GECCO 2023 Tutorial on Benchmarking Multiobjective Optimizers 2.0

Dimo Brockhoff
dimo.brockhoff@inria.fr

Tea Tušar
tea.tusar@ijs.si

The final slides will be made available at
http://www.cmap.polytechnique.fr/~dimobrockhoff/

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
Our plan

- Discuss history, present and future of multiobjective benchmarking
  
With respect to different topics

- Performance assessment / methodology
- Test functions

Finally, recommendations on good algorithms
Disclaimer

This is not an introductory tutorial to multiobjective optimization!

We assume you know basic definitions like

- Objective function
- Pareto dominance/Pareto front/Pareto set
- Ideal/Nadir points
We only consider continuous search spaces
Disclaimer II

We only consider continuous search spaces

We only consider unconstrained problems
Disclaimer II

We only consider continuous search spaces

We only consider unconstrained problems

What we present is highly subjective & selective
  • How important do we find each milestone?
  • Use version numbering and branches
  • What have we learned from the past?
Overview

1 Performance Assessment

2 Test Problems and Their Visualizations

3 Recommendations from Numerical Results
v0.0.1alpha

Performance Assessment
In the Early Beginnings...

... multiobjective EAs were mainly compared visually:

ZDT6 benchmark problem: IBEA, SPEA2, NSGA-II
v0.1beta

Performance Assessment
## Tables

<table>
<thead>
<tr>
<th>Problem</th>
<th>MOCSA</th>
<th></th>
<th></th>
<th>NSGA2</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>((d))</td>
<td>S</td>
<td>CD</td>
<td>ER</td>
<td>((d))</td>
<td>S</td>
</tr>
<tr>
<td>ZDT1</td>
<td>0.0404</td>
<td>0.0055</td>
<td>0.0000</td>
<td>0</td>
<td>0.0270</td>
<td>0.0156</td>
</tr>
<tr>
<td>ZDT2</td>
<td>0.0404</td>
<td>0.0082</td>
<td>0.0000</td>
<td>0</td>
<td>0.0292</td>
<td>0.0146</td>
</tr>
<tr>
<td>ZDT3</td>
<td>0.0438</td>
<td>0.0148</td>
<td>0.0001</td>
<td>0</td>
<td>0.0329</td>
<td>0.0201</td>
</tr>
<tr>
<td>ZDT4</td>
<td>0.0404</td>
<td>0.0097</td>
<td>0.0000</td>
<td>0</td>
<td>0.0328</td>
<td>0.0159</td>
</tr>
<tr>
<td>ZDT6</td>
<td>0.0327</td>
<td>0.0150</td>
<td>0.0000</td>
<td>0</td>
<td>0.0216</td>
<td>0.0119</td>
</tr>
<tr>
<td>DTLZ1</td>
<td>0.1114</td>
<td>0.0068</td>
<td>0.0000</td>
<td>0</td>
<td>0.0615</td>
<td>0.0319</td>
</tr>
<tr>
<td>DTLZ2</td>
<td>0.2319</td>
<td>0.0646</td>
<td>0.0021</td>
<td>0.02</td>
<td>0.1361</td>
<td>0.0683</td>
</tr>
<tr>
<td>DTLZ3</td>
<td>0.2770</td>
<td>0.0225</td>
<td>0.0000</td>
<td>0</td>
<td>0.1139</td>
<td>0.0739</td>
</tr>
<tr>
<td>DTLZ4</td>
<td>0.2478</td>
<td>0.0424</td>
<td>0.0009</td>
<td>0</td>
<td>0.1630</td>
<td>0.0898</td>
</tr>
<tr>
<td>DTLZ5</td>
<td>0.0487</td>
<td>0.0059</td>
<td>0.0000</td>
<td>0</td>
<td>0.0309</td>
<td>0.0176</td>
</tr>
<tr>
<td>DTLZ6</td>
<td>0.0484</td>
<td>0.0156</td>
<td>0.0000</td>
<td>0</td>
<td>0.0306</td>
<td>0.0135</td>
</tr>
<tr>
<td>DTLZ7</td>
<td>0.2897</td>
<td>0.0510</td>
<td>0.0011</td>
<td>0.04</td>
<td>0.1880</td>
<td>0.1322</td>
</tr>
</tbody>
</table>

arXiv, 2012
Numbers have their value. But not only tables, please!
v1.0

Performance Assessment
v1.0: Two Approaches for Empirical Studies

**Attainment function approach**

[Fonseca and Fleming 1996]

- Applies statistical tests directly to the approximation set
- Detailed information about how and where performance differences occur

**Quality indicator approach**

- Reduces each approximation set to a single quality value
- Applies statistical tests to the quality values

<table>
<thead>
<tr>
<th>Indicator</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypervolume indicator</td>
<td>6.3431</td>
<td>7.1924</td>
</tr>
<tr>
<td>e-indicator</td>
<td>1.2090</td>
<td>0.12722</td>
</tr>
<tr>
<td>$R_2$ indicator</td>
<td>0.2434</td>
<td>0.1643</td>
</tr>
<tr>
<td>$R_3$ indicator</td>
<td>0.6454</td>
<td>0.3475</td>
</tr>
</tbody>
</table>

see e.g. [Zitzler et al. 2003]
Empirical Attainment Functions

© Manuel López-Ibáñez
[López-Ibáñez et al. 2010]
Empirical Attainment Functions

© Manuel López-Ibáñez
[López-Ibáñez et al. 2010]
Empirical Attainment Functions

© Manuel López-Ibáñez
[López-Ibáñez et al. 2010]
Empirical Attainment Functions

© Manuel López-Ibáñez
[López-Ibáñez et al. 2010]
Empirical Attainment Functions

© Manuel López-Ibáñez
[López-Ibáñez et al. 2010]
Empirical Attainment Functions

© Manuel López-Ibáñez
[López-Ibáñez et al. 2010]
Empirical Attainment Functions: Definition

The Empirical Attainment Function $\alpha(z)$ "counts" how many solution sets $\mathcal{X}_i$ attain or dominate a vector $z$ at time $T$:

$$\alpha_T(z) = \frac{1}{N} \sum_{i=1}^{N} 1\{\mathcal{X}_i \preceq_T z\}$$

with $\preceq_T$ being the weak dominance relation between a solution set and an objective vector at time $T$. 
Empirical Attainment Functions: Definition

The Empirical Attainment Function $\alpha(z)$ "counts" how many solution sets $X_i$ attain or dominate a vector $z$ at time $T$:

$$\alpha_T(z) = \frac{1}{N} \sum_{i=1}^{N} 1\{x_i \preceq_T z\}$$

with $\preceq_T$ being the weak dominance relation between a solution set and an objective vector at time $T$.

Note that $\alpha_T(z)$ is the empirical cumulative distribution function of the achieved objective function distribution at time $T$ in the single-objective case ("fixed budget scenario").
Empirical Attainment Functions in Practice

latest implementation online at
http://eden.dei.uc.pt/~cmfonsec/software.html

R package: http://lopez-ibanez.eu/eaftools

see also [López-Ibáñez et al. 2010, Fonseca et al. 2011]
Quality Indicator Approach

Idea

• Transfer multiobjective problem into a set problem
• Define an objective function ("unary quality indicator") on sets
• Use the resulting total (pre-)order (on the quality values)
Quality Indicator Approach

Idea

• Transfer multiobjective problem into a set problem
• Define an objective function (“unary quality indicator”) on sets
• Use the resulting total (pre-)order (on the quality values)

Underlying dominance relation should be reflected!

\[ A \preceq B : \iff \forall b \in B \exists a \in A \ a \preceq b \]
Monotonicity and Strict Monotonicity

Monotonicity when quality indicator does not contradict relation

\[ A \preceq B \Rightarrow I(A) \geq I(B) \]
Monotonicity and Strict Monotonicity

Monotonicity when quality indicator does not contradict relation

\[ A \preceq B \Rightarrow I(A) \geq I(B) \]

Strict monotonicity: better = higher indicator

\[ A \preceq B \text{ and } A \neq B \Rightarrow I(A) > I(B) \]
Example: Refinements Using Indicators

\[ I(A) = \text{volume of the weakly dominated area in objective space} \]

\[ I(A,R) = \text{how much } A \text{ needs to be moved to weakly dominate } R \]

- **Unary hypervolume indicator**
  - Strictly monotone
  - [Zitzler and Thiele 1989]
  - [Zitzler et al. 2007]

- **Unary epsilon indicator**
  - Monotone
  - [Zitzler et al. 2003]
v1.0.1 – v1.0.100 and counting

Performance Assessment
Many Indicators Available

22 indicators

Performance Assessment of Multiobjective Optimizers: An Analysis and Review

Eckart Zitzler¹, Lothar Thiele¹, Marco Laumanns¹, Carlos M. Fonseca², and Viviane Gruenert da Fonseca²

¹ Computer Engineering and Networks Laboratory (TIK)
Department of Information Technology and Electrical Engineering
Swiss Federal Institute of Technology (ETH) Zurich, Switzerland
Email: {zitzler, thiele, laumanns}@tik.ee.ethz.ch

²ADEEC and ISR (Coimbra)
Faculty of Sciences and Technology
University of Algarve, Portugal
Email: cmfonseca@ualg.pt, vgruenert@esi.fct.ualg.pt

[Zitzler et al. 2003]
Even More Indicators...

Performance indicators in multiobjective optimization

Charles Audet\textsuperscript{a}, Jean Bigeon\textsuperscript{b}, Dominique Cartier\textsuperscript{c}, Sébastien Le Digabel\textsuperscript{d}, Ludovic Salomons\textsuperscript{e,1}

\textsuperscript{a}GERAD and Département de mathématiques et génie industriel, École Polytechnique de Montréal, C.P. 6079, Succ. Centre-ville, Montréal, Québec, H3C 3A7, Canada.
\textsuperscript{b}Univ. Grenoble Alpes, CNRS, Grenoble INP, G-SCOP, 98000 Grenoble, France.
\textsuperscript{c}Collège de Maisonneuve, 3800 Rue Sherbrooke E, Montréal, Québec, H1X 2A2, Canada.

[Audet et al 2021]

Quality Evaluation of Solution Sets in Multiobjective Optimisation: A Survey

Miqing Li and Xin Yao\textsuperscript{1}

\textsuperscript{1}CERCIA, School of Computer Science, University of Birmingham, Birmingham B15 2TT, U. K.
\textsuperscript{*Email: limitising@gmail.com, x.yao@cs.bham.ac.uk

[Li and Yao 2019]

63 indicators

100 indicators
Many Indicators: What Do We Do?

Focus on indicators which are (strictly) monotone

- All hypervolume-based indicators [Zitzler et al. 2007]
- Unary epsilon indicator [Zitzler et al. 2003]
- R2 [Hansen and Jaszkiewicz 1998]
- IGD+ [Ishibuchi et al. 2015]
v2.0

Performance Assessment
With the right (strictly) monotone indicator, multiobjective optimization is not different from single-objective optimization (!)

We can use our normal tools from single-objective optimization, including

• Reporting of target-based runtimes
• ECDFs of runtimes, performance profiles, data profiles
• Statistical tests, box plots, ...

see for example [Hansen et al. 2021]
Convergence graphs is all we have to start with...
Measuring Performance Empirically

Convergence graphs is all we have to start with...

Advantage of the fixed target view:
Ratio scale (interpretation of #fevals easier than for f-values)
ECDF

Empirical Cumulative Distribution Function of the Runtime
A Convergence Graph
First Hitting Time is Monotonous
15 Runs

[Graph showing the quality indicator over log₁₀(function evaluations).]
15 Runs ≤ 15 Runtime Data Points
Empirical Cumulative Distribution

The **ECDF** of run lengths to reach the target

- Has for each data point a **vertical step of constant size**
- Displays for each x-value (budget) the count of observations to the left (first hitting times)
Aggregation

15 runs
Aggregation

15 runs
50 targets
Aggregation

15 runs

50 targets
Aggregation

15 runs
50 targets
ECDF with 750 steps
Aggregation

50 targets from 15 runs integrated in a single graph
Interpretation

50 targets from 15 runs integrated in a single graph

area over the ECDF curve

= average log runtime

(or geometric avg. runtime)

over all targets (difficult and easy) and all runs
Worth to Note

ECDF graphs

• Should never aggregate over dimension

Dimension is input parameter to algorithm
Worth to Note

ECDF graphs

• Should never aggregate over dimension
  
  Dimension is input parameter to algorithm

• But often over targets and functions

• Can show data of more than 1 algorithm at a time
Worth to Note

ECDF graphs

• Should never aggregate over dimension
  Dimension is input parameter to algorithm

• But often over targets and functions

• Can show data of more than 1 algorithm at a time

• Are an extension of data profiles
  • Introduced by Moré and Wild for single and relative targets
  • But here for multiple and absolute targets

[More and Wild 2009]
Worth to Note

ECDF graphs
• Should never aggregate over dimension
  Dimension is input parameter to algorithm
• But often over targets and functions
• Can show data of more than 1 algorithm at a time
• Are an extension of data profiles
  • Introduced by Moré and Wild for single and relative targets [Moré and Wild 2009]
  • But here for multiple and absolute targets
• Are COCO’s main performance visualization tool
Example ECDF (later more)
Mostly Overlooked: Scaling with Dimension

In single-objective optimization:
• Scaling behavior mandatory to investigate
Mostly Overlooked: Scaling with Dimension

In single-objective optimization:
• Scaling behavior mandatory to investigate

In multiobjective optimization:
• Actually two dimensions: search and objective space
Mostly Overlooked: Scaling with Dimension

In single-objective optimization:
• Scaling behavior mandatory to investigate

In multiobjective optimization:
• Actually two dimensions: search and objective space
• But former almost never looked at right now 😞

~10 papers from EMO’21 and PPSN/GECCO/CEC’20 change dimension but 50+ papers have a “fixed” dimension
Mostly Overlooked: Scaling with Dimension

In single-objective optimization:
• Scaling behavior mandatory to investigate

In multiobjective optimization:
• Actually two dimensions: search and objective space
• But former almost never looked at right now 😞

• But in practice search space scalability almost more important

~10 papers from EMO’21 and PPSN/GECCO/CEC’20 change dimension but 50+ papers have a “fixed” dimension

Number of objectives often fixed
A Few General Recommendations

- Always display everything you have
- Look at single runs
- Do each experiment at least twice

(= look at the variance of your results)
A Few General Recommendations

• Always **display everything** you have
• Look at **single runs**
• Do each experiment **at least twice**

  (= look at the *variance* of your results)

• As quality indicators, use hypervolume, R2, or epsilon indicator

  or any indicator which is at least monotone
A Few General Recommendations

• Always display everything you have
• Look at single runs
• Do each experiment at least twice
  (= look at the variance of your results)

• As quality indicators, use hypervolume, R2, or epsilon indicator
  or any indicator which is at least monotone

• See also the tutorial slides by Nikolaus Hansen on this topic (not restricted to single-objective optimization!)

## Recommended Experimental Setup (w/ or w/o COCO)

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Benchmarking Experiment</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Postprocessing</td>
<td>python -m cocopp resultfolder/ ALG2 ALG3</td>
</tr>
<tr>
<td>4</td>
<td>Displaying and Discussing Summary Results</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Investigating and Discussing Complementary Results</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Processed Data Sharing</td>
<td>Provide html output somewhere</td>
</tr>
<tr>
<td>7</td>
<td>Raw Data Sharing</td>
<td>Easy with COCO archive module &amp; through issue tracker</td>
</tr>
</tbody>
</table>
Overview

1. Performance Assessment

2. Test Problems and Their Visualizations

3. Recommendations from Numerical Results
Test Problems and Their Visualizations

Introduction

Test Problems (1)

Artificial problems (continuous and unconstrained)

- **v0.1**: Individual problems
- **v0.2**: MOP suite (unscalable problems)
- **v0.5**: ZDT suite (scalable number of variables)
- **v1.0**: DTLZ suite (scalable number of variables and objectives)
- **v1.2**: WFG suite
- **v1.3**: Other suites with a bottom-up construction
- **v1.5**: Suites of distance-based problems
- **v2.0**: The bbo-biobj(-ext) suite
Test Problems and Their Visualizations

Visualization of multiobjective landscapes

Low-dimensional search spaces
- Dominance ratio
- Local dominance
- Gradient path length

Any-dimensional search spaces
- Line cuts

Some notes on problem properties
Test Problems and Their Visualizations

Test Problems (2)

Artificial problems (other)

Constrained problems
Mixed-integer problems

Real-world problems

v0.1: Individual problems
v0.2: Suites of unscalable problems
v0.5: Suites of scalable problems (in the number of variables)

Conclusions
Introduction

Why use test problems?

Algorithm A

Algorithm B

© Dimo Brockhoff and Tea Tušar
Benchmarking Multiobjective Optimizers 2.0 @ GECCO 2023, Lisbon, Portugal
Why use test problems?

- Algorithm A: 
  - Expert knowledge
  - Landscape analysis

- Algorithm B: 
  - Expert knowledge
  - Landscape analysis

© Dimo Brockhoff and Tea Tušar
Benchmarking Multiobjective Optimizers 2.0 @ GECCO 2023, Lisbon, Portugal

67
Introduction

Desirable characteristics of a benchmark problem set

[Bartz-Beielstein et al. 2020]

1. Diverse
2. Representative
3. Scalable and tunable
4. Known optima / best performance
5. [Continually updated]
Introduction

Recommendations for multiobjective test suites
Adapted from [Huband et al. 2006]

1. A few “easy” (unimodal) test problems
2. The majority of problems should be hard (multimodal, nonseparable and both multimodal and nonseparable)
3. Diverse Pareto front geometries (including degenerate fronts, disconnected fronts) and disconnected Pareto sets
Introduction

Additional recommendations for multiobjective test problems
Adapted from [Huband et al. 2006]

1. No extremal variables
2. No medial variables
3. Dissimilar variable domains
4. Dissimilar objective ranges
Introduction

Problem Design Approaches

[Deb et al. 2005]

1. Multiple single-objective functions approach
2. Bottom-up approach
   1. Choose a Pareto front
   2. Build the objective space
   3. Construct the search space
Artificial Problems
(Continuous and Unconstrained)
v0.1

Artificial Problems (Continuous and Unconstrained)
Individual Problems

Minimize \[
\begin{align*}
    f_1(x) &= x^2 \\
    f_2(x) &= (x - 2)^2
\end{align*}
\]  
[Schaffer 1985]

Minimize \[
\begin{align*}
    f_1(x) &= \sum_{i=1}^{2} \left[-10 \exp\left(-0.2 \sqrt{x_i^2 + x_{i+1}^2}\right)\right] \\
    f_2(x) &= \sum_{i=1}^{3} \left| x_i^{0.8} + 5 \sin(x_i^3) \right|
\end{align*}
\]  
[Kursawe 1991]

Images licensed under CC BY 2.0
Individual Problems

Minimize = \[
\begin{align*}
  f_1(x) &= 1 - \exp \left[ -\sum_{i=1}^{n} \left( x_i - \frac{1}{\sqrt{n}} \right)^2 \right] \\
  f_2(x) &= 1 - \exp \left[ -\sum_{i=1}^{n} \left( x_i + \frac{1}{\sqrt{n}} \right)^2 \right]
\end{align*}
\]

[Fonseca and Fleming 1995]

Minimize = \[
\begin{align*}
  f_1(x, y) &= \left[ 1 + (A_1 - B_1(x, y))^2 + (A_2 - B_2(x, y))^2 \right] \\
  f_2(x, y) &= (x + 3)^2 + (y + 1)^2 \\
  A_1 &= 0.5 \sin(1) - 2 \cos(1) + \sin(2) - 1.5 \cos(2) \\
  A_2 &= 1.5 \sin(1) - \cos(1) + 2 \sin(2) - 0.5 \cos(2) \\
  B_1(x, y) &= 0.5 \sin(x) - 2 \cos(x) + \sin(y) - 1.5 \cos(y) \\
  B_2(x, y) &= 1.5 \sin(x) - \cos(x) + 2 \sin(y) - 0.5 \cos(y)
\end{align*}
\]

where = \[
\begin{align*}
  A_1 &= 0.5 \sin(1) - 2 \cos(1) + \sin(2) - 1.5 \cos(2) \\
  A_2 &= 1.5 \sin(1) - \cos(1) + 2 \sin(2) - 0.5 \cos(2) \\
  B_1(x, y) &= 0.5 \sin(x) - 2 \cos(x) + \sin(y) - 1.5 \cos(y) \\
  B_2(x, y) &= 1.5 \sin(x) - \cos(x) + 2 \sin(y) - 0.5 \cos(y)
\end{align*}
\]
Individual Problems

Minimize = \begin{align*}
    f_1(x, y) &= 0.5 (x^2 + y^2) + \sin(x^2 + y^2) \\
    f_2(x, y) &= \frac{(3x - 2y + 4)^2}{8} + \frac{(x - y + 1)^2}{27} + 15 \\
    f_3(x, y) &= \frac{1}{x^2 + y^2 + 1} - 1.1 \exp(- (x^2 + y^2))
\end{align*}

[Viennet et al. 1996]
MOP Suite

MOP = Multi-Objective Problem [Van Veldhuizen 1999]

Properties

• A collection of 7 test problems from the literature
• Some problems are both nonseparable and multimodal
• A collection of various Pareto front geometries

Issues

• Most problems have 2 or 3 variables
• Not scalable in the number of objectives
• In many problems the optima lie on the boundary or in the middle of the search space
• The Pareto set is hard to compute for some problems
Artificial Problems (Continuous and Unconstrained)
ZDT Suite

ZDT = Zitzler, Deb, Thiele [Zitzler et al. 2000]

Construction with the bottom-up approach (following Deb’s toolkit [Deb 1999])

Given

\[ x = \{x_1, \ldots, x_n\} \]

Minimise

\[ f_1(y) \]

Distribution f.

Distance f.

Front shape

\[ f_2(y, z) = g(z) h(f_1(y), g(z)) \]

where

\[ y = \{x_1, \ldots, x_j\} \quad \text{Position variable(s) (} j = 1 \text{ for ZDT)} \]

\[ z = \{x_{j+1}, \ldots, x_n\} \quad \text{Distance variables} \]

The separation of variables was done to simplify problem construction.
ZDT Suite

Properties

• Scalable in the number of (distance) variables
• Some problems are multimodal
• Convex, concave and disconnected Pareto fronts
• The Pareto sets and fronts are known

Issues

• Not scalable in the number of objectives (2 objectives)
• 4 problems have optima on the boundary of the search space
• 1 problem has optima in the middle of the search space
• All problems are separable (the first objective depends only on the first variable)
v1.0

Artificial Problems (Continuous and Unconstrained)
DTLZ Suite

DTLZ = Deb, Thiele, Laumanns, Zitzler [Deb et al. 2005]

Improvement over ZDT

• Scalable number of objectives
• Optima do not lie on the boundary of the search space

Remaining issues

• Most problems have optima in the middle of the search space
• Problems still separable in practice (optimizing one variable at a time will yield at least one global optimum)
v1.2

Artificial Problems (Continuous and Unconstrained)
WFG Suite

WFG = Walking Fish Group [Huband et al. 2006]

**Improvement over DTLZ**

- Optima do not lie in the middle of the search space
- Some nonseparable, multimodal, deceptive and biased problems
- Convex, linear, concave, mixed, disconnected and degenerate Pareto fronts

**Remaining issues**

- The Pareto set is linear for 8 of the 9 problems
- Still rely on distance and position variables
v1.3

Artificial Problems (Continuous and Unconstrained)
Other Suites and Problems

Problems constructed with the bottom-up approach [Zapotecas et al. 2019]

- L-ZDT and L-DTLZ problems with linkages [Deb et al. 2006]
- IHR test suite of 5 rotated ZDT problems [Igel et al. 2007]
- ED problems based on Lamé superspheres [Emmerich and Deutz 2007]
- LZ test suite of 9 problems with complicated Pareto sets [Li and Zhang 2009]
- SZDT test suite of 7 scalable problems with complicated Pareto sets [Saxena et al. 2011]
- Convex DTLZ problem [Deb and Jain 2014]
- Inverted DTLZ problem [Jain and Deb 2014]
- MNI test suite of 2 problems with diverse shapes of the Pareto front [Masuda et al. 2016]
- LSMOP test suite of 9 problems for large-scale optimization with variable dependencies [Cheng et al. 2017b]
Other Suites and Problems

Problems constructed with the bottom-up approach

- Minus-DTLZ and Minus-WFG test suites [Ishibuchi et al. 2017]
- MMF test problems with diverse landscapes [Yue et al. 2019]
- GPD (Generalized Position-Distance) benchmark problem generator (problems can have various difficulties) [Meneghini et al. 2020]
- Suite of 10 ZCAT problems with various difficulties [Zapotecas et al. 2023]
CEC Competition Suites

Information about all CEC competitions:
https://www3.ntu.edu.sg/home/EPNSugan/index_files/cec-benchmarking.htm

13 test problems for CEC 2007 [Huang et al. 2007]
• OKA [Okabe et al. 2004], SYM-PART [Rudolph et al. 2007]
• 4 shifted ZDT, 1 rotated ZDT
• 2 shifted DTLZ, 1 rotated DTLZ
• 3 WFG
CEC Competition Suites

13 test problems for CEC 2009 (UF suite) [Zhang et al. 2009]
• 10 with complicated Pareto sets (4 from the LZ suite)
• 2 extended rotated DTLZ
• 1 WFG

15 test problems for CEC 2017 (MaF suite) [Cheng et al. 2017a]
• 7 modified DTLZ problems
• 2 distance minimization problems
• 3 WFG problems
• 1 SZDT problem
• 2 LSMOP problem
CEC Competition Suites

22 test problems for CEC 2019 [Liang et al. 2019]
• 2 SYM-PART
• Omni-test [Deb and Tiwari 2008]
• 19 MMF problems

24 test problems for CEC 2020 [Liang et al. 2020]
• 24 MMF problems
Survey of Recent Papers

64 papers on unconstrained continuous multiobjective optimization from recent conferences (without application papers)

- CEC 2020
- GECCO 2020
- PPSN 2020
- EMO 2021

<table>
<thead>
<tr>
<th>Benchmark</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>DTLZ</td>
<td>60.9%</td>
</tr>
<tr>
<td>WFG</td>
<td>34.4%</td>
</tr>
<tr>
<td>ZDT</td>
<td>20.3%</td>
</tr>
<tr>
<td>DTLZ/WFG/ZDT variants</td>
<td>20.3%</td>
</tr>
</tbody>
</table>

- Not using ZDT/DTLZ/WFG based problems: 23.4%
- Using ZDT/DTLZ/WFG and ZDT/DTLZ/WFG based problems: 40.6%
- Using only ZDT/DTLZ/WFG problems: 35.9%
v1.5

Artificial Problems (Continuous and Unconstrained)
Distance-Based Problems

General idea [Ishibuchi et al. 2010]

• Based on earlier work [Köppen et al. 2005, Rudolph et al. 2007]

Properties

• 2-D test problems that are inherently visualizable
• Pareto set easy to characterize
• Scalable in the number of objectives
• Useful for visualizing the distribution of solutions

Issues

• Simple objective functions
Distance-Based Problems

Extensions

- High-dimensional search spaces [Masuda et al. 2014]
- Distance to lines (instead of points) [Li et al. 2014, 2018]
- Dominance resistance regions [Fieldsend 2016]
- Local Pareto fronts [Liu et al. 2018]
- Problem generator for scalable problems with various properties (local fronts, disconnected Pareto sets and fronts, dominance resistance regions, uneven ranges of objective values, varying density of solutions) [Fieldsend et al. 2019]
v2.0

Artificial Problems (Continuous and Unconstrained)
bbob-biobj Suite

**Motivation** [Brockhoff et al. 2022]

- Real-world problems are not constructed using the bottom-up approach
- Go back to basics – use single-objective functions for each objective
bboṽ-biobj Suite

Construction

• Use the functions from the bboṽ suite [Finck et al. 2009]
  • Well-understood
  • Scalable in the number of variables and parametrized (instances)
  • 24 functions categorized in 5 groups based on their properties
    • Separable
    • Low or moderate conditioning
    • High conditioning and unimodal
    • Multimodal with global structure
    • Multimodal with weak global structure
bobb-biobj(-ext) Suites

<table>
<thead>
<tr>
<th>Seperable</th>
<th>Low or moderate conditioning</th>
<th>High conditioning and unimodal</th>
<th>Multimodal with global structure</th>
<th>Multimodal with weak global structure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **f_1**: Sphere
- **f_2**: Ellipsoid separable
- **f_6**: Attractive sector
- **f_8**: Rosenbrock original
- **f_10**: Rastrigin
- **f_14**: Sharp ridge
- **f_15**: Sum of different powers
- **f_16**: Schaffer F7 (condition 10)
- **f_18**: Schwefel x*sin(x)
- **f_20**: Gallagher 101 peaks
- **f_21**: Gallagher 21 peaks
- **f_24**: Katsuuras
- **f_25**: Lunacek bi-Rastrigin
- **f_26**: Schaefer F7 (condition 1000)
- **f_27**: Schaefer F7 (condition 100)
- **f_28**: Griewank-Rosenbrock f813
- **f_29**: Rosenbrock rotated
- **f_30**: Ellipsoid
- **f_31**: Discus
- **f_32**: Bent cigar
- **f_33**: Sharp ridge
- **f_34**: Sum of different powers

© Dimo Brockhoff and Tea Tušar

Benchmarking Multiobjective Optimizers 2.0 @ GECCO 2023, Lisbon, Portugal
bbob-biobj(-ext) Suites

Properties

• Construction similar as in real-world problems
• Scalability in the number of variables
• Various problem properties (more diverse than existing multiobjective test suites)
• Included in the COCO benchmarking platform [Hansen et al. 2021]
• Problem instances can be quite diverse

Issues

• Only 2 objectives
• Unknown Pareto set and front, but known single-objective optima and available approximations of the Pareto fronts (and sets for lower-dimensional problems)
Visualization of Multiobjective Landscapes
Visualization in Multiobjective Optimization

[Filipič and Tušar 2018, 2020]
Visualization of Multiobjective Problem Landscapes

Low-dimensional search spaces

• Dominance ratio [Fonseca 1995]
• Gradient path length (inspired by gradient plots [Kerschke and Grimme 2017])
• Local dominance [Fieldsend et al. 2019]
• PLOT [Schäpermeier et al. 2020]

Any-dimensional search spaces

• Line cuts [Brockhoff et al. 2022, Volz et al. 2019]
• Optima network [Liefooghe et al. 2018, Fieldsend and Alyahya 2019]
Visualization of Multiobjective Problem Landscapes

Various visualizations of bbo-biobj-ext problems

[https://numbbo.github.io/bbob-biobj/](https://numbbo.github.io/bbob-biobj/)

Visualizations of bbo-biobj and other multi-objective suites using PLOT

[https://schaepermeier.shinyapps.io/moPLOT/](https://schaepermeier.shinyapps.io/moPLOT/)
Visualization of Multiobjective Problem Landscapes

Problems for demonstration

• Double sphere problem bboB-biobj $F_1 = (f_1, f_1)$, instance 1
• Sphere-Gallagher problem bboB-biobj $F_{10} = (f_1, f_{21})$, instance 1
• Double Gallagher problem bboB-biobj $F_{55} = (f_{21}, f_{21})$, instance 1

Gallagher = Gallagher’s Gaussian
101-me Peaks Function
Dominance Ratio

- Discretized search space (501 x 501 grid)
- Dominance ratio = the ratio of grid points that dominate the current point
- All nondominated points have a ratio of zero
- Visualize dominance ratios in logarithmic scale

[Fonseca 1995]
Gradient Path Length

• Compute the bi-objective gradient for all grid points

\[ \nabla v = \frac{\nabla f_1(x)}{\|\nabla f_1(x)\|} + \frac{\nabla f_2(x)}{\|\nabla f_2(x)\|} \]

• From a grid point, follow the path in the direction of this gradient
• Visualize the length of the path to the local optimum

Adapted from [Kerschke and Grimme 2017]
Local Dominance

- **Green: Dominance-neutral local optima regions**
  - Points that are mutually nondominated with all their 8 neighbors
- **Pink: Basins of attraction**
  - Points dominated by at least one neighbor, their dominating paths lead to the same green region
- **White: Boundary regions**
  - Points whose dominating paths lead to different green regions

[Fieldsend et al. 2019]
Line Cuts

Search space projection ($F_1 = (f_1, f_2)$, 5-D, inst. 2)

Unscaled obj. space ($F_{10} = (f_1, f_2, f_3)$, 5-D, inst. 2)

Reference set (3095 of 4658018 points)
cuts through single optima
cut through both optima
two random directions
random cut in plane through optima

cuts through single optima
cut through both optima
two random directions
random cut in plane through optima

cuts through single optima
cut through both optima
two random directions
random cut in plane through optima

cuts through single optima
cut through both optima
two random directions
random cut in plane through optima

Second objective

First objective

Reference set (1870984 points)
cuts through single optima
cut through both optima
two random directions
random cut in plane through optima

Reference set (2216 of 1070984 points)
cuts through single optima
cut through both optima
two random directions
random cut in plane through optima

Reference set (2216 of 1070984 points)
cuts through single optima
cut through both optima
two random directions
random cut in plane through optima

© Dimo Brockhoff and Tea Tušar
Benchmarking Multiobjective Optimizers 2.0 @ GECCO 2023, Lisbon, Portugal
109
Comparison of Problem Landscapes

Two problems where both objectives are separable, first is unimodal and second is multimodal

**ZDT4**

**bbob-biobj-ext** $F_{56}$

- $f_1$ Sphere function
- $f_3$ Rastrigin function
Comparison of Problem Landscapes

Two problems where both objectives are separable and multimodal

**DTLZ1**

**bbob-biobj-ext $F_{62}$**

- $f_3$: Rastrigin function
- $f_4$: Skew Rastrigin-Bueche
Comparison of Problem Landscapes

Two problems where both objectives are nonseparable and multimodal

- WFG9
- bbob-biobj-ext $F_{82}$
  - $f_{17}$ Schaffer F7
  - $f_{19}$ Griewank-Rosenbrock
Some Notes on Problem Properties
Problem Instances

15 instances of the double sphere problem bbob-biobj $F_1$
Problem Instances

15 instances of the double Gallagher problem bbob-biobj $F_{55}$
Real-world problems have instances

- Same problem with varying variable bounds
- Very similar problem (casting steel 31CrV3 vs. 51CrV4)
Scalability in the Number of Objectives

Questionnaire on the properties of real-world problems [van der Blom et al. 2020]
https://sites.google.com/view/macoda-rwp/home
• Only 45 problems (!)
• Of these, only 9% have more than 4 objectives

Are we over-emphasizing many-objective problems?
Scalability in the Number of Objectives

Large number of possible combinations

- ZDT/DTLZ/WFG and other suites have a fixed number of problems
- Example for 10 functions
  - 2 objectives → 55 combinations
  - 3 objectives → 220 combinations
  - 5 objectives → 715 combinations

- Problem selection [Andova et al. 2023]
  - Goal: choose only a limited number of most diverse problems
  - Diversity measured in problem feature space
  - Proof of concept on bbo-biobj problems
Other Artificial Problems
Other Artificial Problems

Suites of multiobjective problems with constraints

- CTP [Deb et al. 2001]
- DTLZ (problems DTLZ8-9) [Deb et al. 2005]
- CF [Zhang et al. 2009]
- C-DTLZ [Jain and Deb 2014]
- NCTP [Li et al. 2016]
- DC-DTLZ [Li et al. 2019]
- LIR-CMOP [Fan et al. 2019]
- MW [Ma and Wang 2019]
- DOC [Liu and Wang 2019]
- DAS-CMOP and DAS-CMaOP [Fan et al. 2020]
- Eq-DLTZ and Eq-IDTLZ [Cuate et al. 2020]
- CLSMOP [He et al. 2021]
Other Artificial Problems

Problem types

Depending on how the constraints affect the Pareto set/front

Problems of Type I not useful for benchmarking constraint handling techniques

Adapted from [Ma and Wang 2019]
Other Artificial Problems

Analysis of multiobjective problems with constraints [Vodopija et al. 2022]

https://vodopijaaljosa.github.io/cmop-web/

Tutorial on multiobjective optimization in the presence of constraints:
https://dis.ijs.si/filipic/wcci2022tutorial/
Other Artificial Problems

Suites of multiobjective mixed-integer problems

- 3 bi-objective problems summing one discrete and one continuous function [Sadowski et al. 2021]
- Exeter suite of 6 problems constructed with the bottom-up approach [McClymont and Keedwell 2011]
- bbo-biobj-mixint suite of 92 bi-objective problems [Tušar et al. 2019]

Pareto set and front approximations for three different instances of the double sphere function
Real-World Problems
v0.1

Real-World Problems
Real-World Problems

Individual problems (white box)

- Water resource planning problem with 3 variables, 5 objectives and 7 constraints [Musselman and Talavage 1980]
- Two bar truss design problem with 2 variables, 2 objectives and 2 constraints [Rao 1987]
- Vibrating platform design problem with 3 variables, 2 objectives and 5 constraints [Ray et al. 2001]
- Welded beam design problem with 4 variables, 2 objectives and 5 constraints [Ray and Lie 2002]
- Multi-speed gearbox design problem with 10 variables, 2 objectives and 38 constraints [Deb and Jain 2003]
- Mineral processing production planning problem with 6 variables, 5 objectives and 9 constraints [Yu et al. 2011]
- Car side impact problem with 11 variables, 3 objectives and 10 constraints [Jain and Deb 2014]
- ...
### Real-World Problems

#### Individual problems (black box)

- Radar waveform design with a varying number of variables and 9 objectives \[\text{Hughes 2007}\]
- HBV problem of calibrating the rainfall-runoff model with 14 variables and 4 objectives \[\text{Reed et al. 2013}\]
- MAZDA car structure design problem with 222 integer variables, 2 objectives and 54 constraints \[\text{Kohira et al. 2018}\]
- Lunar lander landing site selection problem with 2 variables, 3 objectives and 2 constraints \[\text{JSEC and JAXA 2018}\]
- Wind turbine design with 32 variables, 5 objectives and 22 constraints \[\text{JSEC 2019}\]
- Trappist tour planning problem with 34 mixed-integer variables, 2 objectives and 1 constraint \[\text{ESA 2022}\]
- Quantum communications constellations problem with 20 mixed-integer variables, 2 objectives and 2 constraints \[\text{ESA 2023}\]
v0.2
Real-World Problems
Suites of Real-World Problems

Suites of unscalable problems

• DDMOP suite of 7 test problems with a different number of variables (5–17) and objectives (2–10) [He et al. 2020]

• Two suites of previously published problems [Tanabe and Ishibuchi 2020]
  • RE suite of 16 test problems with a different number of variables (2–7) and objectives (2–9)
    • 11 continuous, 1 integer, 4 mixed-integer
  • CRE suite of 8 test problems with constraints a different number of variables (3–7) and objectives (2–5)
    • 6 continuous, 1 integer, 1 mixed-integer

• RCM suite of 50 problems with a different number of variables (2–34), objectives (2–5) and constraints (1–29) [Kumar et al. 2021]
v0.5

Real-World Problems
Suites of scalable problems

- Multi-observable quantum control problems with scalable variables, noise and 2 or 3 objectives [Shir et al. 2012]
- Heat exchanger design problem with scalable variables and 1 or 2 objectives [Daniels et al. 2018]
- Suite of 3 bi-objective TopTrumps problems in multiple dimensions and instances [Volz et al. 2019]
- Suite of 26 bi-objective MarioGAN problems in multiple dimensions and instances [Volz et al. 2019]
- MODAct suite of 20 problems with 2+6k variables, various number of objectives (2–5) and constraints (7–10) [Picard and Schiffmann 2021]
- Scalable multi-agent pathfinding problems
  - Discrete problem formulation with 5 objectives [Weise and Mostaghim 2022]
  - Continuous problem formulation with 2 objectives [Mai et al. 2023]
Conclusions

• We should think about the usefulness of our research
• The questionnaire on the properties of real-world problems has shown their diversity [van der Blom et al. 2020]
• Most research is done on continuous unconstrained problems
• A lot (too much?) research on many-objective problems
• Although the test problems are scalable, most studies use a fixed number of variables
Conclusions

• Problem suites constructed with the bottom-up approach have unrealistic properties

• Algorithms are overfitting to these problems (especially the overused DTLZ and WFG) [Ishibuchi et al. 2017, 2023]

• Using separate functions for the objectives looks like a step in the right direction
Overview

1. Performance Assessment

2. Test Problems and Their Visualizations

3. Recommendations from Numerical Results
python -m cocopp bbo-biobj*
Aggregated Results Over All 55 Functions
Aggregated Results Over All 55 Functions
Multiobjective Benchmarking 3.0?

a.k.a Challenging Open Research Directions
Multiobjective Benchmarking 3.0?

a.k.a Challenging Open Research Directions

• Many-objective problems
  • Problems/suites
  • Indicators
  • Efficient implementations
Multiobjective Benchmarking 3.0?

a.k.a Challenging Open Research Directions

• Many-objective problems
  • Problems/suites
  • Indicators
  • Efficient implementations

• Constraints, mixed-integer, ...
Multiobjective Benchmarking 3.0?

a.k.a Challenging Open Research Directions

• Many-objective problems
  • Problems/suites
  • Indicators
  • Efficient implementations
• Constraints, mixed-integer, ...
• Real-world benchmarking?
  • Simulation crashes
  • Parallelism
  • Dynamic changes
  • Interactive decision making, ...
Multiobjective Benchmarking 3.0?

a.k.a Challenging Open Research Directions

- Many-objective problems
  - Problems/suites
  - Indicators
  - Efficient implementations
- Constraints, mixed-integer, ...
- Real-world benchmarking?
  - Simulation crashes
  - Parallelism
  - Dynamic changes
  - Interactive decision making, ...
- Benchmarking results from more classical approaches
Three “New Year” Resolutions
Three “New Year” Resolutions

1. Show convergence graphs/ECDF

Anything else than tables for fixed budget
Three “New Year” Resolutions

1. Show convergence graphs/ECDF

   Anything else than tables for fixed budget

2. Use “most realistic” problems
Three “New Year” Resolutions

1. Show convergence graphs/ECDF
   Anything else than tables for fixed budget

2. Use “most realistic” problems

3. Showing scaling with (search & objective space) dimension
Three “New Year” Resolutions

1. Show convergence graphs/ECDF

   Anything else than tables for fixed budget

2. Use “most realistic” problems

3. Showing scaling with (search & objective space) dimension
Thank you!
Supplementary Material
Bibliography


Bibliography


Bibliography


Bibliography


© Dimo Brockhoff and Tea Tušar
Bibliography


Bibliography


© Dimo Brockhoff and Tea Tušar
Bibliography


Bibliography


Instructor Biography: Dimo Brockhoff

Dimo Brockhoff
RandOpt team
Inria Saclay - Ile-de-France
CMAP, CNRS, Ecole Polytechnique, IP Paris
Route de Saclay
91128 Palaiseau
France

After obtaining his diploma in computer science (Dipl.-Inform.) from University of Dortmund, Germany in 2005, Dimo Brockhoff received his PhD (Dr. sc. ETH) from ETH Zurich, Switzerland in 2009. After postdoctoral research positions at Inria Saclay Ile-de-France in Orsay and at Ecole Polytechnique in Palaiseau, both in France, Dimo has been a permanent researcher at Inria: from 2011 till 2016 with the Inria Lille - Nord Europe research center and since October 2016 with the Saclay - Ile-de-France research center, co-located with CMAP, Ecole Polytechnique, IP Paris. His most recent research interests are focused on evolutionary multiobjective optimization (EMO) and other (single-objective) blackbox optimization techniques, in particular with respect to benchmarking, theoretical aspects, and expensive optimization.
Tea Tušar

Computational Intelligence Group
Department of Intelligent Systems
Jožef Stefan Institute
Jamova cesta 29
1000 Ljubljana
Slovenia

Tea Tušar is a senior research associate at the Department of Intelligent Systems of the Jožef Stefan Institute, and an assistant professor at the Jožef Stefan International Postgraduate School, both in Ljubljana, Slovenia. After receiving the PhD degree in Information and Communication Technologies from the Jožef Stefan International Postgraduate School for her work on visualizing solution sets in multiobjective optimization, she has completed a one-year postdoctoral fellowship at Inria Lille in France where she worked on benchmarking multiobjective optimizers. Her research interests include evolutionary algorithms for singleobjective and multiobjective optimization with emphasis on visualizing and benchmarking their results and applying them to real-world problems.