Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
`http://siba-ese.unisalento.it/index.php/ejasa/index`
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v15n2p399

An adaptive sampling and weighted ensemble of surrogate models for high dimensional global optimization problems

By Younis, Elbadawy

Published: 20 June 2022

Electronic Journal of Applied Statistical Analysis
Vol. 15, Issue 02, June 2022, 399-420
DOI: 10.1285/i20705948v15n2p399

An adaptive sampling and weighted ensemble of surrogate models for high dimensional global optimization problems

Adel Younis*and Mohamed Elbadawy

Mechanical Engineering Department, Australian University
West Mishref, P.O. Box 1411 Safat 13015, Kuwait

Published: 20 June 2022

The modern engineering design optimization relies heavily on high-fidelity computer. Even though, the computing ability of computers have increased drastically, design optimization based on high-fidelity simulations is still time consuming and impractical. Surrogate modeling is a technique to replace the high-fidelity simulations. This paper presents a novel approach, named weighted ensemble of surrogates (WESO) for computationally intensive optimization problems, The focus is on multi-modal functions to identify its global optima with relatively few function evaluations. WESO search mechanism falls in two steps, explore and fit. The "explore" step is based on exploring the whole design region by generating sample points (agents) using Latin hypercube sampling (LHS) technique to gain prior knowledge about the function of interest (learning phase). The "fit" step is to train and fit a weighted ensemble of surrogate models over the promising region (training phase) to mimic the computationally intensive true function and replace it with a surrogate model (cheap function). The surrogates are then utilized to select candidates' decision variable points at which the true objective function and constraints' functions to be evaluated. Weights are then determined, assigned and an ensemble of surrogate gets constructed using the candidate sample points where optimization can be carried out. WESO has been evaluated on classical benchmark functions embedded in larger dimensional spaces. WESO was also tested on the aerodynamic shape optimization

*Corresponding author: a.younis@ack.edu.kw

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

of turbo-machinery airfoils to demonstrate its ability in handling computationally intensive optimization problems. The results showed to what extent combinations of models can perform better than single surrogate models and provide insights into the scalability and robustness of the approach. WESO can successfully identify near global solutions, faster than other classical global optimization algorithms.

keywords: Surrogate models, optimization, response surface model, radial basis function, Kriging, computational design

1 Introduction

High-fidelity computer simulations such as Finite Element Analysis (FEA) and Computational Fluid Dynamics are used substantially in modern engineering design optimization (CFD). Physical features of modeled systems can now be well predicted, and physical tests are being replaced by virtual and computer simulations, thanks to recent advances in computer simulations and their accuracy in computer analysis. For applications that demand computationally complex computations and analysis, these simulations are critical. A single run of these models might be excessively time-consuming, taking days to produce a solution. On many levels, executing computer simulations with fewer resources and less calculation time helps to improve designs. Surrogate modeling is a popular technique for solving computationally complex engineering design optimization problems. It's tough to decide which surrogate is better for approximation due to a lack of prior knowledge. Ensembles of surrogate models have been gaining traction to fully use many surrogates.

Optimization algorithms that are powerful, robust, and efficient have been introduced in recent years. Surrogate models based on optimization algorithms have been widely used to solve complex, time-consuming, and computationally expensive optimization problems. Identifying good surrogate models and their appropriate locations, on the other hand, is still a task that requires a lot of human effort. Surrogates like these are helpful in acquiring a good understanding of the system's overall behavior. Surrogates, on the other hand, have varying modeling capabilities for systems with varying properties, and many surrogates only work well within a narrow design space of unimodal shape. Response surface function (RSM) Myers et al. (2016) , Kriging methods (Matheron, 1963), artificial neural network (ANN) (Kerh et al., 2008), radial basis function (RBF) (Gutmann, 2001), support vector regression (SVR) (Smola and Schölkopf, 2004), and multivariate adaptive regression splines (MARS) (Friedman, 1991) are examples of surrogates. Each surrogate has its own characteristics that allow it to handle different types of optimization problems.

For an optimization task, a strategy for selecting and adapting the most suited surrogate model from numerous models becomes crucial. Many solutions for dealing with optimization problems involving many surrogates (ensemble of surrogates), in which more than one surrogate is employed to forecast and mimic the genuine or expensive function, have been developed. The weight coefficient of an ensemble of surrogates is often defined by an error index such as root mean square (RMSE) error or R square, which superimposes numerous surrogates together in the form of weights.

The idea of combining different surrogates into a single model can be traced back to Perrone and Cooper's invention of neural networks (Perrone and Cooper, 1992). This approach has been expanded for building ensembles of surrogates by (Zerpa et al., 2005) and (Goel et al., 2007). A noteworthy example is the recent construction of an ensemble of Kalman filters (Evensen, 2003). In statistics, an ensemble technique might be seen as an alternative to model selection (Madigan and Raftery, 1994) and (Buckland et al., 1997). Kaymaz and McMahon (2005) proposed ADAPRES, which replaces normal regression with a weighted regression technique. Regis and Shoemaker (2013) proposed the dynamic coordinate search utilizing Response Surface (DYCORS) models framework for surrogate-based optimization of high-dimensional, expensive, and black-box functions, which integrates a concept from the dynamically Dimensioned Search (DDS) algorithm.

It has been demonstrated that an ensemble of individual surrogates can considerably improve the robustness and accuracy of surrogate model predictions. Using global error measurements, several attempts were made to determine the contribution of a surrogate in the ensemble. In addition, prediction variance is employed to establish the weight components as a local error measure. Friese et al. (2016) evaluated a combination of two surrogates and concluded that combining more than two surrogates could produce superior outcomes. Acar and Rais-Rohani (2009) looked into the effectiveness of employing local error measures and proposed using pointwise cross validation error as a local error measure instead of prediction variance. Zhang et al. (2016) introduced a new point-wise weighted ensemble approach to surrogate models that incorporates error classification and nearest neighbor selection into point-wise weights computation. The prediction variance was proposed by Sanchez et al. (2008) as a local error metric for determining the weight factors of the surrogate models in the ensemble of surrogates. Sanchez et al. (2008) used a local error metric with weight factors that were adaptable across the design space, resulting in more accurate predictions. According to (Viana et al., 2009), cross-validation errors are a stronger predictor of a surrogate model's overall accuracy than prediction variance. The network might have multiple initialization or random weights in neural network ensembles. The accuracy of an aggregate output obtained by combining the outputs of numerous neural networks is often superior to that of any single output. The ensemble network that results often outperforms the individual networks (Perrone and Cooper, 1992). Younis and Dong (2012) devised a mixed metamodeling method in which the algorithm divides the design specification into sub-regions and searches for and fits an ensemble of surrogates with predetermined weights. Acar and Rais-Rohani (2009) proposed an alternative method for constructing a metamodel ensemble by determining the optimal weight factor values that would minimize a certain error measure (e.g., root mean square error). They regarded the ensemble to be made up of five independent stand-alone metamodels to demonstrate the capability of their technique (i.e., PRS, RBF, KR, GP, and SVR). Modeling efficiency and accuracy are intimately related to the size of the design space in surrogate-based global optimization using computationally demanding simulations.

The use of an ensemble of surrogates to solve various optimization problems may be seen in practically every application. Lau et al. (2013) devised a mathematical method known as the minimal description length (MDL) for determining the best artificial neural network (ANN) for reliable demand forecasting. With the goal of permitting many function evaluations without raising algorithm run-time complexity, Akhtar and Shoemaker (2019) integrate the ideas of employing local surrogates and a restart mechanism to increase algorithm runtime and efficiency. Alizadeh et al. (2019) proposed a method for constructing a surrogate ensemble that is both accurate and computationally efficient. The ensemble of surrogates is a weighted average surrogate of response surface models, kriging, and radial basis functions based on overall cross-validation error, according to their proposed method. Ye and Pan (2020) attempted to solve the problem of determining the optimal number and diversity of surrogates for an ensemble. For wind turbine design, Dhamotharan et al. (2018) suggested an ensemble of surrogates-based optimization methodology. To cope with reliability-based design optimization for a vehicle protection system, Gu et al. (2015) used an ensemble of surrogates. Chen and Lu (2020) established a new adaptive reliability analysis method based on ensemble learning of numerous competitive surrogate models.

An ensemble of surrogates-based design optimization approach is presented in this research to address the need for optimization design issues with unknown characteristics, which are computationally costly, time-consuming, and rely on computer simulation. The proposed algorithm's purpose is to strategically arrange weights in the ensemble surrogate to obtain accurate solutions with a small number of function evaluations. The method uses additional design experiment data points from Latin Hypercube Designs to exploit and explore the design region to get prior knowledge of the high-fidelity function, and then fits the surrogate models over the discovered region. The weights are then adjusted based on the performance of each model, and an ensemble surrogate model is built using selected candidate points. Once the termination requirement is met, optimization on the combined surrogate is performed by sampling and assessing the decision variable points using the combined surrogate. Eventually, the algorithm converges on the global solution. Benchmark functions and real-life design applications are used to evaluate the suggested approach's performance.

The goal of this research is to propose a systematic approach to building/fitting a range of surrogates for an ensemble. The fundamental surrogate models for building ensembles are PRS, RBF, and KRG, which are calculated using a weight selection technique detailed in the next section. Prediction accuracy and robustness are used to assess these ensembles' performance on multi-modal benchmark functions and turbomachinery airfoil shape geometry surrogates. As a result, an efficient method is presented for intelligently selecting and constructing appropriate surrogates for expensive computation issues.

2 Surrogate models

In this method, a group of surrogates was used in conjunction. To meet the acquired design points, the proposed solution combines three of the most popular/different surrogates. These are the quadratic response function (QRF) Response Surface Model (RSM), the linear radial basis function (RBF), and Kriging. The next sections introduce the surrogates that were used.

2.1 Response Surface Method

Response surface methods (RSM) have been successfully employed as metamodels for over thirty years. It was originally created for the purpose of analysing physical experiments. Polynomial RSM has been successfully employed in a range of applications to build approximations. The least squares approach is used by RSM to approximate functions on a series of points in the design variable space. Low order polynomials, such as the first and second order polynomials in Equations (1) and (2), are commonly employed as approximation functions for the response surface.

$$\hat{y}(x) = \beta_o + \sum_{i=1}^{k} \beta_i x_i \tag{1}$$

$$\hat{y}(x) = \beta_o + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \sum_{j=1}^{k} \beta_{ij} x_i x_j \tag{2}$$

where, parameters, β, are computed using least squares regression by minimizing the sum of the squares of the deviations of predicted function values, y(x), from the actual function values, y(x), using Equation (3).

$$\hat{\beta} = (F^T F)^{-1} F^T \hat{y} \tag{3}$$

where F is the sample data point design matrix and $\hat{y}$ denotes the response values at each sample point. Polynomial response surface models are simple to build, and their smoothing capabilities allows for rapid convergence of noisy functions in optimization. This oversimplification, however, may be problematic for modelling highly nonlinear or irregular behaviour. Robust design, transdisciplinary optimization, adaptive techniques for global optimization, and manufacturing analysis are just a few of the uses for response surfaces.

2.2 Kriging Metamodels

Kriging (KRG) models are a common global approximation technique today. They're particularly appealing for approximating deterministic simulations because of their capacity to precisely interpolate response values collected at sampling points. The original relationship is approximated by a Kriging model.

$$y = f(x) + z(x) + \epsilon \tag{4}$$

where $f(x)$ is a polynomial with free parameters β for the response surface approach, $Z(x)$ denotes the realization of a stationary, normally distributed Gaussian random process with mean zero, variance $\boldsymbol{\sigma}^2$ and non-zero covariance. Because random errors are not included in this formulation, the term (only defines the approximation error (bias).

The unknown parameters β and $\boldsymbol{\sigma}^2$ can be calculated as follows:

$$\hat{\beta} = (F^T \boldsymbol{R}^{-1} F)^{-1} F^T R^{-1} \tilde{y} \tag{5}$$

$$\hat{\sigma}^2 = \frac{1}{n}(\tilde{y} - F\hat{\beta})^T \boldsymbol{R}^{\mathbf{-1}}(\tilde{y} - F\hat{\beta}) \tag{6}$$

In the typical response surface technique, the term $f(x)$ offers a global trend for the system behaviour. Because it causes a "localized deviation" from the polynomial part of the model, the second half of the formulation $Z(x)$ ensures interpolation of the observations $\hat{y}$ the sampling points x_i.

The covariance matrix of $Z(x)$ characterizes a Gaussian random process. It is defined as

$$Cov(Z(x_k), Z(x_l)) = \sigma^2 \boldsymbol{R}; 1 \leq (k, l) \leq 2 \tag{7}$$

with the correlation matrix

$$\boldsymbol{R} = \begin{bmatrix} R(x_1, x_1) & \cdots & R(x_1, x_n) \\ \vdots & \ddots & \vdots \\ R(x_n, x_1) & \cdots & R(x_n, x_n) \end{bmatrix} \tag{8}$$

Kriging methods can produce accurate forecasts of highly nonlinear or irregular behaviour due to the vast range of correlation functions available.

2.3 Radial Basis Function

The radial basis function (RBF) is a useful tool for smoothing and interpolating data from experiments. The approximate model's shape is determined by the Euclidean distance between the sampled data point and the forecasted point. RBF was created as an analytical method for describing irregular surfaces. It builds approximation models by combining linear combinations of radial symmetric functions of the Euclidean distance. The model can be stated mathematically as

$$\hat{y}(x) = \sum_{i=1}^{N} w_i \phi(\| x - c_i \|) \tag{9}$$

where the approximated function $\hat{y}(x))$ is represented as a sum of N radial basis functions ϕ, each associated with a difference center, c_i, and weighted by an appropriate coefficient, ω_i. The RBF approximation can match arbitrary contours of both deterministic and stochastic response functions quite well.

The radial function $\phi(x)$ used in this work is a Gaussian function as shown in Eq. 10.

$$\phi(r) = e^{-(ar)^2} \tag{10}$$

2.4 Weights Selection Method

One of the goals of this project is to discover the weight parameters for generating a surrogate ensemble. Acar (2010) and Lee (2014) published several weight selection strategies. The authors usually choose the fundamental surrogate models based on experience and personal preference when putting together an ensemble of surrogates.

Building a superior ensemble of surrogates necessitates careful weight selection. The weights are carefully assigned to increase the ensemble's overall forecast accuracy. According to Goel et al. (2007), the weights should reflect the surrogates' confidence and filter out the negative effects of surrogates that perform badly in sampling sparse regions. A weights selection technique addressing these two issues is proposed in (Goel et al., 2007) and formulated as:

$$\omega_i = \frac{\omega_i^*}{\sum_{i=1}^{m} \omega_i^*}; \omega_i^* = (E_i + \alpha \bar{E})^{\beta}; \bar{E} = \frac{1}{m}\sum_{i=1}^{m} E_i, \alpha < 1, \beta < 0 \tag{11}$$

Where ω_i is the weight associated with the i^{th} basic surrogate, E_i is the given error measure of the i^{th} basic surrogate, $\bar{E}$ indicates the average value of all surrogates' error measure. α and β are taken as 0.05 and -1 respectively as recommend by (Goel et al., 2007).

2.5 Statistical Validation Methods

Validating the surrogate model is a crucial step. It reflects the surrogate model's ability to mimic the high-fidelity model or black-box function. Before being utilized as approximation models for computation-intensive processes, surrogates need be validated. Many statistical validation techniques were developed in the previous few decades and are still in use today. This section contains the statistical models that were utilized to validate the developed surrogate models in this study.

2.5.1 Root Mean Square Error (RMSE)

The mathematical procedure for computing RMSE is shown in Equation (12)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{12}$$

2.5.2 Relative Maximum Absolute Error (RMAE)

The mathematical formula for computing RMAE is represented by Equation (13)

$$RMAE = \frac{max(\mid y_1 - \hat{y_1} \mid, \mid y_2 - \hat{y_2} \mid, \cdots, \mid y_n - \hat{y_n} \mid)}{n} \tag{13}$$

2.5.3 R-Square

The mathematical formula for calculating R-square is shown in Equation (14).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y_i})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{14}$$

where $\hat{y_i}$ is the corresponding predicted value for the observed value y_i; $\bar{y}$ is the mean of the observed values.

2.6 The proposed approach (Ensemble of surrogates)

The goal of this research is to use an appropriate linear combination of the three surrogate models, QPF, RBF, and KRG, to generate a better weighted ensemble of surrogates that can mimic the high-fidelity model in the design space or specifically in promising regions. The ensemble of surrogate models looks like this:

$$\hat{y}_{ensebmle}(X) = \omega_q Q(X) + \omega_r R(X) + \omega_k K(X) \tag{15}$$

and

$$\omega_q + \omega_r + \omega_k = 1 \tag{16}$$

where $haty_{ensebmle}(X)$ is the metamodel approximate to the analysis/simulation function, $f(X)$, at sample point x; $Q(x)$,$R(x)$ and $K(x)$ represents the QRF, RBF and KRG surrogate models respectively; ω_q, ω_r and ω_k are weight coefficients that determine the contribution of the three models to the mixed metamodel.

The following are the steps that make up the proposed strategy:

1. Randomly generating sample points using LHD in design region

$$\Phi_0 = X_1, X_2, \cdots, X_p, X_i \in S^n$$

2. Evaluate the fitness values using the cost functions

$$y = minf(X_i) : (X_i, f(X_i)) \in S$$

3. Constructing RBF, RSM and KRG surrogate models

4. Identifying and defining a promising region by identifying the upper and lower boundaries [ub, lb]

5. Using LHD to generate additional points to refine the search in the potential region

$$\Phi_j = \{X_1, \ldots X_r, X_{r+1}, \ldots, X_{r+p}\}$$

6. Selecting candidate points for evaluation utilizing created surrogates and removing any overlapping spots

7. Choosing new candidate points and calculating weights to be applied to each surrogate

$$\omega_i = \frac{\omega_i^*}{\sum_{i=1}^{m} \omega_i^*}, \omega_i^* = (E_i + \alpha \bar{E})^{\beta}$$

8. Using the weights calculated in the previous step, create an ensemble of surrogates.

$$\hat{y}_{ensemble}(X) = \omega_q Q(X) + \omega_r R(X) + \omega_k K(X)$$

9. Using the ensemble of surrogates, evaluating fresh candidate points and choosing the best solution

$$\Omega_j = \{X_1, X_2, \cdots, X_q\}, X_i \in S^n$$

$$\hat{y}_{min} = min\{\hat{f}(X_1), \cdots, \hat{f}(X_q) : (X_i, \hat{f}(X_i)) \in S^n\}$$

10. Determining the ensemble of surrogates' optimum point (best choice variables) for all iterations

$$\hat{y}_{optimum(global)} = min(\hat{y}_{min1}, \hat{y}_{min2}, \cdots, \hat{y}_{minj}) j = 1 \cdots N$$

11. If the termination criterion is met, the process will be terminated.

3 Test problems

3.1 Benchmark test functions

Computation experiments were carried out utilizing a range of benchmark test problems to demonstrate the robustness and efficiency of the newly presented approach. In terms of structure and dimensionality, these benchmark problems are deemed representative. Benchmark difficulties with a lot of dimensions are considered. Figure 1 shows the mathematical formula for benchmark functions.

3.2 Turbomachinery blades airfoil shape optimization

In this study, the proposed technique is used to optimize the shape of a common GT compressor blade airfoil (Safari et al., 2015). The goal is to reduce the total pressure loss coefficient in both design and off-design conditions. The geometry input to the optimization process was parameterized using NURBS curves. The placements of non-uniform rational basis spline (NURBS) control points are treated as design variables

Table 1: Benchmark test functions

Category	Function	D	Search space	Analytic global
High dimensional (D > 10)	Ackley	10	$\begin{bmatrix}-32 & 32\end{bmatrix}^{10}$	0.000
High dimensional (D > 16)	Hartmann	16	$\begin{bmatrix}-1 & 1\end{bmatrix}^{16}$	-25.875
High dimensional (D > 10)	Pavianini	10	$\begin{bmatrix}2.0001 & 10\end{bmatrix}^{10}$	-45.7784

in this fashion. This study's compressor airfoil is made up of four NURBS curves for four segments, with nine control points identifying each section. There are a total of 72 design variables because each control point has two x and y coordinates. However, 16 parameters are known and fixed, which are the junction positions of the segments. C2-continuity (requires the second derivative to be zero at the endpoints) is enforced at the intersections of the segments to identify the remaining 16 parameters. As a result, the optimization methods will keep 40 variables as design variables. However, to save CPU time, the geometry of leading edge (LE) and trailing edge (TE) is kept constant to reduce the number of design parameters while preserving a high level of geometric flexibility.

The objective function's formulation has a significant impact on the airfoil optimization process's outcomes. The geometry code generates parameterized profiles, which are then loaded into COMSOL CFD (computational fluid dynamics software) for a 2-D fluid flow simulation. Following the convergence check, the acceptable profiles' post-processed outputs are entered into the fitness calculation section, where the loss values, L, of the airfoils should be reduced regarding any geometry. As seen in Figure 2, the single objective function is as follows:

$$Min\ L\% = (a_1 L_s + a_2 L_d + a_3 L_c + PF) \times 100$$

Subject to:

$$(|y_3 - y_7|\ and\ |y_3 - y_8|) \leq 15\% \text{ of the chord}$$

$$0 \leq \frac{x_i}{Chord} \leq 100\%$$

$$40 \leq \frac{y_i}{Chord} \leq 55\%$$

Where a_i are weighting factors, PF is the penalty function of geometry constraint. The reduction of total pressure loss on the right (Lc) and left (Ls) sides of the design point significantly expands the operating range. y_i and x_i are design variables as shown in Fig. 2.

The following are the weighting factors for the optimization process: $a1 = 0.20$, $a2 = 0.70$, and $a3 = 0.10$.

$$L_s = \frac{(P_{o1} - P_{o2})_{stall}}{(0.5\rho V_1^2)},$$

$$L_d = \frac{(P_{o1} - P_{o2})_{dis}}{(0.5\rho V_1^2)},$$

$$L_c = \frac{(P_{o1} - P_{o2})_{choke}}{(0.5\rho V_1^2)}$$

Where P_{01} is intake total pressure, P_{02} is outlet total pressure, and V_1 is inlet velocity.

As shown in Fig. 3, a comparison of the entire airfoil geometry was performed. The unoptimized or beginning (Datum) airfoil geometry is represented by the dotted line. The dashed line depicts the RC-GA optimizer-generated airfoil geometry, whereas the continuous line depicts the APRI-generated optimized airfoil geometry. The most obvious conclusion from this graph is that the optimization process has considerably altered the shape of the airfoil in the second half of the chord, from maximum thickness to TE. When an ensemble of three surrogates is created, improved airfoil geometry is obtained, which implies better pressure distribution and contributes significantly to raising lifting force and decreasing drag force of the turbomachinery blades.

4 Test results and discussion

The selected representative test problems were subjected to extensive computational simulations, and positive results were produced. The recommended strategy (mixed surrogates) fared comparably well, according to the findings. When compared to other well-known methodologies, the suggested approach is capable of handling and solving difficult benchmark problems, discovering the global optimum with equivalent accuracy, and providing results with lower computation costs.

A summary of the test findings is offered in this section. The Paviani with 10 design variables, Hartmann with 16 design variables, and Ackley with 10 design variables are used as benchmark test problems in the literature to evaluate the performance of global search algorithms. Their dimensions and beginning range of design spece are provided in Table 1. In Table A1, you'll find the mathematical forms. In terms of problem dimensions and objective function shape features, the test problems are representative. Tables 2 to 4 summarize the statistical benchmark test findings. The optimization experiments are carried out by using an ensemble of surrogates (two to three surrogates were studied and evaluated). The ensemble of the three surrogates produced good results, with RBF and PQR being the only combination that can compete with the ensemble of three surrogates. Three surrogates, on the other hand, are still working on Ackley's test function and the turbomachinery airfoil geometry.

Also, to show the advantages of using ensemble of surrogates with three surrogate models over ensemble of only two surrogates, Tables 5 to 7 show the simulation results of using ensemble of surrogate for 10 runs. It can be observed that combining three surrogates gives better results than combining only two surrogates except in few cases.

In some benchmark problems the two approaches yielded the same result and in other benchmark problems the second approach (two surrogates) slightly outperformed the first approach (three surrogates).

The RMSE was calculated to reflect the accuracy and robustness of the prediction. When three surrogates are combined, improved fitting and regression occurs, showing that the RBF, PQR, and KRG combination yields better prediction and acceptable accuracy.

Figure 4 depicts the convergence trend of several surrogate model combinations. When three surrogates are built, the proposed algorithm begins its search from a point close to the optimal design variable point, which explains why the ensemble of surrogates outperforms others due to its speed and robustness of convergence. Figure 4 (a) depicts the Pavinai convergence trend for all created surrogates, while Figure 4 (b) Simply depicts the ensemble of three surrogates' convergence to the global optimum. The same may be said for Ackley and Hartman, as seen in Figures 4 (d), (e) and (f).

Table 2: Summary of Statistical Test Results of the Proposed Algorithm on Paviani function

	Fitness values					
Constructed Surrogates	Best	Mean	Median	Avg no. of evaluations	Avg no. of iterations	RMSE
K+R+P	-41.2990	-43.5706	-41.4778	316.6	49.2	0.122
K+P	-37.2218	-45.0699	-37.7934	210.5	51	0.417
K+R	-38.5587	-39.9133	-39.1370	210.9	46.9	0.526
R+P	-44.826	-40.8055	-41.113	155.8	51	0.191

K= Kriging; R= Radial basis Function; and P= Polynomial quadratic response surface function

5 Conclusion

A new global optimization approach, namely weighted ensemble surrogate (WESO) is introduced. WESO search mechanism falls in two steps, explore and fit. The "explore" step is based on exploring the whole design region by generating sample points (agents) using Latin hypercube sampling (LHS) technique to gain prior knowledge about the function of interest (learning phase). The "fit" step is to train and fit a weighted ensemble of surrogate models over the promising region (training phase) to mimic the computationally intensive true function and replace it with a surrogate model (cheap function). The surrogates are then utilized to select candidates' decision variable points at which the true objective function and constraints' functions to be evaluated. Weights

Table 3: Summary of Statistical Test Results of the Proposed Algorithm on Hartman

Constructed Surrogates	Fitness values Best	Mean	Median	Avg no. of evaluations	Avg no. of iterations	RMSE
K+R+P	27.1408	25.9991	26.1988	280	43.5	0.094
K+P	27.5055	26.4609	27.4181	195.7	48.6	0.128
K+R	26.7742	26.2530	26.6500	220	51	0.149
R+P	26.1536	25.9798	26.1210	166.7	44.1	0.032

Table 4: Summary of Statistical Test Results of the Proposed Algorithm on Ackley

Constructed Surrogates	Fitness values Best	Mean	Median	Avg no. of evaluations	Avg no. of iterations	RMSE
K+R+P	0.0280	0.0012	0.0225	333.2	49.8	0.069
K+P	0.3036	0.0768	0.1631	199.5	47.4	0.397
K+R	0.4146	0.0114	0.0707	214.7	43.3	0.415
R+P	1.8025	0.5814	1.8058	184.7	48	0.773

are then determined, assigned and an ensemble of surrogate gets constructed using the candidate sample points where optimization can be carried out. The new algorithm was tested using a variety of benchmark test problems and found to be performing comparably well. The experiment results showed robust performance, comparable search accuracy and good computation efficiency, making the new approach an excellent tool for computation intensive, computer analysis/simulation and black-box function based global optimization problems. The ensemble of three surrogates (PRS-RBF-KRG) is preferred in view of prediction accuracy and robustness. The objective of this paper is to guide researchers in selecting the appropriate number of surrogates to be combined and variety of surrogate modeling techniques for building ensemble models, rather than concentrating on developing novel ensemble modeling methods.

2009 2010 2019 2019 1997 2020 2018 2003 1991 2016 2007 2001 2015 2005 2008 2013 2014 1994 1963 Myers et al. (2016) 1992 2013 1989 2015 2008 2004 2011 2009 2020 2012 Zhang et al. (2016)

References

Acar, E. (2010). Various approaches for constructing an ensemble of metamodels using local measures. *Structural and Multidisciplinary Optimization*, 42(6):879–896.

Acar, E. and Rais-Rohani, M. (2009). Ensemble of metamodels with optimized weight factors. *Structural and Multidisciplinary Optimization*, 37(3):279–294.

Akhtar, T. and Shoemaker, C. A. (2019). Combining local surrogates and adaptive restarts for global optimization of moderately expensive functions. In *AIP Conference Proceedings*, volume 2070, page 020048. AIP Publishing LLC.

Alizadeh, R., Jia, L., Nellippallil, A. B., Wang, G., Hao, J., Allen, J. K., and Mistree, F. (2019). Ensemble of surrogates and cross-validation for rapid and accurate predictions using small data sets. *AI EDAM*, 33(4):484–501.

Buckland, S. T., Burnham, K. P., and Augustin, N. H. (1997). Model selection: an integral part of inference. *Biometrics*, pages 603–618.

Cheng, K. and Lu, Z. (2020). Structural reliability analysis based on ensemble learning of surrogate models. *Structural Safety*, 83:101905.

Dhamotharan, V., Jadhav, P. D., Ramu, P., and Prakash, A. K. (2018). Optimal design of savonius wind turbines using ensemble of surrogates and cfd analysis. *Structural and Multidisciplinary Optimization*, 58(6):2711–2726.

Evensen, G. (2003). The ensemble kalman filter: Theoretical formulation and practical implementation. *Ocean dynamics*, 53(4):343–367.

Friedman, J. H. (1991). Multivariate adaptive regression splines. *The annals of statistics*, 19(1):1–67.

Friese, M., Bartz-Beielstein, T., and Emmerich, M. (2016). Building ensembles of surrogates by optimal convex combination. *Bioinspired optimization methods and their applications*, pages 131–143.

Goel, T., Haftka, R. T., Shyy, W., and Queipo, N. V. (2007). Ensemble of surrogates. *Structural and Multidisciplinary Optimization*, 33(3):199–216.

Gu, X., Lu, J., and Wang, H. (2015). Reliability-based design optimization for vehicle occupant protection system based on ensemble of metamodels. *Structural and Multidisciplinary optimization*, 51(2):533–546.

Gutmann, H.-M. (2001). A radial basis function method for global optimization. *Journal of global optimization*, 19(3):201–227.

Kaymaz, I. and McMahon, C. A. (2005). A response surface method based on weighted regression for structural reliability analysis. *Probabilistic Engineering Mechanics*, 20(1):11–17.

Kerh, T., Lai, J., Gunaratnam, D., and Saunders, R. (2008). Evaluation of seismic design values in the taiwan building code by using artificial neural network. *Computer Modeling in Engineering and Sciences*, 26(1):1.

Lau, H. C., Ho, G. T., and Zhao, Y. (2013). A demand forecast model using a combination of surrogate data analysis and optimal neural network approach. *Decision*

Support Systems, 54(3):1404–1416.

Lee, Y. and Choi, D.-H. (2014). Pointwise ensemble of meta-models using v nearest points cross-validation. *Structural and Multidisciplinary Optimization*, 50(3):383–394.

Madigan, D. and Raftery, A. E. (1994). Model selection and accounting for model uncertainty in graphical models using occam's window. *Journal of the American Statistical Association*, 89(428):1535–1546.

Matheron, G. (1963). Principles of geostatistics. *Economic geology*, 58(8):1246–1266.

Myers, R. H., Montgomery, D. C., and Anderson-Cook, C. M. (2016). *Response surface methodology: process and product optimization using designed experiments*. John Wiley & Sons.

Perrone, M. P. and Cooper, L. N. (1992). When networks disagree: Ensemble methods for hybrid neural networks. Technical report, Brown Univ Providence Ri Inst for Brain and Neural Systems.

Regis, R. G. and Shoemaker, C. A. (2013). Combining radial basis function surrogates and dynamic coordinate search in high-dimensional expensive black-box optimization. *Engineering Optimization*, 45(5):529–555.

Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P. (1989). Design and analysis of computer experiments. *Statistical science*, 4(4):409–423.

Safari, A., Younis, A., Wang, G., Lemu, H., and Dong, Z. (2015). Development of a metamodel assisted sampling approach to aerodynamic shape optimization problems. *Journal of Mechanical Science and Technology*, 29(5):2013–2024.

Sanchez, E., Pintos, S., and Queipo, N. V. (2008). Toward an optimal ensemble of kernel-based approximations with engineering applications. *Structural and multidisciplinary optimization*, 36(3):247–261.

Smola, A. J. and Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and computing*, 14(3):199–222.

Sun, G., Li, G., Gong, Z., He, G., and Li, Q. (2011). Radial basis functional model for multi-objective sheet metal forming optimization. *Engineering Optimization*, 43(12):1351–1366.

Viana, F. A., Haftka, R. T., and Steffen, V. (2009). Multiple surrogates: how cross-validation errors can help us to obtain the best predictor. *Structural and Multidisciplinary Optimization*, 39(4):439–457.

Ye, P. and Pan, G. (2020). Selecting the best quantity and variety of surrogates for an ensemble model. *Mathematics*, 8(10):1721.

Younis, A. and Dong, Z. (2012). Global optimization using mixed surrogates and space elimination in computationally intensive engineering designs. *International Journal for Computational Methods in Engineering Science and Mechanics*, 13(4):272–289.

Zerpa, L. E., Queipo, N. V., Pintos, S., and Salager, J.-L. (2005). An optimization methodology of alkaline–surfactant–polymer flooding processes using field scale numerical simulation and multiple surrogates. *Journal of Petroleum Science and Engineering*, 47(3-4):197–208.

Zhang, Y., Chen, X., Wang, N., Yao, W., and Du, B. (2016). Ensemble of surrogates based on error classification by unsupervised learning. In *2016 IEEE Congress on Evolutionary Computation (CEC)*, pages 4344–4349. IEEE.

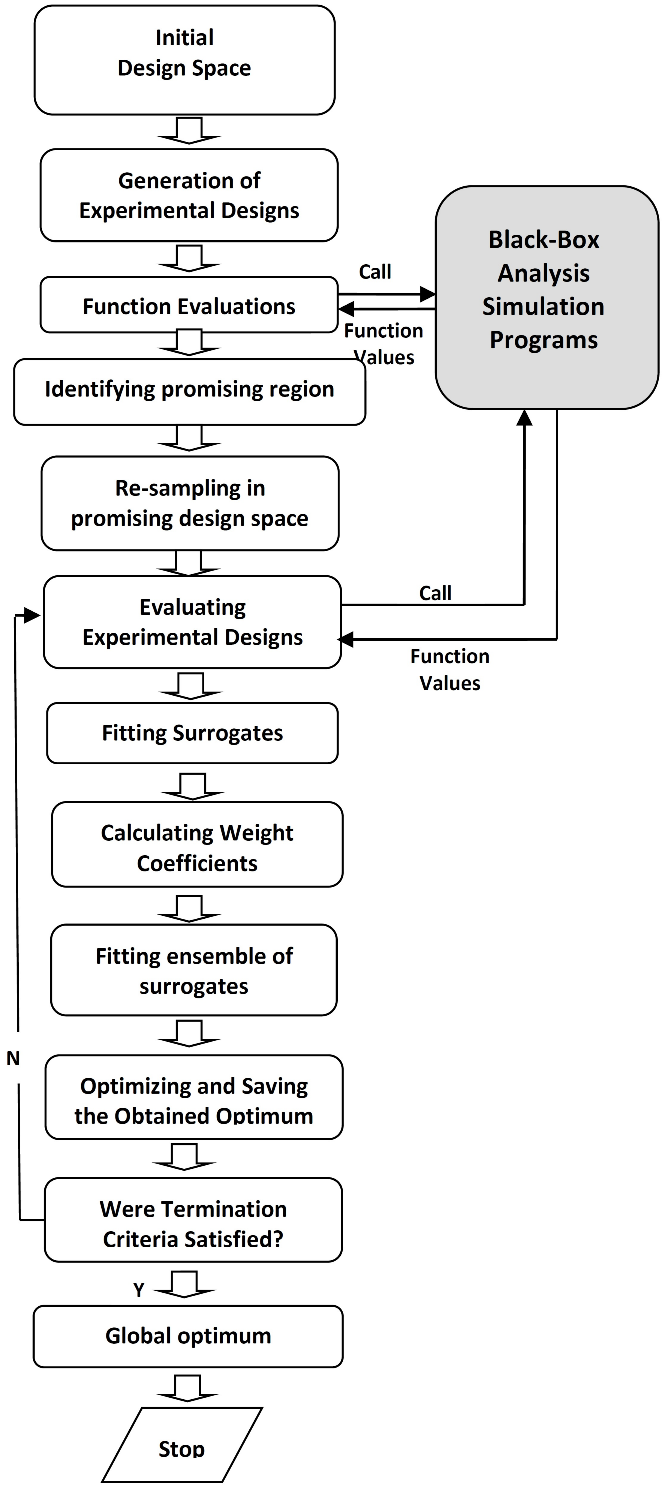

Figure 1: depicts the flow chart for the suggested technique.

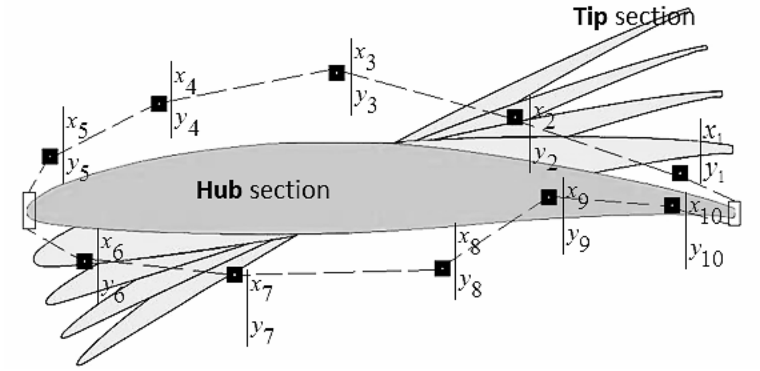

Figure 2: Schematic diagram of twenty shape design variables (Safari et al., 2015)

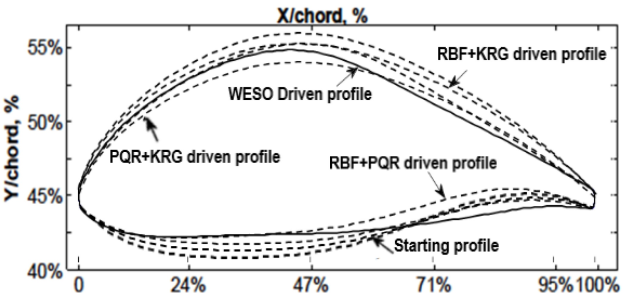

Figure 3: Airfoil shape before and after optimization.

Table 5: Performance comparison of Paviani function with different driven ensemble surrogates

Surrogate assisted model	Run No.	f^*	NOE	# Iterations	Surrogate assisted model	Run No.	f^*	NOE	# Iterations
K+R+P Driven	#1	-40.6644	291	45	K+R+P Driven	#1	-33.4481	205	51
	#2	-40.9998	320	51		#2	-32.9570	201	51
	#3	-36.9048	331	51		#3	-42.3735	212	51
	#4	-41.9558	312	48		#4	-40.0034	210	51
	#5	-39.0343	331	51		#5	-41.4583	201	51
	#6	-43.1459	330	51		#6	-45.0699	210	51
	#7	-40.4866	323	51		#7	-37.5656	222	51
	#8	-42.9396	288	45		#8	-33.9391	216	51
	#9	-43.5706	338	51		#9	-27.3819	215	51
	#10	-43.2882	302	48		#10	-38.0213	213	51
Surrogate assisted model	Run No.	f^*	NOE	# Iterations	Surrogate assisted model	Run No.	f^*	NOE	# Iterations
K+R+P Driven	#1	-39.0893	220	51	K+R+P Driven	#1	-43.1555	153	51
	#2	-33.9908	58	13		#2	-40.2848	164	51
	#3	-39.2194	212	48		#3	-39.1702	161	51
	#4	-39.1846	233	51		#4	-39.05339	154	51
	#5	-39.9133	234	51		#5	-41.94126	156	51
	#6	-38.3937	217	51		#6	-42.5321	153	51
	#7	-39.7890	241	51		#7	-42.60387	151	51
	#8	-38.4542	227	51		#8	-44.82604	147	51
	#9	-39.5528	233	51		#9	-38.6702	161	51
	#10	-38.0004	234	51		#10	-35.8177	158	51

Table 6: Performance comparison of Hartmann's function with different driven ensemble surrogates

Surrogate assisted model	Run No.	f^*	NOE	# Iterations	Surrogate assisted model	Run No.	f^*	NOE	# Iterations
	#1	26.2855	290	45		#1	27.7551	206	51
	#2	26.1160	275	42		#2	28.9615	201	51
	#3	26.1758	306	48		#3	28.5915	201	51
	#4	26.1110	289	45		#4	26.4609	173	42
K+R+P Driven	#5	25.9991	325	51	K+R+P Driven	#5	26.6331	209	51
	#6	26.2218	194	30		#6	27.2154	204	51
	#7	26.6281	309	48		#7	28.0189	212	51
	#8	26.0210	275	42		#8	27.5910	196	51
	#9	26.3396	283	45		#9	27.2452	203	51
	#10	26.0553	254	39		#10	26.5826	151	36
Surrogate assisted model	Run No.	f^*	NOE	# Iterations	Surrogate assisted model	Run No.	f^*	NOE	# Iterations
	#1	-26.6569	215	51		#1	26.2188	155	42
	#2	26.5398	224	51		#2	26.1111	193	51
	#3	26.6431	222	51		#3	26.1146	151	39
	#4	26.2530	220	51		#4	25.9897	196	51
K+R+P Driven	#5	26.8666	220	51	K+R+P Driven	#5	26.3954	190	51
	#6	28.3216	220	51		#6	25.9864	193	51
	#7	26.7158	219	51		#7	25.9798	178	48
	#8	26.3660	224	51		#8	26.3389	136	36
	#9	27.0946	214	51		#9	26.2744	158	42
	#10	26.2843	214	51		#10	26.1273	117	30

Table 7: Performance comparison of Ackley's function with different driven ensemble surrogates

Surrogate assisted model	Run No.	f^*	NOE	# Iterations	Surrogate assisted model	Run No.	f^*	NOE	# Iterations
	#1	0.0096	339	51		#1	0.1052	218	51
	#2	0.0086	304	45		#2	0.2143	190	45
	#3	0.0415	339	51		#3	0.8256	115	27
	#4	0.0308	321	48		#4	0.1251	219	51
K+R+P Driven	#5	-39.0343	331	51	K+R+P Driven	#5	-41.4583	201	51
	#6	0.0134	342	51		#6	0.1517	216	51
	#7	0.0651	349	51		#7	0.9231	211	51
	#8	0.0302	335	51		#8	0.0768	191	45
	#9	0.0012	351	51		#9	0.1429	217	51
	#10	0.0147	333	51		#10	0.1744	207	51
Surrogate assisted model	Run No.	f^*	NOE	# Iterations	Surrogate assisted model	Run No.	f^*	NOE	# Iterations
	#1	0.0203	230	45		#1	1.2643	199	51
	#2	0,0621	249	51		#2	0.5814	195	51
	#3	0.0164	256	51		#3	2.6376	196	51
	#4	0.1240	258	51		#4	2.6538	136	33
K+R+P Driven	#5	0.0114	247	51	K+R+P Driven	#5	1.3503	149	51
	#6	0.7888	71	13		#6	2.6581	150	39
	#7	2.1723	90	18		#7	3.0653	189	51
	#8	0.0793	249	51		#8	0.7937	192	51
	#9	0.0616	259	51		#9	2.2613	200	51
	#10	0.8094	245	51		#10	0.7594	196	51

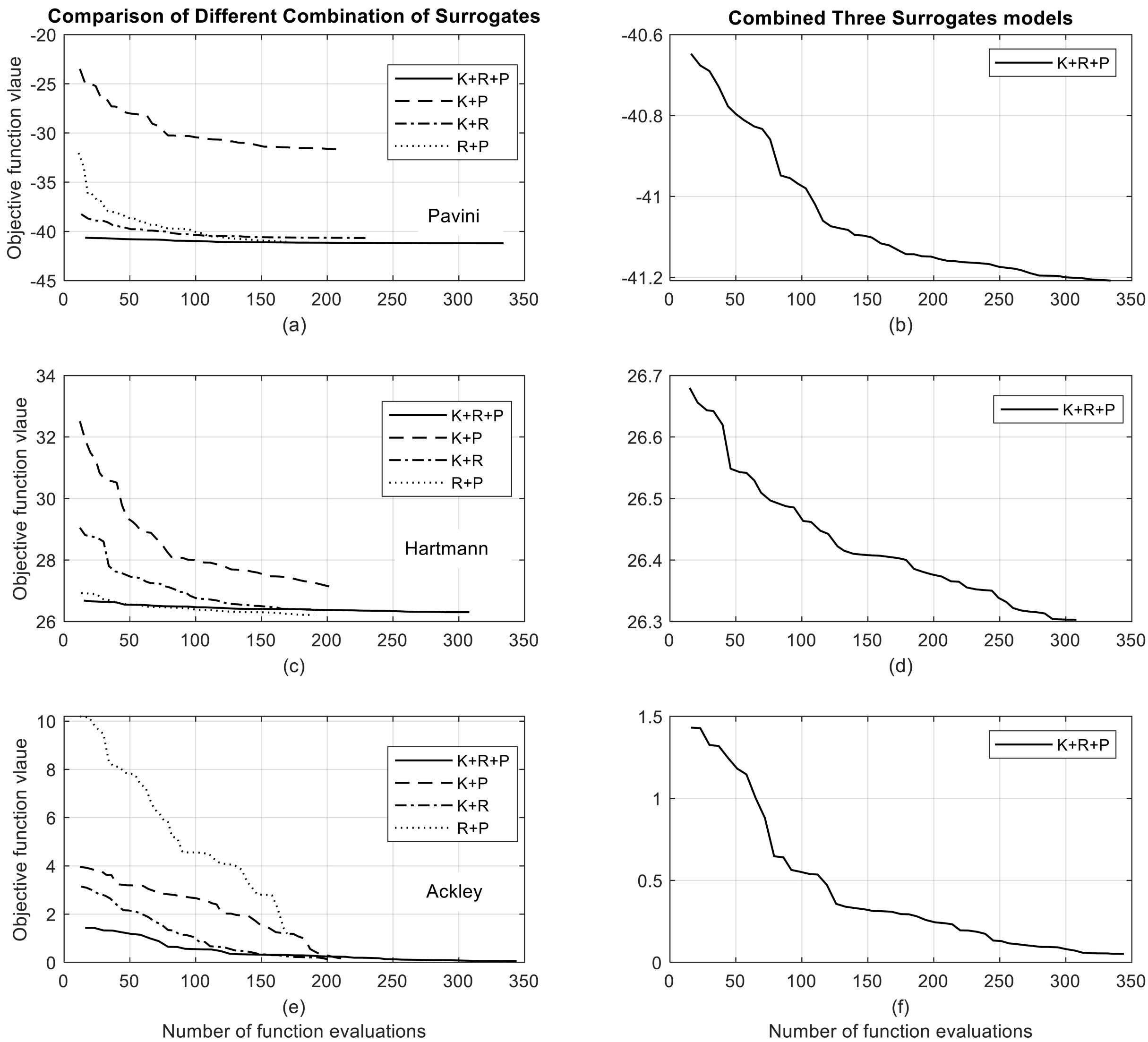

Figure 4: convergence trends and convergence speed of ensemble of surrogates on Paviani, Hartmann and Ackley (a) (c) and (e) respectively. (b), (d), and (f) ensemble of best number of surrogates for Paviani, Hartmann and Ackley respectively.