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 performance, 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

Algorithm 1 cSA-CMA-ES
1: \( t, \lambda_{\text{pre}}, \hat{x}, \hat{y}_{\text{best}} \leftarrow 0, 2\lambda, \) random individual, Null, \( \infty \)
2: DoE \leftarrow Design of Experiments sample
3: \( A \leftarrow (\text{DoE}, f_{h}(\text{DoE}), f_{l}(\text{DoE})) \)
4: \textbf{while} not terminate \textbf{do}
5: \( \rho \leftarrow \text{coefficient of linear regression on } A, \hat{h}, \hat{l} \)
6: \( S_{\text{co}} \leftarrow \text{train surrogate on } A_{\hat{h}} + \rho A_{\hat{l}} \)
7: \( \hat{y}_{t} \leftarrow f_{l}(O_{t}^{\text{pre}}) \)
8: \( \hat{y}_{h} \leftarrow \rho \hat{y}_{t} + S_{\text{co}} \cdot \text{predict}(O_{t}^{\text{pre}}) \)
9: \( O_{t} \leftarrow \text{preselect}(O_{t}^{\prime}, \hat{y}_{h}, A) \)
10: \textbf{if } t \mod 9_{\text{int}} = 0 \textbf{ then}
11: \( \tilde{y} \leftarrow f_{l}(O_{t}^{\prime}) \)
12: \( \hat{A} \leftarrow \hat{A} \cup (O_{t}^{\prime}, \hat{y}_{h}, \tilde{y}) \)
13: \textbf{if } \min(\tilde{y}) < y_{\text{best}} \textbf{ then}
14: \( \hat{y}_{\text{best}} \leftarrow \min(\tilde{y}), O_{t}^{\prime} \arg\min(\tilde{y}) \)
15: \textbf{end if}
16: \textbf{end if}
17: \textbf{else}
18: \( \hat{y}_{h} \leftarrow \hat{y}_{h} \)
19: \textbf{end if}
20: \( P_{t} \leftarrow \text{select}(O_{t}^{\prime}, \hat{y}_{h}, \rho) \)
21: \( \hat{x} \leftarrow \text{recombine}(P_{t}) \)
22: \text{updateInternalParameters(); } t \leftarrow t + 1
23: \textbf{end while}

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.

2 CO-SURROGATE ASSISTED CMA-ES
Co-surrogates are based on the autoregressive model by [6] that describes a high fidelity prediction as \( \hat{f}_{h}(\hat{x}) = \rho \hat{f}_{l}(\hat{x}) + \delta(\hat{x}) \), where \( f_{h} \) and \( f_{l} \) are the high and low fidelity evaluation functions, \( \hat{f}_{h} \) and \( \hat{f}_{l} \) are the high and low fidelity predictions, respectively. The coefficients \( \rho \) and \( \delta(\hat{x}) \) are estimated by solving a linear regression problem on the data points \( \hat{A}_{\hat{h}} + \rho \hat{A}_{\hat{l}} \), where \( \hat{A}_{\hat{h}} \) and \( \hat{A}_{\hat{l}} \) are the data points from a high and low fidelity simulation, respectively.

To this end, Surrogate-Assisted optimization methods such as [3, 7] were developed that use surrogate models (also called meta-models) 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.
We compare it with a similar SA-CMA-ES, that mainly differs in was evaluated as DoE to be used as initial training set for the
generational (line 10) and
3 EXPERIMENTS
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).

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

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).

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).

REFERENCES