[1]Zheng Wang
[1]Yi Qi
[1]Maolin Tang
[1,*]Xiaomin Liu
[1]Bin Bai

Integrated Cost Prediction for Distribution Network Engineering Based on the Novel Multi-objective Coati Algorithm and Laguerre Polynomials

Journal of Electrical Systems

Abstract: - To achieve precise and stable cost prediction for distribution network engineering, a novel multi-objective coati optimization algorithm (MOCOA) is proposed, and based on this algorithm, a new multi-objective integrated prediction model based on Laguerre polynomials (LNN-Adaboost) is developed. Firstly, utilizing the nonlinear approximation properties of Laguerre orthogonal polynomials, a Laguerre neural network is constructed. Secondly, targeting prediction accuracy and stability, the Laguerre neural network is optimized using the multi-objective coati algorithm. Finally, ensemble learning Adaboost is introduced to correct the prediction errors of the model, achieving automatic allocation and recombination of error weights. By comparing with three mainstream multi-objective optimization algorithms on eight test problems, the effectiveness of the proposed algorithm is verified. Taking the cost data of the Ningxia distribution network engineering as the research object, the proposed model MOCOA-LNN-Adaboost is compared with several mainstream models. The results demonstrate that the proposed prediction model exhibits higher prediction accuracy and better prediction stability. Compared to MOCOA-ELM-Adaboost, the proposed prediction model shows an increase of 9.64% in R^2, 2.65% in IA, and a decrease of 53.62% in RMSE, and 61.87% in SDEX..

Keywords:Multi-objective optimization; Laguerre polynomial; Integrated prediction; Distribution network engineering.

I. INTRODUCTION

Distribution network engineering, as a crucial component of energy systems, plays a vital role in urbanization and industrialization processes[1-2] . However, with the continuous growth in energy demand, optimization of energy structure, technological advancements, and societal changes, the construction and operation costs of distribution network engineering have been escalating. Hence, accurate cost prediction for distribution network engineering has become imperative[3-4]. Cost prediction in distribution network engineering is essential for project planning, budgeting, and resource allocation. Accurate cost predictions aid decision-makers in early-stage resource allocation, preventing project delays due to insufficient funding or wastage.

Currently, researchers have made progress in cost prediction for distribution network engineering, mainly focusing on statistical learning[5], linear regression [6], and deep learning [7]. Ye et al. [8] utilized a BP neural network to predict the construction costs of building projects, introducing a particle swarm optimization algorithm to optimize the weights and thresholds of the BP neural network. This enhancement improved the model's generalization ability and enhanced prediction accuracy. Skitmore et al. [9]developed standard regression and cross-validation regression prediction models to forecast the actual construction costs of projects, using data from 93 Australian construction projects and selecting contracts, procurement, and contractors as model inputs. Batselieret al.[10] introduced exponential smoothing prediction methods to forecast project time and costs. By utilizing past performance information and expected management actions as data drivers, a smoothed prediction with similar forecasting concepts was compared with several prediction models, resulting in improved model performance. Zhou et al.[11] aimed to reduce construction costs by employing support vector machine (SVM) to predict tunnel boring machine penetration rates. They optimized SVM hyperparameters using whale optimization algorithm (WOA), grey wolf optimization Algorithm (GWO), and moth-flame optimization Algorithm (MFO). Using 1286 samples from a water conveyance tunnel in Malaysia and evaluating the model with four metrics, the study concluded that the mixed SVM model was suitable for tunnel boring machine problems, demonstrating good prediction accuracy. Xu et al. [12]forecasted construction cost indices by establishing a cointegrated vector autoregressive model, enabling predictions of construction budgets. Fan et al.[13] introduced SVM with least squares support vector machine (LSSVM) to improve engineering cost prediction accuracy. After dimensionality reduction of raw data, the optimal model was selected based on the output of two models, with prediction errors of less than 7%, providing a new approach to cost prediction. KOO et al. [14]developed a case-based reasoning

[1]Economic and Technological ResearchInstitute of State Grid Ningxia ElectricPower Co., Ltd., Yinchuan, 750002, China
*Correspondingauthor:Xiaomin Liu

(CBR) model to mitigate the challenges posed by uncertain information in construction cost prediction. This model combined multiple regression (MR) analysis and artificial neural networks (ANN), with optimization of four model hyperparameters using genetic algorithm (GA). Validated with data from 101 residential projects, the model demonstrated accurate early-stage construction cost predictions. Tatiya et al. [15]observed inaccuracies in current cost prediction methods in the deconstruction domain and proposed integrating artificial intelligence methods for precise cost prediction, providing decision-makers with robust data support. This prediction model also integrated MR, ANN, and CBR components and was validated with deconstruction cases from Michigan, showing a prediction accuracy rate of 95% for the designed model. Additionally, various optimization algorithms were employed to improve the performance of prediction models. However, most existing studies focus on single-objective prediction models aimed at improving prediction accuracy, overlooking the importance of enhancing model stability. Considering the stability of cost prediction results helps improve prediction accuracy at "peak" points, facilitating advanced assessment and effective cost control of distribution network engineering projects. Thus, accurate and stable cost prediction for distribution network engineering is crucial for early-stage resource allocation, necessitating the development of prediction models that balance high prediction accuracy with strong stability.

Furthermore, current research predominantly employs traditional intelligent optimization algorithms to optimize parameters of single-feedforward neural network prediction models, such as BP neural networks, extreme learning machines, and support vectors, lacking exploration of new prediction models for distribution network engineering. Lastly, single-feedforward neural network prediction models themselves suffer from insufficient generalization ability and susceptibility to overfitting, often failing to capture predictive sequence characteristics effectively[16]. Therefore, it is of great significance to conduct precise cost prediction studies for distribution network engineering by combining novel optimization algorithms with complex ensemble models.

Multi-objective optimization algorithms can be broadly categorized into two types, inspired by animal behaviors and strategies based on classic mathematical methods. Among them, a new metaheuristic algorithm proposed in a previous study mimics the unique mating behavior of snakes[17], using competitive mate selection to continuously update individuals, drawing inspiration from snake reproductive strategies. Another approach, inspired by the survival strategies of gazelles[18], analogizes their behaviors of evading predators and seeking refuge to strategies for solving optimization problems, emphasizing the wisdom of gazelles in survival. Additionally, the INFO algorithm[19], introduced in another study, incorporates mathematical weighted averaging techniques to enhance algorithm convergence speed and exploration and development capabilities. This method includes vector fusion and local search operators, which utilize mathematical weighted averaging principles to enable more effective handling of multi-objective optimization problems. However, despite the advantages of these algorithms, the 'no free lunch theorem' emphasizes that no single algorithm is suitable for all optimization problems; thus, problem characteristics and requirements are crucial. Inspired by the hunting and escaping strategies of coati, another study observed these two survival behaviors and proposed the coati optimization algorithm (COA)[20]. Experimental results demonstrate the outstanding performance of the COA algorithm, particularly in simulated experiments. Given that most real-world problems belong to the domain of multi-objective optimization problems, extending the excellent performance of the COA algorithm to the multi-objective domain is essential for better solving complex real-world problems.

To address the aforementioned research challenges, this study first proposes a new multi-objective coati optimization algorithm (MOCOA), and then, based on this algorithm, presents a multi-objective integrated prediction model for distribution network engineering cost based on Laguerre polynomials. To overcome the drawbacks of single-objective prediction models, the proposed model can simultaneously achieve high prediction accuracy and strong prediction stability for distribution network engineering cost. Leveraging the strong nonlinear fitting properties of Laguerre orthogonal polynomials, Laguerre orthogonal basis functions are constructed to further create a Laguerre neural network (LNN) and establish a cost prediction model for distribution network engineering. To address the insufficient generalization ability of feedforward neural networks, ensemble learning Adaboost is introduced to allocate and recombine prediction errors of the model, further enhancing prediction accuracy and obtaining the final prediction model. Compared to current distribution network engineering cost prediction models, the novel integrated prediction model optimized by the proposed optimization algorithm exhibits higher prediction accuracy and better prediction stability.

This research is structured as follows: Section 2 presents the theoretical framework and strategies of the model. Section 3 provides a description of the sample data. Section 4 consists of simulation and discussion, while Section 5 concludes the study.

II. METHOD AND STRATEGY

A. Coati optimization algorithm(COA)

The coati algorithm can be divided into two main stages, each simulating different survival strategies of coati groups. The first stage is the exploration stage. In this stage, the coati algorithm mimics the hunting process of coatis for iguanas. Coatis are intelligent hunters that use clever strategies to catch prey. The focus of this stage is to search for the optimal solution, much like coatis searching for prey while hunting. The second stage is the exploitation stage. In the exploitation stage, the coati algorithm simulates the process of coati groups escaping from predators. Coatis need to constantly evade predators in the wild, so this stage emphasizes strategies for finding shelter and avoiding danger.

1)Exploration phase

Firstly, in the initial stage of hunting, half of the members in the coati pack stealthily climb up to the treetops, preparing to initiate their actions. Their common objective is to force the iguana off the tree, typically achieved through various threats or encirclement methods. This may involve coordinated actions among multiple coatis to create sufficient threat, compelling the prey to react and eventually descend to the ground. The process of climbing up the tree can be defined by the following equation[20]:

$$X_i^{P1}: X_{i,j}^{P1} = x_{i,j} + r \times \left(Iguana_j - I \times x_{i,j} \right), \quad for\ i = 1,2,\ldots,\left\lfloor \frac{N}{2} \right\rfloor, \text{and}\ \ j = 1,2,\ldots,v. \tag{1}$$

Where $P1$ represents the exploration stage, $Iguana$ denotes the position information of the iguana, r is a constant between [0,1], I is a random integer from the set $\{1,2\}$, N is the population size, and v is the dimensionality of decision variables for individuals. Assuming the landing position of the iguana is stochastic, after landing, the other half of the coatis start moving towards the prey for trapping and hunting. This process can be defined by the mathematical formula[20]:

$$Iguana^G: Iguana_j^G = lb_j + r \times (ub_j - lb_j), j = 1,2,....,v. \tag{2}$$

$$X_i^{P1}: X_{i,j}^{P1} = \begin{cases} x_{i,j} + r \times (Iguana_j^G - I \times x_{i,j}), F_{Iguana^G} < F_i, \\ x_{i,j} + r \times (x_{i,j} - Iguana_j^G), F_{Iguana^G} \geq F_i. \end{cases}$$

$$for \quad i = \left\lfloor \frac{N}{2} \right\rfloor + 1, \left\lfloor \frac{N}{2} \right\rfloor + 2,...,N \quad and \quad j = 1,2,...,v. \tag{3}$$

Among them, $Iguana^G$ represents the location where the iguana lands. ub and lb respectively denote the upper and lower bounds of the decision space.

2) Development phase

When encountering predators, coatis swiftly depart from their current location, a process that can be described as[20]:

$$lb_j^{\text{local}} = \frac{lb_j}{t}, ub_j^{\text{local}} = \frac{ub_j}{t}, \text{where } t = 1,2,\ldots,T. \tag{4}$$

$$X_i^{P2}: x_{i,j}^{P2} = x_{i,j} + (1-2r) \cdot \left(lb_j^{\text{local}} + r \cdot \left(ub_j^{\text{local}} - lb_j^{\text{local}} \right) \right), \quad i = 1,2,\ldots,N, j = 1,2,\ldots,v. \tag{5}$$

Where $P2$ represents the developmental stage, t denotes the current iteration number, and T indicates the maximum iteration count. After the coati positions are updated in each stage, it is necessary to conduct a verification to ensure the effectiveness of the update. The verification process is as follows [20]:

$$X_i = \begin{cases} X_i^{P1,P2}, & F_i^{P1,P2} < Fbest_i \\ X_i, & else \end{cases}, \tag{6}$$

Where $F_i^{P1,P2}$ represents the target value, and $Fbest_i$ denotes the current optimal value. For further details on COA, please refer to [20].

B. Multi-objective Coati optimization algorithm (MOCOA)

1) Elite Coati retention Strategy

Unlike single-objective optimization, the optimal solution of multi-objective optimization algorithms is no longer a single solution but a set of Pareto optimal solutions, which means that at any given time, the iguana is not just a single one but a group, as shown in Fig. 1. Introducing the fast nondominated sorting and crowding distance calculation methods from NSGA-II [19], a novel elite coati retention strategy is proposed. Firstly, all individuals in the population undergo a two-stage update, with the updated coatis referred to as offspring individuals and the coatis before the update as parent individuals. All parent and offspring individuals undergo fast nondominated sorting together to determine the nondominated front $Front_l, l = 1,2,...,K$ for each rank. Secondly, the crowding distance of each coati is calculated. Finally, a new population is selected according to the following selection process:

$$P^{G+1} = \begin{cases} P_0 \cup \sum_{i=1}^{K} Front_1, & if \mid P_0 + Front_1 \mid \leq N \\ P_0 \cup Front_1^{|N-P_0|}, & else \end{cases} \tag{7}$$

Where P^{G+1} represents the newly selected population, P_0 denotes the set of selected solutions. When the number of individuals in the set is less than N, the $Front_1$ individuals in the forefront are added to P_0 sequentially. When $\mid P_0 + Front_1 \mid > N$, individuals with the maximum numerical value are selected one by one according to the crowding distance until the number of individuals in P_0 reaches N.

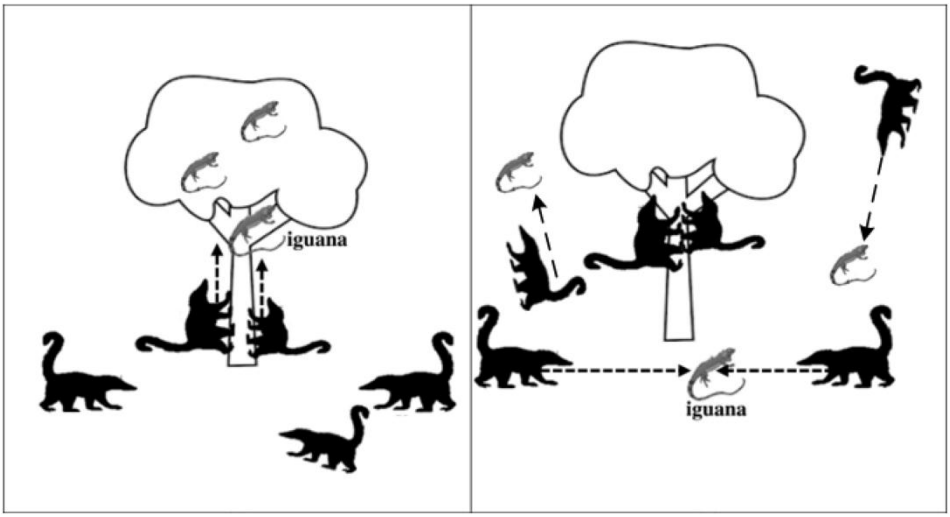

Fig. 1 MOCOA first phase model diagram

2) MOCOA Optimization Procedure

In MOCOA, partial update rules of COA are no longer applicable. In Eq.1, $Iguana$ no longer represents the best individual in each generation, but rather a random individual from the set $Front$. Additionally, MOCOA employing the elite coati preservation strategy no longer requires Eq. 6 to verify the effectiveness of the updates. The specific optimization steps proposed in MOCOA are as follows:

Step1: Initialize the population and set hyperparameters, including population size, bounds for individuals, maximum number of iterations, etc.

Step2: Randomly initialize all coati individuals' information and calculate their corresponding fitness values.

Step3: Perform fast non-dominated sorting on the population and calculate individual crowding distances.

Step4: Randomly select individuals from set $Front_1$ as $Iguana_j$, and complete coati individual updates based on Eq.1 to Eq. 5.

Step5: Utilize the elite coati preservation strategy, as described by Eq. 7, to select the new population P^{G+1}.

Step6: Check if the algorithm satisfies the termination conditions. If not, proceed to step 4 until the termination conditions are met, and output the Pareto front (PF).

These steps constitute the main workflow of the MOCOA algorithm, covering key operations such as initialization, individual updates, multi-objective sorting, crowding distance calculation, and termination conditions. This makes MOCOA an effective multi-objective optimization algorithm for addressing multi-objective optimization problems.

C. Multi-objective Ensemble Forecasting Model Based on Laguerre Polynomials

1) Laguerre Polynomials

Laguerre polynomials are a set of orthogonal polynomials defined on the non-negative real numbers, introduced by Edmond Laguerre, and are related to the gamma distribution density function. The Laguerre polynomials are defined as follows:

$$Q_n(x) = e^x \frac{d^n}{dx^n}\left(x^n e^{-x}\right) x \in [0, +\infty) \tag{8}$$

The mathematical expression for the orthogonality of Laguerre polynomials with respect to the weight function is as follows:

$$\int_0^\infty e^{-x} Q_n(x) Q_m(x) dx = \begin{cases} 0, m \neq n \\ (n!)^2, m = n \end{cases} \tag{9}$$

The recursive relationship expression for Laguerre polynomials is as follows:

$$\begin{cases} Q_0(x) = 1 \\ Q_1(x) = 1 - x \\ Q_{n+1}(x) = (1 + 2n - x)Q(x) - n^2 P_{n-1}(x)(n = 1, 2, \ldots) \end{cases} \tag{10}$$

Based on the theory of Laguerre polynomials, Laguerre orthogonal basis functions are constructed. Utilizing the non-linear fitting characteristics of Laguerre orthogonal basis functions, the prediction curve is fitted. The neural network structure diagram with Laguerre orthogonal basis functions as activation functions is illustrated in Fig. 2.

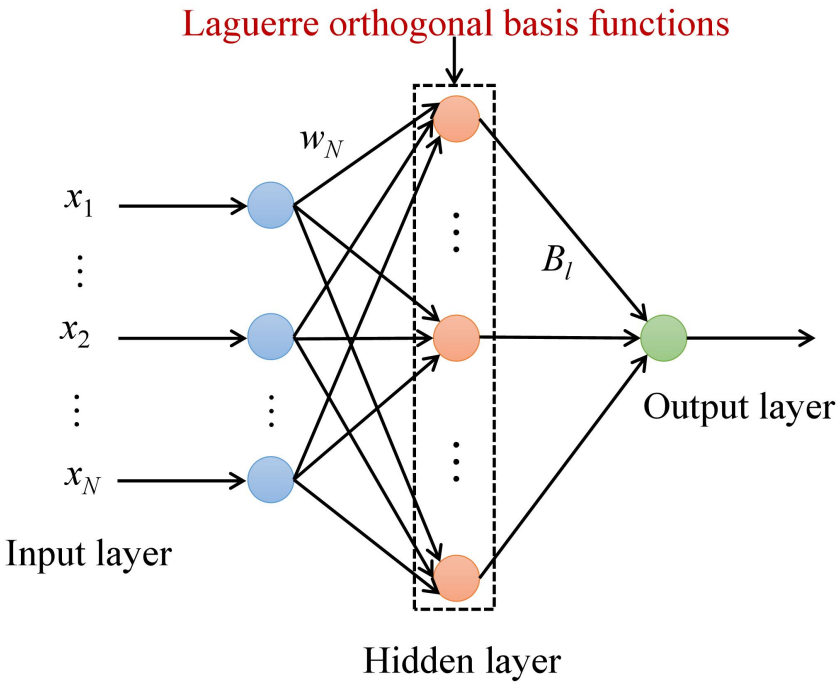

Fig. 2 Laguerre neural network structure diagram

In Fig. 2, x_N represents the input values of the training sample data, while w_N and B_l represent the weights between the input layer and the hidden layer, and between the hidden layer and the output layer, respectively.

2) MOCOA-LNN prediction model

To simultaneously improve the accuracy and stability of the LNN model, the MOCOA multi-objective optimization algorithm is introduced to optimize the weights of the LNN within the prediction model, aiming to achieve more accurate and stable prediction results.

The optimization objectives using MOCOA for the accuracy and stability of LNN weights and thresholds are as follows [16]:

$$F_1 = MSE = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 \quad (11)$$

$$F_2 = std(|\hat{y}_i - y_i|), t=1,2,...,N \quad (12)$$

Wherein, $\hat{y}_h$, y_h , and N represent the predicted value, actual value, and sample size of the distribution network project cost, respectively.

3) Ensemble learning predictive correction model

During the prediction phase, MOCOA assigns optimal weights to the Laguerre feedforward neural network, ensuring both high accuracy and strong stability of the prediction results. However, constrained by the insufficient generalization ability of the Laguerre feedforward neural network itself, the prediction results of the model have certain uncertainties. To reduce the inherent uncertainty of the MOCOA-LNN prediction model and enhance its predictive capability, this study introduces ensemble learning Adaboost. Based on the MOCOA-LNN prediction model, several weak predictors are constructed, and the final MOCOA-LNN strong predictor is built through adaptive allocation and recombination of the weights of weak predictor errors, forming the MOCOA-LNN-Adaboost prediction model to further improve the predictive performance of the model. The process of Adaboost prediction error correction is as follows:

Step1: Construct weak predictors for MOCOA-LNN and calculate errors.

$$e_k = \sum_i D_k(i), \quad i = 1,2,...,n \quad (13)$$

Step2: Calculate the error sequence weights based on the prediction errors of the MOCOA-LNN weak predictors.

$$a_k = \frac{1}{2}\ln(\frac{1-e_k}{e_k}) \quad (14)$$

Step3: Update the weights of weak predictor errors.

$$\beta(i) = g_k(x_i) - y_i \quad (15)$$

$$D_{k+1}(i) = \frac{D_k(i)}{Z_k} \times \begin{cases} a_k & \beta(i) \le \varphi \\ 1 & \beta(i) > \varphi \end{cases} \quad (16)$$

Step4: Construct the final MOCOA-LNN strong predictor.

$$a_k = \frac{a_k}{\sum_{k=1}^{K} a_k} \quad (17)$$

$$\hat{y}_i = \sum_{k=1}^{K} a_k g_k(x_i) \quad (18)$$

In the above formulas, $D_k(i)$ represents the sample weight, $g_k(x_i)$ and y_i respectively denote the predicted value and the actual value of the k th MOCOA-LNN weak predictor, $\beta(i)$ stands for the prediction error of the weak predictor, and $\hat{y}_i$ represents the predicted value of the final MOCOA-LNN strong predictor.

4) The MOCOA-LNN-Adaboost Multi-Objective Ensemble Prediction Model

Firstly, the LNN is constructed through Laguerre polynomials, enabling precise prediction of the cost of distribution network engineering by leveraging the powerful nonlinear fitting characteristics of orthogonal polynomials. Secondly, the MOCOA optimizes the LNN through multi-objective optimization, ensuring that the MOCOA-LNN model achieves both high prediction accuracy and strong predictive stability. Finally, Adaboost is utilized to adaptively recombine the prediction errors of MOCOA-LNN, resulting in the MOCOA-LNN-Adaboost distribution network engineering cost prediction model, further enhancing prediction accuracy. The flowchart of the MOCOA-LNN-Adaboost multi-objective ensemble prediction model is illustrated in Fig. 3.

Wherever Times is specified, Times Roman or Times New Roman may be used. If neither is available on your word processor, please use the font closest in appearance to Times. Avoid using bit-mapped fonts if possible. True-Type 1 or Open Type fonts are preferred. Please embed symbol fonts, as well, for math, etc.

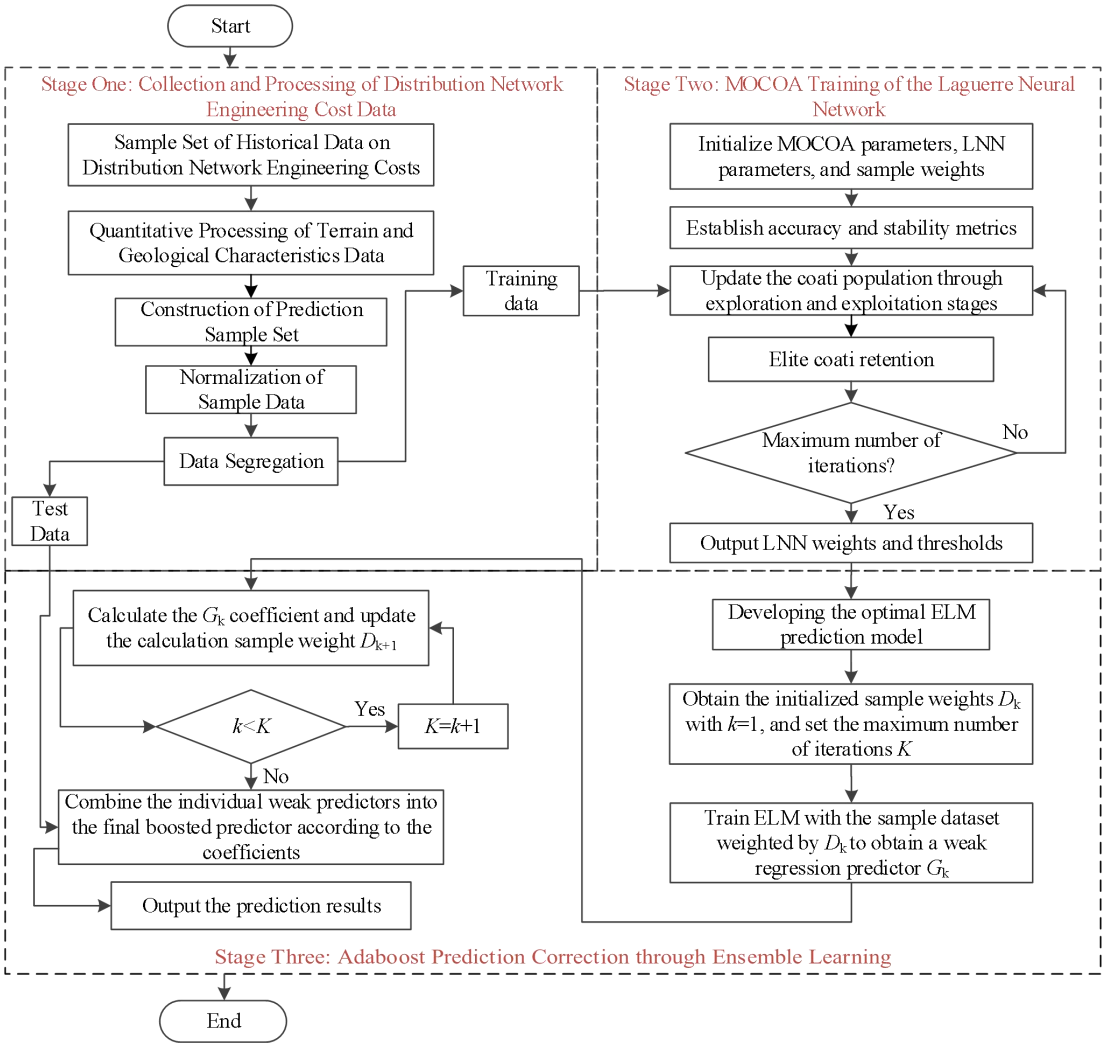

Fig. 3 MOCOA-LNN-Adaboost multi-objective integrated prediction model flow chart

III. RESULTS AND ANALYSIS

This section first evaluates the proposed MOCOA algorithm through extensive simulation experiments. Subsequently, using historical cost data from a distribution network project in Ningxia as the research sample, the performance of the proposed MOCOA-LNN-Adaboost distribution network project cost multi-objective integrated prediction model is validated through a comprehensive set of experiments.

A. MOCOA performance evaluation

To evaluate the superiority of the proposed MOCOA algorithm compared to other mainstream algorithms, various algorithms were compared, including NSGA-III [21], MOSSA [22], and MOMVO [23], along with the proposed MOCOA algorithm. Two different subsets of test problems were utilized: MOP1-MOP4 [24] and ZDT1-ZDT4 [25]. These problems are all bi-objective and represent diverse challenges in multi-objective optimization. The parameters for all four algorithms were uniformly set, with a population size of 100 and a maximum number of iterations of 100. Other parameters were set according to the respective original literature. To ensure the fairness of the experimental results, 30 independent runs were conducted, and the Inverse Generation Distance (IGD) indicator was used to evaluate the performance of the algorithms. The IGD indicator can effectively measure the performance of algorithms in finding approximate solutions to the Pareto front.

As shown in Table 1 and Table 2, the MOCOA algorithm demonstrates excellent performance across eight different test problems, exhibiting superior convergence and diversity. Particularly noteworthy is the significant performance improvement of the MOCOA algorithm compared to other algorithms, notably evident in the MOP4 problem. This further underscores the effectiveness of the algorithm in the field of multi-objective optimization. The corresponding test results depicted in Fig. 4 clearly illustrate that the approximate frontier generated by the MOCOA algorithm is closer to the true frontier and is more uniformly distributed, indicating better performance.

Table 3 presents the p-values of the proposed algorithm relative to the alternative hypotheses of the comparison algorithms. The p-value is less than 0.05 for all problems, indicating a significant difference between the distributions of the proposed algorithm and the selected algorithms. The computational complexity of both the proposed algorithm and the comparison algorithms is , where is the population size, and is the number of objective functions. Table 4 provides the CPU computation time for each algorithm, with time measured in seconds. The results indicate that the proposed algorithm outperforms the comparison algorithms in terms of computation time for each test problem.These experimental findings reinforce the standing of the MOCOA algorithm in multi-objective optimization problems and provide a robust tool for addressing complex multi-objective problems in the real world..

Table 1. IGD values of the proposed algorithm and the comparison algorithm on MOP1-MOP4 problem

OA	MOP1					OA	MOP2				
	Best	Mean	Middle	Variance	Worst		Best	Mean	Middle	Variance	Worst
NSGAIII	0.3867	0.6616	0.8313	0.2442	0.8326	NSGAIII	0.0122	0.0210	0.0218	0.0085	0.0342
MOSSA	0.7652	0.7887	0.7851	0.0132	0.7976	MOSSA	0.0165	0.0224	0.0240	0.0060	0.0303
MOMVO	0.7809	0.7960	0.7870	0.0138	0.8121	MOMVO	0.0203	0.0227	0.0239	0.0014	0.0238
MOCOA	0.3569	0.3601	0.3607	0.0042	0.3630	MOCOA	0.0049	0.0055	0.0055	0.0005	0.0057
OA	MOP3					OA	MOP4				
	Best	Mean	Middle	Variance	Worst		Best	Mean	Middle	Variance	Worst
NSGAIII	0.0126	0.0236	0.0260	0.0102	0.0370	NSGAIII	0.0232	0.0489	0.0579	0.0227	0.0773
MOSSA	0.0183	0.0235	0.0219	0.0053	0.0317	MOSSA	0.0124	0.0177	0.0189	0.0047	0.0241
MOMVO	0.0178	0.0246	0.0243	0.0046	0.0297	MOMVO	0.0097	0.0121	0.0124	0.0014	0.0135
MOCOA	0.0051	0.0056	0.0057	0.0003	0.0060	MOCOA	0.0034	0.0035	0.0035	0.0001	0.0037

Table 2. The IGD values of the MOCOA and the comparison algorithm for ZDT1-ZDT4 problem

OA	ZDT1					OA	ZDT2				
	Best	Mean	Middle	Variance	Worst		Best	Mean	Middle	Variance	Worst
NSGAIII	0.0173	0.0334	0.0316	0.0149	0.0585	NSGAIII	0.0149	0.0451	0.0336	0.0362	0.1080
MOSSA	0.0220	0.0387	0.0399	0.0110	0.0515	MOSSA	0.0328	0.0603	0.0552	0.0210	0.0870
MOMVO	0.0177	0.0229	0.0231	0.0036	0.0271	MOMVO	0.0252	0.0526	0.0348	0.0404	0.1231
MOCOA	0.0038	0.0039	0.0041	0.0001	0.0045	MOCOA	0.0041	0.0041	0.0039	0.0001	0.0048
OA	ZDT3					OA	ZDT4				
	Best	Mean	Middle	Variance	Worst		Best	Mean	Middle	Variance	Worst
NSGAIII	0.0833	0.1658	0.1691	0.0523	0.2265	NSGAIII	0.0835	0.0685	0.1693	0.0331	0.2271
MOSSA	0.0463	0.0742	0.0732	0.0193	0.0996	MOSSA	0.1432	0.3956	0.4824	0.2017	0.5782
MOMVO	0.0210	0.0263	0.0243	0.0055	0.0362	MOMVO	2.0163	3.3843	3.6277	0.8196	4.0798
MOCOA	0.0047	0.0051	0.0044	0.0002	0.0053	MOCOA	0.0042	0.0037	0.0043	0.0001	0.0046

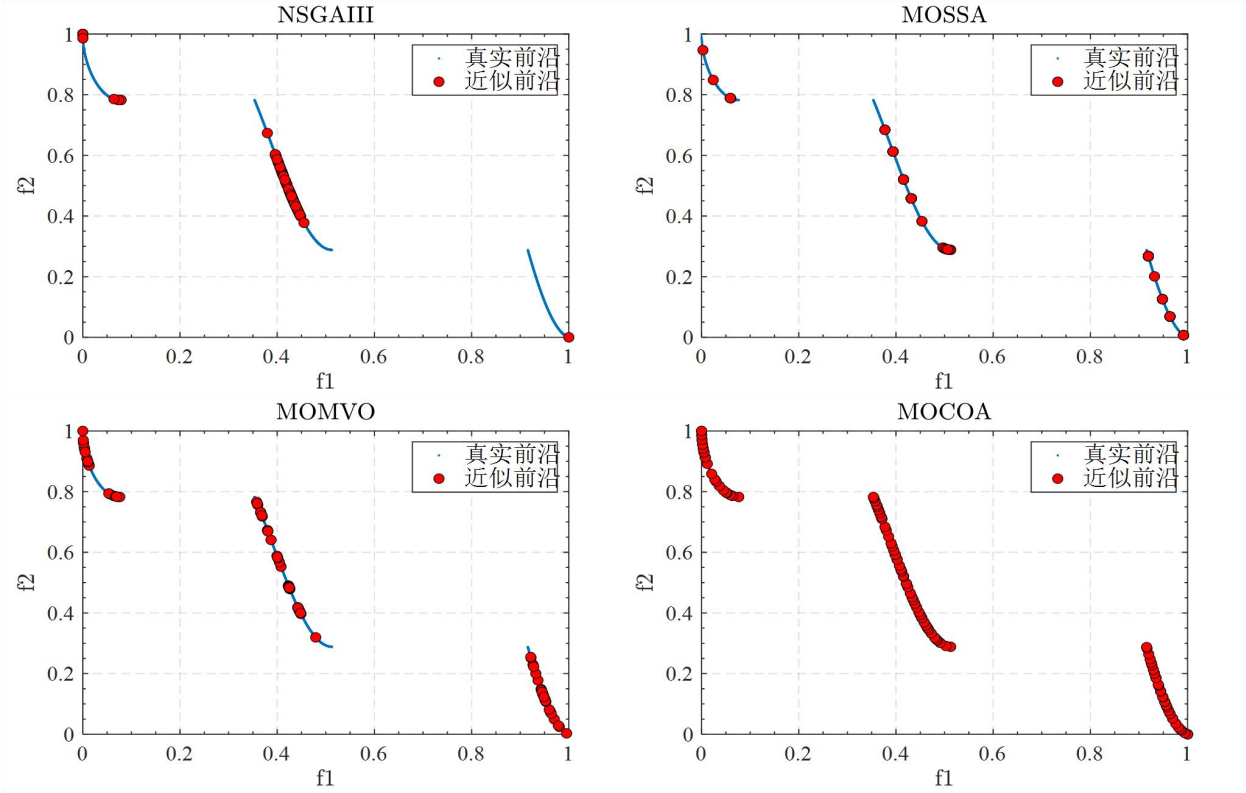

Fig. 4 Test results of four algorithms on MOP4 problem

Table 3 Significance Test of IGD Values between the Proposed Algorithm and Comparison Algorithms on Test Problems

Problem	MOCOA		
	NSGAIII	MOSSA	MOMVO
MOP1	1.21E-12	1.21E-12	1.70E-08
MOP2	4.98E-11	3.02E-11	3.02E-11
MOP3	1.21E-12	1.93E-10	1.21E-12
MOP4	6.07E-11	3.69E-11	3.02E-11
ZDT1	4.18E-02	1.21E-12	1.21E-12
ZDT2	3.34E-01	1.21E-12	1.21E-12
ZDT3	1.21E-12	1.21E-12	1.21E-12
ZDT4	2.63E-11	2.63E-11	2.63E-11

Table 4 The CPU computation time of the proposed algorithm and comparison algorithms on test problems

Problem	NSGAIII	MOSSA	MOMVO	MOCOA
MOP1	10.2462	1.9200	0.7328	0.6456
MOP2	9.9521	6.5610	1.9864	1.7973
MOP3	9.2195	6.2225	1.9712	1.8864
MOP4	9.1906	4.9879	1.7240	1.5362
ZDT1	8.6891	1.5981	0.7109	0.6782
ZDT2	9.4960	1.3283	0.6217	0.6049
ZDT3	8.9806	1.7566	0.7625	0.7457
ZDT4	9.5320	1.2689	0.5625	0.5381

B. Distribution network project cost forecast example verification

1)Collection and Processing of Distribution Network Engineering Cost Data

This study employs the actual cost data of 150 completed distribution network projects from a company in Ningxia as the training set for model training and prediction. The features of the sample data mainly include unit cost of engineering, terrain, geology, conductor cross-section, conductor length, wind speed, ice coating, and unit price of conductor, among others. Quantitative processing is conducted on the feature data of terrain and geology indicators [26]. The sample data for distribution network line engineering costs are presented in Table 5.

Table 5 Data of distribution network project cost

Line number	Unit cost (yuan/ km)	Unit cost of conductor (yuan)	Unit cost of tower material (yuan)	…	landform	geology	Wind speed (m/s)
1	2921157	31306	8509	…	4.96	2.78	26
2	2799351	17850	6035	…	4.72	2.69	24
3	2165983	15725	6035	…	3.45	3.82	26
4	3598887	17850	6035	…	6.32	2.71	26
5	3036849	15725	8509	…	5.19	4.49	26
6	3107696	46257	8509	…	5.34	3.44	26
7	3058003	17850	6035	…	5.24	3.44	26
8	2594365	17850	12515	…	4.31	2.71	24
9	3312861	15725	6035	…	5.75	2.71	24
10	1981443	15130	6035	…	3.09	3.44	23
…	…	…	…	…	…	…	…
149	3777793	52443	8579	…	8.45	6.91	36
150	3996935	55442	9108	…	5.06	10.85	43

2)Prediction and evaluation index

This paper selects several comprehensive evaluation metrics to assess the predictive performance of the model, including the coefficient of determination (R-square, R2), median absolute percentage error (MdAPE), index of agreement (IA), mean absolute error (MAE), root mean square error (RMSE), and stability index (SDEX). Among these six evaluation metrics, larger values of R2 and IA indicate stronger predictive performance of the model, while smaller values of MdAPE, MAE, RMSE, and SDEX suggest stronger predictive performance as well. The mathematical formulas for these evaluation metrics are provided below:

$$\mathrm{R}^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \overline{y}_i)^2} \tag{19}$$

$$\mathrm{MdAPE} = median\left(\left|\frac{\hat{y}_i - y_i}{y_i}\right| \times 100\%\right) \tag{20}$$

$$\mathrm{IA} = 1 - \frac{\sum_{it=1}^{N}\left(\hat{y}_i - y_i\right)^2}{\sum_{it=1}^{N}\left(\left|\hat{y}_i - \overline{y}_i\right| + \left|y_i - \overline{y}_i\right|\right)^2} \tag{21}$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{y}_i - y_i\right| \tag{22}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{y}_i - y_i\right)^2} \tag{23}$$

$$\mathrm{SDEX} = std\left(\left|\hat{y}_i - y_i\right|\right) \quad (24)$$

Where $\overline{y}_i$ represents the average value of the actual distribution network engineering costs.

3) Results and discussion

(a) Compromise choice

Balancing the accuracy and stability of distribution network engineering cost prediction is a multi-objective optimization problem, where accuracy and stability metrics conflict with each other. The obtained solutions form a Pareto front, where all solutions are mutually non-dominated, and each solution corresponds to a position in the population that can be used as the weight of the LNN neural network. In this study, fuzzy decision theory is employed to select the point closest to the origin as the compromise solution for the LNN, achieving a balance between prediction accuracy and stability. The selection of Pareto solutions and compromise solutions in the MOCOA-LNN-Adaboost model is illustrated in Fig. 5.

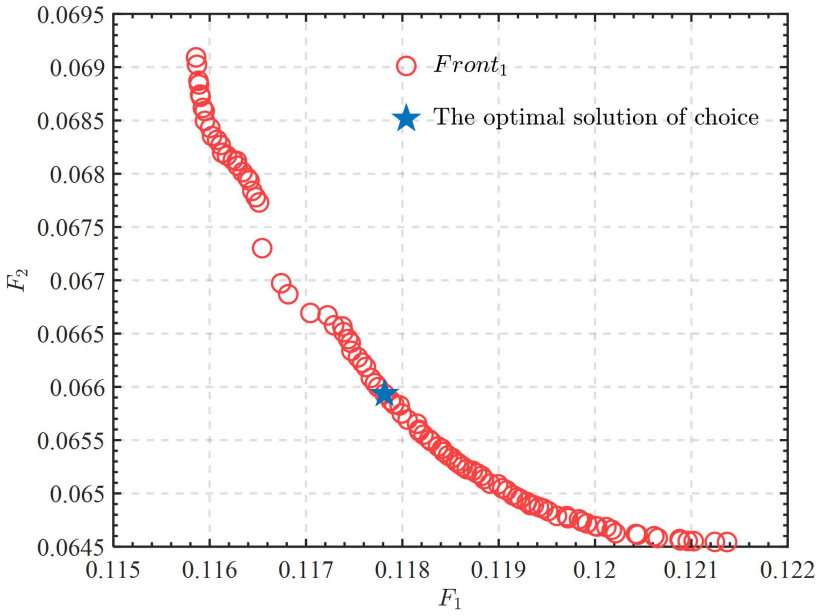

Fig. 5 MOCOA optimization results and compromise solution selection

(b) Comparison of Predictive Results of the MOCOA-LNN-Adaboost Model

To verify the effectiveness of the proposed Adaboost ensemble learning in improving model prediction performance, the predictive results of the MOCOA-LNN-Adaboost model are compared with those of the MOCOA-LNN model. Both models select the same compromise solution as the LNN weights. The evaluation metrics of predictive results are summarized in Table 4, while the predictive result curves are illustrated in Fig. 6, and the comparison of model prediction absolute errors is depicted in Fig. 7.

As shown in Table 6, compared to the MOCOA-LNN prediction model, the proposed prediction model exhibits 11.81% increase in R2, 58.21% decrease in MdAPE, 2.98% increase in IA, 56.46% decrease in MAE, 58.56% decrease in RMSE, and 64.18% decrease in SDEX. Fig. 6 and Fig. 7 indicate that the MOCOA-LNN-Adaboost model outperforms in prediction accuracy, thereby demonstrating higher predictive precision. Consequently, the ensemble learning method Adaboost significantly enhances predictive performance.

Table 6 Self-comparison of prediction evaluation indicators

Model	R^2	MdAPE	IA	MAE	RMSE	SDEX
MOCOA-LNN	0.8511	5.8876	0.9584	1.62E+05	1.92E+05	1.06E+05
MOCOA-LNN-Adaboost	**0.9516**	**2.4607**	**0.9870**	**7.03E+04**	**7.95E+04**	**3.79E+04**

Note: Bold font indicates the optimal value of the evaluation metric.

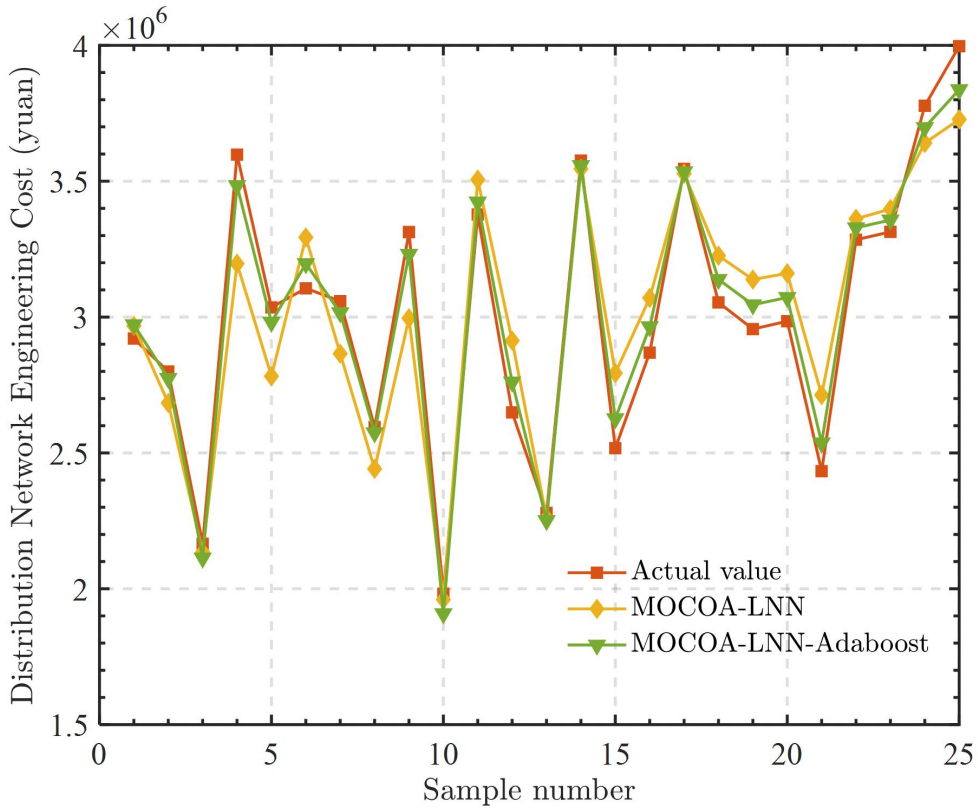

Fig. 6 Self-comparison of model prediction results

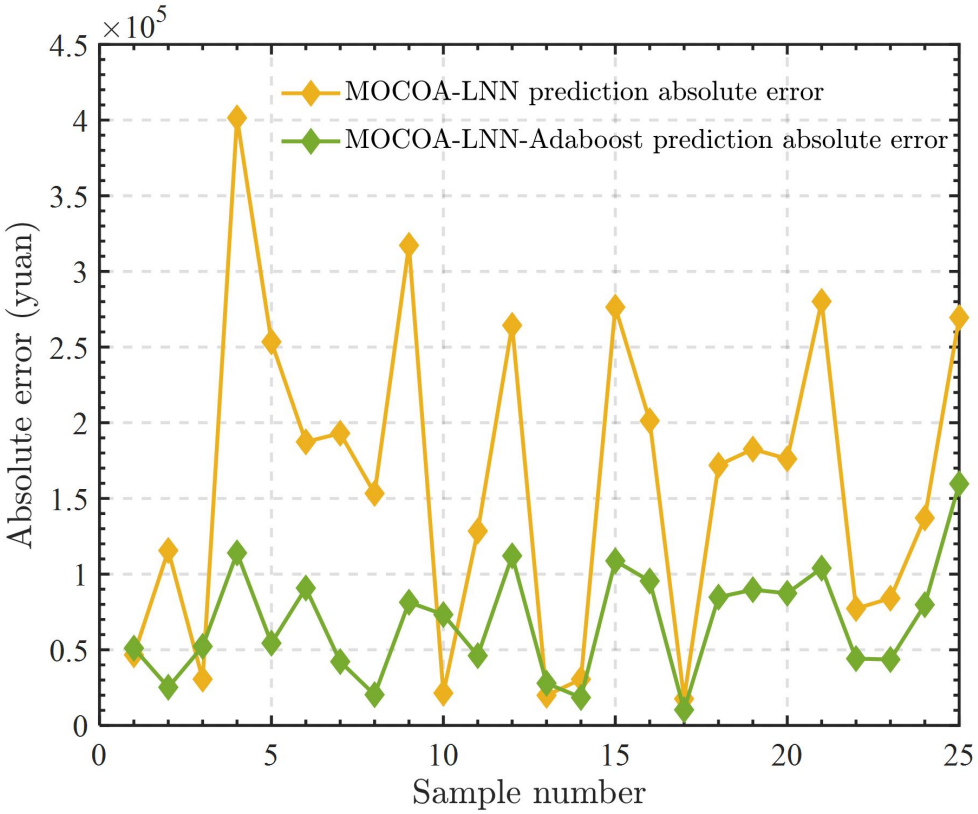

Fig. 7 Comparison of absolute error of model prediction

(c) Comparison of Multi-Model Prediction Results

To verify whether the proposed prediction model is superior to current mainstream prediction models, comparative experiments were conducted between the MOCOA-LNN-Adaboost prediction model and the MOCOA-ELM-Adaboost prediction model, MOCOA-LSTM-Adaboost prediction model, MOCOA-RBF-

Adaboost prediction model, and MOCOA-BPNN-Adaboost prediction model. The comparison of model evaluation indicators is shown in Table 7. The comparison of prediction results is illustrated in Fig. 8.

Table7 Comparison of evaluation indicators of forecasting results of different forecasting models

Model	R^2	MdAPE	IA	MAE	RMSE	SDEX
NSGAIII-ELM-Adaboost	0.8679	4.1204	0.9615	3.51E+04	1.71E+05	9.94E+04
NSGAIII-LSTM-Adaboost	0.9252	3.0047	0.9811	**2.77E+04**	1.36E+05	8.07E+04
NSGAIII-RBF-Adaboost	0.8815	4.5552	0.9668	3.52E+04	1.67E+05	9.12E+04
NSGAIII-BPNN-Adaboost	0.9173	4.3985	0.9781	1.28E+05	1.44E+05	6.65E+04
NSGAIII-LNN-Adaboost	**0.9516**	**2.4607**	**0.9870**	7.03E+04	**7.95E+04**	**3.79E+04**

Note: Bold font indicates the optimal value of the evaluation metric.

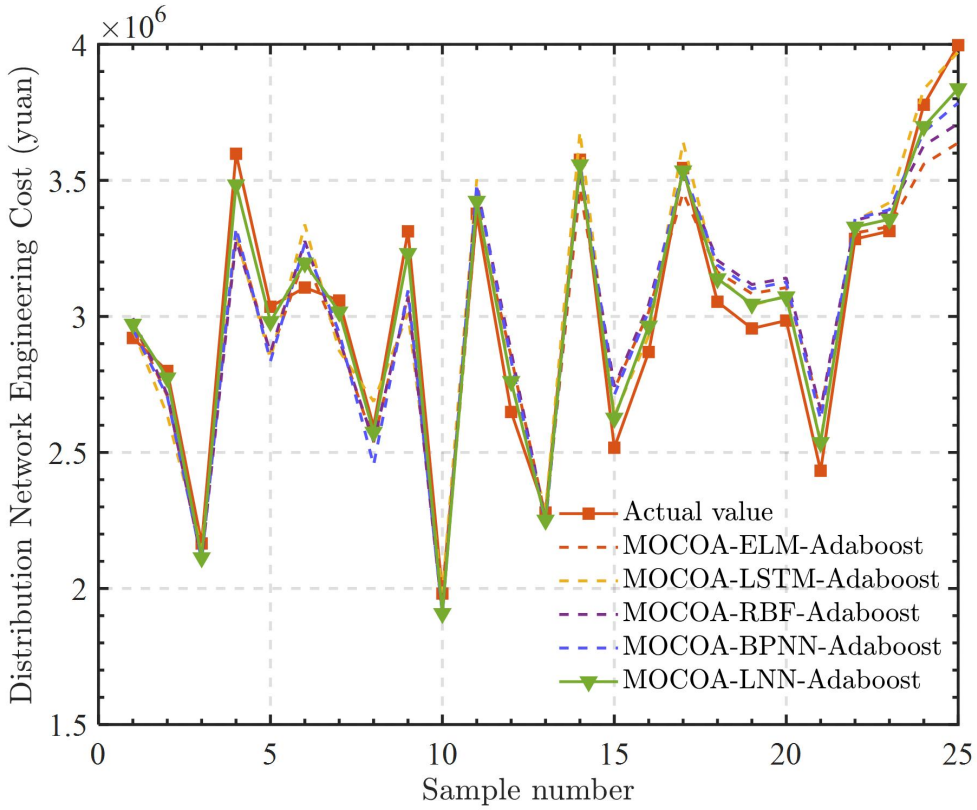

Fig. 8 Self-comparison of prediction results of multiple models

From Fig. 8, it can be observed that compared to the other four prediction models, the proposed MOCOA-LNN-Adaboost model exhibits a higher degree of match with the actual distribution network engineering cost sample curve. It performs better in predicting the "peaks" of the samples, indicating stronger reliability in the model's prediction results.

As shown in Table 7, in the comparative experiments of the six evaluation indicators, except for MAE, the proposed MOCOA-LNN-Adaboost distribution network engineering cost prediction model achieves optimal results compared to the other four mainstream prediction models. Compared to MOCOA-ELM-Adaboost, the proposed model demonstrates 9.64% increase in R2, 2.65% increase in IA, 53.62% decrease in RMSE, and 61.87% decrease in RMSE. Compared to MOCOA-RBF-Adaboost, the proposed model achieves 45.98% reduction in MdAPE. The experimental results indicate that compared to current mainstream prediction models, the proposed prediction model exhibits higher prediction accuracy and better prediction stability.

IV. CONCLUSION

This paper proposes a MOCOA-LNN-Adaboost prediction model that balances high prediction accuracy with strong prediction stability and validates the model using historical data of distribution network engineering costs. Through extensive comparative experiments, the following conclusions are drawn:

(a) The proposed multi-objective optimization algorithm MOCOA exhibits optimal evaluation indicators in both test and real-world problems, demonstrating better diversity and convergence, thus proving the effectiveness and adaptability of the proposed algorithm.

(b) In the selection of Pareto front compromise solutions for multi-objective optimization prediction of distribution network engineering costs, the model's prediction accuracy and stability can be balanced, thereby ensuring both high prediction accuracy and strong prediction stability, thus enhancing the credibility of the model's prediction results.

(c) Ensemble learning with Adaboost effectively improves the prediction performance of multi-objective prediction models, ensuring high prediction stability while further enhancing prediction accuracy, thereby ensuring high precision in prediction results.

Compared to current mainstream prediction models, the proposed MOCOA-LNN-Adaboost multi-objective integrated prediction model demonstrates higher prediction accuracy and better prediction stability, providing new research directions for distribution network engineering cost prediction. However, the effectiveness of the MOCOA-LNN-Adaboost model is more pronounced with large sample data, and for prediction problems with smaller samples, the introduction of data augmentation methods is necessary to enhance the model

ACKNOWLEDGMENT

The authors appreciate the financial supports provided by Natural Science Foundation project of Ningxia Autonomous Region (No. 22022AAC02078).

REFERENCES

[1] Montoya O D, Gil-González W, Grisales-Noreña L F. An exact MINLP model for optimal location and sizing of DGs in distribution networks: A general algebraic modeling system approach[J]. Ain Shams Engineering Journal, 2020, 11(2): 409-418.

[2] Anaya K L, Pollitt M G. The role of distribution network operators in promoting cost-effective distributed generation: Lessons from the United States of America for Europe[J]. Renewable and Sustainable Energy Reviews, 2015, 51: 484-496.

[3] Kai Y, Bo Y, Yanli X, et al. Distribution network project cost prediction based on GA-BP neural network [J]. Automation Instrumentation, 2019, 40(07): 91-93.

[4] Bouhouras A S, Labridis D P, Bakirtzis A G. Cost/worth assessment of reliability improvement in distribution networks by means of artificial intelligence[J]. International Journal of Electrical Power & Energy Systems, 2010, 32(5): 530-538.

[5] Stathopoulos C, Kaperoni A, Galanis G, et al. Wind power prediction based on numerical and statistical models[J]. Journal of Wind Engineering and Industrial Aerodynamics, 2013, 112: 25-38.

[6] Sghir N, Adadi A, Lahmer M. Recent advances in Predictive Learning Analytics: A decade systematic review (2012–2022)[J]. Education and information technologies, 2023, 28(7): 8299-8333.

[7] Wang H, Lei Z, Zhang X, et al. A review of deep learning for renewable energy forecasting[J]. Energy Conversion and Management, 2019, 198: 111799.

[8] Ye D. An algorithm for construction project cost forecast based on particle swarm optimization-guided BP neural network[J]. Scientific Programming, 2021, 2021: 1-8.

[9] Skitmore R M, Ng S T. Forecast models for actual construction time and cost[J]. Building and environment, 2003, 38(8): 1075-1083.

[10] Batselier J, Vanhoucke M. Improving project forecast accuracy by integrating earned value management with exponential smoothing and reference class forecasting[J]. International journal of project management, 2017, 35(1): 28-43.

[11] Zhou J, Qiu Y, Zhu S, et al. Optimization of support vector machine through the use of metaheuristic algorithms in forecasting TBM advance rate[J]. Engineering Applications of Artificial Intelligence, 2021, 97: 104015.

[12] Xu J, Moon S. Stochastic forecast of construction cost index using a cointegrated vector autoregression model[J]. Journal of management in engineering, 2013, 29(1): 10-18.

[13] Fan M, Sharma A. Design and implementation of construction cost prediction model based on svm and lssvm in industries 4.0[J]. International journal of intelligent computing and cybernetics, 2021, 14(2): 145-157.

[14] Koo C W, Hong T H, Hyun C T. The development of a construction cost prediction model with improved prediction capacity using the advanced CBR approach[J]. Expert Systems with Applications, 2011, 38(7): 8597-8606.

[15] Tatiya A, Zhao D, Syal M, et al. Cost prediction model for building deconstruction in urban areas[J]. Journal of Cleaner Production, 2018, 195: 1572-1580.

[16] Wang Jianzhou, Du Pei, Niu Tong, et al. A novel hybrid system based on a new proposed algorithm—Multi-Objective Whale Optimization Algorithm for wind speed forecasting[J]. Applied Energy,2017,208: 344-360.

[17] Hashim F A, Hussien A G. Snake Optimizer: A novel meta-heuristic optimization algorithm[J]. Knowledge-Based Systems, 2022, 242: 108320.

[18] Agushaka J O, Ezugwu A E, Abualigah L. Gazelle optimization algorithm: a novel nature-inspired metaheuristic optimizer[J]. Neural Computing and Applications, 2023, 35(5): 4099-4131.
[19] Ahmadianfar I, Heidari A A, Noshadian S, et al. INFO: An efficient optimization algorithm based on weighted mean of vectors[J]. Expert Systems with Applications, 2022, 195: 116516.
[20] Dehghani M, Montazeri Z, Trojovská E, et al. Coati Optimization Algorithm: A new bio-inspired metaheuristic algorithm for solving optimization problems[J]. Knowledge-Based Systems, 2023, 259: 110011.
[21] Deb K, Jain H. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: solving problems with box constraints[J]. IEEE transactions on evolutionary computation, 2013, 18(4): 577-601.
[22] Mirjalili S, Gandomi A H, Mirjalili S Z, et al. Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems[J]. Advances in engineering software, 2017, 114: 163-191.
[23] Mirjalili S, Jangir P, Mirjalili S Z, et al. Optimization of problems with multiple objectives using the multi-verse optimization algorithm[J]. Knowledge-Based Systems, 2017, 134: 50-71.
[24] Zitzler E, Deb K, Thiele L. Comparison of multiobjective evolutionary algorithms: Empirical results[J]. Evolutionary computation, 2000, 8(2): 173-195.
[25] Liu H L, Gu F, Zhang Q. Decomposition of a multiobjective optimization problem into a number of simple multiobjective subproblems[J]. IEEE transactions on evolutionary computation, 2013, 18(3): 450-455.
[26] YANG Ying, Liu Xiaomin, Youfei et al. Construction and error analysis of distribution network engineering cost prediction Model based on PSO-LSSVM [J]. Automation Technology and Application, 2019,39(02):98-102: