Author(s): Aghaei Pour, Pouya; Bandaru, Sunith; Afsar, Bekir; Emmerich, Michael; Miettinen, Kaisa

Title: A Performance Indicator for Interactive Evolutionary Multiobjective Optimization Methods

Please cite the original version:

Aghaei Pour, P., Bandaru, S., Afsar, B., Emmerich, M., & Miettinen, K. (2023). A Performance Indicator for Interactive Evolutionary Multiobjective Optimization Methods. IEEE Transactions on Evolutionary Computation, 28(3), 778-787. https://doi.org/10.1109/TEVC.2023.3272953

A Performance Indicator for Interactive Evolutionary Multiobjective Optimization Methods

Pouya Aghaei Pour, Sunith Bandaru, *Senior Member, IEEE*, Bekir Afsar,
Michael Emmerich, and Kaisa Miettinen

Abstract—In recent years, interactive evolutionary multiobjective optimization methods have been getting more and more attention. In these methods, a decision maker (DM), who is a domain expert, is iteratively involved in the solution process and guides the solution process toward her/his desired region with preference information. However, there have not been many studies regarding the performance evaluation of interactive evolutionary methods. On the other hand, indicators have been developed for a priori methods, where the DM provides preference information before optimization. In the literature, some studies treat interactive evolutionary methods as a series of a priori steps when assessing and comparing them. In such settings, indicators designed for a priori methods can be utilized. In this article, we propose a novel performance indicator for interactive evolutionary multiobjective optimization methods and show how it can assess the performance of these interactive methods as a whole process and not as a series of separate steps. In addition, we demonstrate the shortcomings of using indicators designed for a priori methods for comparing interactive evolutionary methods.

Index Terms—Decision making, hypervolume indicator, interactive evolutionary algorithms, method comparison, quality indicators.

I. INTRODUCTION

IN MANY real-world problems, we have to optimize multiple conflicting objectives simultaneously. Because of the conflicting nature of the objectives, these problems have a set of optimal solutions, which we refer to as *Pareto optimal solutions*. They represent different tradeoffs and are mathematically incomparable [1]. The set of Pareto optimal solutions is referred to as a *Pareto front* in the objective space.

Many multiobjective optimization methods have been developed that can optimize multiple conflicting objectives simultaneously, e.g., [1], [2], and [3] and references therein. Among the methods, evolutionary ones are getting more attention because, as populations-based methods, they can generate an approximation of the Pareto front. In addition, evolutionary methods can handle different types of decision variables and can be applied to problems with different levels of complexities (e.g., discontinuous or nondifferentiable functions).

For practical problems, one of the Pareto optimal solutions has to be chosen for implementation. Therefore, additional preference information, which usually comes from a domain expert called a decision maker (DM), is required to identify the *most preferred solution*.

Multiobjective optimization methods can be classified based on how and when the DM takes part in the solution process [1]. In *a posteriori methods*, a representative set of Pareto optimal solutions is first generated, and the DM has to find the most preferred solution at the end of the solution process. On the other hand, in *a priori methods*, the DM provides her/his preferences before the solution process begins, and the method tries to generate solutions in the corresponding *desired region* that reflect these preferences. Finally, in *interactive methods*, the DM provides her/his preferences iteratively and guides the search actively. This means that the DM can learn about the interdependencies among the objectives as well as the achievability of the preferences.

In many practical cases, it has been observed that interactive solution processes have two phases [4]. We refer to these phases as the *learning phase* and the *decision phase*. In the learning phase, the DM explores the objective space by providing different preferences and learning about solutions' interdependencies and tradeoffs. In addition, the DM learns about the feasibility of her/his preferences and gradually guides the method toward her/his *region of interest*. Then, in the *decision phase*, the DM fine-tunes solutions in the region of interest until the most preferred solution is found.

Due to the vast number of existing methods, it is important to assess their performance to select the most appropriate method for a given problem. However, unlike single-objective optimization methods, where the assessment is directly connected to the objective function values, assessing the performance of multiobjective optimization methods is not a straightforward task because of the existence of Pareto optimal solutions. Some desirable properties that performance

indicators (indicators for short) for a posteriori, a priori, and interactive methods should ideally possess have been identified in [5], [6], and [7], respectively.

The field of evolutionary multiobjective optimization methods is especially keen on using indicators. In particular, many indicators have been proposed to assess the performance of a posteriori methods [8], [9]. These indicators assess the performance in approximating the whole Pareto front. Furthermore, a few studies have been dedicated to developing indicators for a priori methods [6], [10], [11], [12], [13], [14]. These indicators assess the performance in generating a set of solutions representing parts of the approximated Pareto front which best reflects the DM's preferences. However, to the best of our knowledge, no indicators for assessing the performance of interactive evolutionary methods have been developed. Note that, for simplicity, in the rest of this article, we simply use the term "methods" when referring to evolutionary multiobjective optimization methods.

There have been some attempts to utilize the indicators developed for a priori methods to assess the performance of interactive ones. For instance, in [15], [16], and [17] interactive methods are viewed as a series of a priori steps to be able to apply the above-mentioned indicators. In such settings, an indicator designed for a priori methods is used for each interaction step, and the average of these indicator values is calculated as the final assessment. However, this way of performance assessment is not ideal. Some examples have been provided in [7] showing how this approach can be misleading. According to [7], indicators should consider the nature of the interactive methods where the DM learns more about the problem and her/his preferences during the solution process. Furthermore, as learning and decision phases have different characteristics [18], we may need different indicators.

In this article, we propose a novel indicator for assessing the performance of interactive methods. Furthermore, we demonstrate how to utilize this indicator to assess the learning and decision phases. Then, an engineering test problem is considered, where we apply two interactive methods and use the proposed indicator to assess their performances. Additionally, we utilize some of the indicators developed for a priori methods to demonstrate why it is important to design indicators explicitly for interactive methods.

The remainder of this article is arranged as follows. In Section II, we provide necessary terminologies and background on this topic. We introduce our novel indicator in Section III. Thereafter in Section IV, we describe the interactive solution process on an engineering test problem and demonstrate how one can utilize the indicator to assess learning and decision phases. Finally, we provide conclusions and future research directions in Section V.

II. BACKGROUND

In this section, we provide necessary terminologies and give some background on attempts that have been made to compare interactive methods. Additionally, we briefly discuss the desirable properties that an indicator for interactive methods should possess.

A. Multiobjective Optimization

The general form of multiobjective optimization problems can be formulated as follows:

$$\text{minimize } f(x) = (f_1(x), \ldots, f_k(x))$$
$$\text{subject } x \in S \tag{1}$$

where $f_i : S \rightarrow \mathbb{R}$ are k (conflicting) objectives to be minimized simultaneously and $f(x) \in \mathbb{R}^k$ in the objective space is an objective vector, which we refer to as a solution. Its components are objective (function) values. Finally, $x = (x_1, \ldots, x_n)^T$, is an n-dimensional decision variable vector in $S \subset \mathbb{R}^n$.

A solution $f(x)$ *dominates* $f(y)$ [written as $f(x) \succ f(y)$] if $f_i(x) \leq f_i(y)$ for $i = 1, \ldots, k$ and $f_j(x) < f_j(y)$ for at least one index j. A solution is Pareto optimal if it is not dominated by any other solution. If $f(x)$ and $f(y)$ do not dominate each other, they are called mutually *nondominated*. Evolutionary methods cannot guarantee Pareto optimality, and we can only make sure that the final population is mutually nondominated. Therefore, we refer to approximated solutions in this article.

A vector that is constructed from the best objective values found in the approximated Pareto front is referred to as an *ideal point* $z^\star$. The point that is constructed from the worst objective values in the approximated Pareto front is referred to as a *nadir point* z^{nad}. In the multiple criteria decision-making literature [19], a point in the objective space that is slightly better than the ideal point is called a *utopian point*. Analogously to this definition, we refer to a point in the objective space that is slightly worse than the nadir point as a *dystopian point* z^{dy}. In this article, we set the components of the dystopian point as follows:

$$z_i^{dy} = z_i^{\text{nad}} + 10^{-4}$$

where $i = 1, \ldots, k$.

B. Interactive Methods

As mentioned earlier, there are many interactive methods. They differ, e.g., in the optimization engine they use, the type of preferences that they incorporate, and the type of information they offer to the DM [4], [20].

Before a solution process can be started, a suitable method needs to be chosen. Here, usually, someone who knows the methods helps the DM in making this choice taking into account the properties of the problem and the desires of the DM. We refer to this person as an *analyst* [1].

After an appropriate method has been chosen, the DM starts the interactive solution process by providing her/his preferences. We refer to the act of providing new preferences as an *interaction*. After each interaction, the method generates a set of solutions that reflects the DM's preferences as well as possible and presents them to the DM. Then, either the DM is satisfied by one of these solutions and terminates the solution process, or (s)he updates the preferences and waits for a new set of solutions to be generated.

As mentioned in Section I, in practical problems, one can often observe distinct learning and decision phases. Distinguishing between learning and decision phases is not a

trivial task during the solution process, as the DM may switch between them at any point in the solution process. In this study, we assume that the DM starts with a learning phase and ends with a decision phase. It is important to clarify the meanings of the terms "desired region" and "region of interest." In what follows, a desired region is a part of the Pareto front that reflects the DM's current preferences, while a region of interest is identified at the end of the learning phase, after which the DM fine-tunes solutions in it and finds the most preferred solution in the decision phase.

C. Related Works

In [18], different ways for assessing the performance of interactive methods have been surveyed. Here, the authors provide desirable properties that interactive methods should have and divide the properties into three categories.

1) General properties (GPs).
2) Learning phase properties (LPs).
3) Decision phase properties (DPs).

According to [18], GPs should be considered in both phases, while LPs and DPs are more sophisticated properties corresponding to the learning and decision phases, respectively. Moreover, the authors emphasize that new indicators for assessing interactive methods are needed.

To be able to develop new indicators for interactive methods, it is important to characterize their desirable properties. This was done in [7]. Following the reasoning in [18], these properties have also been divided into three categories (GPs, LPs, and DPs). In [7], nine GPs, two LPs, and two DPs have been identified. Here, we provide a list of them. An indicator for interactive methods should be able to:

GP1: assess the convergence of solutions in those regions of the approximated Pareto front that reflect the DM's preferences the best (local convergence);
GP2: assess the diversity of solutions in those regions of the approximated Pareto front that reflect the DM's preferences the best (local diversity);
GP3: assess the performance irrespective of the number of objective functions (scalability);
GP4: assess the performance without knowledge of the Pareto front;
GP5: assess the performance by incorporating preferences that are provided in different ways;
GP6: assess the performance in a computationally inexpensive manner;
GP7: assess the performance in a manner that is independent of other interactive methods being compared;
GP8: assess the performance without introducing parameters that have an unclear effect on the performance or are unintuitive to set;
GP9: assess the performance as a whole process and not as a series of independent a priori steps;
LP1: assess how much of the Pareto front has been studied (expedition);
LP2: assess how well/fast the method can adapt to new (even very different) preferences (responsiveness);

DP1: assess the capability of fine-tuning solutions inside the region of interest;
DP2: assess the decision phase by considering the amount of information shown to the DM at each interaction.

The use of an artificial DM (ADM) for comparing interactive methods has been getting more attention in recent years (e.g., [15] and [16]). ADMs have two main parts. First, they use different strategies to provide preferences for the learning and decision phases. Second, they analyze the results and compare the interactive solutions process. In the supplementary materials, we use the ADM proposed in [15] only for the first part (generating preferences) and use our proposed indicator to compare the performance of interactive methods.

The advantage of ADMs is that they can be used to assess interactive methods without involving human DMs, which means that many repetitive comparisons can be done in stable conditions (though ADMs cannot assess all aspects related to human behavior). In the absence of appropriate indicators, the ADMs proposed so far mainly view interactive solution processes as a series of a priori steps and use a priori indicators for comparison.

III. ASSESSING INTERACTIVE METHODS

It can be challenging to design a single indicator that possesses all desirable properties listed in Section II. Instead, it is reasonable to design multiple indicators covering different desirable properties [7]. (Like there are different indicators for a posteriori methods reflecting different desirable properties [8], [9].) In the following, we introduce a new indicator called preference-based hypervolume indicator (PHI) for assessing the performance of interactive methods. PHI possesses some of the previously mentioned desirable properties. We also discuss how to utilize it to assess learning and decision phases. We have provided all the notations we use for the rest of this article in Table I.

A. PHI Description

PHI incorporates the DM's preferences in the form of a reference point to construct a desired region. In this study, we define the *desired region* as a region of the objective space enclosed by a hyperrectangle with corners $\hat{z}$ and z^{dy} (denoted as a blue rectangle in Fig. 1). Moreover, PHI uses the DM's reference point $\hat{z}$ to divide the set P into three subsets: 1) solutions $P^{\succ}$ that dominate $\hat{z}$; 2) solutions $P^{\prec}$ that are dominated by $\hat{z}$; and 3) solutions $P^{=}$ that neither dominate nor are dominated by $\hat{z}$.

First, we can express v^{-} to cover both cases in Fig. 1 as follows:

$$v^{-} = HV\left(P \cup \{\hat{z}\}, z^{dy}\right) - HV\left(P^{\succ} \cup \{\hat{z}\}, z^{dy}\right). \quad (2)$$

Based on (2), we calculate $v^{\prec}$ for each case separately as follows:

$$v^{\prec} = \begin{cases} HV\left(P, z^{dy}\right) - v^{-}, & \text{if } P^{\succ} = \varnothing \\ HV\left(\hat{z}, z^{dy}\right), & \text{otherwise.} \end{cases} \quad (3)$$

TABLE I
NOTATIONS AND SYMBOLS THAT WE HAVE USED THROUGH OUT THIS ARTICLE

| Notation	Description
$\hat{z}$	A reference point consisting of desirable values for each objective function
$\hat{z}^d$	Reference point provided in the last interaction
z^{dy}	Dystopian point
P	A set of nondominated solutions to be assessed
$P^{\prec}$	A set of nondominated solutions that are dominated by $\hat{z}$
$P^{\succ}$	A set of nondominated solutions that dominate $\hat{z}$
$P^{=}$	A set of nondominated solutions do not dominate $\hat{z}$ nor are dominated by it
$HV(P, z^{dy})$	The hypervolume function that calculates (or approximates) the hypervolume indicator measure for a solution set P with regard to the dystopian point z^{dy}
$v^{\prec}$	A part of the hypervolume indicator measure $HV(P, z^{dy})$ that is *inside* the desired region (the green areas in Figure 1)
$v^{\succ}$	A part of the hypervolume indicator measure $HV(P^{\succ}, z^{dy})$ that is *outside* the desired region (the orange area in Figure 1b)
v^{-}	A part of the individual hypervolume indicator contribution of $P^{=}$ that is *outside* the desired region (the red areas in Figure 1)
v^{+}	A part of hypervolume measure of the solution set P that have a positive effect in the value of PHI ($v^{\prec} + v^{\succ}$)
$PHI(P, \hat{z}, z^{dy})$	PHI assessment of a solution set P with regard to $\hat{z}$ and z^{dy}
RS	Assessment of the learning phase for an interactive method
ζ_j	The set of the jth and the last reference point ($\zeta_j = \{\hat{z}^j, \hat{z}^d\}$)
v_j^d	The individual hypervolume indicator contributions of $\hat{z}^d$ to $HV(\zeta_j, z^{dy})$, which is denoted as the yellow area in Figure3
v^j	The shared hypervolume indicator contribution of $\hat{z}^j$ and $\hat{z}^d$ to $HV(\zeta_j, z^{dy})$, which is denoted as the purple area in Figure3
λ^j	The coefficient that indicates the similarity between interaction jth and the last interaction
FD	Assessment of the decision phase for an interactive method
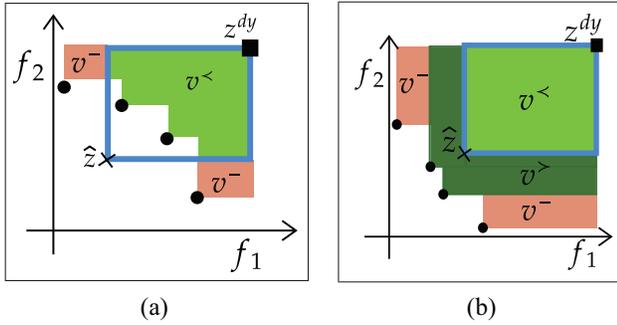
Fig. 1. Examples of calculating PHI. The ideal point is denoted by a purple star, and the dystopian point by a green square. The desired region is shown as a blue rectangle. The areas denoted as $v^{\prec}$, and $v^{\succ}$ have positive effects on PHI assessment, and the areas denoted as v^{-} have negative effects. (a) *Case 1:* $P^{\succ} = \varnothing$. (b) *Case 2:* $P^{\succ} \neq \varnothing$.

Finally, $v^{\succ}$ is obtained as follows:

$$v^{\succ} = \begin{cases} 0, & \text{if } P^{\succ} = \varnothing \\ HV(P^{\succ}, z^{dy}) - v^{\prec}, & \text{otherwise.} \end{cases} \quad (4)$$

The central idea of our proposed indicator is to divide the measured region that determines the hypervolume indicator of P into *positive* and *negative* parts, $HV(P, z^{dy}) = v^{+} + v^{-}$, each representing how they affect the assessment of an interactive method. The positive part v^{+} is formed by solutions that are of immediate interest to the DM, namely, the subsets $P^{\prec}$ or $P^{\succ}$, and is given by

$$v^{+} = v^{\prec} + v^{\succ} = \begin{cases} HV(P, z^{dy}) - v^{-}, & \text{if } P^{\succ} = \varnothing \\ HV(P^{\succ}, z^{dy}), & \text{otherwise.} \end{cases} \quad (5)$$

On the other hand, the negative part, v^{-} calculated in (2), is formed by solutions that are not of immediate interest to the DM, i.e., the solutions that belong to subset $P^{=}$.

The value of v^{+} represent the positive effect of the solutions in calculate of PHI. To punish the performance, we use the value of v^{-} by normalizing $v^{\prec}$ and $v^{\succ}$ as follows:

$$\begin{aligned} PHI(P, \hat{z}, z^{dy}) &= \frac{v^{\prec}}{HV(\hat{z}, z^{dy})} + \frac{v^{\succ}}{HV(P, z^{dy})} \\ &= \begin{cases} \frac{v^{\prec}}{HV(\hat{z}, z^{dy})}, & \text{if } P^{\succ} = \varnothing \\ 1 + \frac{v^{\succ}}{HV(P, z^{dy})}, & \text{otherwise.} \end{cases} \end{aligned} \quad (6)$$

Here, as the value of v^{-} increases, the values of fractions above become smaller (the assessment gets worse). In addition, this normalization ensures that when $P^{\succ} = \varnothing$, we have $PHI(P, \hat{z}, z^{dy}) \in [0, 1]$ as illustrated in Fig. 1(a). In this case, the value corresponds to the extent to which the desired region is covered, reaching a maximum of one when it is fully covered. The normalization also ensures that when $P^{\succ} \neq \varnothing$, as illustrated in Fig. 1(b), we have $PHI(P, \hat{z}, z^{dy}) \in (1, 2)$. Thus, any value greater than one indicates that the reference point is dominated by at least one of the solutions in P, which can be useful to know when dealing with a high number of objectives. In this case, the theoretical maximum of two is only possible when $\hat{z} = z^{dy}$ and $P = \{z^*\}$.

If an analyst wants to study a method's behavior further, PHI can provide more information. For instance, the analyst can get an overview of the proportion of the desired region covered by analyzing v^{+} values. The v^{-} values can also be helpful on how well the method incorporates the DM's preferences.

B. Relation to General Properties

In Section II, we listed desirable properties of indicators for assessing interactive methods. As the value of v^{+} increases, the PHI value gets better, and since PHI is a hypervolume indicator-based indicator, both local convergence

TABLE II
GPs for Indicators Designed for Interactive Methods.
The Symbol ✓ Means That the Indicator Possess
That GP and ✗ Means It Does not

Properties	R-metric	EH-metric	PMOD	PMDA	PHI
$GP1$	✓	✓	✓	✓	✓
$GP2$	✓	✓	✓	✓	✓
$GP3$	✓	✓	✓	✓	✓
$GP4$	✓	✓	✓	✓	✓
$GP5$	✗	✗	✗	✗	✗
$GP6$	✗	✓	✓	✓	✗
$GP7$	✗	✗	✓	✓	✓
$GP8$	✗	✓	✗	✗	✓
$GP9$	✗	✗	✗	✗	✗

Fig. 2. Tracking the generations of an interactive method for one interaction.

(GP1) and local diversity (GP2) are covered by this indicator. Furthermore, PHI can handle many objectives and does not need the knowledge of the Pareto front (GP3) and (GP4).

However, PHI is based on one type of preference information (reference point) and cannot handle different types of preferences (GP5). Furthermore, the computational time complexity of hypervolume indicator calculation, which is almost linear in the size of the population for two and three objectives, increases exponentially with the number of objectives [21]. Therefore, PHI does cover (GP6) only for problems with a small number of objectives.

PHI is independent of other interactive methods being compared. This means that we can assess different interactive methods separately and compare them with each other (GP7). Moreover, PHI does not introduce any *new* parameters (GP8). Setting the dystopian point for hypervolume indicator can significantly impact the assessment. However, there have been many studies on how to set this point [21]. Analyzing the effect of the dystopian point is beyond the scope of this research.

In [7], the ability to assess the interactive solution process as a whole is the last GP for an indicator (GP9). Since the decision and learning phases have different characteristics, it would be challenging to design an indicator that can simultaneously assess both. Therefore, PHI does not cover (GP9).

There are no indicators in the literature designed specifically for comparing interactive methods. However, some studies have utilized indicators designed for a priori methods to assess the performance of interactive methods [17], [22]. Here, we compare the coverage of GPs by PHI and some of the indicators designed for a priori methods. Among these indicators, we have chosen R-metric [10], EH-metric [11], PMOD [12], and PMDA [12] since they possess the desirable properties mentioned in [6]. Note that we did not include UPCF because R-metric can be viewed as an improved version of it. Table II shows the different GPs that are possessed by PHI, R-metric, EH-metric, PMOD, and PMDA. Note that since IGD requires the knowledge of the approximated Pareto front, we only use the hypervolume indicator for R-metric. A more detailed description of these indicators can be found in the supplementary materials.

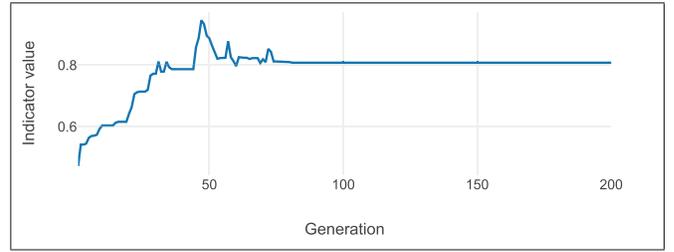In addition to the GPs discussed above, there are desirable properties in [7] for evaluating the learning and decision phases. We cannot use PHI to evaluate these phases as a stand-alone indicator. However, in what follows, we propose ways to utilize PHI to assess the performance of the two phases.

C. Relation to Learning Phase Properties

The list of desirable properties in [7] can be improved. Here, we redefine (LP2) as follows.

LP2: Assess the ability to adapt to new (even very different) preferences (responsiveness) and maintain the best-so-for solutions within interactions (stability).

To assess the performance in the learning phase, we propose the following steps. First, we assess the set P at each generation (t) of the interactive method using the PHI indicator. Then, we assess the learning phase as follows:

$$RS = \sum_{0}^{t_m} \text{PHI}\left(P_t, \hat{z}, z^{dy}\right) \tag{7}$$

where t_m is the number of generations in the whole learning phase, and RS measures responsiveness and stability of an interactive method. Note that since higher values of PHI indicate better performance, the higher the RS values, the better the responsiveness and stability.

Fig. 2 illustrates an example of how we can assess responsiveness and stability in the learning phase for one interaction. We track the PHI values for each generation. In this example, at around generation 48, the PHI values reach the maximum for this interaction. However, the method cannot maintain it, and this value decreases. Now, by using (7), we calculate the area under the curve in Fig. 2, which captures both responsiveness and stability.

D. Relation to Decision Phase Properties

The interaction where the DM chooses the most preferred solution should contribute more than other interactions in assessing the decision phase. Assume we have d interactions during the decision phase. Here, we define a hypervolume indicator-based comparison between the last reference point $\hat{z}^d$ and each reference point $\hat{z}^j$ ($j = 1, \ldots, d$) in the decision phase. Then, we assign a coefficient λ^j for the assessment of each interaction based on the similarity between $\hat{z}^j$ and $\hat{z}^d$. Therefore, the more similar the reference points are to the last reference point, the more contribution they have in the assessment of the decision phase.

We express $v^{d,j}$ as follows:

$$v^{d,j} = HV\left(\zeta_j, z^{dy}\right) - HV\left(\hat{z}^j, z^{dy}\right). \tag{8}$$
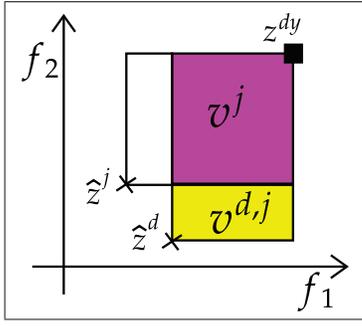
Fig. 3. Similarities of two desired regions corresponding to reference points $\hat{z}^j$ and $\hat{z}^d$ in a bi-objective example. The dystopian point is denoted by a black square.

Then, based on (8), we can express v^j as follows:

$$v^j = HV\left(\hat{z}^d, z^{dy}\right) - v^{d,j}.$$

Next, we can calculate λ^j as follows:

$$\lambda^j = \frac{v^j}{HV\left(\hat{z}^d\right)}. \tag{9}$$

Thus far, we have determined how similar reference points are to $\hat{z}^d$ by calculating λ^j. Now, to assess the decision phase, we propose the following:

$$FD = \frac{\sum_{j=1}^d \lambda^j \mathrm{PHI}\left(P_j, \hat{z}, z^{dy}\right)}{d} \tag{10}$$

where *FD* measures the ability of the method to fine-tune solutions. Similar to *RS*, the higher the value of *FD*, the better is the performance. Figure 3 illustrates a simple example of how we calculate the similarities between two reference points.

It is worth mentioning that we assume the performance assessment process happens after the DM is done with the solution process. The reason for such an assumption is that after the solution process, we have a good estimation of the nadir point, and we can set the dystopian point with more confidence. However, if the nadir point is known, we can assess the performance even during the solution process.

E. Advantages and Limitation of PHI

The main advantages of PHI can be summarized as follows.
1) PHI does not introduce any new parameters whose effect is not already known in the literature.
2) PHI can provide additional information about the search process to the analyst through v^+ and v^- values.
3) PHI can identify if the reference point is dominated and reflect this information in the final assessment [if $\mathrm{PHI}(P, \hat{z}, z^{dy}) > 1$].

A minor limitation of PHI is that under certain special conditions (illustrated in the supplementary materials), improving $\mathrm{PHI}(P, \hat{z}, z^{dy})$ can worsen the value of v^+. This is due to the points outside of the desired region with a positive contribution to v^+. Even commonly used indicators, such as the hypervolume indicator, have certain limitations (e.g., [21] and [23]).

IV. CASE STUDY

In this section, we demonstrate how we can assess the performance of interactive methods with an engineering problem from [24] (RE21). The goal of the problem is to minimize the structural volume f_1 and the joint displacement f_2 of a four-bar truss. For more details, see [25] and the supplementary materials in [24]. We have chosen this bi-objective problem because the visualization of solutions is easy, and we can illustrate the desired region conveniently. We provide more tests on different benchmarks and engineering problems using an ADM [18] in the Supplementary materials.

Among the many interactive methods, we have chosen two for illustration, namely, interactive optimization using preference incorporated space (IOPIS) [26], and interactive reference vector-guided evolutionary algorithm (iRVEA) [27]. As we mentioned earlier, there are no indicators designed explicitly for comparing interactive methods. However, we can utilize indicators for a priori methods the same way we utilized PHI for assessing interactive methods in the learning and decision phases.

To set the parameters of the interactive methods and indicators, we follow the suggestions in the original papers. Moreover, we used the nadir point that is provided in [24] and added a small value to it ($\epsilon_i = 0.2$) to calculate the dystopian point. Moreover, according to [28], there is no exact way to set the number of function evaluations. Therefore, we set a limit of 10 000 function evaluations for both methods. In addition, to have a fair comparison, we provide the same number of generations per interaction. Here, we set it as 300. Next, we need to set the number of interactions in the learning and decision phase. There is no clear way to set these numbers a priori, as they are based on when the DM feels (s)he is ready to move to the decision phase or when to select the most preferred solution. Here, for the sake of comparison, we choose to have three interactions in both the learning and decision phases. Note that studying the effect of these parameters on the performance of the interactive methods is beyond the scope of this research, and here, we are interested in comparing these two methods under the same conditions.

A. Solution Process of RE21

We solve the problem with IOPIS and iRVEA and use different indicators to assess their performances. In the visualizations of the solution process, the reference point is denoted by a black cross, the desired region by a black dashed box, the ideal point by a purple star, the dystopian point by a green square, and the solutions generated by IOPIS by blue dots, and the solutions generated by iRVEA by orange diamonds.

1) Learning Phase: In this phase, the DM wants to learn more about the reachability of solutions on the Pareto front. Additionally, the DM is interested to see if the generated solutions reflect his preferences or not.

a) Interaction 1: Here, the DM provides the reference point $\hat{z}_1 = (1238, 0.0300)$ giving more importance to the first objective than the second one. Fig. 4(a) shows the solutions generated by IOPIS and iRVEA. We can see that IOPIS solutions are all inside the desired region. On the other hand,
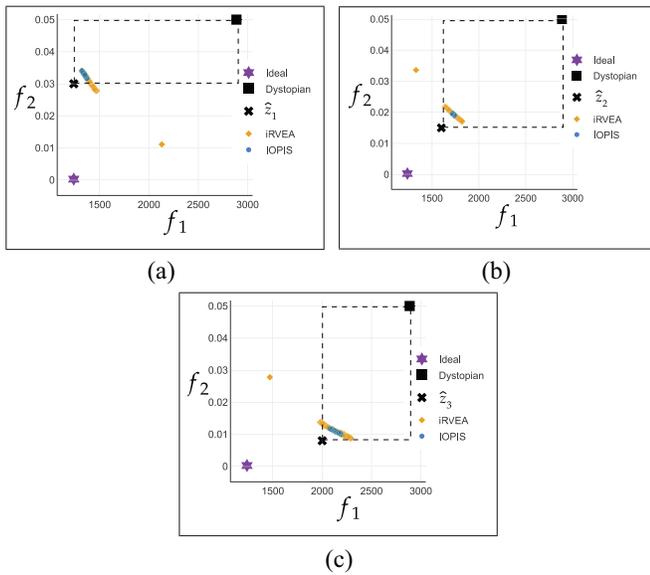
Fig. 4. Results of interactions 1, 2, and 3 in the learning phase for problem RE21. (a) Interaction 1. (b) Interaction 2. (c) Interaction 3.
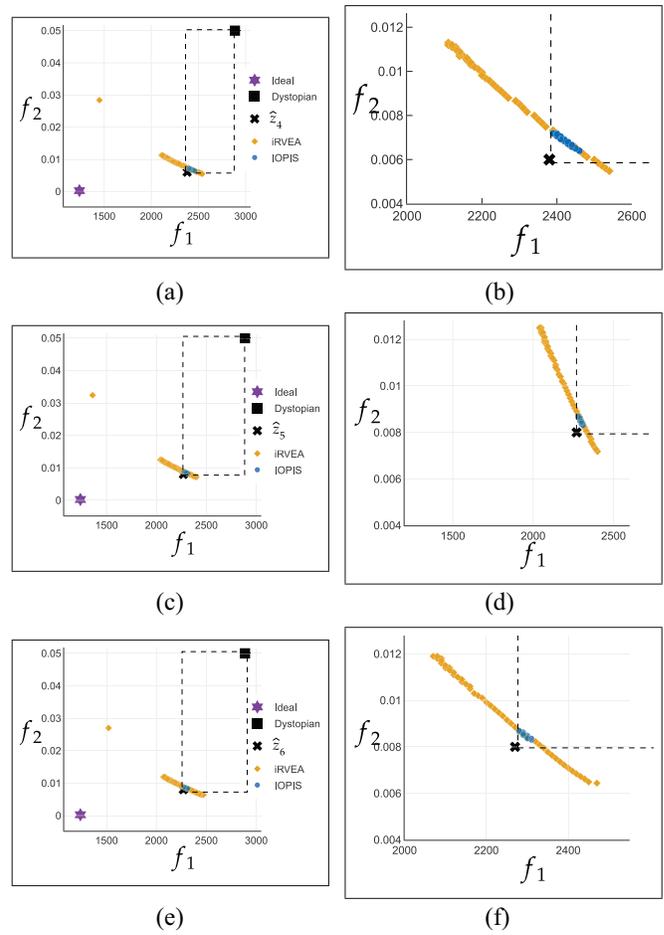

Fig. 5. Results of interactions 4, 5, and 6 in the decision phase for problem RE21. The ideal point is denoted by a purple star and the dystopian point by a green square. (a) Interaction 4. (b) Zoomed-in interaction 4. (c) Interaction 5. (d) Zoomed-in interaction 5. (e) Interaction 6. (f) Zoomed-in interaction 6.

iRVEA has generated some solutions outside of this region, among them one solution that is very far away from the desired region.

b) Interaction 2: Next, the DM moves to the middle of the objective space by providing the second reference point as $\hat{z}_2 = (1600, 0.0150)$. Fig. 4(b) shows the results of the two methods. If we only consider the solutions inside the desired region, we can observe that iRVEA has better local diversity than IOPIS. However, iRVEA still has one solution located outside of the desired region. On the other hand, IOPIS has generated solutions inside the desired region.

c) Interaction 3: Here, the DM provides the third reference point $\hat{z}_3 = (2000, 0.008)$ to learn about the tradeoffs when he wants to give more importance to the second objective. As illustrated in Fig. 4(c), again, iRVEA has one solution that is outside of the desired region, but IOPIS has generated all solutions inside this region.

By now, the DM is confident about the tradeoffs between the objectives. Therefore, he is ready to move to the decision phase.

2) Decision Phase: At the end of the learning phase, the DM has identified his region of interest and tries to fine-tune the solutions within this region, that is, around the solutions corresponding to the last reference point $\hat{z}_3$. Since the changes in the reference points are quite small, we provide *zoomed-in* illustrations of the results in Fig. 5(b), (d), and (f) for different interactions, respectively.

a) Interaction 4: As for the fourth interaction (first in the decision phase), the DM wants to provide a reachable reference point since now he has an idea where the Pareto front lies. Therefore, he provides the fourth reference point as $\hat{z}_4 = (2380, 0.006)$. Fig. 5(a) illustrates the solutions. Here, iRVEA generated solutions outside of the desired region, but IOPIS within this region.

b) Interaction 5: Next, the DM decides to increase the aspiration level for the second objective and expects to see

lower values for the first objective with $\hat{z}_5 = (2270, 0.008)$. Fig. 5(c) shows the results. Again, the solutions have the same behavior as in the previous interaction, where IOPIS sticks to the boundaries of the desired region, but iRVEA generates some solutions outside of it.

c) Interaction 6: Based on the solutions shown in Fig. 5(c), the DM decides to provide a reference point between $\hat{z}_4$ and $\hat{z}_5$ as $\hat{z}_6 = (2310, 0.007)$. Fig. 5(e) illustrates the solutions. At this point, the DM wants to stop the solution process with $(0.008, 2280)$ as the most preferred solution. In fact, both iRVEA and IOPIS were able to reach almost the same solution. Now, we are interested in comparing the two interactive methods.

B. Assessing the Learning and Decision Phases

In Section III, we proposed *RS* (7) for the learning phase and *FD* (10) for the decision phase. To utilize other indicators, we only have to replace the function PHI(.) with the one that calculates the indicator of our choice. For *RS*, we calculate the indicator values for every generation of the interactive methods in the learning phase (interactions 1, 2, and 3). We can assess the learning phase by using (7), which is equivalent to
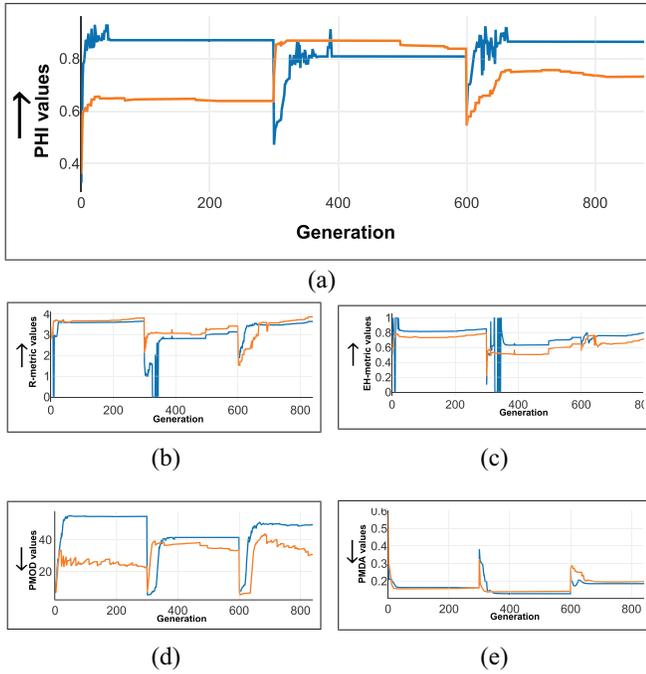
(a)

(b)

(c)

(d)

(e)

Fig. 6. Indicator values for iRVEA (the orange line) and IOPIS (the blue line) for each generation during the learning phase. The arrow ↑ indicates that higher values are better, and the arrow ↓ indicates that lower values are better. (a) PHI values for each generation in the learning phase. (b) R-metric values for each generation in the learning phase. (c) EH-metric values for each generation in the learning phase. (d) PMOD values for each generation in the learning phase. (e) PMDA values for each generation in the learning phase.

TABLE III
VALUES OF *RS* AND *FD* FOR THE LEARNING AND DECISION PHASES BY USING DIFFERENT INDICATORS. THE ARROW ↑ INDICATES THAT HIGHER VALUES ARE BETTER, AND THE ARROW ↓ INDICATES THAT LOWER VALUES ARE BETTER

Indicators	iRVEA		IOPIS	
	RS	FD	RS	FD
PHI ↑	645.54	0.48	**757.57**	**0.92**
R-metric ↑	2804.17	**4.04**	**2957.77**	3.75
EH-metric ↑	599.68	**0.79**	**682.06**	0.77
PMOD ↓	**23864.70**	**25.51**	40995.447	187.18
PMDA ↓	177.90	**0.21**	**146.98**	0.25

(a)

(b)

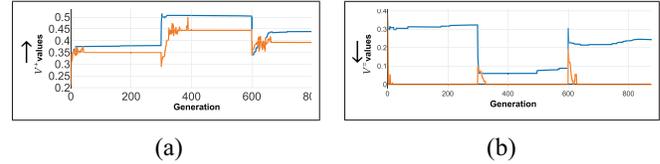Fig. 7. Positive (v^+) and negative (v^-) hypervolume indicator contributions that the solutions had to the desired region. The values of IOPIS are denoted by the blue line, and the values of iRVEA by the orange line. (a) Values of v^+ in the learning phase. (b) Values of v^- in the learning phase.

calculating the area under the curves in Fig. 6. In Fig. 6, the orange line represents the indicator values for iRVEA, and the blue line represents the same for IOPIS.

For PHI values in Fig. 6(a), we can observe that in the first interaction (the first 300 generations), IOPIS has a better performance than iRVEA. In the second interaction (generations 300–600), iRVEA has a better performance for magiority of generations. However, in generation 385, IOPIS obtain a better performance than iRVEA, but it cannot maintain it. In addition, toward the end of the second interaction, iRVEA's performance gets worse. Finally, for the third interaction (generations 600–900), IOPIS retakes the lead.

Moreover, we can see in Fig. 6 that not all other indicators show the same behavior. For example, for R-metric in Fig. 6(b), throughout the majority of generations in all three interactions, IOPIS has a slightly better performance than iRVEA. Another example is PMOD in Fig. 6(d), where except for the beginning of the second interaction, iRVEA has a better performance than IOPIS. In contrast to PMOD, the EH-metric in Fig. 6(c) declares IOPIS as the superior method for all three reference points. Finally, for PMDA, we can observe that for most of generations both methods have very similar performance except for the beginning of second and third interactions.

For *FD*, we calculate λ_4, λ_5, and λ_6. Based on (9), $\lambda_4 = 0.876$, $\lambda_5 = 0.976$, $\lambda_6 = 1$. Then, we get *FD* using (10). See Table III for the *FD* results.

In Table III, we see that the assessments based on R-metric, EH-metric, PMOD, and PMDA declare iRVEA as the winner.

On the other hand, our proposed indicator declares IOPIS as the winner. Fig. 5 shows the solutions in the three interactions of the decision phase. Here, all IOPIS solutions are inside the desired region. However, iRVEA has generated many solutions in all interactions outside of the desired region.

C. Choosing the Preferred Method

After the solution process, the analyst can use the indicators to choose the best method. Here, not all the indicators agree on which method is the winner, and it would be hard to choose the interactive method to be used by the DM. However, except for PHI, other indicators cannot provide easy-to-digest information, which is understandable since they are not designed to assess interactive methods. In PHI, we can easily show the assessment of v^+ (5) and v^- (2) to understand how the interactive method is behaving.

For *RS*, we mentioned that PMOD's assessment differed from other indicators by declaring iRVEA as the winner. In Fig. 7(a), we can observe that for most of the generations, iRVEA has a better $v^=$ values than IOPIS. However, v^- values [see Fig. 7(b)] illustrate that IOPIS converged toward zero quite fast, and at the end of each interaction, none of the solutions generated by IOPIS are outside of the desired region, but this is not the case for iRVEA. This means that iRVEA performs better than IOPIS within the desired region. However, iRVEA generates some solutions that do not reflect the DM's preferences, and all of the solutions generated by IOPIS reflect it (they are all inside the desired region). After analyzing Fig. 7, the analyst feels more comfortable that IOPIS is the better interactive method for the learning phase for RE21.

Obviously, in our case, we could easily see if solutions were inside or outside of the desired region by visualizing the scatter plots. However, in higher-dimensional problems, Fig. 7 can provide valuable insight about the positioning of the solutions.

As for the *FD*, except for PHI, all other indicators declared iRVEA as the winner. We provide information about v^+ and v^-

TABLE IV
v^+ AND v^- VALUES FOR INTERACTION OF iRVEA AND IOPIS IN THE
DECISION PHASE. THE HIGHER THE v^+ VALUE, THE BETTER,
AND THE LOWER THE v^- VALUE, THE BETTER

Interaction	Measure	IOPIS	iRVEA
4	v^+	**0.270**	0.261
	v^-	**0.000**	0.528
5	v^+	0.308	**0.315**
	v^-	**0.000**	0.453
6	v^+	0.291	**0.301**
	v^-	**0.000**	0.471

for each interaction of IOPIS and iRVEA in Table IV. Here, we see that if we only consider solutions inside the desired region, iRVEA had a slightly better performance than IOPIS in the last two interactions, and it would be in line with the results of other indicators. However, IOPIS did not generate any solution outside of the desired region ($v^- = 0$) for all interactions in the decision phase, which is not the case for iRVEA. In fact, we can observe that the contribution of solutions to the hypervolume indicator outside of the desired region is bigger than the contribution to the hypervolume indicator inside it ($v^- > v^+$) for all the interactions of iRVEA in the decision phase. After reviewing these results, the analyst feels that the solutions generated by iRVEA were not reflecting the DM's preferences. Therefore, he chose IOPIS to be best for the decision phase. As a result, he chose IOPIS as the preferred method for this problem's actual interactive solution process because IOPIS was better in both phases.

V. CONCLUSION

In this article, we proposed the first indicator that is explicitly designed for assessing interactive evolutionary multiobjective optimization methods. We also showed how we could utilize this indicator to assess the learning and decision phases with interactive methods. Next, we compared the performance of two interactive methods on an engineering problem and demonstrated how our indicator could help the analyst to choose a suitable method for solving a problem.

The proposed indicator uses DM's reference point to identify the desired region. Then, it divides the hypervolume indicator contribution of solutions into positive and negative contributions. Finally, it uses the positive hypervolume indicator contribution as the main performance assessment and uses the negative hypervolume indicator contribution to penalize the performance of interactive methods. Our proposed indicator covers many of the desirable properties that have been identified for indicators designed for interactive methods. Besides, one of the advantages of this indicator is that all of the calculations are based on the hypervolume indicator, which is a known concept, and it is easy to understand.

Furthermore, we showed that our indicator PHI could provide additional information to the analyst to further analyze the methods being compared. For instance, by only looking at the final performance assessment, one can understand if there exists any solution that dominates the reference point or not. In addition, the indicator can provide information on solutions that are outside of the desired region as well as solutions that are inside it.

There are still some desirable properties that PHI does not cover. As for our next step, we focus on developing an indicator that considers such desirable properties. For example, because we need to calculate the hypervolume indicator multiple times to be able to use our indicator, it makes PHI computationally expensive, and it may not be practical to use it for problems with a high number of objectives, especially during the performance assessment of the learning phase. Developing an indicator that is not computationally expensive and indicators that can consider other types of preferences are future research directions. Moreover, developing an indicator-based interactive evolutionary method (based on PHI) can be considered in the future. In addition, as shown in the supplementary materials, the dystopian point can affect the value of PHI, and coming up with a systematic way of setting it would be an important future research direction. Finally, we assumed that DM starts with the learning phase and moves on to the decision phase. However, in practice, the DM may move back and forth between these two phases. Identifying these phases during interactions with the DM could itself be an interesting line of research.

REFERENCES

[1] K. Miettinen, *Nonlinear Multiobjective Optimization*. Boston, MA, USA: Kluwer Acad. Publ., 1999.
[2] K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms*. Chichester, U.K.: Wiley, 2001.
[3] Y. Jin, H. Wang, and S. Chaoli, *Data-Driven Evolutionary Optimization*. Cham, Switzerland: Springer, 2021.
[4] K. Miettinen, F. Ruiz, and A. P. Wierzbicki, "Introduction to multiobjective optimization: Interactive approaches," in *Multiobjective Optimization: Interactive and Evolutionary Approaches*, J. Branke, K. Deb, K. Miettinen, and R. Slowinski, Eds. Berlin, Germany: Springer, 2008, pp. 27–57.
[5] E. Zitzler, K. Deb, and L. Thiele, "Comparison of multiobjective evolutionary algorithms: Empirical results," *Evol. Comput.*, vol. 8, no. 2, pp. 173–195, Jun. 2000.
[6] A. Mohammadi, M. N. Omidvar, and X. Li, "A new performance metric for user-preference based multi-objective evolutionary algorithms," in *Proc. IEEE Congr. Evol. Comput.*, 2013, pp. 2825–2832.
[7] P. A. Pour, S. Bandaru, B. Afsar, and K. Miettinen, "Desirable properties of performance indicators for assessing interactive evolutionary multiobjective optimization methods," in *Proc. Genet. Evol. Comput. Conf. Companion*, 2022, pp. 1803–1811.
[8] N. Riquelme, C. Von Lücken, and B. Baran, "Performance metrics in multi-objective optimization," in *Proc. Latin Amer. Comput. Conf.*, 2015, pp. 1–11.
[9] C. Audet, J. Bigeon, D. Cartier, S. Le Digabel, and L. Salomon, "Performance indicators in multiobjective optimization," *Eur. J. Oper. Res.*, vol. 292, no. 2, pp. 397–422, 2021.
[10] K. Li, K. Deb, and X. Yao, "R-metric: Evaluating the performance of preference-based evolutionary multiobjective optimization using reference points," *IEEE Trans. Evol. Comput.*, vol. 22, no. 6, pp. 821–835, Dec. 2018.
[11] S. Bandaru and H. Smedberg, "A parameterless performance metric for reference-point based multi-objective evolutionary algorithms," in *Proc. Genet. Evol. Comput. Conf.*, 2019, pp. 499–506.
[12] Z. Hou, S. Yang, J. Zou, J. Zheng, G. Yu, and G. Ruan, "A performance indicator for reference-point-based multiobjective evolutionary optimization," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, 2018, pp. 1571–1578.
[13] G. Yu, J. Zheng, and X. Li, "An improved performance metric for multiobjective evolutionary algorithms with user preferences," in *Proc. IEEE Congr. Evol. Comput.*, 2015, pp. 908–915.
[14] C. S. Y. Wong, A. Al-Dujaili, and S. Sundaram, "Hypervolume-based DIRECT for multi-objective optimisation," in *Proc. Genet. Evol. Comput. Conf. Companion*, 2016, pp. 1201–1208.

[15] B. Afsar, A. B. Ruiz, and K. Miettinen, "Comparing interactive evolutionary multiobjective optimization methods with an artificial decision maker," *Complex Intell. Syst.*, vol. 9, pp. 1165–1181, Apr. 2023.

[16] B. Afsar, K. Miettinen, and A. B. Ruiz, "An artificial decision maker for comparing reference point based interactive evolutionary multiobjective optimization methods," in *Proc. 11th Int. Conf. Evol. Multi-Criterion Optim. (EMO)*, 2021, pp. 619–631.

[17] P. A. Pour, T. Rodemann, J. Hakanen, and K. Miettinen, "Surrogate assisted interactive multiobjective optimization in energy system design of buildings," *Optim. Eng.*, vol. 23, pp. 303–327, Mar. 2022.

[18] B. Afsar, K. Miettinen, and F. Ruiz, "Assessing the performance of interactive multiobjective optimization methods: A survey," *ACM Comput. Surv.*, vol. 54, no. 4, pp. 1–27, 2021.

[19] R. E. Steuer, *Multiple Criteria Optimization: Theory, Computation, and Applications*. New York, NY, USA: Wiley, 1986.

[20] K. Miettinen, J. Hakanen, and D. Podkopaev, "Interactive nonlinear multiobjective optimization methods," in *Multiple Criteria Decision Analysis: State of the Art Surveys*, S. Greco, M. Ehrgott, and J. R. Figueira, Eds. New York, NY, USA: Springer, 2016, pp. 927–976.

[21] K. Shang, H. Ishibuchi, L. He, and L. M. Pang, "A survey on the hypervolume indicator in evolutionary multiobjective optimization," *IEEE Trans. Evol. Comput.*, vol. 25, no. 1, pp. 1–20, Feb. 2021.

[22] G. Lárraga and K. Miettinen, "Interactive MOEA/D with multiple types of preference information," in *Proc. Genet. Evol. Comput. Conf. Companion*, 2022, pp. 1826–1834.

[23] H. Ishibuchi, L. M. Pang, and K. Shang, "Difficulties in fair performance comparison of multi-objective evolutionary algorithms [research frontier]," *IEEE Comput. Intell. Mag.*, vol. 17, no. 1, pp. 86–101, Feb. 2022.

[24] R. Tanabe and H. Ishibuchi, "An easy-to-use real-world multi-objective optimization problem suite," *Appl. Soft Comput.*, vol. 89, Apr. 2020, Art. no. 106078.

[25] F. Y. Cheng and X. S. Li, "Generalized center method for multiobjective engineering optimization," *Eng. Optim.*, vol. 31, no. 5, pp. 641–661, 1999.

[26] B. S. Saini, J. Hakanen, and K. Miettinen, "A new paradigm in interactive evolutionary multiobjective optimization," in *Proc. 16th Int. Conf. Parallel Problem Solving From Nat. (PPSN)*, 2020, pp. 243–256.

[27] J. Hakanen, T. Chugh, K. Sindhya, Y. Jin, and K. Miettinen, "Connections of reference vectors and different types of preference information in interactive multiobjective evolutionary algorithms," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, 2016, pp. 1–8.

[28] A. Kazikova, M. Pluhacek, and R. Senkerik, "How does the number of objective function evaluations impact our understanding of metaheuristics behavior?" *IEEE Access*, vol. 9, pp. 44032–44048, 2021.

Pouya Aghaei Pour received the Ph.D. degree in applied mathematics and computational sciences from the University of Jyväskylä, Jyväskylä, Finland, in 2022.

He is currently working as a Guest Researcher with the University of Jyväskylä. His research interests include multiobjective optimization, machine learning, and addressing computationally expensive problems. He is passionate about developing innovative solutions to complex challenges and contributing to the advancement of knowledge in his areas of expertise.

Sunith Bandaru (Senior Member, IEEE) received the bachelor's degree in mechanical engineering from Jawaharlal Nehru Technological University, Hyderabad, India, 2006, and the master's and Ph.D. degrees from the Indian Institute of Technology Kanpur, Kanpur, India, 2008 and 2013, respectively.

He is an Associate Professor with the Division of Intelligent Production Systems, School of Engineering Science, and the Chair of the Committee for Research Education in Informatics, University of Skövde, Skövde, Sweden. His research interests lie at the intersection of evolutionary computation, machine learning and industrial data analytics, where he has been developing algorithms, methods and tools for knowledge discovery from simulation-based multiobjective industrial optimization problems, knowledge-driven optimization, and intelligent decision support.

Bekir Afsar received the Ph.D. degree in computer engineering from Ege University, Izmir, Turkey, in 2014.

He is a Postdoctoral Researcher in a HORIZON EU project, Biodiversity Digital Twin for Advanced Modelling, Simulation and Prediction Capabilities (BioDT—biodt.eu) with the University of Jyväskylä, Jyväskylä, Finland. His main research interests are multiobjective optimization, data-driven decision making, evolutionary computation, and interactive multiobjective optimization methods and their applications. He is currently focusing on digital twin-enabled multiobjective optimization, where he supports real decision makers and stakeholders in finding the most appropriate forest management strategy considering their preferences as well as the ecological, social, and economic objectives under different climate change scenarios.

Michael Emmerich was born in Coesfeld, Germany. He received the Doctorate of Natural Sciences degree from the Technical University Dortmund, Dortmund, Germany, in 2006.

He is a Dutch-Finnish Researcher of Multiobjective Optimization. He holds the position of an Associate Professor of Multicriteria Optimization and Decision Analysis with Leiden University, Leiden, The Netherlands, and is a Visiting Researcher with the Multiobjective Optimization Group, University of Jyväskylä, Jyväskylä, Finland. Moreover, he has a part-time position as a Lead AI Scientist with the nordic AI company Silo.ai. He has worked in the past at several institutions, including the University of the Algarve, Faro, Portugal; IST Lisbon, Lisbon, Portugal; RWTH Aachen, Aachen, Germany; ACCESS e.V. Aachen, Aachen, Germany; and FOM Amolf Amsterdam, Amsterdam, The Netherlands. His main contributions are in the geometrical and algorithmical analysis of performance indicators in multiobjective optimization, Bayesian multiobjective optimization, navigation methods for many-objective optimization, big-data Gaussian processes, and applications in chemistry, chemical industry, building design, and operational research.

Kaisa Miettinen received the Ph.D. degree in mathematical information technology from the University of Jyväskylä (JYU), Jyväskylä, Finland, in 1994.

She is a Professor of Industrial Optimization with JYU, where she heads the Multiobjective Optimization Group and is the Director of the thematic research Area "Decision Analytics utilizing Causal Models and Multiobjective Optimization." With her group, she develops open-source software. She has worked with IIASA, Laxenburg, Austria; the KTH Royal Institute of Technology, Stockholm, Sweden; and the Helsinki School of Economics, Espoo, Finland. She has authored over 210 refereed journal, proceedings, and collection papers, a monograph on Nonlinear Multiobjective Optimization, and edited 20 proceedings, collections, and special issues. Her research interests include methods and applications of multiobjective optimization, in particular interactive and evolutionary approaches and their hybrids.

Prof. Miettinen has received the Georg Cantor Award of the International Society on Multiple Criteria Decision Making (IS-MCDM). She belongs to editorial boards of seven journals and the steering committee of Evolutionary Multi-Criterion Optimization. She is a member of the Finnish Academy of Science and Letters and has been the President of IS-MCDM.