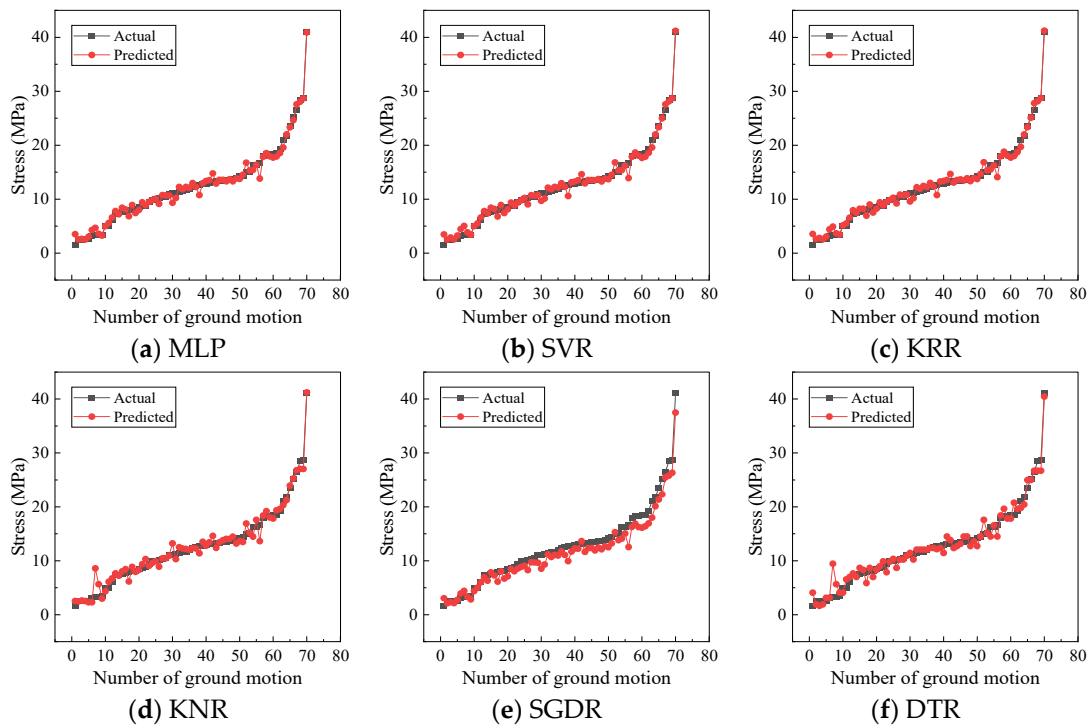
Figure 10. Testing results of the ML models and comparison with actual results.

Table 5. Evaluation indicators of the six ML models.

Indicator	MLP	SVR	KRR	KNR	SGDR	DTR
MAPE	0.0737	0.0771	0.0753	0.1009	0.1146	0.1355
R	0.994	0.994	0.994	0.988	0.993	0.983
MSE	0.6307	0.6340	0.6059	1.2468	2.3710	1.6747
MAE	0.5666	0.5667	0.5593	0.7630	1.2657	0.9227

As previously stated, the estimated error is a crucial factor that reflects the performance of the estimation model. Figure 11 provides the distribution of estimated errors for the six ML models. It is evident that the average errors for all cases are close to zero, with the exception of SGDR. Regarding data divergence, KNR, SGDR, and DTR exhibit more pronounced discreteness. MLP, SVR, and KRR exhibit a higher degree of error concentration, around 0, compared to the aforementioned methods. Based on these situations, it can be determined that the ML models, particularly MLP, SVR, and KRR, are very effective at estimating the stress of the transformer bushing.
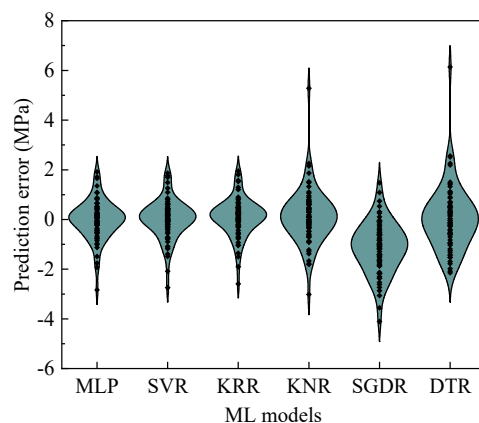

Figure 11. Prediction errors and distributions of the six ML models.

Based on detailed analyses of data points, evaluation indicators, and error distributions, the ML models are deemed to be effective stress prediction models. While a few models may be somewhat less superior, the predicted stresses as a whole align well with the results obtained from the simulation model. The suggested compound ML framework is both straightforward and very efficient, particularly for quick calculation of the maximum stress experienced by the bottom section of a transformer bushing.

5.4. Prediction Performance Evaluation for Response Classifications

In some cases, the on-site workers are more concerned about the responses if they are in a given range, which can help explain the response levels. In this paper, the material strength of the porcelain insulator is 50 MPa, according to the manufacturers. Since IEEE 693 std. stipulates a factor of safety of 2.0 [44], the *CSR* is 25 MPa. Given that we have established the ML regression prediction models, they can also be applied to judge the response classifications. Table 6 shows the confusion matrices of the six ML models. MLP, SVR, SGDR, and DTR models have a 98.6% prediction accuracy, while KRR and KNR models have a 100% prediction accuracy, revealing low errors of prediction response values on which the response classifications exhibiting high accuracy are based.

Table 6. Accuracy evaluation for judging whether the predicted responses are in corresponding classes.

Type	Actual Stress	Predicted Stress <25 MPa	≥25 MPa	Confusion Matrix		Accuracy
MLP	<25 MPa	65	0	TPR = 100%	FNR = 0	98.6%
	≥25 MPa	1	4	FPR = 20%	TNR = 80%	
SVR	<25 MPa	65	0	TPR = 100%	FNR = 0	98.6%
	≥25 MPa	1	4	FPR = 20%	TNR = 80%	
KRR	<25 MPa	65	0	TPR = 100%	FNR = 0	100%
	≥25 MPa	0	5	FPR = 0	TNR = 100%	
KNR	<25 MPa	65	0	TPR = 100%	FNR = 0	100%
	≥25 MPa	0	5	FPR = 0	TNR = 100%	
SGDR	<25 MPa	65	0	TPR = 100%	FNR = 0	98.6%
	≥25 MPa	1	4	FPR = 20%	TNR = 80%	
DTR	<25 MPa	65	0	TPR = 100%	FNR = 0	98.6%
	≥25 MPa	1	4	FPR = 20%	TNR = 80%	

6. Experimental and Simulation Validation

Even though the prediction accuracy of the established ML models has been illustrated using the testing samples, an actual application is still needed to validate the universality. Thus, shaking table tests were conducted for the transformer bushing using a given artificial ground motion. Simulation analysis using artificial ground motion was also implemented. According to the peak acceleration and stress responses, the ML models can be validated further, revealing the actual engineering significance.

6.1. Experimental Validation

The experimental model for the bushing support system during the shaking table tests is illustrated in Figure 12. It was employed to validate the proposed ML models. The detailed structural parameters of the bushing and steel frame were introduced in Section 3. In the tests, the acceleration sensors were installed on the key positions along the vertical orientation of the bushing and strain gauges were installed at the bottom cross-section of the porcelain insulator. Rigid connections were implemented at the bottom of the steel frame to allow for the input of artificial ground motion from this location. The experiment was conducted uniaxially due to the symmetrical nature of the structure [27].
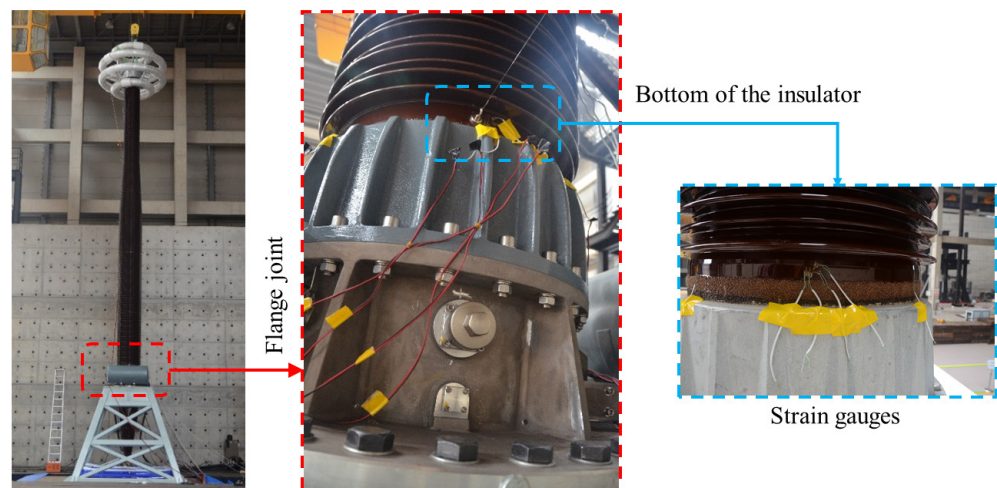
Figure 12. Test model and strain gauge arrangement of the transformer bushing [27].

The test scenarios (TSs) are listed in Table 7. TS 2 is the test scenario where artificial ground motions were inputted to excite the dynamic responses. In TS 1 and TS 3, white noises were inputted to identify the dynamic characteristics of the porcelain bushing before and after ground motion scenarios.

Table 7. Test scenarios for the shaking table tests.

Test	Test Scenario	
	Earthquake Motion	Target PGA/g
TS 1	White noise	0.07
TS 2	Artificial ground motion	0.15
TS 3	White noise	0.07

Based on the examination of dynamic characteristics and observed phenomena, it was determined that the bushing remained undamaged throughout TS 2. The experimental results from TS 2 were used to validate the model. Five acceleration sensors were vertically positioned along the bushing. The specific positions are the same as the above-mentioned key points: the top of the air-side insulator (PA_1), the two junctions with variable cross-sections (PA_2, PA_3), the bottom of the insulator (PA_4), and the top of the steel frame (PA_5). The peak accelerations measured by the sensors are 7.076, 3.807, 3.269, 2.022, and 1.614 m/s^2, respectively. The peak accelerations are inputted into the six ML models and the resulting peak stress is compared to the real peak stress of 5.50 MPa. The actual peak stress was determined by integrating Young's modulus for the porcelain material (106 GPa) with the data collected from strain sensors. Using these procedures, the evaluated results were obtained and are shown in Figure 13. The prediction errors were simultaneously calculated and are shown in this figure. Based on Figure 13, the calculated value of 5.33 MPa obtained from KNR is closest to the measured value of 5.5 MPa, resulting in a relative error of just 3.04%. The six ML models provide precise evaluations of peak stress, with a maximum relative error of just 8.72%. In summary, the ML models possess the capacity to precisely evaluate the maximum stress experienced by the bottom section of the insulator.
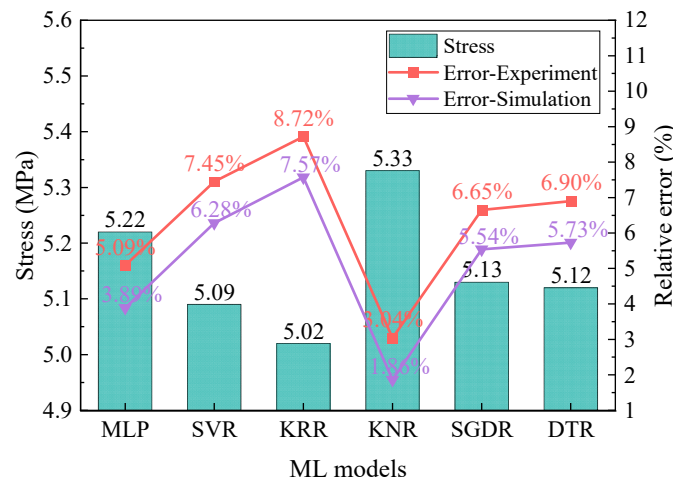
Figure 13. Comparison of the experimental and simulation results obtained using ML models.

6.2. Simulation Validation

To validate the accuracy of ML models using the simulation model, dynamic analysis was performed. The details of the simulation model for the 1100 kV bushing are illustrated in Section 3. The earthquake excitation for this simulation model is the artificial ground motion used in the shaking table test in Section 6.1, as shown in Figure 14. After the dynamic analysis, the stress responses at the bottom section are obtained, with the peak stress being 5.431 MPa. The five peak accelerations are input into the simulation model and the predicted peak stress is then obtained and compared with the simulation results, as shown in Figure 14. The minimum error is 1.86% for KNR and the maximum is 7.57% for KRR. The ML models can be deemed reliable enough based on the results, with the maximum relative error being smaller than that of the experiment.
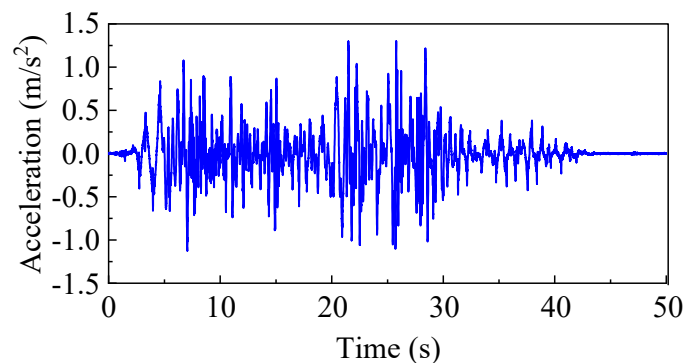
Figure 14. Artificial ground motion for exciting the simulation model.

6.3. Overall Comparison

Shaking table tests and simulation computations were conducted in the preceding two sections. When the same artificial ground motion is used, the two methodologies provide different assessment results for peak root stress. The shaking table test gave a value of 5.500 MPa, while the simulation analysis gave a value of 5.431 MPa. Furthermore, the six ML models were used to evaluate the stress levels using the peak acceleration inputs obtained from the most dependable shaking table experiments. The results of the comparison are shown in Figure 13, where the relative errors are calculated. Based on Figure 13, of the six ML models, KNR predicts the highest peak stress, while KRR predicts the lowest. Compared to the two regularly used methods, KNR consistently demonstrates the lowest error level. Regarding the anticipated maximum stress, dynamic tests provide the highest but most precise results, with the simulation models following closely. The largest discrepancy in results between the ML models and the other two techniques is 8.72%,

while the lowest discrepancy is 1.86%. Once the ML architecture is chosen and tuned, the ML model will have enough accuracy to replace previous approaches in post-earthquake rapid prediction of stress responses of porcelain transformer bushings.

The prediction model based on hybrid ML models has significantly improved computational efficiency compared to traditional methods. Six algorithms were tested on a universal computer platform with an Intel Core i5 CPU, an NVIDIA GeForce GTX 1660 Super GPU, and 16 GB of DDR4 RAM. Taking the earthquake response prediction analysis in Section 6 as an example, the traditional numerical analysis method takes 40 min, while prediction using the hybrid ML models takes 3 s, thus improving the efficiency of rapid emergency warning for post-earthquake damage in substations.

7. Conclusions and Future Work

7.1. Conclusions

This paper established hybrid ML models for predicting seismic responses of porcelain transformer bushings. The developed multiple ML models were compared and validated using shaking table tests and simulation analysis. The conclusions include:

(1) The proposed hybrid machine learning models can quickly forecast the maximum stress experienced by a porcelain bushing and its classifications after earthquakes based on the recorded acceleration data collected during seismic events. By prioritizing pre-earthquake preparations, this approach minimizes the workload after earthquakes and facilitates prompt and effective response to post-earthquake emergencies.

(2) Swarm intelligence evolutionary technologies can quickly and automatically aid in the retrofitting of architecture for ML algorithms, resulting in ML models that demonstrate higher prediction performance for test samples.

(3) Six ML models with optimal architectures were used to predict the maximum stress experienced by a 1100 kV bushing after an earthquake. The accuracies of the six ML models were confirmed using shaking table tests and simulation analysis. Compared with the experimental results, the errors of all the ML models were within 10%, with KNR having the smallest prediction error of 3.04%. The K-nearest neighbor regression model had the best prediction accuracy compared to the other six ML models.

(4) The ML method described in this study has clear benefits over frequently used seismic analytical techniques in terms of speed and accuracy when it comes to post-earthquake emergency relief of electrical equipment in substations.

7.2. Limitations and Future Work

Even though this paper illustrates the feasibility of using ML techniques to predict the seismic responses of electrical equipment in substations, some limitations remain and should be investigated in future work.

(1) This paper mainly focuses on the seismic response prediction of substation equipment using ML techniques where PSO is adopted as the tuning approach for the model's hyperparameters. However, only six ML models were employed in this paper, thus more ML models should be analyzed in the future. Additionally, PSO is a typical metaheuristic optimization algorithm that represents a way of automatically optimizing the model architecture. In the future, other metaheuristic optimization algorithms can also be investigated to predict the accuracy of ML models.

(2) The strong ground motions account for a relatively small proportion of the selected ground motion database, which may influence the training results. In the future, more strong ground motions should be added to the dataset to provide sufficient diversity, further extending the application range of the developed ML models to strong ground motions.

(3) Based on the method proposed in this article, analyzing and simulating more experimental objects under different scenarios will further improve the universality of ML models and facilitate uncertainty analysis. This should be further improved in future work.

References

1. Bu, X.; Cao, Y.; Wu, C.; Zhu, W. Stress Analysis of the Foundation of Transmission Tower-Line System under Wind Load. In Proceedings of the 2022 IEEE/IAS Industrial and Commercial Power System Asia (I&CPS Asia), Shanghai, China, 8–11 July 2022; pp. 233–239.
2. Schiff, A.J. *Northridge Earthquake: Lifeline Performance and Post-Earthquake Response*; American Society of Civil Engineers: New York, NY, USA, 1995.
3. Sezen, H.; Whittaker, A.S. Seismic performance of industrial facilities affected by the 1999 Turkey earthquake. *J. Perform. Constr. Facil.* **2006**, *20*, 28–36. [CrossRef]
4. Eidinger, J.; Davis, C.; Tang, A.; Kempner, L. *M 9.0 Tohoku Earthquake March 11 2011 Performance of Water and Power Systems*; G & E Engineering Systems Inc.: Oakland, CA, USA, 2012.
5. Zhu, W.; Xie, Q.; Liu, X.; Mao, B.; Xue, Z. Towards 500 kV power transformers damaged in the 2022 Luding earthquake: Field investigation, failure analysis and seismic retrofitting. *Nat. Hazards* **2024**, *120*, 6275–6305. [CrossRef]
6. Tang, Q.; Li, Y.; Xu, Y. Land suitability assessment for post-earthquake reconstruction: A case study of Lushan in Sichuan, China. *J. Geogr. Sci.* **2015**, *25*, 865–878. [CrossRef]
7. Ersoy, S.; Saadeghvaziri, M.A. Seismic response of transformer-bushing systems. *IEEE Trans. Power Deliv.* **2004**, *19*, 131–137.
8. Cheng, Y.; Xue, Y.; Du, W.; Lu, Z. Experimental Study on Influence of Soil–Structure Interaction on Seismic Performance of UHV Transformer. *J. Earthq. Eng.* **2023**, *27*, 3024–3059. [CrossRef]
9. Wen, J.; Li, X.; Xie, Q. Cost-effectiveness of base isolation for large transformers in areas of high seismic intensity. *Struct. Infrastruct. Eng.* **2022**, *18*, 745–759. [CrossRef]
10. Xue, Z.; Deng, C.; Zhang, Z.; Zhu, W.; Xie, Q. Seismic Response and Coupling Analysis of 800 kV Connected System of Smoothing Reactors. In Proceedings of the 2023 International Conference on Power System Technology (PowerCon), Jinan, China, 21–22 September 2023; pp. 1–5.
11. Liu, F.; Zhu, W.; Xue, Z.; Xie, Q. Seismic Performance Analysis of 110 kV HV Circuit Breaker and Internal Manipulating Devices. In Proceedings of the 2023 Panda Forum on Power and Energy (PandaFPE), Chengdu, China, 27–30 April 2023; pp. 1350–1354.
12. Yang, Z.; He, C.; Mosalam, K.M.; Xie, Q. Vertical uplift of suspension equipment due to hanger slackening: Experimental and numerical investigation. *Struct. Eng. Mech.* **2022**, *82*, 735–745.
13. Zhu, W.; Wu, M.; Xie, Q.; Xu, J. Floor response spectra and seismic design method of electrical equipment installed on floor in indoor substation. *Soil Dyn. Earthq. Eng.* **2023**, *173*, 108138. [CrossRef]
14. Bai, W.; Moustafa, M.A.; Dai, J. Seismic response of potential transformers and mitigation using innovative multiple tuned mass dampers. *Eng. Struct.* **2018**, *174*, 67–80.
15. Mosalam, K.M.; Günay, S.; Takhirov, S. Response evaluation of interconnected electrical substation equipment using real-time hybrid simulation on multiple shaking tables. *Earthq. Eng. Struct. Dyn.* **2016**, *45*, 2389–2404. [CrossRef]
16. He, C.; He, Z.; Zhu, W. Seismic interconnecting effects of multi-span flexible conductor-post electrical equipment coupling system. *J. Constr. Steel Res.* **2024**, *212*, 108209. [CrossRef]
17. Alessandri, S.; Giannini, R.; Paolacci, F.; Malena, M. Seismic retrofitting of an HV circuit breaker using base isolation with wire ropes. Part 1: Preliminary tests and analyses. *Eng. Struct.* **2015**, *98*, 251–262. [CrossRef]
18. Alessandri, S.; Giannini, R.; Paolacci, F.; Amoretti, M.; Freddo, A. Seismic retrofitting of an HV circuit breaker using base isolation with wire ropes. Part 2: Shaking-table test validation. *Eng. Struct.* **2015**, *98*, 263–274. [CrossRef]
19. Bai, W.; Dai, J.; Zhou, H.; Yang, Y.; Ning, X. Experimental and analytical studies on multiple tuned mass dampers for seismic protection of porcelain electrical equipment. *Earthq. Eng. Eng. Vib.* **2017**, *16*, 803–813. [CrossRef]
20. Zhu, W.; Xie, Q.; Liu, X. Seismic failure risk analysis of $\pm$800 kV coupling filter circuit considering material strength deviation. *Structures* **2023**, *47*, 1566–1578. [CrossRef]

21. He, C.; Jiang, L.Z.; Jiang, L.Q. Seismic failure risk assessment of post electrical equipment on supporting structures. *IEEE Trans. Power Deliv.* **2023**, *38*, 2757–2766. [CrossRef]

22. Bai, W.; Moustafa, M.A.; Dai, J. Seismic fragilities of high-voltage substation disconnect switches. *Earthq. Spectra* **2019**, *35*, 1559–1582.

23. Zhu, W.; Wu, M.; Xie, Q.; Chen, Y. Post-earthquake rapid assessment method for electrical function of equipment in substations. *IEEE Trans. Power Deliv.* **2023**, *38*, 3312–3321. [CrossRef]

24. Zhu, W.; Xie, Q. Post-earthquake rapid assessment for loop system in substation using ground motion signals. *Mech. Syst. Signal Process.* **2024**, *208*, 111058. [CrossRef]

25. Bellorini, S.; Salvetti, M.; Bettinali, F.; Zafferani, G. Seismic qualification of transformer high voltage bushings. *IEEE Trans. Power Deliv.* **1998**, *13*, 1208–1213. [CrossRef]

26. He, C.; Xie, Q.; Yang, Z.; Xue, S. Seismic evaluation and analysis of 1100-kVUHV porcelain transformer bushings. *Soil Dyn. Earthq. Eng.* **2019**, *123*, 498–512. [CrossRef]

27. He, C.; Xie, Q.; Zhou, Y. Influence of flange on seismic performance of 1100-kV ultra-high voltage transformer bushing. *Earthq. Spectra* **2019**, *35*, 447–469. [CrossRef]

28. Lu, J.; Xie, Q.; Zhu, W. Seismic damage detection of ultra-high voltage transformer bushings using output-only acceleration responses. *J. Civ. Struct. Health Monit.* **2023**, *13*, 1091–1104. [CrossRef]

29. Zhu, Q.; Zhang, X.; Gao, Z. Study on the Bending Stiffness of Joints Connecting Porcelain Bushings and Flanges in Ultra-High Voltage Electrical Equipment. *Appl. Sci.* **2022**, *12*, 5899. [CrossRef]

30. Xie, Q.; He, C.; Jiang, B.; Yang, Z. Linear-elastic analysis of seismic responses of porcelain post electrical equipment. *Eng. Struct.* **2019**, *201*, 109848. [CrossRef]

31. Napolitano, R.; Glisic, B. Hybrid physics-based modeling and data-driven method for diagnostics of masonry structures. *Comput.-Aided Civ. Infrastruct. Eng.* **2020**, *35*, 483–494. [CrossRef]

32. Torky, A.A.; Ohno, S. Deep learning techniques for predicting nonlinear multi-component seismic responses of structural buildings. *Comput. Struct.* **2021**, *252*, 106570. [CrossRef]

33. Goulet, J.A.; Michel, C.; Kiureghian, A.D. Data-driven post-earthquake rapid structural safety assessment. *Earthq. Eng. Struct. Dyn.* **2015**, *44*, 549–562. [CrossRef]

34. Ni, P.; Han, Q.; Du, X.; Cheng, X.; Zhou, H. Data-driven approach for post-earthquake condition and reliability assessment with approximate Bayesian computation. *Eng. Struct.* **2022**, *256*, 113940. [CrossRef]

35. Xu, Z.; Chen, J.; Shen, J.; Xiang, M. Recursive long short-term memory network for predicting nonlinear structural seismic response. *Eng. Struct.* **2022**, *250*, 113406. [CrossRef]

36. Liao, Y.; Lin, R.; Zhang, R.; Wu, G. Attention-based LSTM (AttLSTM) neural network for seismic response modeling of bridges. *Comput. Struct.* **2023**, *275*, 106915. [CrossRef]

37. Farahmand-Tabar, S.; Shirgir, S. Synergistic collaboration of motion-based metaheuristics for the strength prediction of cement-based mortar materials using TSK model. In *Handbook of Formal Optimization*; Springer Nature Singapore: Singapore, 2024; pp. 1–30.

38. Abdi, H.; Valentin, D.; Edelman, B. *Neural Networks*; Sage: Thousand Oaks, CA, USA, 1999.

39. Murphy, K.P. *Machine Learning: A Probabilistic Perspective*; MIT Press: Cambridge, MA, USA, 2012.

40. Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceedings of the ICNN'95-International Conference on Neural Networks, Perth, WA, Australia, 27 November–1 December 1995; Volume 4, pp. 1942–1948.

41. Stone, M. Cross-validatory choice and assessment of statistical predictions. *J. R. Stat. Soc. Ser. B (Methodol.)* **1974**, *36*, 111–133. [CrossRef]

42. *GB 50260-2013*; Code for Seismic Design of Electrical Installation. Ministry of Housing and Urban-Rural Development of the People's Republic of China (MOHURD): Beijing, China, 2013. (In Chinese)

43. Pacific Earthquake Engineering Research Center (PEER). PEER Ground Motion Database. 2011. Available online: https://ngawest2.berkeley.edu/ (accessed on 11 January 2023).

44. *IEEE 693 Standard*; Recommended Practice for Seismic Design of Substations. IEEE: New York, NY, USA, 2018.