

Model Uncertainty in Evolutionary Optimization and Bayesian Optimization: A Comparative Analysis

1st Hao Hao

*Institute of Natural Sciences
Shanghai Jiao Tong University
Shanghai, China
haohao@sjtu.edu.cn*

2nd Xiaoqun Zhang

*Institute of Natural Sciences
Shanghai Jiao Tong University
Shanghai, China
xqzhang@sjtu.edu.cn*

3rd Aimin Zhou

*School of Computer Science and Technology
East China Normal University
Shanghai, China
amzhou@cs.ecnu.edu.cn*

Abstract—Black-box optimization problems, which are common in many real-world applications, require optimization through input-output interactions without access to internal workings. This often leads to significant computational resources being consumed for simulations. Bayesian Optimization (BO) and Surrogate-Assisted Evolutionary Algorithm (SAEA) are two widely used gradient-free optimization techniques employed to address such challenges. Both approaches follow a similar iterative procedure that relies on surrogate models to guide the search process. This paper aims to elucidate the similarities and differences in the utilization of model uncertainty between these two methods, as well as the impact of model inaccuracies on algorithmic performance. A novel model-assisted strategy is introduced, which utilizes unevaluated solutions to generate offspring, leveraging the population-based search capabilities of evolutionary algorithm to enhance the effectiveness of model-assisted optimization. Experimental results demonstrate that the proposed approach outperforms mainstream Bayesian optimization algorithms in terms of accuracy and efficiency.

Index Terms—Black-box optimization, Surrogate-Assisted Evolutionary Algorithm, Bayesian Optimization, Model Uncertainty

I. INTRODUCTION

Black-box optimization is a crucial technique used to optimize objective functions in various applications. The term “black-box” emphasizes the lack of access to derivative information, such as gradients, which necessitates the use of algorithms that adaptively select inputs for evaluation. The goal is to efficiently search for the global optimum. This approach is essential in fields like machine learning, where fine-tuning models with opaque internal weights is required [1], as well as in antenna design [2] and optimization of aerodynamic shapes in computational fluid dynamics (CFD), where the relationship between design parameters and performance is complex and highly non-linear [3], [4]. In each case, the optimization process aims to maximize or minimize outputs based on inputs, navigating the search space with a focus on

rapid convergence to the best solution while minimizing the number of evaluations required.

Bayesian Optimization (BO) [5] and Surrogate-Assisted Evolutionary Algorithm (SAEA) [3], [6] both significantly enhance the efficiency of black-box optimization. BO employs probabilistic models, such as Gaussian Processes or Random Forests, to estimate the objective function and uses an acquisition function to decide where to sample next. This technique effectively balances the need for exploration against the need for exploitation. Similarly, SAEA utilize surrogate models to approximate the objective function based on previous evaluations, guiding the evolutionary search toward promising areas and reducing the need for costly function evaluations [7], [8]. Both strategies are designed to converge on optimal solutions with fewer evaluations, which is particularly beneficial for problems where assessing the objective function is expensive.

Surrogate models are integral to both BO and SAEA, serving as replacements for computationally expensive black-box functions. Due to the intrinsic challenge that no model can perfectly replicate the target function, it becomes essential to evaluate and manage model errors. Uncertainty quantification plays a pivotal role in this context, providing a measure of the surrogate model’s error. BO and SAEA, while sharing similar search strategies, face parallel challenges. Among these are the computational burden associated with Gaussian Process (GP) modeling in medium to high-dimensional spaces, the precision of uncertainty estimation in the surrogate models, and the development of effective acquisition functions for model management [9]. Each of these issues is crucial to the performance and efficiency of the optimization algorithms, and addressing them can lead to more robust and faster convergence to the global optimum.

The research and application of model uncertainty have seen a wealth of excellent contributions, yet to this point, no studies have conducted a comparative exploration of this aspect between BO and SAEA. Moreover, an inherent distinction exists between BO and SAEA: BO is typically viewed as a single-point (sequential) search algorithm, necessitating high-quality surrogate models and precise uncertainty assessments

This work is supported by the National Natural Science Foundation of China (No.62306174), the China Postdoctoral Science Foundation (No.2023M74225, No.2023TQ0213) and the Postdoctoral Fellowship Program of CPSF under Grant Number (No.GZC20231588).

to steer its search. Conversely, SAEA operate on a population-based stochastic search paradigm. This distinction prompts the question: Is model uncertainty still a critical component in SAEA? How may we avert an excessive dependence on model accuracy, as well as a reliance on model uncertainty, which are difficult to achieve in high-dimensional spaces? This study aims to delineate the differences between BO and SAEA, as well as to investigate the role of model uncertainty within these two algorithmic frameworks. The contributions of this work are as follows:

- 1) This investigation endeavors to bridge the research gaps within the comparative domain of Bayesian optimization and model-assisted evolutionary computation. By employing a harmonized testing framework, the study meticulously contrasts the performance nuances of quintessential algorithms from each paradigm, thereby providing a detailed exposition of their inherent proficiencies and limitations.
- 2) The research elucidates the significance of model uncertainty, delineating its non-essentiality in the context of evolutionary algorithm (EA) while highlighting its criticality within Bayesian optimization. This discourse illuminates the relative merits and constraints inherent to both methodologies, offering a discerning understanding of their differential applications.
- 3) The work introduces a novel, generalized model management strategy specifically tailored for evolutionary computation, ingeniously conceived to circumvent the “curse of dimensionality” that typically plagues the computation of high-fidelity uncertainties.

The organization of the remainder of this paper is as follows. Section II provides the preliminaries, offering a clear introduction to the problems addressed and the algorithms involved. Section III conducts a comparative analysis of BO and SAEA, highlighting their similarities and differences, and proposes a novel strategy. Section IV details the systematic experiments conducted to evaluate the proposed strategy. Finally, Section V summarizes the findings and discusses potential avenues for future research.

II. PRELIMINARIES

A. Expensive Black-box Optimization Problem

Consider a black-box function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, where each evaluation of f is assumed to be expensive in terms of computational resources or time. The objective of the black-box optimization problem is to find [10]:

$$\mathbf{x}^* = \arg \min_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}) \quad (1)$$

where:

- $\mathbf{x} \in \mathbb{R}^n$ represents a vector of decision variables.
- $\mathcal{X} \subseteq \mathbb{R}^n$ denotes the feasible region within the decision variable space.
- $f(\mathbf{x})$ is the value of the objective function for a given vector \mathbf{x} , which is costly to evaluate.

- \mathbf{x}^* is the optimal solution, satisfying $f(\mathbf{x}^*) \leq f(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{X}$.

The function f is considered a black box as its explicit form is unknown, rendering gradient-based optimization methods inapplicable. Evaluating $f(\mathbf{x})$ typically demands substantial computational effort. The challenge lies in identifying the optimal \mathbf{x}^* with the fewest possible evaluations of f , owing to the significant cost associated with each evaluation.

B. Typical Algorithmic Frameworks

As quintessential methodologies for solving expensive black-box optimization problems, BO and SAEA share considerable common ground in both conceptual ideology and structural framework. This section is dedicated to presenting the fundamental frameworks of these two categories of algorithms, thereby setting the stage for subsequent comparative analysis. Algorithm 1 and Algorithm 2 outline the respective foundational frameworks of BO and SAEA.

Algorithm 1: Bayesian Optimization (BO)

Input: Objective function $f(\mathbf{x})$

Output: The best solution \mathbf{x}^* and its objective value $f(\mathbf{x}^*)$

- 1 Initialize observed data \mathbf{D} , select a surrogate model \mathcal{M} , and choose an acquisition function;
 - 2 **while** *stopping criterion not met* **do**
 - 3 Fit \mathcal{M} to \mathbf{D} ;
 - 4 Select the next query point \mathbf{x}_{next} using the acquisition function;
 - 5 Evaluate $f(\mathbf{x}_{next})$ and augment \mathbf{D} with the new observation $(\mathbf{x}_{next}, f(\mathbf{x}_{next}))$;
 - 6 **end**
-

1) *Bayesian Optimization:* In Algorithm 1, BO begins by initializing a set of observed data points, choosing a surrogate model, and selecting an acquisition function [11](line 1). It then fits the surrogate model to the observed data (line 3). Based on the fitted model, the algorithm computes the acquisition function to identify the next sampling point \mathbf{x}_{next} (line 4). Following this, the objective function f is evaluated at \mathbf{x}_{next} , and the observation is incorporated into the dataset \mathbf{D} . This iterative process of updating the model and acquiring new samples continues until a predefined stopping criterion is fulfilled, such as reaching a maximum number of iterations. The algorithm concludes by returning the best solution \mathbf{x}^* and its associated objective value $f(\mathbf{x}^*)$, which signifies the most favorable result discovered within the explored search space.

2) *Surrogate-assisted Evolutionary Algorithm:* In Algorithm 2, SAEA [] commences with the initialization of a population \mathbf{P} of N individuals, assessing their fitness using the objective function f . The population \mathbf{P} typically forms the initial training set \mathbf{D} for the surrogate model \mathcal{M} (line 1). With each iteration, \mathcal{M} is trained on the dataset \mathbf{D} to approximate the fitness landscape (line 3). Offspring \mathbf{O} are produced from \mathbf{P} using reproductive operators such as selection, crossover,

Algorithm 2: Surrogate-assisted Evolutionary Algorithm (SAEA)

Input: Objective function $f(\mathbf{x})$, population size N

Output: The best solution \mathbf{x}^* and its fitness $f(\mathbf{x}^*)$

- 1 Initialize a population \mathbf{P} of N individuals, evaluate their fitness with f , start surrogate model \mathcal{M} , and form dataset \mathbf{D} ;
 - 2 **while** *stopping criterion not met* **do**
 - 3 Train \mathcal{M} using \mathbf{D} ;
 - 4 Generate offspring \mathbf{O} by applying reproductive operators to \mathbf{P} ;
 - 5 Select a subset of \mathbf{O} for real evaluation based on \mathcal{M} predictions and evaluate their fitness with f ;
 - 6 Integrate the newly evaluated individuals into \mathbf{P} and update the dataset \mathbf{D} ;
 - 7 **end**
-

and mutation (line 4). A subset of these offspring is then selected for true fitness evaluation, guided by the predictions of \mathcal{M} (line 5). The evaluated individuals are integrated into the population \mathbf{P} , and the dataset \mathbf{D} is updated accordingly (line 6). This process is repeated until a stopping criterion is reached, for example, a maximum number of fitness evaluations (FEs). The algorithm yields the best solution \mathbf{x}^* found during the evolutionary process, along with its fitness value $f(\mathbf{x}^*)$.

Both algorithms rely on model-based optimization techniques, which require updating the model at each iteration to inform the subsequent search phase. Although there are similarities in their reliance on a model to guide the search, the specific application and integration of the model within each algorithm differ, which will be explored in detail in the following section.

III. MODEL UNCERTAINTY IN BOS AND SAEAS

A. Comparative Analysis

This subsection delves into the search patterns and the impact of model uncertainty within the frameworks of BO and SAEA through low-dimensional visualization. We consider the function defined in Equation (2), which has its peak at $x = 7.99$, yielding $f(x) = 7.92$, where x is in the range $[0, 10]$. The surrogate models selected are Gaussian Processes (GPs) [12] and Random Forests (RFs) [13], both of which are extensively employed in the domains of BO and SAEA. Seven observational data points are initially provided, each subject to stochastic perturbations, with the random seed being held constant. The resulting visualizations are depicted in Figures 1 and 2.

$$f(x) = x \sin(x) \quad (2)$$

1) *BO Framework:* Fig. 1 illustrates the BO process (Algorithm 1), highlighting how a surrogate model is constructed from observational data (line 3) and how the acquisition

function, specifically Expected Improvement (EI) [14], guides the selection of points for subsequent iterations (line 4).

Fig. 1 reveals that the GP outperforms the RF in terms of function fitting, providing a more accurate estimation of uncertainty. For example, in the interval from 0 to 4, there is a significant discrepancy between the fitted curve and the true function; the GP model accurately estimates this divergence, while the RF model does not. Conversely, in the interval from 6 to 10, where the divergence between the fitted and true curves is minimal, the GP model exhibits lower uncertainty, whereas the uncertainty predicted by the RF model remains relatively high. Further analysis is provided by the green curve in the diagram, which represents the values of EI; a higher value of EI indicates greater exploratory potential at that point. The next point to be sampled, \mathbf{x}_{next} , selected based on the GP model, is closer to the real optimum, whereas the \mathbf{x}_{next} chosen using the RF model deviates from the true optimum. Additionally, by exchanging the uncertainty estimations of the two models, we obtain a comparative curve, EI*. The figure clearly shows that when the GP model is misled by inaccurate uncertainty estimation, there is a distortion in the acquisition function selection. Similarly, the RF model, hindered by its poor fitting capabilities, experiences a shift in acquisition choice, even when it employs a precise estimate of uncertainty.

Therefore, it is evident that accurate function fitting is fundamental to the operation of BO, while a reliable assessment of uncertainty is crucial for the accurate search of the acquisition function.

2) *SAEA Framework:* The essence of EA lies in the execution of selection and generation, methodologies that are fundamentally population-based. Fig. 2 illustrates this process. Within the figure, the blue points (denoted as \mathbf{O}_{all}) represent the top 20% of points, chosen from a uniformly distributed sample of 100 points within the 0-10 interval based solely on the model's predicted values. The orange points represent the optimal points as perceived by the surrogate model. Although the fitting capabilities of the RF is inferior by those of the GP, the RF still captures the essential features of the functional landscape. Consequently, this has a minimal impact on the selection from \mathbf{O}_{all} , which underscores the robustness of population-based selection strategies in overcoming model inaccuracies.

Subsequently, the process of reproduction takes place. We investigate the effects that the number of offspring (\mathbf{O}_{all} or \mathbf{O}_{best}), as selected by the surrogate model, has on the distribution of new solutions in the following iteration. The green point \mathbf{P} represents the current distribution of the parental population. When combined separately with \mathbf{O}_{best} and \mathbf{O}_{all} , and employing the reproduction operator [15], 10,000 new solutions were sampled, as depicted by the orange and blue shaded areas in Fig. 2. The results show that new solutions derived from $\mathbf{P} \cup \mathbf{O}_{all}$ have a significantly higher probability of appearing near the optimal region compared to those from $\mathbf{P} \cup \mathbf{O}_{best}$, regardless of whether the GP or RF model is used for selection.

This indicates that updating the population with $\mathbf{P} \cup \mathbf{O}_{all}$

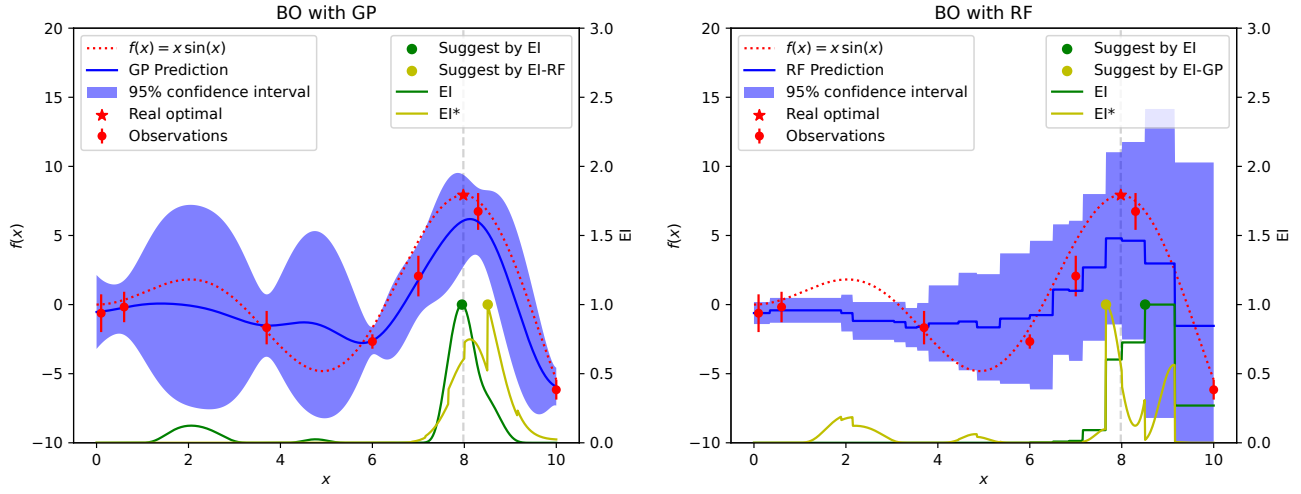


Fig. 1: Performance comparison of GP and RF in uncertainty estimation and function fitting in BO.

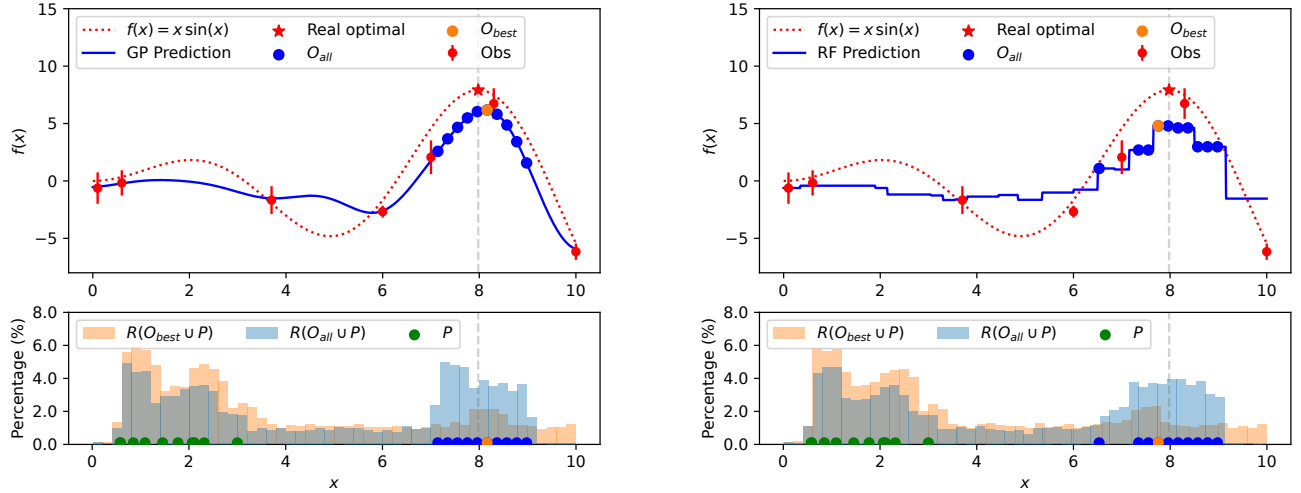


Fig. 2: Performance comparison of GP and RF in function fitting and new solution reproduction in SAEA.

can effectively improve the quality of new solutions emerging from the population, while relying solely on $\mathbf{P} \cup \mathbf{O}_{best}$ may lead to a distribution of new solutions that diverges from the optimal area, thereby reducing the quality of the solutions. Furthermore, during this process, model uncertainty does not seem to significantly influence population iteration, and the difference in modeling accuracy between GP and RF appears to be mitigated, to some extent, by the evolutionary process.

B. New Model Management Strategy

In EAs, the use of a population-based search and generation methodology can mitigate the impact of model accuracy issues, necessitating reliance on a broader population (\mathbf{O}_{all}) rather than on isolated solutions (\mathbf{O}_{best}). Reproduction does not depend on the objective function f , making it a prominent strategy to integrate the subset \mathbf{O}_{all} , selected by the model, into the population without evaluated. This approach is preferred over exclusively incorporating the optimal individuals

\mathbf{O}_{best} . This strategy not only ensures the diversity of the population but also maintains the quality of the population and the continuous updating of the model training data.

As a straightforward and universally applicable strategy, it requires only minor modifications to Algorithm 2:

- Line 5: Select a subset \mathbf{O}_{all} from \mathbf{O} using the surrogate model \mathcal{M} , and identify the most promising candidate \mathbf{O}_{best} within this subset for fitness evaluation.
- Line 6: Use \mathbf{O}_{best} to enhance the dataset \mathbf{D} , and select the top N solutions from \mathbf{D} to merge with \mathbf{O}_{all} to form the new population \mathbf{P} .

In the execution of this strategy, a distribution estimation algorithm known as the variable histogram serves as the reproduction operator, a distribution estimation algorithm (EDA) known as variable-width histogram (VWH) [15] is utilized as the reproduction operator. Due to the inclusion of solutions that have not been evaluated, this approach is termed the unevaluated solution-derived EDA (UEDA). By default, the

population size N is set to 50, and O_{All} is set to $N/2$. The dataset D is pruned to maintain a maximum size of 100 through a survival-of-the-fittest selection process based on f .

IV. EXPERIMENTAL RESULTS

A. Experimental Setup

1) **Comparison Algorithms:** For comparative analysis, this section identifies three representative Bayesian optimization toolkits: the BayesianOptimization toolkit (BOT) [16], Heteroscedastic Evolutionary Bayesian Optimisation (HEBO) [17], and Scikit-Optimize (Skopt) [18]. These toolkits are based on Bayesian optimization algorithms and possess the following principal attributes:

- BOT: Incorporates sequential domain reduction [19] within the standard BO framework, significantly accelerating the search progress and hastening convergence.
- HEBO: Implements input and output warping to reduce the effects of heteroscedasticity and non-stationarity in the objective function. It employs multi-objective optimization to optimize multiple acquisition functions concurrently, enhancing the algorithm's robustness.
- Skopt: Uses the Hedge strategy [20] to probabilistically select among different acquisition functions at each iteration, increasing the robustness of the search process.

Additionally, two surrogate-assisted evolutionary algorithms (SAEAs) are utilized in the experiments: the surrogate-assisted multiswarm optimization (SAMSO) [21] and the fuzzy classification-based preselection for CoDE (FCPS-CoDE) [22], which exemplify EAs assisted by regression and classification models, respectively. An EDA with a local search (EDA/LS) [15] serves as the baseline algorithm for this study and is included in the experimental analysis. A detailed description is as follows:

- SAMSO: Utilizes a dual-swarm approach with cross-swarm intelligence sharing and dynamic swarm size regulation, guided by a Radial Basis Function (RBF) to estimate the evaluation function.
- FCPS-CoDE: Employs a fuzzy K-nearest neighbors (KNN) classifier to predict the solution categories before evaluation, aiding the Composite Differential Evolution (CoDE) algorithm in its evolutionary search.
- EDA/LS: Integrates an Estimation of Distribution Algorithm (EDA) using a Variable Width Histogram (VWH) and includes a local search operator to facilitate the evolutionary search.

The innovations of this work, which incorporate Random Forest (UEDA-RF), Extreme Gradient Boosting [23](UEDA-XG), and Gaussian Process (UEDA-GP), also participate in the experiments.

2) **Test Suites:** The test suites include the LZG [24] and YLL [25] test sets. LZG comprises functions like Ellipsoid, Rosenbrock, Ackley, and Griewank, which present diverse landscapes, including unimodal, gully, and multimodal terrains. YLLF01-13 features characteristics such as unimodal, multimodal, stepwise, and stochastic noise (excluding YLLF10-11 due to overlap with LZG).

3) **Parameter Settings:** To ensure fair comparison, the recommended parameters from the original literature are adopted for each algorithm¹. The specific parameters are as follows:

- Maximum evaluation budget: 500.
- Population size: Set to $N = 30$ for EDA/LS and FCPS-CoDE, $N = 40$ for SAMSO, and $N = 50$ for UEDA-GP, UEDA-XG, and UEDA-RF.
- Dimensionality: All test problems are set with dimensions $n = 20$ and $n = 50$.
- Surrogate model: GPs are used for $n = 20$ in the three BO toolkits. For $n = 50$, RFs are preferred due to lower computational costs. (BOT does not offer an RF option.)

Each algorithm undergoes 30 independent runs on each test instance to accommodate randomness. The Wilcoxon rank-sum test [27] is applied to analyze the results. In the tables, the symbols '+', '-', and '~' indicate comparative performance against the benchmark algorithms (UEDA-GP for $n = 20$ and UEDA-RF for $n = 50$), with '+' denoting significantly better performance, '-' indicating significantly worse, and '~' signifying no significant difference at the 0.05 significance level.

B. Performance Analysis

This section compares the performance of various algorithms across two aspects: solution accuracy and runtime. Table I displays the problem-solving precision of each algorithm at varying dimensions. Observations from the table suggest that for lower-dimensional problems ($n=20$), GPs and XGBoost-assisted UEDAs surpass the RF model in effectiveness but fall short when compared to HEBO-GP. Among nine algorithms, HEBO-GP achieved a mean rank of 1.73 across 15 test problems, demonstrating a distinct advantage over the algorithm proposed in this chapter, UEDA-GP, which resulted in 12 superior outcomes, 2 inferior, and one approximate to HEBO-GP. UEDA-GP, while behind HEBO-GP, nonetheless showed significant superiority over other BO algorithms and also outperformed the compared EA methods. Notably, UEDA-RF has proven its prowess by outperforming BOT-GP, thereby confirming that a solution without evaluation can compensate for RF's lack of modeling accuracy, and in some instances, reach the level of BO. In the case of higher-dimensional problems ($n=50$), UEDA-RF achieved a mean rank of 2.8, which, from the viewpoint of statistical significance, indicates a clear advantage. Specifically, RF-UEDA was superior to GP-UEDA within the same framework and outperformed all BOs tested (even BOT using a GP) and was markedly better than EAs. Alternatively, HEBO when switching to an RF model, experienced a significant decline in performance, attaining the worst mean rank of 8.53. This suggests that the unreliable uncertainty estimates of RF models can mislead the exploratory direction of the algorithm.

This section compares the performance of various algorithms in terms of solution accuracy and runtime. Table I

¹The Bayesian optimization packages are used with default settings; SAMSO is implemented within Platemo [26]; FCPS-CoDE and EDA/LS are realized by us, following the original documentation.

presents the problem-solving precision of each algorithm at different dimensions. Observations from the table indicate that for lower-dimensional problems ($n = 20$), GP and XGBoost-assisted UEDAs outperform the RF in effectiveness but do not match the performance of HEBO-GP. Among nine algorithms, HEBO-GP achieved a mean rank of 1.73 across 15 test problems. UEDA-GP resulted in 12 superior outcomes, 2 inferior, and one similar to HEBO-GP. While UEDA-GP trails behind HEBO-GP, it still shows significant superiority over other BOs and also outperforms the compared EAs. Notably, UEDA-RF has proven its effectiveness by outperforming BOT-GP, confirming that solutions without evaluation can compensate for RF's lack of modeling accuracy and, in some cases, achieve the level of BO. For higher-dimensional problems ($n = 50$), UEDA-RF achieved a mean rank of 2.8, which statistically signifies a clear advantage. Specifically, RF-UEDA was superior within the same framework to GP-UEDA and outperformed all tested BOs (including BOT using a GP) and was markedly better than EAs. Conversely, when HEBO switched to an RF model, there was a significant decline in performance, with the worst mean rank of 8.53. This indicates that the unreliable uncertainty estimates of RF models can mislead the exploratory direction of the algorithm.

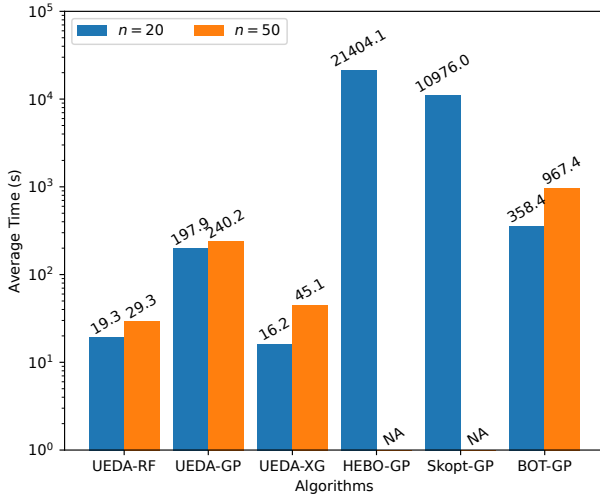


Fig. 3: Runtime comparison of three BOs and UEDA with different surrogates on LZG test suits with $n = 20, 50$ (NA signifies an impractical computational cost for experiment completion).

Fig. 3 presents the runtime statistics for three BOs and the UEDAs incorporating different surrogate models. The evaluations were uniformly conducted on the same hardware configuration (CPU: i9-13900k, Memory: 64GB, Ubuntu 22.04), addressing LZG01-04 problems at both 20 and 50 dimensions. This approach allowed for the determination of the average resolution times for each algorithm at different problem scales. The absence of data for HEBO-GP and Skopt-GP at $n = 50$ signifies the computational cost was infeasible, hence it is denoted as ‘NA’.

The statistical results reveal that the UEDA algorithm significantly reduces runtime compared to the BO algorithms, particularly in comparison with HEBO-GP and Skopt-GP, by more than two orders of magnitude. This efficiency is attributed to UEDA’s use of a more lightweight surrogate model and a reduced amount of model training data.

The experiments lead to the conclusion that while BO algorithms maintain a significant edge in low-dimensional problems, the modeling process involving GPs requires a substantial computational budget. In higher-dimensional, GPs are less applicable due to the curse of dimensionality, while RFs can somewhat mitigate running costs. However, RFs’ imprecision in fitting and uncertainty assessment may mislead the optimization process, leading to a significant decline in BO performance. Conversely, the UEDA framework proposed in this study exploits the characteristics of population-based searches by incorporating information from unevaluated solutions to enhance search efficiency. It also leverages the strengths of RFs to reduce running costs and compensates for issues with modeling accuracy, while avoiding reliance on inaccurate uncertainty assessments, thus exhibiting robust advantages across low to moderate dimensional spaces.

C. Ablation Study

This section substantiates the effectiveness of the presented UEDA framework (utilizing RF as the surrogate model) through ablation studies. Variants are described as presented in Table II. The ablation experiments investigate the advantages of UEDA in comparison to the baseline EDA/LS, as well as the impact on the algorithm’s performance of evaluating all solutions selected by the surrogate model and the exclusion of unevaluated solutions in reproduction. The experiments are conducted on the LZG test suite, with problem dimension settings of $n = 20, 50$. Each experimental set is independently replicated 30 times to mitigate randomness.

The resulting curves of the objective values versus the number of fitness evaluations (FEs) are illustrated in Fig. 4, where the lines represent the mean values of multiple trials, and the shaded areas indicate the standard deviation.

The experimental results are manifestly evident; UEDA-RF outperforms both variants as well as the baseline across all problems. This underscores the efficacy of UEDA and the significance of surrogate assisted selection and reproduction strategies. The performance of UEDA-RF-AL surpasses that of UEDA-RF-NS, and both excel over the baseline algorithm. Upon further analysis, it is apparent that EDA/LS was not tailored for expensive optimization problems, as a fixed number of individuals per generation is evaluated, incurring computational expenses. With limited computational resources (FEs = 500), the search is insufficient. UEDA-RF-ALL evaluates only half the number of individuals selected by the surrogate model per generation, effectively increasing the number of evolutionary generations and henceforth delivering superior performance. UEDA-RF-NS evaluates one individual per generation, which further economizes on computational resources; however, the loss of information from unevaluated

TABLE I: Statistics of mean and standard deviation results obtained by nine comparison algorithms on LZG and YLL test suites with $n = 20, 50$, adhere to a maximum evaluation budget of 500.

$n = 20$									
Problem	UEDA-GP	UEDA-XG	UEDA-RF	HEBO-GP	Skopt-GP	BOT-GP	EDA/LS	SAMSO	FCPS
Ellipsoid	8.63e-03 [1] (8.36e-03)	8.13e+00[5](−) (4.14e+00)	1.18e+01[6](−) (5.22e+00)	1.41e-01[3](−) (3.84e-02)	6.58e-02[2](−) (1.75e-02)	1.55e-01[4](−) (2.83e-02)	7.17e+01[8](−) (1.56e+01)	1.87e+01[7](−) (2.52e+01)	1.30e+02[9](−) (3.13e+01)
Rosenbrock	4.79e+01 [3] (2.55e+01)	9.78e+01[6](−) (3.99e+01)	9.68e+01[5](−) (3.63e+01)	2.18e+01[1](+) (1.10e+01)	5.42e+01[4](≈) (1.25e+01)	1.25e+02[7](−) (4.24e+01)	2.37e+02[8](−) (4.02e+01)	3.57e+01[2](≈) (2.47e+01)	3.22e+02[9](−) (1.05e+02)
Ackley	5.29e+00 [3] (1.37e+00)	8.17e+00[5](−) (1.17e+00)	8.69e+00[6](−) (1.11e+00)	9.10e-01[1](+) (3.43e-01)	7.12e+00[4](−) (4.46e-01)	5.19e+00[2](≈) (2.39e+00)	1.33e+01[7](−) (7.37e-01)	1.83e+01[9](−) (1.33e+00)	1.48e+01[8](−) (1.00e+00)
Griewank	2.12e+01 [7] (5.28e+00)	4.05e+00[4](+) (1.39e+00)	5.33e+00[5](+) (1.62e+00)	7.82e-01[1](+) (9.12e-02)	1.02e+00[2](+) (2.18e-02)	1.43e+00[3](+) (1.46e-01)	2.96e+01[8](−) (7.62e+00)	2.06e+01[6](≈) (1.33e+01)	5.46e+01[9](−) (1.33e+01)
YLLF01	2.08e+03 [7] (5.25e+02)	3.83e+02[4](+) (2.31e+02)	4.09e+02[5](+) (1.96e+02)	5.30e+00[2](+) (1.70e+00)	3.66e+00[1](+) (1.16e+00)	1.01e+01[3](+) (2.70e+00)	3.17e+03[8](−) (5.92e+02)	6.73e+02[6](+) (6.43e+02)	5.37e+03[9](−) (1.86e+03)
YLLF02	1.79e+00 [1] (9.15e-01)	6.07e+00[3](−) (1.15e+00)	3.89e+00[2](−) (9.78e-01)	2.04e+01[4](−) (6.62e+00)	4.62e+05[9](−) (9.03e+05)	3.89e+04[8](−) (1.21e+05)	2.67e+01[6](−) (3.93e+00)	3.27e+01[7](−) (1.72e+01)	2.41e+01[5](−) (3.67e+00)
YLLF03	1.76e+04 [7] (3.51e+03)	1.42e+04[4](+) (3.54e+03)	8.81e+03[2](+) (3.59e+03)	1.68e+04[5](≈) (4.33e+03)	7.47e+03[1](+) (2.69e+03)	3.50e+04[9](−) (6.98e+03)	2.32e+04[8](−) (4.21e+03)	1.71e+04[6](≈) (1.38e+04)	1.26e+04[3](+) (3.48e+03)
YLLF04	2.74e+01 [2] (3.22e+00)	2.96e+01[5](≈) (7.88e+00)	2.97e+01[6](≈) (5.41e+00)	2.14e+01[1](+) (1.14e+01)	2.78e+01[3](−) (1.72e+01)	2.82e+01[4](≈) (9.43e+00)	3.28e+01[7](−) (3.26e+00)	7.06e+01[9](−) (7.40e+00)	4.17e+01[8](−) (6.55e+00)
YLLF05	3.82e+04 [2] (3.34e+04)	4.63e+04[3](≈) (3.51e+04)	8.77e+04[4](−) (9.39e+04)	3.90e+02[1](+) (6.20e+02)	1.77e+05[6](−) (8.96e+04)	2.34e+06[8](−) (1.04e+06)	1.13e+06[7](−) (5.76e+05)	1.12e+05[5](−) (8.27e+04)	4.71e+06[9](−) (4.71e+06)
YLLF06	2.08e+03 [7] (5.49e+02)	3.32e+02[4](+) (1.57e+02)	5.12e+02[5](+) (2.19e+02)	6.97e+00[2](+) (2.09e+00)	5.30e+00[1](+) (1.53e+00)	9.80e+00[3](+) (3.19e+00)	3.20e+03[8](−) (5.04e+02)	8.59e+02[6](+) (1.24e+03)	6.26e+03[9](−) (1.28e+03)
YLLF07	1.39e-01 [2] (5.67e-02)	2.43e-01[3](−) (1.13e-01)	2.66e-01[4](−) (1.57e-01)	9.19e-02[1](+) (3.87e-02)	2.67e-01[5](−) (9.91e-02)	1.05e+00[8](−) (4.24e-01)	6.67e-01[7](−) (2.75e-01)	3.23e-01[6](−) (1.56e-01)	2.09e+00[9](−) (9.17e-01)
YLLF08	4.97e+03 [5] (3.75e+02)	3.47e+03[3](+) (5.53e+02)	2.49e+03[2](+) (5.24e+02)	1.86e+03[1](+) (3.48e+02)	5.17e+03[6](≈) (3.29e+02)	5.54e+03[8](−) (5.52e+02)	4.66e+03[4](+) (3.32e+02)	5.67e+03[9](−) (2.83e+02)	5.52e+03[7](−) (2.88e+02)
YLLF09	1.12e+02 [3] (1.63e+01)	1.32e+02[6](−) (1.55e+01)	1.28e+02[5](−) (1.60e+01)	8.66e+01[1](+) (1.96e+01)	1.63e+02[8](−) (1.86e+01)	1.65e+02[9](−) (3.53e+01)	1.57e+02[7](−) (1.31e+01)	1.16e+02[4](≈) (5.14e+01)	1.02e+02[2](+) (2.09e+01)
YLLF12	5.21e+03 [5] (8.76e+03)	4.28e+03[3](+) (1.76e+03)	1.87e+03[4](+) (9.07e+03)	2.47e+00[1](+) (1.61e+00)	6.68e+04[7](−) (1.21e+05)	8.77e+06[9](−) (5.76e+06)	6.09e+04[6](−) (8.76e+04)	1.23e+02[2](+) (5.49e+02)	2.46e+06[8](−) (3.04e+06)
YLLF13	1.09e+11 [8] (4.44e+10)	6.96e+09[5](+) (8.11e+09)	1.24e+10[6](+) (1.32e+10)	3.17e+04[1](+) (6.92e+04)	2.72e+10[7](+) (1.23e+10)	8.77e+11[9](−) (2.20e+11)	1.10e+06[3](+) (5.84e+05)	6.82e+04[2](+) (1.94e+05)	1.08e+07[4](+) (8.07e+06)
mean rank + / − / ≈	4.20 /	4.20 7/6/2	4.47 7/7/1	1.73 12/2/1	4.40 5/8/2	6.27 3/10/2	6.80 2/13/0	5.73 4/7/4	7.20 3/12/0
$n = 50$									
Problem	UEDA-RF	UEDA-XG	UEDA-GP	HEBO-RF	Skopt-RF	BOT-GP	EDA/LS	SAMSO	FCPS
Ellipsoid	8.48e+02 [3] (1.63e+02)	1.08e+03[4](−) (3.13e+02)	2.81e+02[2](+) (6.92e+01)	3.59e+03[8](−) (4.78e+02)	5.87e+03[9](−) (4.37e+02)	7.52e+00[1](+) (3.19e+00)	1.52e+03[6](−) (2.23e+02)	1.31e+03[5](≈) (1.13e+03)	1.61e+03[7](−) (3.39e+02)
Rosenbrock	1.06e+03 [2] (2.00e+02)	1.41e+03[4](−) (5.21e+02)	7.00e+02[2](+) (1.21e+02)	5.26e+03[8](−) (7.33e+02)	9.59e+03[9](−) (1.03e+03)	5.02e+02[1](+) (1.10e+02)	1.78e+03[6](−) (2.84e+02)	1.58e+03[5](≈) (1.70e+03)	1.90e+03[7](−) (5.11e+02)
Ackley	1.65e+01 [3] (5.42e-01)	1.69e+01[4](≈) (8.67e-01)	1.63e+01[2](+) (4.27e-01)	2.00e+01[8](−) (2.71e-01)	2.06e+01[9](−) (1.40e-01)	8.36e+00[1](+) (6.22e-01)	1.76e+01[6](−) (4.24e-01)	1.86e+01[7](−) (1.18e+00)	1.75e+01[5](−) (6.00e-01)
Griewank	1.52e+02 [2] (2.58e+01)	1.82e+02[4](−) (5.42e+01)	1.72e+02[3](−) (2.49e+01)	5.52e+02[7](−) (7.86e+01)	8.83e+02[9](−) (5.97e+01)	3.74e+00[1](+) (5.77e-01)	2.41e+02[5](−) (2.86e+01)	6.19e+02[8](−) (3.66e+02)	2.69e+02[6](−) (6.72e+01)
YLLF01	1.53e+04 [2] (2.38e+03)	1.99e+04[4](−) (5.57e+03)	1.94e+04[3](−) (1.99e+03)	6.17e+04[7](−) (9.04e+03)	9.56e+04[9](−) (8.96e+03)	6.58e+01[1](+) (8.57e+00)	2.77e+04[5](−) (3.92e+03)	6.69e+04[8](−) (3.70e+04)	2.95e+04[6](−) (5.86e+03)
YLLF02	4.81e+01 [1] (4.05e+00)	8.01e+01[3](−) (8.83e+00)	5.73e+01[2](−) (1.15e+01)	3.31e+04[6](−) (1.71e+05)	8.70e+17[8](−) (4.40e+18)	3.06e+17[7](−) (1.18e+18)	2.79e+04[5](−) (9.04e+04)	1.05e+18[9](−) (3.77e+18)	8.92e+01[4](−) (8.24e+00)
YLLF03	1.06e+05 [2] (1.86e+04)	1.32e+05[3](−) (2.88e+04)	1.35e+05[5](−) (3.00e+04)	1.33e+05[4](−) (2.57e+04)	1.83e+05[7](−) (3.44e+04)	1.85e+05[8](−) (3.21e+04)	1.58e+05[6](−) (2.51e+04)	2.97e+05[9](−) (9.31e+04)	8.24e+04[1](+) (1.68e+04)
YLLF04	6.15e+01 [4] (3.24e+00)	6.17e+01[5](≈) (6.71e+00)	5.42e+01[1](+) (3.58e+00)	8.08e+01[7](−) (4.20e+00)	9.05e+01[9](−) (2.23e+00)	7.28e+01[6](−) (1.52e+01)	5.80e+01[2](+) (2.78e+00)	8.74e+01[8](−) (8.06e+00)	5.88e+01[3](≈) (5.58e+00)
YLLF05	1.40e+07 [2] (5.12e+06)	2.36e+07[4](−) (1.50e+07)	1.91e+07[3](−) (4.17e+06)	1.74e+08[8](−) (4.91e+07)	3.36e+08[9](−) (3.76e+07)	8.64e+06[1](+) (3.24e+06)	2.93e+07[5](−) (5.30e+06)	1.55e+08[7](−) (8.04e+07)	4.49e+07[6](−) (1.82e+07)
YLLF06	1.67e+04 [2] (3.02e+03)	2.06e+04[4](≈) (7.01e+03)	1.96e+04[3](−) (2.74e+03)	6.22e+04[7](−) (7.44e+03)	9.65e+04[9](−) (7.35e+03)	7.19e+01[1](+) (1.43e+01)	2.74e+04[5](−) (4.41e+03)	7.82e+04[8](−) (3.06e+04)	2.92e+04[6](−) (7.35e+03)
YLLF07	9.11e+00 [3] (2.64e+00)	1.56e+01[4](−) (7.06e+00)	5.60e+00[1](+) (2.34e+00)	1.25e+02[7](−) (3.54e+01)	2.56e+02[9](−) (3.52e+01)	8.35e+00[2](≈) (2.96e+00)	2.17e+01[5](−) (4.75e+00)	1.31e+02[8](−) (7.91e+01)	2.96e+01[6](−) (1.28e+01)
YLLF08	1.39e+04 [2] (6.52e+02)	1.51e+04[4](−) (6.17e+02)	1.59e+04[6](−) (5.00e+02)	1.09e+04[1](+) (8.40e+02)	1.58e+04[5](−) (6.43e+02)	1.69e+04[9](−) (5.53e+02)	1.51e+04[3](−) (5.84e+02)	1.66e+04[8](−) (6.62e+02)	1.63e+04[7](−) (6.11e+02)
YLLF09	5.24e+02 [7] (2.52e+01)	5.15e+02[4](≈) (3.91e+01)	4.42e+02[2](+) (2.42e+01)	5.28e+02[8](≈) (3.68e+01)	6.98e+02[9](−) (3.04e+01)	5.18e+02[5](≈) (3.64e+01)	5.19e+02[6](≈) (1.78e+01)	4.97e+02[3](≈) (1.11e+02)	3.80e+02[1](+) (3.15e+01)
YLLF12	7.18e+06 [1] (5.94e+06)	1.87e+07[4](≈) (1.88e+07)	8.64e+06[2](−) (4.03e+06)	3.20e+08[7](−) (1.17e+08)	7.46e+08[9](−) (1.24e+08)	2.10e+07[5](−) (1.73e+07)	1.64e+07[3](−) (5.98e+06)	5.10e+08[8](−) (3.16e+08)	5.44e+07[6](−) (3.49e+07)
YLLF13	2.03e+12 [5] (6.41e+11)	3.16e+12[7](≈) (1.83e+12)	2.75e+12[6](−) (5.74e+11)	2.81e+13[8](−) (7.18e+12)	5.42e+13[9](−) (6.88e+12)	1.39e+12[4](+) (6.90e+11)	7.53e+07[1](+) (2.19e+07)	8.32e+08[3](+) (6.39e+08)	1.58e+08[2](+) (1.01e+08)
mean rank + / − / ≈	2.80 /	4.13 0/9/6	2.87 6/9/0	6.73 1/13/1	8.53 0/15/0	3.53 8/5/2	4.60 2/12/1	6.93 1/11/3	4.87 3/11/1

TABLE II: Algorithm variants for ablation study.

Algorithm	Description
UEDA-RF	UEDA with RF surrogate
UEDA-RF-AL	Evaluate all solutions selected by the model
UEDA-RF-NS	Excludes unevaluated solutions from reproduction
EDA/LS [15]	The baseline algorithm for UEDA

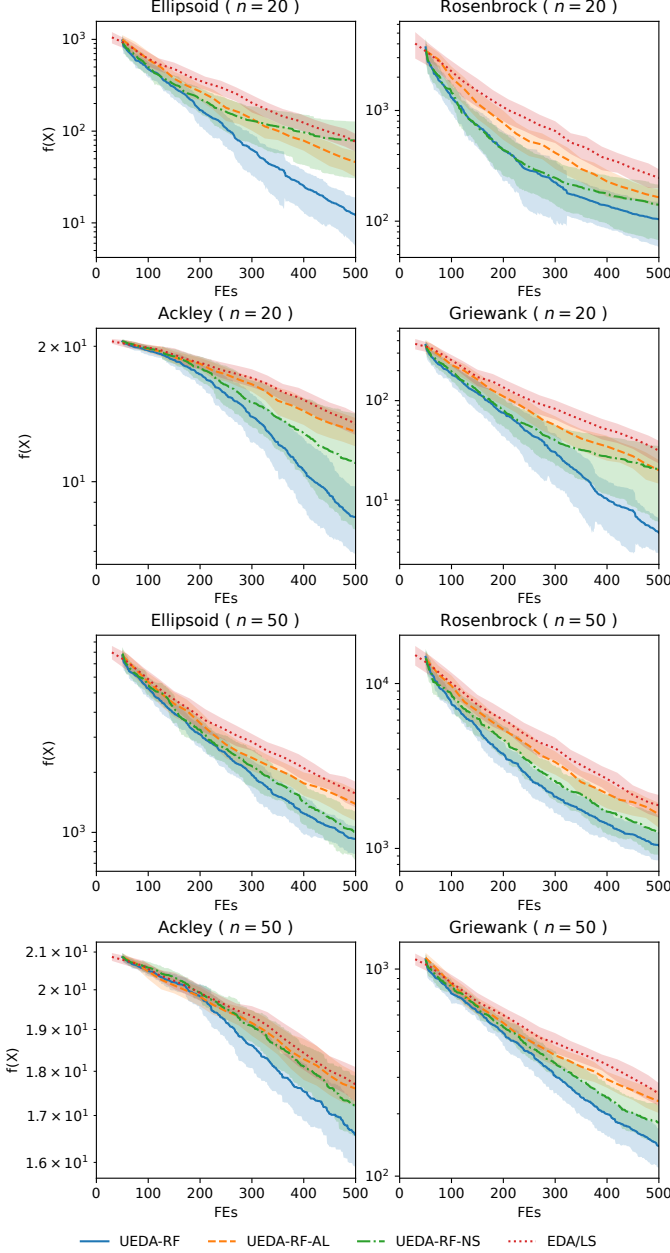


Fig. 4: Comparison of EDA/LS and UEDAs on LZG test suits with $n = 20, 50$.

solutions deteriorates the quality of solutions generated by reproduction, hence underperforming compared to the UEDA presented in this work. In summary, UEDA can effectively conserve computational expenses while ensuring the quality

of new solutions, thereby enhancing search efficiency.

This section substantiates the effectiveness of the presented UEDA framework (utilizing RF as the surrogate model) through ablation studies. Variants are described as presented in Table II. The ablation experiments investigate the advantages of UEDA in comparison to the baseline EDA/LS, as well as the impact on the algorithm’s performance of evaluating O_{all} and the exclusion of unevaluated solutions in reproduction. The experiments are conducted on the LZG test suite, with problem dimensions set at $n = 20$ and $n = 50$. Each experimental set is independently replicated 30 times to mitigate the effects of randomness.

The resulting curves of the objective values versus the number of fitness evaluations (FEs) are illustrated in Fig. 4, where the lines represent the mean values of multiple trials, and the shaded areas indicate the standard deviation.

The experimental results are unequivocally clear; UEDA-RF outperforms both variants as well as the baseline across all problems. This underscores the efficacy of UEDA. The performance of UEDA-RF-AL surpasses that of UEDA-RF-NS, with both outperforming the baseline algorithm. Upon further analysis, it is evident that EDA/LS was not designed for expensive optimization problems, as it evaluates a fixed number of individuals per generation, leading to high computational costs. With limited computational resources (FEs = 500), the search proves to be insufficient. UEDA-RF-ALL evaluates only half the number of individuals selected by the surrogate model per generation, effectively increasing the number of evolutionary generations and thereby delivering superior performance. UEDA-RF-NS evaluates one individual per generation, which further economizes on computational resources; however, the loss of information from unevaluated solutions deteriorates the quality of the solutions generated by reproduction, thus underperforming compared to the UEDA framework presented in this work. In summary, UEDA can effectively conserve computational expenses while ensuring the quality of new solutions, thereby enhancing search efficiency.

V. CONCLUSION

This work began with an exploration of model uncertainty, comparing and contrasting the search behaviors of BO and SAEA. A visual analysis of one-dimensional functions has revealed that an evolutionary algorithm, which employs selection and reproduction based on populations, can mitigate the reliance on acquisition functions derived from model uncertainty. Furthermore, it can sidestep the misdirection caused by inherent model errors. Building on this insight, we proposed a novel evaluation-free iterative strategy that harnesses the strengths of population-based search while reducing evaluation costs. This strategy was implemented as the UEDA, which outperformed three mainstream BOs and two SAEAs in systematic experimental comparisons, thus demonstrating its effectiveness. Notably, UEDA also surpassed BOs in time efficiency by two orders of magnitude. Ablation studies further highlighted the significance of integrating evaluation-free solutions.

However, several avenues for future research remain open. First, the comparative analysis of BO and SAEA search behaviors was confined to one-dimensional functions, which may not capture the potential disparities in high-dimensional search spaces and complex functions. Therefore, a more systematic analysis is warranted. Second, as a general strategy, the application of evaluation-free solutions to other evolutionary search operators, such as genetic algorithms (GA) and differential evolution (DE), deserves further investigation. Lastly, the impact of different surrogate models on the performance of UEDA presents another promising area of study.

REFERENCES

- [1] M. Gasse, S. Bowly, Q. Cappart *et al.*, “The machine learning for combinatorial optimization competition (ml4co): Results and insights,” in *Proceedings of the NeurIPS 2021 Competitions and Demonstrations Track*, ser. Proceedings of Machine Learning Research, vol. 176. PMLR, 06–14 Dec 2022, pp. 220–231.
- [2] H. Hao, A. Zhou, H. Qian, and H. Zhang, “Expensive multiobjective optimization by relation learning and prediction,” *IEEE Transactions on Evolutionary Computation*, vol. 26, no. 5, pp. 1157–1170, 2022.
- [3] Y. Jin, “Surrogate-assisted evolutionary computation: Recent advances and future challenges,” *Swarm and Evolutionary Computation*, vol. 1, no. 2, pp. 61–70, 2011.
- [4] P. Yang, L. Zhang, H. Liu, and G. Li, “Reducing idleness in financial cloud services via multi-objective evolutionary reinforcement learning based load balancer,” *Science China Information Sciences*, vol. 67, no. 2, pp. 1–21, 2024.
- [5] C. W. Ahn, R. S. Ramakrishna, and D. E. Goldberg, “Real-coded bayesian optimization algorithm: Bringing the strength of BOA into the continuous world,” in *Genetic and Evolutionary Computation - GECCO 2004, Genetic and Evolutionary Computation Conference, Seattle, WA, USA, June 26-30, 2004, Proceedings, Part I*, ser. Lecture Notes in Computer Science, vol. 3102. Springer, 2004, pp. 840–851.
- [6] H. Hao, X. Zhang, and A. Zhou, “Enhancing saeas with unevaluated solutions: a case study of relation model for expensive optimization,” *Science China Information Sciences*, vol. 67, no. 2, pp. 1–18, 2024.
- [7] H. Hao, J. Zhang, and A. Zhou, “A comparison study of surrogate model based preselection in evolutionary optimization,” in *Intelligent Computing Theories and Application: 14th International Conference, ICIC 2018, Wuhan, China, August 15-18, 2018, Proceedings, Part II 14*. Springer, 2018, pp. 717–728.
- [8] H. Hao, J. Zhang, X. Lu, and A. Zhou, “Binary relation learning and classifying for preselection in evolutionary algorithms,” *IEEE Transactions on Evolutionary Computation*, vol. 24, no. 6, pp. 1125–1139, 2020.
- [9] X. Wang, Y. Jin, S. Schmitt, and M. Olhofer, “Recent advances in bayesian optimization,” *ACM Comput. Surv.*, vol. 55, no. 13s, pp. 287:1–287:36, 2023.
- [10] D. Golovin, B. Solnik, S. Moitra, G. Kochanski, J. Karro, and D. Sculley, “Google vizier: A service for black-box optimization,” in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, August 13 - 17, 2017*. ACM, 2017, pp. 1487–1495.
- [11] J. T. Wilson, F. Hutter, and M. P. Deisenroth, “Maximizing acquisition functions for bayesian optimization,” in *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, December 3-8, 2018, Montréal, Canada*, 2018, pp. 9906–9917.
- [12] C. E. Rasmussen, “Gaussian processes in machine learning,” in *Advanced Lectures on Machine Learning, ML Summer Schools 2003, Canberra, Australia, February 2-14, 2003, Tübingen, Germany, August 4-16, 2003, Revised Lectures*, ser. Lecture Notes in Computer Science, vol. 3176. Springer, 2003, pp. 63–71.
- [13] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [14] J. Mockus, “On bayesian methods for seeking the extremum,” in *Optimization Techniques, IFIP Technical Conference, Novosibirsk, USSR, July 1-7, 1974*, ser. Lecture Notes in Computer Science, G. I. Marchuk, Ed., vol. 27. Springer, 1974, pp. 400–404.
- [15] A. Zhou, J. Sun, and Q. Zhang, “An estimation of distribution algorithm with cheap and expensive local search methods,” *IEEE Trans. Evol. Comput.*, vol. 19, no. 6, pp. 807–822, 2015.
- [16] F. Nogueira, “Bayesian Optimization: Open source constrained global optimization tool for Python,” 2014–. [Online]. Available: <https://github.com/fmfn/BayesianOptimization>
- [17] A. Cowen-Rivers, W. Lyu, R. Tutunov, Z. Wang, A. Grosnit, R.-R. Griffiths, A. Maravel, J. Hao, J. Wang, J. Peters, and H. Bou Ammar, “Hebo: Pushing the limits of sample-efficient hyperparameter optimisation,” *Journal of Artificial Intelligence Research*, vol. 74, 07 2022.
- [18] T. scikit-optimize contributors, “scikit-optimize: Sequential model-based optimization in python,” 2020. [Online]. Available: <https://github.com/scikit-optimize/scikit-optimize>
- [19] N. Stander and K. Craig, “On the robustness of a simple domain reduction scheme for simulation-based optimization,” *Engineering Computations*, vol. 19, no. 4, pp. 431–450, 2002.
- [20] E. Brochu, M. W. Hoffman, and N. de Freitas, “Portfolio allocation for bayesian optimization,” 2011.
- [21] F. Li, X. Cai, L. Gao, and W. Shen, “A surrogate-assisted multiswarm optimization algorithm for high-dimensional computationally expensive problems,” *IEEE Trans. Cybern.*, vol. 51, no. 3, pp. 1390–1402, 2021.
- [22] A. Zhou, J. Zhang, J. Sun, and G. Zhang, “Fuzzy-classification assisted solution preselection in evolutionary optimization,” in *The Thirty-Third AAAI Conference on Artificial Intelligence*, 2019, pp. 2403–2410.
- [23] T. Chen and C. Guestrin, “Xgboost: A scalable tree boosting system,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13-17, 2016*, B. Krishnapuram, M. Shah, A. J. Smola, C. C. Aggarwal, D. Shen, and R. Rastogi, Eds. ACM, 2016, pp. 785–794.
- [24] B. Liu, Q. Zhang, and G. G. E. Gielen, “A gaussian process surrogate model assisted evolutionary algorithm for medium scale expensive optimization problems,” *IEEE Trans. Evol. Comput.*, vol. 18, no. 2, pp. 180–192, 2014.
- [25] X. Yao, Y. Liu, and G. Lin, “Evolutionary programming made faster,” *IEEE Transactions on Evolutionary Computation*, vol. 3, no. 2, pp. 82–102, 1999.
- [26] Y. Tian, R. Cheng, X. Zhang, and Y. Jin, “PlatEMO: A MATLAB Platform for Evolutionary Multi-Objective Optimization [Educational Forum],” *IEEE Computational Intelligence Magazine*, vol. 12, no. 4, pp. 73–87, 2017.
- [27] M. Hollander, D. A. Wolfe, and E. Chicken, *Nonparametric Statistical Methods*. John Wiley & Sons, 2013, vol. 751.