

## Research Article

# An Analysis of the RBF Hyperparameter Impact on Surrogate-Assisted Evolutionary Optimization

Yoel Tenne 

*Mechanical Engineering Department, Ariel University, Ariel, Israel*

Correspondence should be addressed to Yoel Tenne; [y.tenne@post.com](mailto:y.tenne@post.com)

Received 1 November 2021; Revised 15 May 2022; Accepted 5 October 2022; Published 20 December 2022

Academic Editor: Shah Nazir

Copyright © 2022 Yoel Tenne. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Computationally expensive optimization problems are often solved using surrogates and a common variant is the radial basis functions (RBF) model. It aggregates several basis functions which all depend on a hyperparameter affecting their individual outputs and consequentially the overall surrogate prediction. However, the optimal value of the hyperparameter is typically unknown and should therefore be calibrated. This raises the question how does the hyperparameter affect the overall optimization search effectiveness (and not just the stand-alone surrogate accuracy) and to what extent is such a calibration beneficial, which is an important consideration both for end-users and algorithm researchers alike. To rigorously address this issue this paper presents an analysis based on an extensive set of numerical experiments with an RBF surrogate-assisted evolutionary algorithm. It follows that the hyperparameter strongly affected performance and that the extent of its impact varied depending on the basis function, objective function modality, and problem dimension. Overall, calibration of the hyperparameter was typically highly beneficial to the search performance while dynamically optimizing the hyperparameter during the search yielded additional performance gains.

## 1. Introduction

Industrial optimization problems often involve functions whose evaluation is computationally expensive, for example, in engineering design optimization candidate products could be evaluated by a computer simulation (instead of real-world experiments) to improve the design process efficiency. This simulation-driven setup introduces several optimization challenges: (i) the simulation acts as a “black-box function”, namely, it assigns an output value to each input vector (candidate product) through involved numerical procedures or proprietary codes but there is no analytic expression for this functional relation thereby necessitating derivative-free optimization methods, (ii) each simulation run is typically computationally intensive so only a small number of vectors can be evaluated, and (iii) such “black-box functions” often have complicated features (such as multimodality) which exacerbate the search difficulty.

A framework which has shown to be effective in such problems combines an evolutionary algorithm (EA) with a

surrogate model (also termed in the literature as a “meta-model”) yielding the surrogate-assisted EA (SAEA). The evolutionary algorithm is a heuristic search algorithm which seeks an optimum by using operators inspired by the process of evolution and applies them to a population of candidate solutions, while the surrogate model approximates the true expensive black-box function and provides estimated function values at a much lower computational cost when compared to the true expensive function. A common surrogate variant is the radial basis functions (RBFs) model which uses a set of functions whose values are linearly aggregated. RBFs rely on a hyperparameter (HP) which affects their individual outputs and consequentially the surrogate prediction and the overall SAEA performance. The optimal HP value is typically unknown, and therefore it needs to be calibrated, which raises the following question: *how significant is the impact of the HP on the search and does the numerical effort spent on such calibration yield significant performance gains?* The impact of the HP has been examined in the literature with respect to the stand-alone surrogate

prediction accuracy but hardly so in the context of the overall surrogate-assisted search, which is a significantly more involved relation [1]. To address these open questions this study presents an extensive set of numerical experiments and analysis which examine how the HP affects the overall SAEA performance both when a fixed HP is selected a-priori and when it is dynamically calibrated during the search. Based on the large volume of results, the analysis clarifies in which instance does the HP strongly affect the SAEA performance. The remainder of this paper is organized as follows: Section 2 surveys the relevant literature, Section 3 explains the methods used, Section 4 presents the numerical experiments and their analysis, and lastly Section 5 concludes the paper.

## 2. Literature Survey

SAEAs have shown their effectiveness across a variety of problems ranging from drug formulation to aircraft design [2, 3]. The success of such algorithms stems from the combination of the EA, which is both derivative-free and capable of handling functions with complicated features, with the surrogate which provides approximate function values efficiently. In this setup the EA searches for an optimum of the surrogate while the latter is updated during the search. A basic SAEA is presented in Algorithm 1.

Various SAEA implementations have been explored in the literature, for example in [4], a multifidelity setup was proposed where several simulations of different accuracy were combined for design optimization of hydraulic turbines. In [5] the authors used an RBF-based SAEA for design optimization of a mechanical controller, while in [6] a multiobjective SAEA based on a Kriging surrogate was proposed. In [7] the authors used a SAEA to optimize the layout of a wind turbine farm with the goal of maximizing energy production while accounting for wind regimes and turbine wakes. More intricate frameworks include [8] which combined an EA with machine-learning techniques to handle vectors in which the simulation fails and [9, 10] in which ensembles of multiple surrogates were used concurrently to improve the prediction accuracy. Elaborate surrogate updating procedures were studied in [11, 12], while a dynamic in-search selection of the surrogate type was studied in [9, 13]. Recently in [14] the authors proposed a dual surrogate-assisted algorithm which uses a global surrogate and a group of local surrogates for approximation and a dual-population cooperative particle swarm algorithm to explore the surrogates to better handle highly multimodal functions. Other studies have proposed ensembles of surrogates [15] and more recently in [16] the authors proposed using an ensemble of local surrogates and selecting the subset that provides the most accurate approximation. Also related is [17] in which the authors proposed a surrogate-assisted particle swarm algorithm in which the sampled points are split into multiple subsets and a different surrogate is trained based on each sample. The surrogates are then managed to improve the global and local search capabilities of the algorithm. Also using multiple surrogates, in [18] the authors proposed a

surrogate-assisted algorithm which used both a Kriging surrogate for localized modelling and a global quadratic model to approximate the overall function trend. The algorithm also uses two optimization algorithms (“grey wolf” and a multistart search) to identify good solutions from each surrogate.

Various studies have also examined the impact of the HP on the stand-alone surrogate. In [19] a Kriging surrogate was used along with several HP calibration methods to achieve different balances between global and local prediction accuracies. In [20], the authors proposed estimating the optimal Kriging HP by using a multistart gradient search on the cross-validation error function, while in [21], calibrating the HP was found to affect the accuracy of various surrogates. In [22], the authors found that the size of the initial sample affected the calibration of the HP. However, the mentioned studies examined the impact of the HP only on the prediction accuracy of the stand-alone surrogate while this study analyzes the impact on the overall SAEA performance.

## 3. Methods

*3.1. The RBF Surrogate.* The RBF surrogate used in this study is a linear combination (superposition) of basis functions (BFs), namely,

$$m(\vec{x}) = \sum_{i=1}^n \alpha_i \phi_i(\vec{x}), \quad (1a)$$

$$\phi_i(\vec{x}) = \phi(\|\vec{x} - \vec{x}_i\|_2), \quad (1b)$$

where  $m(\vec{x})$  is the surrogate prediction, while  $\vec{x}_i$ ,  $i = 1, \dots, n$ , and  $\phi_i(\vec{x})$  are the sampled vectors and basis functions, respectively. The scalar weights  $\alpha_i$  are determined from the Lagrange interpolation conditions

$$m(\vec{x}_i) = f(\vec{x}_i), i = 1, \dots, n. \quad (2)$$

Two widely used BFs are the Gaussian and multiquadric (MQ) [23] which are defined as

$$\text{Gaussian: } \phi_i(\vec{x}) = \exp\left(-\sigma(\|\vec{x} - \vec{x}_i\|_2)^2\right), \quad (3)$$

$$\text{MQ: } \phi_i(\vec{x}) = \sqrt{1 + \sigma(\|\vec{x} - \vec{x}_i\|_2)^2}, \quad (4)$$

where  $\sigma$  is the HP which determines the BF shape and consequentially the overall surrogate prediction  $m(\vec{x})$ . This relation is visualized in Figure 1.

Given its impact, various studies have examined how to calibrate the HP based on the stand-alone surrogate accuracy. In [24], the authors used an empirical fixed value derived from the mean distance between the sampled vectors but without accounting for the observed function values. In [25], the HP was calibrated by minimizing the root mean square error (RMSE) of the RBF prediction by using a gradient descent. A related leave-one-out cross-validation procedure was proposed in [26–29] with various

Input: initial sample size, max evaluations, EA parameters;  
 Generate an initial set of vectors and evaluate with the true function;  
 Repeat  
     Train a surrogate model by using the vectors evaluated so far;  
     Search for an optimum of the surrogate by using an EA;  
     Evaluate the predicted solution (and possibly additional vectors) with the true function;  
 until *max evaluations*;  
 Output: the best solution found;

ALGORITHM 1: A basic SAEA.

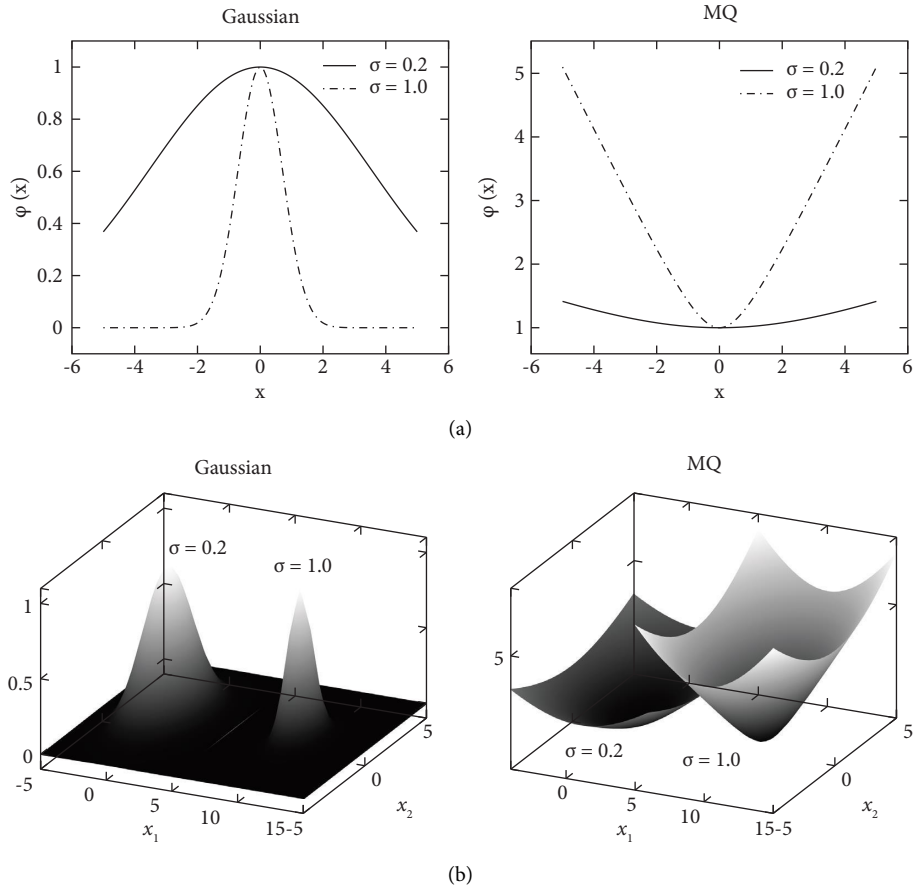


FIGURE 1: The impact of the hyperparameter on (a) univariate and (b) bivariate Gaussian and MQ basis functions.

modifications explored in [30–33]. In [34] the HP was calibrated based on a least-squares fit procedure, while in [35], MARS interpolating polynomials were used. More recently, [36] proposed using an EA for the HP calibration. In the numerical experiments performed in this study, both a fixed and a dynamically optimized HP were used as explained in Section 4.

**3.2. The SAEA Algorithm.** A representative SAEA was implemented for the numerical experiments. It begins by sampling an initial random population and training the RBF surrogate. The main loop then begins in which a real-coded EA seeks an optimum of the surrogate. Every  $g$

generations the best  $q$  percent of the population members are evaluated with the true objective function and are added to the cache of sampled vectors. The surrogate is retrained based on the updated cache and the EA population is re-evaluated with the new surrogate. Following [37] the settings  $g = 4$  and  $q = 20\%$  were used. The loop repeats until either the true (expensive) function evaluations reach  $50d$  (where  $d$  is the function dimension), or 1000 EA generations have occurred. Algorithm 2 provides the complete SAEA workflow.

For the real-coded EA a representative variant with a population size of 20 was used which operates as follows: (i) vectors are ranked by fitness, (ii) parents are selected by stochastic universal selection (SUS) and recombined with a

```

Input: initial sample size, max evaluations, EA parameters,  $g$  and  $s$  parameters;
generate and sample an initial random population;
repeat
  train an RBF surrogate based on the evaluated vectors;
  run one generation of the EA;
  for each  $g^{\text{th}}$  generation do
    evaluate with the best  $q\%$  vectors with the true function;
    retrain the surrogate and re-evaluate the population;
  end
until max function evaluations or max EA generations;
output: the best solution found;

```

ALGORITHM 2: The implemented RBF-based SAEA.

uniform crossover operator with a probability of 0.7 to produce offspring, and (iii) the latter are mutated with the breeder genetic algorithm operator with a probability of 0.1.

## 4. Experiments

**4.1. Fixed HP.** A first set of numerical experiments was performed to analyze if and to what extent different fixed HP values, namely, which are chosen a-priori, affect the overall SAEA performance. Following the Design-of-Experiments (DOE) framework [38] the numerical experiments were formulated by using a single *variable*, namely the HP, which was assigned different fixed values, and 3 *factors* (function type, function dimension, and BF type). Accordingly, in each combination of factors the numerical experiments were repeated with the different HP values so that the HP impact could be analyzed across a wide range of problem settings. The DOE setup is summarized in Table 1.

The established set of test functions of [39] was used as it includes functions with diverse features as shown in Figure 2.

In each test case (namely, a combination of factors), the SAEA was ran with four different HP settings and with 30 runs per setting to obtain an adequate sample, thereby yielding a total of 2400 runs (5 functions  $\times$  2 dimensions  $\times$  2 BF  $\times$  4 HP settings  $\times$  30 repeats).

Tables 2–5 present the resultant statistics (mean, median, standard deviation, min, and max) of the final function values for each batch of 30 runs across the different HP values and test cases. Also provided is the  $p$ -value of the Mann–Whitney statistical significance test. It was obtained by comparing the results corresponding to the HP associated with the best mean to the results associated with the 3 remaining HP settings. Statistical significance was measured at the 0.05 level, namely, for  $p \leq 0.05$ . As such in each of the 20 combinations of factors (function-dimension-BF), the number of significant performance differences observed (denoted  $s$ ) varied between 0 (no significant differences observed) which implies that the HP had no significant impact to 3 (all comparisons had significant differences) which implies that the HP had a strong impact.

The corresponding results are summarized in Table 6 from which it follows that in 1 combination of factors the

HP had no significant impact, in 3 cases it had only a minor impact, and in the remaining 16 cases it had a significant impact. The cases showing a lessened HP impact are low-dimensional and involve the Gaussian basis function. These results are attributed to the relative simplicity of the lower-dimensional cases which allowed for a good approximation across a range of HP values. Also, the output of a Gaussian BF surrogate approaches the sample mean faster than a MQ surrogate due to the Gaussian BF behaviour and therefore the RBF output was similar for different HP values which results in fewer performance differences, as observed. In contrast, with the MQ basis function and with the high-dimensional cases the HP consistently had a significant impact. This is attributed to the MQ basis function, whose output varies for a wider range of HP values. Also, in high-dimensional cases, generating an accurate surrogate is more challenging hence, different HP values resulted in significantly different prediction accuracies. These aspects are further addressed in length in the analysis which follows.

Expanding on the preceding analysis an additional fine-grained analysis was also performed to highlight how each factor (BF, modality, dimension) affected performance and the results are summarized in Table 7. It follows that the per-factor influence is as follows:

**4.1.1. Basis Function.** For a Gaussian BF in 4 cases varying the HP had a negligible impact while the MQ BF variations were always significant, namely, the sensitivity to the HP was lower for a Gaussian BF than it was for a MQ one. This is attributed to the mathematical formulation of each BF, namely, for a Gaussian BF

$$\varphi(\vec{x}) = \exp\left(-\sigma\|\vec{x} - \vec{x}_i\|_2^2\right). \quad (5)$$

the expression decays to 0 rapidly for a wide range of HP values which results in similar surrogate predictions for different HP values and consequentially a lower HP impact. In contrast, with a MQ BF,

$$\varphi(\vec{x}) = \sqrt{1 + \sigma\|\vec{x} - \vec{x}_i\|_2^2}. \quad (6)$$

TABLE 1: Design of the numerical experiments.

Type	Name	Values
Factor	Function Dimension BF	Ackley, Griewank, Powell, Rastrigin, Rosenbrock 5, 10 Gaussian, MQ
Variable	Hyperparameter	0.01, 0.1, 1, 10

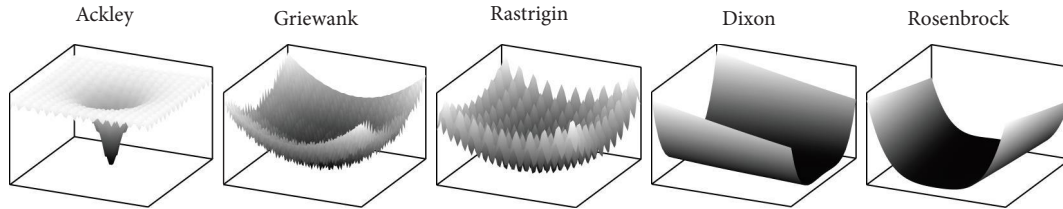


FIGURE 2: The test functions used.

TABLE 2: Statistics 5D, Gaussian.

Function	Statistic	$\kappa = 0.01$	$\kappa = 0.1$	$\kappa = 1$	$\kappa = 10$
Ackley-05	Mean	<b>1.610e + 01</b>	1.794e + 01	1.811e + 01	1.726e + 01
	SD	4.476e + 00	2.037e + 00	1.810e + 00	2.225e + 00
	Median	1.770e + 01	1.848e + 01	1.872e + 01	1.780e + 01
	Min	6.060e + 00	1.199e + 01	1.276e + 01	1.231e + 01
	Max	2.033e + 01	2.008e + 01	1.996e + 01	2.009e + 01
	<i>p</i> -value			9.412e - 02	7.791e - 02
Dixon-05	Mean	1.958e + 02	<b>5.127e + 01</b>	6.533e + 03	3.398e + 03
	SD	2.114e + 02	4.044e + 01	6.211e + 03	3.428e + 03
	Median	1.160e + 02	3.938e + 01	5.316e + 03	2.334e + 03
	Min	1.566e + 00	2.819e + 00	1.831e + 02	6.586e + 00
	Max	7.909e + 02	1.554e + 02	2.792e + 04	1.292e + 04
	<i>p</i> -value	5.428e - 04		1.436e - 11	2.201e - 10
Griewank-05	Mean	4.513e + 01	<b>2.996e + 01</b>	3.317e + 01	3.225e + 01
	SD	2.180e + 01	1.440e + 01	1.911e + 01	1.855e + 01
	Median	4.488e + 01	2.803e + 01	3.068e + 01	2.675e + 01
	Min	5.703e + 00	6.611e + 00	3.571e + 00	8.263e + 00
	Max	8.845e + 01	5.749e + 01	7.990e + 01	6.709e + 01
	<i>p</i> -value	1.411e - 03		2.722e - 01	4.529e - 01
Rastrigin-05	Mean	4.009e + 01	<b>3.891e + 01</b>	4.491e + 01	4.131e + 01
	SD	9.180e + 00	1.245e + 01	1.033e + 01	9.774e + 00
	Median	4.004e + 01	3.789e + 01	4.705e + 01	4.243e + 01
	Min	2.457e + 01	1.497e + 01	2.215e + 01	1.978e + 01
	Max	5.783e + 01	6.491e + 01	6.130e + 01	5.762e + 01
	<i>p</i> -value	4.296e - 01		2.639e - 02	2.254e - 01
Rosenbrock-05	Mean	2.000e + 01	<b>9.075e + 00</b>	6.598e + 01	1.609e + 02
	SD	2.152e + 01	1.520e + 01	9.548e + 01	1.169e + 02
	Median	1.455e + 01	4.323e + 00	1.706e + 01	1.353e + 02
	Min	1.188e + 00	8.286e - 01	3.115e + 00	3.038e + 01
	Max	1.023e + 02	6.534e + 01	3.413e + 02	5.017e + 02
	<i>p</i> -value	2.247e - 05		9.656e - 06	2.886e - 11

the output varies over a wider range of HP values which in turn affects the surrogate prediction and results in a more significant HP impact. As an example Figure 3(a) shows a plot of the RMSE of the surrogate prediction when using Gaussian and MQ BFs with the Ackley-5D function. The error quickly rises for the Gaussian BF with HP values since the individual BFs decay to zero and the surrogate prediction approaches the sample mean. In contrast, there is a much

higher sensitivity to the HP values with the MQ BFs as observed in the experiments.

4.1.2. *Modality.* The HP impact was significant for the low modality functions (Dixon, Rosenbrock, 2-3 significant differences per test case) but its impact was diminished for the high modality functions (Ackley, Griewank, and

TABLE 3: Statistics 10D, Gaussian.

Function	Statistic	$\kappa = 0.01$	$\kappa = 0.1$	$\kappa = 1$	$\kappa = 10$
Ackley-10	Mean	2.018e+01	2.014e+01	<b>1.971e+01</b>	1.977e+01
	SD	4.794e-01	3.787e-01	5.736e-01	5.120e-01
	Median	2.026e+01	2.025e+01	1.975e+01	1.995e+01
	Min	1.851e+01	1.911e+01	1.767e+01	1.862e+01
	Max	2.070e+01	2.080e+01	2.054e+01	2.047e+01
	<i>p</i> -value	1.229e-04	3.187e-04		2.436e-01
Dixon-10	Mean	<b>3.845e+03</b>	1.043e+05	1.234e+05	6.127e+04
	SD	5.651e+03	5.375e+04	5.635e+04	3.290e+04
	Median	2.158e+03	9.943e+04	1.128e+05	5.259e+04
	Min	4.009e+02	2.091e+03	4.255e+04	6.235e+03
	Max	2.456e+04	2.012e+05	2.567e+05	1.335e+05
	<i>p</i> -value		1.027e-10	1.436e-11	2.614e-11
Griewank-10	Mean	1.546e+02	1.171e+02	<b>1.129e+02</b>	1.170e+02
	SD	3.277e+01	2.659e+01	3.387e+01	3.271e+01
	Median	1.510e+02	1.137e+02	1.129e+02	1.159e+02
	Min	1.012e+02	7.288e+01	4.410e+01	3.683e+01
	Max	2.346e+02	1.883e+02	1.624e+02	1.920e+02
	<i>p</i> -value	2.551e-05	3.503e-01		2.973e-01
Rastrigin-10	Mean	<b>1.053e+02</b>	1.187e+02	1.220e+02	1.109e+02
	SD	1.923e+01	2.090e+01	1.541e+01	1.389e+01
	Median	1.097e+02	1.154e+02	1.229e+02	1.102e+02
	Min	6.244e+01	7.789e+01	8.795e+01	8.355e+01
	Max	1.325e+02	1.546e+02	1.598e+02	1.370e+02
	<i>p</i> -value		2.066e-02	5.718e-04	2.166e-01
Rosenbrock-10	Mean	1.907e+02	<b>8.191e+01</b>	1.215e+03	1.092e+03
	SD	1.141e+02	6.346e+01	4.582e+02	4.177e+02
	Median	1.513e+02	6.880e+01	1.234e+03	9.760e+02
	Min	5.047e+01	9.250e+00	3.200e+02	5.115e+02
	Max	5.253e+02	3.307e+02	2.156e+03	2.018e+03
	<i>p</i> -value	4.590e-07		1.588e-11	1.436e-11

TABLE 4: Statistics 5D, multiquadric.

Function	Statistic	$\kappa = 0.01$	$\kappa = 0.1$	$\kappa = 1$	$\kappa = 10$
Ackley-05	Mean	1.489e+01	1.083e+01	<b>1.032e+01</b>	1.134e+01
	SD	3.841e+00	4.928e+00	2.644e+00	2.648e+00
	Median	1.589e+01	9.730e+00	1.041e+01	1.203e+01
	Min	5.115e+00	3.771e+00	5.193e+00	5.197e+00
	Max	1.971e+01	1.851e+01	1.447e+01	1.477e+01
	<i>p</i> -value	2.829e-06	4.238e-01		4.888e-02
Dixon-05	Mean	1.228e+02	<b>4.939e+00</b>	8.916e+01	3.662e+02
	SD	1.155e+02	7.272e+00	1.376e+02	6.810e+02
	Median	8.671e+01	2.390e+00	5.086e+01	1.050e+02
	Min	7.622e+00	4.486e-01	1.659e+00	1.686e+00
	Max	5.362e+02	3.944e+01	7.063e+02	3.443e+03
	<i>p</i> -value	5.208e-11		2.883e-08	5.554e-08
Griewank-05	Mean	<b>3.526e+00</b>	9.077e+00	1.900e+01	1.316e+01
	SD	2.545e+00	8.885e+00	1.387e+01	9.310e+00
	Median	2.610e+00	6.687e+00	1.620e+01	1.282e+01
	Min	6.409e-01	2.856e-01	1.048e+00	2.213e+00
	Max	1.007e+01	4.335e+01	5.363e+01	5.079e+01
	<i>p</i> -value		6.053e-05	1.748e-08	2.654e-08
Rastrigin-05	Mean	3.636e+01	3.500e+01	<b>1.314e+01</b>	1.367e+01
	SD	1.156e+01	1.030e+01	8.408e+00	8.820e+00
	Median	3.753e+01	3.636e+01	1.109e+01	1.101e+01
	Min	1.281e+01	8.895e+00	2.029e+00	2.426e+00
	Max	5.083e+01	5.855e+01	3.887e+01	3.784e+01
	<i>p</i> -value	1.528e-09	5.277e-09		4.238e-01

TABLE 4: Continued.

Function	Statistic	$\kappa = 0.01$	$\kappa = 0.1$	$\kappa = 1$	$\kappa = 10$
Rosenbrock-05	Mean	$2.763e + 01$	$6.443e + 00$	<b><math>3.542e + 00</math></b>	$1.604e + 01$
	SD	$2.164e + 01$	$8.784e + 00$	$1.333e + 00$	$2.136e + 01$
	Median	$1.593e + 01$	$3.464e + 00$	$3.613e + 00$	$6.037e + 00$
	Min	$3.526e + 00$	$6.503e - 02$	$7.370e - 01$	$1.503e - 01$
	Max	$8.159e + 01$	$4.343e + 01$	$6.196e + 00$	$9.069e + 01$
	<i>p</i> -value	$6.146e - 10$	$4.122e - 01$		$1.257e - 05$

TABLE 5: Statistics 10D, multiquadric.

Function	Statistic	$\kappa = 0.01$	$\kappa = 0.1$	$\kappa = 1$	$\kappa = 10$
Ackley-10	Mean	$1.648e + 01$	<b><math>1.224e + 01</math></b>	$1.470e + 01$	$1.480e + 01$
	SD	$3.290e + 00$	$3.485e + 00$	$1.406e + 00$	$1.286e + 00$
	Median	$1.751e + 01$	$1.268e + 01$	$1.464e + 01$	$1.487e + 01$
	Min	$8.904e + 00$	$6.471e + 00$	$1.158e + 01$	$1.213e + 01$
	Max	$1.993e + 01$	$1.722e + 01$	$1.766e + 01$	$1.691e + 01$
	<i>p</i> -value	$5.271e - 06$		$1.014e - 02$	$5.497e - 03$
Dixon-10	Mean	$2.079e + 03$	<b><math>1.294e + 03</math></b>	$1.114e + 04$	$2.838e + 04$
	SD	$2.267e + 03$	$1.340e + 03$	$6.722e + 03$	$2.301e + 04$
	Median	$1.320e + 03$	$9.800e + 02$	$9.620e + 03$	$2.347e + 04$
	Min	$1.576e + 02$	$2.490e + 02$	$1.263e + 03$	$6.460e + 02$
	Max	$1.097e + 04$	$6.601e + 03$	$2.647e + 04$	$9.430e + 04$
	<i>p</i> -value	$3.450e - 02$		$1.244e - 10$	$3.866e - 10$
Griewank-10	Mean	<b><math>1.929e + 01</math></b>	$7.998e + 01$	$1.000e + 02$	$1.073e + 02$
	SD	$9.413e + 00$	$3.160e + 01$	$3.888e + 01$	$2.896e + 01$
	Median	$1.962e + 01$	$7.887e + 01$	$9.844e + 01$	$1.087e + 02$
	Min	$4.407e + 00$	$2.108e + 01$	$2.651e + 01$	$5.450e + 01$
	Max	$4.322e + 01$	$1.444e + 02$	$1.720e + 02$	$1.720e + 02$
	<i>p</i> -value		$1.822e - 10$	$6.329e - 11$	$1.436e - 11$
Rastrigin-10	Mean	$1.026e + 02$	$9.370e + 01$	<b><math>4.610e + 01</math></b>	$5.373e + 01$
	SD	$1.695e + 01$	$1.896e + 01$	$2.132e + 01$	$1.790e + 01$
	Median	$1.036e + 02$	$9.604e + 01$	$4.194e + 01$	$5.095e + 01$
	Min	$5.341e + 01$	$4.806e + 01$	$1.508e + 01$	$2.598e + 01$
	Max	$1.338e + 02$	$1.277e + 02$	$8.833e + 01$	$1.011e + 02$
	<i>p</i> -value	$7.685e - 11$	$2.000e - 09$		$3.926e - 02$
Rosenbrock-10	Mean	$1.150e + 02$	<b><math>4.843e + 01</math></b>	$6.869e + 01$	$1.815e + 02$
	SD	$6.459e + 01$	$3.006e + 01$	$3.263e + 01$	$1.160e + 02$
	Median	$9.499e + 01$	$3.763e + 01$	$6.322e + 01$	$1.751e + 02$
	Min	$2.865e + 01$	$8.866e + 00$	$2.229e + 01$	$4.291e + 01$
	Max	$2.552e + 02$	$1.298e + 02$	$1.450e + 02$	$6.382e + 02$
	<i>p</i> -value	$1.492e - 06$		$2.849e - 03$	$1.528e - 09$

TABLE 6: Test cases grouped by significant differences (*s*).

<i>s</i>	Case	Total
0	G-Ack-05	1
1	G-Gri-05, G-Ras-05, G-Gri-10	3
2	G-Ack-10, G-Ras-10, M-Ack-05, M-Ras-05, M-Ros-05	5
3	G-Dix-05, G-Ros-05, G-Dix-10, M-Dix-05, M-Gri-05, M-Ack-10, M-Dix-10, M-Gri-10, M-Ras-10, M-Ros-10	11

Rastrigin). This is attributed to the fact that as the modality was increased the surrogate prediction accuracy degraded due to the more complicated function features up to a point where the prediction accuracy was negligibly affected by different HP values. To demonstrate this, numerical experiments were performed with a modified Rastrigin function

$$f(\vec{x}) = \sum_{i=1}^d \{(x_i)^2 - 10 \cos(k\pi x_i)\}, \quad (7)$$

where *k* is a prescribed parameter which determines the function modality (higher values increase modality). The surrogate prediction accuracy was examined for a 5D function

TABLE 7: Significant performance differences by factor.

(a) Significant differences (s) by BF		
BF	s	Test case
Gaussian	0	Ack-05
	1	Gri-05, Ras-05, Gri-10
	2	Ack-10, Ras-10
	3	Dix-05, Ros-05, Dix-10, Ros-10
MQ	0	
	1	
	2	Ack-05, Ras-05, Ros-05
	3	Dix-05, Gri-05, Ack-10, Dix-10, Gri-10, Ras-10, Ros-10
(b) Significant differences (s) by modality		
Modality	s	Test case
High	0	G-Ack-05
	1	G-Gri-05, G-Ras-05, G-Gri-10
	2	G-Ack-10, G-Ras-10, M-Ack-05, M-Ras-05
	3	M-Gri-05, M-Ack-10, M-Gri-10, M-Ras-10
Low	0	
	1	
	2	M-Ros-05
	3	G-Dix-05, G-Ros-05, G-Dix-10, G-Ros-10, M-Dix-05, M-Dix-10, M-Ros-10
(c) Significant differences by function dimension (d)		
Dimension	s	Test cases
5	0	G-Ack-05
	1	G-Gri-05, G-Ras-05
	2	M-Ack-05, M-Ras-05, M-Ros-05
	3	G-Dix-05, G-Ros-05, M-Dix-05, M-Gri-05
10	0	
	1	G-Gri-10
	2	G-Ack-10, G-Ras-10
	3	G-Dix-10, G-Ros-10, M-Ack-10, M-Dix-10, M-Gri-10, M-Ras-10, M-Ros-10

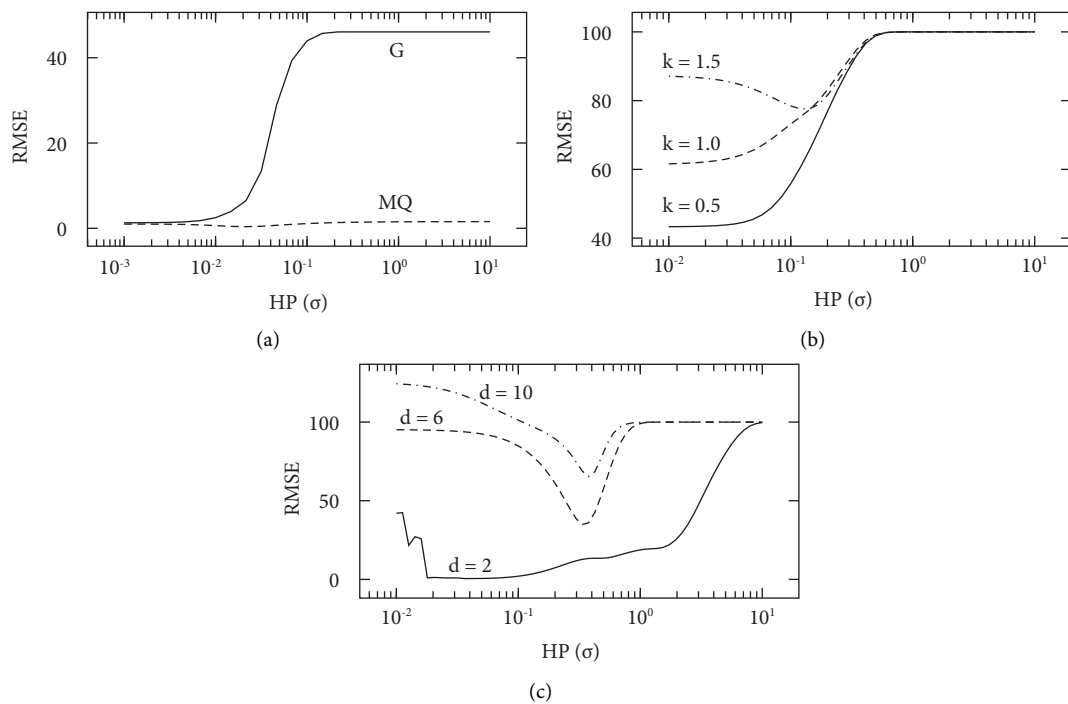


FIGURE 3: Variation of the RMSE by factors. (a) Prediction RMSE by BF (Ackley-5). (b) Prediction RMSE by modality (modified Rastrigin function,  $k$ ). (c) Prediction RMSE by dimension (Rosenbrock,  $d$ ).



TABLE 8: Test cases grouped by significant differences ( $s$ ) and evaluations limits.

Evaluations	$s$	Case	Total
50 <i>d</i>	0	G-Ack-05	1
	1	G-Gri-05, G-Ras-05, G-Gri-10	3
	2	G-Ack-10, G-Ras-10, M-Ack-05, M-Ras-05, M-Ros-05	5
	3	G-Dix-05, G-Ros-05, G-Dix-10, M-Dix-05, M-Gri-05, M-Ack-10, M-Dix-10, M-Gri-10, M-Ras-10, M-Ros-10	11
100 <i>d</i>	0	Ack-05	1
	1	G-Gri-05, G-Gri-10, M-Ack-05, M-Ack-10	4
	2	G-Ras-10, M-Ras-05, M-Ros-05, M-Ros-10	4
	3	G-Dix-05, G-Ras-05, G-Ros-05, G-Ack-10, G-Dix-10, G-Ros-10, M-Dix-05, M-Gri-05, M-Dix-10, M-Gri-10, M-Ras-10	11
200 <i>d</i>	0	0	0
	1	G-Gri-05, G-Gri-10	2
	2	G-Ack-05, G-Ras-10, M-Ack-05, M-Ros-05, M-Ack-10	5
	3	G-Dix-05, G-Ras-05, G-Ros-05, G-Ack-10, G-Dix-10, G-Ros-10, M-Dix-05, M-Gri-05, M-Ras-05, M-Dix-10, M-Gri-10, M-Ras-10, M-Ros-10	13

TABLE 9: Statistics, Gaussian.

Function	Statistic	5D		10D	
		Optimized	Best fixed	Optimized	Best fixed
Ackley	Mean	1.452 <i>e</i> + 01	1.610 <i>e</i> + 01	1.400 <i>e</i> + 01	1.971 <i>e</i> + 01
	SD	4.558 <i>e</i> + 00	4.476 <i>e</i> + 00	4.551 <i>e</i> + 00	5.736 <i>e</i> − 01
	Median	1.497 <i>e</i> + 01	1.770 <i>e</i> + 01	1.435 <i>e</i> + 01	1.975 <i>e</i> + 01
	Min	5.417 <i>e</i> + 00	6.060 <i>e</i> + 00	3.632 <i>e</i> + 00	1.767 <i>e</i> + 01
	Max	2.049 <i>e</i> + 01	2.033 <i>e</i> + 01	2.046 <i>e</i> + 01	2.054 <i>e</i> + 01
	$p$ -value		1.099 <i>e</i> − 01		1.693 <i>e</i> − 07
Dixon	Mean	1.445 <i>e</i> + 01	5.127 <i>e</i> + 01	1.422 <i>e</i> + 01	3.845 <i>e</i> + 03
	SD	4.593 <i>e</i> + 00	4.044 <i>e</i> + 01	4.157 <i>e</i> + 00	5.651 <i>e</i> + 03
	Median	1.434 <i>e</i> + 01	3.938 <i>e</i> + 01	1.498 <i>e</i> + 01	2.158 <i>e</i> + 03
	Min	7.511 <i>e</i> + 00	2.819 <i>e</i> + 00	6.394 <i>e</i> + 00	4.009 <i>e</i> + 02
	Max	2.032 <i>e</i> + 01	1.554 <i>e</i> + 02	2.021 <i>e</i> + 01	2.456 <i>e</i> + 04
	$p$ -value		3.950 <i>e</i> − 05		1.436 <i>e</i> − 11
Griewank	Mean	1.469 <i>e</i> + 01	2.996 <i>e</i> + 01	1.565 <i>e</i> + 01	1.129 <i>e</i> + 02
	SD	4.021 <i>e</i> + 00	1.440 <i>e</i> + 01	4.288 <i>e</i> + 00	3.387 <i>e</i> + 01
	Median	1.470 <i>e</i> + 01	2.803 <i>e</i> + 01	1.696 <i>e</i> + 01	1.129 <i>e</i> + 02
	Min	6.577 <i>e</i> + 00	6.611 <i>e</i> + 00	6.330 <i>e</i> + 00	4.410 <i>e</i> + 01
	Max	2.049 <i>e</i> + 01	5.749 <i>e</i> + 01	2.046 <i>e</i> + 01	1.624 <i>e</i> + 02
	$p$ -value		2.133 <i>e</i> − 06		1.436 <i>e</i> − 11
Rastrigin	Mean	1.618 <i>e</i> + 01	3.891 <i>e</i> + 01	1.653 <i>e</i> + 01	1.053 <i>e</i> + 02
	SD	3.720 <i>e</i> + 00	1.245 <i>e</i> + 01	4.337 <i>e</i> + 00	1.923 <i>e</i> + 01
	Median	1.738 <i>e</i> + 01	3.789 <i>e</i> + 01	1.862 <i>e</i> + 01	1.097 <i>e</i> + 02
	Min	8.667 <i>e</i> + 00	1.497 <i>e</i> + 01	6.259 <i>e</i> + 00	6.244 <i>e</i> + 01
	Max	2.050 <i>e</i> + 01	6.491 <i>e</i> + 01	2.057 <i>e</i> + 01	1.325 <i>e</i> + 02
	$p$ -value		1.165 <i>e</i> − 09		1.436 <i>e</i> − 11
Rosenbrock	Mean	1.583 <i>e</i> + 01	9.075 <i>e</i> + 00	1.407 <i>e</i> + 01	8.191 <i>e</i> + 01
	SD	4.182 <i>e</i> + 00	1.520 <i>e</i> + 01	4.255 <i>e</i> + 00	6.346 <i>e</i> + 01
	Median	1.662 <i>e</i> + 01	4.323 <i>e</i> + 00	1.270 <i>e</i> + 01	6.880 <i>e</i> + 01
	Min	8.272 <i>e</i> + 00	8.286 <i>e</i> − 01	5.874 <i>e</i> + 00	9.250 <i>e</i> + 00
	Max	2.044 <i>e</i> + 01	6.534 <i>e</i> + 01	2.040 <i>e</i> + 01	3.307 <i>e</i> + 02
	$p$ -value		1.000 <i>e</i> + 00		7.386 <i>e</i> − 10

and results are presented in Figure 3(b). It follows that the surrogate prediction RMSE increased with modality for all HP values as observed in the numerical experiments. Consequently the impact of the HP was more pronounced for the lower modality settings and diminished as the modality was increased, as observed in the numerical experiments.

**4.1.3. Dimension.** In the lower dimensional cases ( $d = 5$ ), the HP had a significant impact in 7 cases, while in the higher dimensional cases ( $d = 10$ ) in 9 cases. These results are attributed to the “curse of dimensionality,” namely in the lower dimensional problems a more accurate surrogate could be trained since the problems are simpler and hence

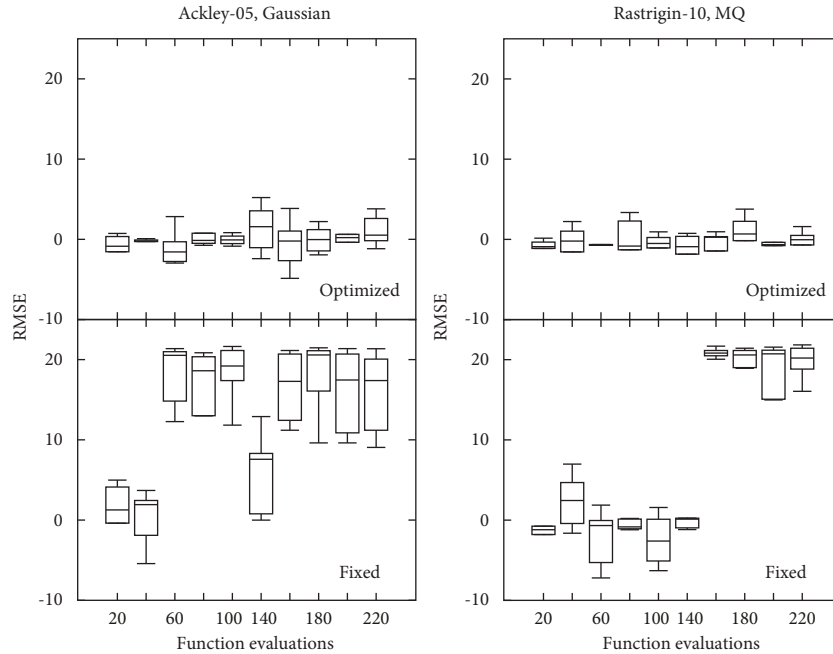


FIGURE 4: Prediction error variation for a fixed vs. optimized HP.

the sensitivity to the HP was lower. In contrast, in the higher dimensional cases training an accurate surrogate was more challenging and required more precise HP values hence the higher sensitivity to the HP. To demonstrate this Figure 3(c) shows the variation of the prediction RMSE for different dimensions with the Rosenbrock function. It follows that for lower dimensions changing the HP affected the RMSE much less than for higher dimensions, as observed in the experiments.

Beyond the above per-factor analysis, an additional evaluation was performed to examine if and to what extent the impact of the HP varied depending on the main algorithm parameter of the limit of true (expensive) function evaluations. Accordingly, beyond the baseline  $50d$  evaluations limit the full set of numerical experiments was repeated also for a limit of  $100d$  and  $200d$ . Table 8 summarizes the obtained results for the number of cases in which the HP had a significant impact. It follows that the impact was consistent across all sets and that the number of cases with a statistically significant performance difference and the test cases in which they occurred remained largely unchanged across the three settings. Accordingly, the remainder of the analysis is presented for the  $50d$  setting.

Overall, the analysis shows that except for specific cases (Gaussian BF/low modality functions/low dimensions), the HP had a significant impact on the overall SAEA performance.

**4.2. Dynamically Optimized HP.** The preceding analysis examined cases where the HP value was unchanged. This raises the following question: *would the overall SAEA performance vary if the HP is dynamically calibrated during the search?* For a rigorous evaluation an additional set of experiments was performed with a modified algorithm in which the HP was dynamically

calibrated based on a cross-validation procedure with an 80-20 training-testing split ratio. The HP optimization was achieved by a golden search algorithm in the interval  $[10^{-4}, 1]$ . Tests were conducted as in the preceding section and results are presented in Table 9 and compared to those from the best-performing variant (mean-wise) from the preceding section. It follows that optimizing the HP consistently improved performance as evident both from the mean and  $p$ -values where the variant with the optimized HP had a statistically significant performance advantage at 0.05 level ( $p$ -value  $< 0.05$ ) in 8 out of 10 cases (all except the Ackley-5D and Rosenbrock-5D functions).

Furthermore, the prediction accuracies of the optimized and fixed HP variants were compared and the results are presented in Figure 4. It follows that the optimized HP variant had significantly smaller prediction errors and consequently achieved better final results. This trend was observed across different test cases, as shown. Overall, the analysis shows that dynamically optimizing the HP consistently improved the SAEA performance.

## 5. Conclusion

Surrogate-assisted evolutionary algorithms (SAEA) are often used to solve simulation-driven optimization problems and a common surrogate is the radial basis functions (RBF) model. The latter relies on a hyperparameter (HP) which affects both the individual basis functions and the overall surrogate prediction. The HP needs to be set by some numerical procedure and this raises the question how significant is its impact on the overall SAEA performance and not only on the stand-alone surrogate, namely, is the computational effort spent on the HP calibration justified in terms of the overall SAEA performance. This is an important consideration both for end-users and algorithm researchers.

To address this issue, this paper has presented an extensive set of experiments involving a representative RBF-based SAEA. Analysis of the large volume of results shows that a fixed HP typically had a significant impact on performance except for specific cases (Gaussian basis function/low dimension/high modality). Analysis also shows that dynamically optimizing the HP during the search significantly improved performance. These results indicate that a careful a-priori selection of the HP can significantly improve the SAEA effectiveness while additional gains could be achieved with dynamic HP calibration.

## Data Availability

Log files of the numerical experiments are available from the corresponding author upon request.

## Conflicts of Interest

The author declares that there are no conflicts of interest.

## References

- [1] Y.-S. Ong, Z. Zhou, and D. Lim, "Curse and blessing of uncertainty in evolutionary algorithm using approximation," in *Proceedings of the 2006 IEEE Congress on Evolutionary Computation (CEC)*, IEEE, Vancouver, Canada, 2006.
- [2] K. Deb, H. Rayan, R. Proteek, and G. Toscano, "Classifying metamodeling methods for evolutionary multi-objective optimization: first results," in *Proceedings of the 2017 International Conference on Evolutionary Multi-Criterion Optimization*, Münster, Germany, 2017.
- [3] Y. Tenne and C. K. Goh, "Computational intelligence in expensive optimization problems," *Evolutionary Learning and Optimization*, Vol. 2, Springer, Berlin, Germany, 2010.
- [4] E. Kontoleon, M. Zormpa, S. Nichtawitz, D. Mack-Sahl, and S. Weissenberger, "Multilevel design optimization of hydraulic turbines based on hierarchical metamodel-assisted evolutionary algorithms," in *Proceedings of the 29th IAHR Symposium on Hydraulic Machinery and Systems*, Kyoto, Japan, 2019.
- [5] T. X. Dinh and K. K. Ahn, "Radial basis function neural network based adaptive fast nonsingular terminal sliding mode controller for piezo positioning stage," *International Journal of Control, Automation, and Systems*, vol. 15, no. 6, pp. 2892–2905, 2017.
- [6] E. G. Baquela and A. C. Olivera, "A novel hybrid multi-objective metamodel-based evolutionary optimization algorithm," *Operations Research Perspectives*, vol. 6, Article ID 100098, 2019.
- [7] J. Shin, S. Baek, and Y. Rhee, "Wind farm layout optimization using a metamodel and EA/PSO algorithm in Korea offshore," *Energies*, vol. 14, no. 146, 2020.
- [8] Y. Tenne, "Machine-learning in optimization of expensive black-box Functions," *International Journal of Applied Mathematics and Computer Science*, vol. 27, no. 1, pp. 105–118, 2017.
- [9] Y. Tenne, "Online ensemble topology selection in expensive optimization problems," *International Journal of Control, Automation, and Systems*, vol. 18, no. 4, pp. 955–965, 2020.
- [10] M. Yousefi, M. Yousefi, R. P. M. Ferreira, J. H. Kim, and F. S. Fogliatto, "Chaotic genetic algorithm and adaboost ensemble metamodeling approach for optimum resource planning in emergency departments," *Artificial Intelligence in Medicine*, vol. 84, no. 10, pp. 23–33, 2018.
- [11] R. G. Regis, "Particle swarm with radial basis function surrogates for expensive black-box optimization," *Journal of Computational Science*, vol. 5, no. 1, pp. 12–23, 2014.
- [12] B. Talgorn, S. Le Digabel, and M. Kokkolaras, "Statistical surrogate formulations for simulation-based design optimization," *Journal of Mechanical Design*, vol. 137, no. 2, 2015.
- [13] H. B. Nguyen, B. Xue, and P. Andreae, "PSO with surrogate models for feature selection: static and dynamic clustering-based methods," *Memetic Computing*, vol. 10, no. 3, pp. 291–300, 2018.
- [14] X. Ji, Y. Zhang, D. Gong, and X. Sun, "Dual-surrogate-assisted cooperative particle swarm optimization for expensive multimodal problems," *IEEE Transactions on Evolutionary Computation*, vol. 25, no. 4, pp. 794–808, 2021.
- [15] Y. Tenne, "An adaptive-topology ensemble algorithm for engineering optimization problems," *Optimization and Engineering*, vol. 16, no. 2, pp. 303–334, 2015.
- [16] H. D. Wang, Y. C. Jin, C. L. Sun, and J. Doherty, "Offline data-driven evolutionary optimization using selective surrogate ensembles," vol. 91, pp. 203–216, 2020.
- [17] C. D. Fan, B. Hou, J. H. Zheng, L. Xiao, and L. Yi, "A surrogate-assisted particle swarm optimization using ensemble learning for expensive problems with small sample datasets," *Applied Soft Computing*, vol. 91, Article ID 106242, 2020.
- [18] H. C. Dong, B. W. Song, P. Wang, and Z. Dong, "Surrogate-based optimization with clustering-based space exploration for expensive multimodal problems," *Structural and Multidisciplinary Optimization*, vol. 57, no. 4, pp. 1553–1577, 2018.
- [19] D. J. J. Toal, N. W. Bressloff, and A. J. Keane, "Kriging hyperparameter tuning strategies," *AIAA Journal*, vol. 46, no. 5, pp. 1240–1252, 2008.
- [20] W. Song and A. J. Keane, "Parameter screening using impact factors and surrogate-based ANOVA techniques," in *Proceedings of the 11th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, Reston, VA, USA, 2006.
- [21] X. D. Du, H. Xu, and F. Zhu, "Understanding the effect of hyperparameter optimization on machine learning models for structure design problems," *Computer Aided Design*, vol. 135, no. 5, Article ID 103013, 2021.
- [22] S. Bagheri, W. Konen, and D. Anders, "Surrogate-assisted optimization for augmentation of finite element techniques," in *Proceedings of the 2014 IEEE Congress on Evolutionary Computation (CEC 2014)*, Beijing, China, 2020.
- [23] M. J. D. Powell, "Radial basis function methods for interpolation of functions of many variables," in *Proceedings of the 5th Hellenic-European Conference on Computer Mathematics and Its Applications (HERCMA-01)*, Athens, Greece, 2001.
- [24] R. L. Hardy, "Theory and applications of the multiquadric-biharmonic method 20 years of discovery 1968–1988," *Computers & Mathematics with Applications*, vol. 19, no. 8–9, pp. 163–208, 1990.
- [25] E. J. Kansa and R. E. Carlson, "Improved accuracy of multiquadric interpolation using variable shape parameters," *Computers & Mathematics with Applications*, vol. 24, no. 12, pp. 99–120, 1992.
- [26] C. Durantin, J. Rouxel, J.-A. Désidéri, and A. Glière, "Multifidelity surrogate modeling based on radial basis functions," *Structural and Multidisciplinary Optimization*, vol. 56, no. 5, pp. 1061–1075, 2017.
- [27] F. J. Hickernell and Y. C. Hon, "Radial basis functions approximation of the surface wind field from scattered data,"

- International Journal of Applied Science and Computation*, vol. 4, no. 3, pp. 221–247, 1998.
- [28] S. Rippa, “An algorithm for selecting a good value for the parameter  $c$  in radial basis function interpolation,” *Advances in Computational Mathematics*, vol. 11, no. 2-3, pp. 193–210, 1999.
- [29] A. Sóbester, S. J. Leary, and A. J. Keane, “On the design of optimization strategies based on global response surface approximation models,” *Journal of Global Optimization*, vol. 33, no. 1, pp. 31–59, 2005.
- [30] N. Benoudjit and M. Verleysen, “On the kernel widths in radial-basis function networks,” *Neural Processing Letters*, vol. 18, no. 2, pp. 139–154, 2003.
- [31] M. A. Golberg, C. S. Chen, and S. R. Karur, “Improved multiquadric approximation for partial differential equations,” *Engineering Analysis with Boundary Elements*, vol. 18, no. 1, pp. 9–17, 1996.
- [32] F. Marchetti, “The extension of Rippa’s algorithm beyond LOOCV,” *Applied Mathematics Letters*, vol. 120, Article ID 107262, 2021.
- [33] M. J. Milroy, G. W. Vickers, and C. Bradley, “An adaptive radial basis function approach to modeling scattered data,” *Journal of Applied Science and Computation*, vol. 16, pp. 319–349, 1994.
- [34] R. E. Carlson and T. A. Foley, “The parameter  $R^2$  in multi-quadric interpolation,” *Computers & Mathematics with Applications*, vol. 21, no. 9, pp. 29–42, 1991.
- [35] P. G. Bengisen, “A new regression-based approach to estimate the shape parameter of MQ-RBFS in a free convection problem,” *Journal of Computing and Information Science in Engineering*, vol. 20, no. 1, p. 10, 2019.
- [36] G. Q. Zhou, W. P. Zhao, and S. Xiang, “Shape parameter optimization of Gaussian radial basis function using genetic algorithm for natural frequencies of laminated composite plates,” in *Designing and Researching of Machines and Technologies for Modern Manufacture*, pp. 153–156, Trans Tech Publications Ltd, Zurich, Switzerland, 2015.
- [37] A. Ratle, “Kriging as a surrogate fitness landscape in evolutionary optimization,” *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, vol. 15, no. 1, pp. 37–49, 2001.
- [38] D. C. Montgomery, *Design and Analysis of Experiments*, Wiley, Hoboken, NJ, USA, 8th edition, 2012.
- [39] P. N. Suganthan, N. Hansen, J. J. Liang et al., *Problem Definitions and Evaluation Criteria for the CEC 2005 Special Session on Real-Parameter Optimization*, NUS, Singapore, 2005.