

# Feature Selection for Surrogate Model-Based Optimization

Frederik Rehbach  
frederik.rehbach@th-koeln.de  
TH Köln  
Cologne, Germany

Lorenzo Gentile  
lorenzo.gentile@th-koeln.de  
TH Köln  
Cologne, Germany

Thomas Bartz-Beielstein  
thomas.bartz-beielstein@th-koeln.de  
TH Köln  
Cologne, Germany

## ABSTRACT

We propose a hybridization approach called Regularized-Surrogate-Optimization (RSO) aimed at overcoming difficulties related to high-dimensionality. It combines standard Kriging-based SMBO with regularization techniques. The employed regularization methods use the least absolute shrinkage and selection operator (LASSO). An extensive study is performed on a set of artificial test functions and two real-world applications: the electrostatic precipitator problem and a multilayered composite design problem. Experiments reveal that RSO requires significantly less time than Kriging to obtain comparable results. The pros and cons of the RSO approach are discussed and recommendations for practitioners are presented.

## CCS CONCEPTS

• **Mathematics of computing** → **Discrete optimization**; • **Theory of computation** → **Continuous optimization**; **Gaussian processes**; • **Computing methodologies** → **Modeling and simulation**;

## KEYWORDS

Optimization, Surrogates, Modeling, Dimensionality Reduction, Real-World

### ACM Reference Format:

Frederik Rehbach, Lorenzo Gentile, and Thomas Bartz-Beielstein. 2019. Feature Selection for Surrogate Model-Based Optimization. In *Genetic and Evolutionary Computation Conference Companion (GECCO '19 Companion)*, July 13–17, 2019, Prague, Czech Republic. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3319619.3322020>

## 1 INTRODUCTION

Real-world optimization problems often have multiple characteristics that clearly distinguish them from artificial test-functions. Among others, two of the most concerning criteria are being high dimensional and being costly to evaluate. A popular approach for the optimization of costly functions is surrogate model-based optimization (SMBO). A more detailed explanation of SMBO and its applications can be found in [1]. Two research questions arise regarding the use of SMBO on high-dimensional costly black-box functions:

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

*GECCO '19 Companion*, July 13–17, 2019, Prague, Czech Republic

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-6748-6/19/07...\$15.00

<https://doi.org/10.1145/3319619.3322020>

(Q-1) How can SMBO efficiently be applied to high-dimensional objective functions?

The second research question stems from our ongoing research cooperations with industry. Simply delivering a black-box type algorithm which somehow manages to find good solutions for a given problem is neither ideal nor even accepted. In order to reach acceptance for a proposed new methodology in industry, algorithms have to deliver interpretable results. High-dimensional surrogate models are not well interpretable.

(Q-2) How to increase the interpretability of high-dimensional SMBO for its application in the industry?

## 2 PROBLEMS AND EXPERIMENTS

**The Electrostatic Precipitator Problem** is an industrial optimization problem that is part of ongoing research. A detailed explanation of the ESP-problem is given in [3].

**Sandwich-Structured Composite Plate Design Problem - (SCPD)** Sandwich panels are composite structures consisting of two thin laminate outer skins and a lightweight, e.g., honeycomb, thick core structure. Owing to the core structure, such composites are distinguished by stiffness. The objective of the optimization is then to minimize the displacement in the out-of-plane direction of one of the vertices of the plate, loaded by lifting and torquing loads, at the free edge varying the plies' lamination angle.

**Artificial Functions** In addition to the two discussed real-world problems, a multitude of artificial test-functions is utilized in this study. All chosen test-functions have a subset of important variables as well as a set of variables which is completely irrelevant for the optimization.

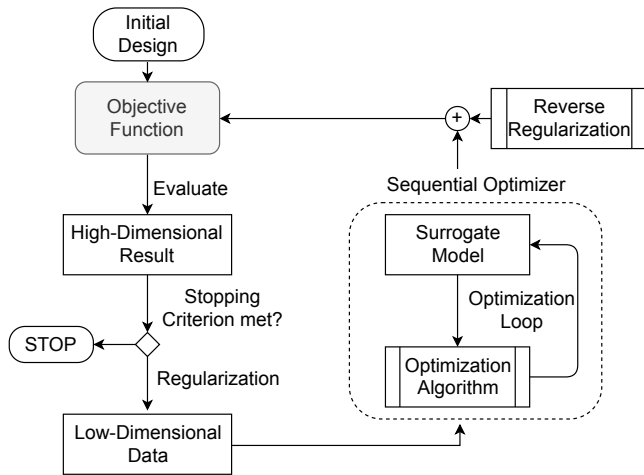
**Experiments** Each optimization run starts with an initial design size of five candidate solutions, randomly sampled from the design space. Since both of the described real-world problems are expensive simulation tasks, only a small budget is feasible. Each algorithm is given a total budget of 200 function evaluations and applied for 20 repetitions.

## 3 REGULARIZED SURROGATE OPTIMIZATION

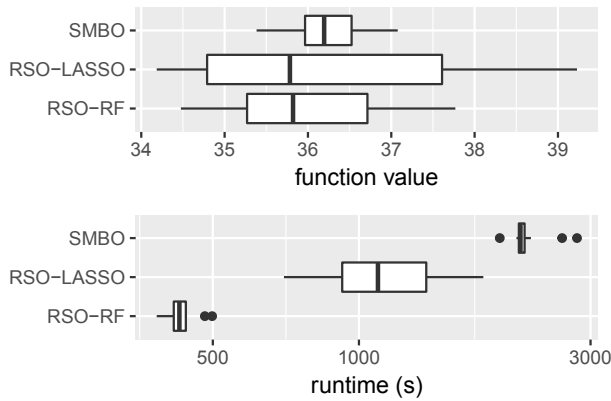
In order to apply SMBO efficiently to high-dimensional objective functions, we propose a hybrid algorithm: Regularized Surrogate Optimization (RSO). It consists of standard SMBO coupled with an additional regularization stage. The implementation of RSO is explained in Figure 1. In this study, two regularization methods are compared: LASSO [4] and a random forest [2] implementation.

## 4 RESULTS AND CONCLUSION

The results of the experiments on one of the real-world problems are shown in Figure 2. For both problems, the results indicate that



**Figure 1: Implementation of the RSO algorithm:** After the evaluation of an initial design, the resulting data is reduced through regularization. The lower-dimensional data is fitted with a surrogate model. An optimization algorithm searches for the best candidate solution on the surrogate model. The proposed candidate is merged with information from a reverse regularization method, back to the original dimensionality. The candidate is evaluated and the process is iterated until some stopping criterion is met.



**Figure 2: Algorithm comparison on the SCPD-Problem.** Upper plots show the performance in terms of best found function value, the lower plots show the total runtime of the algorithms. Lower values are better, thus more to the left is better in the figures.

switching to the RSO approach does not cause loss in the quality of the final function value. Yet, it is clearly visible that especially on the SCPD-problem the RSO approach results in a significant decrease in required algorithm runtime. The random forest aided variant RSO-RF seems to select fewer variables in the regularization step than RSO-LASSO. Therefore, it also requires less runtime. On the SCPD-Problem a more than 5x speedup can be observed. An

**Table 1: Performance Overview of all algorithms on each test-function.** Numbers shown are determined by rank-based pairwise multiple-comparison tests regarding the objective function value. The best result on each problem is marked in bold.

|               | SMBO        | RSO-LASSO   | RSO-RF      |
|---------------|-------------|-------------|-------------|
| ESP-Problem   | <b>1.00</b> | <b>1.00</b> | <b>1.00</b> |
| SCPD-Problem  | <b>1.00</b> | <b>1.00</b> | <b>1.00</b> |
| linketal06dec | <b>1.00</b> | <b>1.00</b> | 2.00        |
| linketal06sin | <b>1.00</b> | <b>1.00</b> | <b>1.00</b> |
| moon10hd      | <b>1.00</b> | <b>1.00</b> | 2.00        |
| moon10hdc1    | <b>1.00</b> | 2.00        | 2.00        |
| moon10hdc2    | <b>1.00</b> | <b>1.00</b> | 2.00        |
| moon10hdc3    | <b>1.00</b> | 2.00        | 2.00        |
| oakoh04       | <b>1.00</b> | 2.00        | 2.00        |
| morretal06    | <b>1.00</b> | 2.00        | 2.00        |
| Mean-Rank     | <b>1.00</b> | 1.40        | 1.70        |

overview of the statistical analysis of the results can be found in Table 1. A review of the results on the artificial test-functions shows that, while there was no quality difference on the two real-world problems, there is one for some of the test-functions. When there is a statistically significant quality difference between standard SMBO and the RSO approach, then often RSO-LASSO delivered better quality results than RSO-RF.

**(Q-1):** The results show that the proposed hybrid algorithm RSO requires significantly less computation time. At the same time it delivers, dependent on the respective optimization problem, the same or only slightly lower quality results. Most importantly, RSO is applicable even to very high-dimensional problems without the same risk of modeling failure that one would have compared to standard SMBO. The reduced modeling run-time can make the RSO algorithm feasible in situations where SMBO would be a bad choice due to time constraints.

**(Q-2):** It is often hard to suggest models or algorithms in industry, if they are acting in a black-box manner. Thus, if the model itself is based on a complex high-dimensional system, gaining support from the field engineers for a new optimization technique is hard. In contrast, a model which is able to correctly identify the few most important variables is easier to explain. The selected variables should mostly coincide with the practitioners field knowledge, confirming his understanding of a system and making it easy to gain support for the new algorithm implementation. The RSO algorithm delivers exactly this interpretability benefit for industry partners.

**REFERENCES**

[1] Thomas Bartz-Beielstein and Martin Zaefferer. 2017. Model-based methods for continuous and discrete global optimization. *Applied Soft Computing* 55 (2017), 154 – 167. <https://doi.org/10.1016/j.asoc.2017.01.039>

[2] Leo Breiman. 2001. Random forests. *Machine learning* 45, 1 (2001), 5–32.

[3] Frederik Rehbach, Martin Zaefferer, Joerg Stork, and Thomas Bartz-Beielstein. 2018. Comparison of Parallel Surrogate-Assisted Optimization Approaches. In *GECCO '18: Genetic and Evolutionary Computation Conference*.

[4] Robert Tibshirani. 2011. Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 73, 3 (2011), 273–282.