Multi-Fidelity Surrogate Model Approach to Optimization

Sander van Rijn ✉
LIACS, Leiden University
Leiden, the Netherlands
s.j.van.rijn@liacs.leidenuniv.nl

Sebastian Schmitt
Honda Research Institute Europe
Offenbach am Main, Germany
sebastian.schmitt@honda-ri.de

Markus Olhofer
Honda Research Institute Europe
Offenbach am Main, Germany
markus.olhofer@honda-ri.de

Matthijs van Leeuwen
LIACS, Leiden University
Leiden, the Netherlands
m.van.leeuwen@liacs.leidenuniv.nl

Thomas Bäck
LIACS, Leiden University
Leiden, the Netherlands
t.h.w.baeck@liacs.leidenuniv.nl

ABSTRACT

Recently the use of Radial Basis Functions (RBF) has been introduced as an optional alternative to co-Kriging in the context of multi-fidelity surrogate modeling. In this paper, we compare the performance of Random Forest-based *co-surrogates* to the previously introduced co-Kriging and co-RBF using a set of bi-fidelity benchmark problems in 2, 4 and 8 dimensions. Our results show that there is a minimal overall difference between the different co-surrogate models with regards to final performace, although the training of Random Forests takes much less time compared to the Kriging and RBF methods.

CCS CONCEPTS

• Theory of computation → Random search heuristics; **• Computing methodologies → Machine learning approaches**;

KEYWORDS

Evolution Strategies, Empirical study, Surrogate Modeling

ACM Reference Format:
Sander van Rijn ✉, Sebastian Schmitt, Markus Olhofer, Matthijs van Leeuwen, and Thomas Bäck. 2018. Multi-Fidelity Surrogate Model Approach to Optimization. In *GECCO '18 Companion: Genetic and Evolutionary Computation Conference Companion, July 15–19, 2018, Kyoto, Japan.* ACM, New York, NY, USA, Article 4, 2 pages. https://doi.org/10.1145/3205651.3205757

1 INTRODUCTION

Optimization problems in engineering domains often rely on simulations to determine the quality of candidate solutions. Numerical optimization methods such as CMA-ES [5] allow automation of such processes, but require many hundreds or thousands of evaluations. As the required amount of runtime per simulation increases, more advanced optimization methods are needed that can work with fewer simulations.

To this end, Surrogate-Assisted optimization methods such as [3, 7] were developed that use surrogate models (also called metamodels) as a less computationally expensive substitute for the simulations. Co-Kriging was created by combining surrogate modeling with multi-fidelity simulations by [4], assuming that a lower fidelity simulation contains some global fitness landscape information.

In this paper we introduce co-Random Forests to the set of co-surrogates consisting only of co-Kriging and co-RBF [2]. We compare the performance of these three co-surrogates on eight benchmark functions as used in co-Surrogate-Assisted CMA-ES (cSA-CMA-ES) optimization.

Algorithm 1 cSA-CMA-ES

$t, \lambda_{\text{pre}}, \vec{x}, \vec{x}_{\text{best}}, y_{\text{best}} \leftarrow 0, 2\lambda$, random individual, Null, ∞
DoE $\leftarrow$ Design of Experiments sample
$\mathbb{A} \leftarrow (\textbf{DoE}, f_h(\textbf{DoE}), f_l(\textbf{DoE}))$
while not terminate **do**
    $\rho \leftarrow$ coefficient of linear regression on $\mathbb{A}_h, \mathbb{A}_l$
    $S_{\text{co}} \leftarrow$ train surrogate on $\mathbb{A}_h - \rho\mathbb{A}_l$
    $O^t_{\text{pre}} \leftarrow \text{mutate}(\vec{x}, \lambda_{\text{pre}})$
    $\vec{y}_l \leftarrow f_l(O^t_{\text{pre}})$
    $\hat{\vec{y}}_h \leftarrow \rho\vec{y}_l + S_{\text{co}}.\text{predict}(O^t_{\text{pre}})$
    $O^t \leftarrow \text{preselect}(O'^t, \hat{\vec{y}}_h, \lambda)$
    if $t \mod g_{\textbf{int}} = 0$ **then**
        $\vec{y} \leftarrow f_h(O^t)$
        $\mathbb{A} \leftarrow \mathbb{A} \cup (O^t, \vec{y}_h, \vec{y}_l))$
        if $\min(\vec{y}) < y_{\text{best}}$ **then**
            $y_{\text{best}}, \vec{x}_{\text{best}} \leftarrow \min(\vec{y}), O^t_{\text{argmin}(\vec{y})}$
        end if
    else
        $\vec{y}_h \leftarrow \hat{\vec{y}}_h$
    end if
    $P^t \leftarrow \text{select}(O^t, \vec{y}_h, \mu)$
    $\vec{x} \leftarrow \text{recombine}(P^t)$
    updateInternalParameters(); $t \leftarrow t + 1$
end while

2 CO-SURROGATE ASSISTED CMA-ES

Co-surrogates are based on the autoregressive model by [6] that describes a high fidelity prediction as $\hat{f}_h(\vec{x}) = \rho f_l(\vec{x}) + \delta(\vec{x})$, where f_h and f_l are the high and low fidelity evaluation functions, $\hat{f}_h$

GECCO '18 Companion, July 15–19, 2018, Kyoto, Japan

ACM ISBN 978-1-4503-5764-7/18/07…$15.00
https://doi.org/10.1145/3205651.3205757

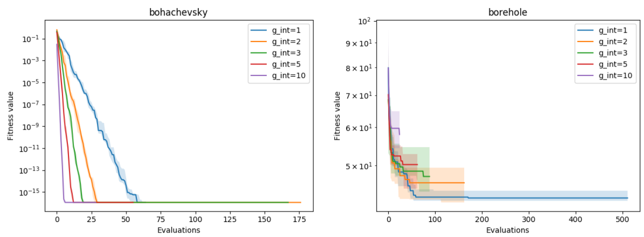

Figure 1: Comparison of generational evolution control intervals *g_int* in cSA-CMA-ES using Random Forests for two example benchmark functions (2d and 8d).

is the predicted value for f_h based on the co-surrogate model, ρ is a scaling parameter and $\delta(\vec{x})$ is an error term for the difference between the high and low fidelity results. Co-surrogates are used to predict this δ, which is calculated as $\delta(\vec{x}) = f_h(\vec{x}) - \rho f_l(\vec{x})$ to create training data for the co-surrogate.

Algorithm 1 shows the cSA-CMA-ES that we use in this paper. We compare it with a similar SA-CMA-ES, that mainly differs in lines 5–9 where the low fidelity function f_l is used for the high fidelity surrogate predictions. They make use of both *pre-selection* (line 10) and *generational* evolution control (lines 11–19).

3 EXPERIMENTS

For all experiments, an initial Latin-Hypercube-Sample of 20 points was evaluated as DoE to be used as initial training set for the surrogates. For pre-selection, λ_{pre} is fixed to 2λ. Each run has been repeated ten times. For RBF, we use the multiquadric basis function, and for Random Forests we create 100 trees. The used benchmark functions were taken from [1, 8].

Figure 1 shows an example of the influence of generational control intervals, while Figure 2 shows the comparison between Kriging, RBF and Random Forest surrogate models for both SA-CMA-ES and cSA-CMA-ES, with a generational interval fixed to 10. In these plots, the solid lines indicates the 50$^{\text{th}}$ percentile of the fitness values over 10 runs, while the shaded regions indicate the spread of the fitness values between the 25$^{\text{th}}$ and 75$^{\text{th}}$ percentile. Only high fidelity evaluations are counted.

4 CONCLUSIONS AND OUTLOOK

In this paper, we have compared (co-)surrogate-assisted optimization performance on a collection of analytical multi-fidelity benchmark functions from literature, with surrogates based on Kriging, RBF and Random Forests. Although Kriging is generally considered to be the most stable choice for a surrogate model, our results show that the performance of the various co-surrogates is generally similar, while any significant differences dependent on the target problem specifically, as is already known in the surrogate-assisted optimization literature.

ACKNOWLEDGMENTS

This work is part of the research program DAMIOSO with project number 628.006.002, which is partly financed by the Netherlands Organisation for Scientific Research (NWO).

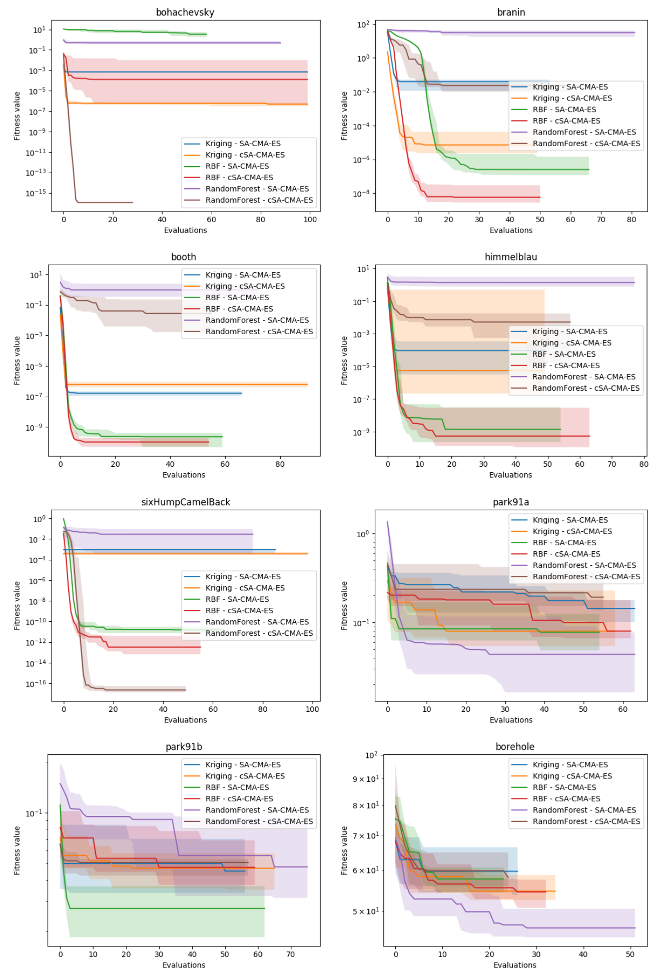

Figure 2: Comparison of Kriging, RBF and Random Forest based cSA-CMA-ES for all benchmark functions.

REFERENCES

[1] Huachao Dong, Baowei Song, Peng Wang, and Shuai Huang. 2015. Multi-fidelity information fusion based on prediction of kriging. *Structural and Multidisciplinary Optimization* 51, 6 (June 2015), 1267–1280. https://doi.org/10.1007/s00158-014-1213-9

[2] Cédric Durantin, Justin Rouxel, Jean-Antoine Désidéri, and Alain Glière. 2017. Multifidelity surrogate modeling based on radial basis functions. *Structural and Multidisciplinary Optimization* 56, 5 (Nov. 2017), 1061–1075. https://doi.org/10.1007/s00158-017-1703-7

[3] Michael Emmerich, Alexios Giotis, Mutlu Özdemir, Thomas Bäck, and Kyriakos Giannakoglou. 2002. Metamodel-Assisted Evolution Strategies. In *PPSN VII*. Springer Berlin Heidelberg, 361–370. https://doi.org/10.1007/3-540-45712-7_35

[4] Alexander I. J. Forrester, AndrÃąs SÃşbester, and Andy J. Keane. 2007. Multi-fidelity optimization via surrogate modelling. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 463, 2088 (Dec. 2007), 3251–3269. https://doi.org/10.1098/rspa.2007.1900

[5] N. Hansen and A. Ostermeier. 1996. Adapting arbitrary normal mutation distributions in evolution strategies: the covariance matrix adaptation. In *Proceedings of IEEE CEC*. 312–317. https://doi.org/10.1109/ICEC.1996.542381

[6] M. C. Kennedy and A. O'Hagan. 2000. Predicting the output from a complex computer code when fast approximations are available. *Biometrika* 87, 1 (March 2000), 1–13. https://doi.org/10.1093/biomet/87.1.1

[7] P. K. S. Nain and K. Deb. 2003. Computationally effective search and optimization procedure using coarse to fine approximations. In *The 2003 Congress on Evolutionary Computation, 2003. CEC '03*, Vol. 3. 2081–2088 Vol.3. https://doi.org/10.1109/CEC.2003.1299929

[8] S. Surjanovic and D. Bingham. [n. d.]. Virtual Library of Simulation Experiments: Test Functions and Datasets. Retrieved 2018-02-07: http://www.sfu.ca/~ssurjano. ([n. d.]).