Data-driven Surrogate-assisted Method for High-dimensional Multi-area Combined Economic/Emission Dispatch

Chenhao Lin, Huijun Liang, Aokang Pang, Jianwei Zhong, and Yongchao Yang

***Abstract*—Multi-area combined economic/emission dispatch (MACEED) problems are generally studied using analytical functions. However, as the scale of power systems increases, existing solutions become time-consuming and may not meet operational constraints. To overcome excessive computational expense in high-dimensional MACEED problems, a novel data-driven surrogate-assisted method is proposed. First, a cosine-similarity-based deep belief network combined with a back-propagation (DBN + BP) neural network is utilized to replace cost and emission functions. Second, transfer learning is applied with a pretraining and fine-tuning method to improve DBN + BP regression surrogate models, thus realizing fast construction of surrogate models between different regional power systems. Third, a multi-objective antlion optimizer with a novel general single-dimension retention bi-objective optimization policy is proposed to execute MACEED optimization to obtain scheduling decisions. The proposed method not only ensures the convergence, uniformity, and extensibility of the Pareto front, but also greatly reduces the computational time. Finally, a 4-area 40-unit test system with different constraints is employed to demonstrate the effectiveness of the proposed method.**

***Index Terms*—Multi-area combined economic/emission dispatch, high-dimensional power system, deep belief network, data driven, transfer learning.**

I. Introduction

With the expansion of power systems and the decentralization of load centers, multi-area power systems consisting of multiple interconnected load centers are essential for safe and stable operation [1]. Additionally, as the electricity industry achieves more autonomy in its business practices, improved economic performance and reduced emissions are expected [2]. Based on these two factors, multi-area combined economic/emission dispatch (MACEED) provides power systems with greater stability and better economic performance [3]. In multi-area power systems, the customer load is the collective responsibility of generators distributed in different areas, offering richer possibilities for the economic and stable operation of power systems. Therefore, MACEED systems warrant further study to minimize fuel costs and emissions.

Traditional methods for solving MACEED problems include the linear weighted-sum method [4], Jacobian-based algorithm [5], and Lagrange algorithm [6]. Many evolutionary algorithms can also be employed to solve MACEED problems such as the multilayer distributed multi-objective consensus [7], fuzzified squirrel search [8], and multi-objective squirrel search [9] algorithms.

Existing methods for solving MACEED problems have been applied to low-dimensional small-scale power systems. However, as the power industry continues to grow, power systems have become increasingly complex. This has resulted in high-dimensional MACEED problems in large-scale power systems, which are computationally expensive [10]. Traditional methods and evolutionary algorithms have struggled to solve such problems [11]. When a method is said to be "computationally expensive", this means that the evaluation itself requires considerable time, computing power, money, and other expenses [12]. However, other problems will become "computationally expensive" under certain circumstances [13]. For high-dimensional large-scale MACEED problems, high-dimensional optimization requires a significant amount of time. However, due to the 15 min dispatch cycle, MACEED problems have become more expensive. Therefore, a computationally efficient method suitable to address high-dimensional MACEED problems is urgently required.

Data-driven surrogate-assisted models [14] have been widely employed to solve computationally expensive problems [15]. In [16], a support vector regression (SVR) surrogate-assisted optimization method was proposed to solve MACEED problems. The original cost and emission functions in MACEED problems were replaced by two SVR surrogate models to reduce the execution time. However, the divided multi-area system was treated as a whole during the optimization process. This implies that the transfer-learning method may not work well in other systems. Accordingly, applications of this method are limited. In [17], a novel com-

Manuscript received: March 1, 2023; revised: March 27, 2023; accepted: June 2, 2023. Date of CrossCheck: June 2, 2023. Date of online publication: September 12, 2023.

This work was supported in part by the National Natural Science Foundation of China (No. 62163013) and in part by the National Natural Science Foundation of Hubei Province (No. 2021CFB542).

C. Lin, H. Liang (corresponding author), A. Pang, J. Zhong, and Y. Yang are with the College of Intelligent Systems Science and Engineering, Hubei Minzu University, Enshi, China (e-mail: linchenhao@hbmzu.edu.cn; lhj@hbmzu.edu.cn; pakdeyouxiang@hbmzu.edu.cn; 1994009@hbmzu.edu.cn; 2003039@hbmzu.edu.cn).

DOI: 10.35833/MPCE.2023.000128

bined deep belief network (DBN) and self-organizing map model was proposed to denoise and combine vibration signals. This model removed the noise from the vibration signals of a wind turbine, enabling the evaluation of the performance degradation process. In [18], a DBN combined with an extreme learning machine model was developed to predict heart disease. The extreme learning machine was regarded as the top layer of the DBN, and the DBN was then transformed into a regression model. This model achieved good performance in medical data classification across various datasets. In [19], expensive multimodal optimization problems were solved using a multitasking particle swarm optimization (PSO) algorithm assisted by a surrogate model. The multiple surrogate models were integrated by using a multitasking optimization method. However, the convergence of the algorithm was improved by the mixed application of migration and the surrogate model management and trust region search strategies. This algorithm obtained a greater number of improved global optimal solutions for benchmark test functions and energy conservation design problems. In [20], a dual-surrogate-assisted cooperative PSO algorithm was proposed to search for optimal solutions. The algorithm combines a dual-population cooperative PSO with a dual-layer cooperative surrogate model guided by a model. Multiple low-computational-cost-based and highly competitive optimal solutions were obtained using the proposed algorithm. In short, DBN models perform well in solving high-dimensional optimization problems.

From a modeling perspective, consistently rebuilding the surrogate model is extremely time-consuming. Transfer-learning method can be applied to reduce the time required to build models. In [21], a new pretraining and fine-tuning transfer-learning method was applied to problems with numerous parameters by adjusting the deep representation scaling (DRS) layers. During network training, the parameters in DRS layers were trained to improve the pretrained convolutional neural network (CNN). Through this pretraining and fine-tuning transfer-learning method, the number of retrained parameters was reduced, and the classification performance was enhanced as compared with typical transfer-learning methods. In [22], a multiscale fusion method with CNNs by pretraining and fine-tuning was proposed for skin-disease classification utilizing dermoscopic images. The results showed that the integration of the three pretraining and fine-tuning CNNs significantly improved the classification performance. Thus, pretraining and fine-tuning transfer-learning method can ensure surrogate model accuracy and significantly reduce the computational load.

If transfer-learning method is used to construct a surrogate model for MACEED problem, finding a superior heuristic algorithm for iterative optimization is necessary. In [23], novel evolutionary programs were used to solve MACEED problems using valve points. This evolutionary technique ensures a better convergence, higher-quality solutions, and low-computational costs. In [24], a combined differential evolution and PSO algorithm was proposed to solve MACEED problems. This algorithm is effective in terms of balancing the convergence characteristics with a global search capability. Overall, heuristic algorithms have outstanding capabilities in solving decision-making problems, and selecting an appropriate algorithm to solve MACEED problems is entirely feasible.

In general, the aforementioned heuristic algorithms are less effective as the dimensions of the decision variables increase. Thus, a new method for solving MACEED problems is urgently required. The previous analysis reveals three key factors in solving high-dimensional MACEED problems in large-scale power systems: ① constructing surrogate models using historical data to replace the fuel cost and emission functions; ② applying transfer learning to enable the rapid development of the surrogate model; and ③ constructing an algorithm that can match different surrogate models.

In this paper, a method for data-driven high-dimensional MACEED optimization, including constraints, is proposed for large-scale complex power systems. Unlike the research in [16], in which the multi-area system was treated as a whole in the optimization process, in this paper, the original multi-area system is divided into several sub-area systems. This method combines a surrogate model with pretraining and fine-tuning through transfer-learning method and an improved multi-objective antlion optimizer (MOALO). A DBN combined with a back-propagation (DBN + BP) regression surrogate model based on the cosine similarity index is introduced to replace the original objective function. Pretraining and fine-tuning technologies are applied through transfer learning, allowing surrogate models to be constructed as necessary, and the traditional MOALO is improved to match the proposed surrogate models. The main contributions of this paper are as follows.

1) A novel cosine-similarity-based DBN + BP regression surrogate model is proposed for high-dimensional MACEED problems with expensive computing. Surrogate models are employed to replace the original objective functions in the MACEED problems to reduce execution time. Unlike in the classical DBN + BP regression model, the cosine similarity is treated as an additional convergence condition, where the cosine similarity significantly improves the accuracy of MACEED solution.

2) A novel pretraining and fine-tuning transfer-learning method is developed for the cosine-similarity-based DBN+ BP regression surrogate model. The surrogate model of one area can be transferred to another using a dimensional transformation scheme. The DBN part of the proposed surrogate model is copied to the DBN parts of other areas, and the data in the corresponding areas are used to fine-tune the DBN. This transfer-learning method reduces the time required to rebuild DBN + BP models in other areas.

3) An improved MOALO matching surrogate model is proposed to solve computationally expensive MACEED problems. To enhance the performance of MOALO, a novel general single-dimension retention (SDR) policy for bi-objective optimization is used to accelerate convergence. In addition, the optimization process of the sub-areas is executed in parallel. This scheme significantly reduces the operational time for solving MACEED.

The remainder of this paper is organized as follows. Sec-

tion II describes the problem formulation. Section III introduces the data-driven surrogate-assisted method for MACEED problems. Section IV presents the simulation experiments and results. Finally, Section V presents the conclusion and future research.

II. Problem Formulation

This section introduces the objective functions and constraints of a typical MACEED problem [25]. High-dimensional MACEED problems are solved using the surrogate-assisted method and transfer-learning method.

A. Typical Mathematical Descriptions of MACEED Problems

1) Optimization Objectives of MACEED Problems

1) Fuel-cost function:

$$C(P)=\sum_{i=1}^{N_a}\sum_{j=1}^{N_g}(a_{ij}P_{ij}^2+b_{ij}P_{ij}+c_{ij})+|d_{ij}\sin(e_{ij}(P_{ij}^{\min}-P_{ij}))| \quad (1)$$

where N_a is the number of sub-area systems; N_g is the number of units in one sub-area; P_{ij} is the power output of the j^{th} unit in the i^{th} area; $P_{ij}^{\min}$ is the floor level of the j^{th} unit in the i^{th} area; a_{ij}, b_{ij}, and c_{ij} are the coefficients in the fuel-cost functions of the j^{th} unit in the i^{th} area; and d_{ij} and e_{ij} are the valve-point effect coefficients of the j^{th} unit in the i^{th} area [26], [27].

2) Emission function:

$$E(P)=\sum_{i=1}^{N_a}\sum_{j=1}^{N_g}[10^{-2}(\alpha_{ij}P_{ij}^2+\beta_{ij}P_{ij}+\gamma_{ij})+\varepsilon_{ij}\exp(\lambda_{ij}P_{ij})] \quad (2)$$

where α_{ij}, β_{ij}, γ_{ij}, ε_{ij}, and λ_{ij} are the coefficients of the pollutant emission functions of the j^{th} unit in the i^{th} area.

2) Constraints

1) Generation capacity limit:

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

where $P_{ij}^{\max}$ is the upper limit of the j^{th} unit in the i^{th} area.

2) Power balance limit:

$$P_i=P_{di}+P_{lossi}+\sum_{p=1,p\ne i}^{N}T_{ip} \quad (4)$$

where P_i, P_{di}, and P_{lossi} are the total output, power load, and power loss in the i^{th} area, respectively.

3) Tie-line capacity limit:

$$-T_{ip}^{\max}\le T_{ip}\le T_{ip}^{\max} \quad (5)$$

where $T_{ip}^{\max}$ is the maximum capacity of tie line between the i^{th} and p^{th} areas.

4) Spinning reserve constraint between different areas:

$$\sum_{j=1}^{M_p}S_{pj}\ge S_{p,\text{req}}+RC_{ip} \quad (6)$$

where S_{pj} is the surplus capacity of the j^{th} unit in the p^{th} area; $S_{p,\text{req}}$ is the reserve requirement in the p^{th} area; and RC_{ip} is the transferable reserve from the i^{th} area to the p^{th} area.

5) Tie-line capacity limit with shared spinning reserve:

$$-T_{ip}^{\max}\le T_{ip}+RC_{ip}\le T_{ip}^{\max} \quad (7)$$

In addition, as described in (4), P_i represents the sum of all units in the i^{th} area:

$$P_i=\sum_{j=1}^{M_i}P_{ij} \quad (8)$$

where M_i is the number of units in the i^{th} area. The transmission loss P_{lossi} is calculated using B coefficient method:

$$P_{lossi}=\sum_{k=1}^{M_i}\sum_{l=1}^{M_i}P_{ik}B_{kl}^iP_{il}+\sum_{j=1}^{M_i}B_{0l}^iP_{il}+B_{00}^i \quad (9)$$

where B_{kl}^i is the $(k,l)^{th}$ element in the loss-coefficient matrix of the i^{th} area; B_{0l}^i is the l^{th} element in the loss-coefficient vector of the i^{th} area; and B_{00}^i is the loss constant of the i^{th} area. Note that all other specific descriptions of the proposed variables in (3)-(8) can be found in [16].

B. Additional Tasks of Large-scale MACEED Problems

For high-dimensional MACEED problems in large-scale power systems, the construction of surrogate models of fuel cost and emission functions must be considered:

$$\begin{cases}Cost(\cdot)\approx F_{\text{cost}}(S_1,S_2,\ldots,S_n)\\ Emission(\cdot)\approx F_{\text{emission}}(S_1,S_2,\ldots,S_n)\end{cases} \quad (10)$$

where $Cost(\cdot)$ and $Emission(\cdot)$ are the fuel cost and emission objective functions, respectively; $(S_1,S_2,\ldots,S_n)$ represent n feasible scheduling solutions, in which n is the scale of the surrogate model training set; and $F_{\text{cost}}(S_1,S_2,\ldots,S_n)$ and $F_{\text{emission}}(S_1,S_2,\ldots,S_n)$ are the surrogate models of $Cost(\cdot)$ and $Emission(\cdot)$, respectively. In addition, fast modeling can be achieved using transfer learning, as shown in Fig. 1.

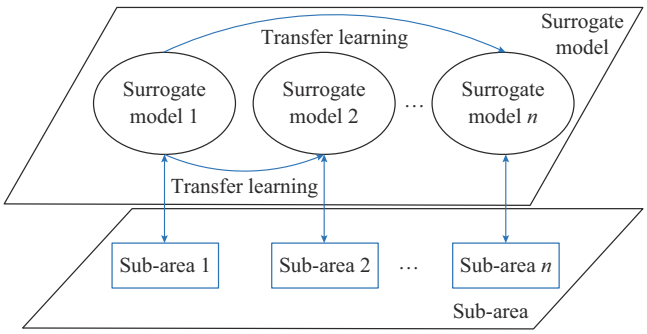

Fig. 1. Transfer learning between different areas.

III. Data-driven Surrogate-assisted Method for MACEED Problems

This section introduces the cosine-similarity-based DBN+BP regression surrogate models to replace the fuel cost and emission functions (1) and (2). The transfer-learning method for pretraining and fine-tuning is used to train the surrogate models rapidly in different areas. Finally, the MOALO is improved and introduced to match surrogate models for high-dimensional MACEED problems.

A. Construction of Cosine-similarity-based DBN+BP Regression Surrogate Model

Most existing methods for addressing high-dimensional MACEED problems require long computational time to determine scheduling strategies and may in fact even fail to find meaningful scheduling strategies [28]. The "curse of dimensionality" is a result of numerous units and constraints

introduced by power system interconnections meaning that the computational time increases exponentially. To overcome this drawback, cosine-similarity-based DBN + BP regression surrogate models are constructed online to replace the mathematical expressions of optimization objectives [29], where the structure is shown in Fig. 2. The settings of the cosine-similarity-based DBN + BP regression surrogate model are presented in Table I. Specifically, the fourth layer of the DBN involves logistic regression, the output of which provides the input for the BP neural network. First, unlabeled data are extracted from the five layers of the DBN. The corresponding labels are then added, and logistic regression is applied to obtain the primary estimation values. Finally, regression evaluation is performed using the BP neural network.

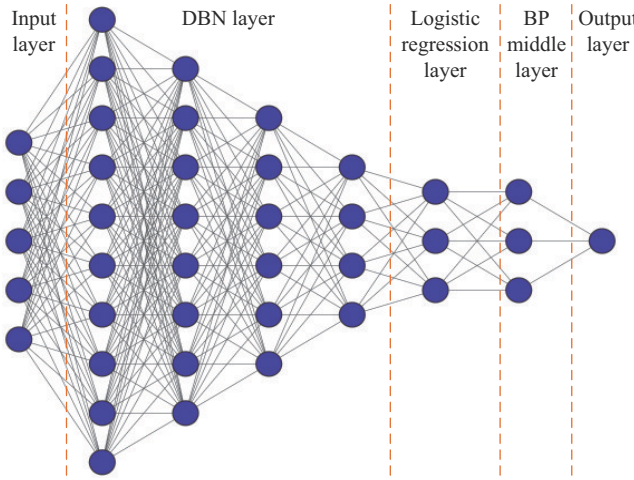

Fig. 2. Structure of proposed cosine-similarity-based DBN + BP regression surrogate model.

TABLE I
SETTINGS OF COSINE-SIMILARITY-BASED DBN + BP REGRESSION SURROGATE MODEL

Section	Floor	Number of node
Input layer	1	10
DBN layer	1	80
	2	60
	3	40
	4	20
Logistic regression layer	1	20
BP middle layer	1	10
Output layer	1	1

The chief operation in the surrogate model is that the convergence criterion combines the minimum mean square error (MSE) with the maximum cosine similarity of the result vector, which is described by:

$$C_s = \frac{\sum_{i=1}^{n_s} A_{i_s} B_{i_s}}{\sqrt{\sum_{i_s=1}^{n_s} A_{i_s}^2} \sqrt{\sum_{i_s=1}^{n_s} B_{i_s}^2}} \tag{11}$$

where C_s is the cosine similaring between A_{i_s} and B_{i_s}; A_{i_s} and B_{i_s} are the i^{th} elements in the sequences of fitting results and of true results, respectively; and n_s is the number of elements in the sequence.

The goal of an optimization task is to find the minimum value that differs from that of the prediction tasks. However, employing only the minimum MSE to evaluate the model is not appropriate because correct scheduling results can be obtained with a relatively large error if the size relationship between elements remains consistent. Conversely, even if the MSE is very small, the optimal solution will be inaccurate if the original size relationship cannot be guaranteed. Based on these points, the cosine similarity between the fitting result and real value vectors should be as large as possible. This provides the convergence criterion for the cosine-similarity-based DBN + BP regression surrogate models and represents the main improvement to the traditional DBN regression model.

To further examine the time-saving advantage of the DBN + BP regression surrogate model, comparison results of the proposed model, original objective functions, and piecewise linear functions [30] are presented in Table II. Specifically, a 4-area 40-unit test system (corresponding to an actual Taiwan power grid, China) [31] is simulated as an example. The segments of the piecewise linear function are set to be 5.

TABLE II
COMPARISON RESULTS OF ACCURACY AND TIME

Area	Type	Method	Accuracy	Time (s)
Area 1	Cost	Original objective functions		7.94×10^{-3}
		Piecewise linear functions	99.51	7.37×10^{-3}
		Proposed model	99.73	1.20×10^{-5}
	Emission	Original objective functions		8.18×10^{-3}
		Piecewise linear functions	99.38	7.68×10^{-3}
		Proposed model	99.45	1.19×10^{-5}
Area 2	Cost	Original objective functions		9.68×10^{-3}
		Piecewise linear functions	99.49	8.96×10^{-3}
		Proposed model	99.13	1.32×10^{-5}
	Emission	Original objective functions		1.03×10^{-2}
		Piecewise linear functions	99.01	9.89×10^{-3}
		Proposed model	99.25	1.71×10^{-5}
Area 3	Cost	Original objective functions		6.38×10^{-3}
		Piecewise linear functions	98.98	5.44×10^{-3}
		Proposed model	99.36	1.59×10^{-5}
	Emission	Original objective functions		7.09×10^{-3}
		Piecewise linear functions	97.36	5.74×10^{-3}
		Proposed model	99.02	2.36×10^{-5}
Area 4	Cost	Original objective functions		5.85×10^{-3}
		Piecewise linear functions	98.65	4.17×10^{-3}
		Proposed model	98.69	1.88×10^{-5}
	Emission	Original objective functions		6.58×10^{-3}
		Piecewise linear functions	99.09	5.97×10^{-3}
		Proposed model	99.61	1.39×10^{-5}

In contrast to piecewise linear functions, the execution time is reduced using the proposed model. In addition, the accuracy is improved by a small degree. The accuracy of a piecewise linear function can be improved by dividing the function into additional segments. However, the execution time is also improved. Therefore, the application of piecewise linear functions is limited. This further illustrates the advantages of the proposed model, where both a significantly reduced execution time and a higher degree of accuracy are achieved.

The execution of the MACEED optimization process is facilitated by the accuracy of the high surrogate model. On the one hand, the quality of the solution benefits from the accuracy of the surrogate model. Thus, compared with the piecewise linear function, the proposed model with a higher degree of accuracy is preferable for solving MACEED problems. On the other hand, the error of the piecewise linear function represents the cumulative result of each generator. The error of the emission function in each generator also accumulates because a piecewise linear function is employed. Let us consider a scenario in which a very low emission of polluting gas is obtained by the cumulative error of the piecewise linear function when multi-objective optimization is performed. However, in practice, this solution is not feasible, as it causes the final Pareto frontier to collapse into an unfeasible domain. Even if the sampling mechanism is added to perform real function fitness, the results calculated using a real function do not fully provide individuals with guiding significance for the algorithm. This undoubtedly makes the algorithm non-convergent and even impossible to solve. Using the proposed model ensures better scheduling schemes with reduced computational time.

B. Fast Transfer-learning Method for Building Surrogate Models in Different Areas

The construction of online cosine-similarity-based DBN+BP regression surrogate models requires massive amounts of data, and processing these data is time-consuming [32]. To more closely match the surrogate model to the original model, considerable data are required to train the surrogate model. Unfortunately, this time-consuming process goes against the original intention of reducing computational time.

To overcome this drawback, a high-dimensional large-scale power system is decomposed into several dispatching sub-areas. By learning the surrogate models constructed for one area, surrogate models can be developed for other areas using the transfer-learning method. A schematic of the pretraining and fine-tuning processes of transfer learning is shown in Fig. 3.

First, the surrogate model trained in area 1 is divided into five layers as introduced in Section III-A. Second, the entire DBN section of the surrogate model in area 1 is copied to the DBN of the surrogate model in other areas; that is, the DBN of area 1 is treated as the pretrained DBN of other areas. In addition, the data in the original domain must be adequately labeled to access the target domain. If the data in the original domain are not accessible to the target domain, dimensionality reduction schemes can be utilized.

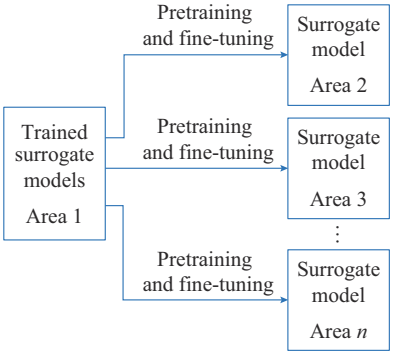

Fig. 3. Pretraining and fine-tuning processes of transfer learning.

Finally, the data from other areas are used to fine-tune the copied DBN and reconstruct the logistic regression layer and BP neural network. Based on this transfer-learning method, cosine-similarity-based DBN+BP regression surrogate models are quickly constructed for different areas. The mean square errors and cosine similarities of different surrogate models are guaranteed.

Different areas are used to verify the advantages of transfer learning. Specifically, a 4-area 40-unit test system is employed. The surrogate models of area 1 are used as the source domain, with the other three areas used as the target domains. This means that the dimensions of the source and target domains are identical. Additionally, the transfer learning with different dimensions is simulated as another case. The different combinations of the four areas are used as target domains. In the second case, the dimensions of the source and target domains are different.

1) Transfer Learning with the Same Dimensions

In this case, the surrogate models of areas 2, 3, and 4 are built using the proposed transfer-learning method. The accuracies of surrogate models with the same dimensions are listed in Table III. Specifically, a better surrogate model accuracy is obtained using the transfer-learning method because the characteristics of the objective functions are included in the DBN structure. Then, the common characteristics of the two objective functions are inherited through the proposed transfer-learning method. Thus, better performances are achieved by the other surrogate models. In other words, the surrogate models are based on an improved DBN structure.

TABLE III
ACCURACIES OF SURROGATE MODELS WITH THE SAME DIMENSIONS

Surrogate model	Objective function	Accuracy without transfer learning (%)	Accuracy with transfer learning (%)
Area 2	Cost	94.51	99.13
	Emission	94.90	99.25
Area 3	Cost	97.43	99.36
	Emission	92.46	99.02
Area 4	Cost	95.77	98.69
	Emission	94.38	99.61

2) Transfer Learning with Different Dimensions

In this case, the surrogate models of the combinations of areas 2 and 3, of areas 2, 3, and 4, and of all four areas are

built using the proposed transfer-learning method. The degrees of accuracy of surrogate models with different dimensions are presented in Table IV.

TABLE IV
DEGREES OF ACCURACIES OF SURROGATE MODELS WITH DIFFERENT DIMENSIONS

Surrogate model	Objective function	Accuracy without transfer learning (%)	Accuracy with transfer learning (%)
Areas 2 and 3	Cost	93.63	99.69
	Emission	93.15	99.11
Areas 2, 3, and 4	Cost	96.79	99.61
	Emission	91.54	97.80
All areas	Cost	94.36	99.20
	Emission	95.33	95.84

Compared with the last case, the dimensions of the surrogate models are different; that is, the dimensions of the domains have changed. Thus, dimension reduction is used for data preprocessing. Many familiar methods of dimension reduction, i.e., linear discriminant analysis (LDA), multidimensional scaling (MDS), isometric mapping (Iso map), landmark isometric mapping (landmark Iso map), locally linear embedding (LLE), Laplacian eigenmaps (Laplacian), local tangent space alignment (LTSA), diffusion maps, Kernel principal component analysis (Kernel PCA), stochastic neighbor embedding (SNE), and autoencoders using evolutionary optimization (autoencoder EA), are tested in this paper. The simulation results are listed in Table V.

TABLE V
ACCURACIES OF SURROGATE MODELS WITH DIFFERENT DIMENSIONS USING DIFFERENT DIMENSION REDUCTION METHODS

Dimension reduction mothod	Accuracy of areas 2 and 3 (%)		Accuracy of areas 2, 3, and 4 (%)		Accuracy of all areas (%)	
	Cost	Emission	Cost	Emission	Cost	Emission
LDA	95.28	90.70	96.09	91.67	96.11	91.02
MDS	99.69	99.11	99.61	97.80	99.20	95.84
Iso map	98.20	95.05	96.52	93.44	98.48	93.70
Landmark Iso map	96.16	90.99	97.60	91.99	98.52	96.02
LLE	99.50	97.86	99.22	95.90	98.37	96.49
Laplacian	98.67	96.33	98.46	94.85	97.82	94.25
LTSA	96.03	89.51	95.79	90.52	99.22	97.68
Diffusion maps	99.18	96.86	97.36	91.74	99.01	95.84
Kernel PCA	94.27	88.48	95.14	88.22	96.75	91.78
SNE	95.25	90.57	95.78	89.07	96.50	92.24
Autoencoder EA	97.42	93.15	97.09	91.21	95.58	89.34

In these simulations, the best performance is obtained using the MDS method. MDS is a classical data-dimensional reduction method and reconstructs the Euclidean distance coordinates between samples using a similarity matrix. In other words, MDS reconstructs the relative positions of samples in a low-dimensional space by utilizing the distances between samples in a high-dimensional space. This method preserves more high-dimensional information. Thus, MDS is suitable for dimension reduction of multiple generators. In general, this method can realize transfer learning in MACEED when the dimensions of the source and target domains are different.

C. Improved MOALO for Matching Surrogate Model

One drawback of the original MOALO method is its low convergence speed. MOALO chooses the shortest niche radius as the optimization direction, and the direction information introduced by numerical changes is not fully utilized. To overcome this drawback, a general bi-objective optimization strategy based on SDR is proposed, which enhances the convergence speed by moving to the current optimal individual. The best single dimension is saved to reduce the search uncertainty and enhance convergence speed. A flowchart of the improved MOALO is shown in Fig. 4. The SDR optimization policy is a general module for bi-objective optimization. This policy consists of eight steps.

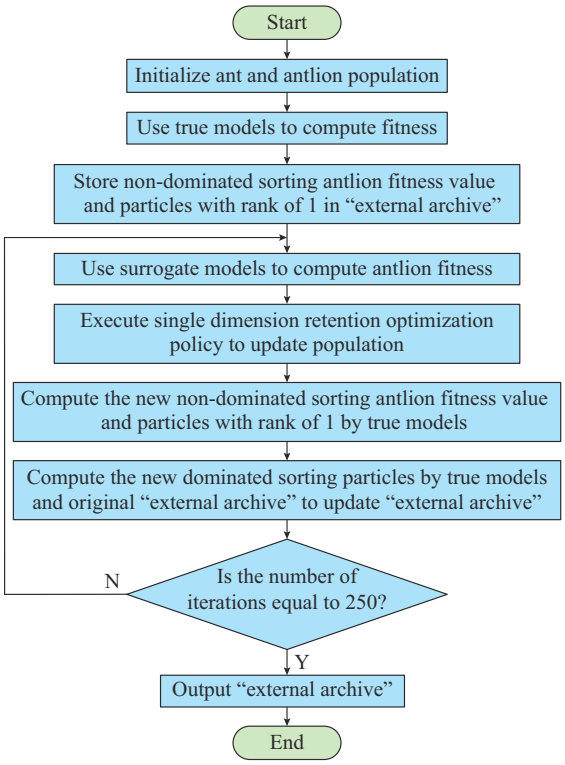

Fig. 4. Flowchart of improved MOALO.

1) The population is divided into four regions, as shown in Fig. 5 and the specific descriptions are shown in Table VI.

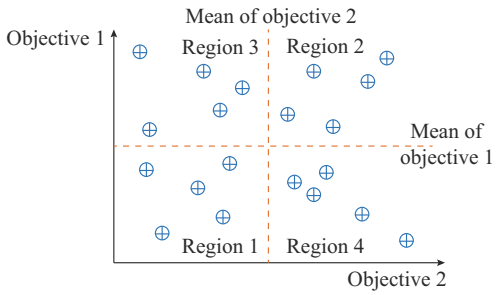

Fig. 5. Population divided into four regions.

TABLE VI
DESCRIPTIONS OF FOUR REGIONS

Region	Description
Region 1	Both fitness values are less than the corresponding current average fitness values
Region 2	Both fitness values are greater than the corresponding current average fitness values
Region 3	The fitness value of objective 2 is less than the current average fitness value of objective 2 and the fitness value of objective 1 is greater than the current average fitness value of objective 1
Region 4	The fitness value of objective 1 is less than the current average fitness value of objective 1 and the fitness value of objective 2 is greater than the current average fitness value of objective 2

2) The minimum values of the two objective functions of min_{y1} and min_{y2} are found, and the corresponding individuals are named $\boldsymbol{min}_{x1}$ and $\boldsymbol{min}_{x2}$, respectively.

3) The individuals in each region execute different operations and move to a new position $\boldsymbol{P}_1(t+1)$.

a) Individuals in region 1 execute random walking in the direction of the vector sum between $\boldsymbol{min}_{x1}$ and $\boldsymbol{min}_{x2}$.

b) Individuals in region 2 execute random walking in the direction of the vector sum between $\boldsymbol{min}_{x1}$ and $\boldsymbol{min}_{x2}$.

c) Individuals in region 3 execute random walking in the direction of $\boldsymbol{min}_{x1}$.

d) Individuals in region 4 execute random walking in the direction of $\boldsymbol{min}_{x2}$.

The mathematical model of the random walk process is:

$$X(t)=[0,\ csumu(2r(t_1)-1),\ \ldots,\ csumu(2r(t_n)-1)] \quad (12)$$

where $X(t)$ is the set of steps that individuals walk randomly; $csumu(\cdot)$ is an accumulation formula; n is the maximum number of iterations; t is the number of steps of a random walk, which can also be understood as the current number of iterations; and $r(\cdot)$ is described as:

$$r(\cdot)=\begin{cases}1 & rand>0.5\\ 0 & rand\le 0.5\end{cases} \quad (13)$$

4) The maxima of two objective functions (labeled max_{y1} and max_{y2}) are found, and the corresponding individuals are named $\boldsymbol{max}_{x1}$ and $\boldsymbol{max}_{x2}$.

5) All original populations exploit their advantages and avoid their disadvantages. This operation is described as:

$$\begin{cases}\boldsymbol{P}_2(t+1)=q_1\boldsymbol{P}(t)+q_2\boldsymbol{min}_{xU}-q_3\boldsymbol{max}_{xU}\\ q_1=0.7-0.7\cdot iter/iter_{\max}\\ q_2=0.1+0.9\cdot iter/iter_{\max}\\ q_3=0.2-0.2\cdot iter/iter_{\max}\end{cases} \quad (14)$$

where $\boldsymbol{P}_2(t+1)$ is the new position after the advantages are exploited and the disadvantages are avoided; $\boldsymbol{P}(t)$ is the original position of the individual; q_1, q_2, and q_3 are the coefficients that vary with the number of iterations; $iter$ is the current iteration number; $iter_{\max}$ is the maximum number of iterations; and $\boldsymbol{min}_{xU}$ and $\boldsymbol{max}_{xU}$ represent different vectors, where their descriptions are given in Table VII.

At the beginning of the iteration process, a considerable amount of information regarding the initial individual is retained. Thus, the initial performance is unstable, and the guidance is not obvious. As the iterations proceed, the information retained from the initial state decreases, and the information learned by high-performing particles is increasingly used. At the end of the iteration process, the operation is effectively optimized around high-performing particles. Specifically, the optimization direction is away from the worst particle and closer to the best particle.

TABLE VII
SPECIFIC DESCRIPTIONS OF $\boldsymbol{min}_{xU}$ AND $\boldsymbol{max}_{xU}$ IN FOUR REGIONS

Region	$\boldsymbol{min}_{xU}$	$\boldsymbol{max}_{xU}$
Region 1	Vector sum of $\boldsymbol{min}_{x1}$ and $\boldsymbol{min}_{x2}$	Vector sum of $\boldsymbol{max}_{x1}$ and $\boldsymbol{max}_{x2}$
Region 2	Vector sum of $\boldsymbol{min}_{x1}$ and $\boldsymbol{min}_{x2}$	Vector sum of $\boldsymbol{max}_{x1}$ and $\boldsymbol{max}_{x2}$
Region 3	$\boldsymbol{min}_{x1}$	$\boldsymbol{max}_{x1}$
Region 4	$\boldsymbol{min}_{x2}$	$\boldsymbol{max}_{x2}$

6) The individual with the maximal niche radius is used to guide other individuals through a random walking process. The third updated position for the entire population is denoted as $\boldsymbol{P}_3(t+1)$.

7) The final location $\boldsymbol{P}_4(t+1)$ is obtained by integrating $\boldsymbol{P}_1(t+1)$, $\boldsymbol{P}_2(t+1)$, and $\boldsymbol{P}_3(t+1)$ according to:

$$\begin{cases}\boldsymbol{P}_4(t+1)=f\boldsymbol{P}_1(t+1)+g\boldsymbol{P}_2(t+1)+h\boldsymbol{P}_3(t+1)\\ f=C_1+iter/iter_{\max}\\ g=0.5(C_2-iter/iter_{\max})\\ h=0.5(C_3-iter/iter_{\max})\\ C_1+C_2+C_3=1\end{cases} \quad (15)$$

where C_1, C_2, and C_3 are the constants between 0 and 1; and f, g, and h are the coefficients that vary with the iteration number. Equation (15) indicates that the information of $\boldsymbol{P}_1(t+1)$ becomes increasingly important as the iterations progress.

8) Finally, the one-dimensional retention mechanism is activated through the pseudo code in Algorithm 1, and $\boldsymbol{P}_{\text{old}}$ is the position of the last iteration. The result is $\boldsymbol{P}_{\text{new}}$, which is the final updated individual position.

Algorithm 1: one-dimensional retention mechanism

Input: $\boldsymbol{P}_4(t+1)$ and $\boldsymbol{P}_{\text{old}}$
Output: $\boldsymbol{P}_{\text{new}}$
1: **for** $j=1$: length (dimension)
2: Generate a random number between 0 and 1
3: **if** the random number is no larger than 0.2, **then**
4: $\boldsymbol{P}_{\text{new}}(j)=\boldsymbol{P}_4(t+1)(j)$
5: **else**
6: $\boldsymbol{P}_{\text{new}}(j)=\boldsymbol{P}_{\text{old}}(j)$
7: **end if**
8: **end for**
9: **return** $\boldsymbol{P}_{\text{new}}$

To further illustrate the effectiveness of the SDR-MOALO, the original MOALO [33] along with a multi-objective particle swarm optimization (MOPSO) [34], multi-objective evolutionary algorithm based on decomposition (MOEA/D)

[35], and non-dominated sorting genetic algorithm II (NSGA-II) [36] are employed. Specifically, the population sizes are set to be 100, where the other parameter settings are listed in Table VIII. In addition, the population sizes and iterations are set to be 100 and 300, respectively. Hypervolume (HV), inverted generational distance (IGD), and spread index are then utilized to compare the algorithm performances. The results of different algorithms are presented in Table IX.

TABLE VIII
PARAMETERS OF SDR-MOALO, MOALO, MOPSO, MOEA/D, AND NSGA-II

Algorithm	Parameter
SDR-MOALO	Mutation rate is set to be 0.02 and cardinality of Pareto archive is set to be 100
MOALO	Mutation rate is set to be 0.02 and cardinality of Pareto archive is set to be 100
MOPSO	The size of adaptive grid is 30, inertia weight is 0.5, learning factor c_1 is 1, and learning factor c_2 is 2
MOEA/D	Sub-problem number is set to be 20
NSGA-II	Mutation and crossover rates are 0.02 and 0.7, respectively

TABLE IX
RESULTS OF DIFFERENT ALGORITHMS ON ZDT1, ZDT2, AND ZDT3

Dataset	Algorithm	HV value	IGD value	Spread index value
ZDT1	SDR-MOALO	4.397×10^{-1}	7.814×10^{-3}	1.559×10^{-1}
	MOALO	4.704×10^{-1}	3.661×10^{-1}	1.170×10^{0}
	MOPSO	6.974×10^{-1}	2.089×10^{-2}	7.018×10^{-1}
	MOEA/D	7.035×10^{-2}	8.870×10^{-1}	1.009×10^{0}
	NSGA-II	5.562×10^{-1}	2.069×10^{-1}	8.603×10^{-1}
ZDT2	SDR-MOALO	4.398×10^{-2}	7.919×10^{-3}	1.736×10^{-1}
	MOALO	1.167×10^{-1}	4.912×10^{-1}	1.040×10^{0}
	MOPSO	2.753×10^{-1}	3.374×10^{-1}	8.300×10^{-1}
	MOEA/D	0	2.112×10^{0}	1.000×10^{0}
	NSGA-II	2.435×10^{-1}	1.942×10^{-1}	9.093×10^{-1}
ZDT3	SDR-MOALO	2.854×10^{-1}	1.129×10^{-2}	3.129×10^{-1}
	MOALO	6.748×10^{-1}	2.010×10^{-1}	1.287×10^{0}
	MOPSO	5.799×10^{-1}	3.720×10^{-2}	7.323×10^{-1}
	MOEA/D	3.597×10^{-1}	5.051×10^{-1}	1.025×10^{0}
	NSGA-II	6.924×10^{-1}	1.807×10^{-1}	8.959×10^{-1}

This section illustrates the superiority of the SDR-MOALO when tested on the ZDT1, ZDT2, and ZDT3 datasets. With the application of the one-dimensional retention mechanism, the convergence and diversity of the Pareto fronts are excellent. Compared with other algorithms, the SDR-MOALO achieves optimal results for all three indices. Specifically, a common drawback is found to exist in the four contrastive algorithms, namely, powerless local search ability in later iterations. In the one-dimensional retention mechanism, some dimensions of superior solutions are retained. Because of this mechanism, more search opportunities are obtained near the superior solutions. Undoubtedly, sufficient and precise local search capabilities emerge to improve the algorithm performance.

IV. SIMULATION EXPERIMENTS AND RESULTS

The superiority of the proposed method is demonstrated using a test system that includes 40 generators and four areas. The specific settings for the constraints in Cases 1, 2, and 3 are listed in Table X. The units of emission, fuel cost, unit output, and time are ton/hour, $/hour, MW, and s, respectively. In addition, the solutions from several previous studies are used to illustrate the effectiveness of the proposed method. For the three simulation cases, Tables I, Supplementary Material A Table SAI, and Table SAII list the settings of the cosine-similarity-based DBN + BP regression surrogate model, the coefficients of the fuel cost function [23], and the coefficients of the emission function [37], respectively.

TABLE X
SPECIFIC SETTINGS FOR CONSTRAINTS

Constraint	Case 1	Case 2	Case 3
Generation capacity constraint	√	√	√
Power balance constraint	√	√	√
Tie-line capacity limit	√	√	√
Transmission loss	×	√	√
Spinning reserve constraint between different areas	×	×	√
Tie-line capacity limits with shared spinning reserve	×	×	√

A. Case 1: Comparison and Simulation in Proposed Method Considering Power Transmission Between Different Areas

In this case, the power load demand is set to be 10500 MW. For Case 1, 20 non-dominated solutions in four areas are shown in Fig. 6.

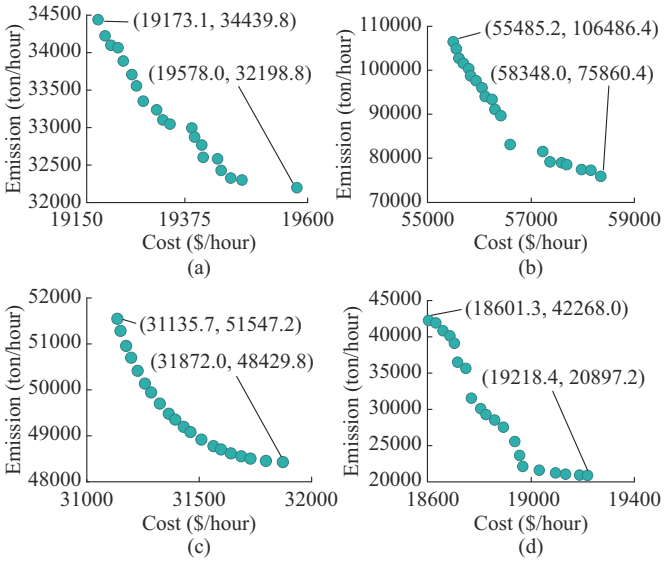

Fig. 6. Pareto fronts in Case 1. (a) Pareto front of area 1. (b) Pareto front of area 2. (c) Pareto front of area 3. (d) Pareto front of area 4.

In addition, the iterative process for the four areas is parallel, and therefore, the Pareto front of each area is generated instead of the Pareto front of all 40 generators. Feasible scheduling decisions are obtained in every sub-area. More flexible scheduling commands can then be introduced by dis-

patchers. The minimum fuel cost and emissions, obtained by summing the minimum values of the four areas, are 124395.29 $/hour and 177386.30 ton/hour, respectively. The algorithm coefficients of the symbiotic organisms search (SOS) [23], a new and efficient variety of the SOS algorithm (NSOS) [23], multi-objective competitive swarm optimizer (MOCSO) [38], and exponent-logistic-modulo map-based NSGA-III (ELM-NSGA-III) [16], are listed in Table XI. The comparisons of the results are presented in Supplementary Material A Table SAIII. The execution time of SOS and NSOS is obtained by evaluating one point on the Pareto front. Thus, the operational time is increased by 20 times to generate the entire Pareto front. The maximum, minimum, and average results for Case 1 are presented in Table XII.

TABLE XI
PARAMETERS OF SDR-MOALO, SOS, NSOS, MOCSO, AND ELM-NSGA-III

Algorithm	Parameter
SDR-MOALO	The maximum iteration is 300, mutation rate is 0.02, and cardinality of Pareto archive is 20
SOS	The maximum iteration is 100000, trial values for O are [50, 60, 70], trial values for k_{max} are [100, 150, 200], selected O is 70, and selected k_{max} is 200
NSOS	The maximum iteration is 100000, trial values for O are [50, 75, 100], trial values for k_{max} are [100, 150, 200], selected O is 60, and selected k_{max} is 100
MOCSO	The number of iteration is 6000, crossover probability P_v is 0.85, penalty coefficients p_p/p_{p2} is 40/78, $p_{p,1}/p_{p2,1}$ is 18/80, $p_{p,2}/p_{p2,2}$ is 28/50, $p_{p,3}/p_{p2,3}$ is 28/50, and $p_{p,4}/p_{p2,4}$ is 18/80
ELM-NSGA-III	The number of iterations, crossover probability, mutation probability, crossover distribution index, variation step, and the size of the elite archive are 500, 1.0, 0.01, 30, 20, and 20, respectively

TABLE XII
THE MAXIMUM, MINIMUM, AND AVERAGE RESULTS FOR CASE 1

Item	Best fuel cost ($/hour)	Best emission (ton/hour)	Time (s)
The maximum	125768.36	185864.58	36.69
The minimum	124395.29	177386.30	29.68
Average	124789.04	180169.68	32.55

First, compared with SOS, NSOS, and MOCSO, the fuel costs are reduced by 0.77%, 0.84%, and 0.05%, respectively, and the emissions are reduced by 15.01%, 12.83%, and 24.20%, respectively. More importantly, the operational time is reduced by 92.30%, 92.60%, and 82.52%, respectively. Compared with ELM-NSGA-III, the emission index is reduced by 21112.28 ton/hour (10.63%), the execution time is reduced by 0.81 s (2.51%), and the cost is nearly the same. Thus, the proposed method produces the lowest costs, emissions, and operational time. In addition, it provides more flexible and extensive decision-making space for scheduling in other areas.

We also consider how superior solutions are obtained. First, compared with the other algorithms, the fuel cost value shows a slight improvement due to the sufficient search capabilities of using a one-dimensional retention mechanism in the SDR-MOALO. Better solutions can be found by retaining the values in several dimensions. Thus, better solutions can be obtained by one-dimensional retention mechanism.

Next, we focus on the execution time, where only 31.47 s are consumed in Case 1. Compared with SOS, NSOS, MOCSO, and ELM-NSGA-III, the execution time is improved by 92.30%, 92.60%, 82.52%, and 2.51%, respectively. The computational time of the objective values is considerably shortened by the surrogate models. This also results from using transfer learning. Transfer learning then further reduces the modeling time. Thus, the operational time for MACEED is drastically reduced.

B. Case 2: Comparison and Simulation in Proposed Method Considering Power Transmission and Losses

The power demand for Case 2 is set to be 10500 MW. To further verify the effectiveness of the proposed method, active loss is considered as an additional constraint. For Case 2, 20 non-dominated solutions in four areas are shown in Fig. 7. The outputs of 40 units are listed in Supplementary Material A Table SAIV. Again, the optimization process for the four areas is parallel, and therefore the Pareto front of each area is generated instead of that of all 40 generators. The best fuel cost and emissions, obtained by summing the best values of the four areas, are 125035.51 $/hour and 186439.80 ton/hour, respectively. The losses are listed in Supplementary Material A Table SAIV. The execution time is only 31.85 s. More importantly, compared with SOS, NSOS, MOCSO, and ELM-NSGA-III for Case 1, the execution time is reduced by 92.21%, 92.51%, 82.31%, and 1.33%, respectively. Even when more complex constraints are considered, the proposed method can solve MACEED problems in a short time.

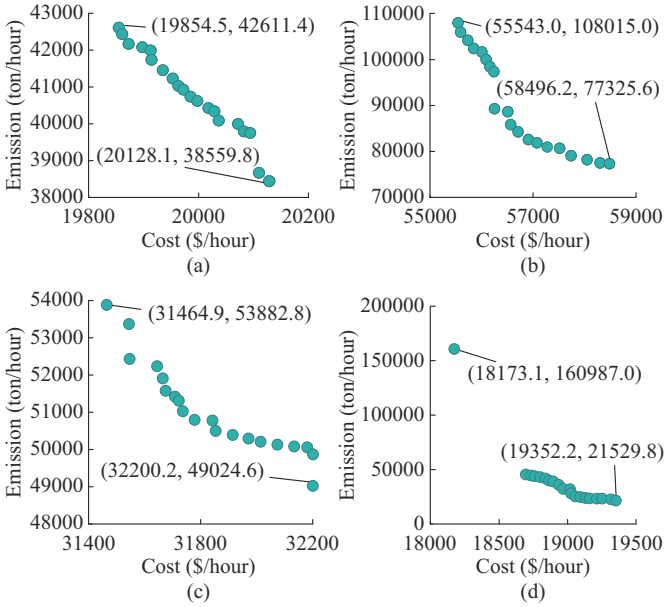

Fig. 7. Pareto fronts in Case 2. (a) Pareto front of area 1. (b) Pareto front of area 2. (c) Pareto front of area 3. (d) Pareto front of area 4.

Next, we examine the performance of the SDR-MOALO. The maximum, minimum, and average results for Case 2 are listed in Table XIII. These values reveal that high-dimensional MACEED problems with complex constraints in large-

scale power systems can be solved accurately and in a stable manner using the proposed method. Because of the search capabilities of the SDR-MOALO, a convergent Pareto front can be obtained. The computational time of the DBN + BP surrogate model is shorter than that of the original mathematical function. As shown in Table XIV, compared with ELM-NSGA-III, the cost, emissions, and average time are reduced by 0.37%, 2.60%, and 9.47%, respectively. With parallel computing, convergence is improved, and the execution time is reduced.

TABLE XIII
THE MAXIMUM, MINIMUM, AND AVERAGE RESULTS FOR CASE 2

Item	Best fuel cost ($/hour)	Best emission (ton/hour)	Time (s)
The maximum	126003.84	188897.19	38.73
The minimum	125035.51	186439.80	31.28
Average	125558.08	187454.64	33.08

TABLE XIV
COMPARATIVE SOLUTIONS FOR CASE 2

Algorithm	Best fuel cost ($/hour)	Best emission (ton/hour)	Average time (s)
SDR-MOALO	125035.51	188897.19	33.08
ELM-NSGA-III	125509.09	193929.85	36.54

C. Case 3: Comparison and Simulation in Proposed Method Considering Spinning Reserve and All Mentioned Constraints

A total of 10 non-dominated solutions in four areas of Case 3 are shown in Fig. 8, and the best outputs are listed in Supplementary Material A Table SAV. Again, the iterative process for the four areas is parallel, and therefore the Pareto front of each area is generated instead of that of all 40 generators. The best fuel cost and emissions are 108353.46 $/hour and only 166302.27 ton/hour, respectively. The corresponding losses are both 62.76 MW. The execution time is only 32.46 s. Compared with the SOS, NSOS, and MOCSO of Case 1, with a higher power load, the execution time is reduced by 92.10%, 92.37%, and 81.97%, respectively. This case represents a real scheduling scenario and demonstrates that the use of surrogate models and transfer learning can significantly reduce the execution time of computationally expensive MACEED problems while deriving excellent feasible solutions.

We next consider the performance of the SDR-MOALO. The results are presented in Table XV and show that the SDR-MOALO is effective at solving high-dimensional MACEED problems with all proposed constraints. Because of the superior search capabilities of the SDR-MOALO, the convergent Pareto front in MACEED problems with all constraints can be obtained. In addition, the short execution time illustrates that the model is also feasible in addressing MACEED problems with all constraints.

Furthermore, transfer learning is feasible for quickly building surrogate models for MACEED problems with flexible constraints. Compared with the results of Case 3 in [16], the best fuel cost is reduced by 0.34%. Moreover, even in the most time-consuming state, the execution time is reduced by 10.55%. This means that the proposed method can solve MACEED problems under all constraints in a stable manner. In summary, MACEED problems under all constraints can be solved in a timely manner.

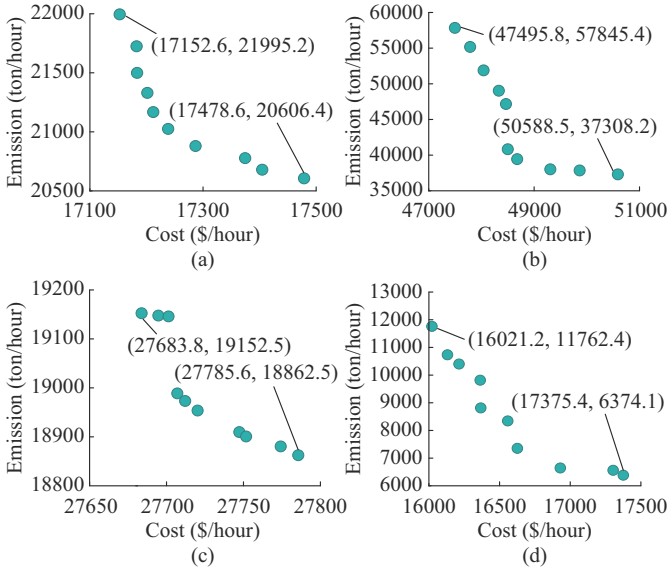

Fig. 8. Pareto fronts in Case 3. (a) Pareto front of area 1. (b) Pareto front of area 2. (c) Pareto front of area 3. (d) Pareto front of area 4.

TABLE XV
THE MAXIMUM, MINIMUM, AND AVERAGE RESULTS FOR CASE 3

Item	Best fuel cost ($/hour)	Best emission (ton/hour)	Time (s)
The maximum	110525.77	169974.31	35.77
The minimum	108353.46	166302.27	31.98
Average	109025.87	167288.84	33.19

D. Case 4: Comparison and Simulation with Gurobi and Cplex Optimizers

To further illustrate the effectiveness of the proposed method, two commercial solvers, i.e., Gurobi [39] and Cplex [40], are applied to the aforementioned three cases.

The two optimizers have two limitations:

1) They cannot directly obtain a real Pareto front for the multi-objective optimization problem [41], [42]. To achieve non-dominated solutions, the commonly used method is to reconstruct objective functions using the weight-sum method:

$$F(P)=\omega C(P)+\gamma(1-\omega)E(P) \tag{16}$$

where $F(P)$ is the single objective function obtained by the weight sum method; $C(P)$ and $E(P)$ are the fuel cost and emission functions, respectively; and ω and γ are the weight and scaling factors, respectively. In Cases 1 and 2, to obtain a Pareto front with 20 non-dominated solutions, ω changes from 1 to 0 with a step size of 0.05, and γ is set to be 0.5. In Case 3, to obtain a Pareto front with 10 non-dominated solutions, ω changes from 1 to 0 with a step size of 0.1, and γ is set to be 0.5.

2) They do not directly support general non-linearities, whether in objective functions or constraints [43]. Thus, the

cost function with the valve-point effect should be transformed into piecewise linear functions. To obtain feasible solutions with high accuracy, the number of segments in piecewise linear functions should be as large as possible. We test the number of segments from 5 to 30. If the number of segments is less than 19, generating a Pareto front is nearly impossible. Thus, in these three cases, the number of segments is set to be 19. The piecewise linear function for these three cases is described by:

$$\Delta P=(P_{\max}-P_{\min})/19 \tag{17}$$

$$C_{\text{piece}}(P)=\begin{cases} C_{\text{picce},1}(P) & P_{\min}\le P\le P_{\min}+\Delta P \\ C_{\text{piece},2}(P) & P_{\min}+\Delta P<P\le P_{\min}+2\Delta P \\ \vdots & \\ C_{\text{picce},i}(P) & P_{\min}+(i-1)\Delta P<P\le P_{\min}+i\Delta P \\ \vdots & \\ C_{\text{picce},19}(P) & P_{\min}+18\Delta P<P\le P_{\max} \end{cases} \tag{18}$$

where ΔP is the length of P in a segment; $C_{\text{piece}}(P)$ is the piecewise linear fuel cost function; and $C_{\text{picce},i}(P)$ is the i^{th} segment of $C_{\text{piece}}(P)$. Then, $C(P)$ can be expressed as:

$$\begin{cases} C(P)=\sum_{i=1}^{19} b_i C_{\text{piece},i}(P) \\ b_i\in\{0,1\} \\ \sum_{i=1}^{19} b_i=1 \end{cases} \tag{19}$$

where b_i is a 0/1 variable, $b_i=1$ means that $C_{\text{piece},i}(P)$ is selected, and $\sum_{i=1}^{19} b_i=1$ means that only one b_i can be set to be 1, where the other b_i must be set to be 0. The Gurobi and Cplex optimizers can then be utilized to simulate the aforementioned three cases, where the results are listed in Table XVI.

TABLE XVI
COMPARATIVE RESULTS OF GUROBI AND CPLEX SOLVERS AND PROPOSED METHOD

Case index	Best fuel cost ($/hour)			Best emission (ton/hour)			Execution time (s)		
	Proposed	Gurobi	Cplex	Proposed	Gurobi	Cplex	Proposed	Gurobi	Cplex
1	124395.29	124511.00	124542.92	177386.30	234694.90	248908.89	32.55	52.78	129.67
2	125035.51	125687.30	125592.26	188897.19	244997.10	263642.48	33.08	184.18	116.13
3	108353.46	108446.31	108448.25	166302.27	215621.08	215356.38	33.19	113.73	43.73

Compared with the Gurobi solver, for Case 1, the best fuel cost and emissions as well as average execution time obtained by the proposed method are reduced by 0.09%, 24.41%, and 38.33%, respectively. For Case 2, these three indices are reduced by 0.52%, 22.90%, and 82.04%, respectively. For Case 3, the best fuel cost and emissions as well as average execution time are reduced by 0.09%, 22.87%, and 70.82%, respectively.

Compared with the Cplex solver, for Case 1, the best fuel cost and emissions as well as average execution time obtained by the proposed method are reduced by 0.12%, 28.73%, and 74.90%, respectively. For Case 2, these three indices are reduced by 0.44%, 28.35%, and 71.51%, respectively. For Case 3, the best fuel cost and emissions as well as average execution time are reduced by 0.09%, 22.78%, and 24.10%, respectively. These excellent results demonstrate that the proposed method is effective in solving computationally expensive MACEED problems. In addition, non-convex and multiple optimization objective functions can be solved quickly and accurately. The surrogate model is thus a powerful tool for dealing with non-convex and computationally expensive objective functions. However, the SDR-MOALO is suitable for solving computationally expensive MACEED problems.

Naturally, the accuracy of the Gurobi and Cplex solvers can be improved by dividing ω more fully (for example, ω changes from 1 to 0 with a step size of 0.01, 0.001) or by enhancing the number of segments in piecewise linear functions. However, these changes enhance the execution time, which may exceed the scheduling cycles. Compared with Gurobi and Cplex, the proposed method significantly reduces the execution time without sacrificing the accuracy of the solutions.

V. CONCLUSION AND FUTURE RESEARCH

This paper describes a new method for solving MACEED problems with high-dimensional decision variables in large-scale power systems. Cosine-similarity-based DBN + BP regression surrogate models are proposed to replace the two objective functions in MACEED. Compared with the original numerical objective functions, the execution time is reduced using the DBN + BP models. The transfer-learning method is then utilized to pretrain and fine-tune the cosine-similarity-based DBN + BP models. Based on the initial surrogate model, data from other areas are used for fine-tuning, significantly reducing the time required to build the surrogate models of different areas. An improved MOALO that matches the surrogate models is introduced to obtain optimization results for computationally expensive MACEED problems. General bi-objective optimization is used to accelerate convergence, allowing the improved MOALO to find the Pareto fronts in different areas. The advantages of this method are verified by a 4-area 40-unit test system, where simulation results demonstrate the effectiveness of the proposed method. Building surrogate models to replace time-consuming constraint functions is also feasible and deserves meaningful future study. MACEED problems are intraday scheduling problems. Therefore, another interesting direction is to use a data-driven method to investigate day-ahead unit commitment problems with 0/1 integer variables [44], [45]. Future research should include the study of data-driven mixed-

integer programming in unit commitment problems.

REFERENCES

[1] A. Marot, A. Kelly, M. Naglic *et al.*, "Perspectives on future power system control centers for energy transition," *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 2, pp. 328-344, Mar. 2022.

[2] F. Qi, M. Shahidehpour, F. Wen *et al.*, "Decentralized privacy- preserving operation of multi-area integrated electricity and natural gas systems with renewable energy resources," *IEEE Transactions on Sustainable Energy*, vol. 11, no. 3, pp. 1785-1796, Jul. 2019.

[3] S. Xu, W. Wu, B. Wang *et al*., "Probabilistic energy and reserve co-dispatch for high-renewable power systems and its convex reformulation," *Journal of Modern Power Systems and Clean Energy*, vol. 11, no. 6, pp. 1734-1745, Nov. 2023.

[4] S. Krishnamurthy and R. Tzoneva, "Decomposition-coordinating method for parallel solution of a multi-area combined economic emission dispatch problem," *International Journal of Electrical and Computer Engineering*, vol. 6, no. 5, pp. 2048-2063, Oct. 2016.

[5] F. Jamshidian, K. C. Liu, and R. Lugtu, "Penalty factor calculations incorporating interchange constraints," *IEEE Transactions on Power Apparatus and Systems*, vol. 102, no. 11, pp. 3667-3671, Nov. 1983.

[6] C. Yingvivatanapong, W. J. Lee, and E. Liu, "Multi-area power generation dispatch in competitive markets," *IEEE Transactions on Power Systems*, vol. 23, no. 1, pp. 196-203, Feb. 2008.

[7] Y. Zhang, Y. Han, D. Liu *et al*., "Low-carbon economic dispatch of electricity-heat-gas integrated energy systems based on deep reinforcement learning," *Journal of Modern Power Systems and Clean Energy*, vol. 11, no. 6, pp. 1827-1841, Nov. 2023.

[8] V. P. Sakthivel and P. D. Sathya, "Single and multi-area multi-fuel economic dispatch using a fuzzified squirrel search algorithm," *Protection and Control of Modern Power Systems*, vol. 6, no. 1, pp. 1-13, Apr. 2021.

[9] V. P. Sakthivel and P. D. Sathya, "Multi-area economic environmental dispatch using multi-objective squirrel search algorithm," *Evolving Systems*, vol. 13, no. 2, pp. 183-199, Apr. 2022.

[10] A. Azizivahed, S. E. Razavi, A. Arefi *et al.*, "Risk-oriented multi-area economic dispatch solution with high penetration of wind power generation and compressed air energy storage system," *IEEE Transactions on Sustainable Energy*, vol. 11, no. 3, pp. 1569-1578, Jul. 2019.

[11] V. P. Sakthivel, M. Suman, and P. D. Sathya, "Combined economic and emission power dispatch problems through multi-objective squirrel search algorithm," *Applied Soft Computing*, vol. 100, p. 106950, Mar. 2021.

[12] M. Yu, J. Liang, Z. Wu *et al.*, "A twofold infill criterion-driven heterogeneous ensemble surrogate-assisted evolutionary algorithm for computationally expensive problems," *Knowledge-based Systems*, vol. 236, p. 107747, Jan. 2022.

[13] P. Yang, K. Tang, and X. Yao, "Turning high-dimensional optimization into computationally expensive optimization," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 1, pp. 143-156, Feb. 2017.

[14] H. Wang and Y. Jin, "A random forest-assisted evolutionary algorithm for data-driven constrained multi-objective combinatorial optimization of trauma systems," *IEEE Transactions on Cybernetics*, vol. 50, no. 2, pp. 536-549, Feb. 2020.

[15] C. He, Y. Tian, H. Wang *et al.*, "A repository of real-world datasets for data-driven evolutionary multiobjective optimization," *Complex & Intelligent Systems*, vol. 6, no. 1, pp. 189-197, Nov. 2019.

[16] C. Lin, H. Liang, and A. Pang, "A fast data-driven optimization method of multi-area combined economic emission dispatch," *Applied Energy*, vol. 337, p. 120884, May 2023.

[17] Y. Pan, R. Hong, J. Chen *et al.*, "A hybrid DBN-SOM-PF-based prognostic approach of remaining useful life for wind turbine gearbox," *Renewable Energy*, vol. 152, pp. 138-154, Jun. 2020.

[18] D. S. Irene, T. Sethukarasi, and N. Vadivelan, "Heart disease prediction using hybrid fuzzy *K*-medoids attribute weighting method with DBN- KELM based regression model," *Medical Hypotheses*, vol. 143, p. 110072, Oct. 2020.

[19] X. Ji, Y. Zhang, D. W. Gong *et al.*, "Multisurrogate-assisted multitasking particle swarm optimization for expensive multimodal problems," *IEEE Transactions on Cybernetics*, vol. 53, no. 4, pp. 2516-2530, Nov. 2021.

[20] X. Ji, Y. Zhang, D. W. Gong *et al.*, "Dual-surrogate-assisted cooperative particle swarm optimization for expensive multimodal problems," *IEEE Transactions on Evolutionary Computation*, vol. 25, no. 4, pp. 794-808, Aug. 2021.

[21] M. Byra, "Breast mass classification with transfer learning based on scaling of deep representations," *Biomedical Signal Processing and Control*, vol. 69, p. 102828, Aug. 2021.

[22] A. Mahbod, G. Schaefer, C. Wang *et al.*, "Transfer learning using a multi-scale and multi-network ensemble for skin lesion classification," *Computer Methods and Programs in Biomedicine*, vol. 193, p. 105475, Sept. 2020.

[23] N. Sinha, R. Chakrabarti, and P. K. Chattopadhyay, "Evolutionary programming techniques for economic load dispatch," *IEEE Transactions on Evolutionary Computation*, vol. 7, no. 1, pp. 83-94, Feb. 2003.

[24] M. Ghasemi, J. Aghaei, E. Akbari *et al.*, "A differential evolution particle swarm optimizer for various types of multi-area economic dispatch problems," *Energy*, vol. 107, pp. 182-195, Jun. 2016.

[25] V. Berisha, C. Krantsevich, P. R. Hahn *et al.*, "Digital medicine and the curse of dimensionality," *NPJ Digital Medicine*, vol. 4, no. 1, p.153, Oct. 2021.

[26] Y. Zhou, Q. Zhai, and L. Wu, "Multistage transmission-constrained unit commitment with renewable energy and energy storage: implicit and explicit decision methods," *IEEE Transactions on Sustainable Energy*, vol. 12, no. 2, pp. 1032-1043, Apr. 2021.

[27] N. Yang, Z. Dong, L. Wu *et al.*, "A comprehensive review of security-constrained unit commitment," *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 3, pp. 562-576, May 2021.

[28] M. Zare, M. R. Narimani, M. Malekpour *et al.*, "Reserve constrained dynamic economic dispatch in multi-area power systems: an improved fireworks algorithm," *International Journal of Electrical Power & Energy Systems*, vol. 126, p. 106579, Mar. 2021.

[29] B. Ibrokhimov, C. Hur, H. Kim *et al.*, "A-DBNF: adaptive deep belief network framework for regression and classification tasks," *Applied Intelligence*, vol. 51, no. 7, pp. 4199-4213, Jan. 2021.

[30] H. Sharifzadeh, "Sharp formulations of nonconvex piecewise linear functions to solve the economic dispatch problem with valve-point effects," *International Journal of Electrical Power & Energy Systems*, vol. 127, p. 106603, May 2021.

[31] P. Chen and H. Chang, "Large-scale economic dispatch by genetic algorithm," *IEEE Transactions on Power Systems*, vol. 10, no. 4, pp. 1919-1926, Nov. 1995.

[32] P. Westermann and R. Evins, "Surrogate modelling for sustainable building design – a review," *Energy and Buildings*, vol. 198, pp. 170-186, Sept. 2019.

[33] S. Mirjalili, P. Jangir, and S. Saremi, "Multi-objective ant lion optimizer: a multi-objective optimization algorithm for solving engineering problems," *Applied Intelligence*, vol. 46, pp. 79-95, Jul. 2017.

[34] Y. Cui, X. Meng, and J. Qiao, "A multi-objective particle swarm optimization algorithm based on two-archive mechanism," *Applied Soft Computing*, vol. 119, p. 108532, Apr. 2022.

[35] R. Carvalho, R. R. Saldanha, B. N. Gomes *et al.*, "A multi-objective evolutionary algorithm based on decomposition for optimal design of Yagi-Uda antennas," *IEEE Transactions on Magnetics*, vol. 48, no. 2, pp. 803-806, Jan. 2012.

[36] K. Deb, S. Agrawal, A. Pratap *et al.*, "A fast and elitist multi-objective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 849-858, Apr. 2002.

[37] M. Basu, "Economic environmental dispatch using multi-objective differential evolution," *Applied Soft Computing*, vol. 11, no. 2, pp. 2845-2853, Mar. 2011.

[38] S. Weng, W. Tan, B. Ou *et al.*, "Reversible data hiding method for multi-histogram point selection based on improved crisscross optimization algorithm," *Information Sciences*, vol. 549, pp. 13-33, Mar. 2021.

[39] X. Huo, X. Wu, Y. Fan *et al.*, "A mixed-integer program (MIP) for one-way multiple-type shared electric vehicles allocation with uncertain demand," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 8972-8984, Jul. 2022.

[40] Z. Ma, H. Zhong, T. Cheng *et al.*, "Redundant and nonbinding transmission constraints identification method combining physical and economic insights of unit commitment," *IEEE Transactions on Power Systems*, vol. 36, no. 4, pp. 3487-3495, Jul. 2021.

[41] X. Zuo, C. C. Murray, and A. E. Smith, "Solving an extended double row layout problem using multiobjective tabu search and linear programming," *IEEE Transactions on Automation Science and Engineering*, vol. 11, no. 4, pp. 1122-1132, Oct. 2014.

[42] L. Duan, Y. Liu, H. Li *et al*., "Nondominated sorting differential evolution algorithm to solve the biobjective multi-AGV routing problem in hazardous chemicals warehouse," *Mathematical Problems in Engineering*, vol. 2022, p. 3785039, Sept. 2022.

[43] F. Wu and M. Dong, "Eco-routing problem for the delivery of perishable products," *Computers & Operations Research*, vol. 154, p.

106198, Jun. 2023.
[44] Z. Yang, K. Li, Q. Niu *et al.*, "A comprehensive study of economic unit commitment of power systems integrating various renewable generations and plug-in electric vehicles," *Energy Conversion and Management*, vol. 132, pp. 460-481, Jan. 2017.
[45] Z. Yang, K. Li, Q. Niu *et al.*, "A novel parallel-series hybrid meta-heuristic method for solving a hybrid unit commitment problem," *Knowledge-based Systems*, vol. 134, pp. 13-30, Oct. 2017.

Chenhao Lin received the B.E. degree in electrical engineering and its automation from North China Electric Power University, Baoding, China, in 2018. He is currently pursuing the M.E. degree in electrical engineering in Hubei Minzu University, Enshi, China. His current research interests include computational intelligence and optimization in power systems.

Huijun Liang received the B.E. degree in industrial automation from Wuhan University of Hydraulic and Electric Engineering, Wuhan, China, in 2000, the M.E. degree in control theory and control engineering from Wuhan University, Wuhan, China, in 2004, the Ph.D. degree in control theory and control engineering from Shandong University, Jinan, China, in 2020. He is currently working as a Lecturer in Hubei Minzu University, Enshi, China. His research interests include computational intelligence and optimization in power systems.

Aokang Pang received the B.E. degree in electrical engineering and its automation from Xi'an Kedagaoxin University, Xi'an, China, in 2021. He is currently pursuing the M.E. degree in electrical engineering in Hubei Minzu University, Enshi, China. His current research interests include computational intelligence and its application to economic dispatch of power systems.

Jianwei Zhong received the B.E. degree in physics from Hubei Institute for Nationalities, Enshi, China, in 1994, and the M.E. degree in material processing engineering from Huazhong University of Science and Technology, Wuhan, China, in 2004. He is currently working as a Professor in Hubei Minzu University, Enshi, China. His research interests include power system operation and new energy generation.

Yongchao Yang received the B.E. degree in electrical engineering and its automation from Hubei Institute for Nationalities, Enshi, China, in 2003, the M.E. degree in instruments science and technology from Chongqing University, Chongqing, China, in 2016, and the Ph.D. degree in electrical engineering from Huazhong University of Science and Technology, Wuhan, China, in 2004. He is currently working as an Associate Professor in Hubei Minzu University, Enshi, China. His research interests include high-voltage and insulation technology, power equipment operating condition monitoring and fault diagnosis.