How to Set the Maximum Number of Function Evaluations for the L-SHADE Algorithm with the AS^3D Approach?

Jana Herzog
jana.herzog1@um.si
Faculty of Electrical
Engineering and Computer Science,
University of Maribor
Koroška cesta 46
SI-2000 Maribor, Slovenia

Janez Brest
janez.brest@um.si
Faculty of Electrical
Engineering and Computer Science,
University of Maribor
Koroška cesta 46
SI-2000 Maribor, Slovenia

Borko Bošković
borko.boskovic@um.si
Faculty of Electrical
Engineering and Computer Science,
University of Maribor
Koroška cesta 46
SI-2000 Maribor, Slovenia

ABSTRACT

The control parameter, maximum number of function evaluations plays two important roles during the optimization process. It can determine the population size of some evolutionary algorithms and it also serves as a stopping condition of an optimization process. In this paper, we focus on setting the value of the control parameter for the L-SHADE algorithm for a chosen large-scale benchmark function. For this purpose we utilized a recently proposed approach AS^3D, which enables us to predict a stopping condition with a certain probability for a given solver and optimization problem, while using the fixed-target approach.

KEYWORDS

stopping condition, evolutionary algorithm, target approach

1 INTRODUCTION

Recently proposed algorithms, such as jSO [6], L-SHADE [19], iL-SHADE [5] and MadDE [4] use the maximum number of function evaluations (*maxFEs*) as a control parameter. This has two roles during an optimization process. It influences the population size and it serves as a stopping criteria. When a specific value of function evaluations as a stopping condition is set, this is the fixed-budget approach [13].

In competitions, such as CEC [1], the maximum number of function evaluations is set as a stopping condition. The value of it is predetermined or set according to the past years competitions. However, it can occur that the improper setting of the stopping condition can affect the performance of the evolutionary algorithms. It can happen that the budget is too small, which can result in premature convergence and the solution which is not of optimal quality. Some mishaps during a new experiment can happen if the predetermined budget is different from the one used in the original paper. This can affect the comparison of the algorithms, since a bigger budget can help the algorithm reach a solution of a better quality. However, the stopping criteria does not only terminate the algorithm, but it also plays an important role in the analysis and comparison of the evolutionary algorithms. The stopping condition can produce a significant differences in the ranking of evolutionary algorithms [18]. Before an experiment, it is crucial to set the correct stopping condition due to all aforementioned points. It is unknown how many number of function evaluations an algorithm will need to reach the solution of a wanted quality. The question which we propose in this paper is: how to determine a maximum number of function evaluations for a chosen evolutionary algorithm and problem?

For this purpose, we focus on setting a stopping condition/control parameter for a given evolutionary algorithm on a chosen benchmark function. We chose a 7-nonseparable, 1-separable Shifted and Rotated Elliptic Function from CEC'2013 Large-Scale Global Optimization benchmark functions [16] and L-SHADE. L-SHADE is considered as the state-of-the-art algorithm from the previous CEC competitions with the success-history based parameter adaptation and linear population size reduction [19]. To be able to set the value of a control parameter (*maxFEs*), we will utilize a recently proposed approach AS^3D (Analysis of the Stochastic Solvers based on the Statistical Distribution) [12]. This approach is based on the statistical distribution and parameters of the observed variable. In our case, this is the number of function evaluations needed to reach a specific target. With this approach, we will not only set a stopping condition/control parameter for the higher dimensions of the given optimization problem, but also provide these values with a specific probability. Identifying the statistical distribution and its parameters enables us to establish a predictive model. The predictive model helps in estimating the stopping condition according to the characteristics of the chosen evolutionary algorithm and optimization problem.

The paper is organized as follows. In Section 2, the related work is described. In Section 3, the experiment and analysis are provided. Section 4 concludes our paper.

2 RELATED WORK

In this paper, we focus on three aspects of the evolutionary computation: on setting the stopping condition, the analysis and comparison of evolutionary algorithms, and the statistics behind it all.

Firstly, we focus on setting the control parameter of the optimization process for a specific evolutionary algorithm and optimization problem. In our case, when analyzing the L-SHADE algorithm, the control parameter serves also as a stopping condition. This is the fixed-budget approach [3]. This means that the value of it is predetermined and the algorithm stops, when the budget is spent. Contrary to the fixed-budget approach, with the fixed-target approach [10], we set a certain quality of solutions, which should be reached by the evolutionary algorithm. In this case we observe the number of function evaluations or runtime needed to reach this quality of solutions.

Setting the stopping condition represents a demanding task. In [14], authors argue that setting a higher number of function evaluations as a stopping condition may not present a higher computational cost and should be considered in benchmarking. They

DOI: https://doi.org/10.26493/scores23.11

examine the effect of a higher evaluation budget on the performance, mean, convergence of the algorithms, and population diversity. In [18], the authors show that using different stopping criteria produces different results while comparing the state-of-the-art algorithms. They argue that this fact is often overlooked by the researchers.

One of the most recent approaches, which offers several aspects of the algorithm's comparison and analysis is AS^3D approach, which will be utilized in this paper. This approach does not only rely on statistics, but it takes into consideration the characteristics of an algorithm and also of an optimization problems. It offers a deeper insight into the performance of an algorithm based on the statistical distribution of the observed variables [12]. It does not only provide a statistical distribution of the observed variable as it is described in [17], but also establishes predictive models with which we can predict stopping conditions with any wanted probability.

However, one should not neglect the basis for every fair comparison of the evolutionary algorithms: the parametric and non-parametric statistical tests [8], [9]. Several statistical approaches have been proposed as an answer to only using the statistical tests. Those are the following [7], [15], [20] and [9].

3 EXPERIMENT

In this paper, we focused on analysing the L-SHADE algorithm on a 7-nonseparable, 1-separable Shifted and Rotated Elliptic Function. The main intention is to show how to set the control parameter for L-SHADE for a larger dimension of the chosen benchmark function.

We made 100 independent runs for each chosen dimension $D = \{5, 10, ..., 40\}$ of the large-scale function. We applied the target-approach with the optimal quality of solutions. The proposed approach AS^3D requires that the given evolutionary algorithm reaches the target in all runs, so that the hit ratio is 100%. We show how to set/predict the control parameter/stopping condition *maxFEs* for the dimension $D = 50$ based on the model established from the smaller dimensions. Then we will empirically validate the results by running the L-SHADE for $D = 50$ and comparing the empirical and predicted values.

Firstly, we analysed the statistical distribution of 100 independent runs of each dimension $D = \{5, 10, ..., 40\}$. For this purpose, we used the Shapiro Wilk's statistical test, where the p-value needs to be less than 0.005 ($p < 0.005$). We observed how many number of function evaluations $NFEs$ are needed that L-SHADE reaches the set optimal solution in each of the independent runs. We show that the statistical distribution is normal for each of the chosen dimensions. The Fig. 1 depicts the normal distribution for the dimension $D = 10$. Histogram can be a good visualization tool for determining the statistical distribution of the given sample [11].

Normal distribution has two parameters: mean and standard deviation, which are calculated as shown in Eq. (1) and Eq. (2).

$$\bar{x} = \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right) \tag{1}$$

The x in Eqs. (1) and (2) represents the *NFEs*. The n represents the sample size (number of independent runs), which is in this case 100.

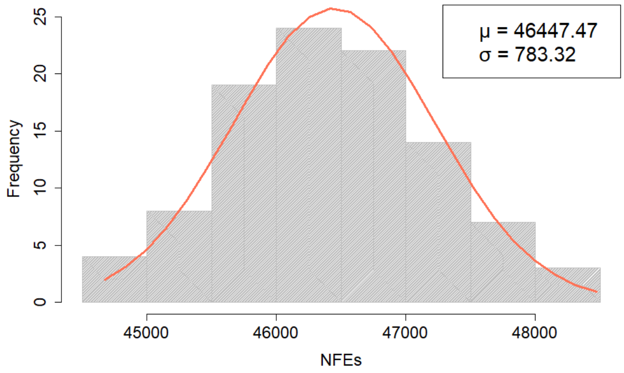

Figure 1: The normal distribution for the dimension $D = 10$.

$$\sigma = \sqrt{\frac{\Sigma(x_i - \mu)^2}{n}} \tag{2}$$

To be able to predict the stopping condition/the control parameter for $D = 50$, we need to establish the predictive model based on the parameters of the statistical distribution. The parameters follow a trend, in our case they follow a polynomial trend shown in Eq. (3). The a, b, and c are the real numbers. To be able to correctly determine, which trend line is being followed by the data we use the R^2 value. This serves as a valuable indicator to determine the optimal curve fit for the provided parameters [2]. If the R^2 is close to 1, this indicates a very good fit. In our case, R^2 was 0.9931. In Eq. (4) the predictive model for the parameter μ is established. The established predictive model is shown in Fig. 2. Eq. (5) shows the trend line fitted to the data and will be used for further calculations.

$$y_{solver} = a \cdot x^2 \quad b \cdot x \quad c \tag{3}$$

$$\mu_{L-SHADE}(D) = 426.31 \cdot D^2 - 2382 \cdot D \quad 24{,}716 \tag{4}$$

The second parameter of the statistical distribution is the standard deviation (σ). We also established a predictive model for (σ) following the same procedure for $D = \{5, 10, ..., 40\}$. The predictive model is shown in Eq. (5) and in Fig. 3.

$$\sigma_{L-SHADE}(50) = 87.774 \cdot D^2 - 2146.2 \cdot D \quad 11{,}957 \tag{5}$$

Firstly, we will predict μ and σ for the $D = 50$. The calculations are shown in Eqs. (6) and (7).

$$\mu_{L-SHADE}(50) = 426.31 \cdot 50^2 - 2382 \cdot 50 \quad 24{,}716 = 971{,}391 \tag{6}$$

$$\sigma_{L-SHADE}(50) = 87.774 \cdot 50^2 - 2146.2 \cdot 50 \quad 11{,}957 = 112{,}136 \tag{7}$$

We predicted both parameters μ and σ. However, to determine the stopping condition, we also need to know with what probability do we want the optimal solutions to be reached. Here, we will need the Z-score table [11], with which we can estimate the probability. In our case, we are calculating the control parameter/stopping

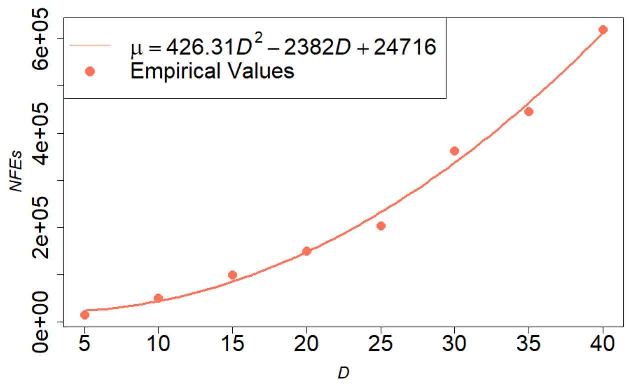

Figure 2: Prediction model for the mean (μ) of *NFEs* for L-SHADE on observed dimensions $D = \{5, 10, ..., 40\}$. The R^2 is 0.99**.**

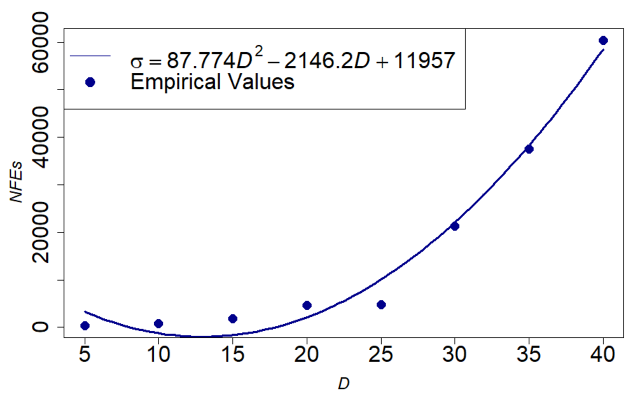

Figure 3: Prediction model for the standard deviation (σ) of *NFEs* for L-SHADE on observed dimensions $D = \{5, 10, ..., 40\}$. The R^2 is 0.98**.**

condition, so we will use the Eq. (8). We want to predict the stopping condition with 99% probability. For this purpose, we will use the Z-score table and check what the value is for 99% probability. The value at 99% probability is 3.1, so we will use this value for our further calculations.

$$Z = \frac{maxFEs - \mu}{\sigma} \tag{8}$$

$$maxFEs(99\%) = \mu \quad Z \cdot \sigma = 971{,}391 \quad 3.1 \cdot 112{,}136 = 1{,}319{,}012 \tag{9}$$

In Eq. (9), we show the prediction of the stopping condition/control parameter *maxFES*. The stopping condition *maxFES* to reach the optimal solutions with L-SHADE on the given benchmark function and 99% probability is 1,356,045 *maxFES*.

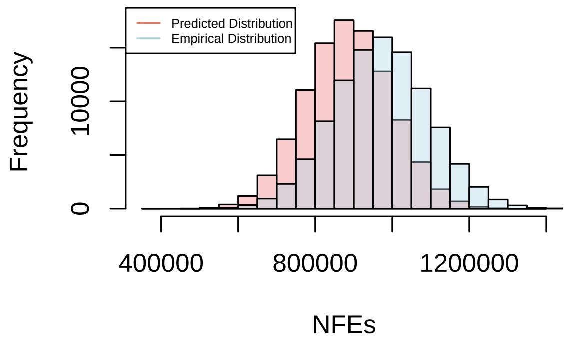

Figure 4: The comparison of the empirical and predicted statistical distribution for $D = 50$.

To empirically validate these results, we measured the hit ratio. Hit ratio represents the relationship between the number of successful runs and all runs. With a 99% hit ratio, we obtained 1,118,419 *maxFEs*. With this, we show the usefulness of the proposed approach. The gap between the empirical and predicted values is 15%. We can apply this approach for any probability. For this purpose, we also show predicted and empirical results for the probability of 50%. The predicted and empirical *maxFEs* for the probability/hit ratio 50% for $D = 50$ are 971,391 and 877,252. The gap between the empirical and predicted value is 10%. In our examples, it is evident that the predicted values are higher than the empirical ones. This shows that our predictive model is slightly pessimistic.

To show how well the predicted and empirical mean (μ) of the *NFEs* match, we will utilize another aspect of the AS^3D approach. Since we predict the parameters of the statistical distribution, we also predict the statistical distribution. In Fig. 4, we show how well the predicted and empirical statistical distributions match.

This experiment indicates that the control parameter for L-SHADE on the chosen benchmark function can be predicted by utilizing the AS^3D approach. However, this approach also enables us to:

- Estimate the probability with which a (sub)-optimal solution can be reached according to the preset stopping conditions.
- Analyze and compare the chosen evolutionary algorithms and optimization problems.

Still, we need to take into the account some limitations of the approach. Firstly, the hit ratio needs to be 100%. This means that the algorithm reaches a given quality of solutions for each independent run. Since the predictive model is established on smaller dimension of the optimization problem and the prediction is made for the larger dimensions, the chosen optimization problem needs to be multidimensional. It can also occur that the parameters of the statistical distribution do not follow any recognizable trend line. This means that the prediction cannot be made.

Overall, the control parameter/stopping condition can be set by the proposed approach.

4 CONCLUSION

In conclusion, the control parameter maximum number of function evaluations (*maxFEs*) holds a significant role in the optimization process. It impacts the parameter population size of an evolutionary algorithm and it serves as a stopping criteria. This paper focused on setting the control parameter *maxFEs* for the L-SHADE algorithm and selected Large-Scale benchmark function. Through the experiment, we show that the recently proposed approach is appropriate for predicting the control parameter of L-SHADE by establishing a predictive model based on smaller dimensions of the chosen benchmark function. The stopping condition is predicted by considering the 99% probability. The calculations were empirically validated by comparing the empirical and predicted values of the control parameter.

In conclusion, by using this approach, we not only set stopping condition, but also provide probability with which the chosen quality of solution can be reached.

ACKNOWLEDGEMENT

This work was supported by the Slovenian Research Agency (Computer Systems, Methodologies, and Intelligent Services) under Grant P2-0041.

REFERENCES

[1] A. W. Mohamed, A. A. Hadi, A. K. Mohamed, P. Agrawal, A. Kumar, P. N. Suganthan. December 2021. *Problem Definitions and Evaluation Criteria for the CEC 2022 Special Session and Competition on Single Objective Bound Constrained Numerical Optimization.* Technical Report. Nanyang Technological University, Singapore. https://github.com/P-N-Suganthan/2022-SO-BO

[2] Sandra Arlinghaus. 1994. *Practical handbook of digital mapping terms and concepts.* CRC Press.

[3] Anne Auger and Nikolaus Hansen. 2023. An Introduction to Scientific Experimentation and Benchmarking. In *Proceedings of the Companion Conference on Genetic and Evolutionary Computation.* 854–877.

[4] Subhodip Biswas, Debanjan Saha, Shuvodeep De, Adam D Cobb, Swagatam Das, and Brian A Jalaian. 2021. Improving Differential Evolution through Bayesian Hyperparameter Optimization. In *2021 IEEE Congress on Evolutionary Computation (CEC).* IEEE, 832–840.

[5] Janez Brest, Mirjam Sepesy Maučec, and Borko Bošković. 2016. iL-SHADE: Improved L-SHADE algorithm for single objective real-parameter optimization. In *2016 IEEE Congress on Evolutionary Computation (CEC).* IEEE, 1188–1195.

[6] Janez Brest, Mirjam Sepesy Maučec, and Borko Bošković. 2017. Single objective real-parameter optimization: Algorithm jSO. In *2017 IEEE Congress on Evolutionary Computation (CEC).* IEEE, 1311–1318.

[7] Borja Calvo, Ofer M Shir, Josu Ceberio, Carola Doerr, Hao Wang, Thomas Bäck, and Jose A Lozano. 2019. Bayesian performance analysis for black-box optimization benchmarking. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion.* 1789–1797.

[8] Joaquín Derrac, Salvador García, Daniel Molina, and Francisco Herrera. 2011. A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms. *Swarm and Evolutionary Computation* 1, 1 (2011), 3–18.

[9] Salvador García, Alberto Fernández, Julián Luengo, and Francisco Herrera. 2010. Advanced Nonparametric Tests for Multiple Comparisons in the Design of Experiments in Computational Intelligence and Data Mining: Experimental Analysis of Power. *Information sciences* 180, 10 (2010), 2044–2064.

[10] Nikolaus Hansen, Anne Auger, Dimo Brockhoff, and Tea Tušar. 2022. Anytime performance assessment in blackbox optimization benchmarking. *IEEE Transactions on Evolutionary Computation* 26, 6 (2022), 1293–1305.

[11] Homer Thornton Hayslett. 2014. *Statistics.* Elsevier.

[12] Jana Herzog, Janez Brest, and Borko Bošković. 2023. Analysis based on statistical distributions: A practical approach for stochastic solvers using discrete and continuous problems. *Information Sciences* 633 (2023), 469–490. https://doi.org/10.1016/j.ins.2023.03.081

[13] Thomas Jansen. 2020. Analysing stochastic search heuristics operating on a fixed budget. *Theory of evolutionary computation: recent developments in discrete optimization* (2020), 249–270.

[14] Anezka Kazikova, Michal Pluhacek, and Roman Senkerik. 2021. How Does the Number of Objective Function Evaluations Impact Our Understanding of Metaheuristics Behavior? *IEEE Access* 9 (2021), 44032–44048. https://doi.org/10.1109/ACCESS.2021.3066135

[15] Antonio LaTorre, Daniel Molina, Eneko Osaba, Javier Poyatos, Javier Del Ser, and Francisco Herrera. 2021. A Prescription of Methodological Guidelines for Comparing Bio-inspired Optimization Algorithms. *Swarm and Evolutionary Computation* (2021). https://doi.org/10.1016/j.swevo.2021.100973

[16] Xiaodong Li, Ke Tang, Mohammad N Omidvar, Zhenyu Yang, Kai Qin, and Hefei China. 2013. Benchmark Functions for the CEC 2013 Special Session and Competition on Large-scale Global Optimization. *gene* 7, 33 (2013), 8.

[17] Mattos, David Issa and Bosch, Jan and Olsson, Helena Holmström. 2021. Statistical Models for the Analysis of Optimization Algorithms With Benchmark Functions. *IEEE Transactions on Evolutionary Computation* 25, 6 (2021), 1163–1177. https://doi.org/10.1109/TEVC.2021.3081167

[18] Miha Ravber, Shih-Hsi Liu, Marjan Mernik, and Matej Črepinšek. 2022. Maximum number of generations as a stopping criterion considered harmful. *Applied Soft Computing* 128 (2022), 109478. https://doi.org/10.1016/j.asoc.2022.109478

[19] Ryoji Tanabe and Alex S. Fukunaga. 2014. Improving the search performance of SHADE using linear population size reduction. In *2014 IEEE Congress on Evolutionary Computation (CEC).* 1658–1665. https://doi.org/10.1109/CEC.2014.6900380

[20] Niki Veček, Marjan Mernik, and Matej Črepinšek. 2014. A chess rating system for evolutionary algorithms: A new method for the comparison and ranking of evolutionary algorithms. *Information Sciences* 277 (2014), 656–679.