A Low-Sample-Count, High-Precision Pareto Front Adaptive Sampling Algorithm Based on Multi-Criteria and Voronoi

Changkun Wu (1094090800@qq.com)
Guangxi University
Ke Liang
Guangxi University
Hailang Sang
Guangxi Yuchai Machine Co.
Yu Ye
Guangxi Yuchai Machine Co.
Mingzhang Pan
Guangxi University

Research Article

Keywords: Pareto front, Voronoi, Maximum crowding criterion, Maximum LOO error criterion, Maximum mean MSE criterion

Posted Date: August 12th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-442735/v1

A low-sample-count, high-precision Pareto front adaptive sampling algorithm based on multi-criteria and Voronoi

Changkun Wu[a], Ke Liang[b], Hailang Sang[c],Yu Ye[c], Mingzhang Pan[*a].

[a] Guangxi Key Laboratory of Manufacturing System & Advanced Manufacturing Technology，School of Mechanical Engineering, Guangxi University, Nanning 530004, China

[b] School of Mechanical Engineering, Guangxi University, Nanning 530004, China

[c] Guangxi Yuchai Machine Co. Nanning 530004, China

* Corresponding author. Tel./Fax: +86 18587857890.

E-mail address: pmz@gxu.edu.cn (Mingzhang Pan)

Abstract

In this paper, a Pareto front (PF)-based sampling algorithm, PF-Voronoi sampling method, is proposed to solve the problem of computationally intensive multi-objectives of medium size. The Voronoi diagram is introduced to classify the region containing PF prediction points into Pareto front cells (PFCs). Valid PFCs are screened according to the maximum crowding criterion (MCC), maximum LOO error criterion (MLEC), and maximum mean MSE criterion (MMMSEC). Sampling points are selected among the valid PFCs based on the Euclidean distance. The PF-Voronoi sampling method is applied to the coupled Kringing and NASG-II models and its

validity is verified on the ZDT mathematical cases. The results show that the MCC criterion helps to improve the distribution diversity of PF. The MLEC criterion and the MMMSEC criterion reduce the number of training samples by 38.9% and 21.7%, respectively. The computational cost of the algorithm is reduced by more than 44.2%, compared to EHVIMOPSO and SAO-MOEA algorithms. The algorithm can be applied to multidisciplinary, multi-objective, and computationally intensive complex systems.

Keywords: Pareto front; Voronoi; Maximum crowding criterion; Maximum LOO error criterion; Maximum mean MSE criterion

Introduction

With the continuous development of simulation technology as well as computational power [1], one can build high-fidelity models to predict complex physicochemical phenomena in the real world, such as the interaction mechanism of fluids with glider geometry and in-cylinder combustion phenomena in internal combustion engines. At the same time, accurate performance evaluation has led to the development of simulation-based multi-objective optimization problems for practical engineering [2]. However, the huge cost of high-fidelity analysis has become a bottleneck for simulation-based optimization, especially for complex systems involving multiple disciplines, multiple objectives, and computationally intensive [3]. It has been reported that it takes 36-160 hours to run a crash simulation for Ford Motor Company [4]. And for ship flow systems, one CFD calculation takes several days [5]. Frequent optimization-seeking iterations and high simulation costs are difficult to be accept in practice.

The application of surrogate models offers the possibility to reduce the cost of optimization [3]. Surrogate models constructed from several training samples are used to approximate computationally expensive functions. And optimization methods search for optimal values on the surrogate models. Clearly, the use of these surrogate model-based optimization design approaches can significantly reduce the number of high-fidelity simulations, reduce time costs, and produce better designs [6].Ibaraki et al [7] used an artificial neural network (ANN) and a genetic algorithm (GA) to form an optimization system for multi-objective optimization of a centrifugal compressor impeller. It was shown that the two optimized impellers located at PF showed significant improvement in efficiency and operating range, respectively, compared to the original impeller. Ekradi et al [8] used the ANN-GA optimization system to study the effect of impeller blade angle on the performance of the centrifugal compressor. The study showed that the optimized impeller improved the isentropic efficiency at the design point by 0.97%, the total pressure ratio by 0.94%, and the mass flow rate by 0.65%. Wang et al [9] applied the Kriging model to the optimal design of an industrial centrifugal impeller, resulting in a 2.49% improvement in the isentropic efficiency of the impeller. Guo et al [10] applied the response surface methodology (RSM) to optimal design of a small centrifugal compressor to optimize the design of a small centrifugal compressor. The results of the study showed that the optimized compressor pressure ratio was improved by 7.5%.

In the literature above, most of the training sample points required for surrogate model construction are obtained by "one-time" sampling methods [11] such as

uniform design [9], Latin hypercube sampling [12], and central composite design [10]. These methods can quickly generate all sample points by a predefined total number of samples. However, in practice, it is often difficult for engineers to determine the appropriate sample size [13], which leads to unnecessary computational costs [14][15]. To achieve good prediction accuracy of surrogate models with a reasonable number of sampling points, several adaptive sampling (also called sequential sampling) methods have been developed in recent years [16]. This class of methods starts with small samples and improves the accuracy of the surrogate model by iteratively selecting samples throughout the design space. Sasena et al [17] compared five sampling criteria in several mathematical cases: expected improvement criterion, improved probability criterion, regional extreme value locus criterion, mean squared error criterion and minimize surprises criterion. It was shown that the criterion with more emphasis on global search required more iterations to locate the optimal values and was less accurate than the criterion with emphasis on local search. Xu et al [16] proposed a sampling criterion, CV-Voronoi, for global surrogate models. This criterion divides the design space into multiple cells based on the current sample points and determines the cell where the next sample point is located by cross-validation, thus reducing the search cost. Guo et al [18] used the EI sampling criterion to improve the global accuracy of the Kriging model and proposed the MBGO global optimization algorithm to optimize a typical axial transonic rotor blade NACA rotor. Li et al [19] proposed a maximum combined expectation (CE) sampling model and compared it with the above sampling criterion. It was shown that the

Pareto optimal frontier could be accurately predicted with a minimum number of sample points using the CE sampling model.

From the above literature, it can be found that most researchers currently design the sampling criterion to find the next sampling point in the whole design space to improve the global accuracy of the surrogate model. Obtaining a more accurate surrogate model with fewer sample points is one of the research directions of most researchers. However, for an optimization problem, engineers are more concerned with the accuracy of the predicted PF. It is conceivable that there exist sample points that are partially far from the region where the PF is located. The existence of such sampling points improves the local prediction accuracy of the surrogate model but does not help much to improve the prediction accuracy of the predicted PF. Based on this idea, it is obvious that avoiding the selection of such sampling points can further reduce the number of sampling points. At present, some scholars have proposed PF-based sampling methods [20][21]. Cheng et al [22] proposed a dynamic expectation super volume improvement sampling algorithm. Several promising prediction points are selected to be added to the training set by comparing the expected over-volume improvement values of the currently predicted PF points. Fan et al [23] used reference vectors and non-dominated ranks to screen PF prediction points to ensure the diversity and convergence of the selected individuals.Gao et al [24] used EIM criterion and distance criterion to screen PF prediction points and proposed an adaptation criterion to balance exploration and exploitation and reduce the sample size.

The aforementioned PF-based sampling methods are keen on adding some selected PF prediction points to the training sample, and different criteria are proposed to balance mining and exploration. However, it is difficult to determine the accuracy of the PF predicted by the surrogate model [25], which increases the number of invalid samples and the calculation cost. Therefore, this paper proposes a new PF-based sampling algorithm, the PF-Voronoi sampling method. Instead of screening the next sampling point from the PF prediction point, the algorithm performs the next sampling point selection in the region where the PF prediction point belongs in the design space. In other words, the PF prediction points function only as a tool to indicate the region where the true Pareto front may exist. The PF-Voronoi sampling method makes reference to the division of the design space by the Voronoi diagram [26] to assign the PF prediction points to different regions. And the appropriate regions are selected for sampling by the three sampling criteria proposed in this paper to make full use of the information of PF prediction points. In this paper, the PF-Voronoi sampling method adaptively updates the Kriging model and couples with NASG-II to construct an effective multi-objective optimization framework - PFVNASG-II.

2. Pareto front (PF)-Voronoi NASG-II algorithm

In this paper, we propose a PF-based optimization algorithm PFVNASG-II. PFVNASG-II constructs a Kriging model for every expensive objective function [27] and uses the NASG-II optimization algorithm to find the optimization for the multi-objective optimization (MOO) problem. The original MOO problem is approximated as

$$\begin{aligned} &\min:\left\{\hat{f}_1(\mathrm{x}),\ \hat{f}_2(\mathrm{x}),\cdots,\ \hat{f}_v(\mathrm{x})\right\} \\ &s.t:g_i(\mathrm{x})\le 0,\ \ i=1,\cdots,t_g \\ &\mathrm{x}=\left\{x_1,x_2,\cdots,x_n\right\}\in[\mathrm{a},\mathrm{b}] \end{aligned} \tag{1}$$

Where $\hat{f}_1(\mathrm{x}),\cdots\hat{f}_v(\mathrm{x})$ are surrogate models of objective functions. n is the number of design variable dimensions. t_g is the number of constraints.

Fig.1 illustrates the flow chart of PFVNASG-II. There exist a global adaptive sampling strategy GEI sampling algorithm in tandem with the PF-Voronoi sampling algorithm. The details of these two sampling algorithms are presented in the next subsection.

In the flow chart, the complete optimization procedure starts with small training samples. The initial training samples are derived from the optimal Latin hypercube (OLH) [28] experimental design. t is the number of iterations. R^2_{PF} is defined as the coefficient of determination of the PF prediction point. When R^2_{PF} = 1, it means that the PF predicted by the surrogate model has zero error to the true function value. It should be noted that R^2_{PF} only represents the prediction accuracy of the local design space, so it is normal that the global R^2 of the surrogate model is relatively low. This is consistent with the understanding of practical engineering optimization design in this paper: the surrogate model only needs to ensure satisfactory prediction accuracy in the local area where the predicted PF is located. R^2_{PF} is defined as

$$R^2_{PF}=1-\sum_{i=1}^{m}(f_i^{PF}-\hat{f}_i^{PF})^2\Big/\sum_{i=1}^{m}(f_i^{PF}-\overline{f}^{PF})^2,\ \overline{f}^{PF}=\frac{1}{m}\sum_{i=1}^{m}f_i^{PF} \tag{2}$$

Where, m is the population size of NASG-II; f_i^{PF} is the observed value of the i-th PF prediction point; $\overline{f}^{PF}$ is the average of the observed values f_i^{PF}。 $\hat{f}_i^{PF}$ is the

predicted value of the i-th PF prediction point。

To judge the convergence and distribution of the multi-objective optimizer proposed in this paper, Inverter generational distance (IGD) and Maximum Spread (MS) are also chosen as performance metrics in this paper.

$$IGD = \frac{\sum_{u \in PF_{true}} d\left(u, PF_{know}\right)}{\left|PF_{true}\right|} \tag{3}$$

$$MS = \left[\frac{1}{v} \sum_{i=1}^{v} \left[\frac{\min(f_i^{\max}, F_i^{\max}) - \max(f_i^{\min}, F_i^{\min})}{F_i^{\max} - F_i^{\min}} \right]^2 \right]^{1/2} \tag{4}$$

Where， PF_{true} is true Pareto front, PF_{know} is the non-dominated solutions found， and $d\left(u, PF_{know}\right)$ is the Euclidean distance between u and its nearest member in PF_{know}， $f_i^{\max}$ and $f_i^{\min}$ are the maximum and minimum of the i-th objective in PF_{know}， $F_i^{\max}$ and $F_i^{\min}$ are the maximum and minimum of the i-th objective in PF_{true}.

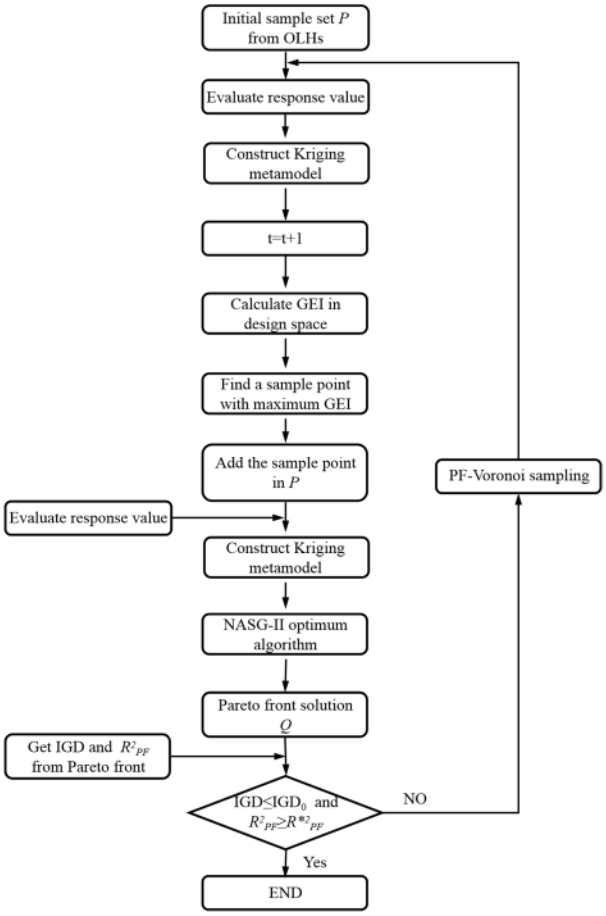


Fig.1 Flow chart of PFVNASG-II.

2.1 Kriging model

Kriging is an efficient way of interpolation and can give probabilistic responses. Its variance depends on the number of available training samples N [29]. For single-objective problems. The Kriging model can be expressed as:

$$\hat{f}(\mathrm{x}) = h^T(\mathrm{x})\hat{\beta} + \mathrm{r}^T(\mathrm{x},\theta)\mathrm{R}^{-1}(\theta)\left(Y - H\hat{\beta}\right) \tag{5}$$

Where $h^T(\mathrm{x}) = \left\{h_1(\mathrm{x}),...,h_{n+1}(\mathrm{x})\right\}$ is the regression function represented by the polynomial function; $Y = \left\{y_1,...,y_N\right\}$ is the observed value of N training samples.; $\hat{\beta}$ are the regression coefficients:

$$\hat{\beta} = \left(H^T\mathrm{R}^{-1}(\theta)H\right)^{-1}H^T\mathrm{R}^{-1}(\theta)Y \tag{6}$$

$$H = \left\{h(\mathrm{x}_1),...,h(\mathrm{x}_N)\right\}^T \tag{7}$$

$\mathrm{R}(\theta) = \left\{R\left(\left|\mathrm{x}_i - \mathrm{x}_j\right|,\theta\right)\right\}_{1\le i,j\le N}$ is the $N\times N$ correlation matrix with the element $R\left(\left|\mathrm{x}_i - \mathrm{x}_j\right|,\theta\right)$. $\mathrm{r}(\mathrm{x},\theta) = \left\{R\left(\left|\mathrm{x} - \mathrm{x}_i\right|,\theta\right),...,R\left(\left|\mathrm{x} - \mathrm{x}_N\right|,\theta\right)\right\}$ is the correlation vector between x and the sample points $\left\{\mathrm{x}_1,...,\mathrm{x}_N\right\}$. $R\left(\left|\mathrm{x}_i - \mathrm{x}_j\right|,\theta\right)$ is spline function model:

$$R\left(\left|\mathrm{x}_i - \mathrm{x}_j\right|,\theta\right) = \prod_{k=1}^{n} R\left(\left|\mathrm{x}_{i,k} - \mathrm{x}_{j,k}\right|,\theta_k\right) \tag{8}$$

$$R\left(\left|\mathrm{x}_{i,k} - \mathrm{x}_{j,k}\right|,\theta_k\right) = \begin{cases} 1 - 1.5d_k^2 + 30d_k^3, d_k \in [0,0.2] \\ 1.25(1-d_k)^3, d_k \in (0.2,1) \\ 0, d_k \in [1,\infty] \end{cases} \tag{9}$$

$$d_k = \theta_k\left|\mathrm{x}_{i,k} - \mathrm{x}_{j,k}\right| \tag{10}$$

Where $\theta = \left\{\theta_1,...,\theta_n\right\}$ is hyperparameters; n is the dimension of the design variable.

2.2 GEI sampling algorithm

It is necessary to concatenate an adaptive global sampling algorithm before implementing a PF-based sampling algorithm. The specific reasons are described in detail in Chapter 3. Since the Kriging model is a surrogate model based on a stochastic process, it can predict both the function value as well as its uncertainty at unobserved points. Based on this property, Jones et al [30] derived the expectation improvement (EI) sampling algorithm to improve the global prediction accuracy of the surrogate model. The EI is

$$EI(\mathrm{x}) = \left(f_{\min} - \hat{f}(\mathrm{x})\right)\Phi\left(\frac{f_{\min} - \hat{f}(\mathrm{x})}{s}\right) + s\varphi\left(\frac{f_{\min} - \hat{f}(\mathrm{x})}{s}\right) \tag{11}$$

Where $\Phi(\cdot)$ and $\varphi(\cdot)$ are the standard normal density and distribution function; $\hat{f}(\mathrm{x})$ is predicted function value by Kriging model; and s is the standard deviation predicted at point x [27].

To apply the EI algorithm to multi-objective optimization problems, this paper uses the generalized expectation improvement (GEI) proposed by Jie et al [31].

$$EI_i^{*}(\mathrm{x}) = \frac{EI_i(\mathrm{x}) - \min(EI_i)}{\max(EI_i) - \min(EI_i)} \tag{12}$$

$$GEI = \sum_{i=1}^{v} EI_i^{*}(\mathrm{x}) \tag{13}$$

Where i is the indicator of the objective function; v is the number of objective functions; The $\max(EI_i)$ and $\min(EI_i)$ are maximal and minimal expected improvement.

2.3 PF-Voronoi sampling algorithm

Fig.2 shows the flow chart of the PF-Voronoi sampling algorithm. As can be seen from the figure, the PF-Voronoi sampling algorithm is divided into two branches. The branch I is connected in series with the main program (Fig.1), while branch II loops independently. P^t is the training sample set in the t-th loop. P_{PF}^t is the non-dominated sample set of P^t, and the number is N_{PF}^t. N_{PF}^* and α are judgment constants. The author assumes that when $N_{PF}^t > N_{PF}^* \ and \ P_{PF}^{t-\alpha} \in P_{PF}^t$, the samples in P_{PF}^t are true PF samples in the design space. When this condition is satisfied, the algorithm enters branch II for further convergence. In other words, the role of the branch I and the main loop is to find enough PF sample points.

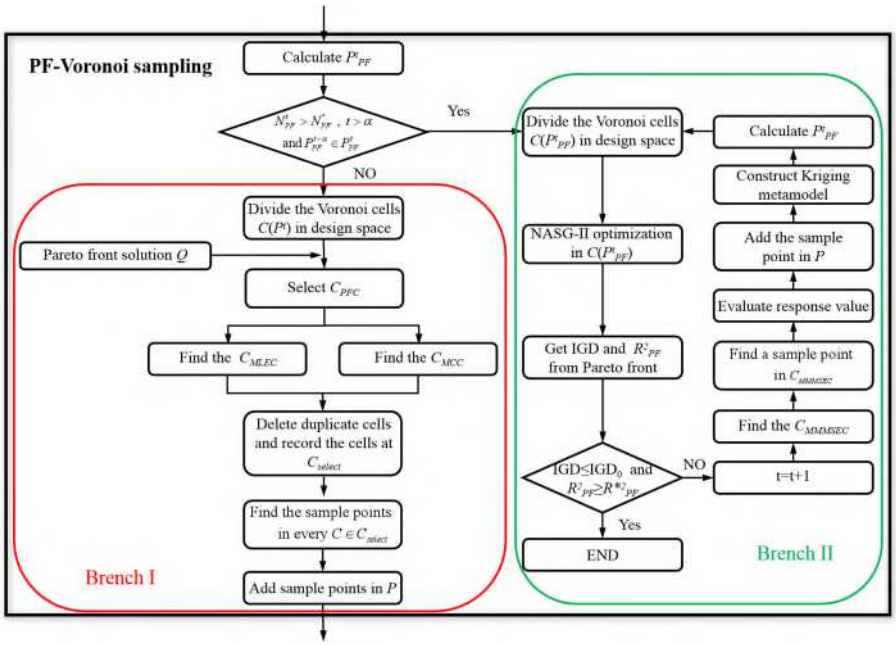


Fig.2 Flow chart of PF-Voronoi sampling algorithm.

The PF-Voronoi sampling algorithm refers to the Voronoi diagram [26] for the partitioning of the design space. The entire design space is divided into a set of Voronoi cells $C(P) = \{C(p_1), C(p_2), ..., C(p_N)\}$ according to the current training sample set $P = \{p_1, p_2, ..., p_N\} \in \mathrm{R}^d$. The i-th Voronoi cell $C(p_i)$ indicates the area

around the sample P_i. The cell $C(p_i)$ can be defined as [26]

$$\mathrm{dom}(p_i, p_j) = \left\{ p \in \mathrm{R}^d \mid \|p - p_i\|_2 \le \|p - p_j\|_2 \right\} \tag{14}$$

$$C(p_i) = \mathrm{I}_{\ p_j \in P \setminus p_i} \mathrm{dom}(p_i, p_j) \tag{15}$$

Where $\mathrm{dom}(p_i, p_j)$ is the vertical bisector between the sample points p_i and p_j. The bisector line separates all points in the plane close to p_i from those close to p_j. Fig.3 shows a set of 2D samples and the corresponding Voronoi cells. The red dots represent the predicted Pareto front $Q = \{q_1, q_2, \cdots, q_m\}$. Clearly, Q is divided into several Voronoi cells. The Voronoi cells containing PF prediction points are referred to here as Pareto frontier cells (PFC). The PF-Voronoi samples in these PFCs.

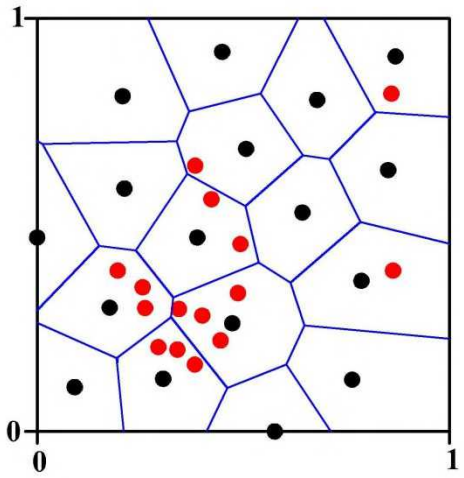


Fig.3 A set of 2D samples and the corresponding Voronoi cells. Black points: Current training sample points; red points: prediction points of Pareto front.

Before performing sampling, the PFCs need to be screened. This is because, after several iterations, the number of samples (NOS) increases significantly, the Voronoi cells increase, and the PFC increases accordingly. Extracting a sample point for every PFCs IS costly and unnecessary. The selection of PFCs is performed by the following three criteria, the first two criterions applied to the branch I and the third one to

branch II.

1、Maximum crowding criterion (MCC): According to equation (1), given the non-dominated solution set $F_{PF}=\left\{f\left(p_{PF}^{1}\right),f\left(p_{PF}^{2}\right),...,f\left(p_{PF}^{k}\right)\right\}$ of the training samples and the corresponding sample set $P_{PF}=\left\{p_{PF}^{1},p_{PF}^{2},...,p_{PF}^{k}\right\}$. Select the PFC C_{MCC} corresponding to the sample with the maximum crowding [32] (removal of boundary points). The main purpose of this criterion is to explore the uncovered real PF regions. The author believes that it is reasonable to assume that it is continuous in the target space when the true PF morphology is not known. The author believes that there are likely to be mispredicted true PFs in the large gaps between the current F_{PF}. It helps to improve the diversity of the distribution of PF prediction points. the crowding degree CD^k of p_{PF}^{k} can be expressed as

$$CD^{k}=\begin{cases}\sum_{j=1}^{v}\left|f_j\left(p_{PF}^{k+1}\right)-f_j\left(p_{PF}^{k-1}\right)\right| & \exists q_i\in C\left(p_{PF}^{k}\right)\wedge f_j\left(p_{PF}^{k}\right)\neq f_j\left(p_{PF}^{b}\right)\\ 0 & else\end{cases} \tag{16}$$

For the j-th objective function, the crowding distance of p_{PF}^{k} is the average distance between its two neighboring solutions $f_j\left(p_{PF}^{k+1}\right)$ and $f_j\left(p_{PF}^{k-1}\right)$. When the Voronoi cell corresponding to p_{PF}^{k} does not contain the PF prediction point q or $f_j\left(p_{PF}^{k}\right)$ is equal to the boundary point $f_j\left(p_{PF}^{b}\right)$, the CD^k is zero.

2、Maximum LOO error criterion(MLEC)：The frontier cells C_{MLEC} with the largest LOO error e_{LOO} [16] is selected. This is in reference to the CV-Voronoi global sampling algorithm proposed by Xu et al. Xu et al. argue that the $e^{\mathrm{i}}{}_{\mathrm{LOO}}$ of sample point i is representative of the prediction accuracy of the region near that sample point. Therefore, improving the accuracy of C_{MLEC} helps to improve the accurate localization of the real PF by the surrogate model. $e^{\mathrm{i}}{}_{\mathrm{LOO}}$ can be defined as:

$$e_{\mathrm{LOO}}^{i}=\left|f\left(p_i\right)-\hat{f}_{P\setminus p_i}\left(p_i\right)\right| \tag{17}$$

Where $f\left(p_i\right)$ is the observed value of p_i at the sample point; $\hat{f}_{P\setminus p_i}\left(p_i\right)$ is the predicted value of p_i using the "leave-one-out" method.

3、Maximum mean MSE criterion(MMMSEC)：Select the frontier cells C_{MMMSEC} with maximum mean MSE $\overline{s}_{\max}$. When the PF-Voronoi sampling algorithm is run to branch II, the optimization search range of NASG-II is restricted to $C(P^{t}_{PF})$. At this time, the PF prediction points are all distributed near the true PF, and using the MSE of the prediction points as an indicator to judge its accuracy can speed up the convergence of PFVNASG-II. $\overline{s}_{\max}$ can be defined as:

$$\overline{s}_i=\sum\nolimits_{j=1}^{l_i}s_{j,i}\Big/l_i \tag{18}$$

$$\overline{s}_{\max}=\max\left(\overline{s}_1,\overline{s}_2,\cdots,\overline{s}_k\right) \tag{19}$$

$s_{j,i}$ is the MSE of the j-th prediction point in the i-th PFC; l_i is the number of predicted points in the i-th PFC.

The selection range of sampling points is limited to the screened PFCs and follows the following guidelines:

- Sampling points should be selected to maximize the accuracy of the PF prediction points.
- To reduce the clustering phenomenon of training samples, the sampling points are avidly selected in the region close to the boundary of Voronoi cells.

Fig.4 shows how the sampling points were selected. The area c_B surrounded by the green dashed circle is the conserved Voronoi cell of sampling point B. The radius of the green dashed circle is

$$r_B = d\left(p_A, p_B\right)/2 \tag{20}$$

Since the sampling point B is subordinate to the Voronoi cell of point A. Therefore, the Euclidean distance $d\left(p_A, p_B\right)$ between A and B is the shortest. It is easy to know that $c_B \in C_B$ according to the division rule of Voronoi cells.

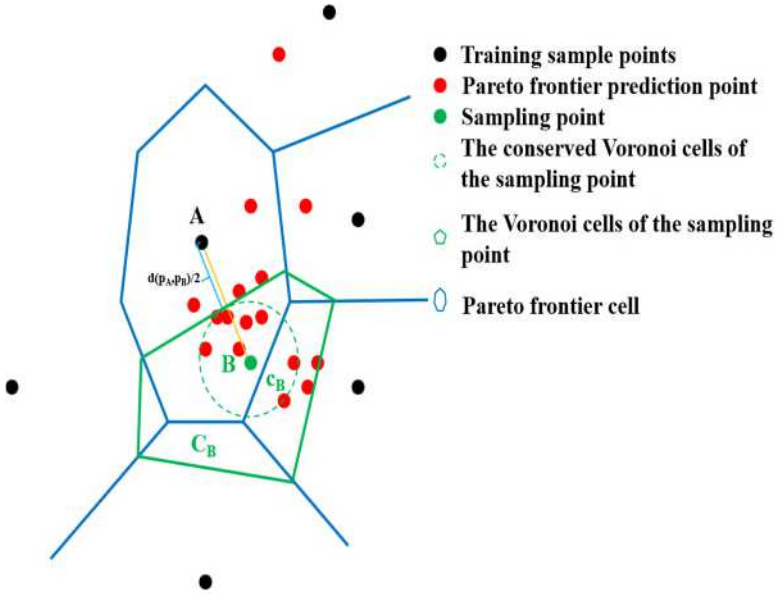


Fig.4 Schematic diagram of sampling point selection method.

Make c_B include more prediction points to improve the accuracy of more prediction points. When there exist multiple sampling points that can contain the same number of prediction points, we prefer to select the sampling point with the largest Euclidean distance from sample point A to reduce the clustering phenomenon of the training samples. The algorithm is described in detail as follows:

1）Set archive P_{PF}^t to store t-th generation of PF sample points.

2）When $t > \alpha$, $N_{PF}^t > N_{PF}^*$ and $P_{PF}^{t-\alpha} \in P_{PF}^t$, go to branch II, otherwise go to branch I。

Branche I:

1）Set the archive Q to store the currently predicted PF points with size m; set the

archive P_{select} to store the selected training samples. Set the archive *l* consisting of zeros and of size *N*.

2）For every individual in *Q*, calculate the minimum Euclidean distance of the i-th individual from the sample in *P*: $d_{\min}^{i} = d\left(q_i, P\right) = d\left(q_i, p_j\right)$, $l_j = l_j + \mathbf{1}$.

3）For every sample in *P*, pick $l_j > \mathbf{0}$ of the samples p_j to deposit in the archive P_{select}.

4）Calculate the samples in P_{select} according to equation (17). Rearrange the samples within P_{select} according to e_{LOO} from largest to smallest. Retain the top sample in P_{select} and remove the remaining samples.

5）Calculate the crowding degree *CD* of sample points in P_{PF}^{t} according to equation (1) (16). Rearrange the samples within P_{PF}^{t} according to *CD* from largest to smallest. Select the top sample within P_{PF}^{t} and deposit it into P_{select}.

6）Remove duplicate samples in P_{select}. For the i-th sample p_{select}^{i} in P_{select} do the following:

a. Initialize archive A.

b. According to equation (14) (15), every individual in *Q* is traversed, when $q_j \in C\left(p_{select}^{i}\right)$, q_j is stored in A.

c. Initialize the iteration counter: t_s=1

d. Cross mutate the individuals in A and deposit subpopulations in A and remove duplicate individuals. For the j-th individual q_A^j in A, when $q_A^j \notin C\left(p_{select}^{i}\right)$, remove q_A^j.

e. Set the archive *l* consisting of zeros of size equal to the number of individuals

in A.

f. For the j-th individual q_A^j in A, calculate its Euclidean distance from the individuals in Q. Go through the individuals in Q, when $d\left(q_A^j, q_k\right) \le d\left(q_A^j, p_{select}^i\right)/\mathbf{2}$, $l_j = l_j + \mathbf{1}$.

g. The individuals in A are sorted according to *l* from largest to smallest. The top *m* individuals were retained and the rest were deleted. If there are less than *m* individuals in A, all individuals are retained. $t_s = t_s+1$.

h. When $t_s<t_s^{max}$, repeat steps d-g。

i. When $l_j=l_{\max}$ and $d\left(p_A^k, p_{select}^i\right)=d\left(A, p_{select}^i\right)$, select p_A^k to call the real function evaluation and add it to the training set P.

7）Return to the main loop.

Branche II:

1）According to equation (14) (15) with the sample in P_{PF}^t, divide the Voronoi cells $C\left(P_{PF}^t\right)$ in design space.

2）Perform NASG-II search in $C\left(P_{PF}^t\right)$. Overwrite the current PF prediction points into *Q*.

3）Calculate IGD and R_{PF}^2 according to equation (2)(3). if IGD ≤ IGD_0 and $R_{PF}^2 \ge R_{PF}^{*2}$ then terminate the program, otherwise execute 4).

4）$t=t+1$, empty P_{select}。

5）Perform branch I steps 1)-3). Determine according to equations (18)(19) and deposit the corresponding sample points into P_{select}.

6）Execute branch I steps a-i.

7）Retrain the Kriging model and update the P_{PF}^{t}.

8）Return to step 1).

3. Simulation experiments

In this section, the performance of PFVNASG-II was verified using the high-dimensional nonlinear mathematical cases ZDT1-ZDT4, ZDT6. The population size of NASG-II is 100, the crossover operator was 0.8, and the variance operator was 0.2. All the codes mentioned above were programmed using MATLAB 2016a. The mathematical cases were shown in Table 1. Other parameter sets were listed in Table 2.

Table 1. The test functions.

Functions	n	Decision space	Objectives
ZDT1	30	$x_i \in [\mathbf{0},\mathbf{1}]$	$\min f_1(\mathrm{x}) = x_1$ $\min f_2(\mathrm{x}) = g(\mathbf{1} - \sqrt{(f_1 / g)})$ $g(\mathrm{x}) = \mathbf{1} + \mathbf{9}\sum_{i=2}^{n} x_i / (n-\mathbf{1})$
ZDT2	30	$x_i \in [\mathbf{0},\mathbf{1}]$	$\min f_1(\mathrm{x}) = x_1$ $\min f_2(\mathrm{x}) = g(\mathbf{1} - (f_1 / g)^{\mathbf{2}})$ $g(\mathrm{x}) = \mathbf{1} + \mathbf{9}\sum_{i=2}^{n} x_i / (n-\mathbf{1})$
ZDT3	30	$x_i \in [\mathbf{0},\mathbf{1}]$	$\min f_1(\mathrm{x}) = x_1$ $\min f_2(\mathrm{x}) = g(\mathbf{1} - \sqrt{(f_1 / g)} - (f_1 / g)\sin(\mathbf{10}\pi f_1))$ $g(\mathrm{x}) = \mathbf{1} + \mathbf{9}\sum_{i=2}^{n} x_i / (n-\mathbf{1})$
ZDT4	10	$x_i \in [\mathbf{0},\mathbf{1}]$	$\min f_1(\mathrm{x}) = x_1$ $\min f_2(\mathrm{x}) = g(\mathbf{1} - \sqrt{(f_1 / g)})$ $g(\mathrm{x}) = \mathbf{1} + \mathbf{10}(n-\mathbf{1}) + \sum_{i=2}^{n}(x_i^{\mathbf{2}} - \mathbf{10}\cos(\mathbf{4}\pi x_i))$

ZDT6	10	$x_i \in [\mathbf{0},\mathbf{1}]$	$\min f_1(\mathrm{x}) = \mathbf{1} - \exp(-\mathbf{4}x_1)\sin^6(\mathbf{6}\pi x_1)$ $\min f_2(\mathrm{x}) = g(\mathbf{1} - (f_1 / g)^2)$ $g(\mathrm{x}) = \mathbf{1} + \mathbf{9}\sum_{i=2}^{n} x_i / (n-\mathbf{1})^{\mathbf{0.25}}$

Table 2. Parameters setting of algorithms.

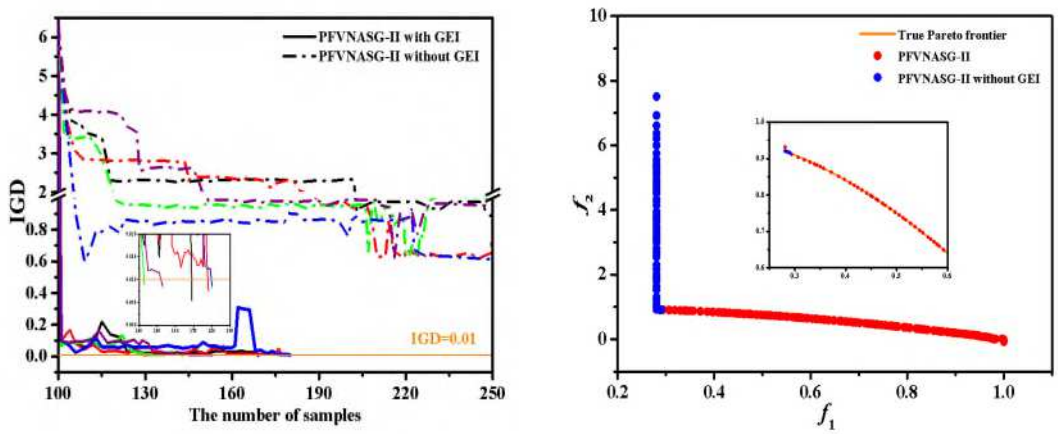
		ZDT1	ZDT2	ZDT3	ZDT4	ZDT6
Initial samples		50	50	50	50	100
$N_{\max}$		200	200	250	250	250
IGD_0		0.0063	0.0052	0.01	0.01	0.01
R^{*2}_{PF}		0.9	0.9	0.9	0.9	0.9
α		3	3	3	3	3
N^*_{PF}		20	15	20	20	20
NASG-II	Population size	100	100	100	100	100
	cross operator	0.8	0.8	0.8	0.8	0.8
	mutation operator	0.2	0.2	0.2	0.2	0.2

3.1 Impact of GEI on PFVNASG-II optimization procedure

First, it needs to be clear that the surrogate model trained with small samples was inaccurate, and difficult to locate the region where the true PF is located. Enriching the training sample set with a PF-based sampling algorithm alone tends to fall into local PFs [25]. Fig.5 showed the effect of tandem GEI on predicting the ZDT6 PF. Each algorithm was run 20 times independently. The resulting data were shown in Table 3.

From Fig. 5(a), it can be seen that the IGD of the optimization program without tandem GEI converged slowly, while the tandem GEI helped the optimization procedure converge quickly and reach the threshold value. In Fig. 5(b), the PF predicted by the optimization procedure without GEI was far from the true PF. The global exploration capability of the PFVNASG-II optimization without GEI is weak, so it is easy to get caught in the local PF. From Table 3, it can be found that the

PFVNASG-II optimization procedure with GEI successfully predicted PF in 20 independent runs, while the prediction success rate of the optimization procedure without GEI was only 25%. Therefore, it is necessary to cascade an adaptive global sampling algorithm to improve the global exploration capability of the surrogate model.



(a) The effect of GEI on IGD (b) The effect of GEI on PF

Fig.5 The PFVNASG-II optimization program predicts the ZDT6 Pareto frontier.

Table 3. Result data of PFVNASG-II with/without GEI.

ZDT6		PFVNASG-II with GEI	PFVNASG-II without GEI
Total run times		20	20
The number of IGD < 0.01		20	4
IGD	Mean.	0.0077	0.6179
	Std.	0.0013	0.4918
R^2_{PF}	Mean	(1,0.9581)	(1,-294.35)
	Std.	(0,0.0364)	(0,845.29)
MS	Mean	0.9992	0.5787
	Std.	0.0019	0.3607
NOS	Mean.	170.8	235.8
	Std.	30.1	29.0

3.2 PFC screening criterion validation

The ZDT3 mathematical case was selected to verify the effect of the MCC criterion and the MLEC criterion on the computational cost of the PFVNASG-II

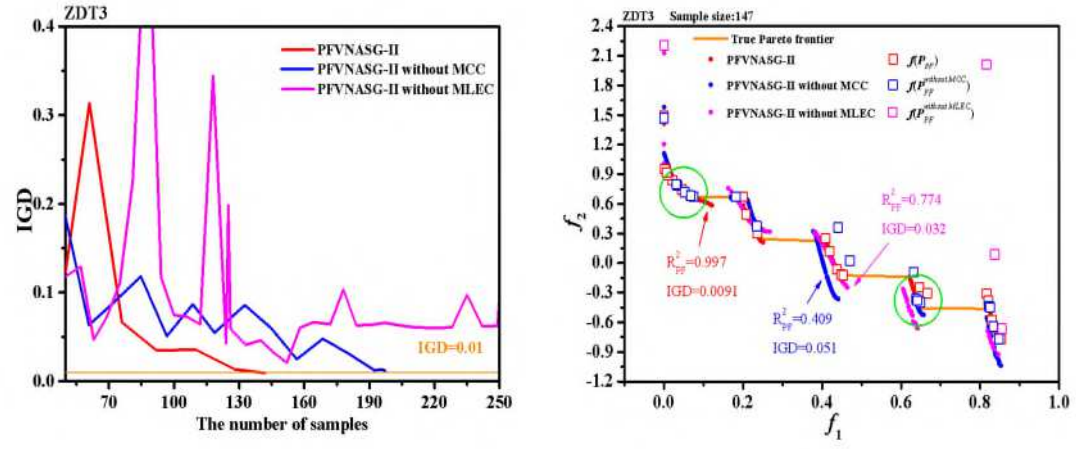
optimization procedure. ZDT3 was chosen because of the significant effect of MCC on the prediction accuracy of it. Each algorithm was run 20 times independently. The resulting data are shown in Table 4.

Fig.6 showed the predictive power of the three optimization programs for PF. Fig.6(a) showed the IGD convergence curves for the PFVNASG-II optimization procedure and the optimization procedures without MCC or MLEC, respectively. Obviously, the IGD of the PFVNASG-II optimization procedure converged to the threshold value faster. While the optimization procedure without the MCC criterion required a larger time cost to reach the convergence threshold. Besides, it can be found that the without MLEC criterion even makes it difficult to converge successfully on a given maximum number of samples (MNOS). Fig.6 (b) compares the predicted PF of the three optimization procedures at a sample size of 147. The IGD and R^2_{PF} of the PFVNASG-II optimization procedure reached the convergence threshold at this sample size. From the figure, it can be found that the R^2_{PF} of PF predicted by PFVNASG-II was (1,0.997). The lack of MCC criterion and MLEC criterion results in poorer prediction accuracy at the same NOS. Fig.6(b) also compared the distributions of the non-dominated solution sets in the target space of the respective sample sets for the three optimization procedures at the current NOS. the non-dominated solution sets of PFVNASG-II and PFVNASG-II without MCC fell on the true PF. But the non-dominated solution sets of PFVNASG-II without MCC were more aggregated, as shown in the green circles. This means that the MCC criterion helped to improve the distributional diversity of the optimization procedure.

The number of non-dominated solution sets of PFVNASG-II without MLEC was small and far from the real PF. obviously, the MLEC criterion facilitated the sampling algorithm to quickly approach the real PF and sped up the convergence rate.

As can be seen from Table 4, all 20 independent operations of PFVNASG-II converge successfully within a given MNOS, and the average NOS was 145.8. While PFVNASG-II without MCC had only an 80% success rate and PFVNASG-II without MLEC had only a 20% success rate. Clearly, the MLEC criterion was crucial in ensuring the successful convergence of the algorithm.

In summary, the MCC criterion and the MLEC criterion showed excellent performance in terms of improving distribution diversity, convergence stability, and reducing time cost.



(a)The effect of MCC and MLEC on IGD (b)The PF of three optimization programs

Fig.6 Predictive power of three optimization procedures for the ZDT3 Pareto frontier.

Table 4. Result data of three different optimization programs.

ZDT3	PFVNASG-II	PFVNASG-II without MCC	PFVNASG-II without MLEC
Total run times	20	20	20
The number of IGD <	20	16	5

0.01				
IGD	Mean.	0.0080	0.0113	0.0845
	Std.	0.0013	0.0060	0.0727
R^2_{PF}	Mean.	(1,0.9874)	(1,0.8978)	(1,-9.098)
	Std.	(0,0.0120)	(0,0.1548)	(0,33.66)
MS	Mean.	0.9966	0.9977	0.9892
	Std.	0.0053	0.0040	0.0312
NOS	Mean.	145.8	180.1	238.8
	Std.	17.8	44.5	22.2

Table 5 showed the percentage of branch II activation by 30 independent operations for different mathematical cases. ZDT1, ZDT2, and ZDT6 were chosen to verify the effect of branch II and the MMMSEC criterion on the PFVNASG-II optimization procedure. Although branch II was activated in 90% of the independent operations in the ZDT4 mathematical case, most of the operations in ZDT4 did not meet the convergence criterion, so ZDT4 was not selected. For comparison, for every independent operation, the computation was suspended when $N_{PF}^{t} > N_{PF}^{*}$ *and* $P_{PF}^{t-\alpha} \in P_{PF}^{t}$. Then the branch I loop as well as the branch II loop were performed independently. A total of 20 independent operations were performed.

Table 5. Activation rate of branch II in different mathematical cases.

	ZDT1	ZDT2	ZDT3	ZDT4	ZDT6
Total run times	30	30	30	30	30
Activation rate of branch II	60%	70%	23.3%	90%	73.3%

Table 6 compared the effects of branch II and the MMMSEC criterion on the convergence speed of the PFVNASG-II optimization program. As can be seen from the table, the optimization procedure without branch II activation had the possibility of convergence failure (under the specified MNOS) in all 20 independent operations. This implied that branch II with MMMSEC criterion is beneficial to improve the

stability of the optimization algorithm. It can also be seen that activating branch II can reduce the computational cost by 21.7%, 18%, and 11.2%, respectively. The branch II and MMMSEC criterion were beneficial to improve the convergence speed of the optimization algorithm.

Table 6. The effects of branch II and MMMSEC criterion on the prediction performance of PFVNASG-II.

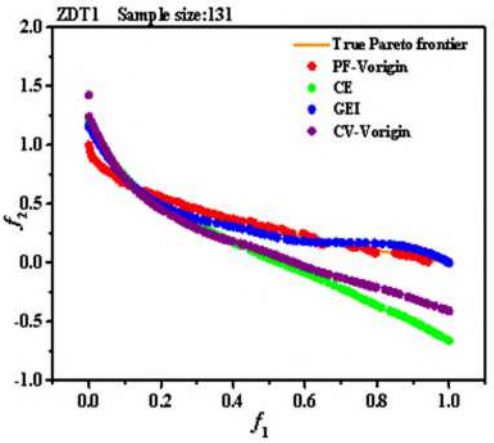
		ZDT1		ZDT2		ZDT6	
		I&II	I	I&II	I	I&II	I
Total run times		20	20	20	20	20	20
The number of IGD < IGD_0		20	16	20	19	20	15
R^2_{PF}	Mean	(1,0.9979)	(1,0.6576)	(1,0.9998)	(1,0.9995)	(1,0.9732)	(1,0.-57.50)
	Std.	(0,0.0032)	(0,0.7772)	(0,0.0005)	(0,0.0009)	(0,0.0333)	(1,0.254.44)
IGD	Mean	0.0058	0.02196	0.0046	0.0047	0.0073	0.2318
	Std.	0.0004	0.0387	0.0002	0.0005	0.0016	0.8322
MS	Mean	0.9929	0.9398	0.9935	0.9922	0.9991	1.0273
	Std.	0.0100	0.1925	0.0061	0.0151	0.0019	04736
NOS	Mean	129.7	165.6	96.50	117.75	182.8	205.9
	Std.	18.2	34.0	11.62	30.15	24.5	33.8

I&II: Call branch I and branch II I: Only branch I is called

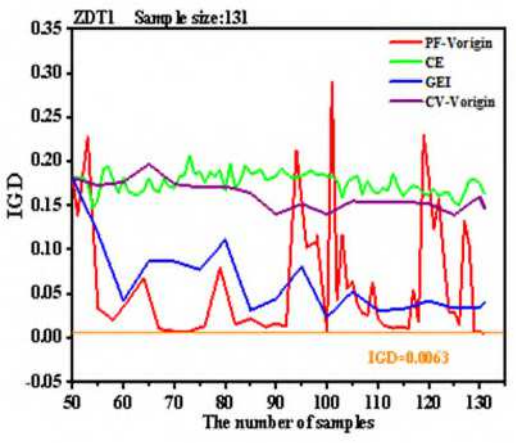
3.3 Comparison with existing algorithms

To examine the performance of the PF-Voronoi sampling algorithm, the CV-Voronoi, GEI, and CE sampling algorithms were selected for comparison with it. To ensure that all algorithms are under the same conditions, these four sampling algorithms were coupled with the Kriging model as well as the NASG-II algorithm. Fig.6 (a)(c)(e)(g) showed the PF predictions of the four sampling algorithms for the same number of sample points. Except for ZDT4, the optimal frontier predicted using the PF-Voronoi sampling algorithm largely overlaps with the true PF. This indicates

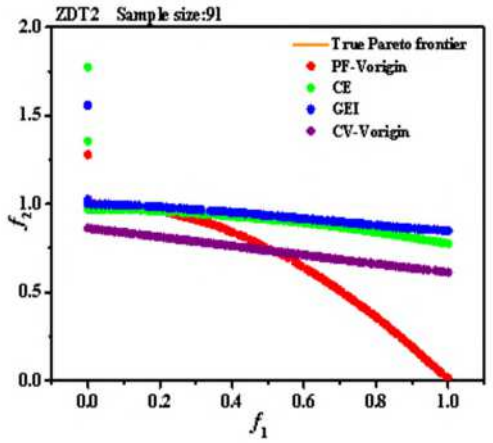
that the PF-Voronoi sampling algorithm can predict PF faster and more accurately. Besides, from Fig.6(b)(d)(f)(h), we can find that the IGD of the PF-Voronoi algorithm converges fastest; followed by the GEI, which can converge quickly in some mathematical cases; while the CV-Voronoi and CE algorithms converge slowly.



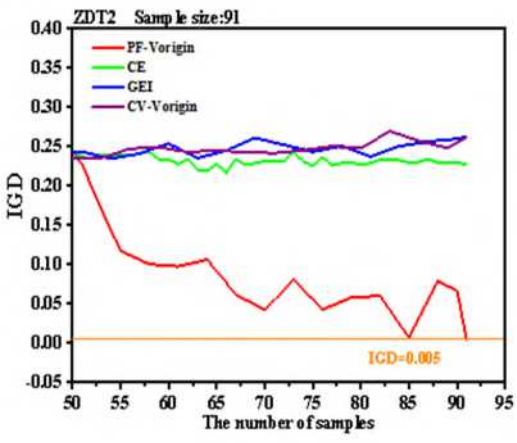
(a)The PF of ZDT1



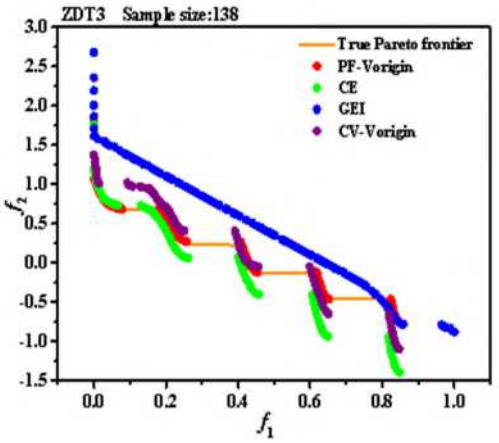
(b)The IGD of ZDT1



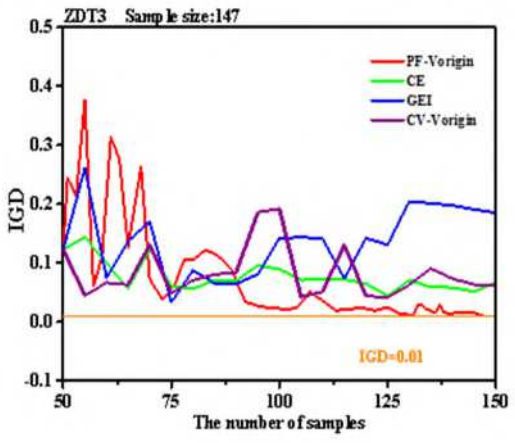
(c)The PF of ZDT2



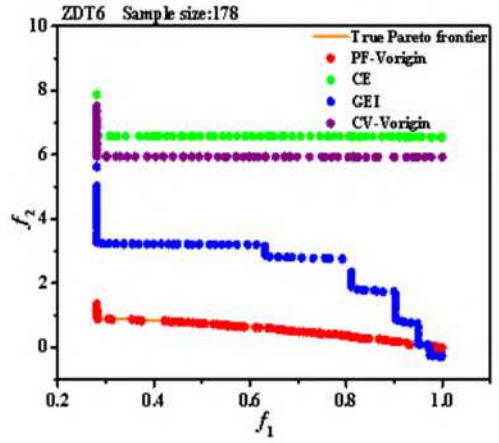
(d)The IGD of ZDT2



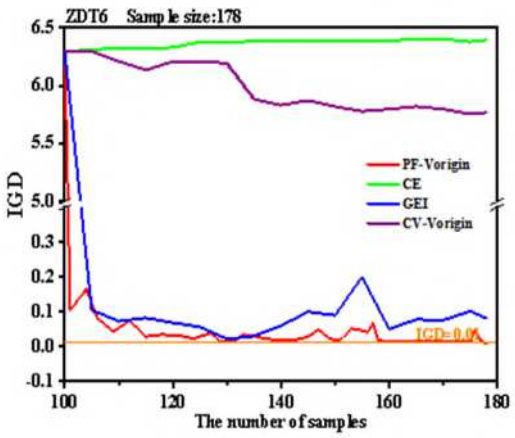
(e)The PF of ZDT3



(f)The IGD of ZDT3



(g)The PF of ZDT6

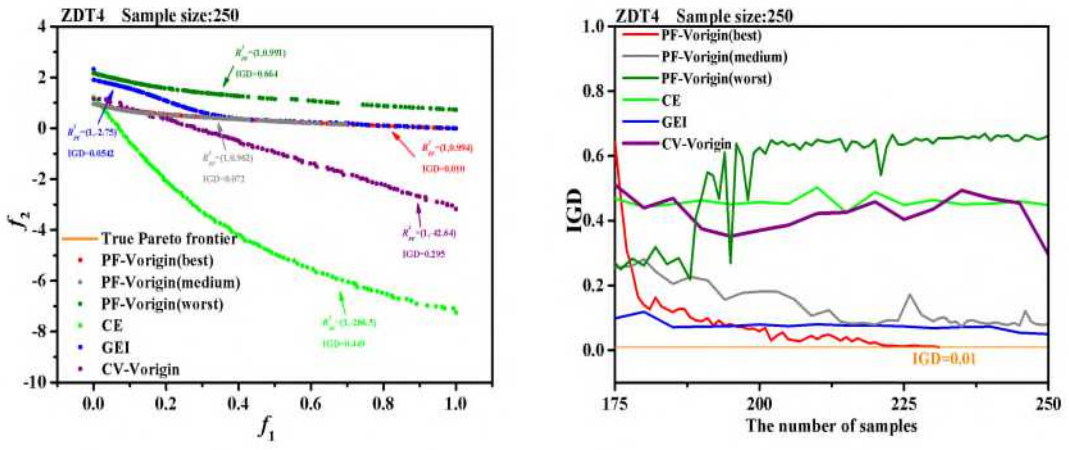


(h)The IGD of ZDT6

Fig.7 Predictive power of three optimization procedures for the ZDT1- ZDT3, ZDT6 Pareto frontier.

Fig.8 compared the PF predictions of the four sampling algorithms with the same number of sample points for the mathematical case of ZDT4. PF-Voronoi (best), PF-Voronoi (medium), and PF-Voronoi (worst) represent the best, medium, and worst IGD convergence of the PF-Voronoi sampling algorithm for 30 independent operations, respectively. From the figure, it can be found that although PF-Voronoi (worst) had the worst IGD convergence curve, its predicted optimal frontier was close to the true Pareto frontier overall, and R^2_{PF}=(1,0.991), the predicted value was close to

the actual value. While the optimal frontier predicted by EI fitted the true Pareto frontier curve more closely than PF-Voronoi (worst), but R^2_{PF}= (1,-2.75), the prediction error was larger and the predicted value was not credible. The optimal frontier predicted by PF-Voronoi (medium) fitted part of the true Pareto frontier curve with poorer distribution diversity. This was the reason for its higher IGD. In summary, PF-Voronoi was more likely to approximate the true Pareto front than the other three sampling algorithms.



(a)The The PF of ZDT4 (b)The IGD of ZDT6

Fig.8 Predictive power of three optimization procedures for the ZDT4 Pareto frontier.

Table 7 compared the prediction performance of the PFVNASG-II optimization procedure with those proposed by researchers in recent years. The data were obtained from the open literature and the dimensionality of the variables was kept consistent. As can be seen from the table, PFVNASG-II exhibited the best prediction performance in the five mathematical cases, while only slightly worse than EHVIMOPSO in ZDT2. Comparing the NOS of every optimization procedure reveals that the PFVNASG-II optimization procedure required the lowest computational cost, which is less than 44.2% of the other optimization procedures.

Table 8 showed the detailed prediction results (mean and standard deviation of 30 independent runs) for PFVNASG-II. As can be seen from the table, the R^2_{PF} is above 0.95 (except for ZDT4), indicating that the error between the predicted and observed values is small. From the IGD, it can be seen that the predictions of PFVNASG-II for the remaining four mathematical cases converge successfully except for ZDT4. The results of IGD and R^2_{PF} showed that PFVNASG-II can accurately predict the PFs of the four mathematical cases and their locations in the design space.

In summary, the PFVNASG-II optimization procedure maintained high prediction accuracy and distribution diversity at a lower computational cost.

Table 7. Mean and standard deviations of IGD and NOS were obtained by different optimizations.

		ZDT1		ZDT2		ZDT3		ZDT4		ZDT6	
		Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
PFVNASG-II	IGD	**0.0056**	**0.0005**	0.0046	0.0002	**0.0080**	**0.0013**	**0.1266**	**0.1418**	**0.0076**	**0.0015**
	NOS	111.6	28.4	95.2	11.6	145.8	17.8	248.6	4.7	170.9	28.2
EHVIMOPSO[22]	IGD	0.0063	0.0017	**0.0044**	**0.0004**	0.0154	0.0162	-	-	0.019	0.015
	NOS	850	0	400	0	850	0	-	-	800	0
SAO-MOEA [23]	IGD	0.0062	0.0009	0.0844	0.1821	0.1320	0.0845	-	-	-	-
	NOS	200	0	200	0	200	0				
TIC-SMEA [24]	IGD	0.0112	0.0027	0.0124	0.0125	0.1124	0.0527	-	-	-	-
	NOS	200	0	200	0	200	0	-	-	-	-
MO-MOEA-C[25]	IGD	35.51	5.61	41.23	6.82	36.26	4.42	-	-	-	-
	NOS	200	0	200	0	200	0	-	-	-	-
K-RVEA[33]	IGD	-	-	-	-	-	-	29.96	13.9	1.467	0.248
	NOS	-	-	-	-	-	-	300	0	300	0
SAEA/ME [33]	IGD	-	-	-	-	-	-	53.98	18	0.1188	0.0551
	NOS	-	-	-	-	-	-	300	0	300	0
TR-NSGA-II [34]	IGD	-	-	-	-	-	-	6.98	-	0.3107	-
	NOS	-	-	-	-	-	-	500	0	500	0

Table 8. Detailed prediction results of PFVNASG-II.

	ZDT1	ZDT2	ZDT3	ZDT4	ZDT6
Total run times	30	30	30	30	30
R^2_{PF} Mean	(1,0.9973)	(1,0.9998)	(1,0.9874)	(1,-1.4684)	(1,0.9590)

	Std.	(0,0.0034)	(0,0.0004)	(0,0.0120)	(0,5.8716)	(0,0.0351)
IGD	Mean	0.0056	0.0046	0.0080	0.1266	0.0076
	Std.	0.0005	0.0002	0.0013	0.1418	0.0015
MS	Mean	0.9798	0.9941	0.9967	0.8136	0.9994
	Std.	0.0217	0.0057	0.0053	0.1843	0.0016
	Max.	179	125	174	250	238
NOS	Mean	111.6	95.2	145.8	248.5	170.9
	Min.	72	78	108	231	121

Finally, it needs to be specified that the PFVNASG-II optimization algorithm had difficulty predicting the mathematical case of ZDT4. In fact, for ZDT4, there were only 3 independent runs where the PFVNASG-II optimization algorithm successfully converges. Besides, there were 12 independent runs where the predicted Pareto frontier lay on the true Pareto frontier but performs poorly in terms of distributional diversity (IGD < 0.1 and MS < 0.9), i.e., the IGD converged slowly. It does not mean that the PF-Voronoi sampling algorithm is seriously flawed, but it is the price that all PF-based sampling algorithms have to pay: weakening the global exploration capacity to reduce the computational cost and improve the local prediction accuracy. Therefore, the PF-Voronoi sampling algorithm needs to be coupled with a sampling algorithm with better global exploration capability to extend its adaptability, which will be the next research direction.

Conclusion

In this work, a PF-based sampling algorithm, the PF-Vorigin sampling algorithm, was proposed. It was applied to the Kriging and NASG-II coupled models and its effectiveness was verified on the ZDT mathematical cases.

The experimental results showed that the MCC criterion was effective in improving the diversity of the optimization procedure. While the MLEC criterion and

MMMSEC criterion were beneficial in accelerating the convergence of the algorithm, reducing the number of training samples by 38.9% and 21.7%, respectively. The application of the GEI sampling algorithm helped the PFVNASG-II optimization procedure to jump out of the local PF. For the ZDT6 Pareto frontier prediction, there existed a 75% probability that the prediction without the tandem GEI fell into the local Pareto frontier.

The PF-Vorigin sampling algorithm had higher prediction performance than the CE, CV-Voronoi, and GEI sampling algorithms on all the ZDT mathematical cases. In comparison with other optimization procedures in the open literature, the number of training samples of PFVNASG-II was only 55.8% of that of optimized procedures such as HVIMOPSO and SAO-MOEA, and $R^2_{PF} > 0.95$ (except ZDT4).

For the mathematical case of ZDT4 with multiple local PFs, PFVNASG-II did not successfully converge for a given MNOS. However, compared to other optimization procedures, PFVNASG-II can bring the IGD down to about 0.12, which was closer to the true Pareto front.

Funding

Not applicable

Conflicts of interest

The authors declare that they have no conflict of interest.

Informed consent

Informed consent was obtained from all individual participants included in the study.

Availability of data and material

The datasets used or analysed during the current study are available from the corresponding author on reasonable request.

Reference

[1] Ibaraki S, Van den Braembussche R, Verstraete T, Alsalihi Z, Sugimoto K, Tomita I. Aerodynamic design optimization of a centrifugal compressor impeller based on an artificial neural network and genetic algorithm. In: Institute Of Mechanical Engineers, 'editors'. 11th International Conference on Turbochargers and Turbocharging. Oxford: Woodhead Publishing; 2014. p. 65-77.

[2] Li Z, Zheng X. Review of design optimization methods for turbomachinery aerodynamics. Prog Aerosp Sci 2017;93:1-23.

[3] Wang GG, Shan S. Review of metamodeling techniques in support of engineering design optimization. J Mech Design 2007;129:370-380.

[4] Gu L. A comparison of polynomial based regression models in vehicle safety analysis. Proceedings of the ASME Design Engineering Technical Conference 2001;2:509-514.

[5] Yang Q, Lin Y, Guan G. Improved sequential sampling for meta-modeling promotes design optimization of SWATH. Ocean Eng 2020;198.

[6] Wang XD, Hirsch C, Kang S, Lacor C. Multi-objective optimization of turbomachinery using improved NSGA-II and approximation model. Comput Method Appl M 2011;200:883-895.

[7] Ibaraki S, Van den Braembussche R, Verstraete T, Alsalihi Z, Sugimoto K,

Tomita I. Aerodynamic design optimization of a centrifugal compressor impeller based on an artificial neural network and genetic algorithm. 2014. p. 65-77.

[8] Ekradi K, Madadi A. Performance improvement of a transonic centrifugal compressor impeller with splitter blade by three-dimensional optimization. Energy 2020;201.

[9] Wang XF, Xi G, Wang ZH. Aerodynamic optimization design of centrifugal compressor's impeller with Kriging model. P I Mech Eng A-J Pow 2006;220:589-597.

[10] Guo S, Duan F, Tang H, Lim SC, Yip MS. Multi-objective optimization for centrifugal compressor of mini turbojet engine. Aerosp Sci Technol 2014;39:414-425.

[11] Crombecq K, Gorissen D, Deschrijver D, Dhaene T. A NOVEL HYBRID SEQUENTIAL DESIGN STRATEGY FOR GLOBAL SURROGATE MODELING OF COMPUTER EXPERIMENTS. Siam J Sci Comput 2011;33:1948-1974.

[12] Chen Q, Ni J, Wang Q, Shi X. Match-based pseudo-MAP full-operation-range optimization method for a turbocharger compressor. Struct Multidiscip O 2019;60:1139-1153.

[13] Liu H, Xu S, Wang X. Sequential sampling designs based on space reduction. Eng Optimiz 2015;47:867-884.

[14] Eason J, Cremaschi S. Adaptive sequential sampling for surrogate model generation with artificial neural networks. Comput Chem Eng 2014;68:220-232.

[15] Viana FAC, Gogu C, Haftka RT. MAKING THE MOST OUT OF

SURROGATE MODELS: TRICKS OF THE TRADE. 2010. p. 587-598.

[16] Xu S, Liu H, Wang X, Jiang X. A Robust Error-Pursuing Sequential Sampling Approach for Global Metamodeling Based on Voronoi Diagram and Cross Validation. J Mech Design 2014;136.

[17] Sasena MJ, Papalambros P, Goovaerts P. Exploration of metamodeling sampling criterion for constrained global optimization. Eng Optimiz 2002;34:263-278.

[18] Guo Z, Song L, Li J, Li G, Feng Z. RESEARCH ON META-MODEL BASED GLOBAL DESIGN OPTIMIZATION AND DATA MINING METHODS. 2015.

[19] Li X, Zhao Y, Liu Z. A novel global optimization algorithm and data-mining methods for turbomachinery design. Struct Multidiscip O 2019;60:581-612.

[20] Emmerich MTM, Giannakoglou KC, Naujoks B. Single-objective and multiobjective evolutionary optimization assisted by Gaussian random field metamodels. Ieee T Evolut Comput 2006;10:421-439.

[21] Pilát M, Neruda R. Aggregate meta-models for evolutionary multiobjective and many-objective optimization. Neurocomputing 2013;116:392-402.

[22] Cheng S, Zhan H, Shu Z, Fan H, Wang B. Effective optimization on Bump inlet using meta-model multi-objective particle swarm assisted by expected hyper-volume improvement. Aerosp Sci Technol 2019;87:431-447.

[23] Li F, Gao L, Shen W, Cai X, Huang S, "A Surrogate-Assisted Offspring

Generation Method for Expensive Multi-objective optimization Problems," in 2020 IEEE Congress on Evolutionary Computation (CEC), (IEEE, 2020), pp. 1-8.

[24] Li F, Gao L, Garg A, Shen W, Huang S. Two infill criterion driven surrogate-assisted multi-objective evolutionary algorithms for computationally expensive problems with medium dimensions. Swarm Evol Comput 2021;60.

[25] Li F, Gao L, Garg A, Shen W, Huang S. A comparative study of pre-screening strategies within a surrogate-assisted multi-objective algorithm framework for computationally expensive problems. Neural Comput Appl 2020.

[26] Aurenhammer F. Voronoi diagrams—a survey of a fundamental geometric data structure. ACM Computing Surveys (CSUR) 1991;23:345-405.

[27] Lophaven SN, Nielsen HB, Sondergaard J. DACE—a Matlab Kriging toolbox; version 2; informatics and mathematical modelling. Technical University of Denmark, Copenhagen 2002.

[28] Park J. Optimal Latin-hypercube designs for computer experiments. J Stat Plan Infer 1994;39:95-111.

[29] Zhou Y, Lu Z. An enhanced Kriging surrogate modeling technique for high-dimensional problems. Mech Syst Signal Pr 2020;140.

[30] Jones DR, Schonlau M, Welch WJ. Efficient global optimization of expensive black-box functions. J Global Optim 1998;13:455-492.

[31] Jie H, Wu Y, Zhao J, Ding J, Liangliang. An efficient multi-objective PSO algorithm assisted by Kriging metamodel for expensive black-box problems. J Global Optim 2017;67:399-423.

[32] Deb K, Agrawal S, Pratap A, Meyarivan T, "A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II," in International conference on parallel problem solving from nature, (Springer, 2000), pp. 849-858.

[33] Ruan X, Li K, Derbel B, Liefooghe A. Surrogate Assisted Evolutionary Algorithm for Medium Scale Multi-Objective Optimisation Problems. 2020. p. 560-568.

[34] Roy PC, Hussein R, Blank J, Deb K, "Trust-Region Based Multi-objective Optimization for Low Budget Scenarios," in International Conference on Evolutionary Multi-Criterion Optimization, (Springer, 2019), pp. 373-385.

Nomenclature

ANN	Artificial neural network	MNOS	Maximum number of samples
CE	Combination of expectations	MOO	Multiobjective optimization
EI	Expected improvement	MS	Maximum Spread
GA	Genetic algorithm	NOS	Number of samples
GEI	Generalized expected improvement	OLH	Optimal Latin Hypercube
IGD	Inverter generational distance	PF	Pareto frontier
MCC	Maximum crowding criterion	PFC	Pareto frontier cell
MLEC	Maximum LOO error criterion	RSM	Response surface method
MMMSEC	Maximum mean MSE criterion		