



OPEN ACCESS

EDITED BY

Arnab Mukherjee,
Indian Institute of Science Education and
Research, India

REVIEWED BY

Gourhari Jana,
Michigan State University, United States
Kun Yao,
Schrodinger, United States

*CORRESPONDENCE

Laurent E. Dardenne,
✉ dardenne@lncc.br

[†]These authors have contributed equally
to this work and share first authorship

RECEIVED 04 September 2023

ACCEPTED 27 November 2023

PUBLISHED 18 December 2023

CITATION

Angelo JS, Guedes IA, Barbosa HJC and
Dardenne LE (2023), Multi- and many-
objective optimization: present and
future in *de novo* drug design.
Front. Chem. 11:1288626.
doi: 10.3389/fchem.2023.1288626

COPYRIGHT

© 2023 Angelo, Guedes, Barbosa and
Dardenne. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](#). The use, distribution or
reproduction in other forums is
permitted, provided the original author(s)
and the copyright owner(s) are credited
and that the original publication in this
journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Multi- and many-objective optimization: present and future in *de novo* drug design

Jaqueline S. Angelo[†], Isabella A. Guedes[†], Helio J. C. Barbosa and
Laurent E. Dardenne*

Coordenação de Modelagem Computacional, Laboratório Nacional de Computação Científica,
Petrópolis, Brazil

de novo Drug Design (dnDD) aims to create new molecules that satisfy multiple conflicting objectives. Since several desired properties can be considered in the optimization process, dnDD is naturally categorized as a many-objective optimization problem (ManyOOP), where more than three objectives must be simultaneously optimized. However, a large number of objectives typically pose several challenges that affect the choice and the design of optimization methodologies. Herein, we cover the application of multi- and many-objective optimization methods, particularly those based on Evolutionary Computation and Machine Learning techniques, to enlighten their potential application in dnDD. Additionally, we comprehensively analyze how molecular properties used in the optimization process are applied as either objectives or constraints to the problem. Finally, we discuss future research in many-objective optimization for dnDD, highlighting two important possible impacts: i) its integration with the development of multi-target approaches to accelerate the discovery of innovative and more efficacious drug therapies and ii) its role as a catalyst for new developments in more fundamental and general methodological frameworks in the field.

KEYWORDS

drug discovery, *de novo* drug design, evolutionary algorithms, multi-objective optimization, many-objective optimization

1 Introduction

Life involves choices, reaching decisions, and seeking compromises. The major challenge lies in managing the conflict between the various goals and objectives [Miettinen \(1999\)](#). As in many real-world problems, the discovery of a new drug with desired pharmacological and pharmacokinetic properties has several objectives to be considered. For instance, in the search for new therapeutic drugs, the maximization of (i) the potency of the drug, (ii) the structural novelty, (iii) pharmacokinetic profile, the minimization of (iv) synthesis costs, and (v) unwanted side effects are desired goals to be optimized [Rosenthal and Borschbach \(2017\)](#); [Lambrinidis and Tsantili-Kakoulidou \(2021\)](#). Thus, designing new effective and safe drugs is inherently a problem with diverse objectives to be optimized concurrently.

Many computational tools have been developed to assist in the design of novel drug-like molecules, such as quantitative structure-activity relationship (QSAR) [Jana et al. \(2020\)](#); [Wang Z. et al. \(2021\)](#); [Socha et al. \(2023\)](#), molecular docking and affinity prediction through machine learning-based scoring functions [Guedes et al. \(2014\)](#) [Santos et al. \(2020\)](#); [Guedes et al. \(2021\)](#). Flurbiprofen, vaborbactam, and atazanavir are commercially approved drugs

discovered by computer-aided drug design Sabe et al. (2021). The term “*de novo*” in Latin means “anew”, “afresh”, or “from the beginning”. The goal of *de novo* drug design (dnDD) is to create novel molecules with desirable properties from scratch. In this context, multiple properties mean multiple objectives to be optimized.

The problem of optimization refers to the task of discovering feasible solutions until no better solution can be found. The quality of a solution is evaluated based on an objective, while the feasible region represents a set of conditions or constraints that limit the solutions to the problem. In dnDD, an objective can be expressed by various properties of interest, such as a similarity score to a known ligand or a binding score with a target receptor Nicolaou et al. (2009). At the same time, the constraints may be any useful function, such as chemical stability and synthetic feasibility Nicolaou et al. (2012).

For a long time, researchers in dnDD neglected the presence of multiple conflicting objectives Nicolaou et al. (2009), such as simultaneously maximizing the potency of a drug to a specific target and minimizing known side effects, which are naturally present in this type of problem. Due to the complexity of designing a new molecule, a single objective may not cover the multi-faceted design issues. Thus, researchers seek new techniques or design strategies that simultaneously consider the multiple aspects of this class of problems.

While in a single-objective optimization problem (SingleOOP), the goal is to optimize only one objective function, in a multi-objective optimization problem (MultiOOP), more than one objective must be simultaneously optimized. If the objective functions are not conflicting, a solution can be found where each objective reaches its optimum value. However, in MultiOOPs, those objectives are frequently conflicting (*i.e.*, the improvement of one objective leads to the degradation of another objective), and also non-commensurable (*i.e.*, when dealing with objectives that have different units or scales of measurement). In this case, there is usually no optimal solution but a set of trade-off solutions representing a compromise between the conflicting objectives. Such solutions are called *non-dominated* solutions, which form the *Pareto (optimal) set*, which is the set of solutions that are all equally optimal concerning the considered objectives.

An automatic dnDD can accelerate the overall drug discovery process but can be complex and computationally demanding. Since LEGEND, the first dnDD technique proposed in 1991 Nishibata and Itai (1991), numerous other methods have been developed to assist researchers in drug discovery. In particular, Evolutionary Algorithms (EAs) have been widely used to find the optimal solution(s) in *de novo design* Devi et al. (2015); Le and Winkler (2015). EAs are population-based metaheuristics in which a collection of candidate solutions evolve under specified selection rules to a state that minimizes/maximizes a general cost function. In contrast to classical search methods that usually aggregate the objective functions into one objective, EAs can be easily applied to multi-objective optimization problems due to their population-based nature, allowing EAs to find a set of non-dominated solutions in a single run. Such methods are called multi-objective EAs (MultiOEs).

Research in evolutionary computation has primarily focused on problems having two or three objectives Coello et al. (2007).

However, many real-world problems have several (more than three) objectives in their formulation, *e.g.*, car side-impact Deb et al. (2009), mechanical engineering problems Ursem and Justesen (2012), water resource management Asafuddoula et al. (2015), routing planning in agricultural mobile robots Zhang et al. (2022), wireless sensor network deployment Ben Amor et al. (2022), among others. Particularly, dnDD has intrinsically various objectives to optimize, clearly more than three Rosenthal and Borschbach (2017); Lambrinidis and Tsantili-Kakoulidou (2021).

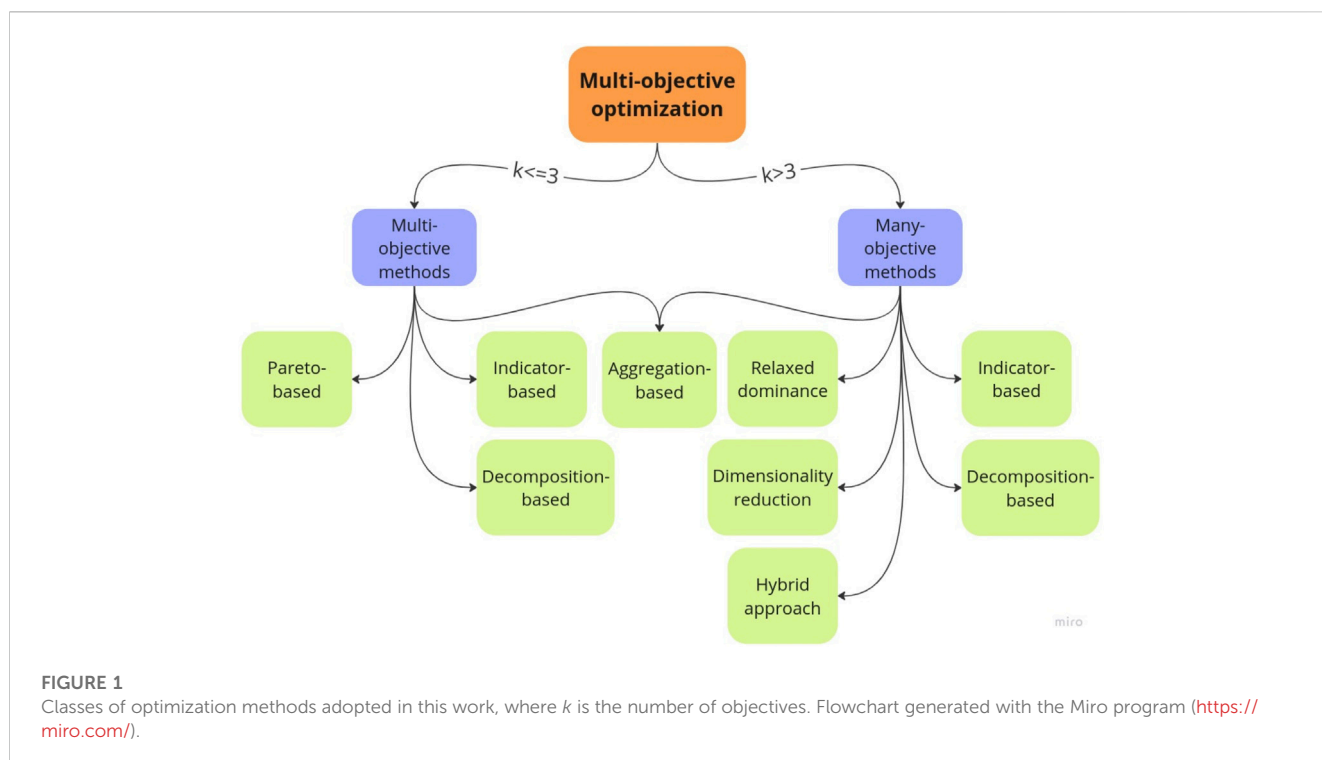
“Multi-objective” refers to scenarios involving three objective functions at most, while “many-objective” is usually adopted to specify problems with more than three objectives. The growing interest in the area of many-objective optimization (ManyOO) motivated the study and the design of new optimization techniques capable of solving problems with four to twenty or even more objective functions, as occurs in the nurse scheduling problem, with 25 objectives to be considered Burke et al. (2004); Süßflow et al. (2007). However, problems with many objectives present additional challenges compared to low-dimensional problems, *e.g.*, finding a good approximation of the Pareto set. Many-objective optimization in real-world applications has many decision design components that users commonly undertake Deb et al. (2023).

Over the past three decades, many-objective EAs (ManyOEs) have been the subject of extensive research and practical implementation in various real-world applications, making them a widely studied and applied field Safi et al. (2018); Sato and Ishibuchi (2023).

This paper aims to review and enlighten the potential application of ManyOO in dnDD, with particular emphasis on EAs. There have been a limited number of examples where ManyOEs are applied to dnDD and published in scientific literature. Thus, the main contributions of this paper are:

- introduce the definition of ManyOOPs and discuss the main challenges that impact the choice and the design of evolutionary techniques when the number of objectives of an optimization problem increases;
- review the different classifications of EAs for solving multi-objective and many-objective optimization problems;
- present and classify some existing MultiOEs for dnDD;
- enumerate and categorize several ManyOEs, from the most representative to the most recent state-of-the-art techniques, which have the potential to be applied in dnDD;
- present recent research that employs machine learning techniques, as they are emerging as a promising class of methods for multi-objective and possibly for many-objective dnDD;
- provide a comprehensive analysis of how the various molecular properties used in the optimization process are applied as either objectives or constraints to the problem.

As far as we know, this is the first review covering the application of ManyOO methodologies for the dnDD in addition to the MultiOO approaches. To organize and classify the different methodologies employed in these two main areas, “multi-objective” and “many-objective”, herein, we classify the optimization algorithms into two classes: *multi-objective methods*, in which problems are defined with two or three objectives, and



many-objective methods, where problems contain four or more objectives, as illustrated in Figure 1.

The paper is organized as follows. Section 2 provides basic concepts and definitions regarding multi-objective optimization. Section 3 presents solutions methods for solving MultiOOPs, including classical techniques and MultiOEAs. Section 4 discuss the main challenges faced when dealing with ManyOOPs. A classification of ManyOEAs is provided, presenting recent and representative methods for each approach. In Section 5, we describe applications in dnDD, focusing on the two classifications of methods proposed (multi-objective and many-objective), presenting not only techniques based on EAs but also based on Machine Learning (ML) methods. Recently, hybrid approaches have emerged, combining EAs with ML techniques to increase the potential of these two classes of methods. Additionally, we will discuss how researchers distinguish objectives and constraints among the many molecules' properties in dnDD. Perspectives concerning the ManyOO approach in dnDD are given in Section 6. The conclusions are presented in Section 7.

2 Key concepts in the multi-objective optimization problem

A MultiOOP can be written as

$$\begin{array}{ll} \text{Minimize/Maximize} & \mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_k(\mathbf{x}))^T \\ \text{subject to} & g_j(\mathbf{x}) \leq 0, \quad j = 1, 2, \dots, J; \\ & h_p(\mathbf{x}) = 0, \quad p = 1, 2, \dots, P; \\ & x_i^l \leq x_i \leq x_i^u, \quad i = 1, 2, \dots, n \end{array} \quad (1)$$

where we have k (≥ 2) objective functions $f_i: \mathbb{R}^n \rightarrow \mathbb{R}$ that must be minimised or maximized. For convenience, we will treat the

MultiOOP as a minimization problem; however, if an objective function $f_k(\mathbf{x})$ needs to be maximized, then it is equivalent to minimizing the function $-f_k(\mathbf{x})$. The vector of objective functions is denoted by $\mathbf{F}(\mathbf{x})$, where $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ is a solution (decision) vector with n variables. Associated with the problem, there are J inequality and P equality constraints, such that $g: \mathbb{R}^n \rightarrow \mathbb{R}^J$ and $h: \mathbb{R}^n \rightarrow \mathbb{R}^P$, respectively. The last set of constraints in (1) is the bound constraints that restrict each variable of the decision vector \mathbf{x} to a value within a lower x_i^l and an upper x_i^u bound, with $i = 1, \dots, n$. A feasible solution is any solution \mathbf{x} that satisfies all constraints, and the set of all feasible solutions constitutes the feasible (set) region $S \subset \mathbb{R}^n$, which is a subset of the *decision space*.

The notable difference between single-objective and multi-objective optimization is that in MultiOOP, the objective functions constitute a multi-dimensional space called the *objective space* $Z \subset \mathbb{R}^k$. The elements of Z correspond to the objective vector $\mathbf{z} = \mathbf{F}(\mathbf{x}) = (z_1, z_2, \dots, z_k)^T$, where $z_i = f_i, \forall i = 1, \dots, k$. For each solution vector \mathbf{x} in the decision space, there is an objective vector \mathbf{z} in the objective space. Such mapping relates an n -dimensional solution vector (in the decision space) to a k -dimensional objective vector (in the objective space). An illustrative example of these two spaces is shown in Figure 2. For clarity, suppose that A , B , and C are three different compounds and f_1 and f_2 are affinity predictions given in pK_i of each molecule against two different targets (e.g., one therapeutic protein and one protein considered as an off-target related to side effects). Our objectives are to maximize the affinity value of these compounds against "target 1" and to minimize the affinity against "target 2". Hence, for each solution in the decision space (e.g., represented by the SMILES of a particular molecule), there is a corresponding point in the

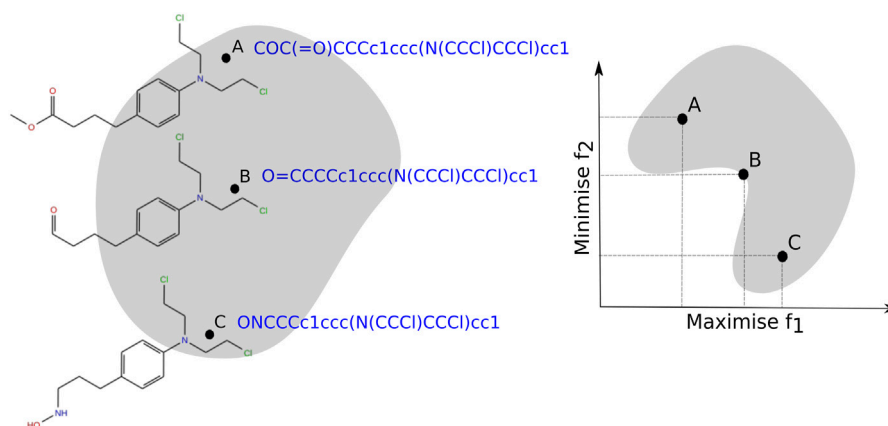


FIGURE 2 Hypothetical decision space (left) and objective space (right) associated with objectives f_1 and f_2 for the molecules (A, B and C) represented by the SMILES strings.

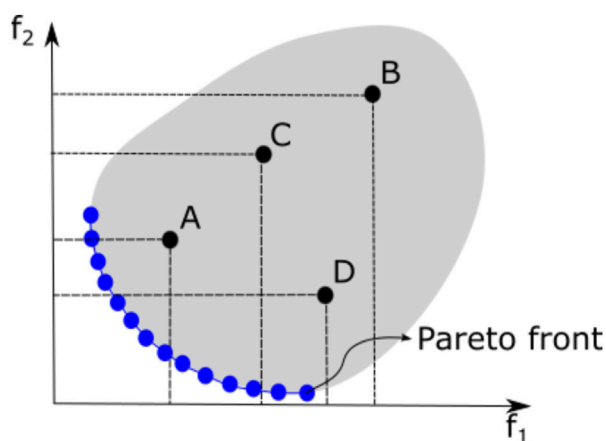


FIGURE 3 Pareto-optimal solutions for a two-objective minimization problem.

objective space (represented by the affinity prediction of the compound against each target).

2.1 Pareto optimality

In MultiOOP, several objectives must be simultaneously optimized. Usually, no single solution would give the best values for all the objective functions. Instead, in a typical MultiOOP with conflicting objectives, a set of solutions is superior to the others when all objectives are considered. Such solutions are those where none of the objectives can be improved without deteriorating at least one of the other objectives. Those solutions are called *Pareto-optimal solutions* due to the Italian economist Vilfredo Pareto, who introduced this theory in 1896 [Pareto \(1964\)](#). The optimal solution for a MultiOOP is based on the Pareto optimality concept, in which two solutions are compared based on whether one dominates the other.

Definition 2.1. A decision vector $\mathbf{x} \in S$ dominates another vector $\mathbf{y} \in S$ ($\mathbf{x} < \mathbf{y}$) if and only if

$$f_i(\mathbf{x}) \leq f_i(\mathbf{y}) \quad \forall i \in \{1, \dots, k\} \quad \text{and} \quad \exists j \in \{1, \dots, k\}: f_j(\mathbf{x}) < f_j(\mathbf{y})$$

where k is the number of objectives.

Definition 2.2. A decision vector $\mathbf{x}^* \in S$ is a Pareto-optimal solution if no other $\mathbf{x} \in S$ dominates \mathbf{x}^* . The set of all Pareto optimal solutions in the decision space is called Pareto-optimal set, or simply Pareto set (PS), and its image in the objective space is called Pareto-optimal front, or Pareto front (PF).

The Pareto-optimal set is the best collection of solutions to the MultiOOP. Research in the area of MultiOOP is concerned with the problem of how to identify the PS or at least a good approximation of it. The ideal approach would be to find (i) a set of solutions as close as possible to the PF and (ii) as diverse as possible along that front [Deb \(2001\)](#). A typical PF is illustrated in [Figure 3](#), where f_1 and f_2 are two objective functions that must be simultaneously minimized. Four different points (solutions values) are shown in this figure: solution A dominates solution B, as $f_1(A) < f_1(B)$ and $f_2(A) < f_2(B)$; A also dominates solution C for the same reason. However, solutions A and D are non-dominated by each other since $f_1(A) < f_1(D)$ but $f_2(A) > f_2(D)$. The Pareto-optimal solutions (dots in blue) that form the PF are those in which no objective can be improved without making at least one other objective worse.

The concept of dominance is of the utmost importance in this type of optimization problem, as it allows the comparison of two different solutions concerning distinct objective functions. Hence, to find the non-dominated set of solutions, the dominance relation $<$ is used to identify the best between two given solutions.

3 Solution methods

There are different ways to deal with a MultiOOP, usually consisting of three stages: model building, optimization, and decision-making [Branke et al. \(2008\)](#). First, we formulate the

optimization problem in which the decision variables, objectives, and constraints are specified. Second, an optimization technique is used to find the best compromise solutions. Since the Pareto optimal solutions are equally acceptable, a decision maker (DM), who usually has expertise in the problem domain, must decide which solution(s) best suits your preferences. Following the approaches that distinguish when the DM interacts with the optimization procedure, the methods can be classified as [Miettinen \(1999\)](#):

- **No-preference:** no information about the importance of the objectives is assumed, and the DM's preferences are not considered. The problem can be solved by any method used to find a single optimum solution. The solution obtained is presented to the DM, which will accept or reject it.
- **A priori:** the hopes and opinions of the DM are taken into consideration before the solution process. Usually, one preferred Pareto-optimal solution is obtained. Those methods, known as *preference-based*, require the DM to know the priority of each objective beforehand.
- **A posteriori:** the PS is first generated, and then the DM is supposed to select the most preferred solution from this set of alternatives. The preference information is considered after the optimization process.
- **Interactive:** the DM preferences are progressively used during the search procedure and are adjusted as the search continues.

The most intuitive and simple way to solve a MultiOOP is to convert it into a SingleOOP. The so-called *classical methods* mainly propose different ways of scalarising the objectives. These are commonly used methods that have existed for many decades. Most of them are *aggregation-based* techniques, which aggregate the objective functions into one objective. Only one Pareto-optimal solution can be found in these methods at each execution. To obtain different Pareto-optimum solutions, the SingleOOP must be solved several times, with different parameters for finding solutions in the entire Pareto-optimum region. The weighted sum, the ϵ -constraint, and the weighted Tchebycheff [Miettinen \(1999\)](#) methods are examples of such classical techniques. The weighted sum method is probably the most used classical approach due to its simplicity and ease of use. In this process, a weighting coefficient is associated with each objective function, and a weighted sum of the objectives is minimized. For each function f_i , there is a weight w_i associated, such that $0 \leq w_i \leq 1$ for all $i = 1, \dots, k$, with $\sum_{i=1}^k w_i = 1$ (weights are normalised). By using this technique, the MultiOOP given in Eq. 1 is converted into a SingleOOP as follows

$$\begin{array}{ll} \text{Minimise} & \sum_{i=1}^k w_i f_i(\mathbf{x}) \\ \text{subject to} & \mathbf{x} \in S \end{array} \quad (2)$$

Considering that the ideal approach for solving MultiOOPs would be to find many different trade-off solutions as close as possible to the PF and as diverse as possible along that front, it becomes clear that classical methods need a great effort to meet these

goals. As they combine multiple objectives into one, some knowledge of the problem is required. Note that the optimization of a single objective results in a single-point solution. In this way, multiple runs must be performed to generate different alternatives to the DM. Moreover, if the variable space is discontinuous and some objectives have many local minima, these methods may not work properly [Srinivas and Deb \(1994\)](#).

Differently, *multi-objective evolutionary methods* seek to optimize the problem in its original form, with independent objectives, providing in each execution a set of Pareto-optimal solutions. EAs have become extremely popular over the last years as a non-classical, stochastic search technique to solve multiOOPs. Next, we present a short overview of this class of methods and their main differences compared to the classical approaches.

3.1 Multi-objective evolutionary algorithms (MultiOEAs)

The process of evolution of species inspires EAs, and they differ from classical methods in various ways. The most prominent is using a population of candidate solutions instead of a single solution, as in classical methods. This characteristic allows those methods to find many Pareto-optimal solutions in a single run. Although they do not guarantee to find the optimal trade-off solutions, they can provide a satisfactory approximation set, which is (hopefully) not too far away from the true PF.

In the field of evolutionary computation, Genetic Algorithms (GAs) are the most popular ones. Developed by J. H. Holland in the 60s ([Holland, 1962](#); [Holland, 1975](#)), they are based on the evolution of a population of individuals. Initially, a population of candidate solutions is randomly generated. For every individual, an objective function associates a fitness value indicating its suitability to the problem. At each iteration, individuals are selected to form the parents. Those parents are reproduced using different operators (*e.g.*, crossover, mutation) to generate new offsprings. Then, a replacement scheme is applied to determine which individuals of the population will survive from the offsprings and the parents. This process is repeated until a stopping criterion is reached. A general scheme of an EA is given in [Algorithm 1](#).

```

1 Set  $k = 0$ ;
2 Randomly generate an initial population of solutions;
3 Evaluate each solution in the initial population;
4 while  $k < G$  do
5   Selection of individuals;
6   Apply mutation and crossover operators to obtain
   new solutions;
7   Evaluate solutions;
8   Select individuals for the next-generation;
9   Set  $k = k + 1$ ;

```

Algorithm 1. Pseudocode of a general EA.

Following these baseline steps, GAs are known to be very efficient in solving real problems in several fields [Slowik and](#)

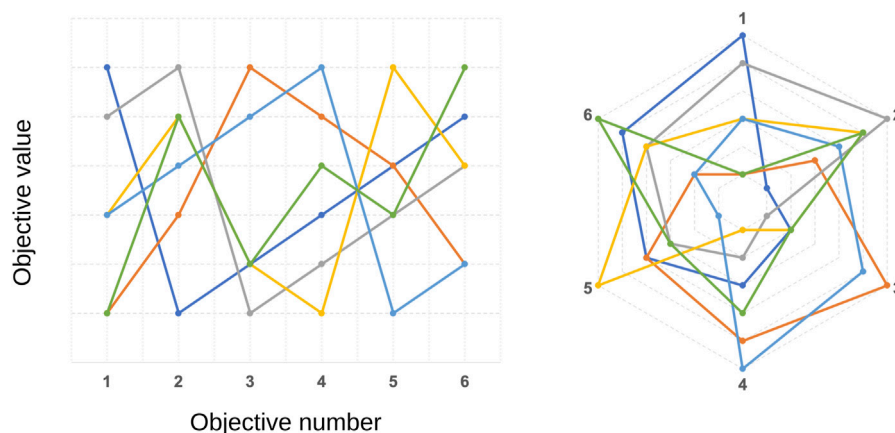


FIGURE 4

Illustration of parallel coordinates (left) and radar (right) plots for different solutions in a 6-objective optimization problem. Each coloured solid line represents one non-dominated solution.

[Kwasnicka \(2020\)](#). As they do not require additional information about the problem, like continuity or differentiability, they are also well-suited to solve black-box optimization problems¹.

The need to find trade-off solutions as close as possible to the Pareto optimal front (good convergence) and as diverse as possible along that front (good diversity) are the most important issues in MultiOOP. Therefore, search algorithms must be designed to obtain multiple solutions, each offering a different trade-off for the objective functions. In a simplistic case, this may be achieved by storing each found solution in an “archive” that maintains only non-dominated solutions. The way this archive is maintained, how individuals are selected and recombined, whether elitism is used or not, and how fitness assignment is applied characterize the various MultiOEs. They can be classified into three main categories according to the different strategies employed [Emmerich and Deutz \(2018\)](#).

3.1.1 Pareto-based

The earliest attempts for solving MultiOOPs with MultiOEs were based on the Pareto dominance relation. The main idea is that the fitness value is assigned to individuals based on the Pareto dominance principle to achieve good convergence. An explicit diversity preservation scheme is also employed to maintain the diversity of solutions. Some of the most representative techniques of this class of methods are the Strength Pareto EA (SPEA/SPEA-II) [Zitzler and Thiele \(1998, 1999\)](#); [Zitzler et al. \(2001\)](#), the Pareto Archived Evolution Strategy (PAES) [Knowles and Corne \(1999, 2000\)](#), the Niche Pareto GA (NPGA/NPGA2) [Horn et al. \(1994\)](#); [Erickson et al. \(2001\)](#) and the Non-dominated Sorting GA (NSGA/NSGA-II) [Srinivas and Deb \(1994\)](#); [Deb et al. \(2002\)](#).

The primary concern with the Pareto-based approach arises when the number of objectives increases, particularly beyond three objectives. In such cases, it becomes increasingly challenging for these techniques to select solutions, as most of the solutions in the population tend to be non-dominated by one another. [Section 4](#) discusses some of the main difficulties that arise when the number of objectives increases.

3.1.2 Indicator-based

In this approach, performance metrics, also known as quality indicators, are employed to define the selection mechanisms. Performance metrics are used to assess the quality of an approximation set generated by an algorithm. It assigns a real value to one or more approximation sets depending on certain quality aspects, such as (i) convergence toward the Pareto optimal region and (ii) diversity of solutions along the PF. The underlying idea of those techniques is to optimize the indicator value of the non-dominated set generated throughout the evolutionary process.

From the literature, it is not difficult to notice that the HV indicator is the most widely adopted metric for evaluating indicator-based optimizers and assessing non-dominated sets' quality. This metric calculates the volume of the dominated region by the obtained solution. Although there are many efforts to reduce the computational complexity of the HV computation, it is also known that the high computational cost involved when the number of objectives increases could limit the use of this metric [Beume et al. \(2009\)](#); [Guerreiro and Fonseca \(2018\)](#). However, the theoretical properties of the HV justify its widespread acceptance [Zitzler et al. \(2003\)](#).

A recent survey on indicator-based MultiOEs can be found in [Falcón-Cardona and Coello \(2020\)](#), where the authors presented solution methods from their origins up to their applications by current state-of-the-art approaches.

3.1.3 Decomposition-based

The pioneering work on decomposition-based methods was due to [Zhang and Li \(2007\)](#), who first proposed in 2007 the Multiobjective EA Based on Decomposition (MOEA/D). In this

¹ In black-box optimization, the objective function and/or constraints are approximated by computational methods that provide an answer (output) to the problem, given a set of input values. In practice, they often occur when the computation of the objective functions and/or constraints requires an expensive computer simulation or when their functional form are unknown.

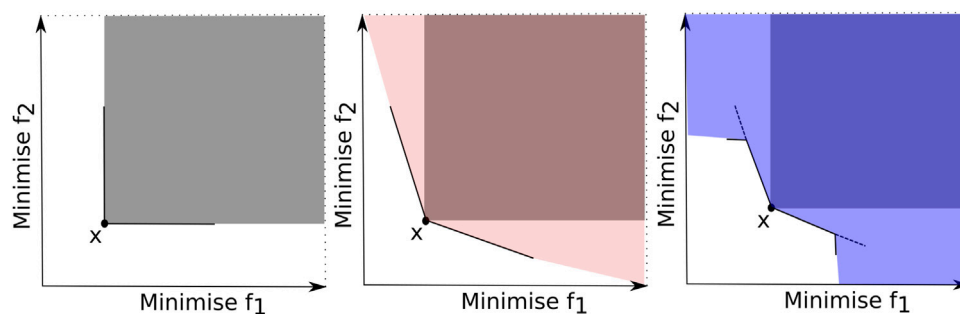


FIGURE 5

Illustration of dominated areas by solution x using three dominance relations: (left) Pareto dominance, (middle) α -dominance and (right) $CN\alpha$ -dominance.

approach, the MultiOOP with k objectives is decomposed into M single-objective subproblems through some aggregation technique. These M subproblems, represented by M -weighted vectors, are simultaneously optimized in a single run. Different weight vectors need to be used to generate a set of Pareto-optimal solutions. The MOEA/D explores the neighborhood relationship between these subproblems to solve the original problem in (1) efficiently.

A variety of methods was proposed for improving the design of the components of the MOEA/D, motivated by the existing limitations of the method. Later investigations tried to improve its performance by seeking new ways of, for instance, (i) decomposing the problem, (ii) generating the weight vectors, (iii) improving the efficiency of genetic operators, and iv) enabling the application of the method to ManyOOPs Trivedi et al. (2017); Xu et al. (2020).

4 Dealing with many-optimization problems

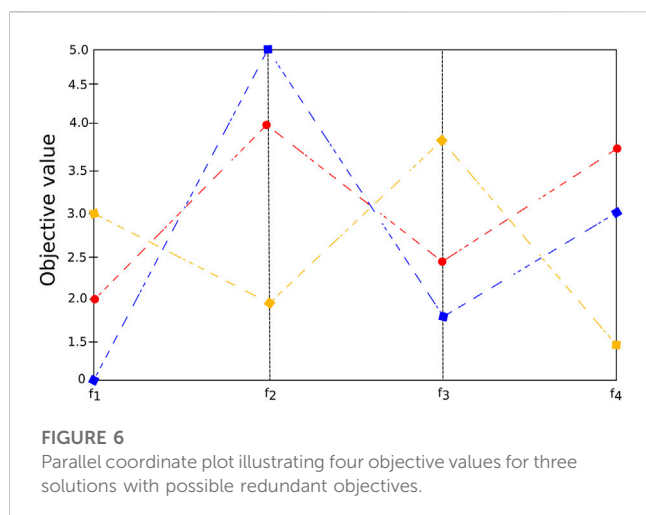
ManyOOPs are those having more than three objectives. As previously mentioned, drug design problems usually have more than three objectives to be optimized simultaneously, making them potential candidates for problems to be treated with ManyOOP approaches.

4.1 Main challenges

The presence of many objectives introduces several challenges that affect the choice and the design of evolutionary techniques. The difficulty of optimizing a large number of objectives is due to the so-called *curse of dimensionality*. The PF takes different forms depending on the number of objectives, e.g., a two-objective problem results in a PF that is a curve or line, while a three-objective problem generates a PF that is a surface. As the number of objectives increases beyond four, the PF may be represented by a hypersurface or other higher-dimensional shape. Therefore, the number of solutions needed to represent the entire PF grows exponentially with the number of objectives Saxena et al. (2013). When the number of objective functions increases, one has to deal with the following issues Ishibuchi et al. (2008):

- almost all solutions in the population become non-dominated by each other, leading to a phenomenon called *dominance resistance* Purshouse and Fleming (2007). Such characteristic severely deteriorates the convergence of MultiOEsAs;
- scaling issues may be encountered in choosing the appropriate size of the population since the number of points needed to approximate the entire Pareto front increases exponentially with the number of objectives Saxena et al. (2013);
- beyond 3D space, visualizing the solution set in the objective space becomes less intuitive and needs special techniques, such as parallel coordinates or radar plots (Figure 4). Moreover, that makes it harder for the DM to choose the best solution.
- performance metrics are also affected by the curse of dimensionality, e.g., the HV and the IGD metrics have their performance compromised Ishibuchi et al. (2016);
- classical aggregation techniques (presented in Section 3) are not limited to problems with up to three objectives. However, the difficulties pointed out in such methodology grow increasingly, e.g., setting the weights of a large number of coefficients.

Given the difficulties mentioned above, the importance of treating ManyOOPs as a distinct class of problem is evident. In Allmendinger et al. (2022), the authors reflect on the question: “What if we increase the number of objectives?” In their paper, theoretical implications are presented on how the presence of many-objectives can impact the performance of MultiOEsAs when solving the NK-landscapes problem Verel et al. (2013). They identified a series of drawbacks of dominance-based (e.g., NSGA-II Deb et al. (2002)), indicator-based (e.g., IBEA Zitzler and Künzli (2004)), and decomposition-based (e.g., MOEA/D Zhang and Li (2007)) MultiOEsAs, along with recommendations for enhancing these techniques in the context of ManyOOPs. Thus, it is clear that MultiOEsAs encounter performance limitations when applied to problems with more than three objectives. Therefore, evolutionary techniques that can be effectively applied to ManyOOPs are often referred to as ManyOEsAs and represent the cutting edge of multi-objective optimization research. The following sections will cover different EAs strategies to deal with ManyOOPs efficiently.



4.2 Many-objective evolutionary algorithms (ManyOEs)

Unsurprisingly, new methods have been proposed to improve existing MultiOO techniques, and new methodologies have been developed in recent years. For readers interested in exploring existing many-objective algorithms across various domains, we recommend referring to the extensive list provided in the articles [Mane and Narasinga Rao \(2017\)](#); [von Lücken et al. \(2019\)](#). We will reference a few relevant strategies and introduce more recently published approaches. Herein, we classified ManyOEs into five categories according to the approach adopted to treat the problem in a ManyOO context: *relaxed dominance*, *indicator*, *decomposition*, *dimensionality reduction* and *hybrid techniques*. The [Supplementary Material](#) includes pseudocodes for some of the referenced methods.

4.2.1 Relaxed dominance

As previously mentioned, when the number of objectives increases, almost all solutions in the population become non-dominated by each other. Hence, for MultiOEs based on the Pareto dominance relation, the selection pressure towards the PF is compromised. A path explored to circumvent the scalability issue of dominance-based methods is to use relaxed forms of Pareto dominance to reduce the impact of dominance resistance. Such an approach can enhance the selection pressure toward the PF and provide a way of regulating the convergence of MultiOEs.

Relaxed Pareto dominance (RPD) relations modify the Pareto dominance concept for better-discriminating solutions for selecting the best ones with enhanced selection pressure. The α -dominance [Ikeda et al. \(2001\)](#), ϵ -dominance [Laumanns et al. \(2002\)](#), cone ϵ -dominance [Batista et al. \(2011\)](#), θ -dominance [Yuan et al. \(2016\)](#), CN-dominance [Dai et al. \(2014\)](#), CN α -dominance [Liu J. et al. \(2019\)](#), and MultiRPD [Zhu et al. \(2022\)](#) are some examples of RPDs. [Figure 5](#) illustrates the dominated area of a solution x using three different dominance relations: (a) the Pareto dominance, (b) the α -dominance and (c) the CN α -dominance for a bi-objective problem. Observe that the α -dominance and the CN α -dominance expand the domination area compared to the Pareto dominance. Under these relaxed definitions, a solution is expected to

have a greater chance of being dominated by other solutions and the selection pressure towards the PF is enhanced.

In [Li et al. \(2015a\)](#), one can find diverse RPD-based methods, each applying a different RPD approach. Recently, the CN α -dominance proposed in [Liu J. et al. \(2019\)](#) combines two existing dominance relations, the α -dominance [Ikeda et al. \(2001\)](#) and the CN-dominance [Dai et al. \(2014\)](#). The idea was to strengthen the selection pressure by expanding the dominated area by combining two other RPDs, as shown in [Figure 5](#). (c) Computational experiments considered a non-dominated neighbor Immune Algorithm (NNIA) as the baseline method. The NNIA with the proposed CN α -dominance was compared against the NNIA with four different dominance relations. The results demonstrated the superiority of the CN α -dominance against different dominance relations in terms of solution quality and selection pressure on benchmark problems with five to fifty-five objectives.

4.2.2 Indicator

To improve the convergence ability of ManyOEs, extensive research has been conducted on fitness evaluation mechanisms based on quality indicators ([Falcón-Cardona and Coello, 2020](#)). As mentioned previously, the high-dimensional problem encountered in ManyOO may turn the application of HV-based MultiOEs impracticable despite the many efforts to mitigate its computational cost. The first method that attempted to accelerate the HV computation and make it scalable for a large number of objectives was proposed by [Bader et al. \(2010\)](#); [Bader and Zitzler \(2011\)](#). The main idea of the hypervolume estimation algorithm (HypE) was not to calculate the exact values of the HV but instead to provide an estimate of this value through Monte Carlo simulations. Experimental results showed that HypE achieved competitive performance regarding the average HV on benchmark problems with up to fifty objectives.

The Two-archive methodology (Two-Arch/Two-Arch2) [Praditwong and Yao \(2006\)](#); [Wang et al. \(2015\)](#) was the first to divide the non-dominated solution set into two archives, one that promotes convergence (CA) and another that emphasizes diversity (DA). CA and DA employ different updating rules that reflect their respective roles in the optimization process. In Two-Arch2, the update rule of CA is based on the quality indicator $I_{\epsilon+}$ from [IBEA Zitzler and Künzli \(2004\)](#), and the DA archive is updated based on the Pareto dominance rule. This method falls under the hybrid approach category. The results demonstrated that Two-Arch and Two-Arch2 outperformed other ManyOEs in terms of convergence with comparable diversity quality in problems having two to eight objectives for the Two-Arch and up to twenty objectives for the Two-Arch2.

4.2.3 Decomposition

In decomposition-based methods, the ManyOOP is divided into several SingleOO subproblems using a set of weight vectors. The basic idea is to find a set of well-distributed, non-dominated solutions along the PF using generated weight vectors so the diversity of the population is controlled explicitly by weight vectors. Ideally, each solution in the population is associated with a subproblem.

The most representative algorithms of this class are the MOEA/DD [Li K. et al. \(2015\)](#) and the NSGA-III [Deb and Jain \(2014\)](#); [Jain](#)

and Deb (2014), which are extensions of the MOEA/D Zhang and Li (2007) and NSGA-II Deb et al. (2002), respectively. Those methods may also be classified as hybrid, employing a decomposition-based approach to maintain population diversity, while the Pareto dominance rule controls the algorithm's convergence. Both methods outperform many contemporary MultiOEs. It is worth noting that the performance of those methods strongly depends on the shape of the PF Ishibuchi et al. (2017).

The MOEA/D with Update when Required (MOEA/D-UR) method de Farias and Araújo (2022) proposed a new scheme to adapt the weight vectors depending on whether they show signs of convergence. An additional technique for partitioning the objective space was proposed to increase the spread of individuals in the population to estimate the level of regularity of the PF's shape. The computational experiments show that the MOEA/D-UR has competitive performance compared to ten state-of-the-art ManyOEs in test problems and real-world problems with up to fifteen objectives.

4.2.4 Dimensionality reduction

The visualization of objectives in large dimensional space can be difficult. See, for example, Figure 6 (a), in which seven objective values are presented for many different solutions (each represented by a color). The more pronounced the slope of the line connecting two solutions, the higher the likelihood of a possible conflict between them. In dimensionality reduction-based methods, the idea is to decrease the problem's difficulty by reducing the number of objectives by identifying redundant objectives Brockhoff and Zitzler (2006, 2007). For many problems, a smaller set of m ($m < k$) conflicting objectives exist that can generate the same PF as the original problem Yuan et al. (2018). For instance, Figure 6 shows the parallel coordinates plot of three solutions x_1 (blue), x_2 (red) and x_3 (yellow) and four objectives. The figure indicates that the objective functions f_1 and f_3 are redundant, since $f_1(x_1) < f_1(x_2) < f_1(x_3)$ as well as $f_3(x_1) < f_3(x_2) < f_3(x_3)$, that is, when f_1 is minimized f_3 is also minimized. Therefore, we can consider $\{f_1, f_2, f_4\}$ as the minimum objective set that preserves the dominance structure.

A list of many dimensionality reduction frameworks and ManyOEs based on them is presented in Li B. et al. (2015), in which the methods are classified as *online* and *offline*. Online methods reduce the number of objectives gradually during the search process. In contrast, offline methods reduce the number of objectives after obtaining the PS. These methods can reduce the computational load of ManyOEs and assist DMs in distinguishing points based on non-redundant objectives. However, an open question is whether the loss of information can cause problems in the optimization process.

A different approach for reducing the number of objectives was proposed by de Freitas et al. (2015), where the authors introduced the Aggregation Tree (AT). The AT tool allows the visualization of redundancies and conflict between objectives in the form of a tree by using the concept of *harmony* to reduce the number of objectives. The more harmonic the two objectives are, the more suitable they are to be aggregated into a group of objectives without much loss in the representation of the PS. This technique was used in a multi-objective GA, the GAPF algorithm, to solve a protein structure prediction problem in which the AT was used to arrange seven terms of the energy function into a three-objective problem Rocha et al. (2017).

4.2.5 Hybrid approach

Hybrid techniques have been proposed to balance convergence and diversity in ManyOEs. The method proposed in Zou et al. (2021), MaOEA/DRA, is based on dominance and decomposition approaches. An elitism mechanism is exploited to balance the convergence and diversity of the evolutionary process. Simultaneously, a reference point adaptation scheme is designed to "learn" the true PF shape of different problems. Computational results showed that the MaOEA/DRA outperformed seven state-of-the-art algorithms on various test problems with up to fifteen objectives.

The IDEA algorithm Xia et al. (2023) couples indicator-based and decomposition-based mechanisms. The decomposition-based approach promotes population diversity, while the I_{∞}^r Yuan et al. (2021) metric is used as the indicator value for achieving population convergence and distinguishing individuals. The performance of the method was evaluated on test problems with up to fifteen objectives. The results demonstrated that IDEA is effective in solving ManyOEs.

5 Applications in *de novo* drug design

In the field of Computational Chemistry, MultiOO has been adopted for decades to obtain trade-offs among the objectives considered Handl et al. (2007); Ekins et al. (2010); Nicolotti et al. (2011); Parikh et al. (2023). A multitude of multi-objective approaches can be found in the literature, and we do not intend to describe or cite them exhaustively. This section describes some reported studies that explicitly apply multi-objective and many-objective strategies in dnDD. Following the previously defined classifications, the methodologies reviewed are mainly classified as aggregation-based and Pareto-based methods.

Recalling the preferences of the DM, the aggregation methods that will be cited in the next sections are mainly based on (i) *a priori* information, when the DM's preferences are considered before the optimization process, and thus only a single final solution is generated, and (ii) *a posteriori* information, when a set of Pareto optimal solutions are generated for the DM to choose the best-suited solution.

5.1 MultiOO methods: optimizing up to three objectives

5.1.1 Aggregation-based

The GANDI algorithm Dey and Caflisch (2008) is a fragment-based approach that generates molecules by joining pre-docked fragments with a list of fragments provided by the user. In the paper, GA Holland (1975) and Tabu Search (TS) Glover and Laguna (1997) algorithms are used in combination to match those fragments. While the GA was used to generate new fragments by applying genetic operators, the TS was adopted to link those encoded fragments. GANDI is an *a priori* method that scalarises the scoring functions into one, providing a single optimal solution. The weighted sum approach was used to aggregate three scoring functions to be minimized: a force field-based binding energy and two measures of similarity to a user-defined structure. Forbidden

connections were avoided to prevent the generation of unstable molecules. When applied to the CDK2 kinase, the proposed method generated 1,809 molecules, of which eight were present in the ZINC database and commercially available Irwin and Shoichet (2005).

The method proposed by Devi et al. (2014) used a GA guided by the scalarisation of two objective functions, drug-likeness (by Lipinski's Rule of 5 Lipinski (2004)) and similarity (by Tanimoto similarity Loving et al. (2010)) to a known reference molecule from the e-Drug3D database. The weights of the objectives are set *a priori*. Two experiments were conducted, one using a local anesthetic drug (lidocaine) and another using an anti-cancer molecule (furanopyrimidine) as reference molecules. For each reference molecule, the authors analyzed the proposed method using one (Tanimoto similarity alone) and two objectives (using a weighted sum of the similarity and drug-likeness functions). For the furano-pyrimidine molecule, the two-objective version of the method could generate drug-like molecules more efficiently than the one-objective formulation.

The MoGADdrug method, introduced in Devi et al. (2021), is a fragment-based GA that constructs new molecules from a set of chemical fragments and a reference molecule as inputs. Building upon previous work Devi et al. (2014), which considered only two fragments (acid and amine), this method incorporates a variable-length representation to construct new solutions considering double amine fragments, thus allowing three fragment types. The objective function is a weighted sum of the oral bioavailability score Lipinski (2004), and the 2D similarity based on the Tanimoto coefficient Brown (2009), with the weights being set *a priori*. The MoGADdrug could design drug-like molecules similar to lidocaine, furano-pyrimidine derivative, imatinib, atorvastatin, and glipizide.

5.1.2 Pareto-based

The Compound Generator (CoG) method Brown et al. (2004) represents the first instance of a MultiOEA being applied in dnDD. The authors used Genetic Programming (GP), in which the topology of molecules is represented by graphs, where the graph nodes represent the atoms. Special mutation and crossover operators were implemented to generate new molecules, and no constraints were imposed in the search space. The problem was formulated as a two-objective optimization problem, where they simultaneously maximize the evolved solutions regarding the Tanimoto similarity Bajusz et al. (2015) between the candidate molecule and two compounds representative of different chemical classes. The authors performed two experiments: one using two similar molecules and another using two diverse molecules. The results demonstrated that the graph-based CoG could generate a set of novel molecules that are visibly similar to the target compounds in both experiments.

Another graph-based framework, MEGA Nicolaou et al. (2009), tested the creation of new selective estrogen receptor (ER) compounds, aiming to maximize the docking score for the "positive" target receptor ER- β and minimize scores for the "negative" but closely related target ER- α . To generate molecules with desired chemical profiles similar to the known ligand tamoxifen, they used filters based on the Rule-of-Five (Ro5) and Tanimoto similarity scores as constraints. The authors pointed out the uneven nature of the objectives, *i.e.*, identifying solutions with

reduced binding affinity for ER- α is substantially simpler than designing compounds with increased binding affinity for ER- β . This may lead to intense search space exploration towards the "easier" objective. To circumvent this situation, a mechanism called *niching* was used to preserve the diversity of the molecules, ensuring that the objectives were treated equally and there were no dominance conditions in favor of a specific objective. In addition, they applied *Pareto-elitism* to prevent good solutions from being "lost" during the generations. The authors reported that both mechanisms generated many non-dominated solutions while increasing the PF extension.

In Daeyaert and Deem (2017), synthesizable molecules were generated using a GA named Synopsis Vinkers et al. (2003). The authors proposed an improved version of the *de novo* program Synopsis that incorporates the non-dominated sorting procedure from NSGAI Deb et al. (2002). The proposed algorithm was evaluated to optimize two objectives: the docking scores, computed by the Autodock Vina program Trott and Olson (2010), associated with the fibroblast growth factor (FGFR) and the vascular endothelial growth factor (VEGFR), aiming to generate dual selective inhibitors for cancer. Moreover, thirteen additional scores were used as constraints to avoid generating unwanted molecules. The obtained solutions exhibit good predicted binding energies to their targets and possess structural and physicochemical parameters falling within the typical range for drug-like molecules. The use of the Pareto dominance approach enabled the generation of high-affinity compounds within the imposed restrictions, unlike the aggregated single-objective approach, which rarely produces good-quality solutions.

In the works of Devi et al. (2019, 2020), two multi-objective methods were explored for dnDD of new drug-like molecules: the monkey algorithm (MoMADrug) Devi et al. (2019) and the biofilm algorithm (MOBifi) Devi et al. (2020). MoMADrug is inspired by the behavior of monkeys, while MOBifi is inspired by the life cycle of bacteria in a biofilm. Both methods were adopted in the context of fragment-based *de novo* design and considered the Tanimoto similarity to known compounds Brown (2009) and oral-bioavailability scores Lipinski (2004) as the objectives to be maximized. In addition to these two objectives, MOBifi also employed the Veber score Veber et al. (2002) as a third objective related to oral bioavailability. In Devi et al. (2019), the authors compared the MoMADrug against the MoGADdrug Devi et al. (2021), which uses a weighted sum of the objectives. Their results showed that MoMADrug could produce a more diverse set of solutions due to its multi-objective nature and its use of Pareto dominance criteria.

The MOBifi Devi et al. (2020) was originally evaluated in three unconstrained benchmark problems from CEC 2009 Zhang et al. (2008) with three objectives against five other multi-objective methods. It was found that MOBifi is a competitive method in terms of the inverted generational distance (IGD) and the maximum spread (MS) performance metrics. The MOBifi method was applied to generate drug-like molecules based on reference anti-diabetic compounds from herbal plants. These generated molecules were further docked against the therapeutic targets tyrosine phosphatase 1B (PTP1B) and α -glucosidase (AGS) associated with the diabetes treatment. Two generated compounds exhibited docking scores similar to the reference inhibitor (rutin), indicating their

potential as anti-diabetic agents. Since the fragments used in the dnDD strategy were generated from the commercially available databases Enamine and e-Lead3d, based on the predetermined coupling reactions concept Yuan et al. (2011), the designed molecules have a high probability of being successfully synthesized and purchased from the chemical vendors.

The proposed approach known as Deep Evolutionary Learning (DEL) introduced in Grantham et al. (2022) combines a deep generative model (DGM) with multi-objective evolutionary algorithms (MultiOEA) for dnDD. DEL leverages a fragment-based variational autoencoder (FragVAE) and NSGAI components, including non-domination rank and crowding distance, to design new molecules. By operating in the continuous latent representation space generated by the neural generative model, DEL avoids the limitations of a discrete structural space. The DGM is iteratively fine-tuned based on the newly generated populations of samples with better properties. The quantitative estimation of drug-likeness (QED), synthetic accessibility score (SAS), and logP were selected as objectives in DEL. Extensive validation of the approach was performed to assess its population validity, novelty, and diversity across various benchmarking sets, including those in the MOSES framework.

Based on the DEL framework, in Mukaidaisi et al. (2022), a graph-based DGM, called JTVAE (Junction Tree Variational AutoEncoder), is integrated into DEL to provide a latent representation space for the MultiOEA exploration. Unlike FragVAE, which utilizes SMILES fragmentation, the JTVAE employs graph fragmentation following the subgraph-by-subgraph strategy. Binding affinity score (BAS) predicted with the docking program QVina, SAS, and logP of the generated molecules are the three objectives to be optimized. During the optimization process, non-dominated ranking is performed. The VAE model is refined by selecting the high-quality generated molecules possessing significant BAS, SAS, and logP properties for the next-generation. Computational experiments were conducted comparing DEL + FragVAE with the JTVAE approach on the ZINC dataset and on a variant of ZINC that included drug molecules from the DrugBank database (ZINC + DrugBank dataset). The results demonstrated that both methods improved the properties of the molecules along the generations and that the JTVAE has a higher HV value than FragVAE. These experimental results confirm that the JTVAE approach, when compared to the FragVAE approach, improves the properties of molecules during the optimization process and leads to a higher HV value.

The DrugEx is a ligand-based approach that applies RNN-based reinforcement learning to generate new chemical structures in the SMILES format. The first version of DrugEx Liu X. et al. (2019) was designed to generate ligands by performing single-objective optimization using the predicted affinity against the human adenosine A_{2A} receptor as the objective through a random forest-based quantitative structure-activity relationship (QSAR) model.

In the second version, DrugEx v2 Liu et al. (2021) expanded its usage to include multi-objective optimization by applying the concept of mutation and crossover into the Reinforcement Learning (RL) framework and a Pareto ranking procedure to handle the different objectives. The performance of DrugEx v2 was compared to ORGANIC Sanchez-Lengeling et al. (2017)

and REINVENT Olivecrona et al. (2017) methods by considering as objectives the affinity prediction against either multiple targets or a single target while considering off-target effects. For multiple targets, the desired molecules should exhibit a high affinity towards A_{1A} and A_{2A} receptors. In contrast, in the target-specific scenario, the designed molecules should display high affinity towards A_{2A} while maintaining low affinity to A₁. In addition, low affinity to the hERG channel (Ether-à-gogo-Related Gene, the α subunit of a potassium ion channel) is required in both cases. The generated molecules showed a large percentage of validity, low duplication, and similarity to known ligands.

In the latest version, DrugEx v3 Liu et al. (2023) adopted a graph-based transformer model as the generative model, considering user-defined scaffolds as inputs to create new molecules with desired chemical profiles. A novel encoding scheme for atoms and bonds was proposed based on an adjacency matrix to enable the transformer model to handle molecular graph representations. Unlike the previous versions, DrugEX v3 was evaluated considering two objectives: (i) the drug-likeness using the QED score and (ii) the affinity score towards the A_{2A} receptor predicted with the random forest-based QSAR model. A Pareto-based ranking scheme was employed to rank molecules based on the average Tanimoto distance instead of the commonly employed crowding distance. SMILES and graph representations were tested on four deep learning (DL) architectures. According to the results, all the molecules generated by the newly proposed method using the provided scaffolds were valid, and most exhibited a high predicted affinity towards A_{2A}.

5.2 ManyOO methods: dealing with more than three objectives

5.2.1 Aggregation-based

In a recent study, Elend et al. (2022) enhanced a Pareto-based ManyOEA proposed in Cofala et al. (2020) by incorporating a neural language model trained on the ZINC database to improve the quality of generated molecules. The goal was to generate molecules that inhibit the therapeutic target Mpro of SARS-CoV-2 by considering multiple objectives such as predicted binding affinity (BA), quantitative estimate of drug-likeness (QED), natural product-likeness (NP), toxicity filter (TF), and synthetic accessibility (SA). A weighted sum approach was used to aggregate the objective values in which weights were set *a priori*. The proposed Evolutionary Molecular Generation Algorithm (EMGA) designed new molecules based on the SMILES representation. The transformer architecture in the neural language model was used as a mutation operator to generate molecule fragments, while a $(\mu + \lambda)$ Evolutionary Strategy (ES) was employed to perform a randomized search in the molecular structure search space. From the molecules generated by EMGA, twenty-one chemically valid molecules were selected for molecular dynamic (MD) simulations. Among them, two were identified as stable and had the potential to inhibit Mpro.

The ATOM Generative Molecular Design (ATOM GMD) proposed in McLoughlin et al. (2023) involves a two-stage process of a variational autoencoder (VAE) and a many-objective GA operating in the latent space. The VAE is used to map molecular

structures to the continuous latent space. Specifically, a JTVAE maps a population of structures to a learned continuous latent space. The encoder component of the JTVAE converts each SMILES structure into a continuous latent vector, and the decoder component performs the opposite transformation. The GA-based approach searches for optimal molecules in the latent space, employing crossover and mutation operations to generate new molecules with desirable properties. The cost function of this ManyOO approach is a weighted sum of terms based on 12 predicted properties, including efficacy and safety (binding affinities against the therapeutic target histamine H1 receptor and the off-targets M2 and hERG predicted with QSAR models), pharmacokinetic profile (Ro5), and developability properties (SAS). The weights are set *a priori*. Focusing on developing potent and selective H1 antagonists, the ATOM GMD approach generated several molecules, of which 106 compounds were further synthesized and experimentally evaluated. Six tested compounds were found to bind H1 at nanomolar concentrations selectively. During the training, the authors included molecules from Neurocrine Biosciences, ChEMBL, and GoStar with known affinity against H1 and compounds structurally similar to the known H1 antagonists from the Enamine REAL Space. This was done to bias the generation of new structures towards the desired profile. Despite being a preprint at the time of our review, to the best of our knowledge, this is the first published work that successfully applied ManyOO, EAs, and ML methods with experimental validation. This highlights the powerful application of Multi/ManyOO strategies in the context of dnDD.

5.2.2 Pareto-based

Recently, Verhellen (2022) evaluated the performance of graph-based implementations of NSGA-II and NSGA-III against a weighted sum method on case studies from the GuacaMol benchmark suite for dnDD Brown et al. (2019) and from datasets constructed to simulate polypharmacology scenarios. The number of optimized objectives ranged from four to five, depending on the benchmark molecule analyzed, including molecular weight, logP, similarity, affinity to a target, and/or blood–brain–barrier permeability. The optimization procedure removed undesired molecules from the population based on structural ADMET filters. Both approaches outperform the weighted sum method in terms of the HV indicator. Regarding efficiency, NSGA-III outperformed NSGA-II by performing fewer function evaluations in all benchmarks. However, both approaches showed similar performance on the analyzed chemical benchmarks. By conducting this comparison, the authors provided valuable insights for effectively applying Multi/ManyOO approaches in dnDD.

Molecules for anti-breast cancer were produced in Mei and Wu (2022) using an enhanced ManyOEA variant, AGM-MOEA Panichella (2019). The proposed method applies the crossover operator from Differential Evolution Price et al. (2005), non-dominated sorting from NSGA-II Deb et al. (2002), and a normalization method from NSGA-III Deb and Jain (2014) for generating and selecting molecules. Six objectives were simultaneously optimized: pIC₅₀ and ADMET properties (Caco-2, CYP3 A4, hERG, HOB, MN), and three performance metrics measured the quality of the PF obtained. To highlight the

importance of adopting a ManyOO approach, the authors demonstrated the conflicting nature of the objectives and emphasized that a simple aggregation of the objectives could be a difficult task. A comparative analysis was conducted between the proposed framework and three other methods: NSGA-II, NSGA-III, and the original AGE-MOEA. The results indicated the superiority of the proposed method regarding the performance metrics.

Cofala et al. (2020) proposed a method that combined a ($\mu + \lambda$) ES and an NSGA-II-based method to design an inhibitor for SARS-CoV-2's main protease. The ManyOEA proposed uses the SELFIES representation for designing new molecules. Five molecular properties were used in the optimization process: binding affinity (BA) computed by the QuickVina2 binding scoring function, QED, SAS, natural product-likeness (NP), and medical chemical filters (MC). Two experiments used the N3 inhibitor and lopinavir, an HIV main protease inhibitor, as ligands targeting the main protease Mpro receptor. First, a SingleOO experiment was conducted, where the fitness function was composed of a weighted sum of the five properties. In preliminary experiments, the authors observed that (i) molecules with high binding scores suffer from low QED scores, and (ii) defining the weights for the objectives could be difficult, concluding that a many-objective approach might be more appropriate. Hence, a many-objective analysis was also conducted, in which the HV metric was used to evaluate the final set of solutions. For the SingleOO experiment, the authors highlighted the conflicting nature of the objectives, which resulted in a trade-off between QED and NP scores *versus* SA and BA scores. Even so, according to the binding scores, the best molecule found in the SingleOO approach achieved better scores than those of N3 and lopinavir. On the other hand, the ManyOO provided a higher diversity of molecules when compared with the SingleOO case, achieving satisfactory values for all properties.

A Generative Adversarial Network (GAN), proposed by Abbasi et al. (2022), combines an autoencoder with a GAN to convert SMILES strings into latent space vectors and use them as real data in GAN training. To generate molecules that exhibit multiple desired properties, an optimization step based on feedbackGAN Gupta and Zou (2019) is applied, incorporating the NSGA-II method to generate non-dominated solutions that will be included in the training set. The case study analyzed aimed to find ligands that bind both to the Kappa Opioid Receptor (KOR) and the A2A receptor (ADORA2A). The other properties to be optimized include the binding affinity pIC₅₀, the molecular topological polar surface area (TPSA), the solubility (LogP), and the SAS. The strategy adopted was evaluated based on several metrics, such as validity, uniqueness and novelty. The proposed framework generated molecules with a high level of diversity (over 0.88) and 100% uniqueness but a low percentage of validly generated molecules (30.2%).

5.3 Objectives or constraints?

In implementing different dnDD strategies, determining which of the desired properties should be considered as objectives or constraints is a very important aspect. Nicolaou et al. (2009) classified their objectives as primary (objectives used to guide the

TABLE 1 Multi-objective and many-objective methods for *de novo* drug design.

Method	Molecular generation	Objectives	Constraints	Constraints handling technique	Multi/ManyOO approach (final result)	REF.
Multi-objective methods						
GANDI	GA ^b and TS ^b	Force field-based binding energy and 2D/3D similarity measure to a known ligand	Chemical validity	Special move operators and elimination of compounds with steric clashes	Aggregation-based (single solution)	Dey and Cafilisch (2008)
MOGA	GA ^b	Oral bio-availability and Tanimoto similarity with known ligands	Synthetic accessibility	Special move operators	Aggregation-based (single solution)	Devi et al. (2014)
MoGADdrug	GA ^b	Oral bio-availability and Tanimoto similarity with known ligands	Synthetic accessibility	Special move operators	Aggregation-based (single solution)	Devi et al. (2021)
CoG	GP ^b	Tanimoto similarity between two compounds	Chemical validity	Special move operators	Pareto-based	Brown et al. (2004)
MEGA	GA-based ^b	Docking score against multiple targets	Tanimoto similarity with known ligands and Ro5	Special move operators	Pareto-based	Nicolaou et al. (2009)
Synopsis	GA ^b	Docking score against multiple targets	RotB, longest sp ³ -sp ³ chain, reactive groups, volume, logP, PSA, molar refractivity, interaction with specific regions of the targets (hinge, backpocket, hydrophobic)	Feasibility rules	Pareto-based	Daeyaert and Deem (2017)
MoMADrug	Monkey Algorithm	Tanimoto similarity with known ligands and Ro5	-	-	Pareto-based	Devi et al. (2019)
MOBifi	Biofilm Algorithm	Tanimoto similarity with known ligands, Ro5 and Veber rules	-	-	Pareto-based	Devi et al. (2020)
DEL	DGM ^b (VAE)	QED ^a , SAS ^a and logP	Chemical validity	Elimination of invalid and duplicated SMILES strings	Pareto-based	Grantham et al. (2022)
DEL-based	DGM ^b (JTVAE)	Affinity prediction, SAS ^a and logP	Chemical validity	Elimination of invalid and duplicated SMILES strings	Pareto-based	Mukaidaisi et al. (2022)
DrugEx v2	DRL ^b	Affinity prediction towards two targets and one off-target	-	-	Pareto-based	Liu et al. (2021)
DrugEx v3	Graph Transformer	QED ^a and affinity prediction to a pre-defined receptor	Chemical validity	Repair technique	Pareto-based	Liu et al. (2023)
Many-objective methods						
EMGA	NLM ^b (Transformer)	Predicted affinity, QED ^a , natural product-likeness, toxicity scores and synthetic accessibility	-	-	Aggregation-based (single solution)	Elend et al. (2022)
ATOM-GMD	DGM ^b (JTVAE)	12 predicted properties, including terms related to efficacy and safety (on and off-target affinities predicted with QSAR models), Ro5 and accessibility scores	-	-	Aggregation-based (non-dominated solutions)	McLoughlin et al. (2023)
NSGAI and NSGAI-based	GA-based ^b	Ranged from four to five (molecular weight, logP, similarity, affinity prediction to a target of interest, and/or blood-brain-barrier permeability)	ADMET filters	Elimination of undesirable compounds	Pareto-based	Verhellen (2022)

(Continued on following page)

TABLE 1 (Continued) Multi-objective and many-objective methods for *de novo* drug design.

Method	Molecular generation	Objectives	Constraints	Constraints handling technique	Multi/ManyOO approach (final result)	REF.
AGM-MOEA	EA-based ^b	Affinity prediction and ADMET properties (Caco-2, CYP3 A4, hERG, HOB, MN)	-	-	Pareto-based	Mei and Wu (2022)
EMOA	EA-based ^b	Affinity prediction, QED ^a , natural product-likeness, toxicity scores and SAS ^a	Toxicity filters	Elimination of toxic molecules	Pareto-based	Cofala et al. (2020)
GAN-based	GAN ^b	Affinity prediction, topological polar surface area (TPSA), LogP and SAS ^a	-	-	Pareto-based	Abbasi et al. (2022)

^aQED, Quantitative estimation of drug-likeness; SAS, Synthetic accessibility score.

^bGA, Genetic Algorithm; TS, Tabu Search; GP, Genetic Programming; EA, Evolutionary Algorithm; DGM, Deep Generative Model; GAN, Generative Adversarial Network; DRL, Deep Reinforcement Learning; NLM, Neural Language Model.

search procedure) and secondary (objectives acting as constraints that can be used as filters to restrict the search space). Properties serving as objectives are typically related to the similarity to a known ligand (e.g., Tanimoto similarity) or binding affinity scores with one or more receptor target(s) of interest. Constraints usually involve descriptors related to pharmacokinetic prediction (e.g., Lipinski's Ro5) and synthetic accessibility. However, it is important to highlight that the researchers define the premises of development, not existing a rule to state which properties must be defined as an objective or a constraint.

Binding affinity scores are widely used as an objective in the optimization procedure to consider the potency of a compound against the therapeutic targets of interest. However, ignoring other information (objectives) important for a lead compound to pass clinical trials and reach the market can affect the quality of the designed molecules Fu et al. (2022). Therefore, several works also include pharmacokinetics-related descriptors in a Multi/ManyOO context, such as QED, logP and SAS.

As previously mentioned, when the number of objectives increases (> 3), the number of Pareto-optimal solutions becomes intractable. Hence, in selecting the number of objectives, the researchers must also consider this issue when choosing an appropriate technique to deal with the different objectives and constraints. To alleviate this drawback, in Daeyaert and Deem (2017), it was proposed to define some objectives as constraints so that the dimension of the problem could be restricted to two or three objectives, allowing the application of MultiOEA as the search procedure. On the other hand, many works adopted a scalarization procedure to aggregate the many objectives into one, using an *a priori* optimization method, thus transforming a multi-objective problem into a single-objective optimization process. In both approaches, detailed knowledge of the problem is required to define the most suitable properties that should be used as objectives or to set the weight vectors to obtain a Pareto-optimal solution (hopefully) in a desired region of the objective space.

It is also crucial to avoid the use of redundant objectives that would increase the complexity of the problem without providing a real increase in accuracy. For example, in the MOBifi Devi et al. (2020), they adopted two correlated objectives for drug-likeness (Veber and Lipinski's rules). Despite these metrics adopts different

parameters, the Veber's rule is proposed as an improved model for oral bioavailability, thus probably not being conflicting with the Lipinski's rule. As already described elsewhere, the MultiOO approaches aim to optimize only conflicting objectives. Therefore, the selection of the objectives can also be guided by dimensionality reduction methods to reduce the number of effective objectives, discarding correlated metrics.

EAs are originally unconstrained search techniques, requiring additional mechanisms to deal with constraints. A simple approach is to discard unwanted solutions during the search process. However, this strategy may not be adequate since the usual assumption is that the first generations of EAs may contain diverse infeasible but still promising solutions. Some examples of alternative approaches to deal with constraints include Mezura-Montes and Coello (2011); Rahimi et al. (2023):

- penalty functions to add a penalty value to the objective function for infeasible solutions;
- move operators and special representation schemes that guarantee the generation of feasible solutions;
- repair techniques to move unfeasible points back to the feasible space;
- feasibility rules for selecting feasible and infeasible individuals; and
- hybrid approaches that combine different strategies to treat constraints

Among those techniques, special move operators and repair techniques are used mainly to guarantee the generation of feasible solutions. For instance, the GANDI Dey and Cafilisch (2008) methodology enforces feasibility by forbidding certain connections on the molecule. CoG Brown et al. (2004) applies special mutation and crossover operations on the graph representation. Synopsis Daeyaert and Deem (2017) enforces feasibility rules in which constraint violations are considered for selecting a new compound. Invalid and toxic molecules are discarded in DEL Grantham et al. (2022) and during the optimization process of EMOA Cofala et al. (2020). A repair mechanism is used in DrugEx v3 Liu et al. (2023) to correct chemically invalid molecules.

Recently, a helpful tool proposed by Schoenmaker et al. (2023) for molecular correction through a generative deep learning method could be used as a repair technique to correct invalid but interesting compounds.

Table 1 summarises the multi-objective and many-objective techniques cited in the previous sections, including their choice of objectives and constraints, the constraint handling techniques applied, and the Multi/ManyOO approach adopted.

It is important to highlight that even considering only objectives, such as machine learning predicted binding affinities, related to diverse targets through many-objectives methodologies in dnDD could *per se* provide a robust framework for designing and optimizing more effective drug candidates. In fact, there is a crescent interest in the development of new compounds targeting multiple targets in a polypharmacology context (*on*-targets), for example, to reduce bacterial resistance, whereas avoiding the interaction with receptors related to side effects (*off*-targets) Zieba et al. (2022); Raghavendra et al. (2018); Wang H. et al. (2021); Dias et al. (2017). Furthermore, by incorporating diverse objectives and constraints related to potency, selectivity, and pharmacokinetics, ManyOO methodologies might enable the design of compounds that exhibit improved overall performance and a higher probability of success in clinical trials.

6 Discussion

The application of multi-objective techniques has become well-established in the field of dnDD. This strategy improves the solutions' quality by considering multiple objectives simultaneously, leading to a more realistic representation of the problem. It is worth noting that a few papers Luukkonen et al. (2023); Liu et al. (2021) emphasize the importance of addressing problems with more than three objectives in diverse areas. However, the proper exploration of ManyOO approaches in dnDD is still scarce despite its significant potential.

The issue of dimensionality in dnDD has been recognized by researchers as an important topic that needs consideration Nicolaou et al. (2009); Devi et al. (2014); Grantham et al. (2022). Despite the use of ManyOEA in Mei and Wu (2022); Verhellen (2022), none of these papers explicitly address the significantly more complex and challenging nature of problems that arise when dealing with more than three objectives. It is relevant to observe that in Cofala et al. (2020), five objectives are employed in the optimization process using the NSGA-II method, a MultiOEA more suitable for handling problems with up to three objectives. As previously mentioned, NSGA-II encounters additional difficulties as the number of objectives increases. In such scenarios, NSGA-III or other ManyOEA could be more appropriate alternatives for handling the problem. Hence, it is important to highlight the following reflections.

1. ManyOO methods still have not received enough attention from the dnDD community. Among all the papers cited herein, only a few of them Elend et al. (2022); McLoughlin et al. (2023); Verhellen (2022); Mei and Wu (2022); Cofala et al. (2020) simultaneously optimized more than three objectives. One is a preprint, and the others were recently published. Although dnDD

research has increased rapidly over the years, its application to the many-objective case seems recent.

2. Although ManyOEA were used in those papers, rarely is there an emphasis on the fact that they are dealing with a much more complex and challenging class of problems compared to cases with up to three objectives.
3. We observed that many papers adopt aggregation-based techniques, most of which apply the weighted sum aggregation approach. It is important to note that other aggregation-based techniques are more efficient and equally simple to implement. The assumption that the optimization problems are always convex to justify the efficient use of weighted sum aggregation is not always valid in a complex and challenging problem such as the dnDD.
4. Some papers considered multi-target evaluations in the context of polypharmacology and/or off-targets, but they are often limited to two or three targets simultaneously. However, there are important contexts in which binding affinity against multiple targets could be considered at the same time, for example, when developing new compounds against infectious diseases (multiple targets for the same pathogen could improve drug efficacy) or in the case of kinase inhibitors (off-targets panel containing dozens of kinases to evaluate selectivity).
5. The main purpose of Multi- and ManyOO methods is to generate trade-off solutions considering conflicting objectives, not to improve the predictions of the objectives independently. For that reason, as in the single-objective problems, one should care about the accuracy of the adopted methods for calculating/predicting the objective values.

Aggregation-based and weight-based approaches are widely used because they simplify the problem by reducing its dimensionality. However, it is important to highlight that methods based on *a priori* aggregation provide only one solution among the several possibilities of existing non-dominated solutions. As stated earlier, using aggregating techniques is the easiest way to approach the problem, but it oversimplifies and fails to inform the user about the trade-offs between the objectives. Besides, with the increase in the number and diversity of objectives, choosing and normalizing weight vectors becomes progressively difficult. Thus, to overcome these issues and take advantage of a more diverse set of non-dominated solutions, we believe many-objective Pareto-based approaches will increase prevalence in dnDD and be the center of novel methodological developments. We also are convinced that, given its intrinsic attributes—namely, the quantity and variety of objectives and constraints, along with the diversity of possible approaches in their prioritization—many-objective dnDD optimization might indeed serve as a source of inspiration for new developments in more fundamental and general methodological frameworks within the field of ManyOO.

Although this paper does not focus on machine learning (ML) techniques, it is important to mention that recent papers have demonstrated a notable increase in the number of works incorporating ML techniques and evolutionary computing algorithms and concepts in their models, particularly in generative models. For those interested in this field, comprehensive reviews of ML techniques for dnDD can be found

in Mouchlis et al. (2021); Bilodeau et al. (2022); Wang et al. (2022); Luukkonen et al. (2023). While few of them briefly mention papers exploring MultiOO, with the majority utilizing MultiOEOs, only the recently published paper Luukkonen et al. (2023) gives special attention to ML methods for multi-objective dnDD. They also list multi-objective ML-based methods, including EAs, reinforcement/conditional learning, and recurrent neural networks. We believe new developments involving ML techniques and ManyOEOs should generate powerful tools for dealing with dnDD.

7 Conclusion

This work presented an overview of MultiOO and ManyOO approaches applied in dnDD, particularly those based on evolutionary computation and machine learning techniques. We provide a general review of the definitions involved in MultiOEOs and ManyOEOs, emphasizing the main challenges that appear when the number of objectives of an optimization problem increases. Our review could trace possible improvements and drawbacks in designing new optimization techniques by examining how the molecular properties are utilized in the dnDD problem to define objectives and constraints.

The increasing interest in applying ManyOO in dnDD is evident, in which evolutionary computation, coupled with ML methods, has continuously strengthened. Those approaches evolve toward efficiently solving the dnDD problem regarding the number of objectives and/or constraints considered. Still, given the multitude and diverse dnDD's characteristics as a ManyOO process, it may serve as a catalyst for new developments in more fundamental and general methodological frameworks within the ManyOO field.

Finally, integrating multi-target drug development and many-objectives optimization approaches has great potential for accelerating the discovery of innovative and more efficacious drug therapies.

Author contributions

JA: Conceptualization, Investigation, Visualization, Writing–original draft, Writing–review and editing. IG:

Conceptualization, Funding acquisition, Investigation, Supervision, Writing–review and editing, Writing–original draft. HB: Funding acquisition, Supervision, Writing–review and editing. LD: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing–original draft, Writing–review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Brazilian agencies Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (grant numbers 301524/2023-8, 309744/2022-9); and the Fundação Carlos Chagas Filho de Apoio à Ciência (FAPERJ) (grant numbers E-26/010.001415/2019, E-26/211.357/2021, E-26/200.393/2023, E-26/200.608/2022, E-26/210.372/2022).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fchem.2023.1288626/full#supplementary-material>

References

- Abbasi, M., Santos, B. P., Pereira, T. C., Sofia, R., Monteiro, N. R. C., Simões, C. J. V., et al. (2022). Designing optimized drug candidates with generative adversarial network. *J. Cheminformatics* 14 (1), 40. doi:10.1186/s13321-022-00623-6
- Allmendinger, R., Jaszkiwicz, A., Liefoghe, A., and Tammer, C. (2022). What if we increase the number of objectives? Theoretical and empirical implications for many-objective combinatorial optimization. *Comput. Operations Res.* 145, 105857. doi:10.1016/j.cor.2022.105857
- Asafuddoula, M., Ray, T., and Sarker, R. (2015). A decomposition-based evolutionary algorithm for many objective optimization. *IEEE Trans. Evol. Comput.* 19 (3), 445–460. doi:10.1109/tevc.2014.2339823
- Bader, J., Deb, K., and Zitzler, E. (2010). "Faster hypervolume-based search using Monte Carlo sampling," in *Multiple criteria decision making for sustainable energy and transportation systems*. Editors M. Ehrgott, B. Naujoks, T. J. Stewart, and J. Wallenius (Berlin, Heidelberg: Springer Berlin Heidelberg), 313–326.
- Bader, J., and Zitzler, E. (2011). HypE: an algorithm for fast hypervolume-based many-objective optimization. *Evol. Comput.* 19 (1), 45–76. doi:10.1162/evco_a_00009
- Bajusz, D., Rácz, A., and Héberger, K. (2015). Why is tanimoto index an appropriate choice for fingerprint-based similarity calculations? *J. Cheminformatics* 7, 20. doi:10.1186/s13321-015-0069-3
- Batista, L. S., Campelo, F., Guimarães, F. G., and Ramirez, J. A. (2011). "Pareto cone ϵ -dominance: improving convergence and diversity in multiobjective evolutionary algorithms," in *Evolutionary multi-criterion optimization*. Editors R. H. C. Takahashi, K. Deb, E. F. Wanner, and S. Greco (Berlin Heidelberg: Springer), 76–90.
- Ben Amor, O., Chelly Dagdia, Z., Bechikh, S., and Ben Said, L. (2022). "Many-objective optimization of wireless sensor network deployment," in *Evolutionary intelligence* (Berlin Heidelberg: Springer).

- Beume, N., Fonseca, C. M., López-Ibáñez, M., Paquete, L., and Vahrenhold, J. (2009). On the complexity of computing the hypervolume indicator. *IEEE Trans. Evol. Comput.* 13 (5), 1075–1082. doi:10.1109/tevc.2009.2015575
- Bilodeau, C., Jin, W., Jaakkola, T., Barzilay, R., and Jensen, K. F. (2022). Generative models for molecular discovery: recent advances and challenges. *WIREs Comput. Mol. Sci.* 12 (5), e1608. doi:10.1002/wcms.1608
- J. Branke, K. Deb, K. Miettinen, and R. Slowinski (Editors) (2008). *Multiobjective optimization: interactive and evolutionary approaches* (Berlin Heidelberg: Springer).
- Brockhoff, D., and Zitzler, E. (2006). “Are all objectives necessary? On dimensionality reduction in evolutionary multiobjective optimization,” in *Parallel problem solving from nature - PPSN IX*. Editors T. P. Runarsson, H.-G. Beyer, E. Burke, J. J. Merelo-uerovs, L. D. Whitley, and X. Yao (Berlin, Heidelberg: Springer), 533–542.
- Brockhoff, D., and Zitzler, E. (2007). “Dimensionality reduction in multiobjective optimization: the minimum objective subset problem,” in *Operations research proceedings 2006*. Editors K.-H. Waldmann and U. M. Stocker (Berlin, Heidelberg: Springer Berlin Heidelberg), 423–429.
- Brown, N. (2009). Chemoinformatics — an introduction for computer scientists. *ACM Comput. Surv.* 41 (2), 1–38. doi:10.1145/1459352.1459353
- Brown, N., Fiscato, M., Segler, M. H., and Vaucher, A. C. (2019). Guacamol: benchmarking models for *de novo* molecular design. *J. Chem. Inf. Model.* 59 (3), 1096–1108. doi:10.1021/acs.jcim.8b00839
- Brown, N., McKay, B., Gilardoni, F., and Gasteiger, J. (2004). A graph-based genetic algorithm and its application to the multiobjective evolution of median molecules. *J. Chem. Inf. Comput. Sci.* 44 (3), 1079–1087. doi:10.1021/ci034290p
- Burke, E. K., De Causmaecker, P., Berge, G. V., and Van Landeghem, H. (2004). The state of the art of nurse rostering. *J. Sched.* 7 (6), 441–499. doi:10.1023/b:bjosh.0000046076.75950.0b
- Coello, C. A. C., Lamont, G. B., and Veldhuizen, D. A. V. (2007). *Evolutionary algorithms for solving multi-objective problems*. New York, NY: Springer.
- Cofala, T., Elend, L., Mirbach, P., Prellberg, J., Teusch, T., and Kramer, O. (2020). “Evolutionary multi-objective design of sars-cov-2 protease inhibitor candidates,” in *Parallel problem solving from nature - PPSN XVI*. Editors T. Bäck, M. Preuss, A. Deutz, H. Wang, C. Doerr, M. Emmerich, et al. (Cham: Springer International Publishing), 357–371.
- Daeyaert, F., and Deem, M. W. (2017). A pareto algorithm for efficient *de novo* design of multi-functional molecules. *Mol. Inf.* 36 (1–2), 1600044. doi:10.1002/minf.201600044
- Dai, C., Wang, Y., and Ye, M. (2014). A new evolutionary algorithm based on contraction method for many-objective optimization problems. *Appl. Math. Comput.* 245, 191–205. doi:10.1016/j.amc.2014.07.069
- Deb, K. (2001). *Multi-objective optimization using evolutionary algorithms*. New Jersey, United States: John Wiley and Sons.
- Deb, K., Fleming, P., Jin, Y., Miettinen, K., and Reed, P. M. (2023). *Key issues in real-world applications of many-objective optimisation and decision analysis*. Cham: Springer International Publishing.
- Deb, K., Gupta, S., Daum, D., Branke, J., Mall, A. K., and Padmanabhan, D. (2009). Reliability-based optimization using evolutionary algorithms. *IEEE Trans. Evol. Comput.* 13 (5), 1054–1074. doi:10.1109/tevc.2009.2014361
- Deb, K., and Jain, H. (2014). An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, Part I: solving problems with box constraints. *IEEE Trans. Evol. Comput.* 18 (4), 577–601. doi:10.1109/tevc.2013.2281535
- Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* 6 (2), 182–197. doi:10.1109/4235.996017
- de Farias, L. R., and Araújo, A. F. (2022). A decomposition-based many-objective evolutionary algorithm updating weights when required. *Swarm Evol. Comput.* 68, 100980. doi:10.1016/j.swevo.2021.100980
- de Freitas, A. R., Fleming, P. J., and Guimarães, F. G. (2015). Aggregation trees for visualization and dimension reduction in many-objective optimization. *Inf. Sci.* 298, 288–314. doi:10.1016/j.ins.2014.11.044
- Devi, R., Sathya, S., Kumar, N., and Coumar, M. (2019). Multi-objective monkey algorithm for drug design. *Int. J. Intelligent Syst. Appl.* 3, 31–41. doi:10.5815/ijisa.2019.03.04
- Devi, R. V., Sathya, S. S., and Coumar, M. S. (2014). Multi-objective genetic algorithm for *de novo* drug design. *Int. J. Soft Comput. Eng. (IJSCE)* 4 (2), 92–96.
- Devi, R. V., Sathya, S. S., and Coumar, M. S. (2015). Evolutionary algorithms for *de novo* drug design – a survey. *Appl. Soft Comput.* 27, 543–552. doi:10.1016/j.asoc.2014.09.042
- Devi, R. V., Siva Sathya, S., and Coumar, M. S. (2020). Multi-objective biofilm algorithm (mobif) for *de novo* drug design with special focus to anti-diabetic drugs. *Appl. Soft Comput.* 96, 106655. doi:10.1016/j.asoc.2020.106655
- Devi, V. R., Sathya, S. S., and Coumar, S. M. (2021). Multi-objective genetic algorithm for *de novo* drug design (MoGADdrug). *Curr. Computer-Aided Drug Des.* 17 (3), 445–457. doi:10.2174/15734099166662006201914143
- Dey, F., and Cafilisch, A. (2008). Fragment-based *de novo* ligand design by multiobjective evolutionary optimization. *J. Chem. Inf. Model.* 48 (3), 679–690. doi:10.1021/ci700424b
- Dias, K. S. T., de Paula, C. T., dos Santos, T., Souza, I. N., Boni, M. S., Guimarães, M. J., et al. (2017). Design, synthesis and evaluation of novel feruloyl-donepezil hybrids as potential multitarget drugs for the treatment of alzheimer’s disease. *Eur. J. Med. Chem.* 130, 440–457. doi:10.1016/j.ejmech.2017.02.043
- Ekins, S., Honeycutt, J. D., and Metz, J. T. (2010). Evolving molecules using multi-objective optimization: applying to ADME/Tox. *Drug Discov. Today* 15 (11), 451–460. doi:10.1016/j.drudis.2010.04.003
- Elend, L., Jacobsen, L., Cofala, T., Prellberg, J., Teusch, T., Kramer, O., et al. (2022). Design of sars-cov-2 main protease inhibitors using artificial intelligence and molecular dynamic simulations. *Molecules* 27 (13), 4020. doi:10.3390/molecules27134020
- Emmerich, M. T. M., and Deutz, A. H. (2018). A tutorial on multiobjective optimization: fundamentals and evolutionary methods. *Nat. Comput.* 17 (3), 585–609. doi:10.1007/s11047-018-9685-y
- Erickson, M., Mayer, A., and Horn, J. (2001). “The Niche Pareto Genetic Algorithm 2 applied to the design of groundwater remediation systems,” in *Evolutionary multi-criterion optimization*. Editors E. Zitzler, L. Thiele, K. Deb, C. A. Coello Coello, and D. Corne (Berlin Heidelberg: Springer), 681–695.
- Falcón-Cardona, J. G., and Coello, C. A. C. (2020). Indicator-based multi-objective evolutionary algorithms: a comprehensive survey. *ACM Comput. Surv.* 53 (2), 1–35. doi:10.1145/3376916
- Fu, T., Gao, W., Coley, C. W., and Sun, J. (2022). “Reinforced genetic algorithm for structure-based drug design,” in 36th Conference on Neural Information Processing Systems (NeurIPS 2022), New Orleans USA, 28 November - 9 December 2022, 1–14.
- Glover, F., and Laguna, M. (1997). *Tabu search*. New York, NY: Springer.
- Grantham, K., Mukaidaisi, M., Ooi, H. K., Ghaemi, M. S., Tchagang, A., and Li, Y. (2022). Deep evolutionary learning for molecular design. *IEEE Comput. Intell. Mag.* 17 (2), 14–28. doi:10.1109/mci.2022.3155308
- Guedes, I. A., Barreto, A. M. S., Marinho, D., Krempser, E., Kuenemann, M. A., Sperandio, O., et al. (2021). New machine learning and physics-based scoring functions for drug discovery. *Sci. Rep.* 11 (1), 3198. doi:10.1038/s41598-021-82410-1
- Guedes, I. A., de Magalhães, C. S., and Dardenne, L. E. (2014). Receptor-ligand molecular docking. *Biophys. Rev.* 6 (1), 75–87. doi:10.1007/s12551-013-0130-2
- Guerreiro, A. P., and Fonseca, C. M. (2018). Computing and updating hypervolume contributions in up to four dimensions. *IEEE Trans. Evol. Comput.* 22 (3), 449–463. doi:10.1109/tevc.2017.2729550
- Gupta, A., and Zou, J. (2019). Feedback gan for dna optimizes protein functions. *Nat. Mach. Intell.* 1 (2), 105–111. doi:10.1038/s42256-019-0017-4
- Handl, J., Kell, D. B., and Knowles, J. (2007). Multiobjective optimization in bioinformatics and computational biology. *IEEE/ACM Trans. Comput. Biol. Bioinforma.* 4 (2), 279–292. doi:10.1109/tcbb.2007.070203
- Holland, J. H. (1962). Outline for a logical theory of adaptive systems. *J. ACM* 9 (3), 297–314. doi:10.1145/321127.321128
- Holland, J. H. (1975). *Adaptation in natural and artificial systems*. United States: The University of Michigan Press.
- Horn, J., Nafpliotis, N., and Goldberg, D. (1994). “A niched Pareto genetic algorithm for multiobjective optimization,” in Proceedings of the First IEEE Conference on Evolutionary Computation. IEEE World Congress on Computational Intelligence, Orlando, Florida, USA, June 27–29, 1994, 82–87.
- Ikeda, K., Kita, H., and Kobayashi, S. (2001). Failure of Pareto-based MOEAs: does non-dominated really mean near to optimal? *Proc. 2001 Congr. Evol. Comput.* 2, 957–962. doi:10.1109/CEC.2001.934293
- Irwin, J. J., and Shoichet, B. K. (2005). ZINC-A free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* 45 (1), 177–182. doi:10.1021/ci049714+
- Ishibuchi, H., Setoguchi, Y., Masuda, H., and Nojima, Y. (2016). “How to compare many-objective algorithms under different settings of population and archive sizes,” in 2016 IEEE Congress on Evolutionary Computation (CEC), British Columbia, Canada, 24–29 July 2016, 1149–1156.
- Ishibuchi, H., Setoguchi, Y., Masuda, H., and Nojima, Y. (2017). Performance of decomposition-based many-objective algorithms strongly depends on pareto front shapes. *IEEE Trans. Evol. Comput.* 21 (2), 169–190. doi:10.1109/tevc.2016.2587749
- Ishibuchi, H., Tsukamoto, N., and Nojima, Y. (2008). “Evolutionary many-objective optimization: a short review,” in 2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence), Hong Kong, June 1–6, 2008, 2419–2426.
- Jain, H., and Deb, K. (2014). An evolutionary many-objective optimization algorithm using reference-point based nondominated sorting approach, Part II: handling constraints and extending to an adaptive approach. *IEEE Trans. Evol. Comput.* 18 (4), 602–622. doi:10.1109/tevc.2013.2281534

- Jana, G., Pal, R., Sural, S., and Chattaraj, P. K. (2020). Quantitative structure-toxicity relationship: an “*in silico* study” using electrophilicity and hydrophobicity as descriptors. *Int. J. Quantum Chem.* 120 (6), e26097. doi:10.1002/qua.26097
- Knowles, J. D., and Corne, D. W. (1999). The Pareto archived evolution strategy: a new baseline algorithm for Pareto multiobjective optimisation. *Proc. 1999 Congr. Evol. Comput.* 1, 98–105.
- Knowles, J. D., and Corne, D. W. (2000). Approximating the nondominated front using the Pareto archived evolution strategy. *Evol. Comput.* 8 (2), 149–172. doi:10.1162/106365600568167
- Lambrinidis, G., and Tsantili-Kakoulidou, A. (2021). Multi-objective optimization methods in novel drug design. *Expert Opin. Drug Discov.* 16 (6), 647–658. doi:10.1080/17460441.2021.1867095
- Laumanns, M., Thiele, L., Deb, K., and Zitzler, E. (2002). Combining convergence and diversity in evolutionary multiobjective optimization. *Evol. Comput.* 10 (3), 263–282. doi:10.1162/106365602760234108
- Le, T. C., and Winkler, D. A. (2015). A bright future for evolutionary methods in drug design. *ChemMedChem* 10 (8), 1296–1300. doi:10.1002/cmdc.201500161
- Li, B., Li, J., Tang, K., and Yao, X. (2015a). Many-objective evolutionary algorithms: a survey. *ACM Comput. Surv.* 48 (1), 1–35. doi:10.1145/2792984
- Li, K., Deb, K., Zhang, Q., and Kwong, S. (2015b). An evolutionary many-objective optimization algorithm based on dominance and decomposition. *IEEE Trans. Evol. Comput.* 19 (5), 694–716. doi:10.1109/tevc.2014.2373386
- Lipinski, C. A. (2004). Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discov. Today Technol.* 1 (4), 337–341. doi:10.1016/j.dttc.2004.11.007
- Liu, J., Wang, Y., Wang, X., Guo, S., and Sui, X. (2019a). A new dominance method based on expanding dominated area for many-objective optimization. *Int. J. Pattern Recognit. Artif. Intell.* 33 (03), 1959008. doi:10.1142/s0218001419590080
- Liu, X., Ye, K., van Vlijmen, H. W. T., Emmerich, M. T. M., Ijzerman, A. P., and van Westen, G. J. P. (2021). DrugEx v2: *de novo* design of drug molecules by Pareto-based multi-objective reinforcement learning in polypharmacology. *J. Cheminformatics* 13, 85. doi:10.1186/s13321-021-00561-9
- Liu, X., Ye, K., van Vlijmen, H. W. T., Ijzerman, A. P., and van Westen, G. J. P. (2019b). An exploration strategy improves the diversity of *de novo* ligands using deep reinforcement learning: a case for the adenosine A_{2A} receptor. *J. Cheminformatics* 11 (1), 35. doi:10.1186/s13321-019-0355-6
- Liu, X., Ye, K., van Vlijmen, H. W. T., Ijzerman, A. P., and van Westen, G. J. P. (2023). DrugEx v3: scaffold-constrained drug design with graph transformer-based reinforcement learning. *J. Cheminformatics* 15 (1), 24. doi:10.1186/s13321-023-00694-z
- Loving, K., Alberts, I., and Sherman, W. (2010). Computational approaches for fragment-based and *de novo* design. *Curr. Top. Med. Chem.* 10 (1), 14–32. doi:10.2174/156802610790232305
- Luukkonen, S., van den Maagdenberg, H. W., Emmerich, M. T., and van Westen, G. J. (2023). Artificial intelligence in multi-objective drug design. *Curr. Opin. Struct. Biol.* 79, 102537. doi:10.1016/j.sbi.2023.102537
- Mane, S. U., and Narasinga Rao, M. (2017). Many-objective optimization: problems and evolutionary algorithms - a short review. *Int. J. Appl. Eng. Res.* 12 (20), 9774–9793.
- McLoughlin, K. S., Shi, D., Mast, J. E., Bucci, J., Williams, J. P., Jones, W. D., et al. (2023). *Generative molecular design and experimental validation of selective histamine h1 inhibitors*. bioRxiv.
- Mei, Y., and Wu, K. (2022). Application of multi-objective optimization in the study of anti-breast cancer candidate drugs. *Sci. Rep.* 12 (1), 19347. doi:10.1038/s41598-022-23851-0
- Mezura-Montes, E., and Coello, C. A. C. (2011). Constraint-handling in nature-inspired numerical optimization: past, present and future. *Swarm Evol. Comput.* 1 (4), 173–194. doi:10.1016/j.swevo.2011.10.001
- Miettinen, K. (1999). *Nonlinear multiobjective optimization*. Norwell, Massachusetts: Kluwer Academic.
- Mouchlis, V. D., Afantitis, A., Serra, A., Fratello, M., Papadiamantis, A. G., Aidinis, V., et al. (2021). Advances in *de novo* drug design: from conventional to machine learning methods. *Int. J. Mol. Sci.* 22 (4), 1676. doi:10.3390/ijms22041676
- Mukaidaisi, M., Vu, A., Grantham, K., Tchagang, A., and Li, Y. (2022). Multi-objective drug design based on graph-fragment molecular representation and deep evolutionary learning. *Front. Pharmacol.* 13, 920747. doi:10.3389/fphar.2022.920747
- Nicolaou, A. C., Kannas, C., and Loizidou, E. (2012). Multi-objective optimization methods in *de novo* drug design. *Mini-Reviews Med. Chem.* 12 (10), 979–987. doi:10.2174/138955712802762284
- Nicolaou, C. A., Apostolakis, J., and Pattichis, C. S. (2009). *De novo* drug design using multiobjective evolutionary graphs. *J. Chem. Inf. Model.* 49 (2), 295–307. doi:10.1021/ci800308h
- Nicolotti, O., Giangreco, I., Introcaso, A., Leonetti, F., Stefanachi, A., and Carotti, A. (2011). Strategies of multi-objective optimization in drug discovery and development. *Expert Opin. Drug Discov.* 6 (9), 871–884. doi:10.1517/17460441.2011.588696
- Nishibata, Y., and Itai, A. (1991). Automatic creation of drug candidate structures based on receptor structure. starting point for artificial lead generation. *Tetrahedron* 47 (43), 8985–8990. doi:10.1016/s0040-4020(01)86503-0
- Olivecrona, M., Blaschke, T., Engkvist, O., and Chen, H. (2017). Molecular *de-novo* design through deep reinforcement learning. *J. Cheminformatics* 9 (1), 48. doi:10.1186/s13321-017-0235-x
- Panichella, A. (2019). “An adaptive evolutionary algorithm based on non-euclidean geometry for many-objective optimization.” in Proceedings of the Genetic and evolutionary computation conference, GECCO '19 (New York, NY, USA: Association for Computing Machinery). page 595–603.
- Pareto, V. (1964). *Cours d'Economie politique*. Genève: Librairie Droz.
- Parikh, P. K., Savjani, J. K., Gajjar, A. K., and Chhabria, M. T. (2023). *Bioinformatics and cheminformatics tools in early drug discovery*. United States: John Wiley and Sons, Ltd, 147–181.
- Praditwong, K., and Yao, X. (2006). A new multi-objective evolutionary optimisation algorithm: the two-archive algorithm. *Int. Conf. Comput. Intell. Secur.* 1, 286–291. doi:10.1109/ICCIAS.2006.294139
- Price, K., Storn, R. M., and Lampinen, J. A. (2005). *Differential evolution: a practical approach to global optimization (natural computing series)*. Berlin, Heidelberg: Springer-Verlag.
- Purshouse, R. C., and Fleming, P. J. (2007). On the evolutionary optimization of many conflicting objectives. *IEEE Trans. Evol. Comput.* 11 (6), 770–784. doi:10.1109/tevc.2007.910138
- Raghavendra, N. M., Pingili, D., Kadasi, S., Mettu, A., and Prasad, S. (2018). Dual or multi-targeting inhibitors: the next generation anticancer agents. *Eur. J. Med. Chem.* 143, 1277–1300. doi:10.1016/j.ejmech.2017.10.021
- Rahimi, I., Gandomi, A. H., Chen, F., and Mezura-Montes, E. (2023). A review on constraint handling techniques for population-based algorithms: from single-objective to multi-objective optimization. *Archives Comput. Methods Eng.* 30 (3), 2181–2209. doi:10.1007/s11831-022-09859-9
- Rocha, G. K., Angelo, J. S., Santos, K. B., Custódio, F. L., Dardenne, L. E., and Barbosa, H. J. (2017). “Using an aggregation tree to arrange energy function terms for protein structure prediction,” in 2017 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), United Kingdom, August 23–25, 2017, 1–7.
- Rosenthal, S., and Borschbach, M. (2017). “Design perspectives of an evolutionary process for multi-objective molecular optimization,” in *Evolutionary multi-criterion optimization*. Editors H. Trautmann, G. Rudolph, K. Klamroth, O. Schütze, M. Wiecek, Y. Jin, et al. (Cham: Springer International Publishing), 529–544.
- Sabe, V. T., Ntombela, T., Jhamba, L. A., Maguire, G. E., Govender, T., Naicker, T., et al. (2021). Current trends in computer aided drug design and a highlight of drugs discovered via computational techniques: a review. *Eur. J. Med. Chem.* 224, 113705. doi:10.1016/j.ejmech.2021.113705
- Safi, H. H., Ucan, O. N., and Bayat, O. (2018). “On the real world applications of many-objective evolutionary algorithms,” in Proceedings of the first international Conference on data science, E-Learning and information systems, DATA '18 (New York, NY, USA: Association for Computing Machinery). page 6.
- Sanchez-Lengeling, B., Outeiral, C., Guimaraes, G., and Aspuru-Guzik, A. (2017). *Optimizing distributions over molecular space. An objective-reinforced generative adversarial network for inverse-design chemistry (ORGANIC)*. ChemRxiv.
- Santos, K. B., Guedes, I. A., Karl, A. L. M., and Dardenne, L. E. (2020). Highly flexible ligand docking: benchmarking of the dockthor program on the leads-pep protein-peptide data set. *J. Chem. Inf. Model.* 60 (2), 667–683. doi:10.1021/acs.jcim.9b00905
- Sato, H., and Ishibuchi, H. (2023). Evolutionary many-objective optimization: difficulties, approaches, and discussions. *IEEJ Trans. Electr. Electron. Eng.* 18 (7), 1048–1058. doi:10.1002/tee.23796
- Saxena, D. K., Duro, J. A. A., Tiwari, A., Deb, K., and Zhang, Q. (2013). Objective reduction in many-objective optimization: linear and nonlinear algorithms. *IEEE Trans. Evol. Comput.* 17 (1), 77–99. doi:10.1109/tevc.2012.2185847
- Schoenmaker, L., Béguignon, O. J. M., Jespers, W., and van Westen, G. J. P. (2023). Uncorrupt SMILES: a novel approach to *de novo* design. *J. Cheminformatics* 15 (1), 22. doi:10.1186/s13321-023-00696-x
- Slowik, A., and Kwasnicka, H. (2020). Evolutionary algorithms and their applications to engineering problems. *Neural Comput. Appl.* 32 (16), 12363–12379. doi:10.1007/s00521-020-04832-8
- Socha, B. N., Pandya, S. B., Patel, U. H., Patel, R. H., Bhatt, B. S., Bhakhar, S., et al. (2023). 1-D MOF [Ag₂(C₁₀H₁₀N₃O₃S)₂(C₄H₈N)₂]_n: photocatalytic treatment, crystallographic evaluation, ADMET parameters, CT-DNA and anticancer activity. *J. Biomol. Struct. Dyn.* 2023, 1–16. doi:10.1080/07391102.2023.2236732
- Srinivas, N., and Deb, K. (1994). Multiobjective optimization using nondominated sorting in genetic algorithms. *Evol. Comput.* 2 (3), 221–248. doi:10.1162/evco.1994.2.3.221

- Sülflow, A., Drechsler, N., and Drechsler, R. (2007). "Robust multi-objective optimization in high dimensional spaces," in *Evolutionary multi-criterion optimization*. Editors S. Obayashi, K. Deb, C. Poloni, T. Hiroyasu, and T. Murata (Berlin, Heidelberg: Springer Berlin Heidelberg), 715–726.
- Trivedi, A., Srinivasan, D., Sanyal, K., and Ghosh, A. (2017). A survey of multiobjective evolutionary algorithms based on decomposition. *IEEE Trans. Evol. Comput.* 21 (3), 440–462. doi:10.1109/TEVC.2016.2608507
- Trott, O., and Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31 (2), 455–461. doi:10.1002/jcc.21334
- Ursem, R. K., and Justesen, P. D. (2012). Multi-objective distinct candidates optimization: locating a few highly different solutions in a circuit component sizing problem. *Appl. Soft Comput.* 12 (1), 255–265. doi:10.1016/j.asoc.2011.08.048
- Veber, D. F., Johnson, S. R., Cheng, H.-Y., Smith, B. R., Ward, K. W., and Kopple, K. D. (2002). Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* 45 (12), 2615–2623. doi:10.1021/jm020017n
- Verel, S., Liefvooghe, A., Jourdan, L., and Dhaenens, C. (2013). On the structure of multiobjective combinatorial search space: MNK-landscapes with correlated objectives. *Eur. J. Operational Res.* 227 (2), 331–342. doi:10.1016/j.ejor.2012.12.019
- Verhellen, J. (2022). Graph-based molecular Pareto optimisation. *Chem. Sci.* 13, 7526–7535. doi:10.1039/d2sc00821a
- Vinkers, H. M., de Jonge, M. R., Daeyaert, F. F. D., Heeres, J., Koymans, L. M. H., van Lenthe, J. H., et al. (2003). Synopsis: synthesize and optimize system *in silico*. *J. Med. Chem.* 46 (13), 2765–2773. doi:10.1021/jm030809x
- von Lücken, C., Brizuela, C., and Barán, B. (2019). An overview on evolutionary algorithms for many-objective optimization problems. *WIREs Data Min. Knowl. Discov.* 9 (1), e1267. doi:10.1002/widm.1267
- Wang, H., Jiao, L., and Yao, X. (2015). Two_arch2: an improved two-archive algorithm for many-objective optimization. *IEEE Trans. Evol. Comput.* 19 (4), 524–541. doi:10.1109/tevc.2014.2350987
- Wang, H., Wang, M., Xu, X., Gao, P., Xu, Z., Zhang, Q., et al. (2021a). Multi-target mode of action of silver against staphylococcus aureus endows it with capability to combat antibiotic resistance. *Nat. Commun.* 12 (1), 3331. doi:10.1038/s41467-021-23659-y
- Wang, M., Wang, Z., Sun, H., Wang, J., Shen, C., Weng, G., et al. (2022). Deep learning approaches for *de novo* drug design: an overview. *Curr. Opin. Struct. Biol.* 72, 135–144. doi:10.1016/j.sbi.2021.10.001
- Wang, Z., Chen, J., and Hong, H. (2021b). Developing QSAR models with defined applicability domains on PPAR γ binding affinity using large data sets and machine learning algorithms. *Environ. Sci. Technol.* 55 (10), 6857–6866. doi:10.1021/acs.est.0c07040
- Xia, Y., Huang, J., Li, X., Liu, Y., Zheng, J., and Zou, J. (2023). A many-objective evolutionary algorithm based on indicator and decomposition. *Mathematics* 11 (2), 413. doi:10.3390/math11020413
- Xu, Q., Xu, Z., and Ma, T. (2020). A survey of multiobjective evolutionary algorithms based on decomposition: variants, challenges and future directions. *IEEE Access* 8, 41588–41614. doi:10.1109/access.2020.2973670
- Yuan, J., Liu, H.-L., Gu, F., Zhang, Q., and He, Z. (2021). Investigating the properties of indicators and an evolutionary many-objective algorithm using promising regions. *IEEE Trans. Evol. Comput.* 25 (1), 75–86. doi:10.1109/tevc.2020.2999100
- Yuan, Y., Ong, Y.-S., Gupta, A., and Xu, H. (2018). Objective reduction in many-objective optimization: evolutionary multiobjective approaches and comprehensive analysis. *IEEE Trans. Evol. Comput.* 22 (2), 189–210. doi:10.1109/tevc.2017.2672668
- Yuan, Y., Pei, J., and Lai, L. (2011). Ligbuilder 2: a practical *de novo* drug design approach. *J. Chem. Inf. Model.* 51 (5), 1083–1091. doi:10.1021/ci100350u
- Yuan, Y., Xu, H., Wang, B., and Yao, X. (2016). A new dominance relation-based evolutionary algorithm for many-objective optimization. *IEEE Trans. Evol. Comput.* 20 (1), 16–37. doi:10.1109/tevc.2015.2420112
- Zhang, Q., and Li, H. (2007). MOEA/D: a multiobjective evolutionary algorithm based on decomposition. *IEEE Trans. Evol. Comput.* 11 (6), 712–731. doi:10.1109/tevc.2007.892759
- Zhang, Q., Zhou, A., Zhao, S., Suganthan, P., Liu, W., and Tiwari, S. (2008). "Multiobjective optimization test instances for the CEC 2009 special session and competition," in *Mechanical engineering* (New York: AMSE International).
- Zhang, X., Guo, Y., Yang, J., Li, D., Wang, Y., and Zhao, R. (2022). Many-objective evolutionary algorithm based agricultural mobile robot route planning. *Comput. Electron. Agric.* 200, 107274. doi:10.1016/j.compag.2022.107274
- Zhu, S., Xu, L., Goodman, E., Deb, K., and Lu, Z. (2022). A general framework for enhancing relaxed Pareto dominance methods in evolutionary many-objective optimization. *Nat. Comput.* 22, 287–313. doi:10.1007/s11047-022-09889-z
- Zieba, A., Stępnicki, P., Matosiuk, D., and Kaczor, A. A. (2022). What are the challenges with multi-targeted drug design for complex diseases? *Expert Opin. Drug Discov.* 17 (7), 673–683. doi:10.1080/17460441.2022.2072827
- Zitzler, E., and Künzli, S. (2004). "Indicator-based selection in multiobjective search," in *Parallel problem solving from nature - PPSN VIII*. Editors X. Yao, E. K. Burke, J. A. Lozano, J. Smith, J. J. Merelo-Guervós, J. A. Bullinaria, et al. (Berlin Heidelberg: Springer), 832–842.
- Zitzler, E., Laumanns, M., and Thiele, L. (2001). *SPEA2: improving the strength Pareto evolutionary algorithm*. Switzerland: ETH Zurich, Computer Engineering and Networks Laboratory. Technical report.
- Zitzler, E., and Thiele, L. (1998). *An evolutionary algorithm for multiobjective optimization: the strength Pareto approach*. Switzerland: ETH Zurich, Computer Engineering and Networks Laboratory. Technical report.
- Zitzler, E., and Thiele, L. (1999). Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach. *IEEE Trans. Evol. Comput.* 3 (4), 257–271. doi:10.1109/4235.797969
- Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C., and da Fonseca, V. (2003). Performance assessment of multiobjective optimizers: an analysis and review. *IEEE Trans. Evol. Comput.* 7 (2), 117–132. doi:10.1109/tevc.2003.810758
- Zou, J., Zhang, Z., Zheng, J., and Yang, S. (2021). A many-objective evolutionary algorithm based on dominance and decomposition with reference point adaptation. *Knowledge-Based Syst.* 231, 107392. doi:10.1016/j.knsys.2021.107392