frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 06 September 2023
DOI 10.3389/fenrg.2023.1273760

OPEN ACCESS

EDITED BY
Zhengmao Li,
Aalto University, Finland

REVIEWED BY
Yikui Liu,
Stevens Institute of Technology, United States
Xiaodong Zheng,
Southern Methodist University, United States

*CORRESPONDENCE
Huijun Liang,
lhj@hbmzu.edu.cn

RECEIVED 07 August 2023
ACCEPTED 23 August 2023
PUBLISHED 06 September 2023

CITATION
Lin C, Liang H, Pang A and Zhong J (2023), Data-driven method of solving computationally expensive combined economic/emission dispatch problems in large-scale power systems: an improved kriging-assisted optimization approach. *Front. Energy Res.* 11:1273760. doi: 10.3389/fenrg.2023.1273760

Data-driven method of solving computationally expensive combined economic/emission dispatch problems in large-scale power systems: an improved kriging-assisted optimization approach

Chenhao Lin, Huijun Liang*, Aokang Pang and Jianwei Zhong

College of Intelligent Systems Science and Engineering, Hubei Minzu University, Enshi, China

Combined economic/emission dispatch (CEED) is generally studied using analytical objective functions. However, for large-scale, high-dimension power systems, CEED problems are transformed into computationally expensive CEED (CECEED) problems, for which existing approaches are time-consuming and may not obtain satisfactory solutions. To overcome this problem, a novel data-driven surrogate-assisted method is introduced firstly. The fuel cost and emission objective functions are replaced by improved Kriging-based surrogate models. A new infilling sampling strategy for updating Kriging-based surrogate models online is proposed, which improves their fitting accuracy. Through this way, the evaluation time of the objective functions is significantly reduced. Secondly, the optimization of CECEED is executed by an improved non-dominated sorting genetic algorithm-II (NSGA-II). The above infilling sampling strategy is also used to reduce the number of evaluations for original mathematic fitness functions. To improve their local convergence ability and global search abilities, the individuals that exhibit excellent performance in a single objective are cloned and mutated. Finally, information about the Pareto front is used to guide individuals to search for better solutions. The effectiveness of this optimization method is demonstrated through simulations of IEEE 118-bus test system and IEEE 300-bus test system.

KEYWORDS

combined economic/emission dispatch, data-driven, genetic algorithm, surrogate model, computationally expensive optimization problems

1 Introduction

In power systems, combined economic/emission dispatch (CEED) problem involves the minimizations of cost and emission while meeting load demand and satisfying operation constraints (Xu et al., 2022). It is an optimization problem with strongly real-time requirements (Li et al., 2020). Mathematical methods for solving it include nonlinear and linear programming methods (Lai et al., 2022). Meta-heuristic algorithms are employed to solve it. For example, multi-objective hybrid bat algorithm (MHBA) (Liang

et al., 2018), non-dominated sorting genetic algorithm (NSGA) (Deb et al., 2002), and the improved non-dominated sorting genetic algorithms (Muthuswamy et al., 2015), *etc.*

However, the above approaches are better suitable to solve CEED problems in low-dimension, small-scale power systems. Along with the expansion of the power grid, the complexity of a power system is greatly increased (Wang et al., 2023), which increases the decision-making dimension of CEED problems (Qu et al., 2023). Unfortunately, existing methods become inefficient when applied to computationally expensive combined economic/emission dispatch (CECEED) problems (Zhang and Yu, 2023). Existing methods are often time-consuming for real-time dispatching, and their solutions generally have low accuracies or may not even converge (Li and Xu, 2018). Therefore, for large-scale, high-dimension power systems, traditional CEED problems are transformed into CECEED problems (Li et al., 2022). New method that is suitable for solving CECEED problems is urgently required.

Data-driven surrogate-assisted schemes are widely applied for solving computationally expensive problems by employing surrogate models (Yu et al., 2022). Accurate surrogate model is indispensable in data-driven optimization (Gao et al., 2023). Many kinds of surrogate models are applied in industrial community including support vector regression (SVR) models, artificial neural network (ANN), linear regression models, *etc.* In Ref. (Lin et al., 2023), SVR-based NSGA was applied for solving dispatching problems. Objective functions were replaced by SVR models while reducing the computing time. However, it was powerless for solving high-dimension dispatching problems. In Ref. (Pang et al., 2023), a data-driven bat algorithm was proposed for solving economic dispatch (ED) problems. It allowed the ED problems to be solved, but the computing time was spent still relatively long. Ref. (Linka et al., 2021). proposed a data-driven ANN. Few data were used to build ANN with good performance. Though it reduced computing time, the accuracy was low. Based on the above analysis, the data-driven surrogate-assisted method shows its feasibility and superior performance in solving complex scheduling problems.

Effective meta-heuristic algorithms are also the keys to reducing the execution time. For example, a dynamic crowding distance was added to NSGA-II (Muthuswamy et al., 2015), and prohibited operating zones (POZs) were considered for improving the accuracy. However, it did not incorporate sufficient guidance for searching, and the computing time remained a problem. MHBA (Liang et al., 2018) searched each single dimension for excellent solutions in CEED problems. To reduce computation time, parallel computing was used. However, MHBA could only deal with the CEED problem in terms of improving the performance of the algorithm, and does not take into account the time-consuming objective function of CEED model. Thus, it is important to build a meta-heuristic algorithm for fast and accurate convergence in CECEED problems.

Based on the above analyses, CECEED problems need to be solved at two levels. The first level involves employing accurate surrogate models to replace computationally expensive objective functions. The second level involves enhancing convergence and search ability of a meta-heuristic algorithm. In this paper, a novel data-driven surrogate-assisted NSGA-II is proposed for solving CECEED problems. The approach combines a Kriging surrogate model and an enhanced NSGA-II to reduce the execution time. Firstly, an online data-driven Kriging-based surrogate model is introduced for replacing objective functions. Secondly, the original NSGA-II is improved to enhance its performance and to determine what data points should be used for updating Kriging model. The major outcomes of this paper are.

(i) An improved online Kriging surrogate model is proposed. Mathematic fuel cost and emission functions in CECEED problems are replaced by online Kriging-based surrogate models to reduce evaluation time. Specifically, initial Kriging surrogate model is constructed by considering random candidate solutions. As the optimization proceeds, additional sampling points, selected from the Pareto front, are added into the initial data set which is utilized for updating Kriging surrogate models. By this way, evaluation time, accuracy, and stability of the proposed surrogate model are improved.

(ii) Two improved strategies are applied to enhance convergence and uniformity of Pareto front. Furthermore, more accurate dispatching decisions of CECEED problems can be obtained. Convergence and uniformity of the Pareto front are intuitively reflected in the Euclidean distance between individuals. The convergence is improved by reducing the distance between individuals far from the Pareto front and individuals on the Pareto front. The uniformity is then improved by increasing the Euclidean distance between individuals that are too close to each other.

(iii) Two novel methods for enhancing the global and local search capabilities of NSGA-II are proposed. Furthermore, more diverse dispatching decisions of CECEED problems can be obtained. Firstly, a cloned edge particles search approach is proposed to improve both capabilities. Secondly, because the search performance can be visualized in terms of Pareto front extensibility, a linear combination method based on the position relationship between current population extreme values and Pareto front extreme values is proposed for improving the extensibility.

The remainder of this paper is organized as follows. Section 2 describes traditional CEED problem and CECEED problems. Section 3 describes an online Kriging-based data-driven surrogate model and Kriging-based NSGA-II (K-NSGA-II). Section 4 presents simulation results from the proposed method and existing approaches on IEEE 118-bus and 300-bus test systems. Section 5 summarizes this paper by showcasing its contributions and suggests some research directions for future work.

2 Problems description

This section describes CEED problems and CECEED problems. It is also important to build data-driven surrogate models for CECEED problems (Li et al., 2021). The objective functions and constraints are given as follows.

2.1 Objective functions and constraints

The objective functions (Sheng et al., 2023) of fuel cost and pollutant emission are described as follows. The first objective is minimizing the total cost:

$$C\left(P_{gi}\right) = \sum_{gi=1}^{N_g}\left[a_{gi}P_{gi}^2 + b_{gi}P_{gi} + c_{gi} + \left|d_{gi}\sin\left(e_{gi}\left(P_{gi}^{\min} - P_{gi}\right)\right)\right|\right., \quad (1)$$

where C (P_{gi}) represents the fuel cost function; N_g represents the total number of units; P_{gi} represents the active power of gi^{th} unit; $P_{gi}^{\min}$ represents the minimum active power of gi^{th} unit; a_{gi}, b_{gi}, c_{gi}, d_{gi}, and e_{gi} represent the coefficients of C (P_{gi}). ($a_{gi}P_{gi}^2 + b_{gi}P_{gi} + c_{gi}$) is the quadratic consumption characteristic curve of gi^{th} unit. $|d_{gi}\sin(e_{gi}(P_{gi}^{\min} - P_{gi})|$ represents the valve point effect of gi^{th} unit. When the inlet valve of steam turbine is suddenly opened, the phenomenon of wire drawing will superimpose a pulsation effect on the consumption characteristic curve. This is the valve point effect.

The second objective is minimizing the emission which is caused by fuel burning. Atmospheric pollutants such as sulfur oxides (SO_x) and nitrogen oxides (NO_x) produced by units can be simulated separately. However, for comparison purposes, the total emission of these pollutants (a comprehensive pollution emission model) is the sum of the quadratic and exponential functions (the exponential function provides more accurate representation):

$$E\left(P_{gi}\right) = \sum_{gi=1}^{N_g}\left[10^{-2}\left(\alpha_{gi}P_{gi}^2 + \beta_{gi}P_{gi} + \gamma_{gi}\right) + \varepsilon_{gi}\exp\left(\lambda_{gi}P_{gi}\right)\right], \quad (2)$$

where E (P_{gi}) represents emission function; α_{gi}, β_{gi}, γ_{gi}, ε_{gi}, and λ_{gi} represent coefficients of E (P_{gi}).

The constraints are described as follows.

(1) Active/reactive power constraints

$$P_{gi}^{\min} \le P_{gi} \le P_{gi}^{\max} \quad (3)$$

where $P_{gi}^{\min}$ and $P_{gi}^{\max}$ are the minimum and maximum active power of gi^{th} unit, respectively.

$$Q_{gi}^{\min} \le Q_{gi} \le Q_{gi}^{\max} \quad (4)$$

where $Q_{gi}^{\min}$ and $Q_{gi}^{\max}$ are the minimum and maximum reactive power of gi^{th} unit, respectively.

(2) Power balance constraints and load flow calculation

The power balance can be satisfied with load flow calculation:

$$P_{gi} - P_{dgi} - V_{gi}\sum_{gj=1}^{N_{\text{buses}}} V_{gj}\left(G_{gi,gj}\cos\theta_{gi,gj} + B_{gi,gj}\sin\theta_{gi,gj}\right) = 0, \quad (5)$$

where P_{dgi} represents the power load at gi^{th} bus; V_{gi} represents the voltage at gi^{th} bus; V_{gj} represents the voltage at gj^{th} bus; $\theta_{gi,gj} = \theta_{gi} - \theta_{gj}$ (θ_{gi} and θ_{gj} represent the voltage angles of gi^{th} and gj^{th} buses, respectively.); $G_{gi,gj}$ is transfer conductance within gi^{th} and gj^{th} buses; $B_{gi,gj}$ is transfer susceptance within gi^{th} and gj^{th} buses; N_{buses} represents that total N_{buses} buses are contained in a power system.

$$Q_{gi} - Q_{dgi} - V_{gi}\sum_{gj=1}^{N_{\text{buses}}} V_{gj}\left(G_{gi,gj}\sin\theta_{gi,gj} - B_{gi,gj}\cos\theta_{gi,gj}\right) = 0, \quad (6)$$

where Q_{dgi} represents the reactive load at gi^{th} bus.

(3) Voltage constraints

$$V_{gi}^{\min} \le V_{gi} \le V_{gi}^{\max} \quad (7)$$

where $V_{gi}^{\min}$, $V_{gi}^{\max}$, and V_{gi} represent the minimum, maximum, and current voltage of gi^{th} bus, respectively.

(4) Line flow constraints

$$S_l \le S_l^{\max}, l = 1, \ldots, N_{\text{lines}}, \quad (8)$$

where S_l and $S_l^{\max}$ represent line flow and upper limit line flow of l^{th} transmission line.

(5) Ramp rate constraints

$$\begin{cases} P_{gi} - P_{gi}^0 \le UR_{gi}, \\ P_{gi}^0 - P_{gi} \le DR_{gi}, \end{cases} \quad (9)$$

where P_{gi}^0 represents previous active power of gi^{th} unit; UR_{gi} and DR_{gi} represent up and down ramp rate of gi^{th} unit, respectively.

(6) Prohibited operating zones (POZs) constraints

$$\begin{cases} P_{gi}^{\min} \le P_{gi} \le P_{gi,1}^{\text{l}} \\ P_{gi,gk-1}^{\text{u}} \le P_{gi} \le P_{gi,gk}^{\text{l}}, gk = 2, \ldots, K_{gi}, \\ P_{gi,K_{gi}-1}^{\text{u}} \le P_{gi} \le P_{gi}^{\max} \end{cases} \quad (10)$$

where K_{gi} represents the number of prohibited operation zones of gi^{th} unit; $P_{gi,gk-1}^{\text{u}}$ and $P_{gi,gk}^{\text{l}}$ represent upper and lower limit of k^{th} POZ.

2.2 Descriptions of CECEED problems

Computationally expensive optimization problem refers to a class of optimization problems that require expensive or even unaffordable costs when evaluating alternative solutions. This class of problems exists widely in many importantly practical application scenarios. On one hand, "computationally expensive" means that the evaluation itself needs to consume a lot of time or other expensive costs. On the other hand, it also includes some problems can be transferred to computationally expensive optimization problem within a time-sensitive scene.

As to CECEED problems, on one hand, the computationally expensive challenge is caused by the high-dimension decision variables. On the other hand, due to the time-sensitive dispatching cycle, the CEED problems become computationally expensive. Thus, it is urgent to find a suitable method for solving CECEED problems quickly and accurately.

To reduce execution time of CECEED problems, surrogate-assisted technology can be employed. Original objective functions are replaced by trained surrogate models. The replacements of objective functions reduce evaluation time.

3 Description of Kriging-assisted optimization approach

This section describes the Kriging-assisted optimization approach for CECEED problems. Kriging models are constructed

for replacing objective functions. To enhance accuracy of surrogate model, a novel infilling sampling strategy is proposed to update Kriging-based model online. The proposed K-NSGA-II method is employed to execute the optimization of CECEED. As the iterations proceed, the selected candidate solutions from the Pareto front are used not only for the actual evaluation, but also for the updating of the Kriging-based surrogate model.

3.1 Kriging-based surrogate model

3.1.1 Modeling of surrogate model

Existing approaches for solving CECEED problems usually require long time (Sharifian and Abdi, 2023), which is too long for scheduling cycle. Some methods suffer from "curse of dimensionality", which results in non-convergence. To overcome this drawback, Kriging model is improved for replacing original objective functions with short evaluation time (Li et al., 2022).

The input data of Kriging surrogate model is a certain candidate solution (x), where x represents the power generating capacities of N_g units ($P_1, ..., P_{gi}, ..., P_{N_g}$). The output value of Kriging surrogate model is the total fuel cost ($C^*(x)$) or emission ($E^*(x)$) when the power generating capacities are $P_1, ..., P_{gi}, ..., P_{N_g}$. The modeling method of Kriging-based surrogate model can be described as follows. For training a Kriging-based surrogate, the initial data set contains total n kinds of different dispatching decisions $x_1, x_2, \ldots, x_n$. Taking the fuel cost function as an example, the corresponding labels are $C(x_1), C(x_2), \ldots, C(x_n)$, where $C(x_1), C(x_2), \ldots, C(x_n)$ are real fuel cost values corresponding to $x_1, x_2, \ldots, x_n$. Kriging model aims to compute the prediction value of unknown points through known points. In other words, Kriging model is an interpolation model whose expected result is a linearly weighted sum of known points. For another candidate solution (dispatching decision) x, the expected fuel cost value ($C^*(x)$) can be obtained by:

$$C^*(x) = \sum_{k=1}^{n} \lambda_k b(x) + Z(x), \tag{11}$$

where $C^*(x)$ is the expected fuel cost value (obtained by Kriging surrogate model) of prediction point x; $b(x)$ represents the basic function; λ_k is the unknown weight coefficient of the k-th basic function; $Z(x)$ is the random error whose mean value is zero. According to Eq. 11, modeling the Kriging-based surrogate model is equal to determining the weight coefficients $(\lambda_1, ..., \lambda_k, ..., \lambda_n)$ (Liu et al., 2022). One-order universal Kriging models are used to fitting the objective functions in CECEED problems. The Kriging-based surrogate model satisfy:

$$\sum_{k=1}^{n} \lambda_k \mathrm{Cov}\left[C^*(x_k), C^*\left(x_j\right)\right] + \delta b\left(x_j\right) = \sum_{k=1}^{n} \mathrm{Cov}\left[C^*(x_k), C^*\left(x_j\right)\right], \tag{12}$$

where δ is a Lagrange multiplier.

$$\mathrm{Cov}\left[C^*(x_k), C^*\left(x_j\right)\right] = \sigma^2 \mathbf{R}\left[R\left(x_k, x_j\right)\right], \tag{13}$$

where σ^2 represents the variance of $C^*(x)$, $\mathbf{R}[\cdot]$ represents a $n \times n$ symmetric correlation matrix; $R(x_k, x_j)$ represents the spatial correlation function of x_k and x_j which can be described as Gaussian covariance function:

$$R\left(x_k, x_j\right) = \exp\left[-\sum_{gi=1}^{N_g} \theta_{gi}\left(x_{k,gi} - x_{j,gi}\right)^2\right], \tag{14}$$

where N_g is the dimension of x (equal to the number of units); θ_{gi} represents the gi^{th} parameter of covariance function. The θ_{gi} is obtained by the maximum likelihood estimation. Then, all λ_k in Eq. 12 can be obtained. Finally, the objective value of the expected point x can be obtained by:

$$C^*(x) = \sum_{k=1}^{n} \lambda_k C^*(k), \tag{15}$$

Compared with original objective functions, Kriging-based surrogate models significantly reduce the evaluation time. Different original objective functions (fuel cost function without valve point effect, fuel cost function with valve point effect, and emission function) for the CECEED in IEEE 118-bus test system are chosen for verifying the effectiveness of Kriging-based surrogate models. Total runtimes (sum of 50 runtimes) of original functions are 0.001982 s, 0.001878 s, and 0.003166 s, respectively. Total runtimes (sum of 50 runtimes) of surrogate models are 0.000792 s, 0.000458 s, and 0.002062 s, respectively. The Kriging model reduce the evaluation time of three objective functions by 60.04%, 59.64%, and 34.87%, respectively. This demonstrates that Kriging model greatly reduces the evaluation time of objectives.

3.1.2 Update of Kriging-based surrogate model

Due to the remarkable fluctuation of data, the accuracy of the Kriging model might not be sufficient (Qian et al., 2023). Kriging model is established on the principle that expected value is similar to that of nearby points. This principle leads to inaccurate predictions of volatile points by the Kriging-based surrogate model. To overcome this drawback, a novel infilling sampling strategy is proposed. The schematic diagram of the proposed infilling sampling strategy is shown as Figure 1A. As the iteration goes on, Pareto front is constantly updated. Some points (circular marks in Figure 1A) with better convergence and diversity occur. However, these points are not included in the data set of the trained Kriging surrogate model. This results in a decrease in the accuracy of the trained surrogate model (The fitting accuracy near the circular marks is lower than that near the fork marks). Thus, these points with circular marks can be added to the data set for updating the Kriging surrogate model.

Figure 1A Schematic diagram of the proposed Figure 1B. Infilling sampling strategy of Kriging infilling sampling strategy. model in IEEE 118-bus test system.

To further illustrate the effectiveness of the proposed infilling sampling strategy, a CECEED optimization of IEEE 118-bus system is taken as an example. The process of adding points in this strategy is shown in Figure 1B. This strategy enhances the accuracy degree of Kriging model. The inspiration comes from changes in the Pareto front. As the iteration proceed, Pareto front is continuously updated. Assume that there are 100 iterations, and the current Pareto front is recorded every 20 iterations. Candidate solutions, which are inconsistent with the last recorded Pareto front, are selected from the current Pareto front. They are then computed by the original objective functions. This means that the real fitness values are generated, representing the latest optimization information

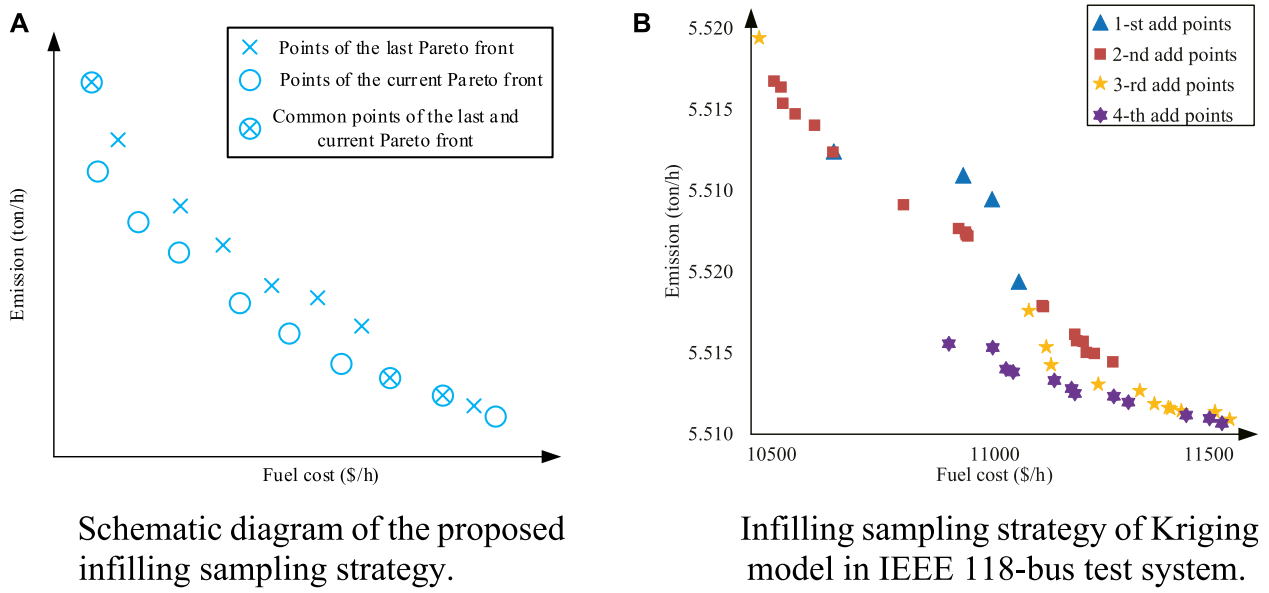

FIGURE 1
The proposed infilling sampling strategy. **(A)** Schematic diagram of the proposed infilling sampling strategy. **(B)** Infilling sampling strategy of Kriging model in IEEE 118-bus test system.

(sampling). The selected points are added to update Kriging-based surrogate models (infilling).

The infilling sampling strategy is executed four times during the optimization process, with 4, 19, 12, and 12 new data points added, respectively. Generally, the added points have a higher degree of convergence than the existing ones. This means that the current Pareto front is closer to the real one. The principle of Kriging models is that things that are closer to each other are more similar than things that are farther away from each other. Thus, the principle and proposed strategy are essentially consistent. More importantly, the infilling sampling strategy guides direction of optimization, as shown in Figure 1A. The accuracy of Kriging model is guaranteed. Figures 2A–D compare Kriging-based surrogate models with or without updating in terms of the mean absolute percentage error (MAPE) and mean square error (MSE). The number of iterations is 100 and the results are averaged over 50 independent tests. Because update strategy is executed every 20 iterations, average MAPE and MSE of 20 generations are selected for comparison and shown in Table 1.

Figure 2A MAPE of Kriging model (emission). Figure 2B. MAPE of Kriging model (fuel cost). Figure 2C. MSE of Kriging model (fuel cost). Figure 2D. MSE of Kriging model (emission).

For the updated Kriging-based surrogate model, there are two levels of improvement. The four updates improve the MAPE by 23.73%, 9.75%, 2.75%, and 0.45%, respectively. This illustrates that the accuracy of Kriging-based surrogate model improves after each update and that the infilling sampling strategy is effective. Compared with the original model (Kriging-based surrogate model without updating), the improvements in the MAPE are 3.93%, 16.15%, 16.42%, and 12.97% in the four updates, respectively. This illustrates that the infilling sampling strategy contributes to better model accuracy.

As for emission function, the situation is similar. The four updates improve the MAPE by 0.55%, 0.03%, −0.16%, and 0.19%, respectively. Although there is a slight decline in accuracy at the third update, the other updates improve the accuracy. Compared with original model, the MAPE improves by 3.93%, 16.15%, 16.42%, and 12.97% over the four updates, respectively. Thus, this strategy enhances the performance of Kriging-based surrogate model.

3.2 Introduction of K-NSGA-II

To solve CECEED problems, NSGA-II suffers from poor convergence (Wei et al., 2022), uniformity, search ability, and extensibility (Chen et al., 2023). To overcome these drawbacks, optimization strategies are applied to Pareto front and search pattern to enhance performance. The convergence, uniformity, extensibility, and search ability of proposed K-NSGA-II are enhanced through the proposed optimization strategies. Additionally, the usage of Kriging surrogate models reduces evaluation time. The flow tree of K-NSGA-II process is described in Figure 3. It can be seen that original NSGA-II flowchart is shown in the black solid frame and the improvements of K-NSGA-II are shown in red dashed box. The evaluations of objective functions are replaced by the evaluations of Kriging surrogate functions. Two kinds of optimization strategies are added to original NSGA-II. The optimization of Pareto front, the improvement of search ability, and the simulation results are described in Section 3.2.1, Section 3.2.2, and Section 3.2.3, respectively.

3.2.1 Enhanced convergence and uniformity

The application of NSGA-II to CECEED problems would result in insufficient convergence and uniformity (Li et al., 2022). The variant operators in NSGA-II are liable to produce precocious genes in early stages. Furthermore, these genes may constitute a high proportion. This leads to some decreases in the convergence and uniformity. To overcome these drawbacks, two optimal strategies are applied to enhance convergence and uniformity, respectively.

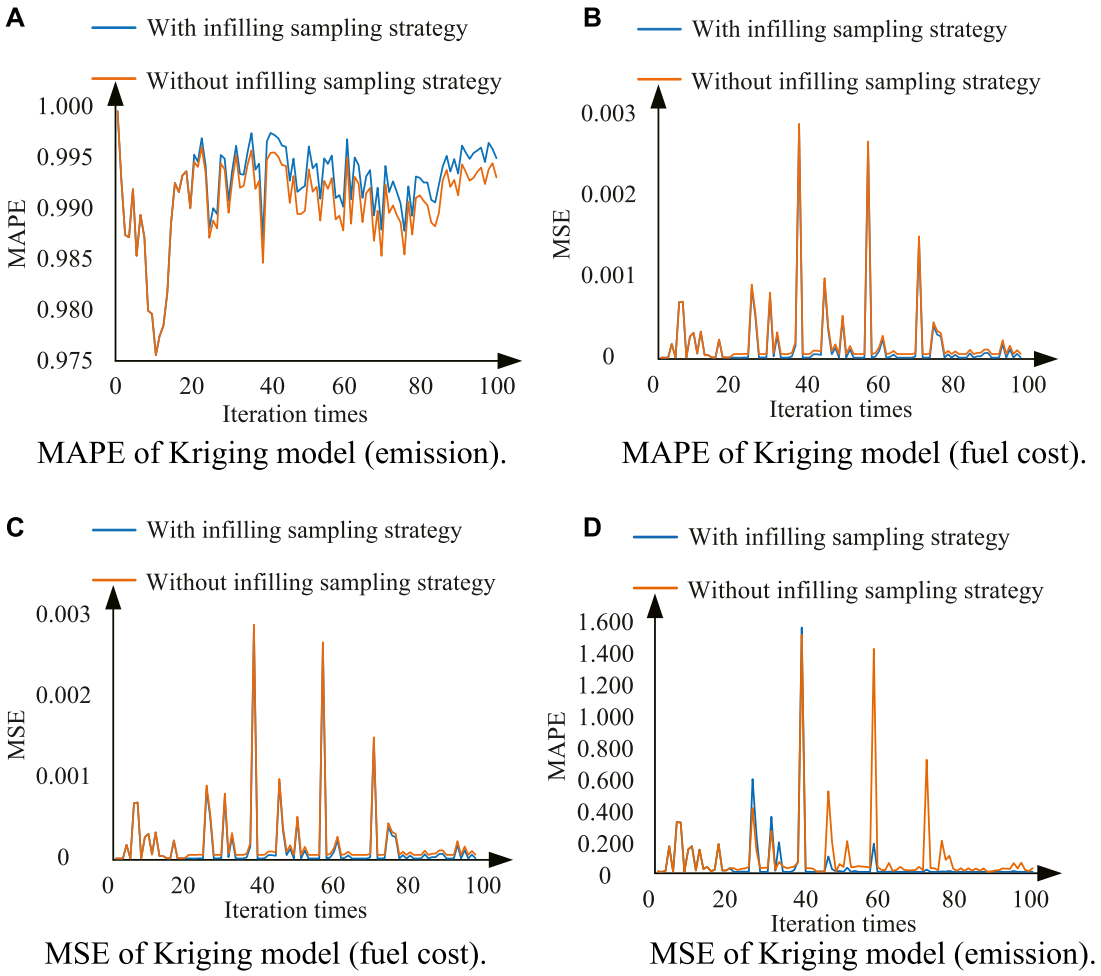

FIGURE 2
Accuracy comparison of Kriging model. **(A)** MAPE of Kriging model (emission). **(B)** MAPE of Kriging model (fuel cost). **(C)** MSE of Kriging model (fuel cost). **(D)** MSE of Kriging model (emission).

TABLE 1 MAPE and MSE comparisons of Kriging-based surrogate model.

Mean value of 20 iterations	Original fuel cost model		Updated fuel cost model		Original emission model		Updated emission model	
Index	MAPE	MSE	MAPE	MSE	MAPE	MSE	MAPE	MSE
1–20 iterations	0.7049	8.13E-02	0.7049	8.13E-02	0.9890	1.58E-04	0.9890	1.58E-04
21–40 iterations	0.8392	1.46E-01	0.8722	1.58E-01	0.9933	3.54E-04	0.9944	3.10E-04
41–60 iterations	0.8241	1.41E-01	0.9572	1.78E-02	0.9929	3.32E-04	0.9947	2.73E-04
61–80 iterations	0.8448	7.26E-02	0.9835	1.50E-03	0.9914	2.31E-04	0.9931	1.60E-04
81–100 iterations	0.8745	1.92E-02	0.9879	1.10E-03	0.9933	3.73E-05	0.9950	3.17E-05

3.2.1.1 Strategy to improve convergence

S1: For each individual $x_i, i = 1, \ldots, N$, where N is the population number in the offspring set, the fitness values $y_{v,i}, i = 1, \ldots, N$ are computed by Kriging model.

S2: For each $y_{v,i}$, the nearest solution $f_{v,i}$ on P_v is identified (where P_v is the current Pareto front).

S3: The distances $d_{1,i}, i = 1, \ldots, N$, between every $y_{v,i}$ and the nearest $f_{v,i}$ are computed.

S4: The average $d_{\text{mean},1}$ of all distances is computed.

S5: If $d_{1,i}$ is greater than or equal to $d_{\text{mean},1}$, x_i is replaced by $x_i^{\text{new},1}$ according to:

$$\begin{cases} R_0 \in [1, 10], \\ x_i^{\text{new},1}(j) = e(j) + R_0 e(j)\Phi, \end{cases} \tag{16}$$

where R_0 is a random constant between 1 and 10, Φ is a random variable that obeys the normal distribution, j ranges from 1 to V

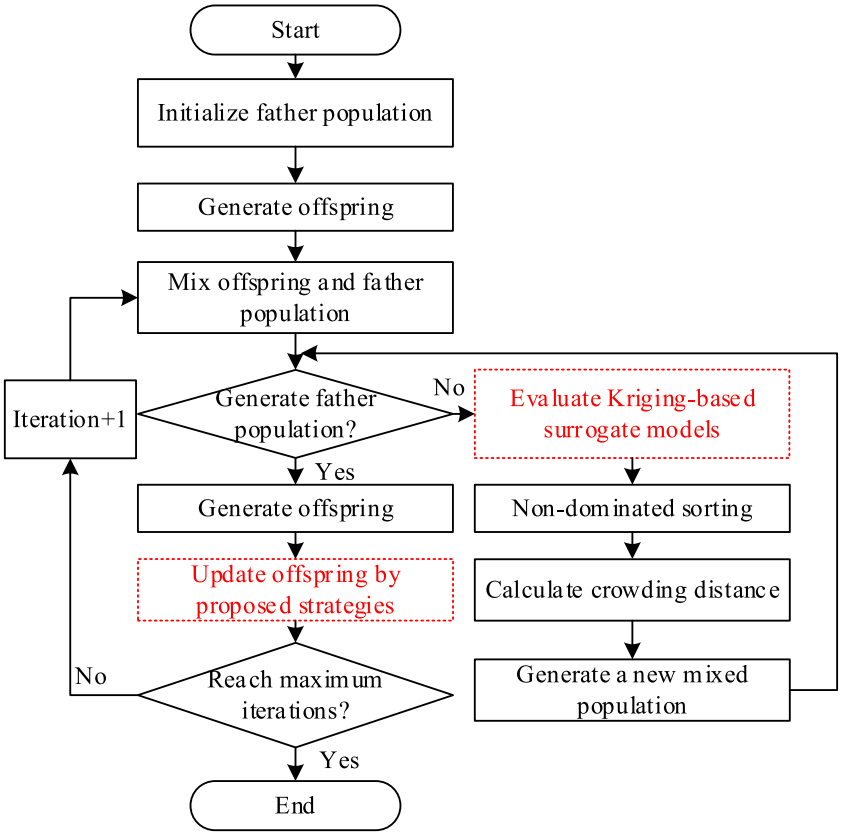

FIGURE 3
Flowchart of K-NSGA-II.

(dimension of individual), and $e(j)$ represents the solution corresponding to fitness value $f_{v,i}$. As shown in Figure 4A, this strategy improves convergence by replacing points far from the Pareto front with points near the Pareto front.

3.2.1.2 Strategy to improve uniformity

S1: For each individual $x_i, i = 1, \ldots, N$, the fitness values $y_{v,i}, i = 1, \ldots, N$, are computed by the surrogate models.

S2: For each $y_{v,i}$, the nearest corresponding fitness value $y_{v,k}, k \neq i$, is identified in the offspring set.

S3: The distances $d_{2,i}, i = 1, \ldots, N$, between $y_{v,i}$ and corresponding $y_{v,k}$ are computed.

S4: The average value $d_{2,\text{mean}}$ of all distances is computed and compared with every $d_{2,i}$.

S5: If $d_{2,i}$ is less than or equal to $d_{2,\text{mean}}$, x_i is replaced by $x_i^{\text{new},2}$ according to:

$$\begin{cases} R_1 \in [1,2], \\ x_i^{\text{new},2}(j) = x_i(j) - R_1(b_1 - b_2)[x_i(j) - x_k(j)], \end{cases} \tag{17}$$

where R_1 is a random constant between 1 and 2, b_1 and b_2 are random variables between 0 and 1 that obey the average distribution, and $j = 1, \ldots, V$ is the dimension of the variable. The schematic diagram is shown in Figure 4B. This strategy improves uniformity by moving particles that are close together further away from each other.

3.2.2 Improving the search ability of NSGA-II

NSGA-II easily become trapped around a local optimal solution (Yadav et al., 2022). Search directions toward better values are retained with high probability, and search directions toward poor values are retained with a lower probability. In the case of a local optimal value, the probability that NSGA-II will move in other directions is extremely low, so the premature convergence appears. To overcome this drawback, two optimization strategies are applied to improve the search ability of NSGA-II.

Figure 4A Strategy to improve convergence. Figure 4B. Strategy to improve uniformity. Figure 4C Strategy to improve extensibility. Figure 4D. Strategy to improve search capability.

3.2.2.1 Strategy to improve uniformity

The search ability can be visualized in terms of Pareto front extensibility. Thus, a linear combination method based on positional relationship between the current population extremes and current Pareto front extremes is used to enhance extensibility. The method is shown as Figure 4C and described as follows:

S1: For each individual $x_i, i = 1, \ldots, N$, the fitness values $y_{v,i}, i = 1, \ldots, N$, are computed by the Kriging surrogate models.

S2: For the fitness value of each objective function, the maximum value is $y_{a,\max}, a \in [1, M]$, where M is the number of objectives.

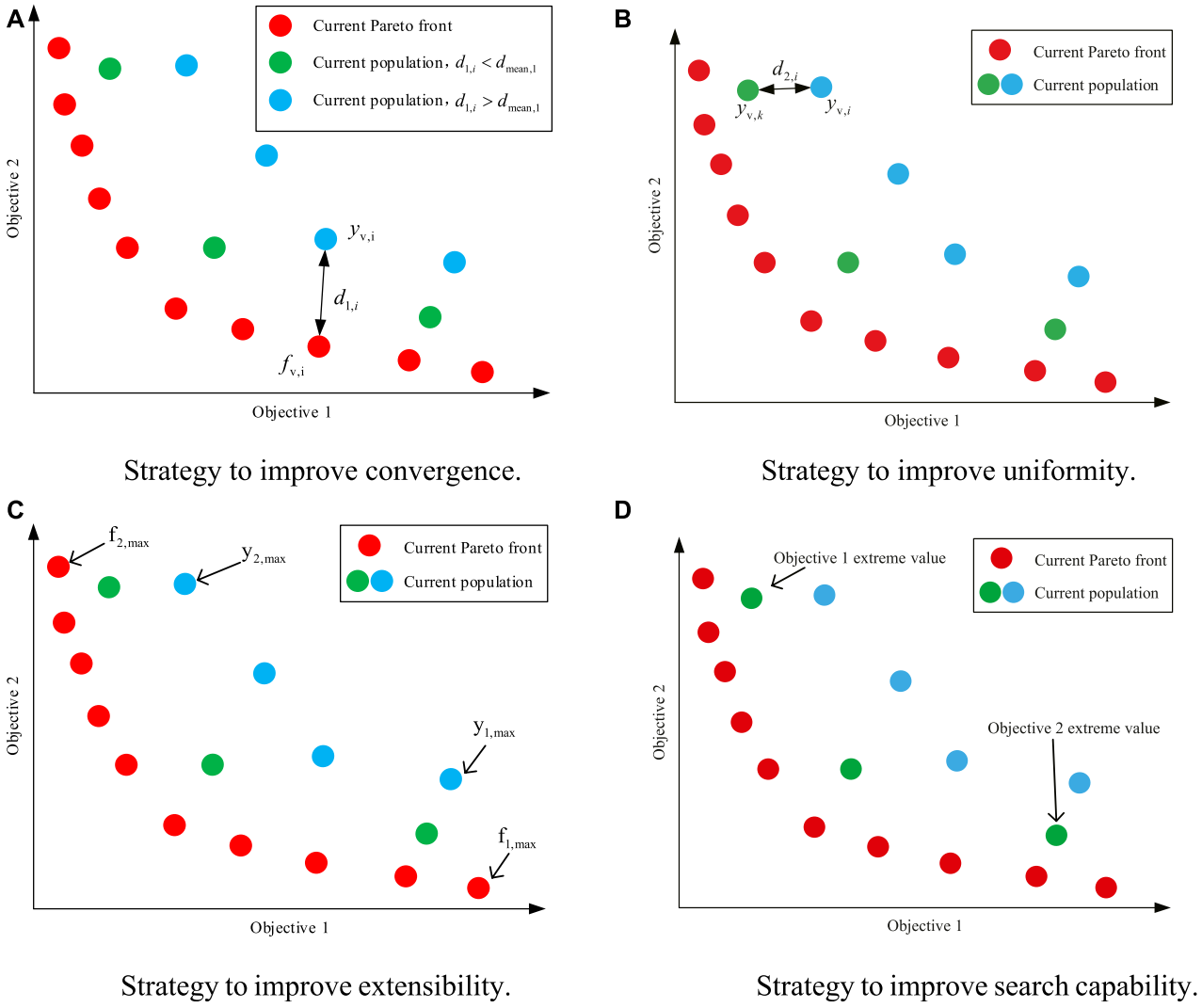

FIGURE 4
Optimization strategies for improving NSGA-II. **(A)** Strategy to improve convergence. **(B)** Strategy to improve uniformity. **(C)** Strategy to improve extensibility. **(D)** Strategy to improve search capability.

S3: For the fitness value of the current P_v, the maximum $f_{a,max}, a \in [1, M]$, of the function in P_v is recorded and the corresponding candidate solution t_a ($a \in [1, M]$) is determined.

S4: For every $y_{a,max}$, the corresponding candidate solution x_a ($a \in [1, M]$) is updated according to:

$$x_a^{new} = x_a + c_1 t_a, \tag{18}$$

where c_1 is a random variable between 0 and 0.5 that obey the average distribution. This strategy improves extensibility of NSGA-II by calculating the vector sum of the scaled population extremum candidate solution and the Pareto front candidate solution. In other words, the strategy mixes all the information of the population extremal candidate solution and part of the information of the Pareto front extremal candidate solution.

3.2.2.2 Mutation strategy to improve search ability

As stated above, NSGA-II is prone to become trapped around local optimal solutions. As another means of overcoming this drawback, a new local and global optimization method based on cloned individuals is proposed. In K-NSGA-II, genes that represent extreme values for a single objective are cloned. A schematic diagram is shown in Figure 4D. Two variation distribution indices are used. The first enables genes to jump away from local optima, and the second is used to improve convergence. The K-NSGA-II with this strategy has stronger ability to jump out of local optimum and enhance global search ability.

3.2.3 Testing of benchmark functions

Three benchmark test functions (ZDT1, 2, and 3) are utilized for illustrating the improvements of K-NSGA-II. The population is set to 40, the number of iterations is 200, cross-distribution

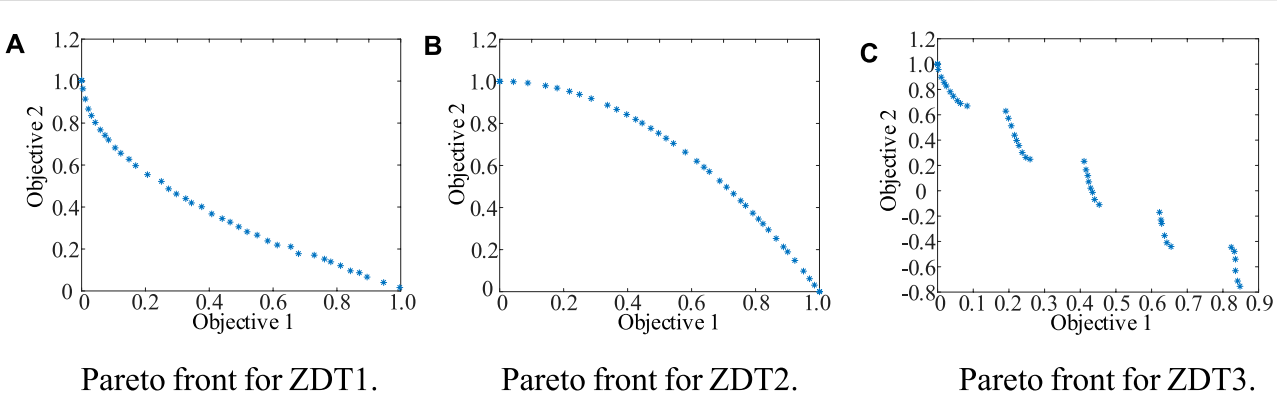

FIGURE 5
Pareto fronts for ZDT1, 2, and 3. **(A)** Pareto front for ZDT1. **(B)** Pareto front for ZDT2. **(C)** Pareto front for ZDT3.

TABLE 2 Statistical results of K-NSGA-II (20 tests).

Algorithm	Index	ZDT1	ZDT2	ZDT3
K-NSGA-II	mean convergence	**0.014340**	**0.001949**	**0.012616**
NSGA-II	mean convergence	0.033482	0.072391	0.114500
K-NSGA-II	MSD of convergence	0.002464	0.000080	**0.003271**
NSGA-II	MSD of convergence	0.004750	0.031689	0.007940
K-NSGA-II	mean diversity	**0.158623**	0.207601	**0.552062**
NSGA-II	mean diversity	0.390307	0.430776	0.738540
K-NSGA-II	MSD of diversity	0.010147	0.014684	0.065412
NSGA-II	MSD of diversity	0.001876	0.004721	0.019706
K-NSGA-II	mean of the worst diversity	**0.186958**	**0.246197**	**0.736580**
NSGA-II	mean of the worst diversity	0.390307	0.430776	0.738540

The bold values are the best comparison results.

index is 20, variation distribution index is 100, and mutation probability is 1/30; benchmark functions are 30- dimensional. For the cloning strategy, variation distribution index for global and local search are 30 and 180, respectively. The results are shown in Figures 5A–C. Performance statistics from the original and improved algorithms are compared in Table 2.

Compared with NSGA-II, the mean convergence of K-NSGA-II on the three benchmark functions improves by 57.17%, 97.31%, and 88.98%, respectively. The mean square deviations (MSD) are improved by 48.13%, 99.75%, and 58.80%, respectively. These improvements illustrate that K-NSGA-II performs much better than NSGA-II of convergence and stability. The mean diversity of K-NSGA-II on benchmark functions has improved by 59.36%, 81.81%, and 25.25%, respectively. It indicates that K-NSGA-II outperforms NSGA-II of diversity.

Although the mean square deviation of diversity is less than that for NSGA-II, the worst diversity over the 20 tests is better than that of NSGA-II, as presented in Table 2. Compared with NSGA-II, the worst diversity score for each benchmark function improves by 52.10%, 42.85%, and 2.65%, respectively. Overall, K-NSGA-II is able to execute multi-objective optimization effectively.

Figure 5A Pareto front for ZDT1. Figure 5B. Pareto front for ZDT2. Figure 5C. Pareto front for ZDT3.

4 Examples and results

In this section, three simulation cases using IEEE 118-bus and IEEE 300-bus systems are considered to illustrate the effectiveness of the proposed method. For emission, cost, power, and time, the units in following tables are ton/h, $/h, MW, and s, respectively. The population number of K-NSGA-II is 50, the number of repetitions is 20, cross-distribution index is 20, variance distribution index is 100, and variation rate is the reciprocal of the units. The given execution time in the following tables contains the update and evolution time of Kriging surrogate models. Additionally, modified NSGA-II (M-NSGA-II) (Muthuswamy et al., 2015), MHBA (Liang et al., 2018), NSGA-II (Deb et al., 2002), and real coded genetic algorithm (RCGA) (Muthuswamy et al., 2015) are chosen for comparison.

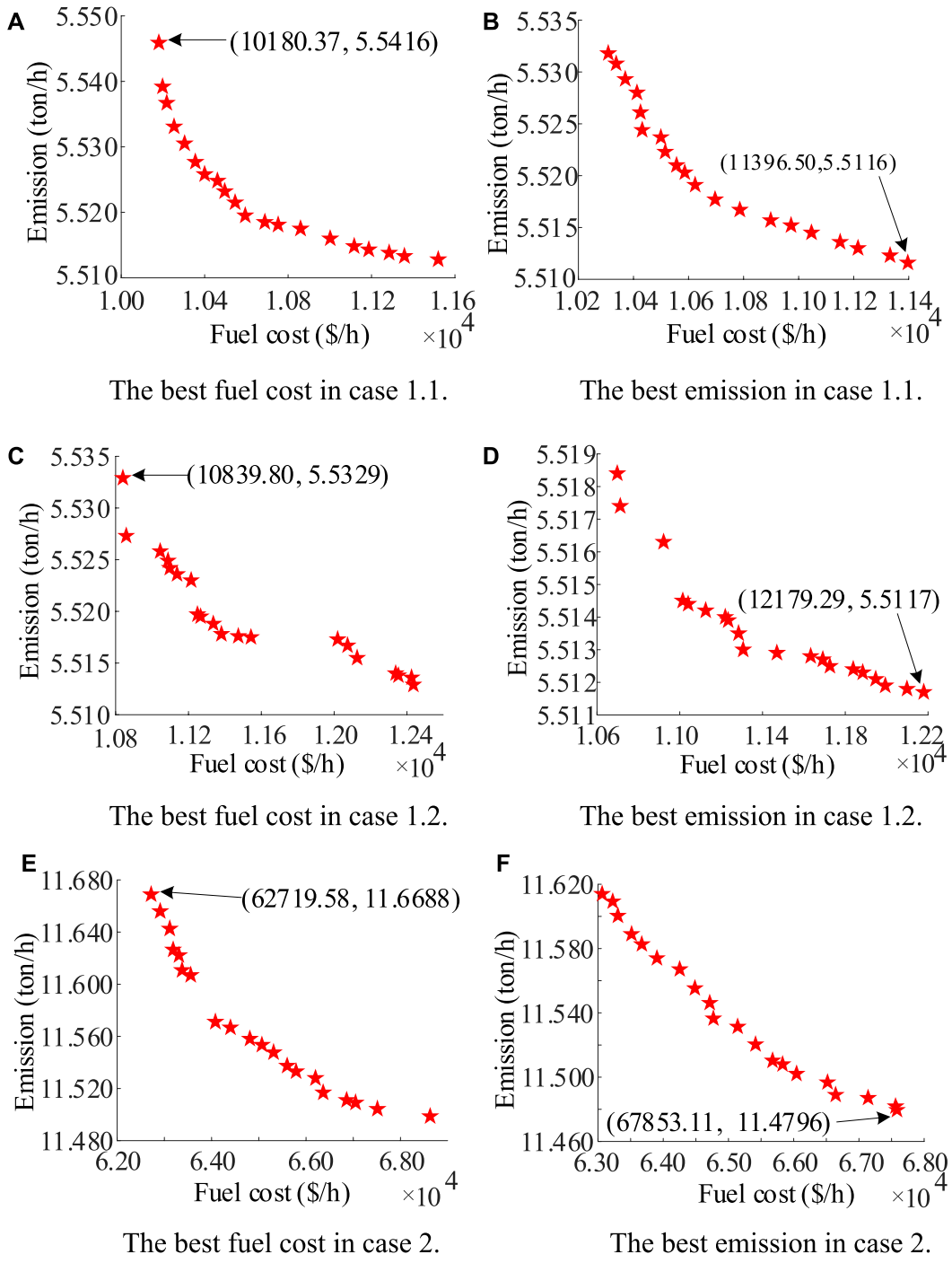

FIGURE 6
Simulation results (Pareto fronts) simulation cases. **(A)** The best fuel cost in case 1.1. **(B)** The best emission in case 1.1. **(C)** The best fuel cost in case 1.2. **(D)** The best emission in case 1.2. **(E)** The best fuel cost in case 2. **(F)** The best emission in case 2.

4.1 Case 1: IEEE 118-bus test system

Total load demand is set as 3668 MW. The parameters of fuel cost and emission come from reference (Paul et al., 2022). The simulations of this test system are split into two different simulation cases according to the restrictions applied: (a) All constraints mentioned above are included, although valve point effects and POZs are neglected. (b) All constraints are considered.

4.1.1 Case 1.1: Without POZs and valve point

Total 20 solutions are produced on the Pareto front, and they are given in Figures 6A, B. When minimizing fuel cost, the solution (fuel cost, emission) is (10180.37, 5.5416), and when minimizing emission, the solution is (11396.50, 5.5116). Comparisons with other algorithms are shown in Table 3. The minimum value of cost is 10180.37 $/h, and it is less than the fuel costs given by other four algorithms. The runtime of K-NSGA-II

TABLE 3 Best solutions in terms of the best fuel cost and emission for case 1.1.

Items	K-NSGA-II	MHBA	M-NSGA-II	NAGA-II	RCGA	K-NSGA-II	MHBA	M-NSGA-II	NAGA-II	RCGA
	In terms of the best fuel cost					In terms of the best emission				
P_1	816.51	787.58	611.29	640.18	673.52	314.73	299.12	330.36	314.74	310.33
P_2	10.00	499.45	62.23	54.08	73.06	12.89	482.28	436.53	396.63	425.52
P_3	62.14	10.00	89.07	83.84	76.07	10.23	10.00	89.83	82.98	89.99
P_4	214.79	30.00	298.65	285.21	299.99	41.18	30.00	299.49	299.49	299.99
P_5	1.04	231.01	40.07	40.49	40.00	1.00	40.00	394.78	395.52	399.99
P_6	3.00	1.00	5.32	1.44	1.02	3.42	1.00	5.68	8.83	1.66
P_7	61.25	3.00	9.50	12.74	17.65	239.67	3.22	8.78	22.66	18.70
P_8	5.00	30.00	31.86	30.03	30.00	5.26	240.00	239.60	236.20	239.99
P_9	20.00	5.00	41.87	48.72	10.39	136.11	5.00	40.01	49.48	49.98
P_{10}	21.30	27.81	195.86	151.96	138.69	194.86	146.19	197.99	199.66	199.95
P_{11}	399.07	20.00	197.05	190.13	199.99	399.95	200.00	194.05	198.92	199.99
P_{12}	396.26	399.21	395.26	394.83	399.59	399.84	399.95	303.40	340.86	291.42
P_{13}	448.95	398.85	398.51	397.63	399.95	499.96	400.00	385.21	386.81	399.83
P_{14}	552.68	598.30	582.46	590.54	599.99	600.00	599.86	157.57	174.59	167.68
P_{15}	1.00	1.00	3.02	3.99	3.68	1.00	1.00	3.78	4.89	1.64
P_{16}	598.57	625.27	658.53	672.04	690.62	699.97	699.49	310.37	303.70	307.75
P_{17}	204.92	152.11	243.36	240.91	228.31	298.71	300.00	291.66	276.89	292.74
P_{18}	5.01	5.00	10.79	36.10	5.31	5.26	5.00	46.15	38.28	49.98
P_{19}	4.01	4.00	5.70	6.96	8.30	4.03	5.01	39.74	39.80	26.84
Losses	157.49	160.58	212.5	213.90	228.50	200.07	199.12	107.1	103	106.0
Cost	10180.37	10186.8	11552.0	11577.5	11509.7	11396.50	11433.0	18311.0	17993.4	18227.5
Emission	5.5416	5.5417	13.7960	14.1910	14.9726	5.5116	5.5117	5.5466	5.4950	5.4876
Time	34.751	104.797	825.875	838.171	31806.1	34.939	104.797	825.875	838.171	31806.1

It is notable that the other algorithms require 20000, 50000, 30000, and 50000 evaluations, respectively, whereas the K-NSGA-II, requires 189 evaluations. Significant time gap and fewer evaluations illustrate the advantages of K-NSGA-II, for CECEED, optimization. The minimum emission achieved by K-NSGA-II, is 5.5116 ton/h. Compared with other algorithms, time is reduced by 66.66%, 95.77%, 95.83% and 99.89%, respectively. K-NSGA-II, requires only 176 evaluations. Thus, K-NSGA-II, is accurate and time-saving for solving CECEED, problems.

is significantly lower than that of the other algorithms. Compared with other four algorithms, the runtime is reduced by 66.84%, 95.81%, 95.85%, and 99.89%, respectively. Further, let us focus on the saved time for evolutions of objective functions. In the optimization process (in terms of the best cost), the surrogate-based evolutions are executed 5,000 (50 × 100 = 5,000) times. For executing two strategies in 3.2.1.1 and 3.2.1.2, due to the number of additional evolutions (based on Kriging surrogate model) is uncertain, the total number of additional evolutions is 9,857. For executing the strategy in 3.2.2.1, the number of additional evolutions is 40000 (400 × 100 = 40000). As for the strategy in 3.2.2.2, the number of additional evolutions is also 40000 (400 × 100 = 40000). Then, the total saved time of evolutions can be computed (94857×(0.001982–0.000792)/50)+94857×(0.003166–0.002062)/50) = 4.35 s). The saved time accounts for 12.52% of the total running time. When obtaining the best emission, the saved time is 4.21 s. It accounts for 12.05% of the total running time. Thus, the proposed Kriging-based surrogate model is time-saving and effective.

Figure 6A. The best fuel cost in case 1.1. Figure 6B. The best emission in case 1.1. Figure 6C. The best fuel cost in case 1.2. Figure 6D. The best emission in case 1.2. Figure 6E. The best fuel cost in case 2. Figure 6F. The best emission in case 2.

Let us see what performance the proposed K-NSGA-II can obtain. Firstly, better convergence and diversity are obtained. Thanks to the K-NSGA-II, better diversity is the reason of more powerful search ability. Secondly, excellent decisions are obtained because of promotion of convergence and uniformity in K-NSGA-II. Finally, runtime is reduced. This is due to the utilization of Kriging-based surrogate model. K-NSGA-II is effective and time-saving to solve CECEED problems.

TABLE 4 Best solutions in terms of the best fuel cost and emission for case 1.2.

Items	K-NSGA-II	MHBA	M-NSGA-II	NAGA-II	RCGA	K-NSGA-II	MHBA	M-NSGA-II	NAGA-II	RCGA
	In terms of the best fuel cost					In terms of the best emission				
P_1	616.42	876.82	647.70	619.24	657.56	300.14	300.05	318.55	309.40	299.24
P_2	10.04	483.11	52.96	95.49	56.72	10.09	499.49	406.68	413.20	429.01
P_3	14.04	10.95	88.71	89.43	83.98	30.28	10.00	80.82	84.30	83.29
P_4	299.53	31.11	295.43	299.84	295.47	40.02	30.00	297.83	284.14	283.87
P_5	1.00	250.00	41.97	40.39	40.36	1.10	40.00	396.48	395.50	399.73
P_6	3.28	1.00	7.01	2.91	3.99	3.07	1.02	8.39	6.53	9.87
P_7	1.05	3.00	16.57	10.56	17.01	240.00	3.00	22.85	22.85	22.16
P_8	5.35	30.00	30.89	33.21	30.11	5.01	239.76	232.41	238.81	236.78
P_9	21.75	5.00	47.87	33.63	46.01	133.83	5.65	43.54	49.86	44.70
P_{10}	21.08	20.00	171.92	198.87	152.79	198.51	126.86	199.39	178.82	199.21
P_{11}	399.97	20.00	192.81	191.02	199.62	399.79	200.00	198.89	194.30	199.48
P_{12}	397.92	385.48	399.75	386.31	399.86	399.79	400.0o	330.96	320.01	362.94
P_{13}	435.62	396.17	399.46	395.33	384.55	495.41	399.49	390.69	389.45	371.46
P_{14}	597.74	599.63	581.20	548.72	597.15	600.00	600.00	177.04	159.18	168.44
P_{15}	1.01	1.00	1.18	1.09	2.93	1.01	1.66	4.80	3.02	1.88
P_{16}	699.99	698.72	602.69	661.36	651.21	699.96	700.00	316.59	347.36	282.91
P_{17}	298.51	30.00	271.81	252.20	239.03	299.88	300.00	263.65	288.18	296.96
P_{18}	6.83	5.00	6.76	5.97	23.03	5.00	6.37	47.78	48.76	49.15
P_{19}	4.09	4.21	6.16	4.80	5.50	4.64	4.65	34.53	37.94	31.69
Losses	167.21	198.07	194.93	202.46	218.98	199.51	200.05	103.95	103.69	104.86
Cost	10839.80	10537.6	11944.0	12043.8	12039.2	12179.29	12222.9	18482.0	18533.6	18795.7
Emission	5.5329	5.5446	13.5930	13.2770	14.2840	5.5117	5.5117	5.5210	5.5390	5.5630
Time	37.60	109.664	870.523	867.342	31938.32	44.42	109.664	870.523	867.342	31938.32

4.1.2 Case 1.2: All proposed constraints

In this case, total 20 solutions are produced on the Pareto front in Figures 6C, D. These solutions when minimizing the fuel cost and emission are (10839.80, 5.53285) and (12179.29, 5.5117), respectively. Comparisons with the other algorithms are presented in Table 4. The minimum fuel cost is only 10839.80 $/h, and it is less than the best results of three other algorithms, i.e., M-NSGA-II, NSGA-II, and RCGA. Again, the runtime is significantly lower than for the other algorithms. Compared with other methods, the runtime is reduced by 65.71%, 95.68%, 95.66%, and 99.88%, respectively. The number of evaluations required by these algorithms are 20000, 50000, 30000, and 50000, respectively, whereas K-NSGA-II requires only 233 evaluations. The time gap and the fewer iterations further illustrate the advantages of K-NSGA-II for CECEED. Table 4 indicates that the minimum emission achieved by the proposed method is 5.5117 ton/h. Compared with four algorithms, runtime is reduced by 59.49%, 94.90%, 94.88%, and 99.86%, respectively. K-NSGA-II requires only 296 evaluations.

In the optimization process (in terms of the best fuel cost), the surrogate-based evolutions are executed 5,000 times. For executing two strategies in 3.2.1.1 and 3.2.1.2, the total number of additional evolutions is 9,765. For executing the strategy in 3.2.2.1, the number of additional evolutions is 40000. As for the strategy in 3.2.2.2, the number of additional evolutions is also 40000. Then, the total saved time of evolutions is 4.78 s. The saved time accounts for 12.72% of the total running time in this case. When obtaining the best emission, the saved time is 4.81 s. It accounts for 10.83% of the total running time. This further demonstrates that the proposed method is time-saving and effective.

Let us see what performance the proposed data-driven surrogate-based method can obtain. In this case, more complex constraints are involved. Proposed method also obtains excellent Pareto fronts for making dispatching decisions in CEED problems. On one hand, the K-NSGA-II is effective to solve optimization problems with complicated constraints. On the other hand, more importantly, the execution time is still very short. This illustrates the Kriging-

TABLE 5 Best solutions in terms of cost for case 2.

Items	K-NSGA-II	MHBA	Items	K-NSGA-II	MHBA	Items	K-NSGA-II	MHBA
P_1	361.87	464.10	P_{20}	1292.13	1191.94	P_{39}	1127.68	1238.36
P_2	211.20	30.01	P_{21}	595.56	600.00	P_{40}	321.08	241.13
P_3	310.94	158.95	P_{22}	1628.62	1927.58	P_{41}	334.24	396.03
P_4	45.69	20.58	P_{23}	376.39	479.90	P_{42}	346.64	373.71
P_5	122.10	25.00	P_{24}	364.47	302.89	P_{43}	110.75	188.83
P_6	1600.86	1551.29	P_{25}	139.86	20.58	P_{44}	406.28	487.19
P_7	161.79	278.32	P_{26}	491.46	526.34	P_{45}	599.98	556.67
P_8	256.25	269.76	P_{27}	175.34	229.43	P_{46}	20.59	20.75
P_9	779.15	751.84	P_{28}	326.15	333.00	P_{47}	1079.02	846.45
P_{10}	140.33	91.81	P_{29}	499.92	372.98	P_{48}	19.66	15.08
P_{11}	215.08	180.91	P_{30}	298.75	296.88	P_{49}	42.92	159.13
P_{12}	82.16	25.58	P_{31}	695.15	686.11	P_{50}	449.27	477.57
P_{13}	482.92	446.96	P_{32}	37.20	225.36	P_{51}	397.64	398.78
P_{14}	82.77	154.78	P_{33}	635.44	610.49	P_{52}	97.77	90.04
P_{15}	97.19	225.79	P_{34}	649.54	540.44	P_{53}	1397.58	1106.03
P_{16}	91.53	222.96	P_{35}	147.64	142.99	P_{54}	766.86	715.88
P_{17}	50.25	131.09	P_{36}	39.55	94.67	P_{55}	643.85	620.79
P_{18}	87.51	188.83	P_{37}	599.44	478.86	P_{56}	18.01	21.42
P_{19}	1202.27	1134.53	P_{38}	479.55	641.20	P_{57}	26.76	21.25
Cost	**62719.58**	62897.2	Losses	535.27	504.58			
Emission	11.6688	11.6701	Time	**114.548**	265.118			

The bold values are the best comparison results.

based surrogate model is capable for replacing original objective functions with complicated constraints. In a word, this data-driven surrogate-assisted algorithm is effective and time-saving for solving CECEED problems with complex constraints.

4.2 Case 2: IEEE 300-bus test system

Total load demand is 23525.85 MW (Paul et al., 2022). In this case, total 20 solutions are produced in Figures 6E, F. Solutions given by K-NSGA-II when minimizing the fuel cost and emission are (62719.58, 11.6688) and (67853.11, 11.4796), respectively. Comparison results given by MHBA are presented in Tables 5, 6. The best cost given by K-NSGA-II is 62719.58 $/h, and it is better than that of MHBA. Runtime of K-NSGA-II is lower, by 56.79%, than that of MHBA. MHBA requires 20000 evaluations, compared with just 304 for K-NSGA-II. This time gap and reduced number of evaluations further illustrate the advantages of K-NSGA-II for CECEED. In Table 5, minimum emission is 11.4796 ton/h. Compared with MHBA, runtime is reduced by 58.59% and 20000 evaluations of MHBA is reduced to 288. Thus, K-NSGA-II has strong potential to be applied in CECEED problems. In this case, the number of iterations is set as 100. Total runtimes (sum of 50 runtimes) of fuel cost and emission function are 0.009092 s and 0.012639 s, respectively. Total runtimes (sum of 50 runtimes) of surrogate models are 0.001631 s and 0.004326 s, respectively. In terms of the best cost, the saved time of evolutions is 29.39 s (the number of surrogate-based evaluations is 93165). The saved time accounts for 25.66% of the total running time. In terms of the best emission, the saved time of evolutions is 29.80 s (the number of surrogate-based evaluations is 94447). The saved time accounts for 26.99% of the total running time. This further demonstrates the advantages of the proposed surrogate-based method in solving high-dimension CECEED problems.

Let us see what performance the proposed data-driven method can obtain. In this case, a power grid with more units is used to verify the effectiveness of the proposed Kriging-based optimization method. The proposed method obtains excellent Pareto fronts of CECEED problems. On one hand, the K-NSGA-II is effective for solving CECEED problems. On the other hand, execution time is dramatically reduced in a CECEED problem. This illustrates the Kriging surrogate model is effective to replace original high-dimension objective functions. In a word, the proposed method is effective and time-saving to solve CECEED problems with a high dimension.

TABLE 6 Best solutions in terms of emission for case 2.

Items	K-NSGA-II	MHBA	Items	K-NSGA-II	MHBA	Items	K-NSGA-II	MHBA
P_1	361.87	464.10	P_{20}	698.92	1238.53	P_{39}	467.41	1097.37
P_2	211.20	30.01	P_{21}	599.42	577.99	P_{40}	370.06	251.65
P_3	310.94	158.95	P_{22}	942.43	1968.67	P_{41}	500.00	407.55
P_4	45.69	20.58	P_{23}	600.00	592.64	P_{42}	437.79	383.19
P_5	122.10	25.00	P_{24}	189.71	322.53	P_{43}	284.17	191.92
P_6	1600.86	1551.29	P_{25}	183.19	109.29	P_{44}	600.00	592.76
P_7	161.79	278.32	P_{26}	600.00	599.37	P_{45}	549.28	581.46
P_8	256.25	269.76	P_{27}	292.97	247.09	P_{46}	134.60	13.70
P_9	779.15	751.84	P_{28}	401.37	332.69	P_{47}	1656.92	1658.71
P_{10}	140.33	91.81	P_{29}	500.00	413.40	P_{48}	18.34	23.60
P_{11}	215.08	180.91	P_{30}	399.99	332.15	P_{49}	151.37	186.94
P_{12}	82.16	25.58	P_{31}	672.40	665.50	P_{50}	467.05	435.70
P_{13}	482.92	446.96	P_{32}	289.21	258.23	P_{51}	500.00	440.77
P_{14}	82.77	154.78	P_{33}	695.77	581.47	P_{52}	33.57	127.46
P_{15}	97.19	225.79	P_{34}	699.82	695.81	P_{53}	922.72	1075.08
P_{16}	91.53	222.96	P_{35}	229.79	192.25	P_{54}	792.89	718.64
P_{17}	50.25	131.09	P_{36}	116.49	96.00	P_{55}	790.95	615.76
P_{18}	87.51	188.83	P_{37}	599.07	589.89	P_{56}	20.90	62.56
P_{19}	1202.27	1134.53	P_{38}	745.77	724.05	P_{57}	66.16	24.12
Cost	67853.11	66317.4	Losses	1156.58	1299.27			
Emission	**11.4796**	11.5385	Time	**110.40**	266.581			

The bold values are the best comparison results.

4.3 Comparative results with Cplex and Guribo solvers

The comparative results with Cplex (Huo et al., 2022) and Guribo (Ma et al., 2021) solvers for solving the case 1.1, case 1.2, and case 2 are displayed in this section. Mathematical programming and precise algorithms are preferred in most of the real-world applications. However mathematical methods (by using Cplex and Guribo solvers) are usually powerless to solve multi-objective optimization problems. To use the two mathematical solvers, original multi-objective CECEED problems are transferred to single-objective optimization problems by weight sum method (Zuo et al., 2014):

$$F(P) = \theta C(P) + (1-\theta)E(P), \tag{19}$$

where θ represents the weight factor. For obtaining Pareto fronts in case1.1, case1.2, and case2, the θ shrinks from 1 to 0 with a step size 1/19. After the proposed transition, due to the non-convex objective functions, the piecewise linear method should be applied (Wu and Dong, 2023):

$$\begin{cases} \Delta P = (P_{\max} - P_{\min})/\mathrm{Pie} \\ \mathrm{F_{piece}}(P) \begin{cases} \mathrm{F_{piece,1}}(P), P_{\min} \le P \le P_{\min} + \Delta P \\ \mathrm{F_{piece,2}}(P), P_{\min} + \Delta P \le P \le P_{\min} + 2\Delta P \\ \cdots \\ \mathrm{F_{piece,}}i(P), P_{\min} + (i-1)\Delta P \le P \le P_{\min} + i\Delta P \\ \cdots \\ \mathrm{F_{piece,Pie}}(P), P_{\min} + (\mathrm{Pie}-1)\Delta P \le P \le P_{\max} \end{cases} \end{cases} \tag{20}$$

where ΔP is the width of P in a piece; $\mathrm{F_{piece}}(P)$ represents the piecewise linear function of original objective function ($F(P)$); $\mathrm{F_{piece,}}i(P)$ represents the i-th piecewise linear function in $\mathrm{F_{piece}}(P)$; Pie represents total Pie pieces are contained in $\mathrm{F_{piece}}(P)$. Then, the original objective function $F(P)$ can be described as:

$$\begin{cases} F(P) = \sum_{i=1}^{\mathrm{Pie}} b_i \mathrm{F_{piece,}}i(P); \\ b_i \in \{0,1\}; \\ \sum_{i=1}^{\mathrm{Pie}} b_i = 1; \end{cases} \tag{21}$$

where b_i is a constant which is equal to 0 or 1; $b_i = 1$ represents that the i-th piecewise linear function ($\mathrm{F_{piece,}}i(P)$) is used as the current

TABLE 7 Comparative results with Cplex and Guribo solvers. (The execution time is in parentheses).

Items	The best fuel cost: $/h			The best emission: ton/h		
Index of cases	This paper	Cplex	Guribo	This paper	Cplex	Guribo
1.1	10180.37 (34.751 s)	10258.50 (51.35 s)	10247.91 (48.36 s)	5.5116 (34.939 s)	5.5426 (52.01 s)	5.5430 (49.33 s)
1.2	10839.80 (37.60 s)	11031.06 (60.17 s)	11031.26 (57.36 s)	5.5117 (37.60 s)	5.5136 (59.27 s)	5.5137 (54.28 s)
2	62719.58 (110.10 s)	66346.96 (531.87 s)	65625.83 (183.21 s)	11.4796 (110.40 s)	11.4975 (521.28 s)	11.4974 (184.56 s)

objective function; $\sum_{i=1}^{\text{Pie}} b_i = 1$ represents that only one piecewise linear function can be chosen (only one b_i is equal to 1). After the following additional processing, the mathematical programming methods (Cplex and Guribo solvers) can be used to solve the proposed CECEED problems.

As shown in Table 7, for case 1.1, compared with Cplex solver and Guribo solver, the best fuel cost obtained by the proposed method is reduced by 0.76% and 0.66%, respectively. As for the execution time, it is reduced by 32.33% and 28.14%, respectively. In terms of pollutant emission, the best emission is reduced by 0.56% and 0.56%, respectively. As for the execution time, it is reduced by 32.83% and 29.17%, respectively. In case 1.2, the improved degrees are similar to the comparative results in case 1.1. In case 2, a large-scale system with more dimensions, compared with Cplex and Guribo solvers, the best fuel cost obtained by the proposed method is reduced by 5.47% and 4.43%, respectively. As for the execution time, it is reduced by 79.30% and 39.91%, respectively. In terms of the best emission in case 2, the best emission is reduced by 0.15% and 0.15%, respectively. As for the execution time, it is reduced by 78.82% and 40.18%, respectively. Compared with the test system with 118 buses, the proposed surrogate-assisted method has a stronger advantage. In other words, it is effective for handling the CECEED problems with more dimensions and larger scale.

5 Concluding remarks

This paper proposes a novel data-driven Kriging-assisted method for solving CECEED problems. The optimization process includes two aspects. Kriging-based surrogate models are used to replace the original computationally expensive objective functions, reducing the evaluation time. Additionally, the NSGA-II method is improved to enhance its ability to handle high-dimensional optimization problems. A new infilling sampling strategy is proposed to update the Kriging-based surrogate models, thus enhancing the accuracy of these models. Novel optimization strategies were added to improve NSGA-II, focusing on convergence, uniformity, extensibility, and search ability, respectively. The effectiveness of this method has been illustrated by conducting simulations of the IEEE 118-bus and 300-bus systems. The results indicate that this data-driven Kriging-assisted method is suitable for solving CECEED problems. Note that time-consuming unit commitment problems could be solved by such a surrogate-assisted approach and left as future research.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

CL: Investigation, Software, Validation, Writing–original draft. HL: Conceptualization, Formal Analysis, Funding acquisition, Project administration, Resources, Supervision, Writing–review and editing. AP: Data curation, Writing–review and editing. JZ: Visualization, Writing–review and editing.

Funding

The authors declare financial support was received for the research, authorship, and/or publication of this article. Project Supported by National Natural Science Foundation of China (NSFC) (NO. 62163013).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Chen, C., Wang, L., and Niu, M. (2023). Research on the application of improved nsga-ii in the structure design of wind turbine blade spar cap. *Front. Energy Res.* 11, 1160423. doi:10.3389/fenrg.2023.1160423

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: nsga-ii. *IEEE Trans. Evol. Comput.* 6, 182–197. doi:10.1109/4235.996017

Gao, Y., Kang, B., Xiao, H., Wang, Z., Ding, G., Xu, Z., et al. (2023). A model for identifying the feeder-transformer relationship in distribution grids using a data-driven machine-learning algorithm. *Front. Energy Res.* 11, 1225407. doi:10.3389/fenrg.2023.1225407

Huo, X., Wu, X., Fan, Y., and Ding, C. (2022). A mixed-integer program (mip) for one-way multiple-type shared electric vehicles allocation with uncertain demand. *IEEE Trans. Intelligent Transp. Syst.* 23, 8972–8984. doi:10.1109/TITS.2021.3088858

Lai, W., Zheng, X., Song, Q., Hu, F., Tao, Q., and Chen, H. (2022). Multi-objective membrane search algorithm: A new solution for economic emission dispatch. *Appl. Energy* 326, 119969. doi:10.1016/j.apenergy.2022.119969

Li, J., Zuo, W., E, J., Zhang, Y., Li, Q., Sun, K., et al. (2022a). Multi-objective optimization of mini u-channel cold plate with sio2 nanofluid by rsm and nsga-ii. *Energy* 242, 123039. doi:10.1016/j.energy.2021.123039

Li, L. L., Liu, Z. F., Tseng, M. L., Zheng, S. J., and Lim, M. K. (2021). Improved tunicate swarm algorithm: solving the dynamic economic emission dispatch problems. *Appl. Soft Comput.* 108, 107504. doi:10.1016/j.asoc.2021.107504

Li, Q., Liu, Z., Xiao, Y., Zhao, P., Zhao, Y., Yang, T., et al. (2022b). An intelligent optimization method for preliminary design of lead-bismuth reactor core based on kriging surrogate model. *Front. Energy Res.* 10, 849229. doi:10.3389/fenrg.2022.849229

Li, Z., Wu, L., Xu, Y., Moazeni, S., and Tang, Z. (2022c). Multi-stage real-time operation of a multi-energy microgrid with electrical and thermal energy storage assets: A data-driven mpc-adp approach. *IEEE Trans. Smart Grid* 13, 213–226. doi:10.1109/TSG.2021.3119972

Li, Z., Xu, Y., Fang, S., Wang, Y., and Zheng, X. (2020). Multiobjective coordinated energy dispatch and voyage scheduling for a multienergy ship microgrid. *IEEE Trans. Industry Appl.* 56, 989–999. doi:10.1109/TIA.2019.2956720

Li, Z., and Xu, Y. (2018). Optimal coordinated energy dispatch of a multi-energy microgrid in gridconnected and islanded modes. *Appl. Energy* 210, 974–986. doi:10.1016/j.apenergy.2017.08.197

Liang, H., Liu, Y., Li, F., and Shen, Y. (2018). A multiobjective hybrid bat algorithm for combined economic/emission dispatch. *Int. J. Electr. Power and Energy Syst.* 101, 103–115. doi:10.1016/j.ijepes.2018.03.019

Lin, C., Liang, H., and Pang, A. (2023). A fast data-driven optimization method of multi-area combined economic emission dispatch. *Appl. Energy* 337, 120884. doi:10.1016/j.apenergy.2023.120884

Linka, K., Hillgärtner, M., Abdolazizi, K. P., Aydin, R. C., Itskov, M., and Cyron, C. J. (2021). Constitutive artificial neural networks: A fast and general approach to predictive data-driven constitutive modeling by deep learning. *J. Comput. Phys.* 429, 110010. doi:10.1016/j.jcp.2020.110010

Liu, Y., Wang, M., Zhang, X., Duan, J., Gao, H., and Liu, J. (2022). Kriging surrogate model enabled heuristic algorithm for coordinated volt/var management in active distribution networks. *Electr. Power Syst. Res.* 210, 108089. doi:10.1016/j.epsr.2022.108089

Ma, Z., Zhong, H., Cheng, T., Pi, J., and Meng, F. (2021). Redundant and nonbinding transmission constraints identification method combining physical and economic insights of unit commitment. *IEEE Trans. Power Syst.* 36, 3487–3495. doi:10.1109/TPWRS.2020.3049001

Muthuswamy, R., Krishnan, M., Subramanian, K., and Subramanian, B. (2015). Environmental and economic power dispatch of thermal generators using modified nsga-ii algorithm. *Int. Trans. Electr. Energy Syst.* 25, 1552–1569. doi:10.1002/etep.1918

Pang, A., Liang, H., Lin, C., and Yao, L. (2023). A surrogate-assisted adaptive bat algorithm for large-scale economic dispatch. *Energies* 16, 1011. doi:10.3390/en16021011

Paul, K., Sinha, P., Mobayen, S., El-Sousy, F. F., and Fekih, A. (2022). A novel improved crow search algorithm to alleviate congestion in power system transmission lines. *Energy Rep.* 8, 11456–11465. doi:10.1016/j.egyr.2022.08.267

Qian, H. M., Wei, J., and Huang, H. Z. (2023). Structural fatigue reliability analysis based on active learning kriging model. *Int. J. Fatigue* 172, 107639. doi:10.1016/j.ijfatigue.2023.107639

Qu, K., Zheng, X., and Yu, T. (2023). Environmental-economic unit commitment with robust diffusion control of gas pollutants. *IEEE Trans. Power Syst.* 38, 818–834. doi:10.1109/TPWRS.2022.3166264

Sharifian, Y., and Abdi, H. (2023). Solving multi-area economic dispatch problem using hybrid exchange market algorithm with grasshopper optimization algorithm. *Energy* 267, 126550. doi:10.1016/j.energy.2022.126550

Sheng, W., Li, R., Yan, T., Tseng, M. L., Lou, J., and Li, L. (2023). A hybrid dynamic economics emissions dispatch model: distributed renewable power systems based on improved coot optimization algorithm. *Renew. Energy* 204, 493–506. doi:10.1016/j.renene.2023.01.010

Wang, R., Wen, X., Wang, X., Fu, Y., and Zhao, Y. (2023). Low-carbon economic dispatch of regional integrated energy system based on carbon-oxygen cycle. *Front. Energy Res.* 11, 1206242. doi:10.3389/fenrg.2023.1206242

Wei, H., Zhang, Y., Wang, Y., Hua, W., Jing, R., and Zhou, Y. (2022). Planning integrated energy systems coupling v2g as a flexible storage. *Energy* 239, 122215. doi:10.1016/j.energy.2021.122215

Wu, F., and Dong, M. (2023). Eco-routing problem for the delivery of perishable products. *Comput. Operations Res.* 154, 106198. doi:10.1016/j.cor.2023.106198

Xu, L., Wei, W., Cai, X., Liu, C., Jiang, X., and Yang, J. (2022). Day-ahead economic dispatch strategy for distribution network considering total cost price-based demand response. *Front. Energy Res.* 10, 870893. doi:10.3389/fenrg.2022.870893

Yadav, A., Mishra, S., and Sairam, A. S. (2022). A multi-objective worker selection scheme in crowdsourced platforms using nsga-ii. *Expert Syst. Appl.* 201, 116991. doi:10.1016/j.eswa.2022.116991

Yu, F., Gong, W., and Zhen, H. (2022). A data-driven evolutionary algorithm with multi-evolutionary sampling strategy for expensive optimization. *Knowledge-Based Syst.* 242, 108436. doi:10.1016/j.knosys.2022.108436

Zhang, R., and Yu, J. (2023). New urban power grid flexible load dispatching architecture and key technologies. *Front. Energy Res.* 11, 1168768. doi:10.3389/fenrg.2023.1168768

Zuo, X., Murray, C. C., and Smith, A. E. (2014). Solving an extended double row layout problem using multiobjective tabu search and linear programming. *IEEE Trans. Automation Sci. Eng.* 11, 1122–1132. doi:10.1109/TASE.2014.2304471