EFFICIENT AERODYNAMIC OPTIMIZATION USING HIERARCHICAL KRIGING COMBINED WITH GRADIENT

Chao SONG, Xudong YANG, Wenping SONG
National Key Laboratory of Science and Technology on Aerodynamic Design and Research, Northwestern Polytechnical University, Xi'an, 710072, China

Keywords: *kriging; hierarchical kriging; gradient; optimization*

Abstract

The gradient information could improve the design efficiency of the original kriging model. A new method combined gradients with Hierarchical Kriging (HK) model is developed in this paper. The model is built in two steps. At first, initial sample points and the gradient with respect to design variables are computed by high-fidelity solvers. Gradient information and a selected step size are used for Taylor approximation to obtain derived sample points. Then these sample point are used for building a low-fidelity kriging model. At last, a high-fidelity model can be obtained by adjusting the low-fidelity using the initial samples. An analytical function test case demonstrates that the gradient enhanced Hierarchical Kriging (GEHK) model has overcome limitations of traditional gradient-based kriging model, and the prediction accuracy of the model can be improved evidently. In the airfoil drag reduction case, the GEHK model improves the optimization efficiency, and could get a better result compared with the ordinary kriging model.

1 Introduction

Surrogate modeling can be used to greatly improve the design efficiency. The kriging model is perhaps the most effective, due to the advantage of capturing multiple local extrama of highly nonlinear functions [1]. With the development of computers, the kriging model has been extensively used in aerodynamic optimizations [2~4]. However, the accuracy of the kriging model is limited by training strategies and the number of samples [5]. When a large number of design variables are involved in the design problem and high-fidelity numerical models are utilized, it is still very time-consuming for the design process.

To enhance the efficiency of a kriging model, gradients of the objective function can be involved in the model. There are two ways to combine kriging model with gradients.[6] The one is Direct Cokriging (D.Cok.). In this method, the samples and the gradient information are conserved through a heritage of the correlation function and parameters. The other one is the Indirect Cokriging (I.Cok). The original kriging formulation can be used with an augmentation sample data derived by gradients. It was found that the parameters of D.Cok. may converge to wrong values, then the model is ill-fitted.[7] For I.Cok, step sizes used to derive new samples are key parameters. A too long step may destroy the accuracy of derived samples, while a too small step cannot introduce much gradient information for the model, due to similarity of samples and derived samples. Laurenceau [7] suggested a fixed step size of 10^{-4} times of design space range. Liu [8] presented an approach to solve the problem by including the step size for the design variables as one of parameters in kriging models.

This objective of the present work paper is to develop a new method to combine gradient information with kriging models. A recently proposed kriging model called Hierarchical Kriging (HK) is utilized and the gradient is solved by the adjoint method. The HK model is briefly introduced in section 2, and the proposed method in this paper is described in section 3. An analytical function optimization case is tested using the I.Cok. and GEHK method in section 4, then an airfoil drag reduction case is demonstrated. Concluding remarks are provided in section 5.

2 Overview of Kriging Model

2.1 Kriging

Here we just present a brief describe about kriging modeling. A more detail information about the model can be found in the reference [5]. Kriging models combine a global model plus localized departures:

$$y(\mathbf{x}) = \mu + Z(\mathbf{x}) \qquad (1)$$

where $\mathbf{x}$ is an m-dimensional vector (m is the number of design variables), and $y(\mathbf{x})$ is the unknown function of interest.

The first term in Eq. (1) provides a "global" model of the design space, which is called regression model. We employed a constant term in this paper. The second is a realization of a stochastic process with zero mean, Z(x). The Z(x) term represents a local deviation from the global model, calculated by quantifying the correlation of x with nearby points. The covariance matrix of Z(x) is given by Eq. (2).

$$Cov[Z(\mathbf{x}^i), Z(\mathbf{x}^j)] = \sigma^2 \mathbf{R}[R(\mathbf{x}^i, \mathbf{x}^j)] \qquad (2)$$

In Eq. (2), R is the correlation matrix, and $R(\mathbf{x}^i, \mathbf{x}^j)$ is the stochastic-process correlations between any two of the n_s sample data points $\mathbf{x}^i, \mathbf{x}^j$. In this paper, the authors use a spline correlation model, which is defined by:

$$R_j = \begin{cases} 1 - 15\xi_j^2 + 30\xi_j^3 & 0 \le \xi_j \le 0.2 \\ 1.25(1-\xi_j)^3 & 0.2 \le \xi_j \le 1 \\ 0 & \xi_j \ge 1 \end{cases} \qquad (3)$$

where $\xi_j = \theta_j |w_j - x_j|$,and $\theta_j(\theta_j \ge 0)$ are the unknown correlation parameters used to fit the model, and w_j and x_j are the j-th components of vectors w and x. The kriging predictor for the values of x is obtained from

$$\hat{y}(\mathbf{x}) = \hat{\mu} + \mathbf{r}^T \mathbf{R}^{-1}(y - \mathbf{1}\hat{\mu}) \qquad (4)$$

where $\hat{\mu}$ is the estimated value of μ, and y is a column vector of length ns that contains the sample values of the response, $\mathbf{r}$ is the correlation vector of length n_s between an untried x and the sample data points.

$$\mathbf{r}(x) = [R(\mathbf{x}, \mathbf{x}^1), R(\mathbf{x}, \mathbf{x}^2), \ \dots \ , R(\mathbf{x}, \mathbf{x}^{ns})]^T \qquad (5)$$

For any given vector $\boldsymbol{\theta}$ that consists of components θ_k , $\hat{\mu}$ and the estimate of the variance $\hat{\sigma}^2$ can be defined as

$$\hat{\mu} = \frac{\mathbf{1}^T \mathbf{R}^{-1} y}{\mathbf{1}^T \mathbf{R}^{-1} \mathbf{1}} \qquad (6)$$

$$\hat{\sigma}^2 = \frac{(y - \mathbf{1}\hat{\mu})^T \mathbf{R}^{-1} (y - \mathbf{1}\hat{\mu})}{n_s} \qquad (7)$$

The unknown correlation parameters $\boldsymbol{\theta}$ of the kriging model are estimated by maximizing the following likelihood function.

$$Ln(\theta) = -[n_s \ln(\hat{\sigma}^2) + \ln|\mathbf{R}|]/2 \qquad (8)$$

2.2 Hierarchical Kriging

In a standard kriging model, the regression model is defined as low-polynomials, typically linear or quadratic. Instead of using polynomials, the HK uses a secondary kriging model as the model trend of the primary kriging model of the function of interest. A kriging model for low-fidelity function is built as the exactly same process of a standard kriging. Refer to Eq.(4), the prediction of the low-fidelity function at any untried point X can be written:

$$\hat{y}_{lf}(\mathbf{x}) = \hat{\mu}_{lf} + \mathbf{r}_{lf}^T R_{lf}^{-1}(y_{lf} - \mathbf{1}\hat{\mu}_{lf}) \qquad (9)$$

There is an assumption that the random process corresponding to the high-fidelity function is of the form:

$$y(\mathbf{x}) = \hat{\mu}\hat{y}_{lf} + Z(\mathbf{x}) \qquad (10)$$

The HK predictor can be written as

$$\hat{y}(\mathbf{x}) = \hat{\mu}\hat{y}_{lf}(\mathbf{x}) + \mathbf{r}^T \mathbf{R}^{-1}(y - \hat{\mu}\mathbf{F}) \qquad (11)$$

where $\hat{\mu} = \frac{\mathbf{F}^T \mathbf{R}^{-1} y}{\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F}}$, $\mathbf{F} = [y_{lf}(x^{(1)}, \ \dots \ ^{n_s}))]^T$.

Compared with Eq.(4), the scaling factor $\hat{\mu}$ of HK can be obtained by replace $\mathbf{1}$ with $\mathbf{F}$. The approximated low-fidelity function $\hat{y}_{lf}$ scaled by an unknown constant factor μ, serving as the global trend function.

The HK model also provides a better mean-squared error (MSE) than traditional cokriging , which is one of the unique features of this model. The readers is refer to [9] for more detail information of building a HK model.

3 Proposed Gradient Enhanced HK model

The GEHK model has two level of fidelity in our study. Firstly, some initial samples are obtained by CFD simulation, including aerodynamic characteristic (e.g., lift coefficients, drag coefficients) and their gradients with respect to design variables. Then new samples are derived by first order Taylor approximation using gradients and a selected step size. Although the samples may be accurate enough when steps are appropriate, these new derived samples are considered as the low-fidelity data. A low-fidelity kriging model is built using derived samples. At last, a high-fidelity model can be obtained by adjust the low-fidelity kriging with initial samples. The process of an optimization using GEHK is shown in Fig. 1.

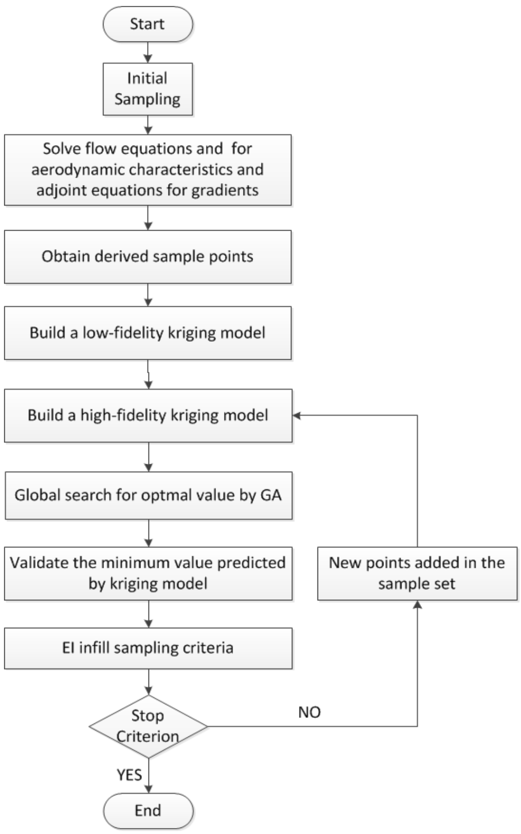


Fig. 1 Flowchart of the optimization using GEHK model

4 Optimization Results

4.1 Analytic Function Test Case

An analytical test case by Forrester [10] is employed here. The analytical function is:

$$y = (6x-2)^2 \sin(12x-4), x \in [0,1] \quad (12)$$

Four sample points was used for building a kriging model. The sample points, the true function and the kriging result are presented in Fig. 2a. Three different step sizes were used for first order Taylor approximation, and two new derived samples were obtained around every initial sample point. Thus, a total number of 8 derived sample points was obtained.

For the I.Cok. model, the initial 4 samples and the 8 derived sample points were used for building the kriging model indiscriminatingly. However, the GEHK was modeled in two steps. At first, the derived samples were used for building a low-fidelity model. Then the low-fidelity was adjusted by the initial sample points.

When the step size $\Delta x = 0.1$, both the I.Cok and GEHK perform better than the kriging without gradient, as shown in Fig. 2a and Fig. 2b. The minimum is overpredicted by the I.Cok and a conservative prediction is obtained by the GEHK. As $\Delta x = 0.3$, the I.Cok. is fitted badly, due to the poor accuracy of derived samples. It will have adverse effect on the process of searching minimum. However, the inaccurate samples are not used directly for building GEHK model. When a larger step size $\Delta x = 0.5$ is used, false minimum is presented by I.Cok (Fig. 2d). The GEHK model continued to be better than the I.Cok. model using a very large step size.

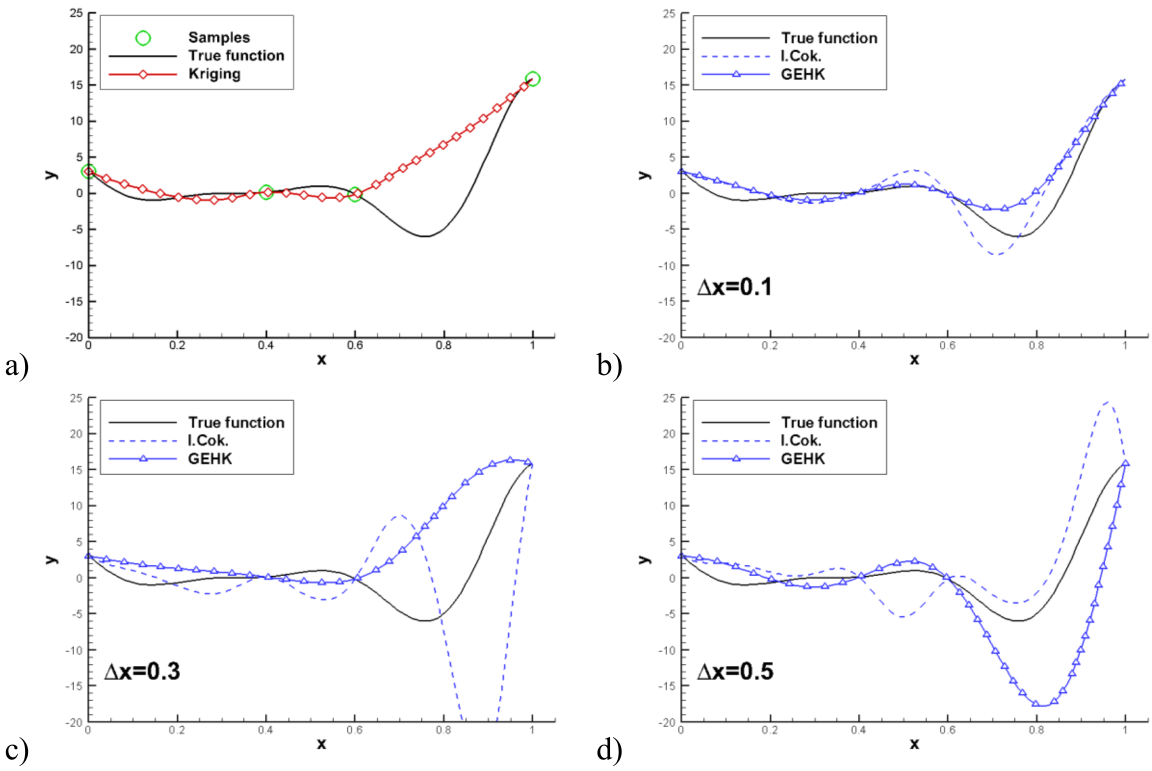


Fig. 2 Comparison of different kriging model for an analytical example

To further demonstrate the advantage of the GEHK method, the expected improvement (EI) [11,12] function was employed to refine the model. The step size was selected as 0.5. The convergence history of objective function is sketched in Fig. 3. After 5 iterations, the GEHK model achieves the minimum. It needs about 10 iterations for the I.Cok model to find the minimum.

It is worth noting that the same sample points are used for both kriging models. For the I.Cok. model, the original kriging formulation is used with an increased number of sample data which located in the proximity of the original sample points. The error could be introduced into the model due to inappropriate step sizes. However, the GEHK does not need very accurate derived samples, because we have the assumption that the derived samples are "low fidelity data" at first. The derived samples provide global trend for the fitting of true function. These samples also serve as if they are gradients because they tend to have strong correlations with the original sample points given the close distances to each other.

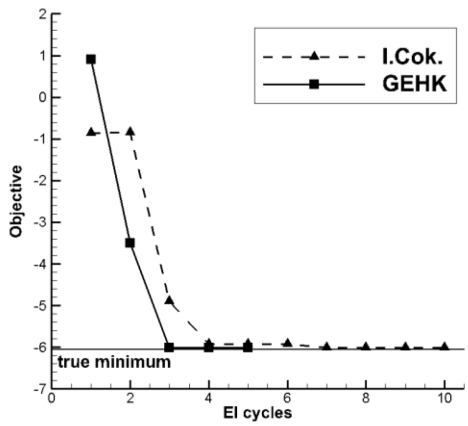


Fig. 3 Convergence history of objective function based on I.Cok. and GEHK method

4.2 Airfoil Drag Reduction Case

In a practical optimization case, multipoint design strategy is often employed to avoid degradation in performance at offdesign conditions. A multipoint design needs more computational cost when additional design conditions are considered. Although the application of kriging model can reduce the computational cost, the optimization is still time-costly due to each design condition requires a simulation. In order to verify the adaptability of the GEHK for multipoint design, this section aims to a three-design point case of RAE2822 airfoil using the GEHK.

Table 1 presents the design points for RAE2822 airfoil in this optimization case. The objective function is a weighted drag coefficient, which was calculated from drag coefficient of each design condition. Weighting coefficient is defined by the relative importance of each design condition.

Table 1 Design points for RAE2822 airfoil multi-point drag reduction optimization

	Ma	Re	α	ω
1	0.73	6.50e6	2.78°	0.4
2	0.726	6.46e6	2.44°	0.4
3	0.6	5.34e6	2.57°	0.2

The aerodynamic and geometrical constraints were imposed by adding penalties to objective function. The optimization problem can be written as:

$$\min \omega_1 \frac{C_{d1}}{C_{d1,0}} + \omega_2 \frac{C_{d2}}{C_{d2,0}} + \omega_3 \frac{C_{d3}}{C_{d3,0}} + c_1\left(1-\frac{A}{A_0}\right) + c_2\left(1-\frac{C_l}{C_{l0}}\right),$$

$$c_1 = \begin{cases} 0 & \frac{A}{A_0} \ge 0.995 \\ 1 & \frac{A}{A_0} < 0.995 \end{cases}, \qquad (13)$$

$$c_2 = \begin{cases} 0 & \frac{C_l}{C_{l0}} \ge 0.995 \\ 1 & \frac{C_l}{C_{l0}} < 0.995 \end{cases}.$$

where A is the area of the airfoil and the subscript 0 indicates the value for the baseline airfoil.

The aerodynamic characteristics (cl, cd) were solved by the SU^2, which was developed by the Aerospace Design Lab (ADL) of Stanford University [13]. The one-equation turbulence model S-A was employed and the JST scheme was used for spatial discretization coupled with an implicit time resolution. Gradients with respect to design variables were solved by the adjoint method. Due to the similarity in the form of adjoint equations and original flow equations, we applied the numerical scheme in a very similar way to compute the adjoint solution. The computational grid is shown in Fig. 4. A hybrid-element mesh has been used with 22,740 total elements. The unstructured portion of the mesh starts at a distance far enough from the airfoil surface. Hence, the boundary layer will remain in the structured portion of the grid. The minimum value of dimensionless wall distance is less than 1 to resolve the boundary layer profile.

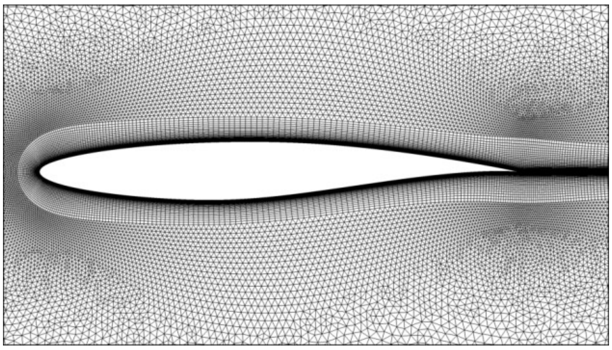


Fig. 4 Convergence history and relative prediction error of multi-point design using the GEHK and kriging model

A 10-variable optimization employing the Hicks-Henne parameterization was considered. The amplitude of each design variable was in the domain ±0.5% of the airfoil chord length.

Before the optimization started, a set of sample points were obtained by Latin Hypercube sampling (LHS). For building a kriging model, the number of samples for each design condition is commonly 10 times of design variables. Then the total number of simulations for building a kriging model is 300. In this particular case, 1/4 samples were selected randomly for GEHK model. A total number of 225 simulations were required, because the gradients of lift and drag were also solved by the adjoint method. The step size is selected to be 10^{-4} times of design space range.

After the initial model has been built, the models needed to be refined by EI infill criteria. The maximum iteration step is limited to 60.

Fig. 5 shows the convergence histories of objective function of GEHK, and that of a kriging models is also shown for comparison.

During the optimization process, the relative error of model prediction at the first design point was evaluated, which is also presented in Fig. 5. When the first iteration completed, the value of objective function obtained by GEHK is obviously lower than that of standard kriging. It is extremely interesting that the value obtained by GEHK is still lower when continuing the standard kriging optimization for a further 80 calculations. It should be noticed that the samples used by GEHK is a subset of standard kriging model. After about 10 iterations, the prediction error of the GEHK model always keeps at a lower level and there are obvious peak values on the standard kriging error curve. Perhaps given enough updates, the standard kriging is expected to approach the global optimum. However, the optimization efficiency could be improved greatly by GEHK model with less computational budget.

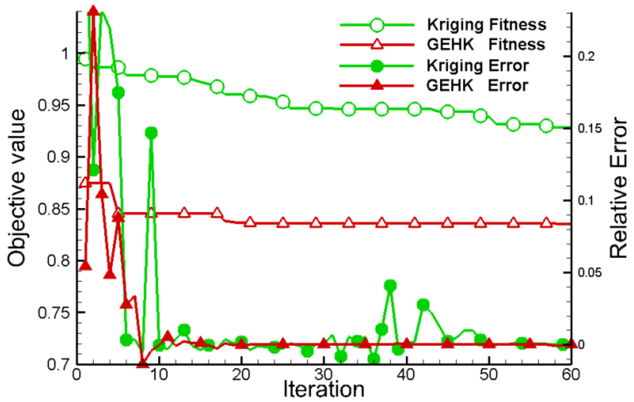


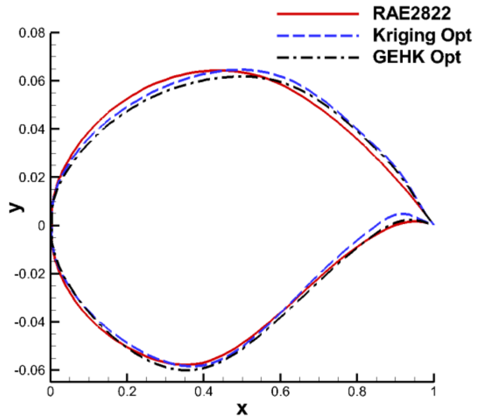
Fig. 5 Convergence history and relative prediction error of multi-point design using the GEHK and kriging model

Table 2 presents the optimal results obtained by the GEHK and the kriging model. Each of the two models attempts to trade off large reduction in drag at $Ma = 0.73$ and $Ma = 0.726$, while simultaneously try to increase the drag at $Ma = 0.6$. Meanwhile, both lift coefficient and area of airfoils are constrained. Compared this two model, the GEHK model offers a better result than ordinary kriging model. At $Ma = 0.73$, the GEHK model reduces the drag coefficient from 0.017120 to 0.012503, and the drag is reduced to 0.014394 by the kriging design. At $Ma = 0.726$, the GEHK model also offers a more substantial reduction than the kriging. The drag increases by 2.6% at $Ma = 0.6$ in the GEHK design case, which is smaller than optimal value obtained by the kriging model. However, at all of the design points, drag reductions is at the expense of a decrease in lift coefficients.

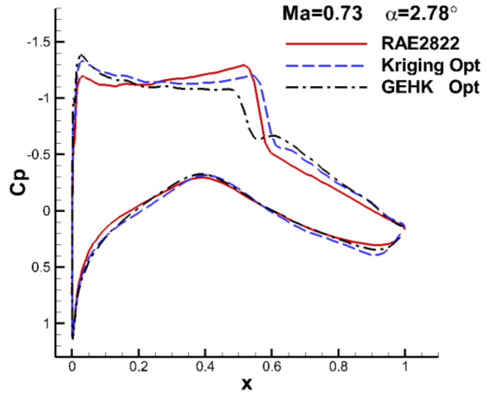
Table 2 Multi-point design result of the GEHK and standard kriging-base optimization

Models	Conditions	Lift increment	Drag reduction	Area increment
	1	9.2%	15.9%	
kriging	2	8.1%	10.9%	-0.48%
	3	10.0%	-3.2%	
	1	1.61%	26.9%	
GEHK	2	-0.51%	15.7%	-0.49%
	3	0.40%	-2.62%	

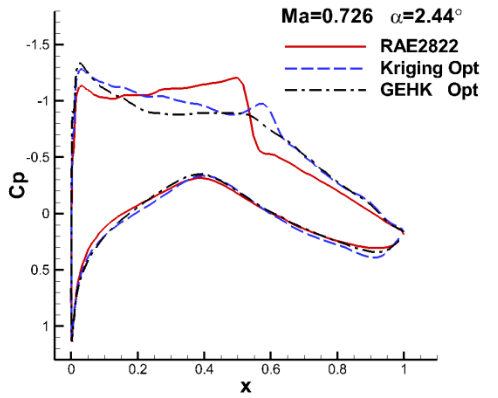
Fig. 6 presents the geometry and pressure distributions at each design points obtained by the GEHK model and kriging model. It is clear from the pressure distribution at $Ma = 0.73$ that optimizations employing the GEHK and kriging model have weaken the shock on the upper surface, and the GEHK model performs better. The shock has been eliminated at $Ma = 0.726$ by the GEHK design, and there is still a weak shock by the kriging design. The pressure distributions of both the GEHK and standard kriging design case change little at $Ma = 0.6$.



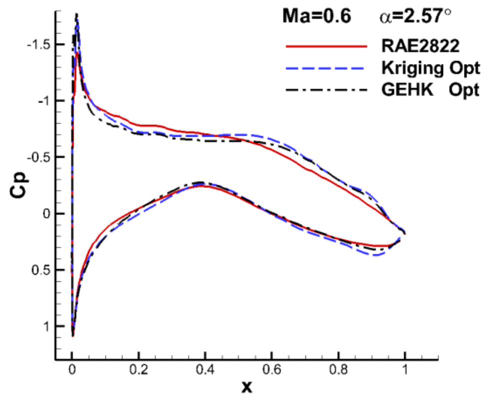
(a) Airfoil geometries of optimization result



(b) Optimized pressure distribution at the condition of 1



(c) Optimized pressure distribution at the condition of 2



(d) Optimized pressure distribution at the condition of 3

Fig. 6 Optimization result of multi-point design using the GEHK and standard kriging model

5 Conclusions

A new method combined gradients with Hierarchical Kriging model is developed in this paper. New samples are derived by Taylor approximation using gradients and selected steps. Then a low-fidelity kriging model is built using derived samples. Finally, a high-fidelity model is obtained by adjust the low-fidelity kriging with initial samples. The new method has overcome limitations of traditional gradient-based kriging model, and the prediction accuracy of the model can be improved by less samples and gradient. An optimization test of an airfoil has proved that the gradient-based Hierarchical Kriging enhanced the optimization efficiency, and got a better result, compared with ordinary kriging model and gradient-based kriging model.

6 Acknowledgments

This research is sponsored by the National Natural Science Foundation of China (NSFC) under grant No. 11272263.

References

[1] Forrester, A. I. J., and Keane, A. J. Recent advances in surrogate-based optimization. *Progress in Aerospace Sciences*, Vol. 45, No. 1-3, pp. 50-79, 2009.

[2] Jeong, S., Murayama, M., and Yamamoto, K. Efficient Optimization Design Method Using Kriging Model. *J Aircraft,* Vol. 42, No. 5, pp. 1375-1375, 2005.

[3] Simpson, T. W., Mauery, T. M., Korte, J., and Mistree, F. Kriging models for global approximation in simulation-based multidisciplinary design

optimization. *AIAA Journal,* Vol. 39, pp. 2233-2241, 2001.

[4] Simpson, T., Mistree, F., Korte, J., and Mauery, T. Comparison of response surface and kriging models for multidisciplinary design optimization. *7th AIAA/ USAF/ NASA/ ISSMO Symposium on Multidisciplinary Analysis and Optimization. American Institute of Aeronautics and Astronautics (AIAA)*, 1998.

[5] J. Toal, D. J., Bressloff, N. W., and Keane, A. J. Kriging Hyperparameter Tuning Strategies. *AIAA Journal* Vol. 46, No. 5, pp. 1240-1252, 2008.

[6] Chung, H. S., and Alonso, J. Using gradients to construct cokriging approximation models for high-dimensional design optimization problems. *40th AIAA Aerospace Sciences Meeting. American Institute of Aeronautics and Astronautics (AIAA)*, 2002.

[7] Laurenceau, J., and Sagaut, P. Building Efficient Response Surfaces of Aerodynamic Functions with Kriging and Cokriging. *AIAA Journal,* Vol. 46, No. 2, pp. 498-507, 2008.

[8] Liu, W., and Batill, S. Gradient Enhanced Response Surface Approximations Using Kriging Models. *9th AIAA/ ISSMO Symposium on Multidisciplinary Analysis and Optimization. American Institute of Aeronautics and Astronautics (AIAA)*, 2002.

[9] Han, Z.-H., and Görtz, S. Hierarchical Kriging Model for Variable-Fidelity Surrogate Modeling. *AIAA Journal* ,Vol. 50, No. 9, pp. 1885-1896, 2012.

[10] Forrester, A. I. J., Sóbester, A., and Keane, A. J. Multi-fidelity optimization via surrogate modeling. Proceedings of the Royal Society A: *Mathematical, Physical and Engineering Sciences*, Vol. 463, No. 2088, pp. 3251-3269, 2007.

[11] Jones DR, Schlonlau M, Welch W J. Efficient Global Optimization Of Expensive Black-Box Functions. *Journal Of Global Optimization*, Vol. 13, No. 4, pp. 455-492, 2008.

[12] Goto, Y., Jeong, S., Obayashi, S., and Kohama, Y. Design Space Exploration of Supersonic Formation Flying Focusing on Drag Minimization. *Journal of Aircraft,* Vol. 45, No. 2, pp. 430-439, 2008.

[13] Palacios, F., Alonso, J., Duraisamy, K., Colonno, M., Hicken, J., Aranake, A., Campos, A., Copeland, S., Economon, T., Lonkar, A., Lukaczyk, T., and Taylor, T. Stanford University Unstructured SU2: An open-source integrated computational environment for multi-physics simulation and design. *51st AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition. American Institute of Aeronautics and Astronautics (AIAA)*, 2013.

Contact Author Email Address

Chao Song, Email:songchao1900@mai.nwpu.edu.cn

Copyright Statement