# Comparison of Response Surface and Kriging Surrogates in Aerodynamic Design Optimization of Hypersonic Spiked Blunt Bodies

M. Y. M. Ahmed[*] and N. Qin[**]

**Abstract:** The high fidelity numerical simulation codes used in aerodynamic design and optimization applications show a number of drawbacks. Of these, the high computational cost, memory and time demands associated with complicated designs. Metamodels or surrogates provide a much cheaper alternative for these codes. The use of spikes in the design of hypersonic vehicles yields a considerable reduction in drag and aerodynamic heating effects. In this work, surrogates are used in the context of design optimization of a spiked blunt body in hypersonic flow conditions. Four different surrogate models including quadratic response surface and kriging were constructed based on the values of drag and heating responses. The evolutionary genetic algorithm is applied to find the optimum design based on each of the surrogates. The structure of the various models was investigated and the main differences were addressed. For the cases investigated in this work, kriging surrogate based on exponential correlation produced a relatively better prediction of new points in the design space and, consequently, better optimized design.

**Keywords:** Aerodynamic design optimization, Surrogate modeling, Design of Experiments.

---

[*] PhD Student, Department of Mechanical Engineering, University of Sheffield,
m.ahmed@sheffield.ac.uk
[**] Professor of Aerodynamics and Fluid Mechanics, Department of Mechanical Engineering,
University of Sheffield, Senior Member AIAA, n.qin@sheffield.ac.uk

## Nomenclature

| | | | |
|---|---|---|---|
| $\mathbf{A}$ | matrix of training points values | $T$ | temperature, $[K]$ |
| $C_D$ | drag coefficient | $x$ | design parameter |
| $D$ | body diameter, $[m]$ | $\mathbf{x}$, $\bar{\mathbf{x}}$ | vector of design parameters |
| $d$ | weighted Euclidian distance | $y$ | exact response |
| $k$ | number of design parameters | $\hat{y}$ | surrogate response |
| $L$ | length, $[m]$ | $\mathbf{y}$ | vector of exact responses |
| $m$ | number of polynomial coefficients | $z(\mathbf{x})$ | Gaussian stochastic function |
| $n$ | number of design points | $\boldsymbol{\alpha}$ | vector of surrogate parameters |
| $p$ | exponent of correlation function | $\beta$ | RSM polynomial coefficient |
| $q$ | number of untried points | $\boldsymbol{\beta}$ | vector of polynomial coefficients |
| $\mathbf{R}$ | correlation matrix | | mean of random field |
| $\mathbf{r}$ | correlation vector | $\varepsilon$ | error |
| $R$ | main body radius, $[m]$ | $\theta$ | correlation weights |
| $R^2$ | error estimate | $\sigma^2$ | error variance |
| $scf(\ )$ | spatial correlation function | $\Omega$ | design space |

*Subscripts*

| | |
|---|---|
| $i$, $j$ | a point in the design space |
| $l$ | a design parameter |

*Abbreviations*

| | | | |
|---|---|---|---|
| *ADO* | aerodynamic design optimization | *LHS* | Latin hypercube sampling |
| *CFD* | computational fluid dynamics | *MAE* | maximum absolute error |
| *EKG* | general-exponential kriging | *QRS* | quadratic response surface |
| *e1ft, e1rt, e2ft, e2rt* | experiments' notation, Table 2 | | |
| *GA* | genetic algorithm | *RMS* | root mean square error |
| *GKG* | Gaussian kriging | *XKG* | exponential kriging |
| *KG* | kriging | | |

## I. Introduction

In aerodynamic design and optimization (*ADO*) problems, a design's performance is estimated using numerical models based on computational fluid dynamics (*CFD*) techniques. These high-fidelity models proved to be a reliable, efficient, flexible, and relatively cheap means of analysis and design especially compared with experimental methods. The main drawback of using numerical analysis codes is that they are computationally expensive, highly memory demanding, and time consuming. Despite the fact that these models use simplified equations, the continuously improving computers, and the implementation of parallel computing made the numerical analysis more manageable; the demand for more accurate calculations and the search for more complicated designs arise as well. Another drawback of high fidelity models is the occurrence of numerical noise. Typical sources of numerical noise include the incomplete numerical convergence, numerical round-off, and discretization errors.

These drawbacks become more severe in the context of optimization since it involves more computations. The conventional gradient-based search techniques require hundreds or thousands of *CFD* calls. Consequently, the optimization cost of complex designs becomes rather expensive. In addition, since gradient-based search techniques deal with one design at a

time, parallel computation that minimizes the cost can not be used. The numerical errors may lead the search algorithm to false optima. Another feature related to these codes is that a majority of them are originally designed as stand-alone modules and complications arise when trying to interfere their performance, sometimes the source codes are not accessible at all. Consequently, an increasing effort has been devoted to search for cheap alternatives of the high fidelity analysis codes in analysis, design, and optimization. One of these alternatives is the use of approximation models, metamodels, or surrogates.

The basic idea of the using the surrogates (approximation models) is to replace the high fidelity, expensive analysis code with a less expensive approximate model. Instead of dealing with one design at a time, a population of $n$ designs (the training points) is simulated in the same time (possibly making use of parallel computation) using the high fidelity model. Based on the estimated responses, a low fidelity surrogate is constructed. The one-shot calculations associated with surrogate construction are relatively cheaper than gradient evaluations involved in *CFD*-based optimization. Once the surrogate is constructed, the high fidelity model is overlooked and the subsequent search is made directly on the surrogate model. An additional step is required in surrogate-based optimization; the optimum design attained by the search technique must be validated using the high fidelity model. In some cases, the surrogate model accuracy is improved by augmenting the training set with the optimum; this can be repeated until no further improvement is achieved. The role of surrogates in aerodynamic design optimization can be visualized in Fig. 1.
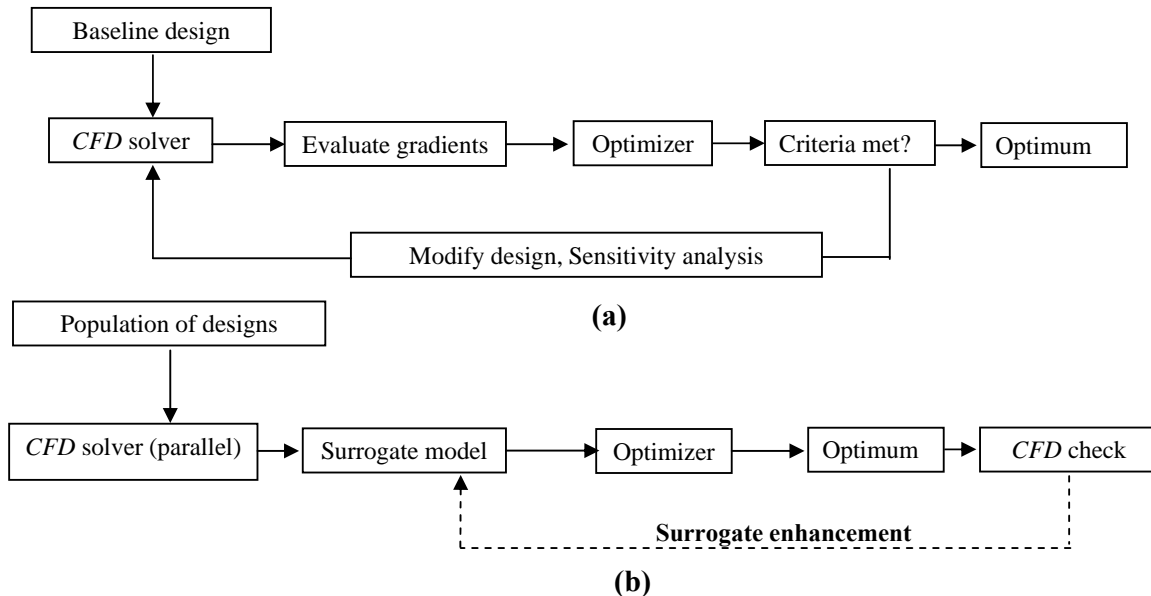


**Fig. 1 Basic Elements of (a) CFD-based and (b) CFD-surrogate-based Optimization**

In addition to lower computation cost, smoothing out the error noise arising from numerical simulation, and no needed modification to the analysis code, surrogate models give better "feeling" of the inputs-outputs functional relations. When used in optimization, surrogates give additional benefits such as smoothing out the noise arising from evaluating the objective and constraint functions, easy adaptation to parallel computations, separating the analysis code from the optimization algorithms, and easy integration of various codes used in multidisciplinary optimization. The high-fidelity model is commonly represented by the functional relation

$$y = f(\mathbf{x}) \tag{1}$$

where $\mathbf{x} = \{x_1, x_2, .... x_p\} \in \mathrm{R}^k$ is the vector of $k$ design inputs (parameters) and $y$ is the scalar output (also called the response, dependent variable). On the other hand, a surrogate model is expressed as

$$\hat{y} = \hat{f}(\mathbf{x}, \boldsymbol{\alpha}) \approx f(\mathbf{x}) \tag{2}$$

where $\boldsymbol{\alpha}$ is a vector of undetermined parameters that must be evaluated prior to applying the surrogate. There is a variety of surrogate models; among them response surface and kriging models are mature enough and are widely used in many applications. In this work, these two models are used in the context of design optimization of spiked blunt body in hypersonic flow.

The remainder of this paper is organized as follows: The structure of the used surrogates is presented in the next section. Then, we illustrate the design problem in concern, namely, the spiked hypersonic vehicle. The main results are discussed near the end of the paper which finalizes with the main conclusions.

## II. Surrogates' Structure

### II.1. Response Surface Model (RSM)

The response surface model, also termed polynomial regression, is a parametric regression model. Parametric means that the model uses the training points' responses to estimate the unknown parameters $\boldsymbol{\alpha}$; once they are known, the training set is no longer used and only the parameters decide the response at new points. Regression means that the training set responses are fit by the model. RSM was originally tailored for physical experiments [1] where the output is characterized by random errors however, it is commonly used in deterministic numerical experiments to account for random numerical errors. In this model, the training points' responses are fit by a polynomial plus an error.

$$\hat{y} = y + \varepsilon \tag{3}$$

where $\varepsilon$ represents the random error which is assumed to be normally distributed with zero mean and variance $\sigma^2$. The approximation polynomial can have any order; however, it is typically of the first order or second order. In this work, a quadratic (second order) polynomial is used. The model is expressed as

$$\hat{y} = \beta_o + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \sum_{j \geq i}^{k} \beta_{ij} x_i x_j \tag{4}$$

respectively, where $\beta_o$ is called the intercept, $\beta_{ij}, i \neq j$ are the interaction coefficients, $x_i$ refers to one of the $k$ design parameters. The polynomial coefficients $\beta' s$ (also known as the regression parameters) represent the undetermined parameters vector $\boldsymbol{\alpha}$. The values of $\beta' s$ reflect the importance (dominance) of the respective design parameter. They are estimated via

minimizing the sum of squares of deviations of predicted values, $\hat{\mathbf{y}}(\mathbf{x})$, from the actual high-fidelity values, $\mathbf{y}(\mathbf{x})$, using the equation:

$$\boldsymbol{\beta} = \left[\mathbf{A}^T \mathbf{A}\right]^{-1} \mathbf{A}^T \mathbf{y} \tag{5}$$

where $\mathbf{y}$ is a column vector that contains the exact values of the response at all training points and $\mathbf{A}^T$ is the transpose of matrix $\mathbf{A}$ expressed as:

$$\mathbf{A}(n \times m) = \begin{bmatrix} 1 & x_{11} & \cdots & x_{11}x_{12} \cdots & x_{1k}^2 \\ 1 & x_{21} & \cdots & \cdots & \cdots & x_{2k}^2 \\ . & . & \cdots & \cdots & \cdots & . \\ . & . & \cdots & \cdots & \cdots & . \\ 1 & x_{n1} & \cdots & x_{n1}x_{n2} \cdots & x_{nk}^2 \end{bmatrix} \tag{6}$$

where $m = (k+1)(k+2)/2$. The response at a new point $\overline{\mathbf{x}}$, $\hat{y}(\overline{\mathbf{x}})$, is directly evaluated by substituting for $\overline{\mathbf{x}}$ in the polynomial equation.

## II.2. Kriging Model (KG)

Kriging, also knows as design and analysis of computer experiments DACE, is a nonparametric interpolation model. Nonparametric means that the training points are involved in estimating the unknown parameters $\boldsymbol{\alpha}$ and predicting the new points' responses as well. Interpolation means that the model exactly interpolates the responses at all training points. In contrast to RSM, kriging was originally tailored for computer experiments characterized by deterministic errors. Kriging imposes a global model that interpolates all design points plus "localized" functions representing the deviations (departure) from the global model at all points. The response is expressed as

$$\hat{y}(\mathbf{x}) = f(\mathbf{x}) + z(\mathbf{x}) \tag{7}$$

where $f(\mathbf{x})$ is a low-order polynomial that interpolates the design points. Typically, a constant value was found sufficient for modelling complex input-output relations [2]. Hence, the output can be viewed as a random field with mean $\boldsymbol{\beta}$,

$$\hat{y}(\mathbf{x}) = \boldsymbol{\beta} + z(\mathbf{x}) \tag{8}$$

$z(\mathbf{x})$ is a Gaussian stochastic function that represents the realization of random process with zero mean, variance $\sigma^2$, and covariance given by

$$Cov(Z) = \sigma^2 \mathbf{R}(\mathbf{x}^i, \mathbf{x}^j) \tag{9}$$

where $\mathbf{R}(\mathbf{x}^i, \mathbf{x}^j)$ is the correlation matrix which is an $n \times n$ symmetric matrix with ones in the diagonal. Other terms in the matrix are given by the spatial correlation function

$$R(\mathbf{x}^i, \mathbf{x}^j) = \exp\left[-d(\mathbf{x}^i, \mathbf{x}^j)\right] \qquad (10)$$

where

$$d\left(\mathbf{x}^i, \mathbf{x}^j\right) = \sum_{l=1}^{k} \theta_l \left(\left|x_l^i - x_l^j\right|^p\right) \qquad (11)$$

$i$ and $j$ denote two training points, $l$ refers to a design parameter, $\theta$ is a positive weight factor related to each design parameter, and $k$ is the number of design parameters. The value of $\theta$ reflects the degree of nonlinearity of temperature response with respect to the corresponding design parameter [2]. The spatial correlation function can be viewed as functions of the "weighted" distance between samples.

Three different correlation functions are used in this work: exponential ($p = 1$), Gaussian ($p = 2$), and general exponential ($0 < p < 2$); a value of $p$ closer to $2$ indicates that the response function is highly differentiable with respect to the design parameters [2]. The mean parameter, $\boldsymbol{\beta}$, is evaluated by minimizing the sum of squares of error using the equation

$$\boldsymbol{\beta} = \left[\mathbf{A}^T \mathbf{R}^{-1} \mathbf{A}\right]^{-1} \mathbf{A}^T \mathbf{R}^{-1} \mathbf{y} \qquad (12)$$

where $\mathbf{A}$ is an $n \times m$ matrix of training set points depending on the choice of the function $f(\mathbf{x})$; in the constant case, $\mathbf{A}$ is a $n \times 1$ vector of all ones. The parameters $\theta$ and $p$ that ensure the "best fit" of the model to the training data are evaluated by maximizing the likelihood estimation MLE expressed as:

$$-\frac{1}{2}\left[n \ln(2\pi) + n \ln \sigma^2 + \ln|\mathbf{R}| + \frac{1}{2\sigma^2}(\mathbf{y} - \mathbf{A}\boldsymbol{\beta})^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{A}\boldsymbol{\beta})\right] \qquad (13)$$

where the maximum likelihood estimation of $\sigma^2$ is expressed as

$$\sigma^2 = \frac{1}{n}(\mathbf{y} - \mathbf{A}\boldsymbol{\beta})^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{A}\boldsymbol{\beta}) \qquad (14)$$

The response at a new point $\bar{\mathbf{x}}$, $\hat{y}(\bar{\mathbf{x}})$, is directly evaluated by applying the equation

$$\hat{y}(\bar{\mathbf{x}}) = \boldsymbol{\beta} + \mathbf{r}^T(\bar{\mathbf{x}})\mathbf{R}^{-1}(\mathbf{y} - \mathbf{A}\boldsymbol{\beta}) \qquad (15)$$

where $\mathbf{r}(\bar{\mathbf{x}})$ is the correlation vector between $\bar{\mathbf{x}}$ and all training points. The main pros and cons and ranges of application of RSM and kriging surrogates were discussed in the literatures, for example, [2, 3, 4]. They can be summarized in Table 1.

**Table 1 Basic Differences between RSM and KG**

| | Advantages | Disadvantages |
|---|---|---|
| **RMS** | - Well-established, easy to implement.<br>- Better performance is expected for low order non-linear response functions. Best suited for small scale applications ($k < 10$).<br>- For optimization purposes, it requires relatively fewer calls of the analysis code. Convergence is achieved relatively faster. | Less efficiency is expected when applied to problems with highly non-linear and irregular performance. Higher order polynomials can be adopted. In this case, a large number of training points is needed to accurately capture the function behavior and estimate a large number of parameters. In addition, instabilities and false optima may appear |
| **KG** | - Well suited for numerical experiments with deterministic errors.<br>- Extremely flexible by virtue of a wide range of correlation functions.<br>- Even when a constant term is used for the global part of the model, the performance is comparable to second-order polynomial regression model.<br>- KG can be sufficiently applied to problems with up to 50 parameters.<br>- KG tends to yield more accurate approximations over a wide range of samples size and design | - Model construction can be time consuming for large problems. A number of complicated calculations are involved including matrices operations and optimization. Such costs may *overweight* the approximation benefits.<br>- If the training points are relatively "close" to each other, the correlation matrix can become singular.<br>- Additional points are needed to assess the model which imposes additional computational cost. |

Surrogate models are involved in a variety of aerodynamic design optimization applications. On the one hand, quadratic RSM was implemented in optimization of transonic airfoil [5], transonic rotor blades [6], subsonic flying wing [7], rocket engine injectors [8], high speed civil transport HSCT [9], unmanned aerial vehicles UAV [10], and high speed stand-off missiles HSSM [11]. On the second hand, Kriging models were involved in the optimization of transonic airfoils [12, 13], three-element airfoil [14, 15], three element wing [16], subsonic engine nacelle [17], small jet aircraft [18], and hypersonic vehicle with air-breathing engine [19]. On the other hand, a number of researchers utilized both models in their work and addressed the differences. For example, the design of a high speed civil transport (HSCT) [20] and rocket engine nozzle [21]. It is noted that all these literatures adopted the Gaussian form of correlation function; other correlation functions such as exponential or general exponential were not used in any work.


## II.3. Models Assessment

The fitness of the surrogate is assessed using a variety of techniques. For example, regression error is used with RMS [1], sample sensitivity and cross validation are used with Kriging [2]. However, when two different surrogates are compared, the fitness is generally assessed using new points other than the training ones. The surrogate performance in predicting these untried points is evaluated using the root mean square error [19], maximum absolute error [19], and $R^2$ error estimate [2, 19]:

$$RMS = \sqrt{\frac{1}{q}\sum_q (y_i - \hat{y})^2} \, , \tag{16}$$

$$MAE = Max_q \left( |y_i - \hat{y}_i| \right), \tag{17}$$

$$R^2 = 1 - \left[ \sum_q (y_i - \hat{y}_i)^2 \, \middle/ \sum_q (y_i - \bar{y})^2 \right] \tag{18}$$

where $\bar{y}$ is the mean of the exact response at the $q$ untried points $y_i$ and $\hat{y}_i$ is the corresponding surrogate response value. Smaller values of *RMS and MSE* and a higher value of $R^2$ indicate better surrogate accuracy.

## III. Case Study

Hypersonic vehicles experience high levels of drag and aerodynamic heating during flight. Many efforts are devoted to minimizing these effects for longer ranges, lower fuel consumption, and safer flights. One approach is to use aerospikes. An aerospike is a needle-like body mounted at the nose tip of the main body. In many cases, a larger disk, called the aerodisk, is fitted at the tip of the spike for better performance. It is believed that the efficiency of the spike in reducing drag and aerodynamic heating effects largely depends on the spike geometry. For some spike designs, undesirable flow oscillations are invoked causing a number of structure and communication problems. Hence, a proper selection of the optimum spike design would yield optimum aerodynamic performance.

In this work, we aim to find the optimum design of a hypersonic spiked body at hypersonic flight velocity (Mach **6** freestream) and zero incidence. The main body is assumed to be a hemisphere of a fixed diameter, $D$, the spike is a thin circular cylinder of a fixed radius, $\delta = 0.05D$ and a variable length, and the aerodisk is a hemispherical disk of a variable radius, Fig. 2.a. Hence, the design parameters are the spike length, $L$, and the aerodisk radius, $r$. Their scaled counterparts are termed $x_1$ and $x_2$, respectively, Fig. 2.b.
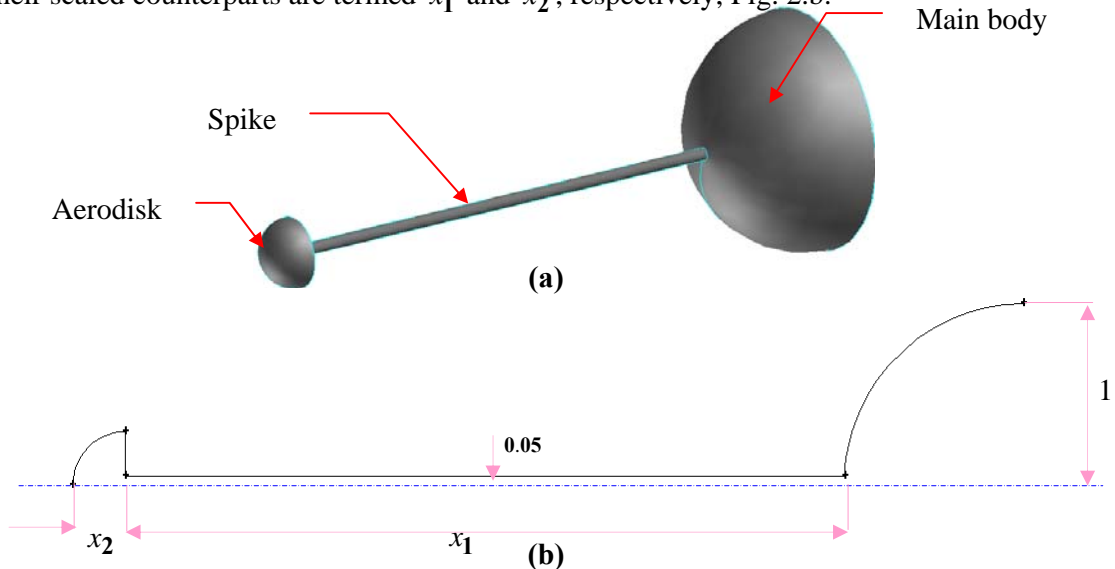


Fig. 2  **Geometry of a generic spiked hypersonic blunt vehicle**

Two numerical experiments are conducted in this work. The first experiment uses **40** design points in the training set and **20** other points for assessment purposes. The other experiment uses **100** training points and **60** other points for assessment purposes. All points are selected using the standard Latin hypercube sampling (*LHS*) technique. Fig. 3 illustrates the design space and the location of training and test points for the two experiments.



**(a)**                      **(b)**
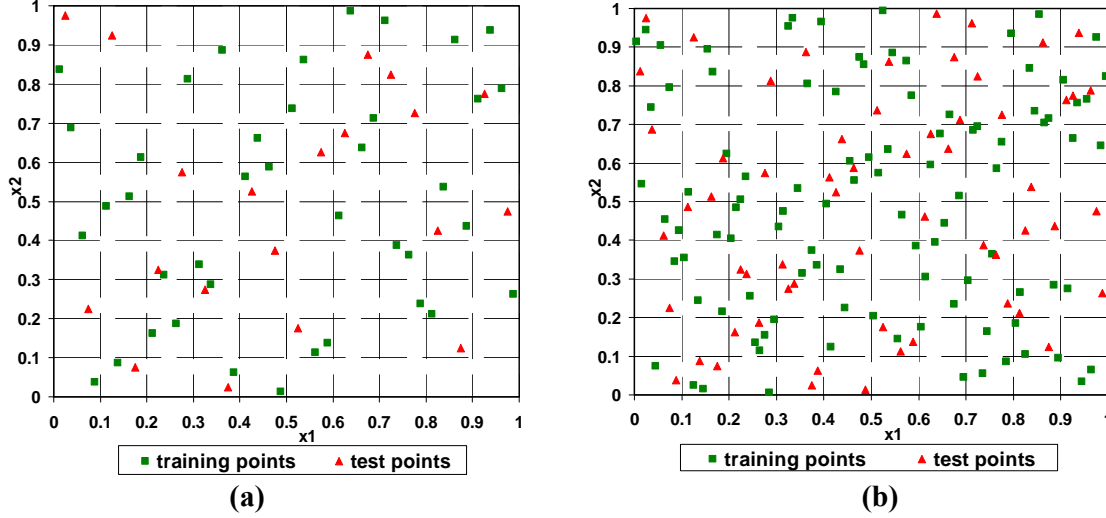
**Fig. 3 Latin Hypercube Sampling for (a) Experiment 1 and (b) Experiment 2**

Each point in the design space represents a unique design. All designs are evaluated using a high fidelity flow solver. In this work, we use FLUENT [22] which solves the Reynolds-averaged Navier-Stokes equations using the cell-centred finite volume approach. Prior to being used in this optimization work, the flow solver is validated by comparing its results with a number of previous experimental works and the resolution of the computational domain associated with the numerical solution is examined as well. Two aerodynamic performance criteria are of a special concern namely, the total drag coefficient $C_D$ and the maximum adiabatic wall temperature on the main body, $T$. These two values for each design are scaled with respect to a reference value. In this work, we use the respective values of unspiked hemispherical body as the reference values. The reference values of total drag coefficient (freestream conditions and hemisphere base area) and adiabatic stagnation temperature are **0.888** and **3243.3** *K*, respectively.

Our preliminary investigation of the aerodynamic performance of these designs reveals that, for designs having $L/D < 1.0$, flow instabilities arise along with deterioration in spike efficiency. Hence, the two experiments are repeated focussing only on the refined design space with coordinates: $0.1625 < x_1 < 1.0$ and $0 < x_2 < 1$. Four surrogate models are examined namely, quadratic *RSM* (*QRS*), exponential Kriging (*XKG*), Gaussian Kriging (*GKG*), and General exponential Kriging (*EKG*). A total of eight distinct experiments are performed. For easier referencing, they are given a special notation, Table 2.

**Table 2. Notation of the Experiments**

| Experiment | Notation | |
|---|---|---|
| | experiment 1 | experiment 2 |
| Temperature response, full design space | e1tf | e2tf |
| Temperature response, refined design space | e1tr | e2tr |
| Drag response, full design space | e1df | e2df |
| Drag response, refined design space | e1dr | e2dr |

The optimization technique used in this work is the genetic algorithm (*GA*). This population-based evolutionary algorithm has a number of appealing advantages over the conventional gradient-based techniques. Only the objective function value is used and its gradients are not needed and they are likely to find the global optimum even with highly nonlinear and multimodal phenomena. In our case, **200** individuals are evolving through infinite number of generations. Fitness of individuals is estimated based on the constructed surrogates. The best **20** individuals in each generation pass directly to the next one whereas the remaining individuals are generated via Gaussian mutation (with probability of **0.8** to enhance diversity of the search) and scattered crossover. Evolution process continues until convergence; a fitness tolerance of **1e-6** is taken as the convergence criteria. In this work, design objectives are treated separately. The genetic algorithm toolbox available in MATLAB [23] is used in this work.

## IV. Results and Discussions

### IV.1. Structure of the Surrogate Models
The three kriging surrogates used in the experiments have the same underlying spatial correlation function of the form:

$$d\left(x_i, x_j\right) = \sum_{l=1}^{k} \theta_l \left( \left| x_i^l - x_j^l \right|^p \right) = \theta_1 \left( x_i^1 - x_j^1 \right)^p + \theta_2 \left( x_i^2 - x_j^2 \right)^p \tag{19}$$

On the other hand, the polynomials used in these experiments have the following form:

$$\hat{y}(\overline{\mathbf{x}}) = \beta_o + \sum_{i=1}^{k} \beta_i \overline{x}_i + \sum_{i=1}^{k}\sum_{j \geq i}^{k} \beta_{ij} \overline{x}_i \overline{x}_j = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \beta_4 x_1^2 + \beta_5 x_2^2 \tag{20}$$

where $i$ and $j$ refer to any training point, $x$. The superscripts **1** and **2** refer to spike length and aerodisk radius with the corresponding weight parameters, $\theta_1$ and $\theta_2$, respectively. $p$ is an undetermined power parameter, and $\beta's$ are the unknown polynomial coefficients.

Temperature response
The values of the undetermined parameters for temperature experiments are listed in Tables 3 and 4; values of $\theta_1$, $\theta_2$, and $p$ are in bold, underlined, and italic numbers, respectively. The values in shaded cells refer to those of the refined space experiments.

**Table 3 Values of Quadratic Polynomial Coefficients
in Temperature Experiments**

|  | $\beta_1$ | $\beta_2$ | $\beta_3$ | $\beta_4$ | $\beta_5$ |
|---|---|---|---|---|---|
| *e1ft* | -.5008 | .1394 | -.0636 | .395 | -.0756 |
| *e1rt* | -.046 | .0335 | .0136 | .0167 | -.0268 |
| *e2ft* | -.7419 | .2322 | -.141 | .6187 | -.1203 |
| *e2rt* | -.0468 | .0251 | .0086 | .0213 | -.0171 |

**Table 4 Values of Kriging Parameters in Temperature Experiments**

|  | *e1ft* | | | *e2ft* | | |
|---|---|---|---|---|---|---|
|  | *e1rt* | | | *e2rt* | | |
| *GKG* | **20** | 18.34 | *2* | **13.4543** | 20 | *2* |
|  | **2.8284** | 0.8409 | *2* | **4** | 2.8284 | *2* |
| *XKG* | **4** | 0.3536 | *1* | **1** | 9.76e-4 | *1* |
|  | **0.1768** | 0.0743 | *1* | **0.1768** | 0.0743 | *1* |
| *EKG* | **1.3899** | 0.2455 | *0.5548* | **2.1435** | 0.134 | *0.7579* |
|  | **0.125** | 0.2874 | *1.7411* | **0.6598** | 0.483 | *1.8661* |

The values $\beta_1$ and $\beta_2$ are larger than those of $\beta_4$ and $\beta_5$, respectively. This indicates that the spike length has a dominant rule in temperature response compared to the aerodisk radius. Adding more training points emphasizes the spike length dominance whereas refining the space reduces it. We may infer that the predicted temperature response becomes less sensitive to spike length by excluding the undesirable designs that yield very high temperatures. In addition, $\beta_1$ and $\beta_4$ are larger than those of $\beta_2$ and $\beta_5$, respectively. This reflects that temperature response tends to be linear with respect to design parameters. Adding more training points emphasizes this trend. However, refining the space adds a degree of nonlinearity to temperature response especially with respect to spike length. The signs of $\beta_1$ and $\beta_2$ indicate that temperature tends to decrease with increasing spike length and decreasing aerodisk radius. These trends can be viewed in the graphical representation of *QRS*, Fig. 5; the green square and red triangular dots refer to training and test points, respectively.



**(a)**        **(b)**
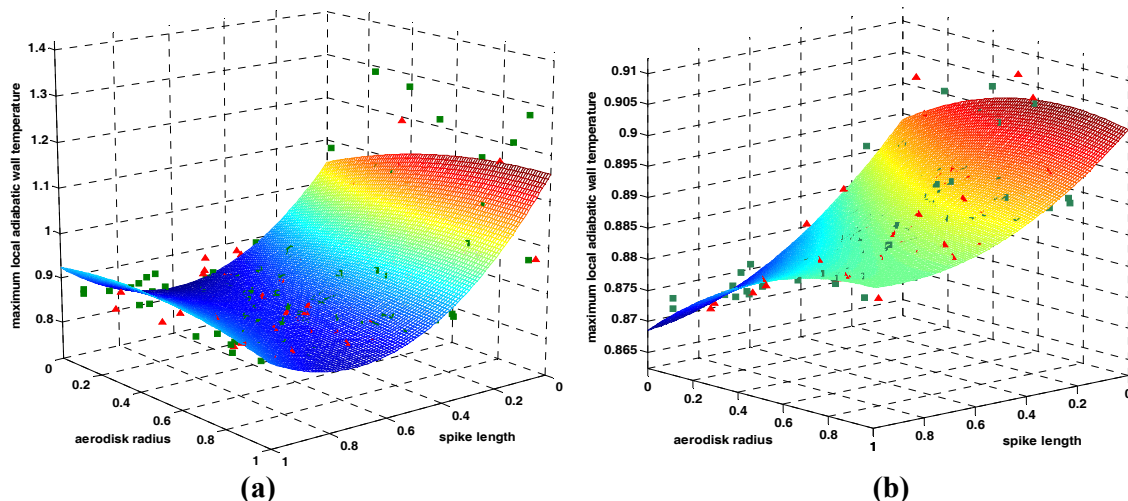
**Fig. 5 Graphical Representation of *QRS* for Experiments (a) *e2ft* and (b) *e2rt***

These nonlinear trends in temperature response can be confirmed from the values of $\theta$ in all Kriging surrogates. *GKG* yields the highest values for $\theta$ in *all* experiments which indicate high nonlinearity; this may be explained by the fact that *GKG* actually *imposes* a nonlinear behavior. However, this nonlinearity diminishes by adding more points and refining the design space, Fig. 6.



(a)                                                              (b)
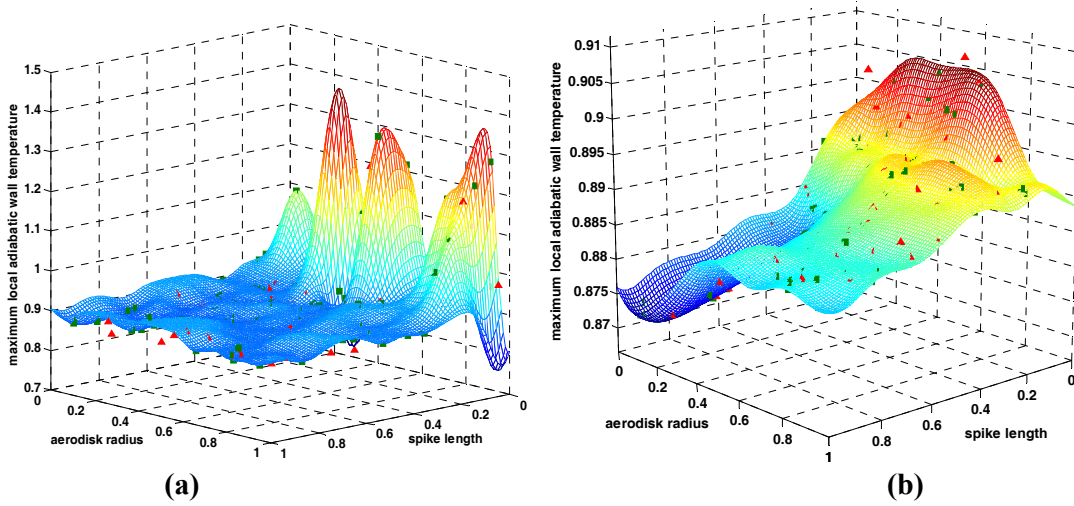
**Fig. 6 Graphical Representation of *GKG* for Experiment (a) *e2ft* and (b) *e2rt***

By refining the space, the value of $p$ for *EKG* approaches $2$. This may indicate that the temperature response becomes more *differentiable* by excluding the designs with undesirable output. Finally, comparing Figures 5 and 6, we can infer the different natures of RSM and KG. While the RSM surrogate fits into responses of training points (green dots), KG surrogate exactly interpolates them.

Drag response

Similarly, the values of the RSM and KG surrogates' undetermined parameters in drag experiments are listed in Table 5 and 6, respectively.

**Table 5  Values of quadratic polynomial coefficients in drag experiments**

|        | $\beta_1$ | $\beta_2$ | $\beta_3$ | $\beta_4$ | $\beta_5$ |
|--------|-----------|-----------|-----------|-----------|-----------|
| *e1fd* | -1.516    | -.1966    | -.3231    | 1.2603    | .4461     |
| *e1rd* | -.3617    | -.0508    | .0903     | .2041     | .0556     |
| *e2fd* | -1.4895   | -.0551    | -.0603    | 1.1348    | .1432     |
| *e2rd* | -.3439    | -.0996    | .0806     | .1978     | .1006     |

**Table 6  Values of kriging surrogates parameters in drag experiments**

| | e1fd | | | e2fd | | |
| | e1rd | | | e2rd | | |
|---|---|---|---|---|---|---|
| GKG | **1** | 2 | 2 | **8** | 2 | 2 |
| | **1.8616** | 0.2726 | 2 | **4** | 2.8284 | 2 |
| XKG | **0.25** | 0.0078 | 1 | **0.25** | 0.0156 | 1 |
| | **0.2292** | 0.0625 | 1 | **0.2293** | 0.0653 | 1 |
| EKG | **0.8409** | 0.483 | 1.8661 | **1.6818** | 0.483 | 1.8661 |
| | **0.5946** | 0.483 | 1.8661 | **0.732** | 0.483 | 1.8661 |

The dominating role of spike length is clear in drag response in both linear and nonlinear parts. This dominance seems to increase with more training points and diminish with refining the design space. Drag response is more nonlinear with respect to spike length. Increasing the number of training points emphasizes this trend whereas refining the design space reduces the degree of nonlinearity. *XKG* and *EKG* show very comparable values by refining the space or adding more training points; this may address their relative robustness in incorporating drag response.

## IV.2. Assessment of Surrogate Models

The values of temperature and drag responses at the test points are predicted by the surrogates. These values are then compared with the corresponding values estimated by the high fidelity CFD code. The values of *MSE*, *MAE*, and $R^2$ are shown in Table 7 and 8 for temperature and drag responses in bold, underlined, and italic numbers, respectively, for all surrogates. The shaded cells refer to the refined space values.

**Table 7 Temperature experiments' *MSE* , *MAE*, and $R^2$**

| | e1ft | | | e2ft | | |
| | e1rt | | | e2rt | | |
|---|---|---|---|---|---|---|
| QRS | **.0504** | .1877 | .5301 | **.0555** | .1753 | .3955 |
| | **.0022** | .0072 | .8968 | **.0018** | .0066 | .9376 |
| GKG | **.0714** | .2769 | .0576 | **.0513** | .2096 | .4832 |
| | **.0018** | .0048 | .9305 | **.0013** | .0052 | .9689 |
| XKG | **.0365** | .1496 | .7541 | **.0445** | .2195 | .601 |
| | **.0014** | .0044 | .9569 | **7.89e-4** | .0022 | .9883 |
| EKG | **.0384** | .1611 | .7274 | **.0271** | .094 | .8552 |
| | **.0012** | .0012 | .9681 | **4.97e-4** | .0014 | .9954 |

**Table 8 Drag experiments' *MSE* and *MAE*, and $R^2$**

|  | *e1fd* | | | *e2fd* | | |
|---|---|---|---|---|---|---|
|  | *e1rd* | | | *e2rd* | | |
| *QRS* | **.0848** | .2892 | .6773 | **.0674** | .2637 | .8249 |
|  | **.0212** | .08 | .8087 | **.0148** | .078 | .8796 |
| *GKG* | **.0653** | .2366 | .8086 | **.0183** | .0753 | .9871 |
|  | **.0116** | .0431 | .9426 | **.0118** | .0702 | .9233 |
| *XKG* | **.0326** | .1044 | .9523 | **.0147** | .0779 | .9917 |
|  | **.0167** | .0661 | .8814 | **7.89e-4** | .0388 | .9883 |
| *EKG* | **.0765** | .2946 | .7374 | **.0085** | .0276 | .9972 |
|  | **.0124** | .0493 | .9349 | **.0052** | .034 | .9854 |

Adding more training points in constructing the surrogate generally reduces the prediction error at untried points. However, a more significant improvement is attained by refining the design space especially in temperature response. The *XKG* and *EKG* surrogates yield relatively better prediction accuracy. Figures 7 and 8 illustrate graphical representation of these two surrogates in temperature and drag experiments, respectively.



**(a)**                                                    **(b)**
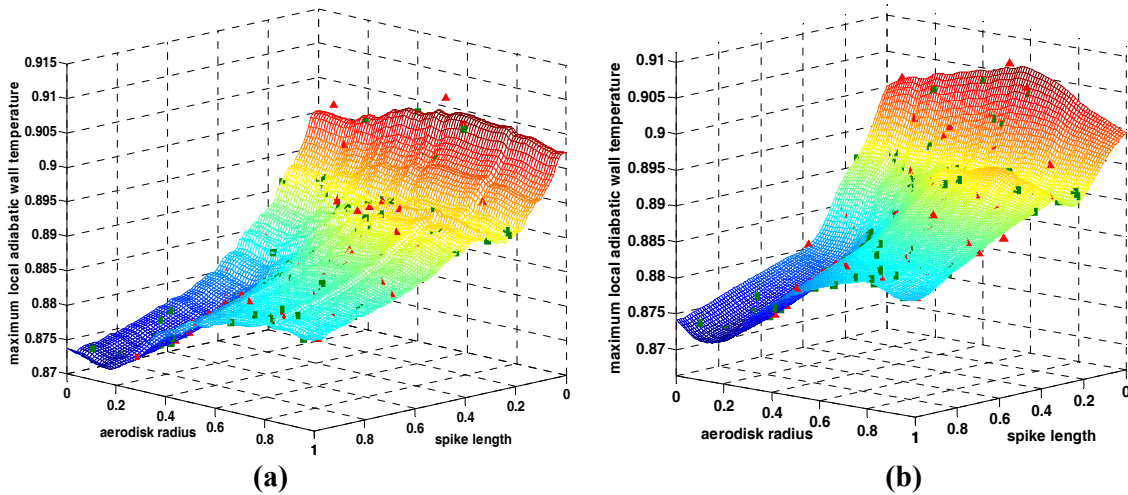
**Fig. 7 Graphical Representation of *e2rt* Experiment for (a) *XKG* and (b) *GKG***
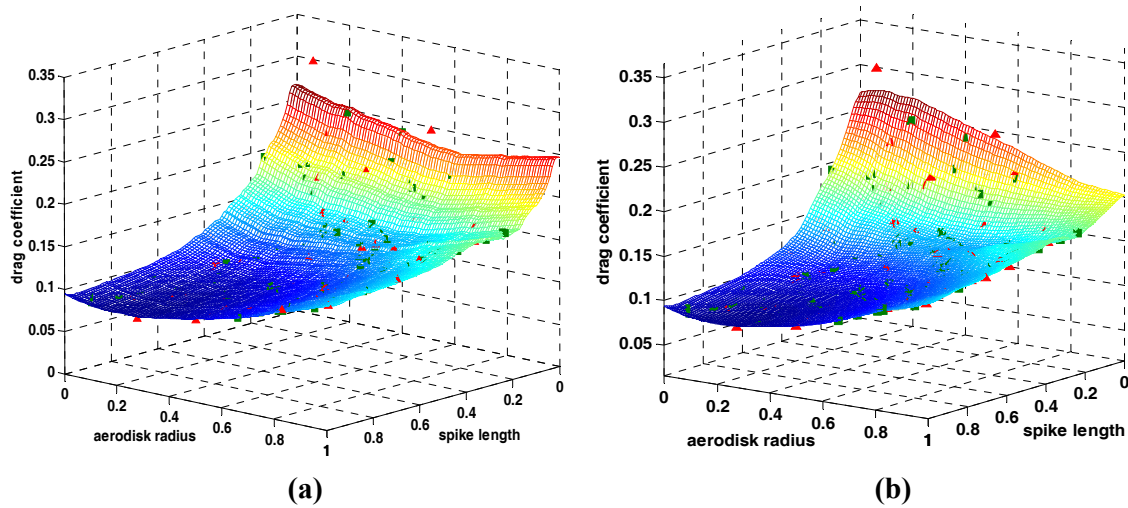
**Fig. 8 Graphical Representation of *e2rd* Experiment for (a) *XKG* and (b) *GKG***

### IV.3. Optimization Results

The optimization algorithm (here, the genetic algorithm) is applied to each of the resulting surrogates to find the predicted optimum design. As illustrated in Fig. 1, these candidate designs must be evaluated using the *CFD* solver to address their exact level of optimality. Table 9 illustrates the optimization results of the four surrogates studied here. The *%error* refers to the difference between the predicted and exact response; the plus and minus signs indicate that the surrogate over-estimated and under-estimated the response, respectively. On the other hand, the *%improvement* indicates the superiority of the exact response of the optimized design compared with the best design in the training set of each experiment; a minus sign indicates that the response of the surrogate-predicted design is worse than the best training design. The bold numbers refer to *exp.1* and the italic numbers refer to *exp.2, exp.3*, and *exp.4*. The shaded cells contain the values corresponding to the refined design space.

**Table 9. Optimization Results**

| | Temperature Response | | | | Drag Response | | | |
|---|---|---|---|---|---|---|---|---|
| | % error | | % improvement | | % error | | % improvement | |
| *QRS* | **-4.8403** | *-5.058* | **-1.542** | *0.4397* | **-74.614** | *-48.18* | **-22.76** | *-13.866* |
| | **-2.6596** | *-0.211* | **-1.69** | *0.2195* | **-18.779** | *1.0808* | **-3.711** | *-3.357* |
| *GKG* | **-8.6097** | *-26.888* | **-3.45** | *-15.169* | **-35.482** | *-8.647* | **-22.1** | *-1.48* |
| | **-0.004** | *-0.0155* | **-0.137** | *0.0548* | **-2.451** | *-2.219* | **-0.4346** | *-1.19* |
| *XKG* | **-0.1036** | *-0.1457* | **-0.3754** | *-0.1328* | **-10.027** | *-3.996* | **-7.5796** | *0.06* |
| | **-0.2102** | *0.0225* | **-0.1512** | *0.0986* | **0.0369** | *-0.159* | **0.0368** | *-0.706* |
| *EKG* | **-4.8919** | *-0.0006* | **-4.037** | *-0.04* | **-8.6589** | *-3.282* | **-1.3756** | *-0.96* |
| | **-0.0843** | *0.0299* | **0.0959** | *0.1015* | **-2.0567** | *-0.2078* | **-0.6116** | *-0.295* |

Generally, adding more training points and refining the design space enhance the accuracy of surrogates. For the case in concern, *QRS* yields undesirable performance especially in drag response considering both error and improvement estimates. On the other hand, *XKG* and *EKG* produce better results which indicate that they incorporate the physical phenomena and the associated design-response functional relations. The minor improvements achieved by the optimized designs may be justified by the existence of near-optimum designs within the randomly-selected training set.

# V. Conclusions

Surrogate models represent a cheap and reliable alternative for expensive high fidelity computational models. In addition to this remarkable advantage, surrogates provide a clear understanding of the design-response functional relations. Many information about the importance of the design parameters and the nature of response can be deduced by investigating the surrogate structure. Consequently, parameters of minor effect can be overlooked for more economic and realistic design and optimization work. Surrogate performance can be improved by incorporating more training points. In addition, refining the design space by eliminating the designs that yield undesirable output enhances the surrogate performance. Thus, a good understanding of the phenomena in concern is recommended to construct a better surrogate. For the case studied here, kriging surrogates based on exponential correlation function give relatively better prediction at the untried points. When these surrogates are coupled with the appropriate search technique, better optimum designs are attained.

# References

[1] Myers, R. H. and Montgomery, D. C., "Response Surface Methodology: Process and Product Optimization Using Designed Experiments," John Wiley and Sons, New York, 2002.

[2] Keane, A. J. and Nair, P. B., "Computational Approaches for Aerospace Design, the Pursuit of Excellence," John Wiley and Sons, Chichester, 2005.

[3] Simpson, T. W., Peplinski, J. D., Koch, P. N., and Allen, J. K., "Metamodels for Computer-based Engineering Design: Survey and recommendations," *Engineering with Computers*, Vol. 17, 2001, pp. 129- 150.

[4] Fang, H., M. Rais-Rohani, M., Liu R.-Z., and Horstemeyer, M. F., "A comparative study of metamodeling methods for multiobjective crashworthiness optimization," Computers and Structures, Vol. 83, 2005, pp. 2121–2136.

[5] Vavalle, A. and Qin, N., "Iterative Response Surface Based Optimization Scheme for Transonic Airfoil Design," *Journal of Aircraft*, Vol. 44, No. 2, 2007, 365- 376.

[6] Lian, Y. and Liou, M., "Multiobjective Optimization Using Coupled Response Surface Model and Evolutionary Algorithm," *AIAA Journal*, Vol. 43, No. 6, 2005, pp. 1316-1325.

[7] Sevant, N. E., Bloor, M. I. G., and Wilson, M. J., "Aerodynamic Design of a Flying Wing Using Response Surface Methodology," *Journal of Aircraft*, Vol. 37, No. 4, 2000, pp. 562- 569.

[8] Goel T., Vaidyanathan, R., Haftka, R. T., Shyy, Y., Queipo, N. V., and Tucker, K., "Response surface approximation of Pareto optimal front in multi-objective optimization," *Comput. Methods Appl. Mech. Engrg.*, Vol. 196, 2007, pp. 879– 893.

[9] Daberkow, D. D. and Dimitri N. Mavris, D. N., "New Approaches to Conceptual and Preliminary Aircraft Design: A Comparative Assessment of a Neural Network Formulation and a Response Surface Methodology," 1998 World Aviation Conference September 28-30, 1998, Anaheim, CA.

[10] Sobieski, I. P. and Kroo, I. M., "Collaborative Optimization Using Response Surface Estimation," *AIAA Journal*, Vol. 38, No. 10, 2000, pp. 1931- 1938.

[11] Sobieski, I. P. and Kroo, I. M., "Collaborative Optimization Using Response Surface Estimation," *AIAA Journal*, Vol. 38, No. 10, 2000, pp. 1931- 1938.

[12] Jeong, S., Murayama, M., and Yamamoto, K., "Efficient Optimization Design Method Using Kriging Model," *Journal of Aircraft*, Vol. 42, No. 2, 2005, 413- 420.

[13]     Keane, A. J., "Statistical Improvement Criteria for Use in Multiobjective Design Optimization," *AIAA Journal*, Vol. 44, No. 4, 2006, pp. 879- 891.

[14]     Kanazaki, M., Tanaka, K., Jeong, S., and Yamamoto, K., "Multi-objective Aerodynamic Optimization of Elements' Setting for High-lift Airfoil Using Kriging Model," 44th AIAA Aerospace Sciences Meeting and Exhibit, January 2006, Reno, Nevada.

[15]     Kanazaki, M., Tanaka, K., Jeong, S., and Yamamoto, K., "Multi-Objective Aerodynamic Exploration of Elements' Setting for High-Lift Airfoil Using Kriging Model," *Journal of Aircraft*, Vol. 44, No. 3, 2007, pp. 858- 864.

[16]     Kanazaki, M., Imamura, T., Jeong, S., and Yamamoto, K., "High-lift Wing Design in Consideration of Sweep Angle Effect Using Kriging Model," 46th AIAA Aerospace Sciences Meeting and Exhibit 7 - 10 January 2008, Reno, Nevada.

[17]     Song, W. and Keane, A. J., "Surrogate-Based Aerodynamic Shape Optimization of a Civil Aircraft Engine Nacelle," *AIAA Journal*, Vol. 45, No. 10, 2007, pp.2565- 2574.

[18]     Kumano, T., Jeong, S., Obayashi, S., Ito, Y., Hatanaka, K., and Morino, H., "Multidisciplinary Design Optimization of Wing Shape for a Small Jet Aircraft Using Kriging Model," 44th AIAA Aerospace Sciences Meeting and Exhibit, January 2006, Reno, Nevada.

[19]     Tsuchiya, T., Takenaka, Y., and Taguchi, H., "Multidisciplinary Design Optimization for Hypersonic Experimental Vehicle," *AIAA Journal*, Vol. 45, No. 7, 2007, pp. 1655- 1662.

[20]     Giunta, A. A., "Aircraft Multidisciplinary Design Optimization Using Design of Experiments Theory and Response Surface Modelling Methods", Ph.D. Dissertation, Faculty of Virginia Polytechnic Inst. And State Univ., Blacksburg, VA., 1997.

[21]     Simpson, T. W., Mauery, T. M., Korte, J. J., and Mistree, F., "Comparison of Response Surface and Kriging Models for Multidisciplinary Design Optimization," *AIAA Paper AIAA-98-4755*, 1998.

[22]     *FLUENT 6.3 User's Guide*, Fluent Inc., Lebanon, NH 03766, USA.

[23]     *MATLAB 7 release 14 User's Guide*, The Mathworks Inc., U.S.A., 2005.