

Comparison of Kriging and Radial Basis Function surrogate models applied to a global optimization framework

Marcela A. Juliani¹, Wellison J. S. Gomes¹

¹*Center for Optimization and Reliability in Engineering (CORE), Department of Civil Engineering, Federal University of Santa Catarina*

*João Pio Duarte da Silva S., 88037-000, Santa Catarina, Brazil
marcelajulianii@gmail.com, wellison.gomes@ufsc.br*

Abstract. A literature survey reveals that many optimization problems present objective and/or constraints functions that demand high computational effort. Optimization algorithms which are able to solve these problems with just a few evaluations of such functions become necessary, in order to avoid prohibitive computational costs. In this context, there are a lot of surrogate models that can be employed to replace objective and/or constraints functions whenever possible, which are much faster to be evaluated than the original functions. In the present paper, a global optimization framework based on surrogate models is investigated, and two different surrogate models are considered: Radial Basis Function and Kriging. The framework consists of three search strategies, which may take place in each iteration of the optimization process: a local search, a global search and a refinement step. This optimization procedure is applied in benchmark problems from the literature and the results obtained by each surrogate model are compared. As a result, the framework was found to be considerably stable and to achieve satisfactory responses with both surrogates. Overall, among the cases analyzed, the framework based on Radial Basis Function showed better performance.

Keywords: Radial Basis Function, Kriging, Global optimization, Metamodels, Surrogate models.

1 Introduction

The application of complex computational models, which better represent the real behavior of a structure, raised challenges in the field of structural optimization. Usually, such models have objective and/or constraints functions that demand high computational effort. Therefore, the number of evaluations of such functions must be limited during the optimization process, to avoid prohibitive computational costs. If the problem presents many local minima, its solution becomes much more challenging, since application of global optimization procedures usually requires too many objective and constraints functions evaluations.

In the literature, it has been common to use surrogate models to deal with this type of optimization problem [1–3]. In this case, computationally expensive functions can be replaced by surrogate models which are much faster to be evaluated. Among the surrogate models used for this purpose, Radial Basis functions (RBFs) [4, 5] and Kriging [6, 7] can be highlighted.

In this context, the present paper compares the performance of these two surrogate models, RBF and Kriging, when applied to a global optimization framework proposed herein. The framework is composed by three different search strategies, which may occur in each iteration of the optimization process: a local search, based on the farthest apart subset concept, employs the surrogate model and looks for multiple local minima along the design space; a global search, which also employs the surrogate model, is performed by using a metaheuristic optimization method in an attempt to find the global minima; and a refinement step, which uses the real objective function and constraints, and aims at improving the best solution found so far. Eight benchmark problems from the literature are evaluated, including test functions and structural optimization of trusses.

The remainder of this paper is organized as follows: section 2 presents the surrogate models considered herein; the proposed global optimization framework is described in section 3; section 4 presents the application of the proposed approach in numerical examples; conclusions about the performance of the surrogate models and the framework are discussed in section 5.

2 Surrogate models

A surrogate model acts as a “curve fit” to a set of points previously defined by a real function. Thus, it is possible to use the surrogate model, *i.e.* the metamodel, in order to predict the results assumed by the original function without using it directly [8]. In the context of optimization, surrogate models can be used, for example, to replace objective and/or constraint functions, when these functions demand high computational efforts to be evaluated. The main idea is to obtain metamodels sufficiently accurate and with construction/evaluation time considerably shorter than the evaluations of the original functions.

Surrogate models are generated through a sampling plan $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_{n_{ED}})$ formed by n_{ED} points of the design space. Each one of these points can be associated with a value of the function to be replaced, such as the objective function $f(\mathbf{x})$ of an optimization problem. In this way, one can calculate the responses vector $\mathbf{y} = (y_1, \dots, y_{n_{ED}})$, where $y_i = f(\mathbf{x}_i)$, with $i = 1, \dots, n_{ED}$. From these data, it is possible to fit a surrogate model and to obtain predictions $\hat{y}(\mathbf{x}) \approx f(\mathbf{x})$ at any point \mathbf{x} from design space, via metamodel. There are several surrogate models that can be used for this purpose. RBF and Kriging are briefly presented in the following.

The prediction of a function at point \mathbf{x} through Radial Basis Function can be given by

$$\hat{y}(\mathbf{x}) = \sum_{j=1}^{n_{ED}} c_j \phi(\|\mathbf{x} - \mathbf{x}_j\|), \quad (1)$$

where $\phi(\|\mathbf{x} - \mathbf{x}_j\|)$ is the value of the radial kernel, $\|\mathbf{x} - \mathbf{x}_j\|$ is usually the Euclidean distance between the observational point, \mathbf{x} , and the center, \mathbf{x}_j , and c_j are the unknown coefficients which are determined by solving a linear system of equations depending on the interpolation conditions [9]. In this paper, the hybrid kernel presented in Mishra et al. [10] is adopted, which is a hybrid basis function that uses a combination of both the Gaussian and the cubic radial basis function. The hybrid kernel is given by $\phi(r) = e^{-(\epsilon r)^2} + \gamma r^3$, where ϵ is the shape parameter for the Gaussian kernel and γ controls the contribution of the cubic kernel. In order to obtain the combination of these parameters for which the performance of the hybrid kernel is the best, the parameters are defined by cross validation, where the Leave-one-out error is minimized, considering the formulation presented by Rippa [11]. This approach proved to be more accurate, stable and faster than many others commonly used.

On the other hand, in the Kriging metamodel the prediction at a point \mathbf{x} is given by

$$\hat{y}(\mathbf{x}) = \hat{\mu} + \boldsymbol{\psi}^T \boldsymbol{\Psi}^{-1} (\mathbf{y} - \mathbf{1}\hat{\mu}), \quad (2)$$

where $\hat{\mu}$ is the estimated mean of the stochastic process, $\boldsymbol{\psi}$ is the vector of correlations between the observed data and the new prediction, $\boldsymbol{\Psi}$ is the correlation matrix of all the observed data and $\mathbf{1}$ is a vector filled with ones. Kriging also requires the definition of a kernel function to compute the correlations, for which the commonly used Gaussian function is chosen herein. In the case of Kriging, the unknown parameters are usually found by using the Maximum Likelihood Estimate (MLE), performed in this paper by Particle Swarm Optimization (PSO) [12].

3 Global optimization framework

The type of optimization problem addressed herein consists of finding a vector of design variables \mathbf{x} that minimizes the objective function $f(\mathbf{x})$, subject to the constraints $g_j \leq 0$ and to the lower and upper bounds of each variable x_i , with $i = (1, \dots, n)$ and $j = (1, \dots, m)$, where n and m are the numbers of design variables and constraints, respectively.

Since the surrogate model is used as an approximation of the objective function and/or constraints, based on sampled points, it is prudent to improve the accuracy of the metamodel during the optimization process, by inserting infill points in regions of the design space that may contain the optimal point. Here, one surrogate model is created to represent the objective function and, if the problem is constrained, another surrogate model is created to represent all constraints. Three different types of searches are used to select the infill points: local search, global search and refinement. The main steps of the global optimization framework proposed here, which was developed in MATLAB [13], are the following:

1. Define a sampling plan: n_{samp} random points are generated uniformly in the design space; among these points, the one closest to the center of the search space is selected to compose the sampling plan; the next selected point is defined as the farthest apart point from those selected previously; the selection proceeds by the criterion of the maximum Euclidean distance between the points, until the n_{ED} points which define the initial sampling plan are obtained;
2. Evaluate the objective and constraint functions for the sampled points, obtaining $\mathbf{y}^f = (y_1^f, \dots, y_{n_{ED}}^f)$ and $\mathbf{y}^g = (y_1^g, \dots, y_{n_{ED}}^g)$, respectively, where $y_i^f = f(\mathbf{x}_i)$ and $y_i^g = \max(\mathbf{g}(\mathbf{x}_i))$, with $i = (1, \dots, n_{ED})$;

3. Fit one surrogate model for the objective function and another for the constraints, considering \mathbf{y}^f and \mathbf{y}^g ;
4. Insert infill points in the sampling plan: In this step, the search strategies are based on the Augmented Lagrangian function, given by $y_i^{augm} = f(\mathbf{x}_i)$ if \mathbf{x}_i is feasible and $y_i^{augm} = f(\mathbf{x}_i) + \rho(y_{max} + \max(\mathbf{g}(\mathbf{x}_i))^2)$ otherwise, where ρ is a penalty parameter and y_{max} is the highest value contained in the vector \mathbf{y}^f . This expression can be used to optimize both the original and the prediction functions, changing $f(\mathbf{x})$ and $g(\mathbf{x})$ to their predictions via surrogate model. The search strategies to select infill points are as follows:
 - a) *Local search*: This search is performed in order to look for local minima over the design space. For this, n_s design subspaces are defined and in each one of these regions, the optimization of the surrogate model is carried out by genetic algorithms [14]. The best design point found in each optimization subproblem is an infill point. To generate these design subspaces, first a point in the sampling plan is randomly selected. The next points are chosen based on the farthest apart subset concept, selecting the point from the sampling plan which is the farthest from the already selected points. The design subspaces are centered in these points and their boundaries are defined by dividing the design space as evenly as possible among them.
 - b) *Global search*: In this paper, two possible types of global search are employed, in an attempt to find the global minima. The first type, which can be applied to both RBF and Kriging, corresponds to the optimization of the surrogate model by genetic algorithms, considering the entire design space. The second type corresponds to one of the most used in Kriging, based on the expected improvement concept, which is related to the probability of finding a minimum lower than that found so far and to the variance of the surrogate model. Thus, one way to choose an infill point in each iteration of the optimization process is to select the one that maximizes the expected improvement function. To incorporate the constraints in this process, the expected improvement can be multiplied by the probability that a design is feasible, which results in the constrained expected improvement.
 - c) *Refinement*: In order to improve the best solution found so far, two refinement strategies are adopted, and each one occurs at a different stage of the optimization process. The first is applied when the previous searches did not find a new minimum in $n_{maxstall}$ iterations or in iterations multiples of 10, and corresponds to the optimization of the original function by the Interior-point method [15], starting from the design that has the lowest value of y_i^{augm} found so far. Whenever the first refinement strategy is not applied, the second one takes place. In this case, the metamodel is evaluated on a grid of n_{grid}^n evenly spaced points over the design space, starting with n_{grid} equal to 2, and the best point from the grid is taken as an infill candidate. The candidate is accepted if it leads to an Augmented Lagrangian value lesser than the average over the sampling plan. Otherwise, the grid is refined for the next iteration by taking $n_{grid} = n_{grid} + 1$. To avoid unnecessary evaluations of the objective and constraint functions, as well as ill-conditioning problems in the metamodels, the infill points are included in the sample only if they are distant enough from all the points of the sample. Also, the value of ρ is doubled every time an infill point results unfeasible. A limit of 1 is applied to rho, so that the constraints are not overly weighted.
5. Update \mathbf{y}^f and \mathbf{y}^g , and go to step 3. The procedure is performed until a stop criterion is reached. The stop criterion used herein is the number of objective function evaluations, limited to n_{OFE} . The optimal design vector is the feasible point of the sampling plan that has the lowest value of the objective function.

Figure 1 illustrates some steps of the framework, considering the Simionescu problem addressed in section 4. It is noticed that the surrogate model (Fig. 1b) results similar to the original surface of the problem (Fig. 1a), and that along the search strategies (Fig. 1c, 1d and 1e), most of the infill points are added in the regions of global minima.

4 Numerical examples

The methodology is applied in the solution of four unconstrained and four constrained problems. Three scenarios of the framework are investigated, where the first applies the RBF and the other two apply Kriging with global search by optimization (Kriging I) or by maximization of the expected improvement (Kriging II), respectively. The input data of the framework used in all examples, except when specified otherwise, is given by: $n_{samp} = 1 \cdot 10^6$; $n_{ED} = 10n$; $\rho = 1 \cdot 10^{-5}$; $n_s = 3$; $n_{OFE} = 100n$; $n_{maxstall} = 3$.

The problems are presented in Table 1. Global minimum, lower and upper bounds and other specific parameters of the functions are: Ackley, $f^*(0, 0) = 0$ and $-5 \leq x_1, x_2 \leq 10$; Cross-in-tray, $f^*(\pm 1.34941, \pm 1.34941) = f^*(\pm 1.34941, \mp 1.34941) = -2.06261$ and $-10 \leq x_1, x_2 \leq 10$; Rosenbrock, $f^*(1, \dots, 1) = 0$ and $-5 \leq x_i \leq 10$, and $n_{ED} = 50$ to $n = 10$; Simionescu, $f^*(\pm 0.84853, \mp 0.84853) = -0.072$ and $-1.25 \leq x_1, x_2 \leq 1.25$; Townsend, $f^*(2.00529, 1.19445) = -2.02399$, $-2.25 \leq x_1 \leq 2.5$ and $-2.5 \leq x_2 \leq 1.75$. The trusses are presented by Spillers and MacBain [16] and illustrated in Fig. 2, where the design variables are the cross-sectional areas of the bars and the objective function corresponds to the volume of the structures, subject to stress and displacement constraints. The first truss has $f^*(0.326, 0.326) = 44.7841 \text{ in}^3$ (734 cm^3), considering that mem-

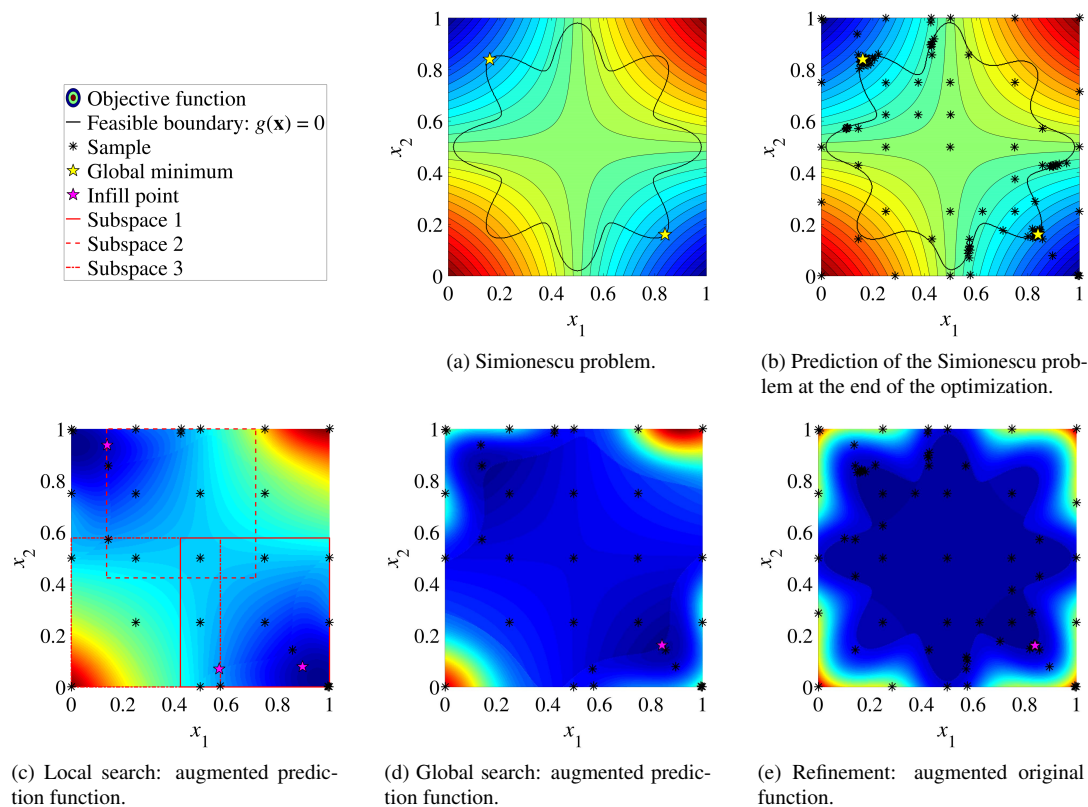


Figure 1. Global optimization framework based in Kriging: Simionescu problem.

bers 1-3 have the same area, modulus of elasticity $E = 29000$ ksi (200 GPA), $P = 1$ kip (4.45 kN) and $0.1 \leq x_1, x_2 \leq 5$ in² (0.65 - 32.26 cm²). The second truss presents $f^*(1.061, 0.707, 1, 0.707, 0.707, 0.5, 0.75) = 9.00$ for $0.45 \leq x_1, x_2 \leq 1.1$, with $P = 1$, $L = 2$, $n_s = 10$, displacement δ on node 4 limited to $\delta_{max} = 4$ and stresses σ in the bars limited to $\sigma_{max} = 1$.

In order to evaluate the performance of the three scenarios of the framework, each example is run 500 times, except for the seven-bar truss, which is run 100 times due to its larger computational demand. Comparisons are made in terms of usual metrics and through a probabilistic metric for comparing metaheuristic optimization algorithms, proposed by Gomes et al. [17], which is based on the idea of population interference, and yields the probability that a given algorithm produces a smaller minimum than an alternative algorithm, in a single run.

Table 1. Numerical examples.

Problem	Mathematical representation
Ackley	$f(x_1, x_2) = -20 \exp\left(-0.2 \sqrt{0.5(x_1^2 + x_2^2)}\right) - \exp(0.5(\cos(2\pi x_1) + \cos(2\pi x_2))) + 20 + e$
Cross-in-tray	$f(x_1, x_2) = -0.0001 \left(\left \sin(x_1) \sin(x_2) \exp\left(\left 100 - \frac{\sqrt{x_1^2 + x_2^2}}{\pi} \right \right) \right + 1 \right)^{0.1}$
Rosenbrock	$f(\mathbf{x}) = \sum_{i=1}^{n-1} \left(100(x_{i+1} - x_i^2)^2 + (1 - x_i)^2 \right)$, with $n = 2$ and $n = 10$
Simionescu	$f(x_1, x_2) = 0.1x_1x_2 \Rightarrow g(x_1, x_2) = x_1^2 + x_2^2 - \left(1 + 0.2 \cos\left(8 \arctan \frac{x_1}{x_2}\right)\right)^2 \leq 0$
Townsend	$f(x_1, x_2) = -(\cos((x_1 - 0.1)x_2))^2 - x_1 \sin(3x_1 + x_2) \Rightarrow g(x_1, x_2) = x_1^2 + x_2^2 - (2\cos(t) - \frac{1}{2}\cos(2t) - \frac{1}{4}\cos(3t) - \frac{1}{8}\cos(4t))^2 - (2\sin(t))^2 \leq 0$, with $t = \text{atan}(x_1, x_2)$
4-bar truss	$f(x_1, x_2) = 3c_1 + \sqrt{3}c_2 \Rightarrow g_1(x_1, x_2) = 6\left(\frac{3}{c_1} + \frac{\sqrt{3}}{c_2}\right) - 3 \leq 0$ $g_2(x_2) = 7.172 - c_2 \leq 0 \Rightarrow g_3(x_1) = 5.721 - c_1 \leq 0$, with $c_1 = \frac{x_1 E}{1000P}$ and $c_2 = \frac{x_2 E}{1000P}$
7-bar truss	$f(\mathbf{x}) = \sum_{i=1}^n x_i L_i \Rightarrow g_i(\mathbf{x}) = \frac{\sigma_i}{\sigma_{max}} - 1 \leq 0 \Rightarrow g_8(\mathbf{x}) = \frac{\delta}{\delta_{max}} - 1 \leq 0$, with $i = (1, \dots, n)$ and $n = 7$

Convergence plots for the usual metrics are shown in Fig. 3 and 4, for unconstrained and constrained prob-

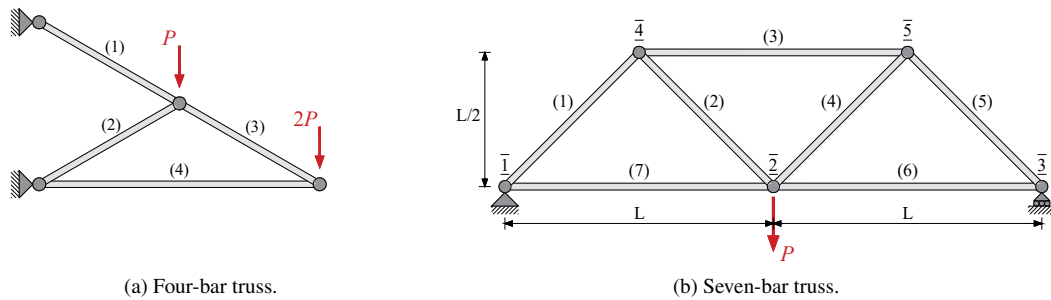


Figure 2. Benchmark trusses.

lems, respectively. It can be considered that the three cases of the framework found the global minimum in all runs of the Rosenbrock $_{n=2}$ and Simionescu functions. Regarding the other problems, the framework achieved a satisfactory mean result, almost always close to the best result, with little variation over the runs. Thus, it is understood that the framework, whose main characteristic is the application of the three search strategies, is stable and has little influence of randomness to achieve good results.

The RBF had the lowest mean and standard deviation for the Ackley function, followed by Kriging I, whereas, for the Cross-in-tray function, the best performances in relation to these parameters were obtained by Kriging II and RBF, consecutively. For the Rosenbrock $_{n=10}$ function, RBF and Kriging II showed similar performances, better than Kriging I. Also, the RBF had a lower mean and standard deviation for both trusses. However, for the other cases of the framework, it is not possible to affirm which is the best, using this metric.

The probabilistic metrics are shown in Table 2. For unconstrained problems, it is clear that the RBF presented the best performance, followed by Kriging II. However, for constrained problems, the RBF performed better than Kriging I and II for the truss problems, and the Kriging I performed better for the Townsend function.

Overall, it can be said that the framework based on the RBF better adapted to these benchmark problems. Global search via the expected improvement maximization (Kriging II) performs well for unconstrained problems. However, in some constrained problems, the constrained expected improvement function presents plateaus in feasible regions and others in infeasible regions, which makes this search procedure inefficient.

Table 2. Probability (in %), that the RBF algorithm produces smaller minima than the Kriging I (KG I) or Kriging II (KG II), and its complements.

	Ackley		Cross-in-tray		Rosenbrock $_{n=10}$		Townsend		4-bar truss		7-bar truss	
	KG I	KG II	KG I	KG II	KG I	KG II	KG I	KG II	KG I	KG II	KG I	KG II
RBF is better	60.20	58.56	51.06	23.68	56.50	52.13	39.91	62.62	93.83	78.22	60.97	72.49
Alg. are equivalent	0.07	0.07	27.37	35.78	2.53	2.27	6.35	8.40	0.00	0.00	0.00	0.00
Alter. is better	39.74	41.36	21.57	40.55	40.98	45.59	53.74	28.98	6.17	21.78	39.03	27.51

* Alg. and Alter. stand for “Algorithm” and “Alternative”, respectively.

5 Conclusions

In this paper, a comparison between Kriging and Radial Basis Function metamodels was presented, in the context of a proposed global optimization framework. In order to identify promising regions of the design space, the framework is composed of three types of search: local search, global search and refinement. Three cases of the framework were evaluated, where the first uses RBF, and the others use Kriging with two different global search schemes. Eight benchmark examples were evaluated. It was found that the proposed framework presents satisfactory results and little influence of randomness. Among the cases evaluated, the framework based on RBF was found to be more efficient and robust.

Acknowledgements. The authors gratefully acknowledge the financial support from Scientific and Technological Research Support Foundation of Santa Catarina State (FAPESC), Coordination of Superior Level Staff Improvement (CAPES) and National Council for Scientific and Technological Development (CNPq, via grant 302489/2017-7).

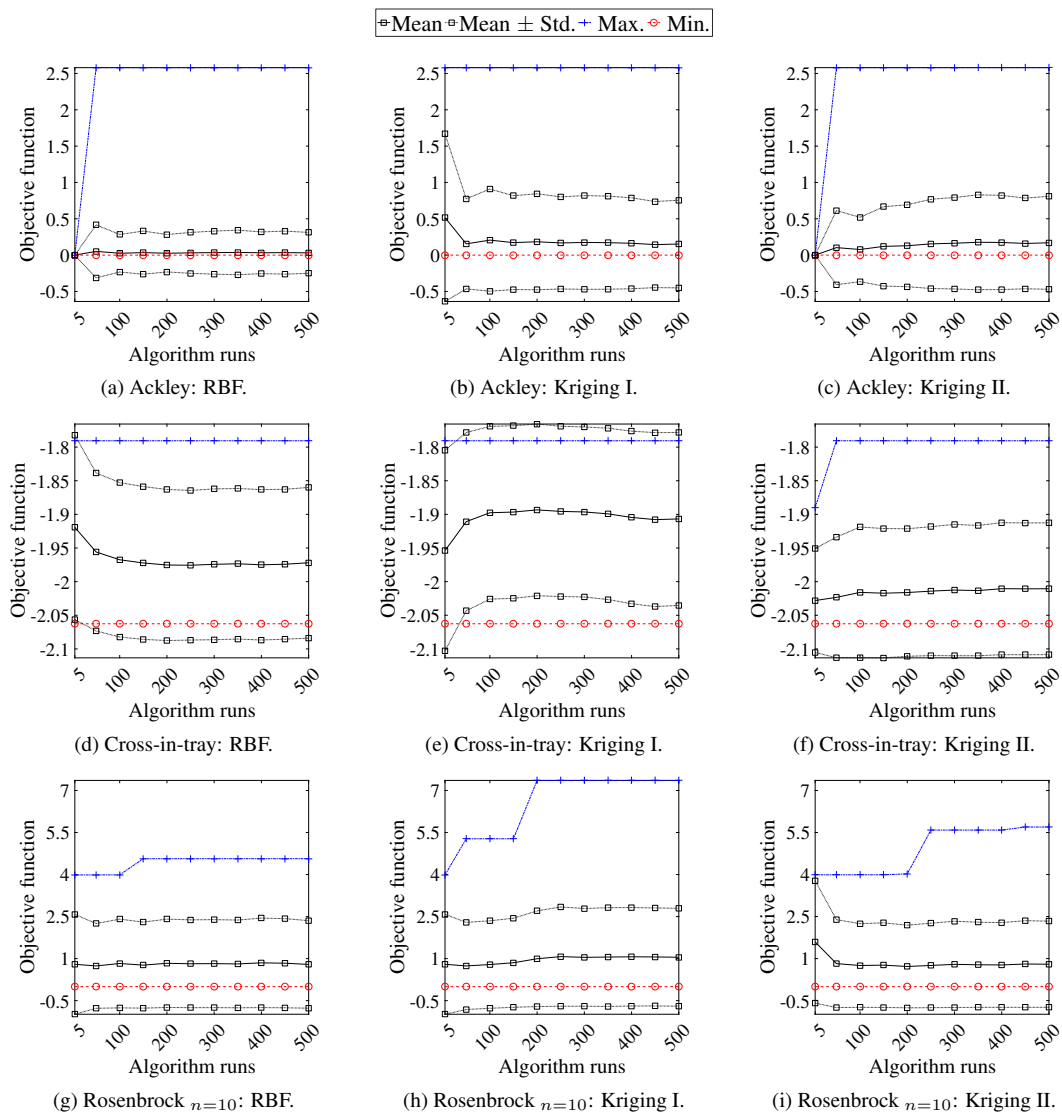


Figure 3. Convergence histories for usual metrics: Unconstrained problems.

Authorship statement. The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

References

- [1] Durantin, C., Marzat, J., & Balesdent, M., 2016. Analysis of multi-objective kriging-based methods for constrained global optimization. *Computational Optimization and Applications*, vol. 63, pp. 903–926.
- [2] Dong, H., Song, B., Dong, Z., & Wang, P., 2018. Scgosr: Surrogate-based constrained global optimization using space reduction. *Applied Soft Computing*, vol. 65, pp. 462–477.
- [3] Qian, J., Yi, J., Cheng, Y., Liu, J., & Zhou, Q., 2020. A sequential constraints updating approach for kriging surrogate model-assisted engineering optimization design problem. *Engineering with Computers*, vol. 36, pp. 993–1009.
- [4] Orr, M. J. L., 1996. *Introduction to Radial Basis Function Networks*. Centre for Cognitive Science: University of Edinburgh.
- [5] Buhmann, M. D., 2004. *Radial Basis Functions*. Cambridge University Press.
- [6] Krige, D. G., 1951. A statistical approach to some basic mine valuation problems on the Witwatersrand. *Journal of the Southern African Institute of Mining and Metallurgy*, vol. 52, n. 6, pp. 119–139.

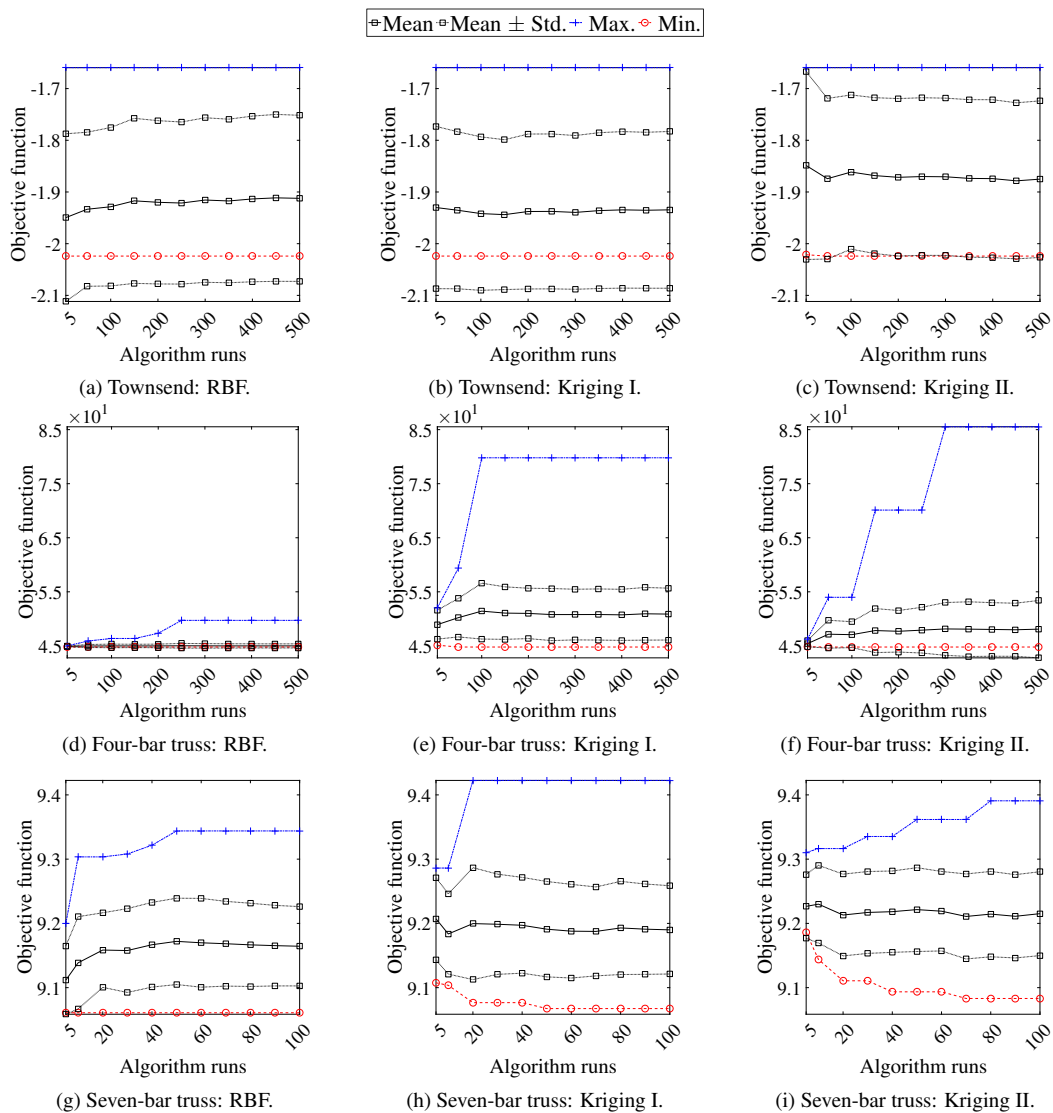


Figure 4. Convergence histories for usual metrics: Constrained problems.

- [7] Jones, D., Schonlau, M., & Welch, W. J., 1998. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, vol. 13, n. 4, pp. 455–492.
- [8] Forrester, A. I. J., Sobester, A., & Keane, A. J., 2008. *Engineering Design via Surrogate Modelling*. John Wiley and Sons.
- [9] Mishra, P. K., Nath, S. K., Sen, M. K., & Fasshauer, G. E., 2018. Hybrid Gaussian-cubic radial basis functions for scattered data interpolation. *Computational Geosciences* (, vol. 22, pp. 1203–1218.
- [10] Mishra, P. K., Fasshauer, G. E., Sen, M. K., & Ling, L., 2019. A stabilized radial basis-finite difference (RBF-FD) method with hybrid kernels. *Computers and Mathematics with Applications*, vol. 77, n. 9, pp. 2354–2368.
- [11] Rippa, S., 1999. An algorithm for selecting a good value for the parameter c in radial basis function interpolation. *Advances in Computational Mathematics*, vol. 11, pp. 193–210.
- [12] Kennedy, J. & Eberhart, R., 1995. Particle swarm optimization. *Proceedings of IEEE - International Conference on Neural Networks*, vol. , pp. 1942–1948.
- [13] MathWorks, 2017. *MATLAB: Primer*.
- [14] Goldberg, D. E., 1989. *Genetic algorithms in search, optimization & machine learning*. Addison-Wesley.
- [15] Coleman, T. F. & Li, Y., 1996. An interior, trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, vol. 6, pp. 418–445.
- [16] Spillers, W. R. & MacBain, K. M., 2009. *Structural Optimization*. Springer.
- [17] Gomes, W. J. S., Beck, A. T., Lopez, R. H., & Miguel, L. F. F., 2018. A probabilistic metric for comparing metaheuristic optimization algorithms. *Structural Safety*, vol. 70, pp. 59–70.