Is Global Sensitivity Analysis Useful to Evolutionary Computation?*

Thomas Chabin, Alberto Tonda, Evelyne Lutton
UMR 782 GMPA, INRA 1 Av. Lucien Brétignères 78850 Thiverval-Grignon, FRANCE
{thomas.chabin,alberto.tonda,evelyne.lutton}@grignon.inra.fr

ABSTRACT

Global Sensitivity Analysis (GSA) studies how uncertainty in the inputs of a system influences uncertainty in its outputs. GSA is extensively used by experts to gather information about the behavior of models, through computationally-intensive stochastic sampling of parameters' space. Some studies propose to make use of the considerable quantity of data acquired in this way to optimize the model parameters, often resorting to Evolutionary Algorithms (EAs). Nevertheless, efficiently exploiting information gathered from GSA might not be so straightforward. In this paper, we present a counterexample followed by experimental results to prove how naively combining GSA and EA can bring about negative outcomes.

Categories and Subject Descriptors

I.2.8 [**Computing Methodologies**]: Artificial Intelligence—*Problem Solving, Control Methods, and Search*

Keywords

Global sensitivity analysis, evolutionary computation, EASEA, real-valued optimization

1. INTRODUCTION

Global Sensitivity Analysis

One of the most common approaches to GSA has been developed by Sobol [3], and provides the impact of each individual decision variable and its interactions with other variables on performance objectives, using sensitivity indices. GSA is mainly used for two goals: *factor prioritizing*, to decide which variable uncertainty to work on, in order to reduce the uncertainty of the output the most; and *factor fixing*, to highlight which variables can be fixed to an arbitrary value without influencing much the output.

*This work has been funded by the French National Agency for research (ANR), under the grant ANR-11-EMMA-0017, EASEA-Cloud Emergence project 2011,

GECCO'15, July 11–15,2015,Madrid,Spain.

ACM ISBN 978-1-4503-3472-3/15/07..

DOI: http://dx.doi.org/10.1145/2739480.2764675

Each of the following types of indices takes values between 0 and 1, and represents a proportion of influence.

First-order sensitivity indices are used for the factor priority problem. They represent the direct influence of the uncertainty of a parameter onto the variance of an output.

Higher-orders sensitivity indices can also be computed for every set of parameters with a high computational cost and therefore will not be considered in this work. Such indices represent how much the combined action of the set of parameters is directly responsible for the variance in the output.

Total-effect sensitivity indices are used for the factor-fixing problem. A total-effect index is attributed to each parameter, and is interpreted as the sum of all n-order indices involving the considered parameter. A parameter with a total-effect index near zero can be fixed to an arbitrary value inside his interval of uncertainty without considerably affecting the variance of the output.

Sensitivity Analysis and Optimization

SA is aiming at finding the parameters whose variation influences the output of a function (or a model) the most. It is therefore not surprising that several attempts have been performed to combine SA with optimization tools.

In [2], the authors use GSA measurements to reduce the problem's dimensionality, first optimizing the values of a sub-set of the most sensitive parameters, and then restarting the evolution from the solutions found in this way, finally optimizing the remaining values. This approach was found to lead to poor results in some cases.

In the following sections, a counterexample to this method is presented allong with an experimental analysis. An explanation to this behaviour is presented in the discution.

2. EXPERIMENTAL ANALYSIS

Isolating a relevant subset of parameters

The tested optimization strategy relies on the following statement (factor fixing approach, see section 1): a low total effect index reveals a non-influent parameter that can be arbitrarily fixed with a small impact on the fitness function.

To decide which parameters are non-influential, a threshold is arbitrarily fixed (a low value in the range $[0, 1]$): parameters that have a total sensitivity index below this threshold are considered non-influential.

Algorithm

The method was tested with an explicit population based EA, programmed thanks to the EASEA package [1].

The approach is based on [2]. Influential parameters are optimized in a first stage, while the non-influential parameters are fixed at the middle of their interval of uncertainty.

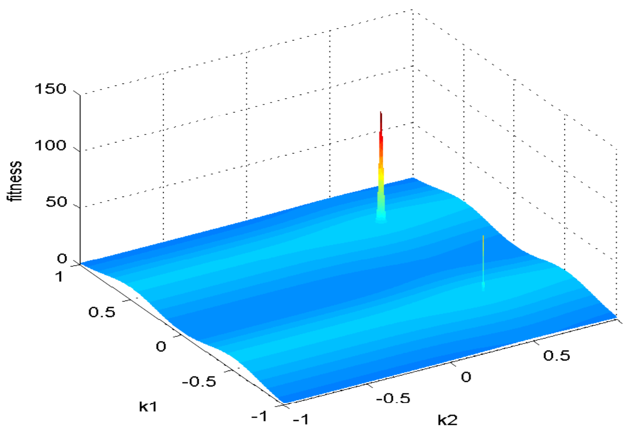

Figure 1: Counterexample II. There are two thin peaks, a very thin one corresponding to a local optimum is located at $(-0.5, 0.5)$ and a larger one, global optimum at $(0.5, 0.5)$.

The best points of the last generation are injected in the initial population of a second optimization using all parameters. The non-influencial parameters who were fixed are initialized with a random value.

Counterexample

We built a counterexample fitness function as follows:

$$fit_2(k1, k2) = g(k1, 10.9, 0.5, 0.25) + g(k1, 11, -0.5, 0.25) + g(k2, 1, 0.5, 0.25) + g2d(k1, k2, 100, 0.5, 0.01, 0.5, 0.01) + g2d(k1, k2, 50, -0.5, 0.0025, 0.5, 0.0025)$$

with : g, a Gaussian function: $g(k, a, b, c) = a \times exp(-\frac{(k-b)^2}{2c^2})$ $g2d$, a 2D Gaussian function: $g2d(k1, k2, a, b, c, d, e) = a \times exp(-(\frac{(k1-b)^2}{2\times c^2} + \frac{(k2-d)^2}{2\times e^2}))$, and $k1, k2 \in \big[-1; 1\big]$

This fitness function is displayed in Fig.1 contains 2 optima : a local optimum, located at $(k1 = -0.5; k2 = 0.5)$, and a global optimum, located at $(k1 = 0.5; k2 = 0.5)$.

A GSA on this function, shows that $k1$ can be considered as an influential parameter (Total effect $k_1 = 0.98$) and $k2$ as a non-influential one (Total effect $k_2 = 0.09$).

A progressive refinement is compared to a plain optimization (full search space) using a classical EA, with the parameter settings reported in Table 1. Over 100 runs, the full search always finds the global optimum whereas the restart strategy always get stuck on the local optimum (Figure 2).

Population size	$\mu = 2000$
Offsprings size	$\lambda = 1800$
Number of generations	full search : 250 Approach 2 : 50 then 200
Tournament selection	$Size = 2$
BLX-α Crossover	$p = 1.$
Log normal self adaptive mutation	$p = 1.$ $\tau = \sqrt{2}$
Number of Runs	100

Table 1: EA parameter setting for the optimization, full search space and the restart strategy, for the counterexample.

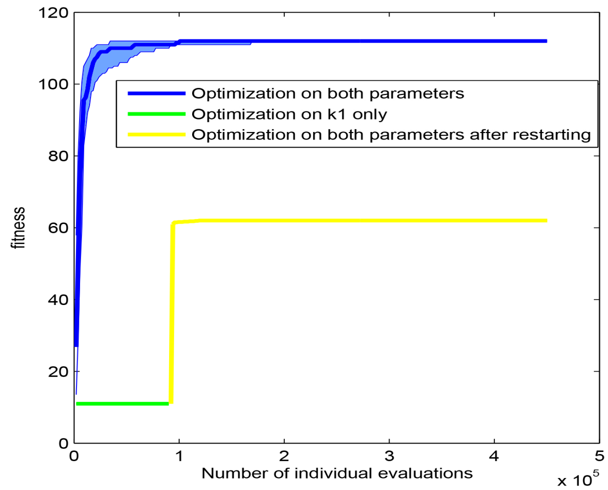

Figure 2: Statistics of 100 runs on the Counterexample with a classical EA.

3. DISCUSSION AND CONCLUSIONS

The counterexample sheds light on the fact that sensitivity analysis may deliver misleading information to the optimization process. A possible explanation is that GSA provides an averaged viewpoint on each parameter. It is clear that averaging may hide interesting irregular areas where optima may be found. Another problem is due to the fact that the results of a GSA may drastically vary with the choice of the parameter range. Additionally, the question of an efficient use of GSA inside an optimization procedure is raised: GSA is, in itself, extremely time consuming, and this cost has not been taken into account in the previous experiments. Finally, the approach presented in Section 2, based on the one proposed in [2], can only be used on a function containing both influencial and non-influencial parameters identified by GSA.

To conclude, we presented a case study, specifically designed to provide deceiving information to sensitivity analysis used during an optimization process. As a result, stochastic optimization biased by this information, based on the method in [2] has been experimentally proven unable to reach the global optimum on a counterexample.

4. REFERENCES

[1] P. Collet, E. Lutton, M. Schoenauer, and J. Louchet. Take it easea. In *Parallel Problem Solving from Nature PPSN VI*, pages 891–901. Springer, 2000.

[2] G. Fu, Z. Kapelan, and P. Reed. Reducing the complexity of multiobjective water distribution system optimization through global sensitivity analysis. *Journal of Water Resources Planning and Management*, 138(3):196–207, 2011.

[3] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. *Global Sensitivity analysis, The Primer*. John Wiley & Sons, 2008.