AIAA 2002–0317

Using Gradients to Construct Cokriging Approximation Models for High-Dimensional Design Optimization Problems

Hyoung-Seog Chung and Juan J. Alonso
Stanford University, Stanford, CA 94305

40th AIAA Aerospace Sciences Meeting and Exhibit
January 14–17, 2002/Reno, NV

Using Gradients to Construct Cokriging Approximation Models for High-Dimensional Design Optimization Problems

Hyoung-Seog Chung* and Juan J. Alonso†
Stanford University, Stanford, CA 94305

The Kriging method is an interpolation scheme that can be used for modeling deterministic computer analyses as the realization of a stochastic process. The technique has been recognized as an alternative to the traditional Response Surface method in generating approximation models of computationally expensive CFD analyses. This is due to its ability to interpolate sample data and to model a function with multiple local extrema. To fully exploit the advantage of the Kriging method, however, a large number of sample data points should be spread out to fill the design space. This can be very costly and even impractical in high-dimensional design optimization. In this work, the Cokriging method, an extension of Kriging, which can incorporate secondary information such as values of gradients in addition to primary function values of the sample points has been utilized for constructing approximation models in a realistic design optimization process. This approach improves on the accuracy and efficiency of using the Kriging method for high-dimensional design problems. Provided that gradient information is available through inexpensive algorithms such as the adjoint method, Cokriging significantly reduces the large computational cost needed for the original Kriging method to accurately capture multiple local extrema of the unknown response function within a relatively large design space. After validating the feasibility of the Cokriging method using simple one- and two-dimensional analytic functions, the approach is applied to the aerodynamic design of a supersonic business jet. The results of these 2- and 5-variable test design problems indicate that great improvements on the efficiency and applicability of the Kriging method in high-dimensional design optimization problems can be achieved.

Nomenclature

β	constant underlying global portion of Kriging model
C_D	drag coefficient
$\mathbf{f}$	constant vector used in Kriging model
$\mathbf{f_c}$	constant vector used in Cokriging model
k	number of design variables
n_s	number of sample points
$\mathbf{r}$	vector of correlation values for Kriging model
$\mathbf{r_c}$	vector of correlation values for Cokriging model
$R(.)$	correlation function for Kriging model
$\mathbf{R}$	correlation matrix for Kriging model
$\mathbf{R_c}$	correlation matrix for Cokriging model
x	scalar component of $\mathbf{x}$
$\mathbf{x}$	vector denoting all locations (sites) in the design space
$\mathbf{x^p}$	vector denoting the p^{th} location in the design space
$y(.)$	unknown function
$\hat{y}(.)$	estimated model of y(.)
$\boldsymbol{\theta}$	vector of correlation parameters for Kriging model
$\hat{\sigma}^2$	estimated sample variance

*Doctoral Candidate, AIAA Member
†Assistant Professor, Department of Aeronautics and Astronautics , AIAA Member

Introduction

THE optimization of aerospace systems is an iterative process that requires computational models embodied in complex and expensive analysis software. This paradigm is well exemplified by the field of Multidisciplinary Design Optimization (MDO) which attempts to exploit the synergism of mutually interacting disciplines in order to improve the performance of a given design, while increasing the level of confidence that the designer places on the outcome of the design itself. MDO methods greatly increase the computational burden and complexity of the design process.[1,2] For this reason, high-fidelity analysis software typically used in single discipline designs may not be suitable for direct use in MDO.[3] Faced with these problems, the alternative of using approximation models of the actual analysis software has received increased attention in recent years. A second advantage of using approximation models during the optimization process is that they can be used with optimization algorithms which do not rely on the computation of sensitivity derivatives.

One of the most common methods for building an approximate model is the response surface method (RSM) in which a polynomial function of varying order (usually a quadratic function) is fitted to a number of sample data points using least squares regression. This method has achieved popularity since it provides an explicit functional representation of the sampled data, and is both computationally inexpensive to run and easy to use. However, response surface models have several key limitations: their accuracy is only guaranteed within a small trust region, and, by design, they are unable to predict multiple extrema. In addition, these methods were originally developed to model data resulting from physical experiments which had a random error distribution. Since the nature of computer experiments is such that random errors are not present (a bias is much more common), the use of these methods for modeling deterministic data has resulted in serious debate within the statistical community.[4] In order to overcome these problems, Sacks, et al.[5] proposed an interpolation modeling technique, known as the Kriging method, developed in the fields of spatial statistics and geostatistics, in order to approximate the results of deterministic computer analyses. The Kriging method is different from the RSM since the interpolation of the sampled data is carried out using a maximum likelihood estimation procedure,[6] which allows for the capturing of multiple local extrema. However, the Kriging method has higher computational cost and is harder to implement. In addition, it does not provide an explicit model function. The Kriging method also suffers from accuracy limitations which are a function of the method used for the selection of the sample points and the total number of these points.

To fully exploit the advantage of capturing multiple local extrema of the unknown response function using a Kriging model, a large number of sample data points should be spread out to fill the design space. This sampling process can be very costly and even impractical in high-dimensional design optimization. In contrast, the Cokriging method is an extension of the Kriging method, which incorporates secondary information such as gradients in addition to primary function values when generating approximation models for unknown functions. Provided that gradient information is available through inexpensive algorithms such as the adjoint method, the Cokriging method can utilize this relatively cheap gradient information in lieu of the computationally expensive functional evaluations (CFD analyses in our case). The approach can approximate the original function with a much smaller number of samples, without significantly sacrificing the accuracy of the approximation; therefore, the method significantly reduces the large computational cost needed for the original Kriging method to accurately capture multiple local extrema of the unknown response function within a relatively large design space.

In this paper, and after an initial validation of the Cokriging method was conducted using simple one- and two-dimensional analytic functions, the technique was applied to a test case of the aerodynamic design of a low boom supersonic business jet. A three-dimensional Euler flow solver and an automatic mesh generation capability were used in the aerodynamic design of a supersonic business jet using a variety of geometric design parameters. This study identifies the efficiency, applicability, and limitations of the proposed method within a typical MDO process.

Overview of Kriging Method

Original Kriging Method

The Kriging technique uses a two component model that can be expressed mathematically as

$$y(\boldsymbol{x}) = f(\boldsymbol{x}) + Z(\boldsymbol{x}), \tag{1}$$

where $f(\boldsymbol{x})$ represents a global model and $Z(\boldsymbol{x})$ is the realization of a stationary Gaussian random function that creates a localized deviation from the global model.[7] $f(x)$ can be considered to be an underlying constant, β ,[6] and then equation (1) becomes

$$y(\boldsymbol{x}) = \beta + Z(\boldsymbol{x}), \tag{2}$$

which is used in this paper. The estimated model of equation (2) is given as

$$\hat{y} = \hat{\beta} + \boldsymbol{r^T(x)R^{-1}(y - f\hat{\beta})}, \tag{3}$$

where $\mathbf{y}$ is the column vector of response data and f is a column vector of length n_s which is filled with ones. $\mathbf{R}$ in equation (3) is the correlation matrix which can be obtained by computing $R(x^i, x^j)$, the correlation function between any two sampled data points. This correlation function is specified by the user. In this work, the authors use a Gaussian exponential correlation function of the form provided by Giunta, et al.[4]

$$R(\boldsymbol{x^i}, \boldsymbol{x^j}) = \exp[-\sum_{k=1}^{n} \boldsymbol{\theta_k |x_k^i - x_k^j|^2}]. \tag{4}$$

The correlation vector between x and the sampled data points is expressed as

$$\boldsymbol{r^T(x) = [R(x, x^1), R(x, x^2), ..., R(x, x^n)]^T}. \tag{5}$$

The value for $\hat{\beta}$ is estimated using the generalized least squares method as

$$\hat{\beta} = \boldsymbol{(f^T R^{-1} f)^{-1} f^T R^{-1} y}. \tag{6}$$

Since R is a function of the unknown variable θ, $\hat{\beta}$ is also a function of θ. Once θ is obtained, equation (3) is completely defined. The value of θ is obtained by maximizing the following function over the interval $\theta > \mathbf{0}$

$$-\frac{[n_s \ln(\hat{\sigma}^2) + \ln|\boldsymbol{R}|]}{2}, \tag{7}$$

where

$$\hat{\sigma}^2 = \frac{(\boldsymbol{y} - \boldsymbol{f}\hat{\boldsymbol{\beta}})^T \boldsymbol{R}^{-1}(\boldsymbol{y} - \boldsymbol{f}\hat{\boldsymbol{\beta}})}{n_s}. \tag{8}$$

This requires the solution of a k-dimensional unconstrained non-linear optimization problem, which can be reduced to a one-dimensional problem by assuming the same correlation parameter for each component of the sample points. Thus, equation (4) can be simplified as

$$R(\boldsymbol{x^i}, \boldsymbol{x^j}) = \exp[-\theta \sum_{k=1}^{n} |\boldsymbol{x_k^i} - \boldsymbol{x_k^j}|^2]. \tag{9}$$

Direct Cokriging Method

The Kriging method estimates are derived using only the values of a single sample function. However, it is always possible that there may be additional information available at the sample points. These additional variable such as function gradients are usually cross-correlated with the primary variable and thus contain useful information about the primary variable. The Cokriging method can estimate the unknown primary function of interest more effectively by utilizing these secondary function values.[8] In this study, the Cokriging method which exploits gradient information at the sample points as the additional information was investigated.

For the original Kriging method, the covariance matrix of $Z(\boldsymbol{x})$ is defined as

$$Cov\left[y(\boldsymbol{x}^i), y(\boldsymbol{x}^j)\right] = \sigma^2 \mathbf{R}\left[R\left(\boldsymbol{x^i}, \boldsymbol{x^j}\right)\right], \tag{10}$$

where $\mathbf{R}$ is the correlation matrix and $R\left(\boldsymbol{x^i}, \boldsymbol{x^j}\right)$ is the correlation function given by equation (4). The correlation matrix $\mathbf{R_c}$ and the correlation vector r_c for the Cokriging method are evaluated not only using function values at the sample point, but also with their gradient. Then, the covariance of the Cokriging method can be modified as follows

$$Cov\left[y(\boldsymbol{x}^i), y(\boldsymbol{x}^j)\right] = \sigma^2 \mathbf{R}\left[R\left(\boldsymbol{x^i}, \boldsymbol{x^j}\right)\right], \tag{11}$$

$$Cov\left[y(\boldsymbol{x}^i), \frac{\partial y(\boldsymbol{x}^j)}{\partial x_k}\right] = \sigma^2 \frac{\partial R(\boldsymbol{x}^i, \boldsymbol{x}^j)}{\partial x_k}, \tag{12}$$

$$Cov\left[\frac{\partial y(\boldsymbol{x}^i)}{\partial x_k}, y(\boldsymbol{x}^j)\right] = -\sigma^2 \frac{\partial R(\boldsymbol{x}^i, \boldsymbol{x}^j)}{\partial x_k}, \tag{13}$$

$$Cov\left[\frac{\partial y(\boldsymbol{x}^i)}{\partial x_k}, \frac{\partial y(\boldsymbol{x}^j)}{\partial x_l}\right] = -\sigma^2 \frac{\partial R(\boldsymbol{x}^i, \boldsymbol{x}^j)}{\partial x_k \partial x_l}. \tag{14}$$

Accordingly, the Cokriging model can be obtained by modifying equation (2) to yield

$$\hat{y_c} = \hat{\beta_c} + \boldsymbol{r_c}^{\boldsymbol{T}}(\boldsymbol{x})\boldsymbol{R_c}^{-1}(\boldsymbol{y_c} - \boldsymbol{f_c}\hat{\boldsymbol{\beta}}_{\boldsymbol{c}}), \tag{15}$$

and

$$\hat{\beta_c} = (\boldsymbol{f_c}^{\boldsymbol{T}}\boldsymbol{R_c}^{-1}\boldsymbol{f_c})^{-1}\boldsymbol{f_c}^{\boldsymbol{T}}\boldsymbol{R_c}^{-1}\boldsymbol{y_c}. \tag{16}$$

where

$$\boldsymbol{y_c} = \left[y(\boldsymbol{x}^1), ..., y(\boldsymbol{x}^{n_s}), \frac{\partial y(\boldsymbol{x}^1)}{\partial \boldsymbol{x}_1}, \frac{\partial y(\boldsymbol{x}^1)}{\partial \boldsymbol{x_2}}, ... \frac{\partial y(\boldsymbol{x}^{n_s})}{\partial \boldsymbol{x}_k}\right],$$

$$\boldsymbol{f_c} = [1, 1, ..., 1, 0, 0, ..., 0]$$

where $\boldsymbol{f_c}$ contains n_s ones and $n_s \times k$ zeros. The reader is referred to Ref[5,7,8] for more details on the development of the Kriging and Cokriging techniques.

Indirect Cokriging

Another way to include gradient information in the Kriging method is to estimate additional function values using the gradients available in a close neighborhood of the sample point. In this approach, the original Kriging formulation can be used with an increased number of sample data located in the proximity of the original sample points. These additional function values serve as if they were gradients because they tend to have strong correlations with the original sample points given the close distances to each other. The authors have given this approach the name Indirect Cokriging only to distinguish it from the Cokriging (Direct Cokriging) method which uses the gradient information directly. Both of these approaches are used in the results presented in this paper.

3. Test Problem : Supersonic Business Jet(SBJ) Design

The design problem in question involves the drag minimization of a supersonic business jet wing-body configuration at a specified lift coefficient. The aircraft geometry and flow conditions were parameterized with a total of 22 potential design variables.

For the initial test case, two geometric design variables were chosen so that the actual objective function and the approximation function modeled by the Cokriging method could be graphically compared and validated. The selection of the wing location along the fuselage and the fuselage radius at the mid-section of the fuselage as design variables produced a well-shaped objective function (C_D) having multiple local extrema. This design space was used as a realistic test function for the visualization of Cokriging models. The test problem was extended to 5 design variables to investigate the Cokriging method further in a slightly more realistic design case: we have been using 8-10 design variable test cases for the design of low-boom supersonic aircraft configurations. The chosen design variables represent the following geometric parameters:

x_1 = wing sweep angle at quarter chord line
x_2 = wing aspect ratio
x_3 = wing position along fuselage
x_4 = fuselage radius at 0.5 $\times$ fuselage length
x_5 = fuselage radius at 0.6 $\times$ fuselage length

Figure 1 shows their graphical interpretation. The airfoil shape for all wing stations was selected as a simple biconvex airfoil of varying thickness. Once the design variables were selected, an automatic mesh generation procedure that was able to handle the geometry variations imposed by the changes in the design variables was utilized to automatically create different sets of meshes needed for the CFD calculations. The three-dimensional Euler solver, FLO87, developed by Jameson [9,10] was used to calculate aerodynamic coefficients at sample design points chosen by incrementing each variable from its baseline value using a design of experiments approach. The free-stream flow conditions were fixed at $M_\infty = 1.5$ and the coefficient of lift, based on the wing planform area, was fixed at $C_L = 0.1$. Kriging and Cokriging models were built using drag output data from the Euler solver, and were incorporated into a nonlinear optimization process using Matlab's **fminsearch** to perform realistic design optimization calculations.

4. Design Tools

4.1 Grid Generator

A grid generator called CH-GRID developed by Reuther et al. was used for mesh generation for supersonic business jet wing-body configurations. CH-GRID is a stand-alone form of the C-H type grid generator for the adjoint-based, single-block wing-body design code, SYN87-SB. Figure 2 shows a typical wing-body mesh. A geometry generation engine was constructed for this work so that the input geometry for CH-GRID could be easily created using an input file that only contains the values of the design variables described above.

4.2 Flow Solver

The CFD flow solver must meet fundamental requirements of accuracy, efficiency, robustness, and fast convergence to be used in high-dimensional design optimization problems. The accuracy is important because the approximation model accuracy and the improvement predicted by the optimization process using these models can only be as good as the accuracy of the flow analysis itself. Efficiency is also required when the number of design variables increases and the required number of sample evaluations for constructing the approximation models increases accordingly. The robustness of the solver, i.e., its ability to obtain a flow solution for a variety of configuration shapes and flow conditions, is particularly critical for the construction of sample databases for the approximation models, in which large variations of the values of the design variables are permitted. In addition, the benefit of aerodynamic optimization lies in obtaining the last few percentage points in aerodynamic efficiency. In such cases, the solution must be highly converged such that the noise in the figure of merit is well below the level of realizable improvement. [11]

Jameson's FLO87 code used in this study easily met all of the previously mentioned criteria. FLO87 solves the steady three-dimensional Euler equations using a modified explicit multistage Runge-Kutta time stepping scheme. FLO87 achieves fast convergence with the aid of multigridding and implicit residual smoothing. Also, a driver program called RS87 was developed to utilize multiprocessor computers for analyzing a number of different configurations simultaneously.

The necessary gradient information needed for the construction of the Cokriging models for this work was obtained using the finite difference method since the preliminary design problems were set up with a relatively small number of design variables. However, the finite difference method results in a very high computational burden in high-dimensional problems since k additional CFD evaluations are needed at each sample point to get directional gradients along k design variables. Therefore, the use of gradient information based on finite differencing to replace function evaluations does not result in improvements in computational performance. However, recently developed adjoint-based methods[12] can be used to compute gradient information of systems governed by PDEs with low computational cost. For the example of a 7-dimensional problem, the complete gradient information at a given design point can be obtained with the equivalent of two function evaluations (one for the flow solution and the other for the adjoint calculation), whereas the finite difference approach requires 8 or N+1 (N being the number of design variables) function evaluations. At a given design point, a flow solver together with an adjoint method provides one function value and seven gradients in the direction of each design variable with the cost of *only* two function evaluations. Six flow calculations can be saved with the use of these seven gradients when constructing a Cokriging model. The main advantage of using the Cokriging method is to alleviate the high computational burden of constructing the original Kriging models for high-dimensional problems by using the gradient information calculated through a relatively inexpensive method instead of using function values. Details of the adjoint methodology have been previously published in the literature and are mathematically involved. The reader is referred to[11,13,14] for more details.

4.3 Optimization Procedure

Drag minimization of wing-body configurations was performed using the **fminsearch** function in Matlab. **fminsearch** is a unconstrained optimization routine based on the Nelder Mead direct search method with line searches that uses only function values. The bounds for the values of the design variables were imposed by adding a quadratic penalty to the objective function. The routine worked fairly well for the test design problems mentioned above. The optimization problem can then be written as

$$\min_{x \in R^m} \; C_D(x), \tag{17}$$

subject to

$$x_{min} \leq x \leq x_{max},$$

5. Results and Discussion

5.1 Graphical Validation of Cokriging Method using Simple Analytic Test Functions

The Cokriging technique was applied to approximate simple one- and two-dimensional analytic functions. The results were compared with those of the original Kriging method for validation purposes only. Figure 3 shows the results of the one-dimensional validation case. The original Kriging model was fitted to three sample data points while the direct Cokriging model was constructed with three function values and three gradients at each sample point. The indirect Cokriging model was constructed with three function values and three additional sample values estimated from the gradient information. As shown in Figure 3 (a) the two Cokriging models approximated the test function much better than the original Kriging model. The inaccuracy of the original model clearly resulted from under-sampling and a lack of correlation between the sample points. The same test was repeated with 5 sample points instead, and the results are shown in Figure 3 (b). Surprisingly, both Cokriging models were nearly identical to the analytic test function across the entire domain of interest, while the original Kriging model still showed poor performance in approximating the function.

For two-dimensional graphical validation purposes, an analytic test function called the *peaks* function was chosen from MATLAB User's Guide[15] as shown in Figure 4 (a). Figure 4 shows the results with 5 sample points while Figure 5 uses 9 sample points. Once again, the original Kriging approximation entirely missed the key features of the test function due to the under-sampling effect. The best approximation that the original Kriging method could achieve with only 5 samples was a nearly planar shape as shown in Figure 4 (b). The Cokriging models were also generated using 5 sample points and their gradients in each direction, which make up for a total of 15 pieces of information (5 function values and 10 gradients) available for surface fitting. The accuracy of the approximation models was greatly improved by adding the gradient information. Both Cokriging models could capture the general features of the test function. Even though the peak values and their locations were not that accurate, the approximation models were fairly accurate within the domain of sample locations (Figure 4 (c), (d)).

The results of the 2-D test with 9 sample points can be seen in Figure (5). While the original Kriging model still did a poor job of predicting the analytic test function, both Cokriging models were able to capture accurately not only the general features of the test function but also the function values and locations of the peaks. Note that there is little difference between the direct Kriging and indirect Kriging results, which implies that both approaches of implementing gradient information are practically identical.

These graphical examples clearly demonstrate the applicability and efficiency of the Cokriging technique for simple 1- and 2-dimensional function approximations. Although in these simple test cases, the computational savings are minimal, when the number of design variables grows, the use of the Cokriging method would become compelling assuming that there exists an accurate and cost effective way to obtain gradient information such as the adjoint method.

5.2 Cokriging Modeling Test on 2-Dimensional SBJ Design Problem

To further validate and investigate their ability to approximate the results of the original CFD code, 2-dimensional Cokriging models were created for the C_D of the supersonic business jet test problem, using sample data obtained from CFD analyses. The two design parameters of the wing-body configuration chosen as design variables were the wing x-location along fuselage(x_1) and the radius at the fuselage mid-point(x_2). 400 CFD calculations were performed varying the design variables to obtain a graphical representation of the actual CFD analyses and the results are shown in Figure 6 (a). The design variables were chosen so as to generate a realistic test function having multiple local extrema to check the ability of Cokriging models to simulate that feature. One important point to note from the figure is that the actual aerodynamic coefficient (C_D in this case) varies smoothly with respect to the geometric design variables. The selection of Gaussian exponential correlation function for the Kriging method was based on the assumption that the response function to be modeled was very smooth in nature. Thus, Figure 6 (a) provides the validation of the assumption and the rationale for using the Gaussian correlation function in this problem.

For the original Kriging method, CFD results at 9 sample points were used to construct the model and for the Cokriging models, 9 sample points together with gradient information at these points were used. The results are compared in Figure 6. As in the case of the analytic test cases, Cokriging models performed much better than Kriging model in predicting the objective function in terms of its general shape and magnitude. Again, very little difference was found between the direct and indirect Cokriging methods and therefore, we feel comfortable that we can use either of the two Cokriging approaches with similar accuracy. The indirect Cokriging method has the advantage of having a simpler formulation, but there is a chance for numerical errors to be introduced while estimating the additional function values from the gradients. The direct method can be more accurate because it uses the gradient information directly, but the formulation becomes more complex as the dimensionality of the problem increases.

It is found that the accuracy of Kriging models largely depends on the values of the correlation parameters defined in Equation 4. However, the determination of θ requires another optimization process for Equation 7, which imposes additional difficulties. This k dimensional unconstrained optimization problem can be reduced to a one-dimensional one assuming the correlation parameter for each variable to be the same and using Equation 9 instead. The effect of this simplification in the *theta* optimization was investigated in Figures 6 and 7. The models of Figure 6 were generated using Equation 9 (1-D θ optimization), while those of Figure 7 used Equation 4 (k-D θ optimization). ¿From comparisons between Figures 6 and 7, it was found that the simplification caused little degradation at least in this 2-D test design problem. However, we can deduce that this result was due to pure coincidence involving two design variables which had almost the same variable range scale and sensitivity to the objective function. The authors' previous study indicated that as the number of design variables increased, simplifying the θ optimization would cause severe degradation of the model accuracy and additional difficulties in selecting the right values for θ. The utilization of the gradient also helps to find the correct values for θ, even though the θ selection process remains as one of the biggest challenges for Kriging and Cokriging techniques in general.

In Figure 8, the contour plots for each of the models are compared. Within the range of sample points, the Cokriging models accurately replicated the CFD calculation results in terms of the general shape and the locations of extrema.

To investigate the applicability of the Cokriging models in 2-dimensional design spaces, C_D optimizations were performed using each of the approximation models generated above. Figure 9 (a) shows the C_D optimization results using the data base constructed from 400 CFD calculations over the design space. The optimization was repeated five different times by varying the starting points. Four of them converged to one local minimum point and one converged to a different one. The optimizations were again repeated using different approximation models generated from Kriging and Cokriging methods, and the results are compared in the subsequent figures. As shown, the predicted optimum design point and optimum value of C_D for these cases were nearly identical. The predicted C_D from the CFD database was 0.0059708 whereas those for the Cokriging models with 5 and 9 sample points (and gradients) were 0.0059759 and 0.0059758 respectively with almost same optimum design points. The %error for the optimized C_D value is within 0.086 %. As we can observe from Figure 9 (b), original kriging's ability to simulate the unknown function was clearly limited without an extensive set of sample data. The optimum point and the predicted C_D value were far off from the actual ones. The optimization results are summarized in Table 1.

5.3 Cokriging Modeling Test on 5-Dimensional SBJ Design Problem

The design test case was extended to the slightly more realistic design problem of 5-dimensional SBJ C_D minimization. The base design point for the SBJ design problem was chosen to be given by the following values of the design variables in the problem: x_1=5.0, x_2=2.7, x_3=0.45, x_4=0.50, x_5=0.45. The design variables are defined in section 3 and in Figure 1. Design optimization was performed using the indirect Cokriging model generated from 11 sample points which consisted of one base design point and 2 extra design points located along each design variable axis. The results from two successive optimization cycles with the following bounds on the design variables are presented in Table 2. These bounds were imposed by adding a quadratic penalty function to the model.

$$
\begin{aligned}
0.00 \leq x_1 \leq 10.0,\\
2.40 \leq x_2 \leq 3.00,\\
0.30 \leq x_3 \leq 0.60,\\
0.44 \leq x_4 \leq 0.56,\\
0.39 \leq x_5 \leq 0.51,
\end{aligned}
$$

The first design cycle started from the base design point and resulted in a predicted C_D of 0.0055332 with the optimum design at $x_1 \sim x_5$=$[5.0521, 2.7401, 0.5022, 0.4495, 0.3904]$. This predicted C_D value was checked by running the CFD calculation at the predicted optimum design. The C_D value from the CFD calculation was 0.0055609, whereas the base design point had C_D of 0.0064787. With only one design cycle using a Cokriging model constructed from 11 samples and their gradients, the C_D of SBJ design can be reduced by 14.17%, and the error between the predicted C_D value and the verified one from CFD was merely 0.498%. As expected, the wing aspect ratio tended to increase and the fuselage radii decreased. The wing sweep angle did not change significantly. This may be due to the fact that its sensitivity for C_D was small compared to other variables within the limits set up for this case. Overall, this test verified the accuracy and efficiency of the Cokriging model within a realistic design environment.

Conclusions

In this study, a method to improve the efficiency of the original Kriging method by using gradient information at the sampling sites was investigated. This method uses Cokriging techniques and its applicability in a realistic design problem were evaluated for a 5-dimensional Supersonic Business Jet Design. The results show that the modified Kriging method can generate approximation models with much higher accuracy than the original Kriging method with greatly improved efficiency in terms of reducing the computational cost of obtaining sample data. Confidence on the fact that the proposed method can be used efficiently in realistic high-dimensional design problems such as in complex multidisciplinary system designs was gained. In particular, we are interested in using this type of procedure to extend the use of adjoint methods to more dramatic geometry changes such as those to be found in planform and geometry optimization. For design problems where larger areas of the design space are investigated, a procedure such as the one described in this work is likely to be more effective than traditional adjoint approaches where small steps for changes in the configuration are required.

Future Work

The presented study identified the areas of further investigation. They include (1) research on a robust method to determine the correlation parameter θ, (2) testing of the Kriging method in higher dimensional design applications, and (3) research on various techniques of design of experiments (DOE) most suitable for the proposed method.

References

[1]Thomas A. Zang and Lawrence L. Green. "Multidisciplinary Design Optimization Technique: Implications and Opportunities for Fluid Dynamics" *30th AIAA Fluid Dynamics Conference*, Norfolk, Virginia, AIAA 99-3798, June 1999

[2]Jaroslaw Sobieszczanski-Sobieski and Raphael T. Haftka "Multidisciplinary Aerospace Design Optimization: Survey of Recent Developments," *34th AIAA Aerospace Sciences Meeting and Exhibit* , Reno, Nevada, AIAA 96-0711, January 1996.

[3]Resit Unal, Roger A. Lepsch, Jr. and Mark L. McMillin "Response Surface Model Building and Multidisciplinary Optimization Using Overdetermined D-Optimal Designs," *7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, AIAA 98-4759, September 1998

[4]Anthony A. Giunta and Layne T. Watson "A Comparison of Approximation Modeling Techniques: Polynomial Versus Interpolating Models," *7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization* , St. Louis, Missouri, AIAA 98-4758, September 1998.

[5]J. Sacks, W. J. Welch, T. J. Michell, and H. P. Wynn "Design and Analysis of Computer Experiments," *Statistical Science Vol 4. No.4*, pp. 409-453, 1989

[6]Timothy W. Simpson, Timothy M. Mauery, John J. Korte and Farrokh Mistree "Comparison of Response Surface and Kriging Models in the Multidisciplinary Design of an Aerospike Nozzle," *7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, St. Louis, Missouri, AIAA 98-4755, September 1998

[7]J. R. Koehler and A. B. Owen "Computer Experiments," *Handbook of Statistics Vol. 13*, pp. 261-308, Elsevier Science, New York, eds. S. Ghosh and C. R. Rao

[8]Edward H. Isaaks and R. Mohan Srivastava *An Introduction to Applied Geostatistics*, Oxford Univ. Press, Oxford, 1989.

[9]A. Jameson, W. Schmidt, and E. Turkel "Numerical Solutions of the Euler Equations by Finite Volume Methods with Runge-Kutta Time Stepping Schemes," *AIAA Paper 81-1259*, Jan. 1981.

[10]A. Jameson "Multigrid Algorithms for Compressible Flow Calculations," *Lecture Note in Mathematics* edited by W. Hackbusch and V. Trottenberg, Vol. 1228, Springer-Verlag, 1986.

[11]James J. Reuther, Antony Jameson, and Juan J. Alonso "Constrained Multipoint Aerodynamic Shape Optimization Using an Adjoint Formulation and Parallel Computers, Parts I and II," *Journal of Aircraft* Vol 36, No 1, January-February, 1999

[12]A. Jameson "Optimum Aerodynamic Design Using CFD and Control Theory" *AIAA 12th Computational Fluid Dynamics Conference* , AIAA paper 95-1729, San Diego, CA, 1995.

[13]A. Jameson, L. Martinelli and N. A. Pierce "Optimum Aerodynamic Design Using the Navier-Stokes Equations," *Theoretical and Computational Fluid Dynamics* Vol 10, pp. 231–237, 1998.

[14]A. Jameson, N. Pierce and L. Martinelli "Optimum Aerodynamic Design Using the Navier-Stokes Equations," *AIAA Paper 97-0101*, Jan. 1997.

[15]*MATLAB User's Guide for Unix Workstations*, The MathWorks, Inc. Mass., August 1992, January-February, 1999

	Using Exact Function Estimated from 400 CFD Data	Using Cokriging Model with 5 Samples and Gradients	Using Cokriging Model with 9 Samples and Gradients
Optimum Design	$X_1 = 0.36543$ $X_2 = 0.43883$	$X_1 = 0.36575$ $X_2 = 0.43839$	$X_1 = 0.36575$ $X_2 = 0.43839$
Predicted Optimum	0.00597063	0.00597588	0.00597508
Verified Optimum	0.00597063	0.00597063	0.00597063
% Error (%)	0.0	0.0879	0.0745

Table 1. Comparison of 2-D SBJ Design Optimization Results
(Base Design Point : X_1=0.4, X_2=0.45)

	1st Design Cycle	2nd Design Cycle
Optimum Design	$X_1 = 5.0521$ $X_2 = 2.7401$ $X_2 = 0.5022$ $X_2 = 0.4495$ $X_2 = 0.3904$	$X_1 = 5.0424$ $X_2 = 2.7154$ $X_2 = 0.5175$ $X_2 = 0.4556$ $X_2 = 0.3906$
Predicted Optimum	0.00553315	0.00550393
Verified Optimum	0.00556087	0.00553665
% Error (%)	0.499	0.591

Table 2. 5-D SBJ Design Optimization Results
(* Calculated using CFD analysis code with predicted optimum design)

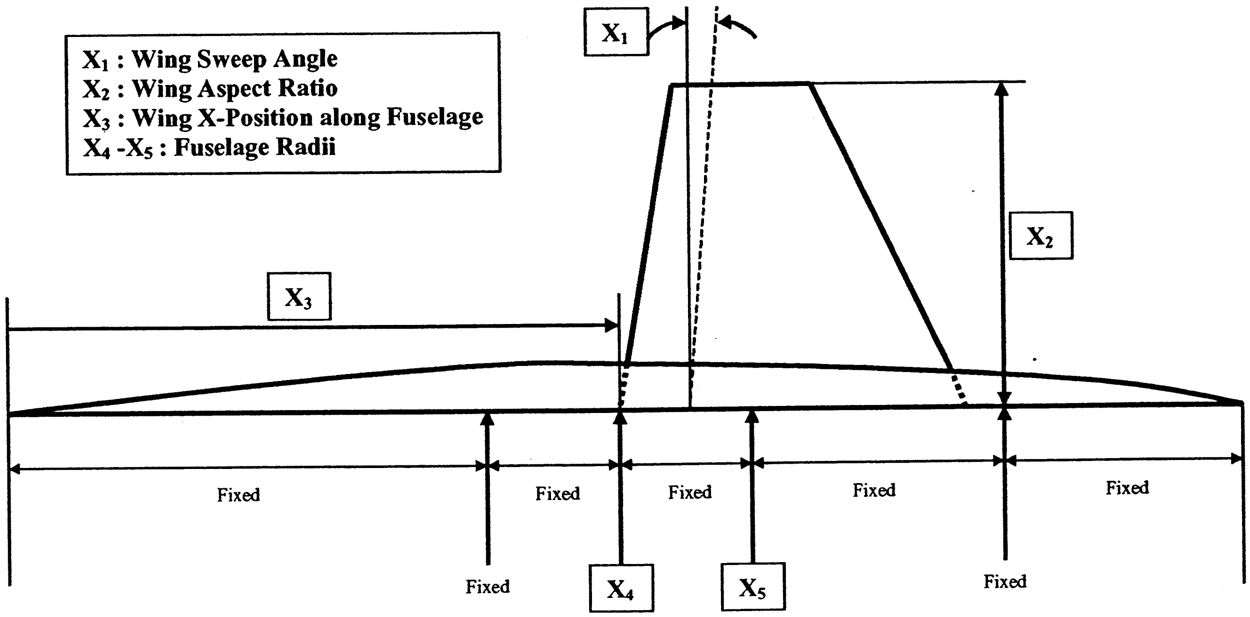


Fig. 1 Definition of Design Variables for SBJ Wing-Body Configuration

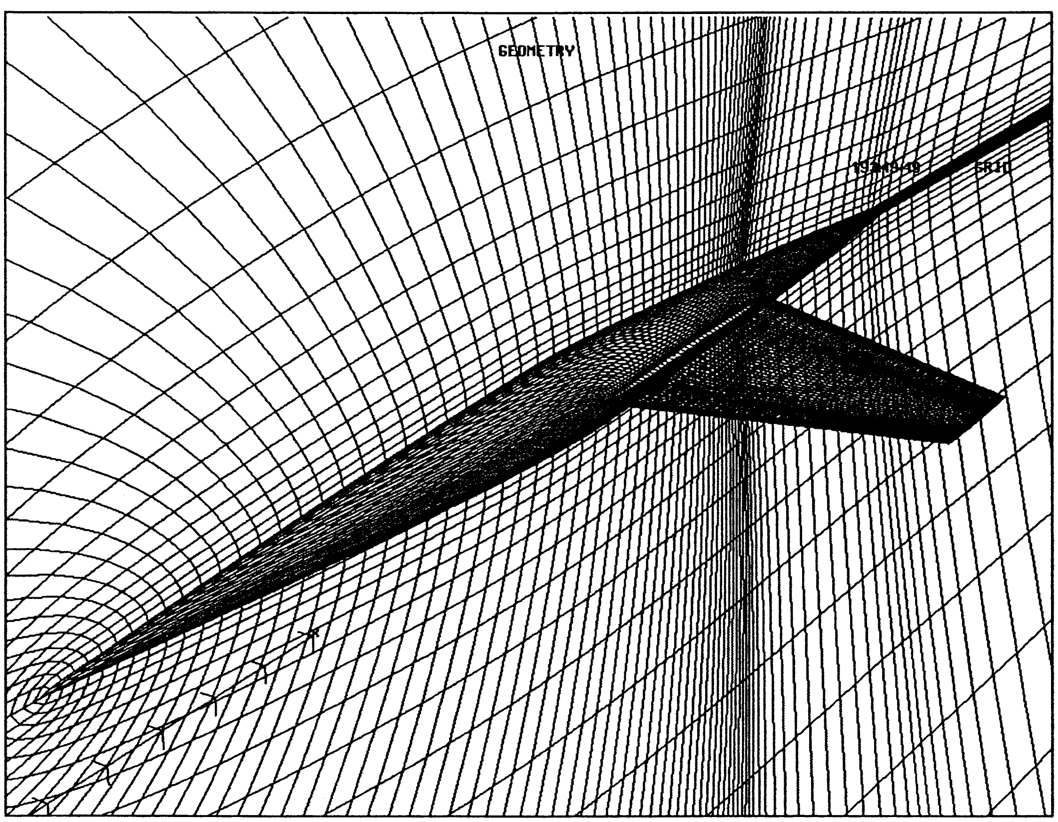
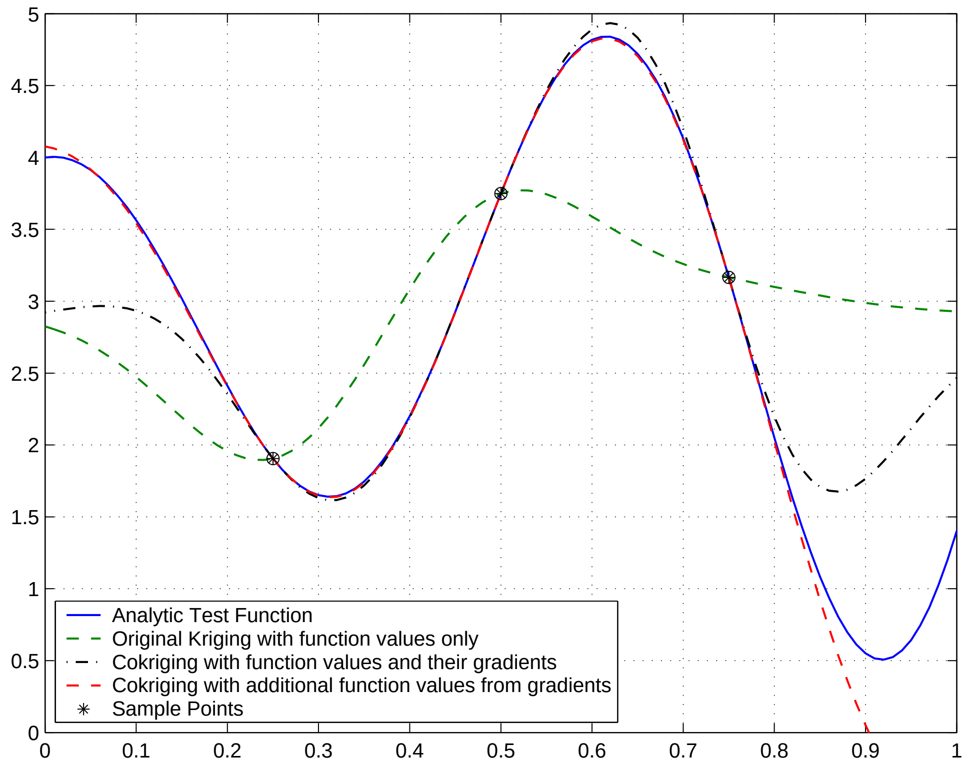
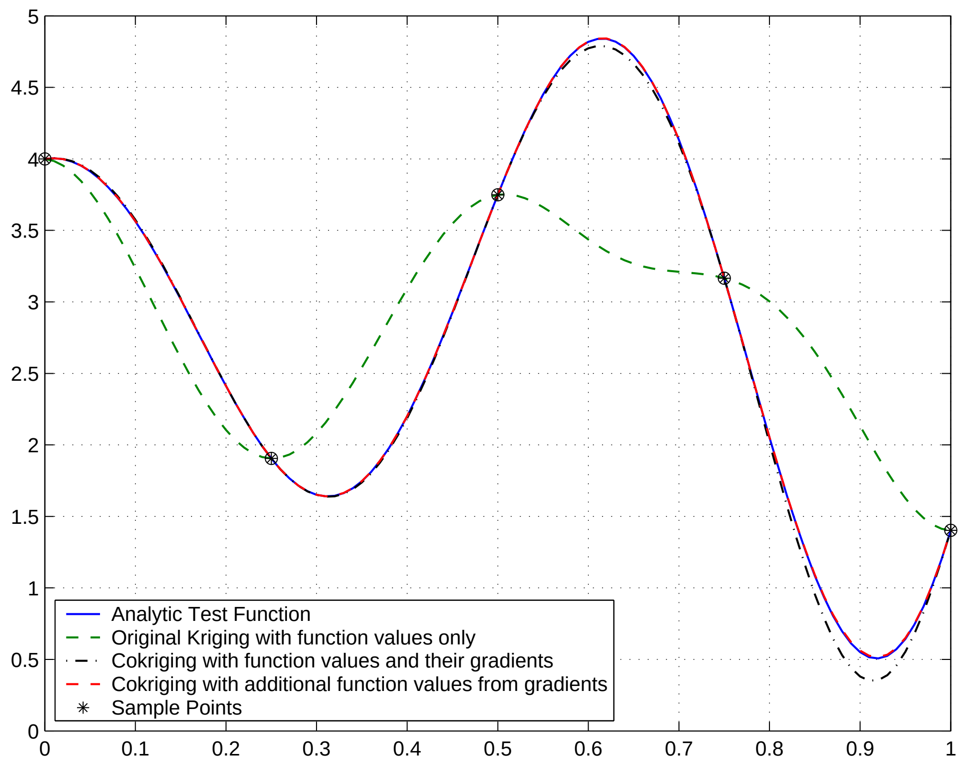


Fig. 2 A Typical Supersonic Business Jet Mesh for CFD Calculation

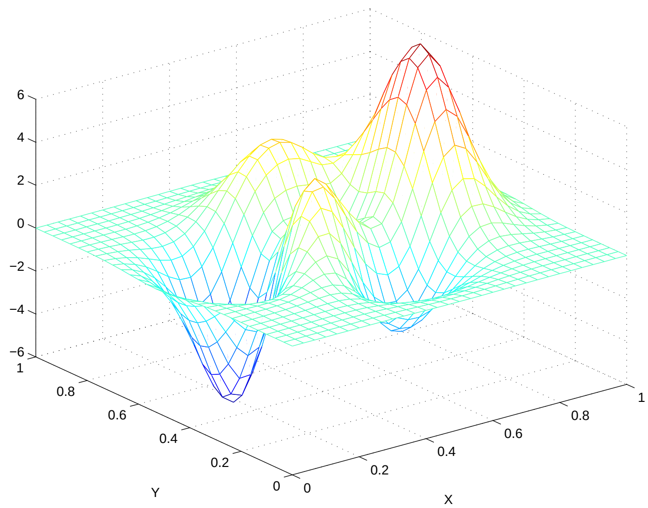


(a) Kriging and Cokriging Models with 3 Sample Values and their Gradients

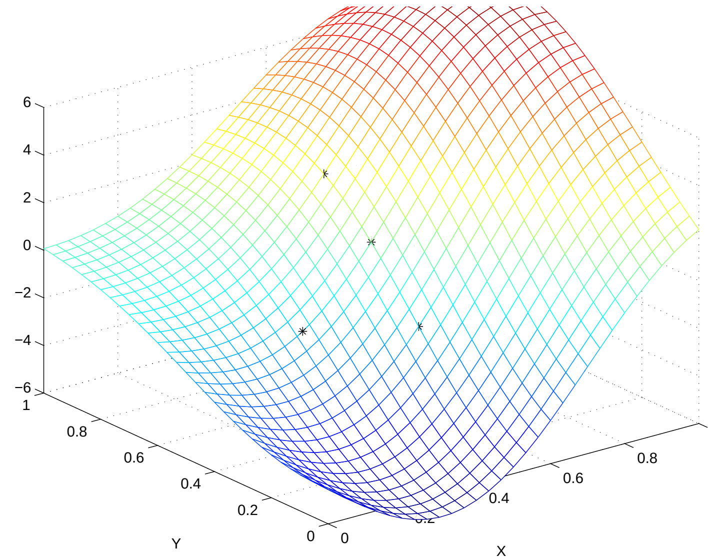


(b) Kriging and Cokriging Models with 5 Sample Values and their Gradients

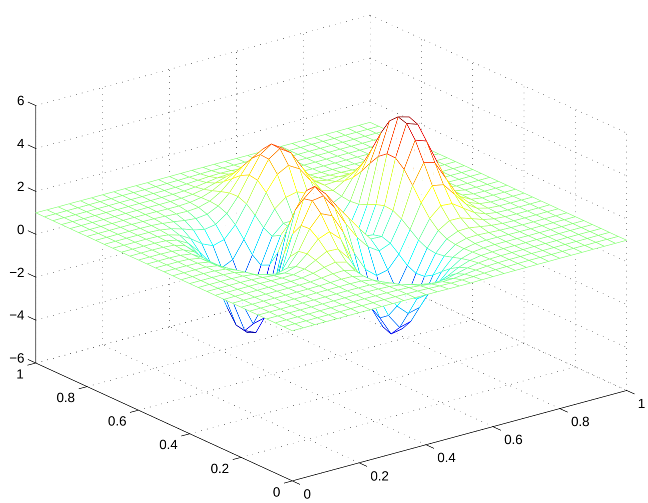
Fig. 3 Cokriging Model Validation on One-Dimensional Analytic Test Function



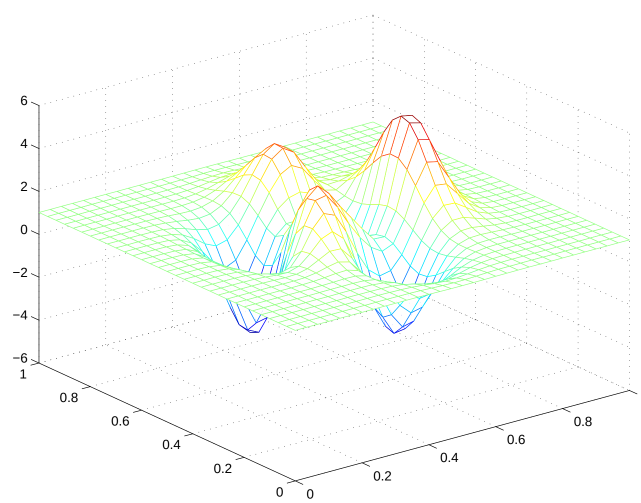
(a) 2-D Analytic Test Function



(b) Original Kriging Model with 5 Sample Values

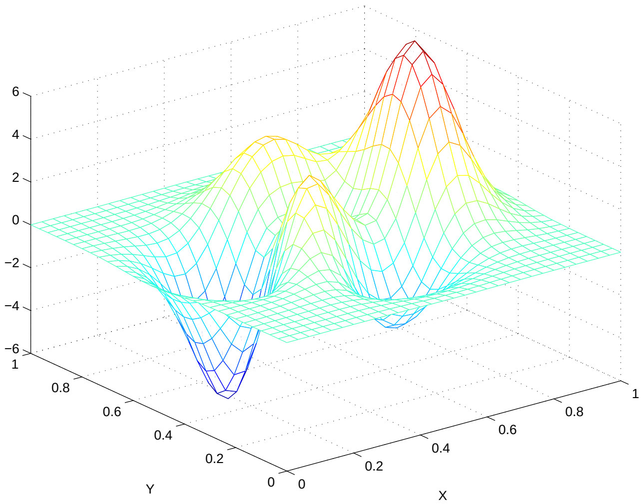


(c) Cokriging Model with 5 Sample Values and their Gradients

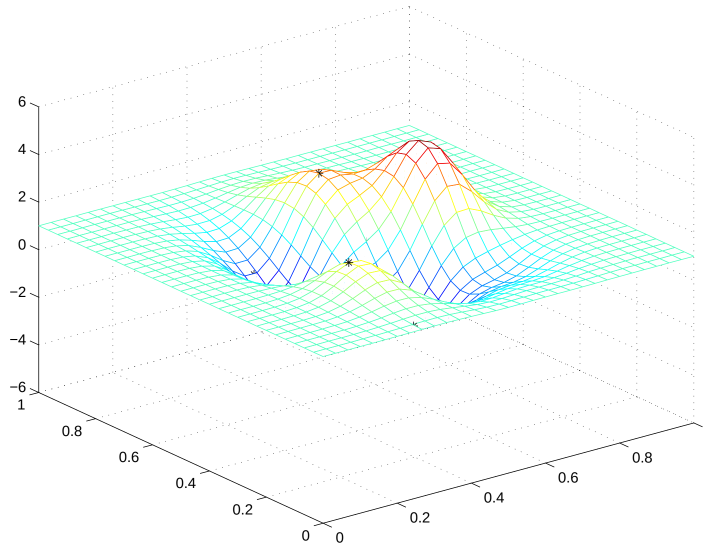


(d) Cokriging Model with 5 Samples and 10 additional Values obtained from gradients

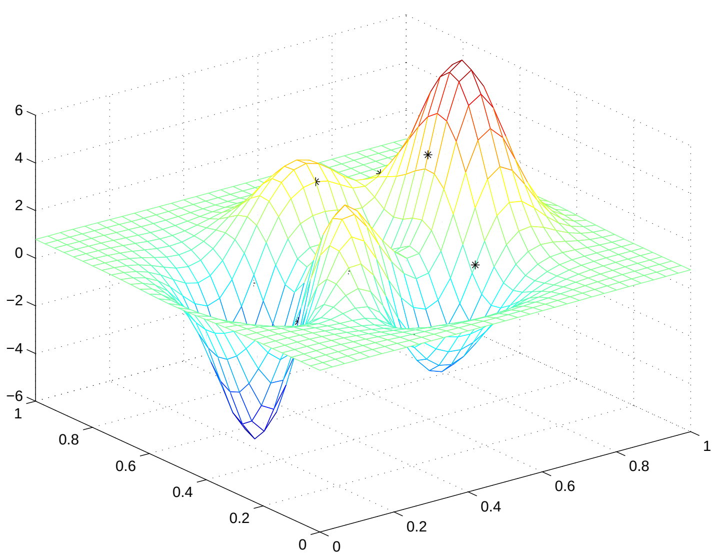
Fig. 4 Cokriging Model Validation on Two-Dimensional Analytic Test Function (5 Samples)



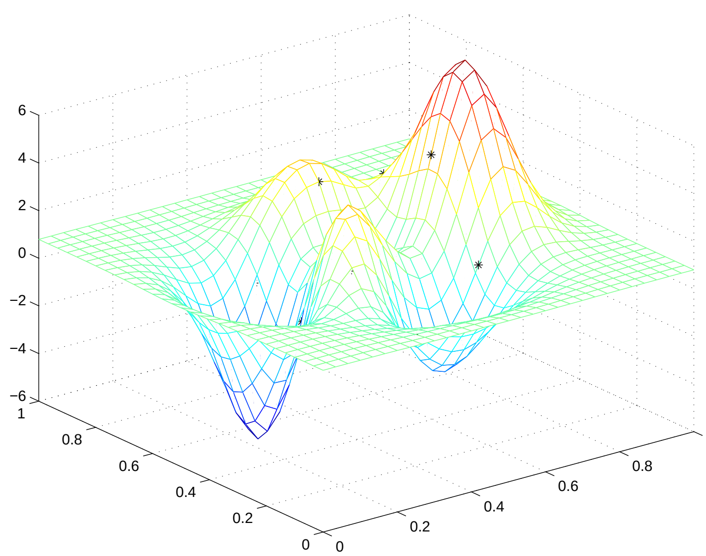
(a) 2-D Analytic Test Function



(b) Original Kriging Model with 9 Sample Values

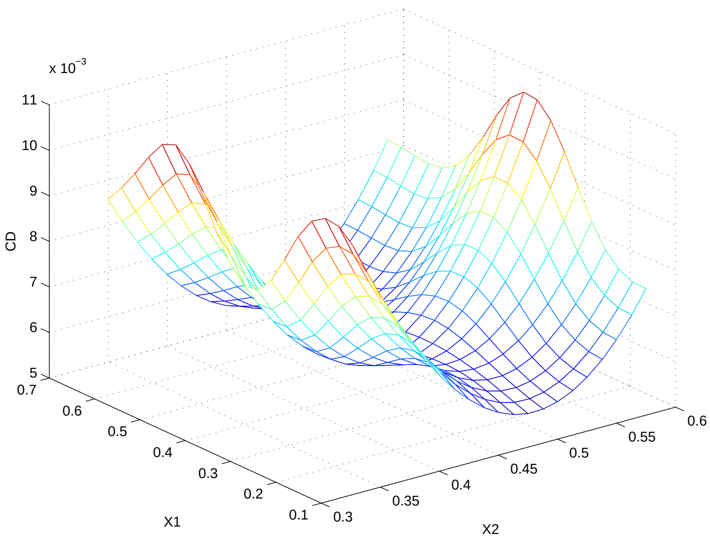


(c) Cokriging Model with 9 Sample Values and their Gradients

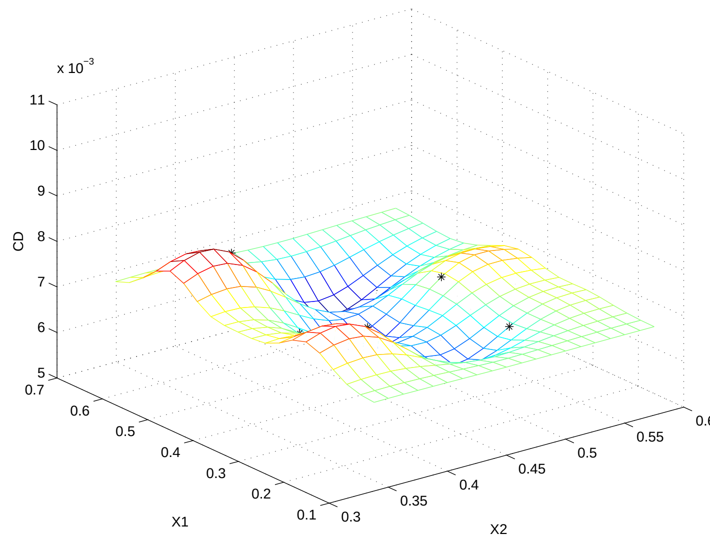


(d) Cokriging Model with 9 Samples and 18 additional Values obtained from gradients

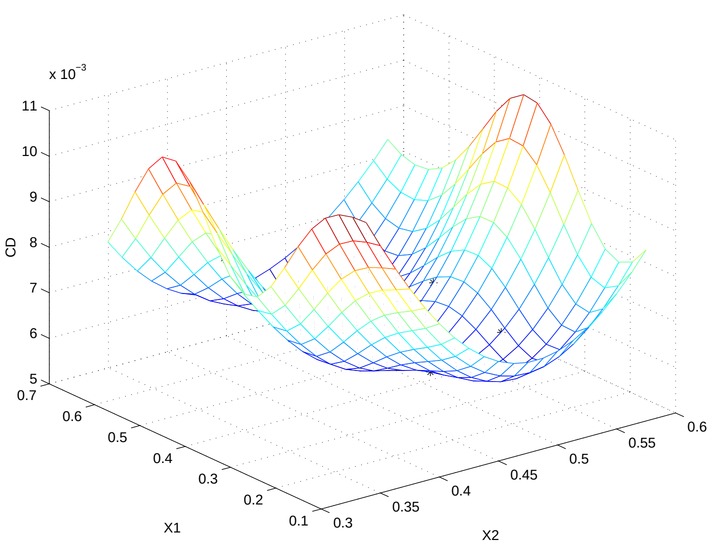
Fig. 5 Cokriging Model Validation on Two-Dimensional Analytic Test Function (9 Samples)



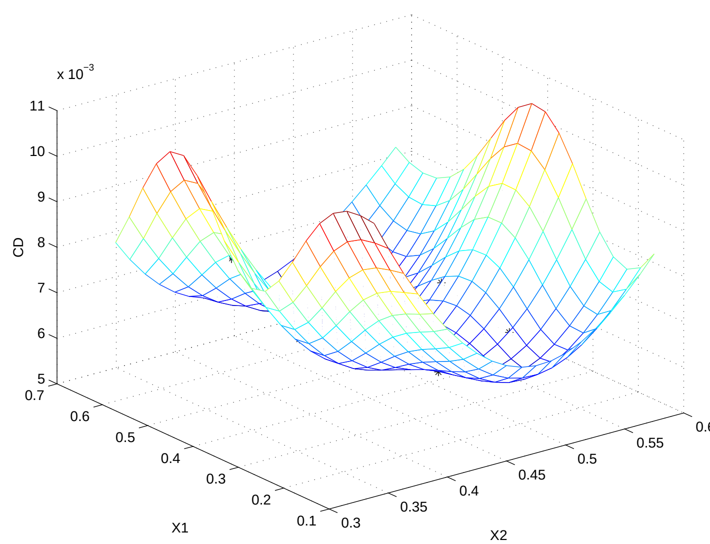
(a) CFD Calculation Results over 400 Design Points



(b) Original Kriging Model with 9 Sample Values

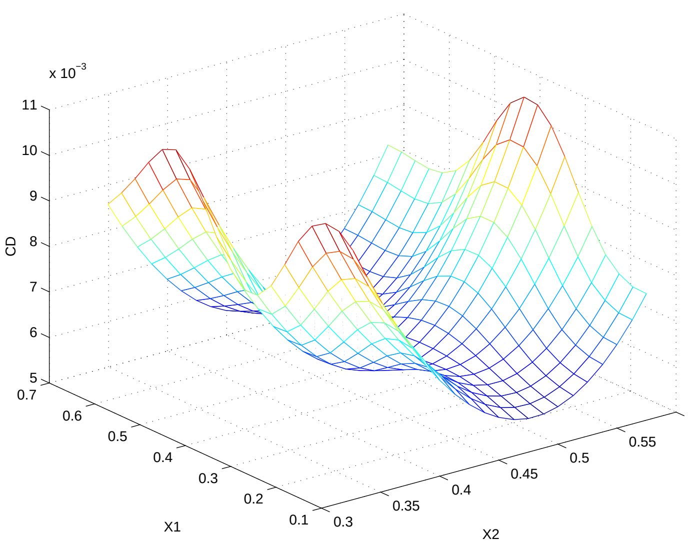


(c) Cokriging Model with 9 Sample Values and their Gradients

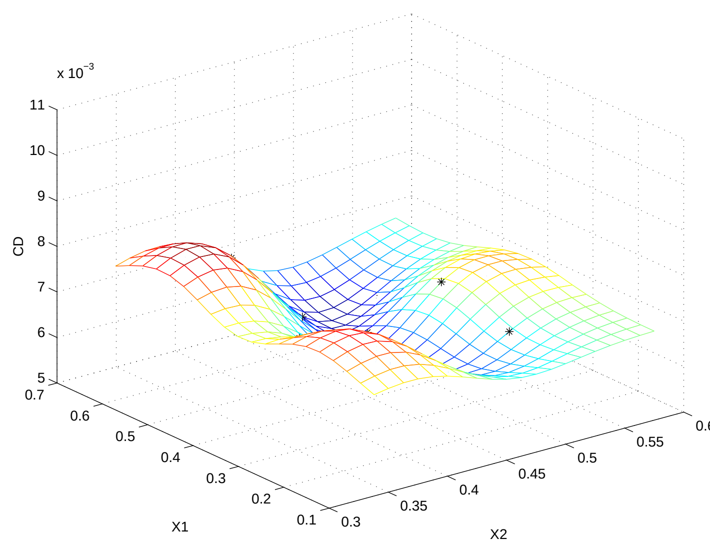


(d) Cokriging Model with 9 Samples and 18 additional Values obtained from gradients

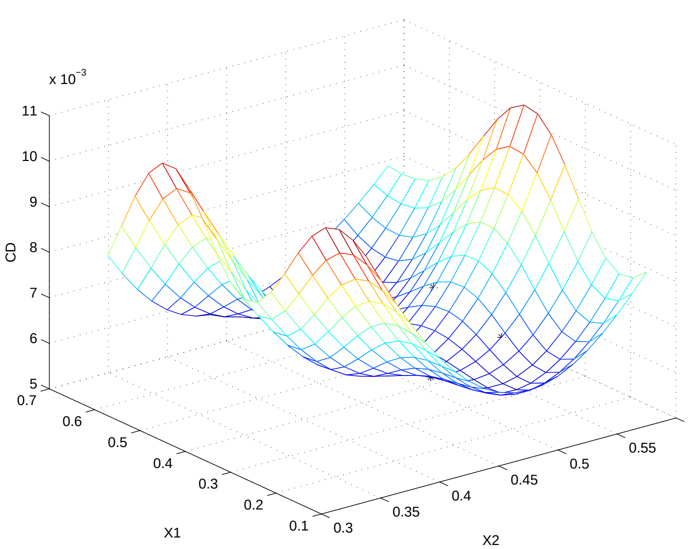
Fig. 6 C_D **Cokriging Models for 2-D SBJ Design Problem (1-D Correlation Parameter)**



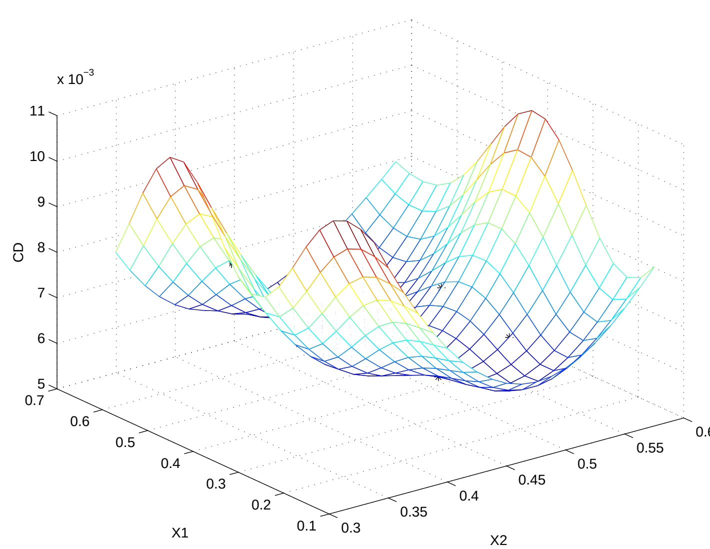
(a) CFD Calculation Results over 400 Design Points



(b) Original Kriging Model with 9 Sample Values

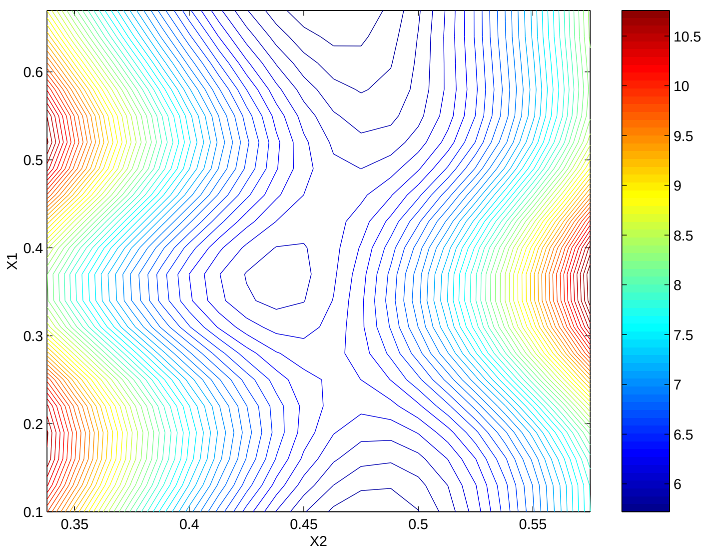


(c) Cokriging Model with 9 Sample Values and their Gradients

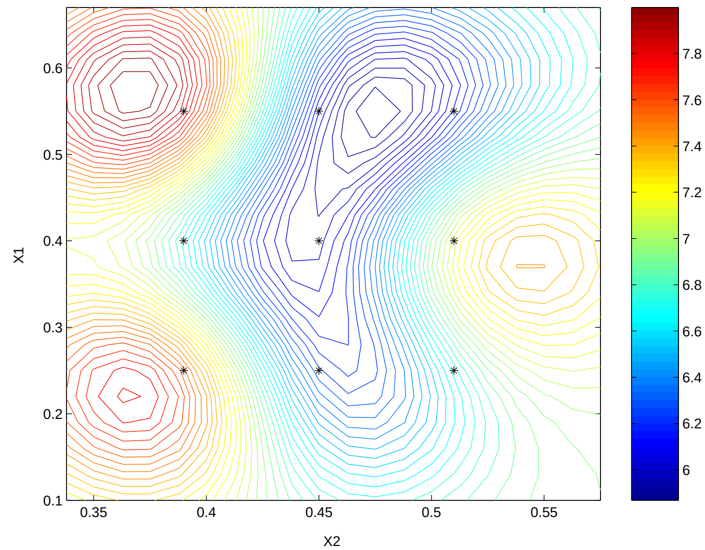


(d) Cokriging Model with 9 Samples and 18 additional Values obtained from gradients

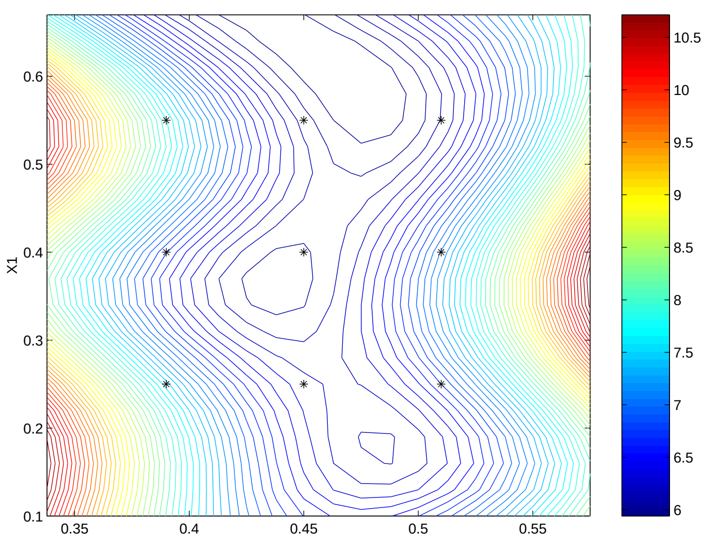
Fig. 7 **C_D Cokriging Models for 2-D SBJ Design Problem (2-D Correlation Parameter)**



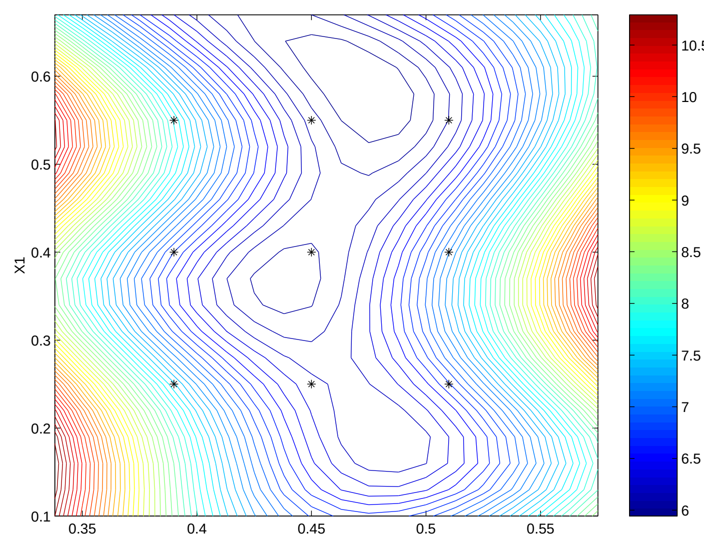
(a) CFD Calculation Results over 400 Design Points



(b) Original Kriging Model with 9 Sample Values

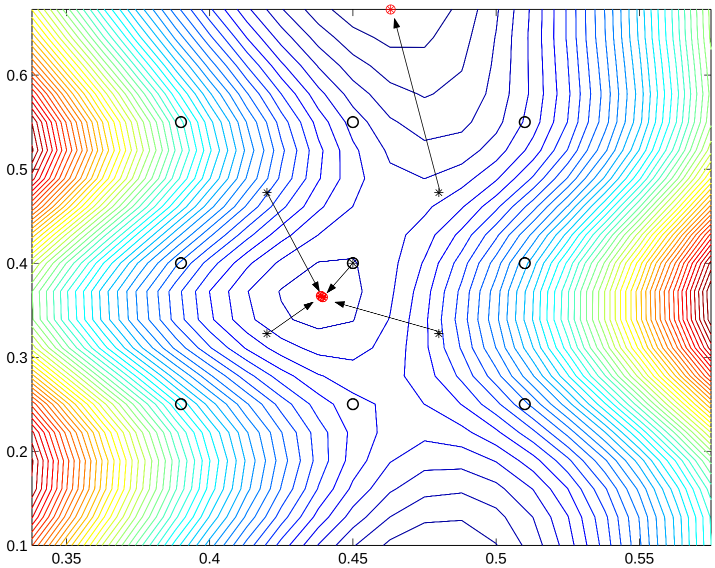


(c) Cokriging Model with 9 Sample Values and their Gradients

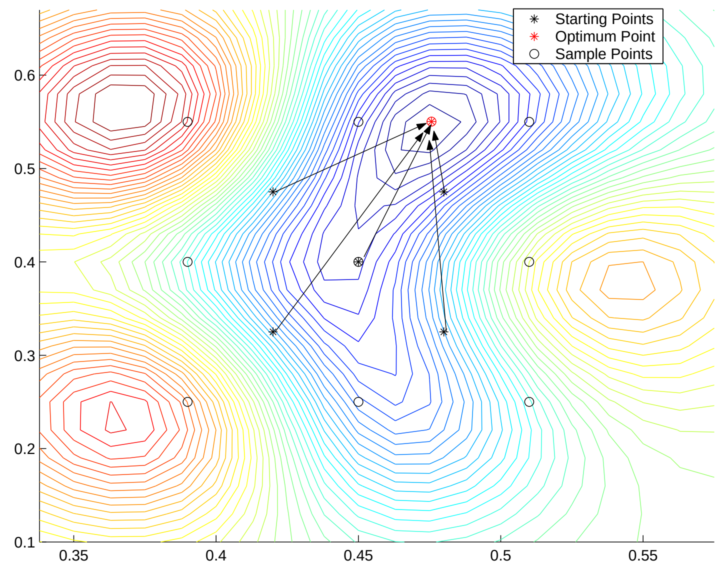


(d) Cokriging Model with 9 Samples and 18 additional Values obtained from gradients

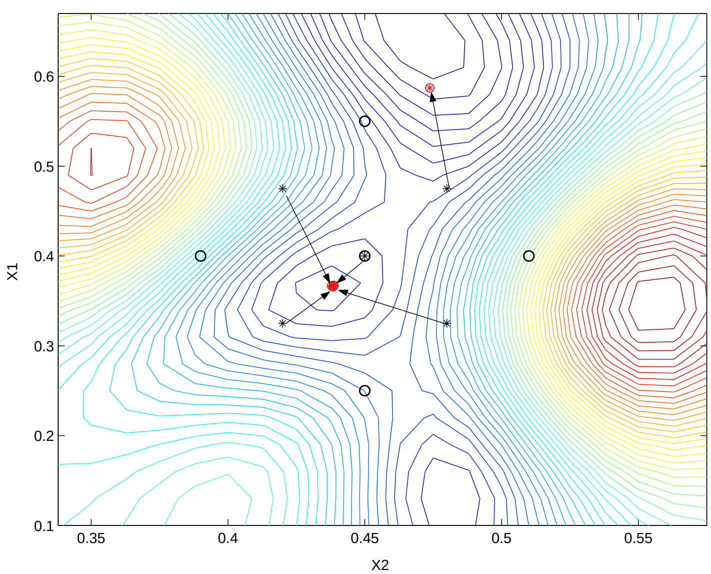
Fig. 8 Contour Plots of C_D Cokriging Models for 2-D SBJ Design Problem



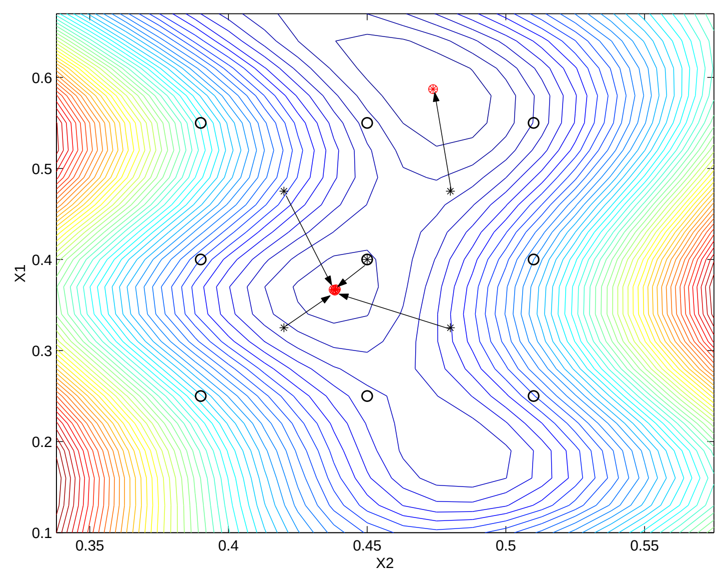
(a) 2-D CD Optimization Results Using CFD Calculation Results over 400 Design Points



(b) 2-D CD Optimization Results Using Original Kriging Model with 9 Sample Values



(c) 2-D CD Optimization Results Using Cokriging Model with 5 Sample Values and their Gradients



(d) 2-D CD Optimization Results Using Cokriging Model with 9 Sample Values and their Gradients

Fig. 9 Contour Plots of 2-D C_D Optimization Results for 2-D SBJ Design Problem