Rapid prediction and optimization of mechanical response of the CEPC detector barrel yoke based on machine learning surrogate models

Shang Xia,[1,2] Wei He,[1,2,*] Quan Ji,[1,3] Fei-Peng Ning,[1,3] Jin-Fan Chang,[1,2] Xiao-Yan Ma,[1] Xiao-Hui Qian,[1,2] Ya-Tian Pei,[1,2,3] Jin-Yu Fu,[1,3] Shao-Jing Hou,[1,3] Jian Wang,[1] and Zhuo Liu[1,2]

[1]*Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, China*
[2]*University of Chinese Academy of Sciences, Beijing 100049, China*
[3]*Center for High Energy Physics, Henan Academy of Sciences, Zhengzhou 450046, China*

The Circular Electron-Positron Collider (CEPC) is a critical large-scale scientific facility for high-energy physics research. The barrel yoke, as the outermost and fundamental detector component, requires continuous design optimization. Conventional finite element analysis (FEA)-based optimization suffers from low computational efficiency. To address this, we propose a rapid mechanical response prediction and optimization method using machine learning surrogate models. Four algorithms—Ridge regression, extreme gradient boosting (XGBoost), support vector regression (SVR), and multilayer perceptron (MLP)—are trained on FEA data to predict maximum deformation and equivalent stress. All nonlinear models outperform the linear Ridge baseline; SVR achieves the best test R^2 of 0.9998 (maximum deformation) and 0.9651 (maximum equivalent stress). Using the SVR surrogates in conjunction with sequential least squares programming (SLSQP), we optimize the end-flange structural parameters. FEA validation yields prediction errors of 1.17% for maximum deformation and 0.52% for maximum equivalent stress, while optimization time is reduced from dozens of hours to minutes. This approach effectively increases optimization efficiency and provides a reusable framework for the rapid optimization of other subdetector mechanical structures.

Keywords: CEPC, Detector, Yoke, Machine learning, Surrogate model

I. INTRODUCTION

With the discovery of the Higgs boson, the focus of particle physics research has shifted to the high-precision measurement of its properties [1, 2]. The proton–proton collision mechanism of the Large Hadron Collider (LHC) suffers from inherent limitations, including uncertainty in the initial state and high background noise, which restrict the effective identification of Higgs events. The Circular Electron-Positron Collider (CEPC) scheme proposed by Chinese scientists adopts electron–positron collisions, which feature a well-defined initial state and low background noise, enabling efficient production of Higgs events [3–5]. The CEPC is designed with a higher luminosity, which imposes more stringent requirements on the performance of its detector system.

The structural layout of the CEPC detector is shown in Fig. 1, where the outermost component is the yoke [6]. The functions of the yoke are to form a return path for the magnetic flux of the solenoid superconducting magnet and provide support for all of the internal subdetectors and related components [7]. In addition, the yoke must provide installation space for the muon detectors and absorb all particles except muons [8]. The barrel yoke is divided into the barrel yoke and the end yokes. The barrel yoke, which provides structural support and fixation for the internal subdetectors and related components, is the most critical large-scale structural component of the CEPC detector.

To meet the precision requirements of the CEPC detector, the barrel yoke requires continuous optimization during the design phase. The conventional optimization workflow relies on repeated Finite Element Analysis (FEA), where every adjustment of design parameters requires a complete re-run of the simulation, resulting in a long optimization cycle and low computational efficiency.

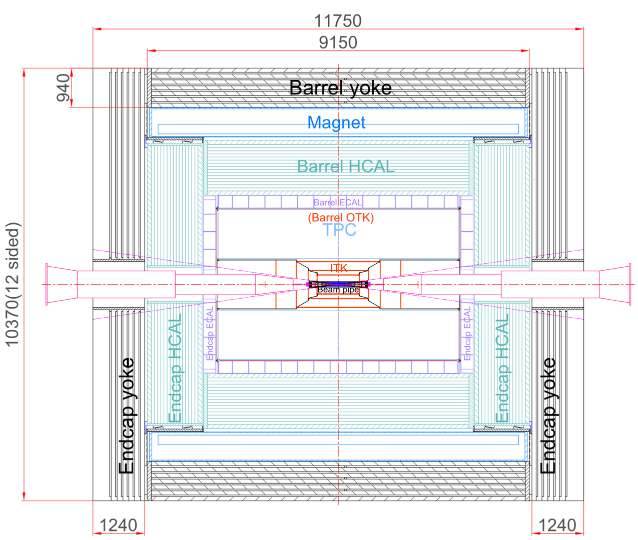

Fig. 1. Structural layout of the CEPC detector.

In recent years, machine learning surrogate models have attracted widespread attention and have been increasingly applied in structural design optimization [9]. By replacing high-fidelity numerical simulations with computationally efficient approximate mathematical models, surrogate models significantly improve computational efficiency while maintaining prediction accuracy, thereby establishing a vital technical pathway for the optimization of complex engineering structures. Various surrogate modeling methods have

* Corresponding author,hewei99@ihep.ac.cn

been adopted globally for structural optimization. For instance, Kriging models and their variants have been extensively used in the lightweight design and uncertainty quantification of aerospace structures [10], while artificial neural networks have demonstrated strong nonlinear fitting capabilities in damage identification and response prediction for civil engineering structures [11]. In addition, support vector regression, leveraging its excellent generalization performance with small sample sizes, has been successfully applied to the parameter optimization of mechanical structures [12]. However, existing research on surrogate-based optimization has predominantly focused on conventional-scale mechanical components or civil engineering structures, with relatively few studies addressing ultra-large-scale scientific facilities. The barrel yoke structure of the CEPC detector is characterized by its ultra-large dimensions, multi-layered staggered connections, and coupled multi-physics constraints. The complex mapping relationship between its mechanical responses and structural parameters imposes higher demands on the prediction accuracy, generalization capability, and optimization efficiency of surrogate models.

To address the aforementioned issues, this paper replaces time-consuming FEA with machine learning surrogate models. A training dataset is constructed through Latin Hypercube Sampling (LHS), and both a linear regression method (Ridge regression) and several nonlinear regression algorithms—including support vector regression (SVR), extreme gradient boosting (XGBoost), and multi-layer perceptron (MLP)—are employed to develop rapid prediction models for the mechanical response of the barrel yoke structure. The predictive performance of each algorithm is systematically compared, and the optimal surrogate models are subsequently integrated with the Sequential Least Squares Programming (SLSQP) gradient-based optimization algorithm to achieve automatic optimization of the structural parameters. The overall flowchart of this study is illustrated in Fig. 2.

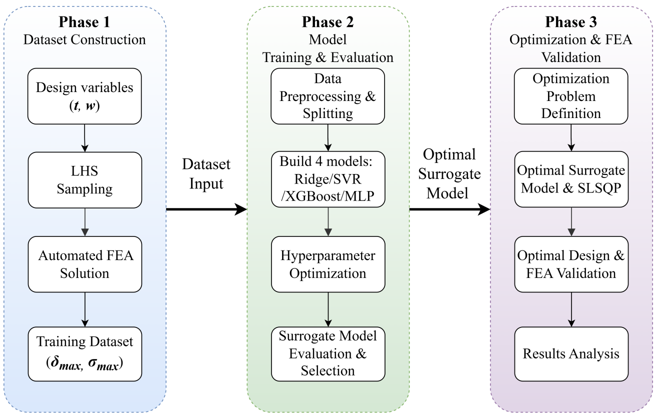

Fig. 2. Study flowchart.

The main features of this work are as follows: (1) the application of machine learning surrogate models to the structural optimization of the CEPC detector barrel yoke extends the applicability of surrogate modeling methods to the domain of large-scale scientific engineering; (2) through a systematic comparative analysis of four representative algorithms (linear model, kernel method, ensemble learning, and deep learning), the basis for surrogate model selection tailored to this class of complex structures is established; (3) a complete technical chain of "parametric modeling–LHS sampling–multi-model comparative evaluation–surrogate-based optimization–FEA validation" is established, providing a reusable methodological framework for the rapid iterative optimization of similar ultra-large-scale detector structures.

II. BARREL YOKE AND DATASET ACQUISITION

A. Parameters of barrel yoke

The barrel yoke structure of the CEPC detector is shown in Fig. 3. It consists of a polygonal staggered laminated structure composed of 12 barrel yoke modules [13] and two end flanges. This structure is referred to as the inter-supporting barrel yoke structure [6]. Through the synergistic bearing of the modules and end flanges, the realization of the structure's inter-supporting function is enabled.

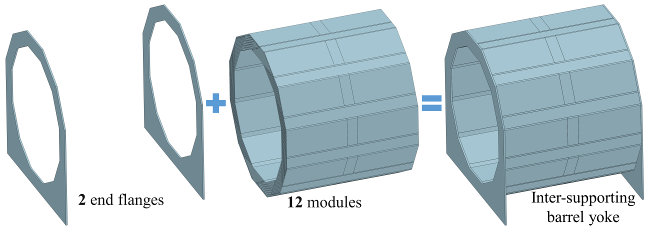

Fig. 3. Inter-supporting structure of the barrel yoke.

The structure and dimensions of the module are shown in Fig. 4 and Fig. 5. Each steel plate has a thickness of 100 mm, totaling 7 layers. The installation gap for the muon detectors in each layer is 40 mm, allowing the installation of 6 layers of muon detectors. The installation gap for the muon detectors in each layer of the module is divided into two sections by the isolation strips [6].

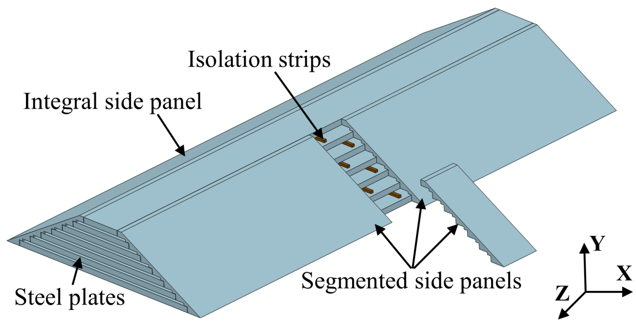

Fig. 4. Structure of the module.

The structure of the end flange is shown in Fig. 6. The structure features a dodecagonal inner boundary and an octag-

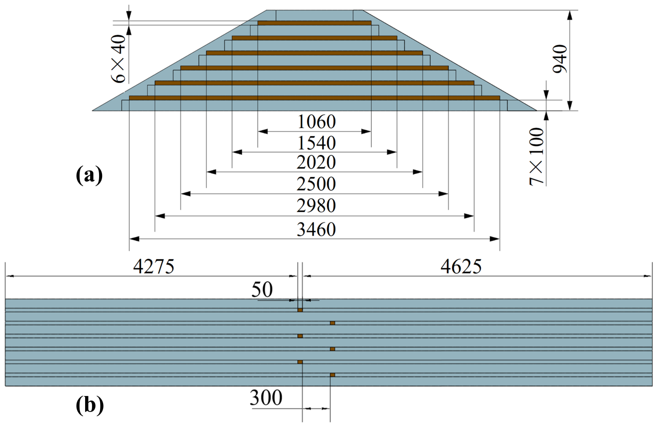

Fig. 5. Dimensional parameters of the module: (a) Z-direction view, (b) YZ-plane cross-sectional view

onal outer boundary, and the lower part of the outer boundary is equipped with a flange that extends to the bottom [6].

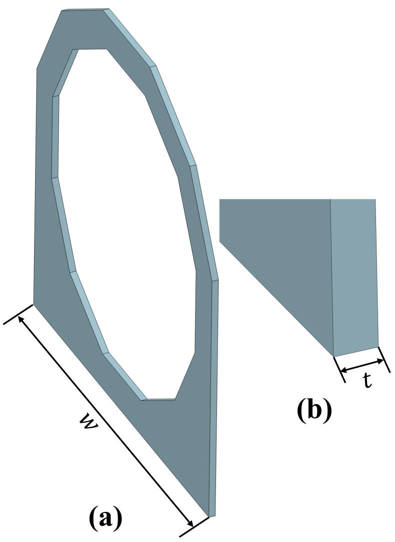

Fig. 6. Structure of the end flange. (a) w is the bottom support width, (b) t is the thickness.

B. Dataset acquisition

The design of the barrel yoke modules already satisfies the physical requirements of the muon detectors and the solenoid superconducting magnet. Parameters such as plate thickness, installation gaps, and the layout of isolation strips are strictly governed by the detector's physical specifications and are therefore excluded from the optimization. Accordingly, the optimization effort is directed toward the end flange structure. As a critical load-transferring component connecting the barrel yoke modules to the bottom support, the end flange has geometric parameters that directly influence the stiffness and strength distribution of the overall barrel yoke structure. The remaining geometric parameters of the end flange, such as the inner boundary dimensions and the outer boundary dimensions of the upper half, are determined by the external dimensions of the barrel yoke and the overall layout of the detector, leaving limited room for independent adjustment. Structural analysis reveals that increasing the bottom support area of the end flange enhances the stiffness and strength of the barrel yoke, thereby reducing the overall structural deformation and stress. The bottom support area of the end flange is primarily determined by two parameters: the thickness (t) and the bottom support width (w). Increasing either the thickness or the bottom support width enlarges the bottom support area, thereby influencing the mechanical performance of the barrel yoke structure. The thickness (t) and the bottom support width (w) of the end flange are selected as design variables. Due to design constraints, t and w are limited within the following ranges: $t \in [100, 300]$ mm and $w \in [10370, 15000]$ mm. The lower bound of t is determined by the geometric dimensions required for fixing the internal subdetectors [6], while its upper bound is governed by the actual manufacturing and fabrication processes. The lower bound of w is set by the dodecagonal outer-profile dimensions resulting from the assembly of the 12 modules of the barrel yoke, and its upper bound is dictated by the physical size constraints of the actual underground experimental hall described in Ref. [6].

A total of 1500 sample points were generated in the feature parameter space using the Latin Hypercube Sampling (LHS) method [14]. The sample size was determined based on the following comprehensive considerations: (1) The design space in this study involves only two feature parameters and is therefore low-dimensional; 1500 sample points can achieve dense coverage of the two-dimensional design space, ensuring the completeness of the training data for the surrogate models. (2) The LHS method partitions the value range of each input feature parameter into several equal subintervals and randomly selects one sample point within each subinterval, thereby ensuring uniform filling of the design space, effectively avoiding the clustering and voids typical of traditional random sampling, and significantly improving sample representativeness. (3) Each sample point requires one complete FEA; the total computational time for 1500 sample points amounts to approximately 250 hours, which provides sufficient training data for the surrogate models within an acceptable one-time time investment. (4) The dataset is divided into a training set and a test set at an 8:2 ratio; the total sample size of 1500 ensures that both the training and test sets possess adequate statistical significance, guaranteeing the objectivity and reliability of the model performance evaluation. The maximum deformation (δ_{max}) and maximum equivalent stress (σ_{max}) of the barrel yoke are selected as the prediction targets.

The FEA is performed using the ANSYS software. The material of the barrel yoke is structural steel (10#), whose properties are listed in Table 1 [15, 16]. The barrel yoke structure is assembled from steel plates, and the actual connections are all complex bolted joints, while the contacts exhibit frictional nonlinearity. These factors make the simulation extremely difficult. Moreover, the current computational resources are inadequate to carry out the relevant computations. To facilitate simulation, the complex internal contacts of the modules are simplified as bonded contact. The contacts between modules and between modules and end flanges are simplified as bonded contact. To ensure the efficiency of FEA, the smallest mesh size is set to $50\,\mathrm{mm}$. The fixed support is the bottom surface of the barrel yoke. The loads mainly include the self-weight of the barrel yoke and the 2145-ton gravitational loads of the internal subdetectors and related components. Among these, the gravitational load of $60\,\mathrm{t}$ from the barrel muon detector is applied to the 12 modules of the barrel yoke, while the gravitational load of $2085\,\mathrm{t}$ from the other internal subdetectors and related components is applied to the dodecagonal inner boundary on the end flange.

Table 1. Material parameters of structural steel (10#).

Property	Value
Density	$7850\,\mathrm{kg\,m^{-3}}$
Elastic modulus	$206\,\mathrm{GPa}$
Poisson's ratio	0.3
Tensile strength	$335\,\mathrm{MPa}$
Yield strength	$205\,\mathrm{MPa}$
Allowable stress	$136.7\,\mathrm{MPa}$

After setting up the 1500 sample points, the software automatically modifies the parametrized barrel yoke model and executes a fully automated FEA using the same settings as the previous analysis. Upon completion, the data from each sample point are extracted, and a training dataset is subsequently built.

III. SURROGATE MODELS OF MACHINE LEARNING

Surrogate models are essential tools in the field of design optimization. By replacing high-fidelity numerical models with computationally inexpensive approximate models, they can significantly improve computational efficiency while guaranteeing prediction accuracy. The prediction principles, model construction, and prediction results of the surrogate models established by four algorithms, Ridge regression, eXtreme Gradient Boosting (XGBoost), Support Vector Regression (SVR), and Multi-Layer Perceptron (MLP), are presented in the subsequent sections.

A. Ridge regression model

1. Prediction principle

Ridge regression is a linear regression method with L2 regularization constraints [17]. To address the inherent limitations of ordinary linear regression, including excessively high variance in coefficient estimation, poor numerical stability, and degraded generalization capability caused by proneness to overfitting, this method imposes an L2 regularization penalty term on the regression coefficients to achieve the optimal trade-off between model bias and variance, thereby endowing the model with stronger anti-overfitting performance and numerical robustness. In this paper, the Ridge regression model is adopted as the baseline model. The mathematical expression of the Ridge regression model is given as follows:

$$y = \omega_1 x_1 + \omega_2 x_2 + \cdots + \omega_n x_n + b \tag{1}$$

$$L(W) = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 + \alpha \sum_{j=1}^{p} (\omega_j)^2 \tag{2}$$

$$W = (X^T X + \alpha I)^{-1} X^T Y \tag{3}$$

where $L(W)$ denotes the loss function, the first term is the sum of squared residuals, and the second term is the L2 regularization term. N is the number of samples. y_i is the ground-truth value. $\hat{y}_i$ is the predicted value. α is the regularization strength. p is the number of weights. ω_j is the weight coefficient of the model. Equation (3) presents the closed-form analytical solution of the Ridge regression model, through which the optimal regression coefficients can be obtained via direct solution. I is the identity matrix of order p. By adding the regularization term αI to the diagonal elements of the $X^T X$ matrix, the permanent invertibility of the matrix can be guaranteed, which further enhances the stability and robustness of the model.

2. Model construction

The core mechanical response indicators of the barrel yoke include δ_{max} and σ_{max}, with significant differences in numerical magnitude and data distribution characteristics between the two types of indicators. To ensure the accuracy of prediction results and the generalization performance of the models, independent Ridge regression models are constructed for δ_{max} and σ_{max}, respectively. Model training, hyperparameter optimization, and performance validation are carried out separately. The Ridge regression models are mainly implemented based on the scikit-learn and are written in Python.

First, the sample data is read and preprocessed. A normalization method is adopted to map the input features and target values to the interval $[0, 1]$, eliminating dimensional differences and numerical magnitude differences between different indicators and ensuring that the contributions of each feature to the model are comparable.

Subsequently, dataset splitting is performed, where the entire sample dataset is randomly divided into a training set and a test set at an 8:2 ratio. The training set is used for model fitting and hyperparameter optimization, while the test set is not involved in the training process at all and is only used for the unbiased evaluation of the model's generalization capability, ensuring the objectivity of performance validation. The data splitting is kept consistent for all models in this paper.

Model hyperparameter optimization is conducted using the grid search method. Taking the maximization of the coefficient of determination on the training set as the optimization objective, the optimal hyperparameters of the models corresponding to the two mechanical response indicators are determined respectively, ensuring that the model achieves an optimal balance between the fitting effect on the training set and the generalization capability on the test set. After grid search, the optimal hyperparameters of the models are obtained as shown in Table 2. The models are trained using the optimal hyperparameters, and then the prediction results are output and subjected to data processing.

Table 2. Hyperparameters of the Ridge regression models.

Hyperparameter	δ_{max} pred. model	σ_{max} pred. model
α	0.1	1
Solver	sag	sag

3. *Result*

As shown in Fig. 7(a), for the maximum deformation prediction model, the coefficient of determination (R^2) and root mean square error (RMSE) of the training set are 0.9614 and $0.0736\,\mathrm{mm}$, respectively. For the test set, the R^2 is 0.9648, and the RMSE is $0.0719\,\mathrm{mm}$. The results of the maximum deformation prediction model indicate that the model has a certain reference value. However, in terms of the fitting performance, the prediction accuracy of the model shows obvious deviations under small and large deformation working conditions, and the model cannot characterize the nonlinear variation law between maximum deformation and input parameters. The overall prediction capability of the model is insufficient, which is difficult to meet the demand for high-precision prediction.

As shown in Fig. 7(b), for the maximum equivalent stress prediction model, the R^2 and RMSE of the training set are 0.8619 and $6.5907\,\mathrm{MPa}$, respectively. For the test set, the R^2 is 0.8935, and the RMSE is $5.7887\,\mathrm{MPa}$. This indicates that the model cannot achieve accurate stress prediction, with its accuracy far below the acceptable threshold. The scatter points in the plot deviate significantly from the diagonal line, which verifies the existence of strong nonlinear mechanical behavior between the maximum equivalent stress and the end flange thickness as well as the bottom support width, necessitating the adoption of nonlinear models for fitting. The predictive capability of the model fails to meet the requirements for high-precision prediction.

B. SVR model

1. *Prediction principle*

SVR is an extension of the Support Vector Machine (SVM) algorithm for regression tasks [18]. By introducing the ε-

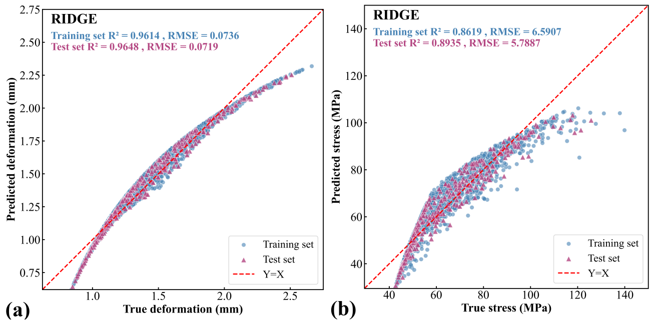

Fig. 7. Scatter plots of ground-truth versus predicted values from the Ridge regression models: (a) δ_{max} prediction model; (b) σ_{max} prediction model.

insensitive loss function, the SVR algorithm allows a maximum deviation of ε between the predicted and ground-truth values without incurring any loss. For samples with deviations exceeding this threshold, the loss of the function is calculated via support vectors, and the best model is finally obtained by minimizing the total loss and maximizing the margin. For nonlinear modeling, features can be mapped into a high-dimensional feature space through a kernel function before regression processing is performed. This algorithm exhibits excellent robustness and generalization performance. The mathematical formulation of the SVR model is given as follows:

$$f(x) = \omega \cdot x + b \tag{4}$$

$$L(\omega, b) = \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{N}\mathcal{L}_\varepsilon(f(x_i) - y_i) \tag{5}$$

$$\mathcal{L}_\varepsilon(z) = \begin{cases} 0, & |z| \leq \varepsilon \\ |z| - \varepsilon, & |z| \geq \varepsilon \end{cases} \tag{6}$$

$$f(x) = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*)K(x_i, x) + b \tag{7}$$

In Equation (4), ω is the weight vector of the model, b is the bias term, and x is the input feature vector. In Equation (5), the first term represents the model complexity, while the second term corresponds to the empirical error. C is the regularization penalty coefficient, $f(x_i)$ is the predicted value, and y_i is the ground-truth value. Equation (6) defines the ε-insensitive loss function. Equation (7) is the prediction function, where α_i and α_i^* are Lagrange multipliers satisfying the constraints $\alpha_i \cdot \alpha_i^* = 0$, $\alpha_i \geq 0$, and $\alpha_i^* \geq 0$. Only the samples corresponding to a small number of non-zero Lagrange multipliers are the support vectors, and $K(x_i, x)$ is the kernel function.

2. Model construction

The data preprocessing, dataset splitting, and grid search workflow for the SVR models are consistent with those described in Section 3.1.2 for the Ridge regression models. The models are likewise built using the scikit-learn and are written in Python. Independent models are constructed for δ_{max} and σ_{max}. After grid search, the optimal hyperparameters of the models are obtained as shown in Table 3.

Table 3. Hyperparameters of the SVR models.

Hyperparameter	δ_{max} pred. model	σ_{max} pred. model
C	10	100
epsilon	0.01	0.5
gamma	0.1	scale
kernel	rbf	rbf

3. Result

As shown in Fig. 8(a), for the maximum deformation prediction model, the R^2 and RMSE of the training set are 0.9998 and $0.0057\,\mathrm{mm}$, respectively. For the test set, the R^2 is 0.9998, and the RMSE is $0.0056\,\mathrm{mm}$. The ground-truth values and the corresponding predicted values for both the training and test sets show excellent agreement, enabling accurate prediction of the maximum deformation of the barrel yoke structure. The prediction error of the model is controlled within $0.006\,\mathrm{mm}$, which fully meets the prediction accuracy requirements for the structural maximum deformation.

As shown in Fig. 8(b), for the maximum equivalent stress prediction model, the R^2 and RMSE of the training set are 0.9500 and $3.9634\,\mathrm{MPa}$, respectively. For the test set, the R^2 is 0.9651, and the RMSE is $3.3134\,\mathrm{MPa}$. The prediction error of the model is controlled within $4.0\,\mathrm{MPa}$, satisfying the prediction accuracy requirements for the structural maximum equivalent stress. All evaluation metrics of the model on the test set outperform those on the training set, indicating that the model is free from overfitting, has successfully captured the core law of the mechanical response of the barrel yoke structure, and exhibits excellent generalization capability.

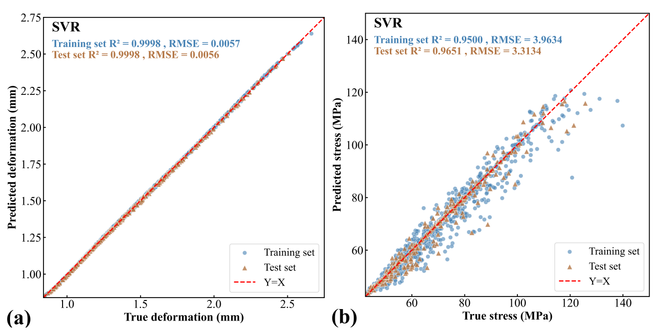

Fig. 8. Scatter plots of ground-truth versus predicted values from the SVR model. (a) δ_{max} prediction model; (b) σ_{max} prediction model.

C. XGBoost model

1. Prediction principle

XGBoost is an efficient optimized ensemble learning algorithm based on the Gradient Boosting Decision Tree (GBDT) [19]. This algorithm optimizes the loss function by introducing second-order Taylor expansion, controls model complexity via regularization terms, and improves computational efficiency through the approximate splitting algorithm and histogram algorithm. Combined with techniques such as column sampling and row sampling, it achieves accurate data prediction, and exhibits the advantages of fast computation speed, high prediction accuracy, strong generalization capability, and native support for missing value handling. The XGBoost model is an additive model consisting of K regression decision trees, with its mathematical formulation given as follows:

$$\hat{y}_i = \phi(x_i) = \sum_{k=1}^{K} f_k(x_i), \quad f_k \in \mathcal{F} \tag{8}$$

$$L(\phi) = \sum_{i=1}^{N} l(\hat{y}_i, y_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{9}$$

$$\Omega(f_k) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} \omega_j^2 \tag{10}$$

where K denotes the number of decision trees, $\mathcal{F}$ represents the function space of all regression decision trees. $f_k(\boldsymbol{x}_i)$ is the predicted value of sample x_i from the k-th tree, and $\hat{y}_i$ is the final predicted value of sample x_i. $L(\phi)$ is the objective function, where the first term is the empirical loss term. N is the number of samples. $l(\hat{y}_i, y_i)$ is the differentiable loss function. The second term is the regularization term $\Omega(f_k)$, where γ is the regularization coefficient controlling the number of leaf nodes. T is the number of leaf nodes of the k-th tree, λ is the L2 regularization coefficient controlling the weights of leaf nodes, and ω_j is the weight of the j-th leaf node.

2. Model construction

The XGBoost models are likewise built using the scikit-learn and are written in Python. The data preprocessing, dataset splitting, and grid search workflow are consistent with Section 3.1.2. Independent models are constructed for δ_{max} and σ_{max}. After grid search, the optimal hyperparameters of the models are obtained as shown in Table 4.

3. Result

As shown in Fig. 9(a), for the maximum deformation prediction model, the R^2 and RMSE of the training set are

Table 4. Hyperparameters of the XGBoost models.

Hyperparameter	δ_{max} pred. model	σ_{max} pred. model
colsample_bytree	0.7	0.7
learning rate	0.1	0.05
max_depth	3	3
n_estimators	300	300
reg_alpha	0.1	1
reg_lambda	5	5
subsample	0.8	0.8

0.9932 and $0.0309\,\mathrm{mm}$, respectively. For the test set, the R^2 is 0.9924, and the RMSE is $0.0334\,\mathrm{mm}$. The ground-truth values and the corresponding predicted values for both the training and test sets show close agreement, and the prediction error of the model is controlled within $0.034\,\mathrm{mm}$, which meets the prediction accuracy requirements for the maximum structural deformation. From the sample distribution, the model exhibits satisfactory fitting performance under small and medium deformation conditions, whereas slight deviations emerge under large deformation conditions with noticeable prediction errors, demonstrating insufficient fitting and extrapolation capability for large deformation scenarios. This is also a prevalent problem inherent to tree-based models in regions with sparse boundary samples.

As shown in Fig. 9(b), for the maximum equivalent stress prediction model, the R^2 and RMSE of the training set are 0.9566 and $3.6964\,\mathrm{MPa}$, respectively. For the test set, the R^2 is 0.9569, and the RMSE is $3.6828\,\mathrm{MPa}$. The prediction error is controlled within $3.7\,\mathrm{MPa}$, satisfying the prediction accuracy requirements for the maximum equivalent stress of the structure.

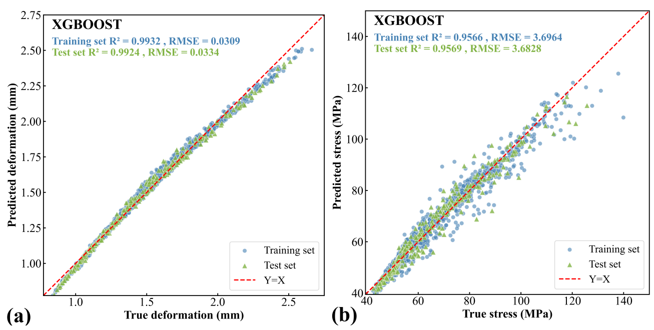

Fig. 9. Scatter plots of ground-truth versus predicted values from the XGBoost model. (a) δ_{max} prediction model; (b) σ_{max} prediction model.

D. MLP model

1. *Prediction principle*

The MLP is the most fundamental feedforward neural network model in the field of deep learning. To address the inherent limitation that traditional linear models struggle to fit complex nonlinear relationships, MLP constructs a hierarchical feature extraction and transformation structure by introducing multiple hidden layers and nonlinear activation functions, which realizes deep nonlinear mapping of input data and exhibits an extremely strong fitting capability for complex functions [20].

The MLP model consists of an input layer, multiple hidden layers, and an output layer. Its forward propagation process can be expressed as follows: for an input sample x, the output $h_1 = \sigma(W_1 x + b_1)$ is obtained first through linear transformation and nonlinear activation in the first hidden layer. If the model contains multiple hidden layers, the signal is propagated layer by layer, and the output h_l of the l-th layer is calculated from the output h_{l-1} of the previous layer: $h_l = \sigma(W_l h_{l-1} + b_l)$. Finally, the model obtains the predicted value $\hat{y} = W_{out} h_{last} + b_{out}$ through the output layer. Here, σ is the nonlinear activation function, which introduces nonlinear representation capability and enables MLP to fit complex nonlinear functional relationships. Both the number of hidden layers and the number of neurons in each layer are hyperparameters of the model, which determine the network capacity.

The parameter optimization of MLP relies on the backpropagation algorithm. This algorithm calculates the gradient of the loss function with respect to the parameters of each layer using the chain rule, and updates the weights and biases via the gradient descent algorithm. The mean squared error (MSE) is usually adopted as its loss function, with the expression shown in Equation (11). In the equation, N is the number of samples, $\hat{y}_i$ is the predicted value, and y_i is the ground-truth value.

$$L(W, b) = \frac{1}{N} \sum_{i=1}^{N} (\hat{y}_i - y_i)^2 \tag{11}$$

2. *Model construction*

The MLP models are built on the PyTorch framework, employing ReLU activation functions, BatchNorm normalization, and Dropout regularization, with MSE as the loss function, Adam as the optimizer, and an early stopping strategy to prevent overfitting. The data preprocessing and dataset splitting scheme are consistent with Section 3.1.2, and independent models are constructed for δ_{max} and σ_{max}. The optimal hyperparameters of the models are shown in Table 5.

Table 5. Hyperparameters of the MLP models.

Hyperparameter	δ_{max} pred. model	σ_{max} pred. model
(fc1, fc2, fc3)	(128, 64, 32)	(128, 64, 32)
batch_size	64	64
Epochs	300	400
Lr	0.001	0.0008
Weight decay	1×10^{-6}	1×10^{-7}
Patience	30	40

3. Result

As shown in Fig. 10(a), for the maximum deformation prediction model, the R^2 and RMSE of the training set are 0.9937 and $0.0296\,\mathrm{mm}$, respectively. For the test set, the R^2 is 0.9933, and the RMSE is $0.0313\,\mathrm{mm}$. The ground-truth values and the corresponding predicted values for both the training and test sets show very close agreement, and the prediction error of the model is controlled within $0.032\,\mathrm{mm}$, which meets the prediction accuracy requirements for the maximum structural deformation. In terms of the fitting performance over the full interval, the model achieves good prediction accuracy under small and medium deformation conditions, with minor prediction deviations occurring under large deformation conditions. This reflects that the model possesses a relatively weak fitting ability for sparse samples at the data boundary and exhibits certain limitations in the prediction of large deformation conditions.

As shown in Fig. 10(b), for the maximum equivalent stress prediction model, the R^2 and RMSE of the training set are 0.9520 and $3.8860\,\mathrm{MPa}$, respectively. For the test set, the R^2 is 0.9661, and the RMSE is $3.2662\,\mathrm{MPa}$. The prediction error of the model is controlled within $3.9\,\mathrm{MPa}$, satisfying the prediction accuracy requirements for the structural stress. All evaluation metrics of the model on the test set outperform those on the training set, indicating that the model is free from overfitting, has successfully captured the core law of the mechanical response of the barrel yoke, and exhibits excellent generalization capability.

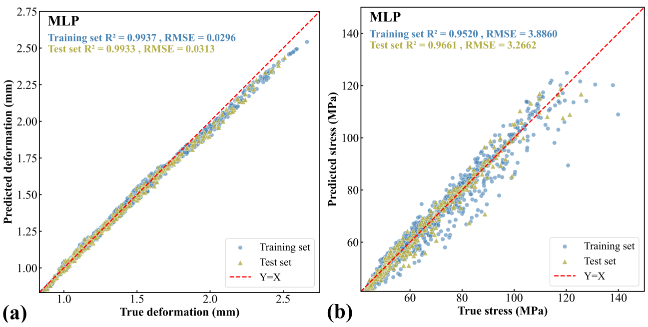

Fig. 10. Scatter plots of ground-truth versus predicted values from the MLP model. (a) δ_{max} prediction model; (b) σ_{max} prediction model.

IV. RESULT COMPARISON

To quantitatively compare the performance differences among the four models, the coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RMSE) are selected as the quantitative evaluation metrics for model performance to facilitate comparative analysis. The formulas for each metric are given as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{12}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{13}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{14}$$

where n is the number of samples, y_i is the ground-truth value, $\hat{y}_i$ is the predicted value, and $\bar{y}$ is the mean of the ground-truth values of the samples. For the MAE and RMSE, smaller values indicate a smaller prediction error and higher prediction accuracy of the model. The R^2 has a value range of $(-\infty, 1]$; the closer the value is to 1, the better the fitting performance of the model on the data, and the stronger the explanatory power of the model for the target variable.

As can be seen from Figs. 7, 8, 9, 10 and 11:

First, there is no significant difference between the evaluation metrics of the training set and test set for all models, and no overfitting is observed, indicating that the generalization capability of all models meets the engineering application requirements.

For the prediction of maximum deformation, all nonlinear models (SVR, XGBoost, MLP) outperform the linear baseline model (Ridge regression) in prediction accuracy. This verifies that a nonlinear mapping relationship exists between the structural deformation and the end flange thickness as well as the bottom support width, and that nonlinear models are essential to achieve high-precision prediction. The SVR model exhibits the optimal performance, with an extremely high R^2 on both the training and test sets, reflecting excellent fitting accuracy. Meanwhile, its MAE and RMSE are extremely low, meaning the model has the lowest average prediction deviation and more robust prediction performance for extreme samples. The MLP and XGBoost models follow, with slightly lower accuracy than the SVR model but still outperforming the Ridge regression model.

For the prediction of the maximum equivalent stress, the nonlinear characteristics are far more pronounced than those for maximum deformation prediction. The accuracy of the linear baseline model (Ridge regression) drops sharply, showing a significant gap with the nonlinear models. This directly confirms the existence of a strong nonlinear mapping relationship between the structural maximum equivalent stress and the end flange thickness as well as the bottom support width, and that nonlinear models are indispensable for high-precision prediction of this indicator. All three nonlinear models (SVR, MLP, and XGBoost) exhibit favorable prediction performance, with the test set R^2 all above 0.95 and the inter-model difference less than 0.01, indicating that all three models have excellent fitting capability for the mechanical response law of structural stress. Among them, the SVR model

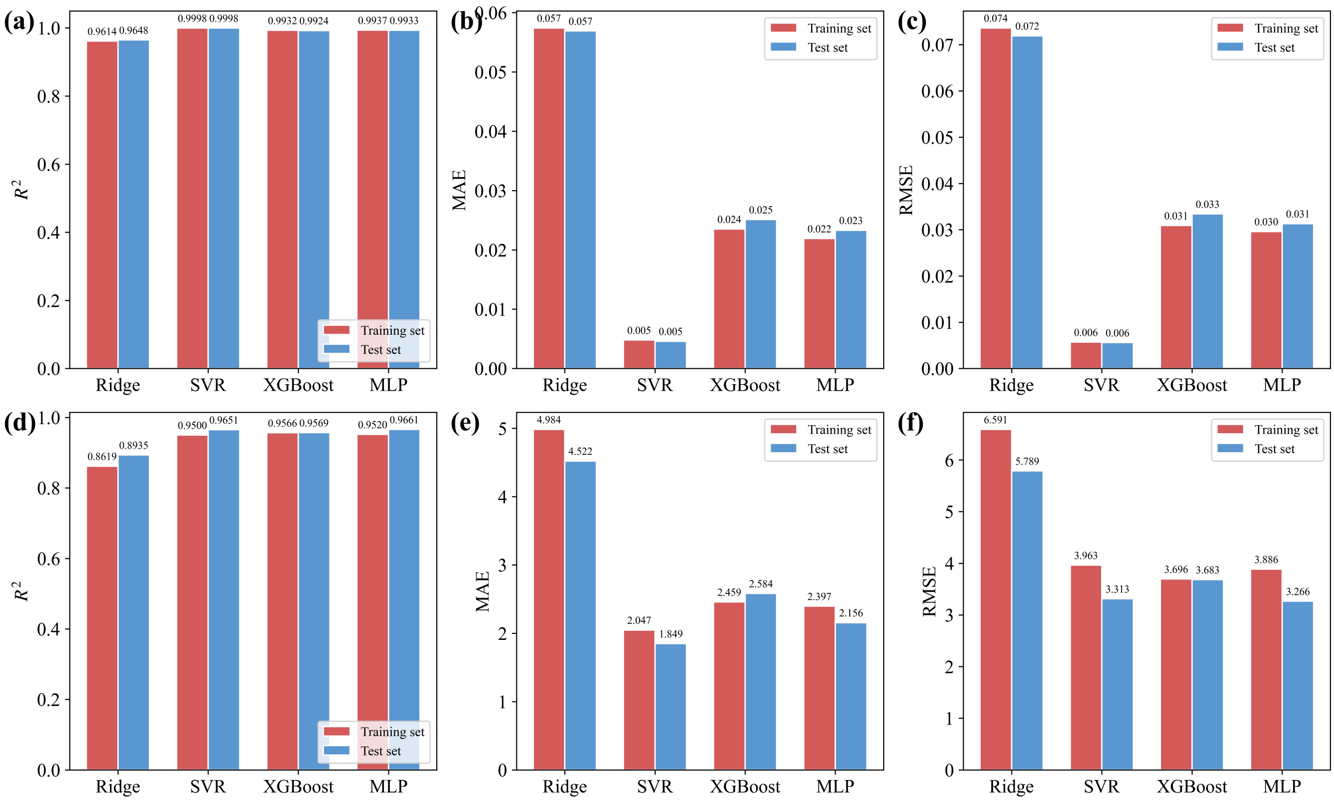

Fig. 11. Bar charts of evaluation metrics for prediction models based on the four algorithms. (a), (b), (c) are evaluation metrics for the δ_{max} prediction models; (d), (e), (f) are evaluation metrics for the σ_{max} prediction models.

achieves the minimum MAE, corresponding to the lowest average prediction deviation, while the MLP model obtains the minimum RMSE, presenting more robust prediction for extreme samples.

From the perspective of prediction accuracy and stability, the SVR model exhibits more remarkable comprehensive performance compared with the other three models. Consequently, the SVR model is the most suitable surrogate model for the prediction of maximum deformation and maximum equivalent stress of the CEPC detector barrel yoke.

V. STRUCTURAL OPTIMIZATION BASED ON THE SVR SURROGATE MODELS AND VALIDATION

According to the design requirements in Reference [6], under the combined loading of the barrel yoke structure's self-weight and the gravitational loads of the internal subdetectors and related components, the maximum deformation of the barrel yoke must not exceed 2 mm, and the maximum equivalent stress must not exceed the allowable stress. Under these two constraints, the mass of the end flange must be minimized to reduce material cost and structural self-weight. In this section, the structural optimization problem is formulated under the above constraints. The SVR surrogate models trained in Section 3 are employed to replace FEA, and the Sequential Least Squares Programming (SLSQP) algorithm is adopted to realize rapid optimization design of the end flange. The optimization results are subsequently validated through FEA.

A. Optimization problem formulation

The mass (m) of the end flange can be expressed as a function of t and w. Based on the geometric model of the end flange, the relationship between mass and the design variables is derived as:

$$m \approx t \cdot (38.9426 + 0.025804 \cdot w) \tag{15}$$

where t and w are in units of mm, and m is in units of kg. Integrating the design constraints and the optimization objective, the mathematical formulation of the end flange structural optimization problem is:

$$\begin{aligned} \min \quad & m(t,w) \approx t \cdot (38.9426 + 0.025804 \cdot w) \\ \text{s.t.} \quad & \delta_{max}(t,w) \leq 2 \text{ mm} \\ & \sigma_{max}(t,w) \leq 136.7 \text{ MPa} \\ & 100 \leq t \leq 300 \text{ mm} \\ & 10370 \leq w \leq 15000 \text{ mm} \end{aligned} \tag{16}$$

where $\delta_{max}(t,w)$ and $\sigma_{max}(t,w)$ denote the maximum deformation and maximum equivalent stress of the barrel yoke structure, respectively. Both are nonlinear implicit functions of the design variables $(t,\ w)$. In conventional optimization workflows, these quantities must be obtained through FEA. They are replaced by the trained SVR surrogate model now.

B. Surrogate-model-based optimization method

For this optimization problem, the SVR surrogate models trained in Section 3 are employed as fast evaluators of the mechanical response functions in lieu of FEA. During the optimization process, the optimization algorithm generates candidate design points $(t,\ w)$ at each iteration, and the SVR models predict the corresponding $\delta_{max}(t, w)$ and $\sigma_{max}(t, w)$, which are returned to the optimizer to evaluate the objective function and constraint satisfaction without invoking an FEA solver. This approach shifts the dominant computational burden of optimization from FEA solving to the trained SVR surrogate models. The optimization process is implemented within the scipy.optimize framework of SciPy and written in Python [21].

The optimization algorithm adopted is SLSQP, a gradient-based constrained optimization method built on Lagrangian–Newton iteration [22]. Its core idea is to approximate the original optimization problem as a quadratic programming subproblem at the current iterate: the Lagrangian function is expanded to second order via a Taylor series, and the constraint functions are approximated to first order linearization. The search direction obtained by solving this subproblem simultaneously accounts for the reduction of the objective function and the satisfaction of the constraints. A one-dimensional line search is then performed along this direction to determine an appropriate step size. The algorithm alternates between subproblem solution and design-point update until the convergence criterion is satisfied.

SLSQP was selected as the optimization solver for the following reasons: (1) the optimization problem involves only two design variables and two inequality constraints, constituting a small-scale continuous optimization problem for which gradient-based methods are best suited; (2) compared with metaheuristic methods such as genetic algorithms and particle swarm optimization, SLSQP leverages gradient information to achieve superlinear convergence rates, significantly reducing the number of objective function evaluations; and (3) SLSQP can simultaneously handle inequality constraints, equality constraints, and variable bound constraints, which fully aligns with the form of the problem at hand.

C. Optimization results, validation, and analysis

The detailed optimization results are presented in Table 6. The SVR model predicts δ_{max} to be exactly $2\,\mathrm{mm}$, indicating that deformation is a tight constraint; σ_{max} is $94.580\,\mathrm{MPa}$, well below the allowable stress, indicating a loose constraint. Therefore, under this loading condition, the stiffness of the barrel yoke structure is the dominant design-driving factor. It should be noted that the optimal thickness $t = 100\,\mathrm{mm}$ lies exactly at the lower bound of the design variable. This is not a convergence error caused by model prediction bias or sparse boundary samples, but rather arises from the geometric constraints imposed on the end flange by the fixed dimensions of the subdetectors. $100\,\mathrm{mm}$ is the minimum allowable thickness [6]. Under this engineering constraint, stiffness, rather than strength, governs the optimal design.

Table 6. Optimization results of the SVR model with SLSQP algorithm.

Parameter	Optimal value
t	100 mm
w	12 491.28 mm
m	36 126.76 kg
δ_{max}	2 mm
σ_{max}	94.580 MPa

To verify the reliability of the SVR surrogate-model-based optimization results, an actual finite element model was built according to the optimal design parameters, and FEA was performed using exactly the same simulation settings as described in Section 2.2. The comparison of the SVR predicted values and the FEA values is shown in Table 7.

Table 7. Comparison between SVR predictions and FEA simulation results.

	δ_{max}	σ_{max}
SVR predicted value	2.0000 mm	94.580 MPa
FEA value	2.0238 mm	94.087 MPa
Absolute error	0.0236 mm	0.493 MPa
Relative error	1.17%	0.52%

As shown in Table 7, the SVR predictions agree well with the FEA results. The relative error of δ_{max} is 1.17%. The FEA value of $2.0238\,\mathrm{mm}$ slightly exceeds the $2\,\mathrm{mm}$ design limit, which is within an acceptable range and can be addressed by appropriately tightening the deformation constraint to reserve a safety margin. The relative error of σ_{max} is 0.52%. The FEA value of $94.087\,\mathrm{MPa}$ remains well below the allowable stress of $136.7\,\mathrm{MPa}$, providing an ample safety margin. Fig. 12 shows the total deformation and equivalent stress contour plots of the optimized barrel yoke structure.

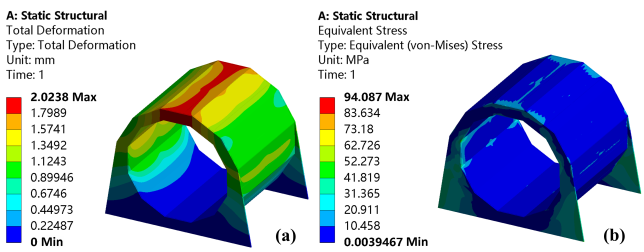

Fig. 12. Total deformation and equivalent stress contour plots of the optimized barrel yoke structure.

The slight prediction deviation of the SVR models at the optimal solution primarily originates from two sources:

(1) the optimal solution lies in a boundary region of the design space where the LHS sample density is relatively low; (2) the SVR model is a smooth regression model that exhibits a smoothing effect on local nonlinear details. Overall, the deviation between the SVR predictions and the FEA results is within an acceptable range, validating the feasibility and reliability of substituting the SVR surrogate models for FEA in structural optimization.

In conventional structural optimization, each design parameter adjustment requires a full FEA evaluation of the mechanical response. A single solve takes approximately 20 minutes including pre- and post-processing. If the optimization results do not meet the design requirements, engineers rely on experience to iteratively adjust parameters and recompute, with total time potentially reaching tens of hours. Moreover, manual judgment is prone to becoming trapped in local optima. In contrast, the proposed SVR surrogate model combined with the SLSQP algorithm can automatically complete the optimization within seconds and directly output the optimal parameters, after which only a single FEA validation of the optimal design scheme is required. Consequently, the proposed method shifts the dominant computational burden from FEA to the surrogate models, reducing the FEA computation time required for structural optimization from tens of hours to the order of minutes, while also avoiding the local-optimum problem inherent in manual optimization.

VI. CONCLUSIONS

Through structural parameterization and automated FEA, a dataset of maximum response values for the CEPC detector barrel yoke under various parameter combinations was obtained. Based on this dataset, four machine learning algorithms, namely Ridge regression, SVR, XGBoost, and MLP, were respectively employed to construct surrogate models, and the models were trained and their performance evaluated. And the optimal surrogate models were combined with the optimization algorithm to optimize the structural parameters of the end flange, leading to the following conclusions:

1. The thickness and bottom support width of the end flange were selected as the input features of the four models, and the maximum deformation and maximum equivalent stress of the barrel yoke were taken as the prediction output. Through comprehensive analysis of the evaluation metrics of the four models, the SVR model is determined to achieve the optimal prediction performance for both the maximum deformation and maximum equivalent stress of the CEPC detector barrel yoke.
2. Based on the trained SVR surrogate models combined with the SLSQP optimization algorithm, automatic optimization of the end flange structural parameters was realized, yielding optimal values of thickness $100\,\mathrm{mm}$ and bottom support width $12\,491.28\,\mathrm{mm}$. The optimal design scheme was validated through FEA, with prediction errors of 1.17% for maximum deformation and 0.52% for maximum equivalent stress, showing that the SVR surrogate model can effectively replace FEA for structural optimization. Compared with the conventional FEA-based optimization method, which requires several to over ten hours of computation, the proposed method reduces the optimization solution time to the minute level while maintaining prediction accuracy, thereby validating the feasibility and efficiency advantages of the SVR surrogate models in the structural optimization of the CEPC detector barrel yoke.
3. Machine-learning-based surrogate models can effectively address the low computational efficiency of FEA in the optimization of the CEPC detector barrel yoke, while satisfying the required prediction accuracy. The technical workflow of "parametric modeling–LHS sampling–multi-model comparative evaluation–surrogate-based optimization–FEA validation" established in this work provides a reusable methodological framework for the rapid optimization of other subdetector structures.

DECLARATIONS

Conflict of interest The authors declare that they have no conflict of interest.
Author contributions Shang Xia: Investigation (equal); Methodology (equal); Writing - original draft (equal); Writing - review & editing (equal). Wei He: Conceptualization (equal); Supervision (equal); Writing - review & editing (equal). Quan Ji: Conceptualization (equal); Writing - review & editing (equal). Fei-Peng Ning: Conceptualization (equal); Supervision (equal); Writing - review & editing (equal). Jin-Fan Chang: Conceptualization (equal); Supervision (equal); Writing - review & editing (equal). Xiao-Yan Ma: Supervision (equal); Writing - review & editing (equal). Xiao-Hui Qian: Supervision (equal); Writing - review & editing (equal). Ya-Tian Pei: Writing - review & editing (equal). Jin-Yu Fu: Writing - review & editing (equal). Shao-Jing Hou: Writing - review & editing (equal). Jian Wang: Writing - review & editing (equal). Zhuo Liu: Writing - review & editing (equal).
Data availability The data that supports the findings of this study are available from the corresponding author

[1] G. Aad, T. Abajyan, B. Abbott *et al.*, Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. Phys. Lett. B **716**, 1–29 (2012). doi: 10.1016/j.physletb.2012.08.020
[2] S. Chatrchyan, V. Khachatryan, A. Sirunyan *et al.*, Observation of a new boson at a mass of 125 GeV with the CMS ex-

periment at the LHC. Phys. Lett. B **716**, 30–61 (2012). doi: 10.1016/j.physletb.2012.08.021

[3] S. Jin, X. C. Lou, M. Q. Ruan *et al.*, The Circular Electron Positron Collider: its physics, technology, and status. PHYSICS **48**(3), 148–158 (2019). doi: 10.7693/wl20190302 (In Chinese)

[4] CEPC Study Group, CEPC Conceptual Design Report: Volume I—Accelerator (2018). doi: 10.48550/arXiv.1809.00285

[5] CEPC Study Group, CEPC Conceptual Design Report: Volume II—Physics & Detector (2018). doi: 10.48550/arXiv.1811.10545

[6] CEPC Study Group, CEPC Technical Design Report—Reference Detector (2025). doi: 10.48550/arXiv.2510.05260

[7] F. P. Ning, CEPC vs FCC: Detector magnet designs for future circular electron–positron colliders. Supercond. **18**, 100260 (2026). doi: 10.1016/j.supcon.2026.100260

[8] Y. F. Wang, *Design, Research and Manufacture of Beijing Spectrometer (BESIII)*, 1st edn. (Shanghai Scientific & Technical Publishers, Shanghai, 2011), pp. 14–17. (In Chinese)

[9] A.I.J. Forrester, A.J. Keane, Recent advances in surrogate-based optimization. Prog. Aerosp. Sci. **45**(1–3), 50–79 (2009). doi: 10.1016/j.paerosci.2008.11.001

[10] Z. H. Han, Kriging surrogate model and its application to design optimization: A review of recent progress. Acta Aeronaut. Astronaut. Sin. **37**(11), 3197–3225 (2016). doi: 10.7527/S1000-6893.2016.0083 (In Chinese)

[11] H. Adeli, Neural networks in civil engineering: 1989–2000. Comput.-Aided Civ. Infrastruct. Eng. **16**(2), 126–142 (2001). doi: 10.1111/0885-9507.00219

[12] S. Funk, A.A. Basmaji, U. Nackenhorst, Globally supported surrogate model based on support vector regression for nonlinear structural engineering applications. Arch. Appl. Mech. **93**, 825–839 (2023). doi: 10.1007/s00419-022-02301-3

[13] S. Xia, Q. Ji, J. S. Zhang *et al.*, Mechanical properties analysis and structural optimization of CEPC detector barrel iron yoke. Radiat. Detect. Technol. Methods **9**, 107–131 (2025). doi: 10.1007/s41605-024-00497-2

[14] M.D. McKay, R.J. Beckman, W.J. Conover, A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics **42**(1), 55–61 (2000). doi: 10.1080/00401706.2000.10485979

[15] Z. Z. Wu, S. G. Luo, *Mechanical Design Manual*, 3rd edn. (Higher Education Press, Beijing, 2006), pp. 26–27. (In Chinese)

[16] X. F. Sun, X. S. Fang, T. L. Guan, *Mechanics of Materials*, 4th edn. (Higher Education Press, Beijing, 2002), pp. 35–37. (In Chinese)

[17] A.E. Hoerl, R.W. Kennard, Ridge regression: Biased estimation for nonorthogonal problems. Technometrics **12**(1), 55–67 (1970). doi: 10.1080/00401706.1970.10488634

[18] A.J. Smola, B. Schölkopf, A tutorial on support vector regression. Stat. Comput. **14**, 199–222 (2004). doi: 10.1023/B:STCO.0000035301.49549.88

[19] T. Chen, C. Guestrin, XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM, 2016), pp. 785–794. doi: 10.1145/2939672.2939785

[20] D. Rumelhart, G. Hinton, R. Williams, Learning representations by back-propagating errors. Nature **323**, 533–536 (1986). doi: 10.1038/323533a0

[21] P. Virtanen, R. Gommers, T.E. Oliphant *et al.*, SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods **17**, 261–272 (2020). doi: 10.1038/s41592-019-0686-2

[22] D. Kraft, A software package for sequential quadratic programming. DFVLR-FB 88–28, Deutsche Forschungs- und Versuchsanstalt für Luft- und Raumfahrt, Köln (1988).