Black-Box Optimization Benchmarking for Noiseless Function Testbed using Artificial Bee Colony Algorithm

Draft version *

Mohammed El-Abd American University of Kuwait melabd@auk.edu.kw

ABSTRACT

This paper benchmarks the artificial bee colony (ABC) algorithm using the noise-free BBOB 2010 testbed.

Categories and Subject Descriptors

G.1.6 [Numerical Analysis]: Optimization—global optimization, unconstrained optimization; F.2.1 [Analysis of Algorithms and Problem Complexity]: Numerical Algorithms and Problems

General Terms

Algorithms

Keywords

Benchmarking, Black-box optimization

ARTIFICIAL BEE COLONY

The ABC algorithm was first proposed in [7]. The algorithms was inspired by the method adopted by a swarm of honey bees to locate food sources. There are two different honey bee groups that share knowledge in order to successfully locate such sources. First, there are the employed bees that are currently exploiting a food source. Second, there are the onlookers that wait at the nest and establish communication with the employed bees.

This algorithm was applied to multidimensional and multimodal function optimization in [7, 2, 11]. The swarm is divided into employed bees, scouts and onlookers. S_n solutions to the problem are randomly initialized in the function domain and referred to as food sources. A number of employed bees, set as the number of the food sources (half the colony size), are used to find new food sources using the following equation:

$$v_{ij} = x_{ij} + \phi_{ij} \times (x_{ij} - x_{kj}), \tag{1}$$

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. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

GECCO'10, July 7–11, 2010, Portland Oregon, United States of America. Copyright 2010 ACM X-XXXXX-XX-X/XX/XX ...\$10.00.

for $j \in \{1 \dots d\}$ where d is the number of dimensions, ϕ_{ij} is a random number uniformly distributed in the range [-1,1], k is the index of a randomly chosen solution, \mathbf{x}_i is the current food source exploited by employed bee i, and \mathbf{v}_i is the new food source to be exploited. Both \mathbf{v}_i and \mathbf{x}_i are then compared against each other and the employed bee exploits the better food source.

Next, each onlooker bee randomly selects a food source to exploit according to the probability given in equation 2. Then, each onlooker bee tries to find a better food source around the selected one using equation 1.

$$p_i = \frac{fit_i}{\sum_{j=1}^{S_n} fit_j},$$
(2)

where fit_i is the fitness of the i^{th} food source.

If a food source cannot be improved for a predetermined number of cycles, referred to as limit, this food source is abandoned. The employed bee that was exploiting this food source becomes a scout and randomly selects another food source in the domain according to:

$$x_{ij} = x_j^{min} + rand \times (x_j^{max} - x_j^{min}), \tag{3}$$

where x_j^{min} and x_j^{max} are the minimum and maximum domain bounds.

The ABC algorithm is shown in Algorithm 1.

Algorithm 1 The ABC algorithm

Require: Max_Cycles, ColonySize, limit.

- 1: Initialize the food sources
- 2: Evaluate the food sources
- 3: Cvcle=1

11:

- 4: while $Cycle \leq Max_Cycles$ do
- 5: for each employed bee i do
- 6: Produce a new solution
- 7: Evaluate the new solution
- Apply Greedy selection choosing the better solution 8: 9: end for
- 10: for each onlooker bee i do
 - Probabilistically choose a solution according to p_i
- 12: Evaluate the new solution
- 13: Apply Greedy selection choosing the better solution 14:
- 15:
 - Re-initialize solutions not improved for *limit* cycles
- 16: Memorize the best solution
- Cvcle = Cvcle + 117:
- 18: end while
- 19: return best solution

Previous studies performed to assess the performance ABC

^{*}Submission deadline: March 25th.

included the work in [12] showing that the ABC algorithm performs better than PSO, an evolutionary algorithm (EA) and DE on a small suite of classical benchmark functions. Another study was carried in [10] that compared ABC against PSO, a Genetic Algorithm (GA), DE and an Evolutionary Strategy (ES) algorithm on a larger number of functions. It was shown that the performance of ABC is better than or at least similar to those algorithms while having a smaller number of parameters to tune. The work in [9] compared ABC to HS and the Bees Algorithm (BA) proposed in [13]. The comparison was based on a small set of classical functions and the ABC showed superior performance over both algorithms while producing reasonable results for higher dimensions.

2. PARAMETER TUNING

For ABC, the work in [1] indicated that the there is no need to have a huge colony size in order to provide good results. We use 40 bees as our previous experiments conducted on the CEC05 benchmarks [14] and repeated using populations of 20, 40 and 100 bees for different problem sizes showed that this setting provided the best results on average. The recommendations in [10] were followed by setting the limit parameters to $S_n \times D$, although recent research [1] indicated that lower values might be needed for more difficult functions.

3. RESULTS

Results from experiments according to [4] on the benchmark functions given in [3, 6] are presented in Figures 1, 2 and 3 and in Tables 1 and 2.

Experiments use the ABC code available at [8]. The ABC algorithm is allowed to perform a maximum of $10^5 \times D$ function evaluations for all test functions and no restart mechanism was used.

4. CPU TIMING EXPERIMENT

For the timing experiment, ABC was run on f8 and restarted until at least 30 seconds had passed (according to Figure 2 in [5]). The experiments have been conducted with an Intel Core 2 Quad 2.4 GHz under Windows Vista using the MATLAB-code provided. The results were 2.0×10^{-4} seconds per function evaluation in dimensions 2 up to 20. A dependency of CPU time on the search space dimensionality is not visible.

5. REFERENCES

- [1] B. Akay and D. Karaboga. Parameter tuning for the artificial bee colony algorithm. In *Proceedings of 1st International Conference on Computational Collective Intelligence*, pages 608–619. Springer-Verlag:Berlin Heidelberg, 2009.
- [2] B. Basturk and D. Karaboga. An artificial bee colony (ABC) algorithm for numeric function optimization. In *Proceedings of the IEEE Swarm Intelligence* Symposium, pages 66–66. IEEE Press, 2006.
- [3] S. Finck, N. Hansen, R. Ros, and A. Auger. Real-parameter black-box optimization benchmarking 2010: Presentation of the noiseless functions. Technical Report 2009/20, Research Center PPE, 2010.

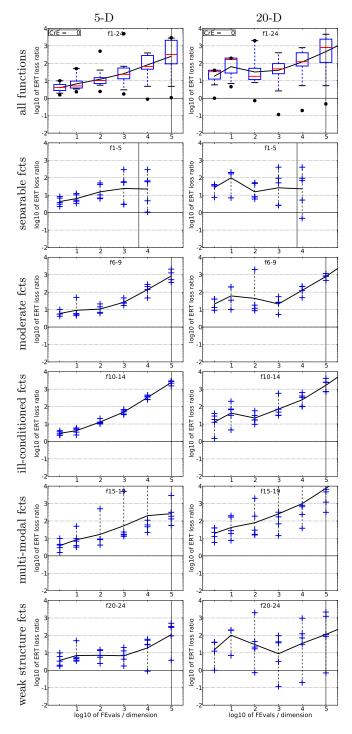


Figure 3: ERT loss ratio versus given budget FEvals. The target value $f_{\rm t}$ for ERT (see Figure 1) is the smallest (best) recorded function value such that ERT $(f_{\rm t}) \leq$ FEvals for the presented algorithm. Shown is FEvals divided by the respective best ERT $(f_{\rm t})$ from BBOB-2009 for functions f_1-f_{24} in 5-D and 20-D. Each ERT is multiplied by $\exp({\rm CrE})$ correcting for the parameter crafting effort. Line: geometric mean. Box-Whisker error bar: 25-75%-ile with median (box), 10-90%-ile (caps), and minimum and maximum ERT loss ratio (points). The vertical line gives the maximal number of function evaluations in this function subset.

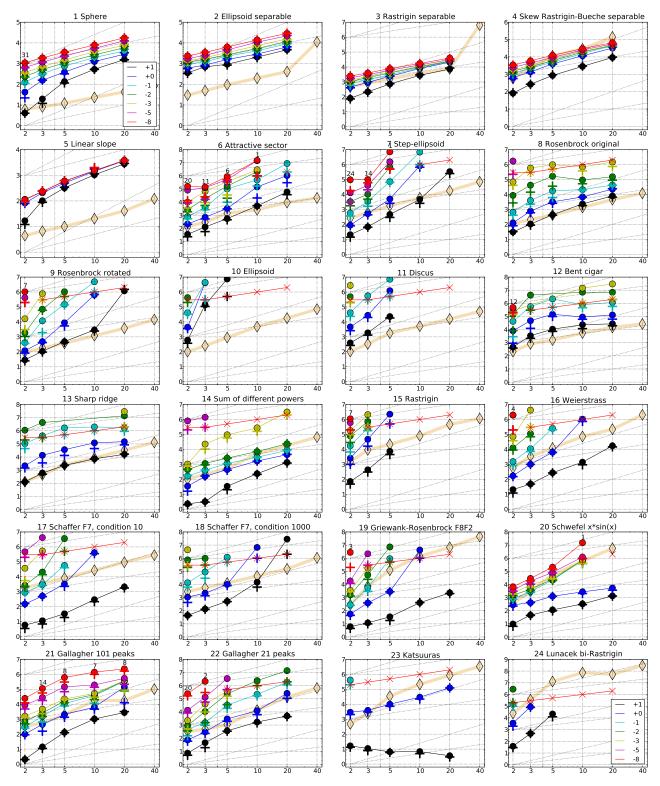


Figure 1: Expected Running Time (ERT, ullet) to reach $f_{\mathrm{opt}}+\Delta f$ and median number of f-evaluations from successful trials (+), for $\Delta f=10^{\{+1,0,-1,-2,-3,-5,-8\}}$ (the exponent is given in the legend of f_1 and f_{24}) versus dimension in log-log presentation. For each function and dimension, $\mathrm{ERT}(\Delta f)$ equals to $\#\mathrm{FEs}(\Delta f)$ divided by the number of successful trials, where a trial is successful if $f_{\mathrm{opt}}+\Delta f$ was surpassed. The $\#\mathrm{FEs}(\Delta f)$ are the total number (sum) of f-evaluations while $f_{\mathrm{opt}}+\Delta f$ was not surpassed in the trial, from all (successful and unsuccessful) trials, and f_{opt} is the optimal function value. Crosses (×) indicate the total number of f-evaluations, $\#\mathrm{FEs}(-\infty)$, divided by the number of trials. Numbers above ERT-symbols indicate the number of successful trials. Y-axis annotations are decimal logarithms. The thick light line with diamonds shows the single best results from BBOB-2009 for $\Delta f=10^{-8}$. Additional grid lines show linear and quadratic scaling.

$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	cc # ERT 10% 90% RT _{Succ} 2 #5 1.6e3 6.5e2 2.7e3 1.6e3 2 15 2.8e3 8.8e2 4.6e3 2.8e3 2 15 4.1e3 1.4e3 6.2e3 4.1e3 3 15 7.6e3 3.5e3 9.7e3 7.6e3 3 15 1.3e4 1.2e4 1.3e4 1.3e4	$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	# ERT 10% 90% RT _{SUCC}	$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	# ERT 10% 90% RT _{succ}	$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	15 3.4e5 5.3e4 9.4e5 3.4e5 0 60e-1 37e-1 83e-1 5.8e5	$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$	13 1.2e6 9.5e4 1.8e6 8.9e5 0 92e-1 67e-1 10e+0 2.0e6	$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	f ₁₃ in 20-D, N=15, mFE=2.00e6 # ERT 10% 90% RT _{succ} 15 1.6e4 8.6e3 2.3e4 1.6e4	$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	f17 in 20-D, N=15, mFE=2.00e6 # ERT 10% 90% RT _{succ} 15 2.1e3 7.0e2 3.9e3 2.1e3	$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	0 40e-1 27e-1 44e-1 1.2e6 	$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	1.580 1.58	$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
$ \begin{array}{llllllllllllllllllllllllllllllllllll$	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	0 64e-2 54e-2 79e-2 9.6e5 	$\begin{array}{cccccccccccccccccccccccccccccccccccc$

Table 1: Shown are, for a given target difference to the optimal function value Δf : the number of successful trials (#); the expected running time to surpass $f_{\rm opt}+\Delta f$ (ERT, see Figure 1); the 10%-tile and 90%-tile of the bootstrap distribution of ERT; the average number of function evaluations in successful trials or, if none was successful, as last entry the median number of function evaluations to reach the best function value (RT_{succ}). If $f_{\rm opt}+\Delta f$ was never reached, figures in *italics* denote the best achieved Δf -value of the median trial and the 10% and 90%-tile trial. Furthermore, N denotes the number of trials, and mFE denotes the maximum of number of function evaluations executed in one trial. See Figure 1 for the names of functions.

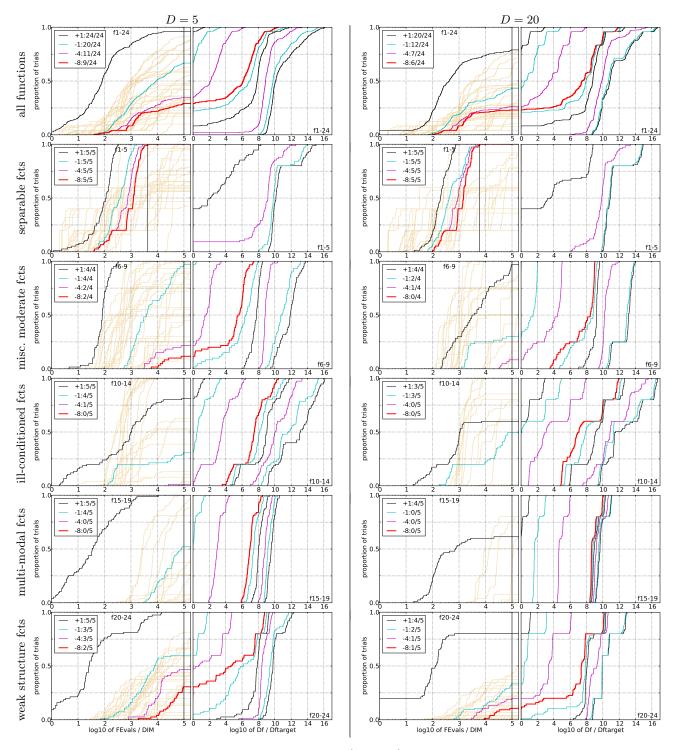


Figure 2: Empirical cumulative distribution functions (ECDFs), plotting the fraction of trials versus running time (left subplots) or versus Δf (right subplots). The thick red line represents the best achieved results. Left subplots: ECDF of the running time (number of function evaluations), divided by search space dimension D, to fall below $f_{\rm opt} + \Delta f$ with $\Delta f = 10^k$, where k is the first value in the legend. Right subplots: ECDF of the best achieved Δf divided by 10^k (upper left lines in continuation of the left subplot), and best achieved Δf divided by 10^{-8} for running times of D, 10D, 100D... function evaluations (from right to left cycling black-cyan-magenta). The legends indicate the number of functions that were solved in at least one trial. FEvals denotes number of function evaluations, D and DIM denote search space dimension, and Δf and Df denote the difference to the optimal function value. Light brown lines in the background show ECDFs for target value 10^{-8} of all algorithms benchmarked during BBOB-2009.

Table 2: ERT loss ratio (see Figure 3) compared to the respective best result from BBOB-2009 for budgets given in the first column. The last row $RL_{\rm US}/D$ gives the number of function evaluations in unsuccessful runs divided by dimension. Shown are the smallest, 10%-ile, 25%-ile, 50%-ile, 75%-ile and 90%-ile value (smaller values are better).

ie value (smaller values are better).									
f_1 - f_{24} in 5-D, maxFE/D=100011									
#FEs/D	best	10%	25%	\mathbf{med}	75%	90%			
2	1.6	1.9	2.5	4.0	5.8	10			
10	2.3	3.3	3.8	5.1	8.3	50			
100	2.4	5.4	6.9	10	15	42			
1e3	1.8	3.0	13	23	52	90			
1e4	0.89	4.4	39	63	2.7e2	4.0e2			
1e5	1.1	4.7	80	3.2e2	1.8e3	2.8e3			
RL_{US}/D	1e5	1e5	1e5	1e5	1e5	1e5			
f_{1} - f_{24} in 20-D, maxFE/D=100002									
#FEs/D	best	10%	25%	\mathbf{med}	75%	90%			
2	1.0	5.4	13	31	40	40			
10	4.6	7.0	23	1.7e2	2.0e2	2.0e2			
100	0.73	7.7	11	18	49	2.0e3			
1e3	0.12	2.6	20	46	95	4.2e2			
1e4	0.20	4.8	59	1.2e2	4.0e2	8.1e2			
1e5	0.47	4.8	1.0e2	7.7e2	2.3e3	4.8e3			
1e6	0.47	9.0	2.7e2	5.3e3	1.6e4	4.3e4			
RL_{US}/D	1e5	1e5	1e5	1e5	1e5	1e5			

- [4] N. Hansen, A. Auger, S. Finck, and R. Ros. Real-parameter black-box optimization benchmarking 2010: Experimental setup. Technical Report RR-7215, INRIA, 2010.
- [5] N. Hansen, S. Finck, R. Ros, and A. Auger. Real-parameter black-box optimization benchmarking 2009: Noiseless functions definitions. Technical Report RR-6829, INRIA, 2009.
- [6] N. Hansen, S. Finck, R. Ros, and A. Auger. Real-parameter black-box optimization benchmarking 2010: Noiseless functions definitions. Technical Report RR-6829, INRIA, 2010.
- [7] D. Karaboga. An idea based on honey bee swarm for numerical optimization. Technical Report TR06, Engineering Faculty, Computer Engineering Department, Erciyes University, 2005.
- [8] D. Karaboga. Artificial bee colony code. http://mf.erciyes.edu.tr/abc/software.htm, 2008.
- [9] D. Karaboga and B. Akay. Artificial bee colony (abc), harmony search and bees algorithms on numerical optimization. In *Proceedings of Innovative Production Machines and Systems Virtual Conference, IPROMS*, 2009.
- [10] D. Karaboga and B. Akay. A comparative study of artificial bee colony algorithm. Applied Mathematics and Computation, 214:108–132, 2009.
- [11] D. Karaboga and B. Basturk. A powerful and efficient algorithm for numerical function optimization: Artificial bee colony (ABC) algorithm. *Journal of Global Optimization*, 39(3):459–471, 2007.
- [12] D. Karaboga and B. Basturk. On the performance of artificial bee colony (ABC) algorithm. Applied Soft Computing, 8(1):687–697, 2008.

- [13] D. T. Pham, A. Ghanbarzadeh, E. Koc, S. Otri, S. Rahim, and M. Zaidi. The bees algorithm - a novel tool for complex optimisation problems. In *Proceedings* of *Innovative Production Machines and Systems* Virtual Conference, IPROMS, pages 451–461, 2006.
- [14] P. N. Suganthan, N. Hansen, J. J. Liang, K. Deb, Y.-P. Chen, A. Auger, and S. Tiwari. Problem definitions and evaluation criteria for the CEC 2005 special session on real-parameter optimization. Technical Report 2005005, ITT Kanpur, India, 2005.