20th International Conference on the Application of Computer
Science and Mathematics in Architecture and Civil Engineering
K. Gürlebeck and T. Lahmer (eds.)
Weimar, Germany, 20–22 July 2015

Benchmarking of Optimization Algorithms

J. Meier *, **T. Schanz**

* *Gruner AG*
Basel, Switzerland
E-mail: joerg.meier@vint.de

Keywords: Optimization, Parameter Back Calculation, Benchmarking.

Abstract. *In this paper, we present an empirical approach for objective and quantitative benchmarking of optimization algorithms with respect to characteristics induced by the forward calculation. Due to the professional background of the authors, this benchmarking strategy is illustrated on a selection of search methods in regard to expected characteristics of geotechnical parameter back calculation problems. Starting from brief introduction into the approach employed, a strategy for optimization algorithm benchmarking is introduced. The benchmarking utilizes statistical tests carried out on well-known test functions superposed with perturbations, both chosen to mimic objective function topologies found for geotechnical objective function topologies. Here, the moved axis parallel hyper-ellipsoid test function and the generalized Ackley test function in conjunction with an adjustable quantity of objective function topology roughness and fraction of failing forward calculations is analyzed. In total, results for 5 optimization algorithms are presented, compared and discussed.*

1 MOTIVATION

For numerical simulations, it is essential to use a model parameter set which generates a realistic system response. In practice, parameter back calculation based on the direct approach is often used for this purpose in which mathematical optimisation algorithms are a critical component. Many different optimization algorithms are known and are available within the literature (for an overview, see e.g. [11] or [7]). These algorithms use a variety of approaches to perform the search for optimal parameter combinations. However, the performance and convergence of the different optimization algorithms itself varies strongly, and also depend strongly on the optimization problem to solve. Exorbitant computational costs or, in the worst case, an improper or random parameter set will be returned if an unsuitable optimization algorithm is selected. Therefore, a suitable optimization algorithm has to be carefully selected for each problem.

For optimization algorithms, the performance by means of "finding the optimum reliable" and "low computational cost" cannot be assessed in a closed mathematical form in most of the cases. However, we may use a statistical analysis of the solution obtained from optimization runs to make conclusions on some properties of the optimization algorithm and the search-performance. For this approach, the success rates of finding the optimum is treated as a stochastic value and statistical measures are applicable. Hereafter the approach used and some results are presented.

2 EMPIRICAL BENCHMARKING APPROACH

The parameter identification approach using the direct back analysis method consists of an iterative procedure controlled by an optimization algorithm. The model parameters are iteratively changed in such a way to achieve better agreement between the model results and the measured values, e.g. the field measurements. This agreement (or disagreement) is measured by the objective function $f(x)$. The aim of the optimization algorithm is therefore to iteratively minimize the objective function value.

The computational cost caused by this iterative process is mainly influenced by the number of forward calculations (usually numerical simulations) requested by the optimization algorithm. The processor usage of the optimization algorithm itself on the other hand is usually neglectable. A typical single forward calculation used in Geotechnics (e.g. Finite Element model of an excavation pit) requires often 5 min calculation time or more while the optimization algorithm itself needs less than 1 sec. Minimizing the computational cost of an optimization sequence is therefore equivalent with minimizing the number of forward calculations.

For the statistical approach applied here, no "real life" forward calculation has been used to avoid high calculation time. Therefore, the normal calculation scheme for an objective function value has been altered by substituting the time-consuming forward calculation by a well-defined and well-known test function. This substitution has no influence on the iterative process controlled by the optimization algorithm nor have the optimization algorithms to be adapted. From view of the optimization algorithm, an objective function value is still calculated based on a parameter vector.

As stated by [2] and also by the experience of the authors (e.g. [6, 8, 14]), many objective functions from the field of (geo-)technics have a globally convex shape, in which often secondary (locally optimal) solutions are present. Furthermore many objective functions show a certain 'roughness' or 'noise' at a smaller scale. Additionally for some parameter vectors a forward calculation can fail reproducible (e.g. no convergence in the Finite Element Method) and no objective function value can be calculated accordingly.

In view of these facts, two test functions have been chosen for the benchmarking presented: Firstly, the Moved axis parallel hyper-ellipsoid function, which has no secondary optima and one global optimum [10]. Secondly, the Ackley test function in its generalized form, which shows several secondary optima of varying objective function values and a single global optimum (see [1] for details).

Both test functions exhibit no roughness or failed parameter vectors. In order to incorporate both characteristics the original test function $f(x)$ is superposed with a noise field $r(x)$ according to Eq. (1) and (2). In Eq. (2) *srnd*(■) is a pseudo-random number generator returning equal-distributed numbers ranging from 0.0 to 1.0, while for one and the same argument (■) always the same number is retuned. The control variable τ is the noise scaling factor. According to Eq. (3) a parameter vector x is considered "failed" if the pseudo-random number for x is smaller or equal to a predefined failure probability p_f.

$$f^*(x) := f(x) + r(x) \tag{1}$$

$$r(x) := \tau\left(\frac{1}{2} - \frac{1}{n}\sum_{i=1}^{n} \mathrm{srnd}(x_i)\right) \tag{2}$$

$$\mathrm{srnd}(x) \leq p_f \tag{3}$$

The majority of optimization algorithms will not find the exact location of the test function global optimum (x^*), but will rather move asymptotically towards x^* due to underlying paradigms. The optimization sequence is considered to be successful, if the parameter set x_{min} with the smallest objective function value is located within Ψ as defined by Eq. (4).

$$\Psi = \left\{ x_{min} \;\middle|\; \|x_{min} - x^*\|_2 \leq d_\Psi \right\} \tag{4}$$

The search range Ω and d_Ψ has been chosen as follows (please note that the relative size of Ψ for both test functions is equal compared to Ω):

- Ackley Test Function: $-1.0 \leq x_i \leq +2.0$ with $d_\Psi = 0.1$
- Moved axis parallel hyper-ellipsoid function: $-10.0 \leq x_i \leq +20.0$ with $d_\Psi = 1.0$

To assess the probability of which an optimization algorithm is able to converge within Ψ on a test function for given values of p_f and τ, a large number of optimization sequences is repeatedly run. For each sequence, the start parameter sets are chosen randomly within Ω and the number of forward calculation is limited to 500. The quotient of successful sequences over the total number of sequences is considered as success rate p. The number of optimization sequences is increased until the success rate is stabilizing, what usually corresponds to some 10'000 runs.

3 RESULTS

The empirical benchmarking approach described above has been applied to 5 selected optimization algorithms. Namely, Monte-Carlo method (MC), a gradient descent method (GD) (e.g. [12, 11]), an evolutionary-genetic algorithm (EG) [e.g. 11], the Simplex-Nelder-Mead optimizer (SNM) [9] and the particle swarm optimizer (PSO) [5, 4]. For each algorithm both test functions have been used with $n = \{2, 3, 4, 6, 8 \text{ and } 10\}$ unknown parameters.

The diagrams of Figure 3 show on the vertical axis the success rate p. over the noise control variable τ and, respectively, the failure rate p_f. The main conclusions are:

- The MC method performs well for $n = 2$. For higher n it clearly suffers from the "curse of dimensionality" [3].

- The GD performs works very nice for smooth objective function topologies with no secondary optima. It performs very badly if secondary optima are present, as it is the case for the Ackley test function. The GD is also not robust to noise and failing forward calculations.
- The SNM optimizer is much more robust than the GD. Nevertheless, due to its local character the success rate for the Ackley test function is ~50% even for the ideal case of $\tau = 0$ and $p_f = 0$.
- The EG algorithm class is very popular among many researchers due to its high robustness. This robustness is also visible in Figure 3. The major drawback of this method is its need for a large number of forward calculations as also stated by [13].
- For both test function the PSO (10 particles) used shows the best performance values. It outperforms clearly all other tested algorithms including the EG method. This finding is in agreement with the experience of other researchers, e.g. [13].

The results of Figure 3 show clearly how different the optimization algorithms behave for the two test functions. This illustrates how the nonlinearity of the objective function has a strong influence to the performance of the optimization algorithm.

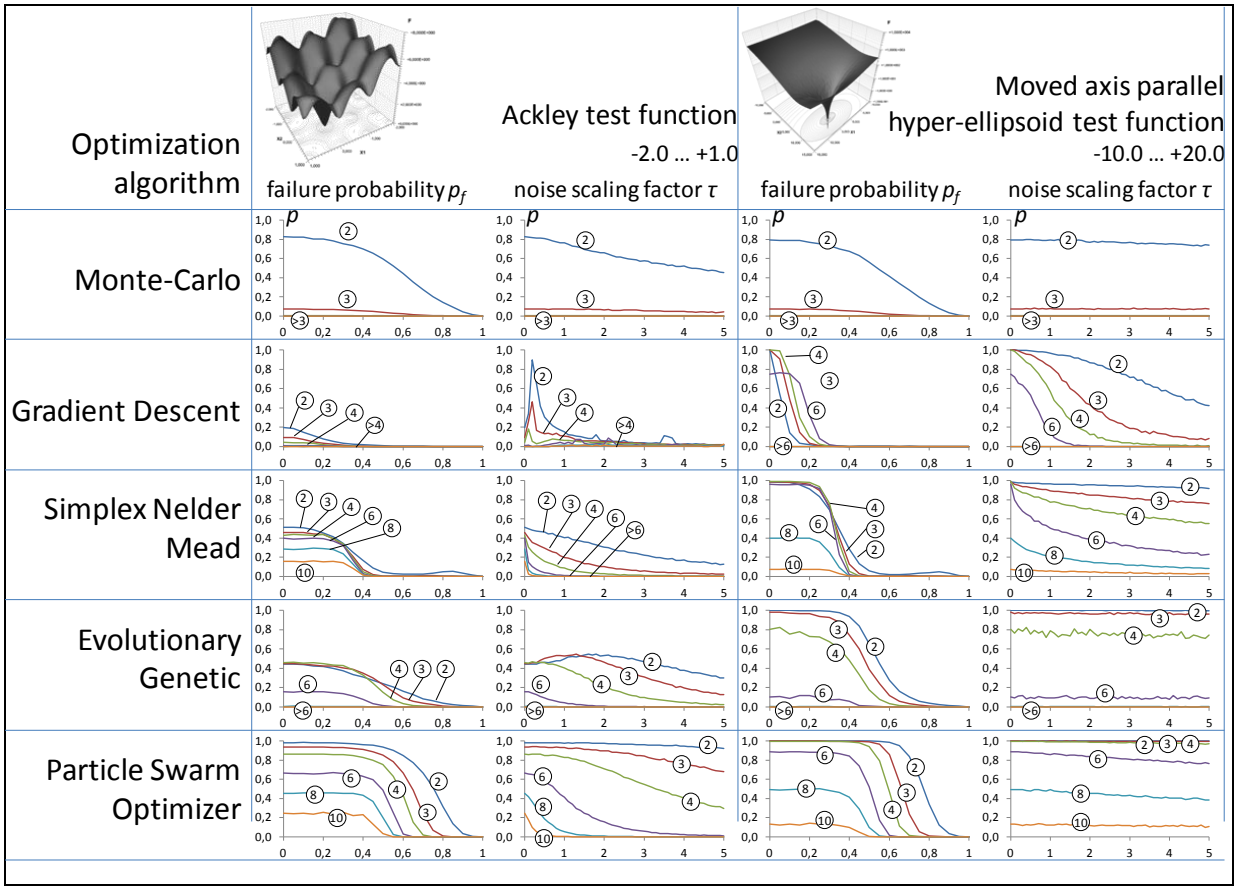

Figure 1: Results of the 5 tested optimization algorithms

4 CONCLUSION AND OUTLOOK

In the present paper an approach for impartial and quantitative benchmarking of optimization algorithms has been briefly presented and applied to 5 selected optimization algorithms. Of all the optimization algorithms tested, the PSO shows the best performance. However, if an objective function with a severely different nonlinearity is present, or if the number of unknown parameters n increases strongly the PSO may be outperformed by other algorithms.

The next step will be to apply the benchmarking to more algorithms and to define a rating function based on the performance profiles to provide an objective rating method for optimization algorithms.

REFERENCES

[1] Ackley, D. H. (1987): A connectionist machine for genetic hill climbing. Boston: Kluwer Academic Publishers.

[2] Chen, S. (2009): Locust Swarms - A New Multi-Optima Search Technique. 2009 IEEE Congress on Evolutionary Computation (CEC 2009), p. 1745 - 1752.

[3] Dixon, L.C.W.; Gomulka, J.; Hersom,S.E. (1976): Reflections on the global optimization problem. In: Dixon, L.C.W. (ed.) Optimization in action, pp. 398-435. London: Academic Press.

[4] EBERHARDT, R. C.; KENNEDY, J. (1995): A new optimizer using particle swarm theory. Proceedings of the Sixth International Symposium on Micro Machine and Human Science, Nagoya, Japan, IEEE Service Center, Piscataway, NJ, 39 – 43.

[5] KENNEDY, J.; EBERHARDT, R. C. (1995): Particle Swarm Optimization. Proceedings of IEEE International Conference on Neural Networks, Volume IV, S. 1942 – 1948, Perth, Australia, IEEE Service Center, Piscataway, NY.

[6] Knabe, T.; Schweiger, H. F.; Schanz, T. (2012): Calibration of constitutive parameters by inverse analysis for a geotechnical boundary problem. Canadian Geotechnical Journal 49(2): 170-183, 10.1139/t11-091

[7] MEIER, J. (2008): Parameterbestimmung mittels inverser Verfahren für geotechnische Problemstellungen. Dissertationsschrift, Fakultät Bauingenieurwesen, Bauhaus-Universität Weimar.

[8] Most, T.; Knabe, T. (2009): Reliability analysis of the bearing failure problem considering uncertain stochastic parameters. Comput. Geotech (2009), doi:10.1016/j.compgeo.2009.11.003.

[9] NELDER, J. A.; MEAD, R. (1965): A simplex method for function minimization. Computational Journal 7, p. 308 - 313.

[10] Polheim, H. (1999): Evolutionäre Algorithmen-Verfahren, Operatoren und Hinweise für die Praxis. Springer.

[11] SPALL, J. C. (2003): Introduction to stochastic search and optimization: estimation, simulation and control. Wiley-Interscience series in discrete mathematics.

[12] SPELLUCCI, P. (1993): Numerische Verfahren der nichtlinearen Optimierung. Birkhäuser Verlag, Basel Boston Berlin, ISBN 3-7643-2854-1.

[13] Vaz, A. I. F.; Vicente, L. N. (2006): A particle swarm pattern search method for bound constrained nonlinear optimization. J Glob Optim (2007) 39:197–219, DOI 10.1007/s10898-007-9133-5.

[14] Zimmerer, M. M. (2010): Identifikation konstitutiver Parameter von weichen feinkörnigen Böden, Beitrag zum Konsolidationsverhalten von Ton, Dissertation, Bauhaus-Universität Weimar, Schriftenreihe des DFG Graduiertenkollegs 1462 Modellqualitäten (1).