Review Article

Institution of MECHANICAL ENGINEERS
Journal of AEROSPACE ENGINEERING

Proc IMechE Part G:
J Aerospace Engineering
227(9) 1516–1532

DOI: 10.1177/0954410012457891
uk.sagepub.com/jaero
SAGE

Efficient global optimization for S-duct diffuser shape design

HyoGil Bae[1], SooHyung Park[2] and JangHyuk Kwon[1]

Abstract

The efficient global optimization method is a global optimization technique based on the stochastic kriging model to efficiently search the global optimum in a design space. Efficient global optimization selects the next sample point in the view of the probability that a global minimum is located. To present the probability, the efficient global optimization method introduces the expected improvement function. The mean and variance at the untried point provided from the kriging model are used to calculate the expected improvement function. Efficient global optimization selects the maximum expected improvement point as the next sample point. After validating the efficient global optimization method by several test functions, we applied it to a diffusing S-duct shape design problem which needs a computationally expensive turbulent computational fluid dynamics analysis. The design objective is to improve the total pressure recovery of the S-duct. The improved S-duct shape was searched globally through the efficient global optimization method. Our results confirmed that the efficient global optimization method can efficiently provide a meaningful engineering result in the S-duct shape design.

Keywords

Efficient global optimization, expected improvement, kriging model, S-duct

Date received: 10 February 2012; accepted: 11 July 2012

Introduction

The diffusing S-duct is a typical aircraft inlet geometry. It decelerates air entering the compressor and increases the static pressure. The Lockheed Martin F-16 has a chin-type S-duct and Korean advanced supersonic trainer T-50 has a bifurcated S-duct.

A well-designed diffusing S-duct can help to enhance engine performance by delivering a nearly uniform flow with a small transverse velocity component at the engine compressor entrance. But the centerline curvature associated with the S-duct determines the centrifugal force and cross-stream pressure gradient between the bottom and top of the S-duct. As a result, the diffusing S-duct integrates the complex three-dimensional (3D) internal flow which includes secondary flows and flow separation. Wellborn et al.[1] performed a compressible subsonic flow experiment through a representative S-duct configuration and investigated the complicated flow mechanism in detail. The baseline diffuser of our design corresponds to the experimental configuration of Wellborn et al.

The aerodynamic performance of subsonic diffusers is principally determined by two parameters; the flow distortion at the outlet of the diffuser and the total pressure recovery (shortly PR). If the flow distortion is reduced and the PR is close to one, it can be considered to be well designed.

To estimate the performance characteristics of an S-duct without experiments, computational studies have been done with different turbulent models. Harloff et al.[2] used the PARC3D[3] computer program to solve the full, 3D Reynolds-averaged Navier–Stokes

[1]Department of Aerospace Engineering, Korea Advanced Institute of Science and Technology, Republic of Korea
[2]Department of Aerospace Information Engineering, Konkuk University, Republic of Korea

Corresponding author:
JangHyuk Kwon, Department of Aerospace Engineering, Korea Advanced Institute of Science and Technology, 291 Daehak-ro, Yuseong-gu, Daejeon, Republic of Korea.
Email: jhkwon@kaist.ac.kr

equations. Harloff et al. computed a subsonic S-duct flow using the Baldwin–Lomax algebraic turbulence model and the k–ε turbulent model. Zhang et al.[4] performed computational analysis of the S-duct diffuser by utilizing the Baldwin–Lomax turbulent algebraic model. Details are presented by Zhang.[5] Lee and Kim[6] analyzed the S-duct computationally and concluded that the k–ω shear stress transport (SST) two-equation turbulent model predicts the flow characteristics better than Spalart–Allmaras and k–ω models.

For the optimization method of the S-duct shape, the sequential quadratic programming[4,7] and the adjoint equations[6] were used. Both are gradient-based optimizers. The multi-objective genetic algorithm was also used as a global search method.[8]

Generally, gradient-based optimization methods provide an accurate local minimum in the design space, but such optimization requires sensitivity information to calculate the search direction. Using the adjoint equations,[6] the sensitivity can be computed at once irrespective of the number of design variables. However, applying the adjoint equation to a specific analysis code is still limited for engineers to use in the field. The genetic algorithm also has the demerit of requiring a lot of evaluating functions.

Thus, surrogate-based optimization has become a noteworthy method[9,10] nowadays. With a surrogate model, the global minimum can be searched for by looking at and investigating the design space, and the surrogate model can be constructed with a few analysis runs. Another advantage is that the previously analyzed data can be reused to build a surrogate model later after it has been stored. However, the difficulty of surrogate-based optimization is in selecting the next sample point that can cover the exploitation and exploration searches in the design space.

For this reason, Jones et al.[11] developed the efficient global optimization (EGO) method based on the stochastic kriging model. EGO selects the next sample point that has the maximum probability that the global minimum exists. The probability is expressed by the expected improvement (EI) function, which consists of the mean value and the variance of the kriging model. The uncertainty (= variance) of the kriging model at untried points is considered as not a risk but an opportunity to search for a good solution. Because the EGO method needs a minimal number of analysis runs to search a global minimum, the EGO method can be a good way for optimization when the analysis is computationally expensive such as in a computational fluid dynamics (CFD) analysis.

The objective of this study is to apply the EGO method to the diffusing S-duct shape design and to investigate the usefulness of the EGO method.

Efficient global optimization

Kriging model

The EGO method is based on the stochastic kriging model. The true function $Y(\mathbf{x})$ is assumed to be comprised of a regression model $\mathbf{f}(\mathbf{x})^{\mathrm{T}}\boldsymbol{\beta}$ and stochastic departure model $Z(\mathbf{x})$. The regression model covers the design space globally but the departure model expresses the design space locally[12]

$$Y(\mathbf{x}) = \mathbf{f}(\mathbf{x})^{\mathrm{T}}\boldsymbol{\beta} + Z(\mathbf{x}) \tag{1}$$

where $\mathbf{x}$ is the design variable vector, $\mathbf{f}(\mathbf{x})$ the vector of regression functions, and $\boldsymbol{\beta}$ the vector of unknown coefficients.

It can be assumed that departure terms are correlated with each other when doing a computer experiment because the computer experiment is deterministic and does not have measurement error and noise. In the kriging model, the covariance between departures $Z(\mathbf{s}_i)$ and $Z(\mathbf{s}_j)$ at sample points is assumed to be expressed as the product of process variance σ_z^2 and correlation matrix $\mathbf{R}$

$$\mathrm{cov}[Z(\mathbf{s}_i), Z(\mathbf{s}_j)] \equiv \sigma_z^2\mathbf{R} \tag{2}$$

The correlation matrix consists of the spatial correlation function R such as equation (3). When the distance between the points is large, the function will be zero. On the other hand, when the distance between the points is small, the function will become near one

$$\mathrm{R}(\boldsymbol{\theta}, \mathbf{s}_i, \mathbf{s}_j) = \exp\left[-\sum_{k=1}^{\mathrm{DV}} \theta_k |s_{i,k} - s_{j,k}|^2\right] \tag{3}$$

where $i, j = 1, \ldots, n$ are indices of the number of sample points, s the coordinate of sample point, k the number of design variables and θ the correlation parameter.

Using the properties of the best linear unbiased predictor, the kriging model is completed. The mean square error (MSE) of prediction is minimized by introducing Lagrange multipliers and unbiasedness constraints. Thus, the kriging model $\hat{\mathrm{Y}}(\mathbf{x})$ is derived in equation (4). The kriging model also is called design and analysis of computer experiments (DACE) predictor[13]

$$\hat{\mathrm{Y}}(\mathbf{x}) = \mathbf{f}(\mathbf{x})^{\mathrm{T}}\hat{\boldsymbol{\beta}} + \mathbf{r}(\mathbf{x})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{Y} - \mathbf{F}\hat{\boldsymbol{\beta}}) \tag{4}$$

$$\mathbf{r}(\mathbf{x}) = [R(\boldsymbol{\theta}, \mathbf{x}, \mathbf{s}_1), R(\boldsymbol{\theta}, \mathbf{x}, \mathbf{s}_2), \ldots, R(\boldsymbol{\theta}, \mathbf{x}, \mathbf{s}_n)]^{\mathrm{T}} \tag{5}$$

$$\hat{\boldsymbol{\beta}} = (\mathbf{F}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{F})^{-1}\mathbf{F}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{Y} \tag{6}$$

where $\mathbf{r}(\mathbf{x})$ is the correlation vector between the untried point $\mathbf{x}$ and the sample point $\mathbf{s}$, $\hat{\boldsymbol{\beta}}$ the generalized least square estimation, and $\mathbf{F}$ the matrix of the regression function. The MSE of the predictor also is derived in equation (7)

$$\mathrm{MSE}[\hat{Y}(\mathbf{x})] = \sigma_z^2\left(1 - \mathbf{r}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{r} + \mathbf{u}^{\mathrm{T}}\left(\mathbf{F}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{F}\right)^{-1}\mathbf{u}\right) \quad (7)$$

$$\sigma_z^2 = (\mathbf{Y} - \mathbf{F}\hat{\boldsymbol{\beta}})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{Y} - \mathbf{F}\hat{\boldsymbol{\beta}})/n \quad (8)$$

$$\mathbf{u} = \mathbf{F}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{r}(\mathbf{x}) - \mathbf{f}(\mathbf{x}) \quad (9)$$

With equations (8) and (9), the correlation parameter vector $\boldsymbol{\theta}$ is still unknown. In the kriging model, the maximum likelihood estimation (MLE) is used to estimate the correlation parameters. Because the response $\mathbf{Y}$ is assumed to have Gaussian distribution, the likelihood function to determine the correlation parameter is defined as equation (10)

$$L \equiv \frac{1}{(2\pi)^{n/2}\sqrt{|\sigma_z^2\mathbf{R}|}}\exp\left[\frac{-(\mathbf{Y} - \mathbf{F}\hat{\boldsymbol{\beta}})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{Y} - \mathbf{F}\hat{\boldsymbol{\beta}})}{2\sigma_z^2}\right] \quad (10)$$

Finally, the MLE is converted to a minimization problem of non-logarithm form, as equation (11).[12] The correlation parameter is determined by minimizing this likelihood function

$$\min_{\theta} \sigma_z^2|\mathbf{R}|^{1/n} \quad (11)$$

In this study, the kriging code implementation was based on the Simpson[14] FORTRAN program and had sophisticated numerical techniques (Cholesky decomposition, QR factorization, the nugget effect, and avoiding underflow).[15]

Kriging model validation

Before we use the kriging model for design, we assess the model accuracy. The kriging model is not a regression model but interpolation, thus R^2 validation method is not useful. A few additional points must be selected as validation samples. If an expensive CFD run is needed to obtain the additional data at the validation point, it is not a practical procedure. However, the cross-validation method allows us to assess the model accuracy without sampling any points beyond those used to fit the kriging model.

In this study, the correlation parameters are computed using all sample points n. The cross-validated prediction of $Y(\mathbf{x}_i)$ is denoted by $\hat{Y}_{-i}(\mathbf{x}_i)$ and the sample point i is not used in computing the correlation matrix $\mathbf{R}$ and the vector $\mathbf{r}$. Since the number of samples is n, we perform the cross-validation test n times. If the kriging model is valid, the line of actual function versus the cross-validated prediction should be plotted on a 45° line.

In addition to the cross-validated prediction at $\mathbf{x}_i$, the cross-validated MSE at $\mathbf{x}_i$ is also calculated and denoted by $\sigma_{-i}^2(\mathbf{x}_i)$. If the kriging model has to be approximately 99.7% confident that $Y(\mathbf{x}_i)$ lies in $\pm 3\sigma_{-i}$, it is convenient to use the standardized cross-validated residual equation (12)

$$\frac{Y(\mathbf{x}_i) - \hat{Y}_{-i}(\mathbf{x}_i)}{\sigma_{-i}(\mathbf{x}_i)} \quad (12)$$

If the kriging model is valid, the value should be roughly in the interval $[-3, +3]$.[11]

EGO method

The next sample point is determined considering the predictor and the MSE (= variance) which are provided from the kriging model. In equation (13), the improvement function is defined to express the probability that the next sample would be less than the current minimum. The current minimum $f_{\min}$ is obtained based on the sampled points so far

$$I = \max[(f_{\min} - Y(\mathbf{x})), 0] \quad (13)$$

$$Y(\mathbf{x}) \sim \mathrm{N}(\hat{Y}(\mathbf{x}), \hat{\sigma}^2(\mathbf{x})) \quad (14)$$

where $Y(\mathbf{x})$ is the true function assuming the normal distribution random variable model. By the probability technique, the expected value of the improvement function is calculated as given in equation (15)

$$E[\mathrm{I}(\mathbf{x})] = \left(f_{\min} - \hat{Y}\right)\Phi\left(\frac{f_{\min} - \hat{Y}}{\hat{\sigma}}\right) + \hat{\sigma}\phi\left(\frac{f_{\min} - \hat{Y}}{\hat{\sigma}}\right) \quad (15)$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ denote the standard normal cumulative distribution function (CDF) and the probability density function (PDF). The first term is the difference between the current minimum and the predicted value multiplied by the probability that the random variable $Y(\mathbf{x})$ is smaller than $f_{\min}$. The first term is large where $\hat{Y}$ is likely smaller than $f_{\min}$. The second term tends to be large where there is a high uncertainty (= variance). Thus, the EGO method selects the next sample point that has the maximum EI value for both regions of improvement and regions of high uncertainty.[16] Then, the kriging model is updated with the selected next sample point. As EGO samples the next point close to the optimum, EI is gradually reduced.

Validation of EGO method

Before applying EGO to the S-duct shape design, we must perform validation step with several test functions. Branin, Matlab peak and Mystery functions are 2D problems. Hartmann-3 and 6 are 3D and 6D ones, respectively. All computations started with the standardized data, where the mean is zero and the variance is one in each coordinate direction.

For stop criteria, Jones et al.[11] proposed that EGO should be terminated when the value of the EI function is less than 1% of the current best minimum. Particularly, in this study, the minimum of all sample points is used as the 'current best minimum'. Additional criteria are considered at the same time. First, the relative error between the optimum based on all sample points and the optimum based on the kriging model is less than 1%. Second, the relative error of the optimum based on sampled points is less than 5%. The relative error of the optimum based on sampled points is calculated only if the function value of the sampled point is less than the 'current best minimum'.

When the three criteria are met simultaneously, EGO is terminated. Figure 1 shows the flow chart of EGO.

In Figures 2 to 4, the red circle points indicate the initial design of experiment (DOE) and 'x' denotes the sampled points by EGO methods.

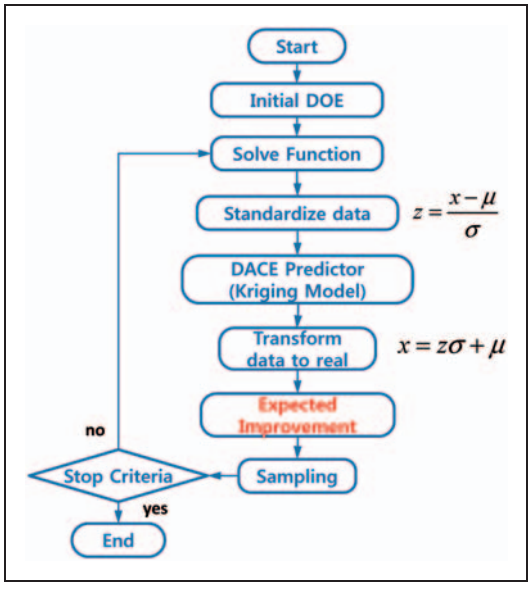

Figure 1. EGO flow chart.
DOE: design of experiment; DACE: design and analysis of computer experiments.

Branin function

$$f = \left(x_2 - \frac{5.1x_1^2}{4\pi^2} + \frac{5x_1}{\pi} - 6\right)^2 + 10\left(1 - \frac{1}{8\pi}\right)\cos x_1 + 10$$
$$-5 \leqslant x_1 \leqslant 10 \quad 0 \leqslant x_2 \leqslant 15$$
$$f(-\pi, 12.275) = f(\pi, 2.275) = f(3\pi, 2.475) = 0.397887$$

The Branin function has three global minima in the design space that are identical but have different locations as in Figure 2. EGO could search the global minimum just after 11 samplings that meet all criteria. However, intentionally 18 samplings were made in order to show that the EGO can search *three* global minima successfully. The detailed results are listed in Table 1.

Matlab peak function

$$f = 3(1-x_1)^2 e^{(-x_1^2-(x_2+1)^2)} - 10\left(\frac{x_1}{5} - x_1^3 - x_2^5\right)e^{(-x_1^2-x_2^2)} - \frac{1}{3}e^{(-(x_1+1)^2-x_2^2)}$$
$$-3 \leqslant x_1, x_2 \leqslant 3$$
$$f(0.2282, -1.6255) = -6.55113$$

The Matlab peak function can be obtained easily by typing *peaks* in the window of MATLAB program. This function has two big upper and lower peaks as shown in Figure 3. Just after 13 samplings, the global minimum was searched for meeting all criteria.

Mystery function

$$f = 2 + 0.01\left(x_2 - x_1^2\right)^2 + (1-x_1)^2 + 2(2-x_2)^2 + 7\sin(0.5x_1)\sin(0.7x_1x_2)$$
$$0 \leqslant x_1, x_2 \leqslant 5$$
$$f(2.5044, 2.5778) = -1.4565$$

The Mystery function has two local minima and one global minimum in the design space, as in Figure 4. The center point (2.5,2.5) of the initial DOE is intentionally excluded because that point is located close to the true global minimum. After 19 samplings, not the

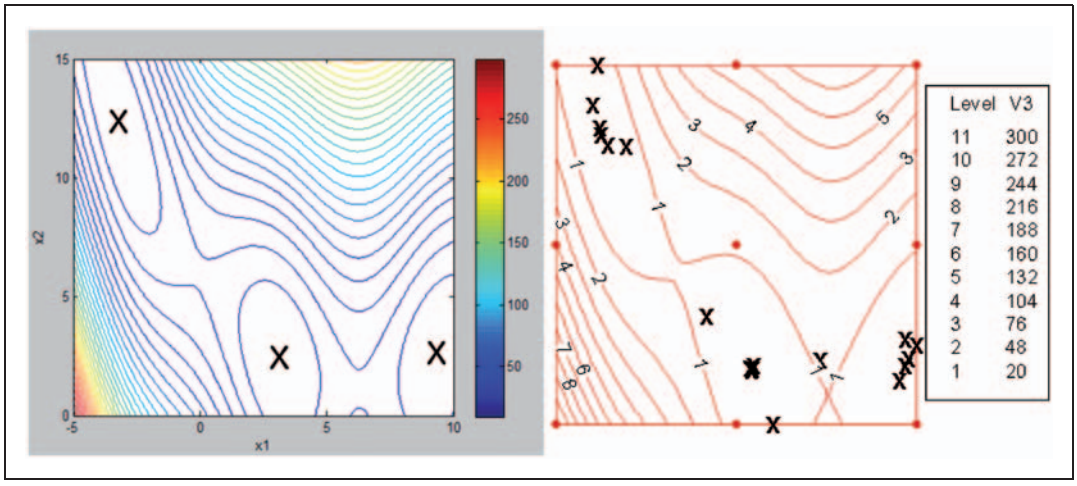

Figure 2. Branin function true solution (left) vs. EGO solution (right).

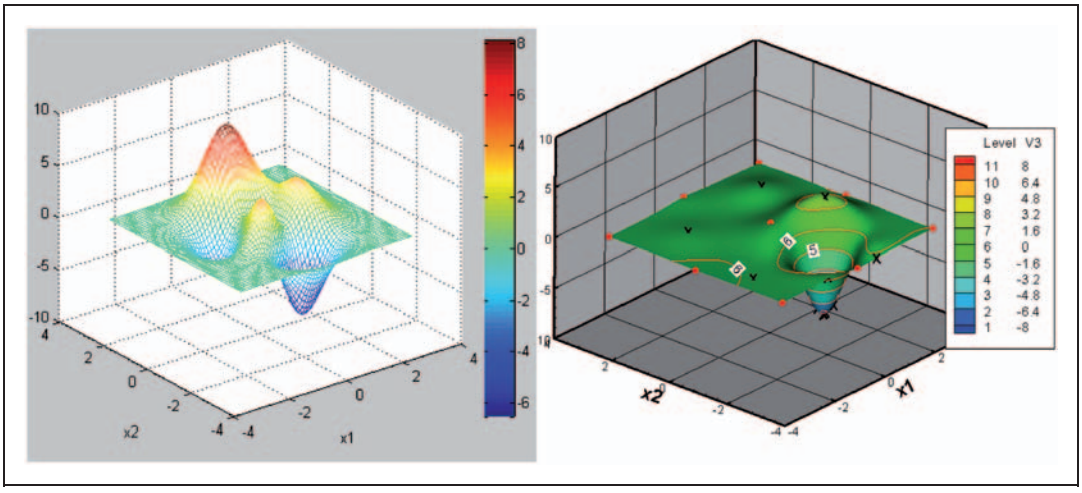

Figure 3. Matlab peak function true solution (left) vs. EGO solution (right).

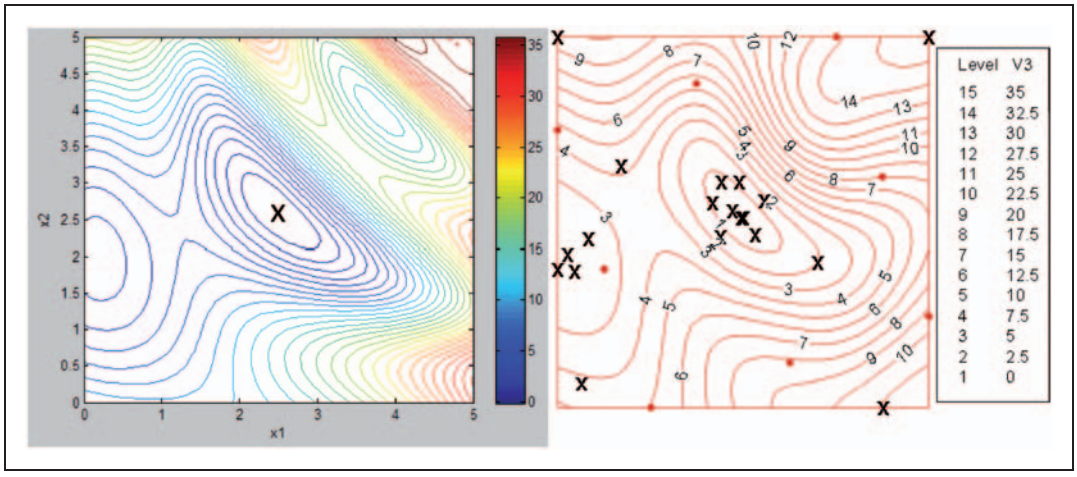

Figure 4. Mystery function true solution (left) vs. EGO solution (right).

Table 1. Test function results for the EGO method.

	Branin	Matlab peak	Mystery	Hartmann-3	Hartmann-6[D]	Criteria
True optimum	0.39789	−6.55113	−1.45650	−3.86278	−1.20068	
Optimum by EGO	0.39807	−6.55038	−1.45290	−3.85672	−1.18475	
Error (%)	0.04	−0.01	−0.25	−0.16	−1.33	
No. of initial DOE	9	9	8	33	65	
No. of sampling	18[C]	13	19	5	80	
EI	4.00719E−008	0.002291	0.000920	0.002540	0.008182	<1% of current best minimum[E]
[A]	0.002959	0.002031	0.005972	0.003158	0.012487	<0.01
[B]	0.036645	0.007361	0.006978	0.000567	0.015702	<0.05

EGO: efficient global optimization; DOE: design of experiment; EI: expected improvement.
[A] Relative error between the optimum based on sampled points and the optimum based on the kriging model; abs((Samp_opt - Krig_opt)/Samp_opt). [B] Relative error of the optimum based on sampled points. [C] intentional 18 iterations to search three global minima. [D] -ln(-y) transformed. [E] 'Current best min' is the minimum of all sample points.

local minimum but the global minimum was searched for meeting all criteria.

Hartmann $H_{3,4}$ *function*

$$H_{3,4}(x) = -\sum_{i=1}^{4} \alpha_i \exp\left[-\sum_{j=1}^{3} A_{ij}(x_j - P_{ij})^2\right]$$

$$\alpha = [1, 1.2, 3, 3.2]^{\mathrm{T}}, \quad A = \begin{bmatrix} 3.0 & 10 & 30 \\ 0.1 & 10 & 35 \\ 3.0 & 10 & 30 \\ 0.1 & 10 & 35 \end{bmatrix}$$

$$P = 10^{-4} \begin{bmatrix} 6890 & 1170 & 2673 \\ 4699 & 4387 & 7470 \\ 1091 & 8732 & 5547 \\ 3815 & 5743 & 8828 \end{bmatrix}$$

$$0 \leqslant x_i \leqslant 1, i = 1, 2, 3$$

$$x^* = (0.114614, 0.555649, 0.852547)$$

$$f(x^*) = -3.86278$$

The Hartmann-3 function (http://www-optima.amp.i.kyoto-u.ac.jp/member/student/hedar/Hedar_files/TestGO_files/Page1488.htm) was considered to validate the EGO method with three design variables. The initial DOE was constructed with 33 Latin hypercube design (LHD) points following the '$10k$' rule.[11] The term k indicates the number of design variables. As shown in Figure 5, the cross-validation and standardized cross-validated residual were plotted. The points were aligned along the 45° line and were laid down between [−3,+3] standardized deviation except just one deviated point. Including the deviated sample point, EGO searched for the minimum just after five samplings satisfying all criteria. The EI and function value converged as shown in Figure 6.

Hartmann $H_{6,4}$ *function*

$$H_{6,4}(x) = -\sum_{i=1}^{4} \alpha_i \exp\left[-\sum_{j=1}^{6} B_{ij}(x_j - Q_{ij})^2\right]$$

$$\alpha = [1, 1.2, 3, 3.2]^{\mathrm{T}},$$

$$B = \begin{bmatrix} 10 & 3 & 17 & 3.5 & 1.7 & 8 \\ 0.05 & 10 & 17 & 0.1 & 8 & 14 \\ 3 & 3.5 & 1.7 & 10 & 17 & 8 \\ 17 & 8 & 0.05 & 10 & 0.1 & 14 \end{bmatrix}$$

$$Q = 10^{-4} \begin{bmatrix} 1312 & 1696 & 5569 & 124 & 8283 & 5886 \\ 2329 & 4135 & 8307 & 3736 & 1004 & 9991 \\ 2348 & 1451 & 3522 & 2883 & 3047 & 6650 \\ 4047 & 8828 & 8732 & 5743 & 1091 & 381 \end{bmatrix}$$

$$0 \leqslant x_i \leqslant 1, i = 1, 2, 3, 4, 5, 6$$

$$x^* = (0.20169, 0.150011, 0.476874, 0.275332, 0.311652, 0.6573)$$

$$f(x^*) = -3.32237, -\ln(-f(x^*)) = -1.20068$$

The Hartmann-6 function (http://www-optima.amp.i.kyoto-u.ac.jp/member/student/hedar/Hedar_files/TestGO_files/Page1488.htm) was considered to validate the EGO method with six design variables. At first the Hartmann-6 function should be log-transformed as $-\ln(-y)$ by investigating the cross-validation test and the initial DOE was 65 points.[11] As shown in Figure 7, the cross-validation and standardized cross-validated

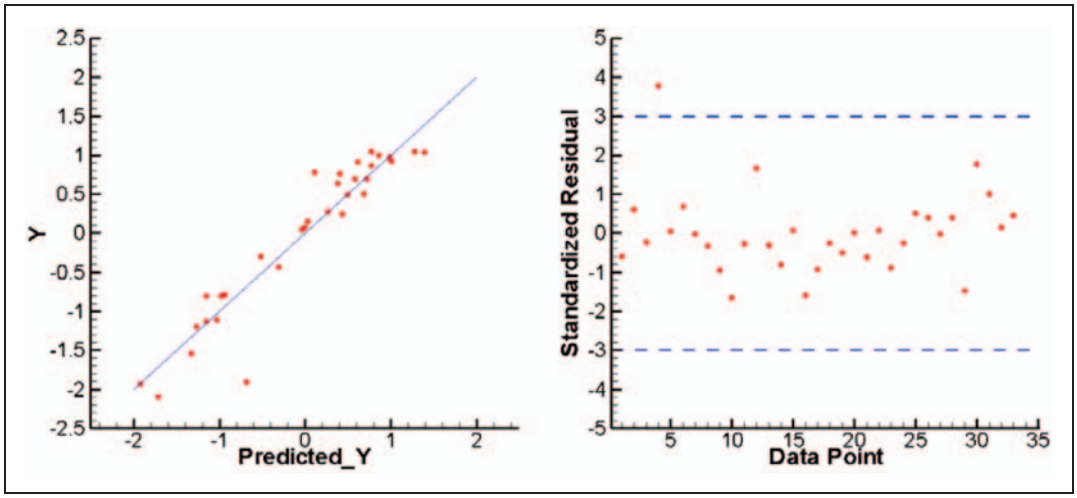

Figure 5. Cross-validation and standardized residual of Hartmann-3.

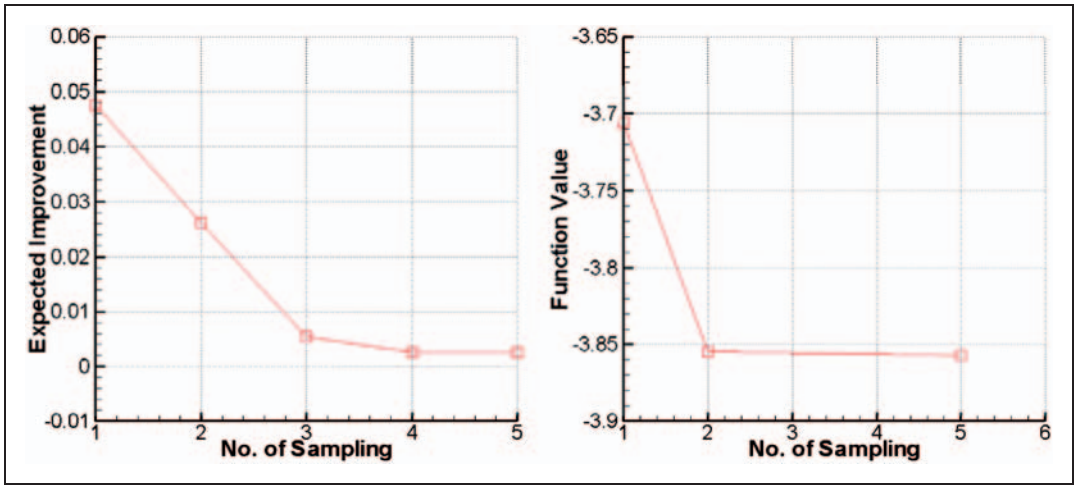

Figure 6. History of EI and function value of Hartmann-3.

residual were plotted. The points were spread out along the 45° line more than the Hartmann-3 case even though the points were laid down between [−3,+3] standard deviation. After running EGO, the EI curve was reduced with fluctuation and the function value was not converged completely even after 80 samplings (Figure 8). The EI criteria proposed by Jones et al. was met but [A] and [B] criteria were not satisfied, as can be observed from Table 1.

In summary, the EGO method with three or less design variables searched the global minimum very efficiently. However, as the number of design variables increases, establishment of an accurate design space becomes difficult because it is complex to find the correlation parameters for all variables in the kriging model. The EGO method is appropriate to the problem which has a small number of design variables. But the screening method can be applied to reduce the number of design variables before starting optimization.[17] In this study, therefore, design optimization for S-duct diffuser shape is performed with two design variables after a sensitivity study for five design variables.

Diffusing S-duct CFD analysis

Geometry of S-duct

The geometry of the diffusing S-duct is shown in Figure 9. The duct centerline (*cl*) is defined by two planar circular arcs with identical radii, R of 102.1 cm and subtended angles, $\theta_{max}/2$ of 30°. The centerline coordinates, given by equation (16), are indicated by the dashed line in Figure 9. All cross sections (plane A, B, C, D, and E) perpendicular to the centerline are circular. The duct inlet radius, r_1 is 10.21 cm. The duct outlet radius, r_2 is 12.57 cm, which produced an area

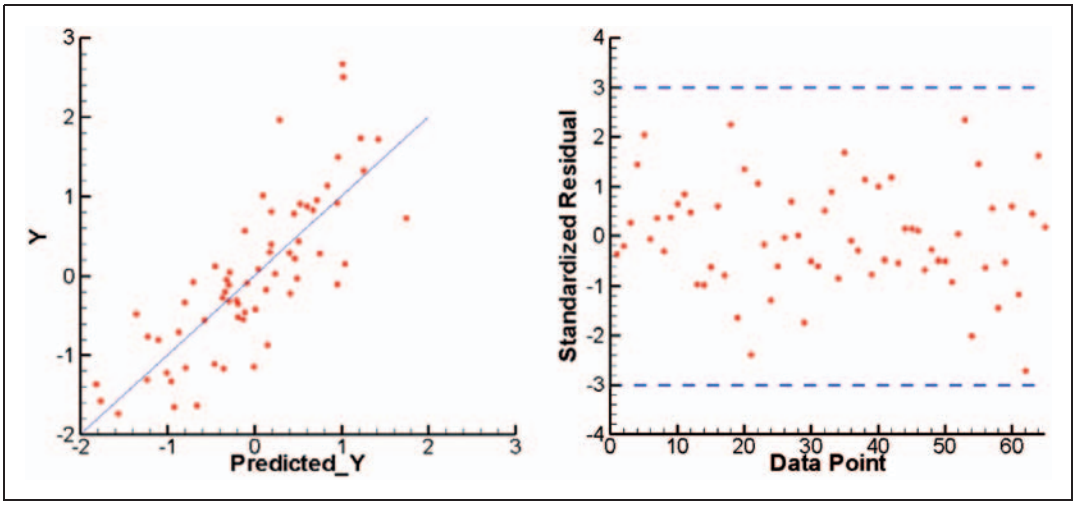

Figure 7. Cross-validation and standardized residual of Hartmann-6.

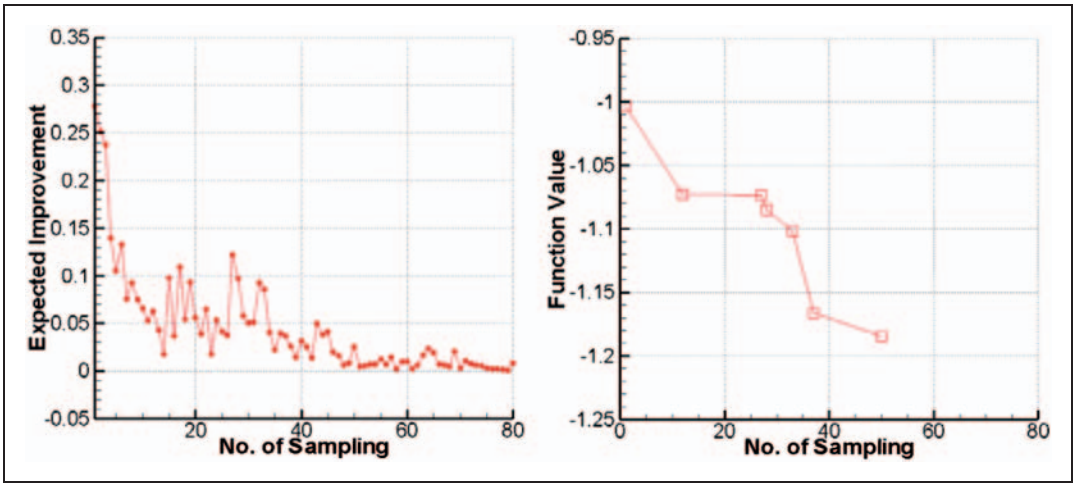

Figure 8. History of EI and function value of Hartmann-6.

ratio, A_1/A_2 of 1.52. The variation of the duct radius as a function of the angle θ is given by equation (17).

For $0 \leqslant \theta \leqslant \theta_{\max}/2$

$$x_{cl} = R\sin\theta$$
$$y_{cl} = 0$$
$$z_{cl} = R\cos\theta - R$$

For $\theta_{\max}/2 \leqslant \theta \leqslant \theta_{\max}$

$$x_{cl} = 2R\sin(\theta_{\max}/2) - R\sin(\theta_{\max} - \theta)$$
$$y_{cl} = 0$$
$$z_{cl} = 2R\cos(\theta_{\max}/2) - R - R\cos(\theta_{\max} - \theta) \quad (16)$$

$$r/r_1 = 1 + 3(r_2/r_1 - 1)(\theta/\theta_{\max})^2 - 2(r_2/r_1 - 1) \times (\theta/\theta_{\max})^3 \quad (17)$$

Grid generation

The CFD grid is generated automatically by the commercial software GridPro (www.gridpro.com). GridPro creates a multi-block structured grid, especially under the topology paradigm. That is, if the topology (consisting of corners and edges) that fills the internal space inside the S-duct surface is built by a user at one time, GridProcon forms the boundary of the S-duct surface

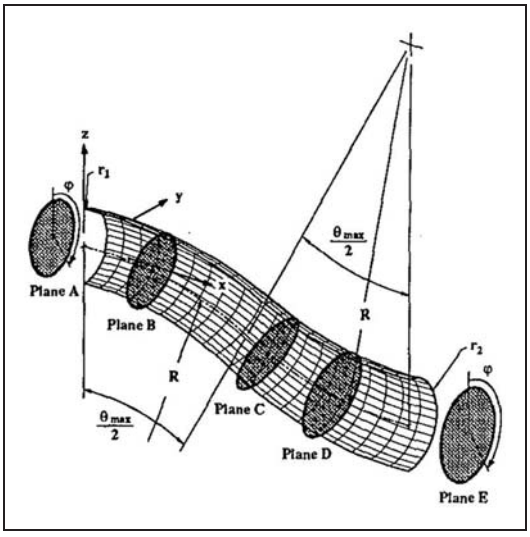

Figure 9. Baseline S-duct diffuser.[1] (Reprinted with permission of the American Institute of Aeronautics and Astronautics).

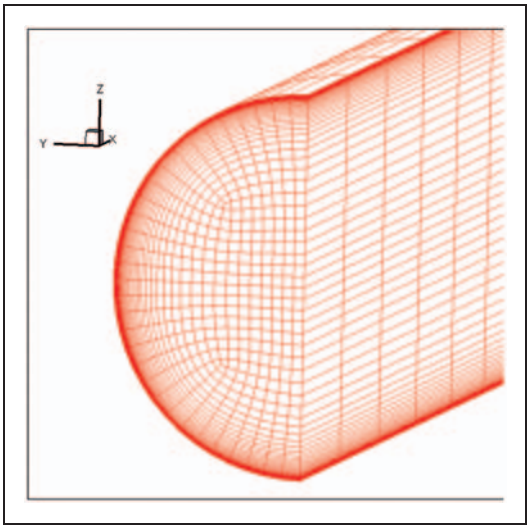

Figure 10. S-duct inlet grid.

and generates the grid automatically. A user has to only pay attention to updating the new S-duct surface data during the design optimization. The S-duct geometry was normalized by the inlet diameter and the first wall distance was 10^{-5} to the normal direction so the non-dimensional first-cell distance is set to around one. The inlet grid of the S-duct is shown in Figure 10. The total grid size was 80 cells streamwise and 32 cells circumferential and a total of 150,000 cells.

CFD analysis

S-duct CFD analysis was conducted by the in-house parallelized Reynolds-averaged Navier–Stokes equation code, KFLOW.[18] The time integration was done by the diagonalized alternating direction implicit method. Spatial numerical flux was discretized using Roe's flux difference splitting and high-ordered by third monotone upstream-centered schemes for conservation laws. The Harten entropy correction function was used to prevent the non-physical expansion wave. The k–ω SST turbulence model which has a good performance for the flow separation and secondary flow was used.[19]

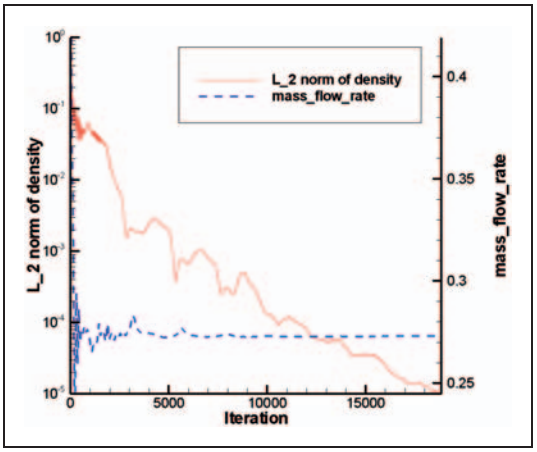

Figure 11. Density L_2 norm and outlet $\dot{m}$.

For boundary conditions, the no-slip condition was applied to the wall and the Riemann invariants were used for the inlet boundary.[5] The static pressure provided from the experiment was fixed at the entire outlet boundary surface, but that static pressure was updated to meet the target mass flow rate $\dot{m}$ at the outlet by a simple Bernoulli equation.[20]

The inlet flow conditions were the centerline Mach 0.6, Reynolds number 2.6×10^6 based on the inlet diameter and boundary layer thickness of about 7% of the radius at the reference point (plane A).[1]

A total of eight Intel x64 Quad CPU 2.93 GHz and 30 cores in those CPUs were used for the present parallel computing. It took about half an hour per one condition to converge up to 10^{-5} of the L_2 norm of density, as shown in Figure 11. The mass flow rate also converged.

Figure 12 presents the surface static pressure coefficient according to the arc length. The top, side, and bottom of the S-duct are at $\phi = 10^\circ, 90^\circ, 170^\circ$, respectively. The static pressure coefficient is defined as equation (18)

$$Cp = \frac{p_0 - p_{cl}}{p_{0,cl} - p_{cl}} \tag{18}$$

where $p_{0,cl}$ and p_{cl} are the inlet centerline total pressure and static pressure. As shown in Figure 12, the CFD

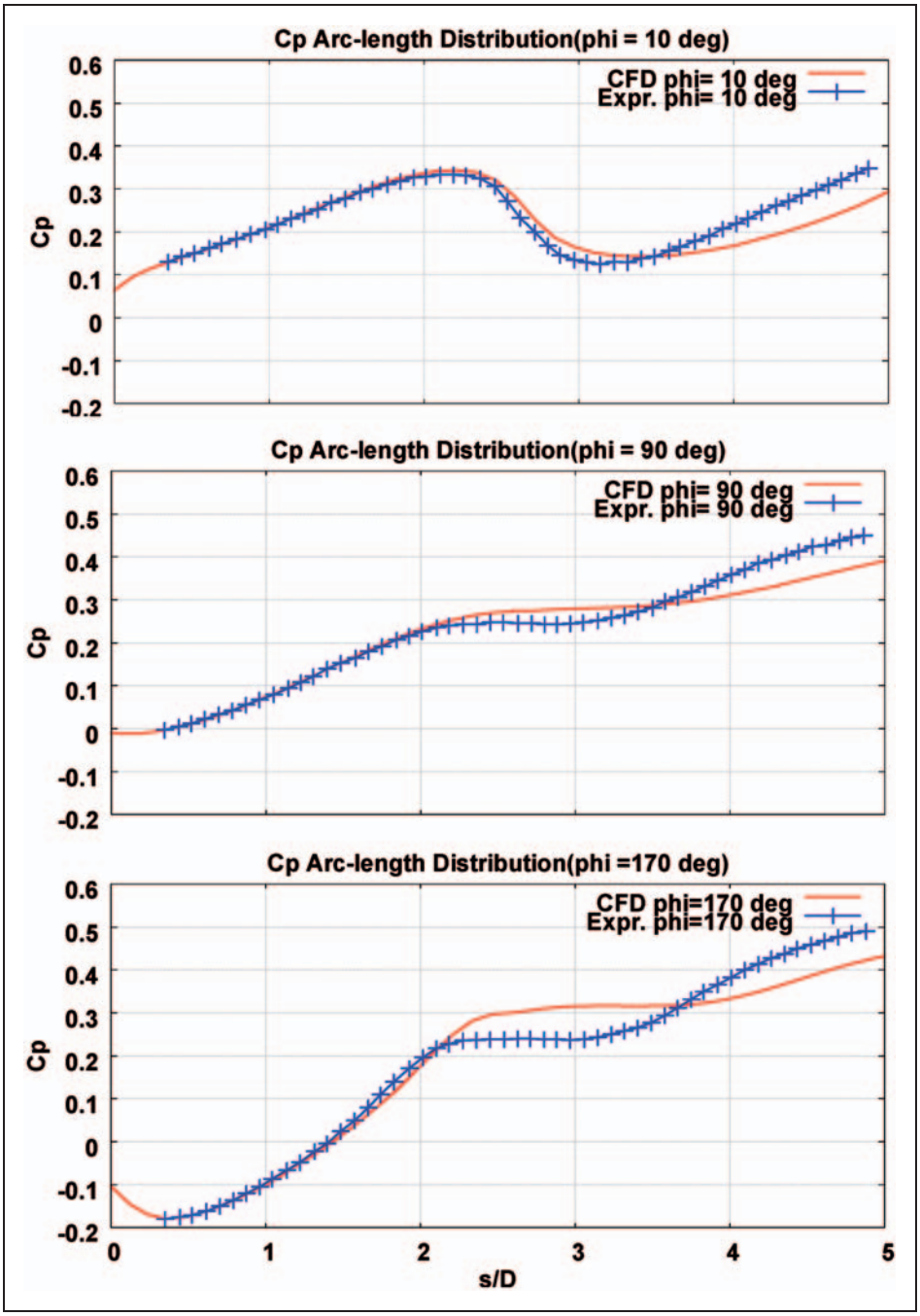

Figure 12. Arc distribution of surface static pressures for three circumferential positions.
CFD: computational fluid dynamics.

results were in good agreement with the experiments up to the first bent ($s/D = 2$) of the S-duct due to no flow separation. After passing the first bent, the flow separated and a cross flow occurred near the surface by the centrifugal force effect and diffusing effect. Therefore, the CFD results slightly differed from the experiment in the separation region, but all the trends of the pressure coefficient followed the experiment.

The surface streamline was compared with the visualized experiment qualitatively[1] and the secondary flow looks similar to the measured one as presented in Figure 13.

In conclusion, the baseline S-duct was analyzed to obtain the PR of 0.9735 and the distortion coefficient (DC60) of 0.6505. The PR and DC are defined in detail in the following section.

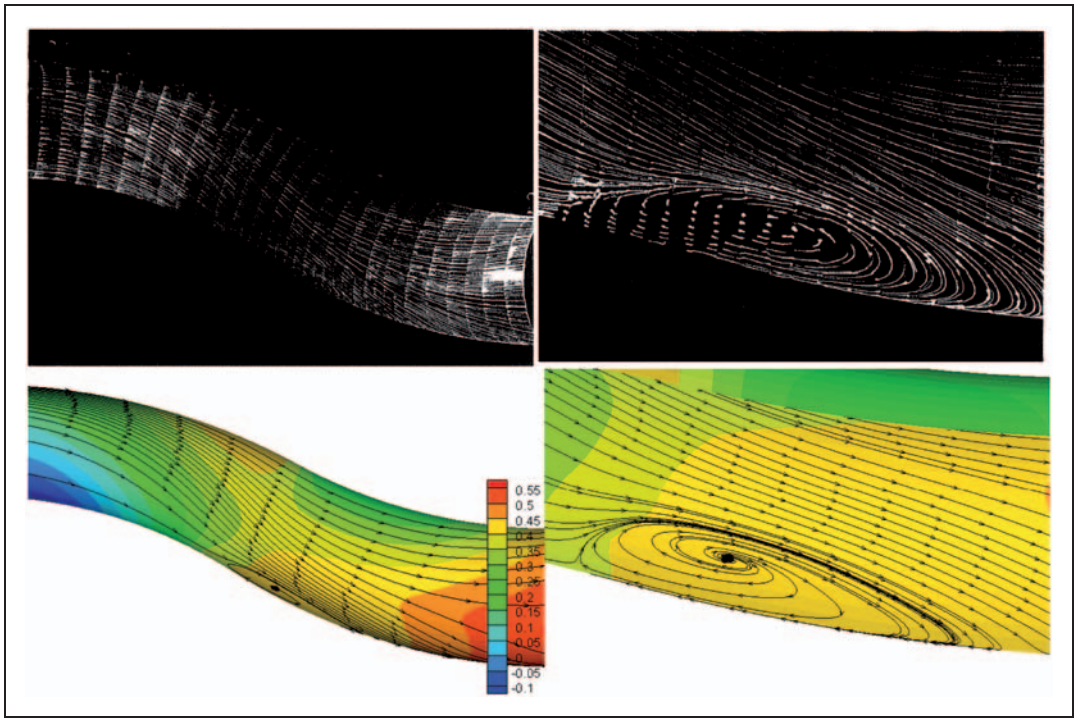

Figure 13. Measured (up) and analyzed (down) surface streamline.

Design objectives and variables

The EGO method was applied to the expensive CFD black box problem, the diffusing S-duct shape design. As the design objective functions, the PR ratio and the DC are defined in equations (19) and (20).[8] If there is little energy loss in the S-duct, PR is closer to one. The DC is a measure of the non-axis-symmetry of the outlet flow. The DC becomes small as the flow progresses uniformly

$$\mathrm{PR} = \frac{\bar{p}_{0,\mathrm{outlet}}}{\bar{p}_{0,\mathrm{inlet}}} \tag{19}$$

$$\mathrm{DC}(\psi) = \max_{0 \leqslant \phi \leqslant 2\pi} \frac{[\bar{p}_0 - \bar{p}_0(\phi, \psi)]}{\bar{q}} \tag{20}$$

$$\bar{p}_0(\phi, \psi) = \frac{\int_0^{\psi} \int_0^{R} p_0(r, \phi + \theta) r \, \mathrm{d}r\mathrm{d}\theta}{\int_0^{\psi} \int_0^{R} r \, \mathrm{d}r\mathrm{d}\theta} \tag{21}$$

where p_0 is the total pressure, $\bar{p}_0$ the mean total pressure at the outlet and $\bar{q}$ the mean dynamic pressure at the outlet. The ϕ of term $\bar{p}_0(\phi, \psi)$ is the starting angle of a pie-shaped slice of ψ over which the total pressure is integrated, as shown in Figure 14. Previous studies showed that design results were similar for different values of ψ.[4] In this study, DC60 was used with $\psi = 60^\circ$.

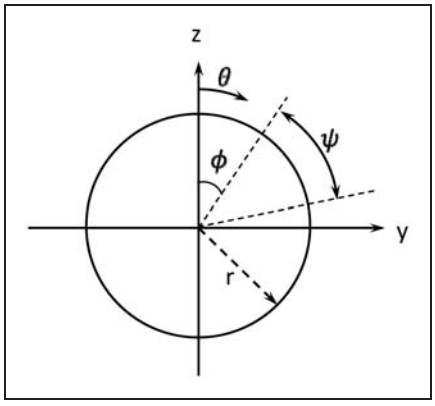

Figure 14. Sample of the pie-slice used to calculate the distortion coefficient.[8]

In this study, only the PR is maximized. The DC was calculated but not used for designing directly. Equation (22) is a simple unconstrained problem. The side constraint of design variable β_n is in equation (23)

$$\max \mathrm{PR} \tag{22}$$

$$-0.05 < \beta_n < 0.05 \tag{23}$$

The radius $r(x)$ of the S-duct was controlled by the Hicks–Henne bump function $H_n(x)$. The design variables β_n are the factors multiplied by the Hicks–Henne

$$r_{new}(x) = r_{old}(x) + \beta_n H_n(x) \quad (24)$$

$$H_n(x) = \sin^3(\pi x^{\ln(0.5)/\ln(x_n)}) \quad (25)$$

$$0 < x < 1$$
$$x_n = 1/6, 2/6, 3/6, 4/6, 5/6$$

Figure 15 shows the exaggerated shape change of the S-duct by four times the maximum design variables to account for the effect of design variable.

Design results

'Validation of EGO method' section confirmed that the EGO method searched for the global minimum efficiently with two design variables, so the S-duct diffuser was designed with two design variables. Two design variables were chosen among five design variables from the pressure recovery sensitivity. In Table 2, DV1 has the largest positive sensitivity and DV5 has the largest negative sensitivity. It is inferred that DV1 and DV2 have the similar sensitivity since they are located close to each other. Finally, DV1 and DV5 were selected as the two design variables.

In this study, EGO was coded for only minimization problems and so the sign of the objective function was converted as in equation (26).

$$\max \mathrm{PR} \rightarrow \min -\mathrm{PR} \quad (26)$$

With nine initial sample points (circle), the kriging model was generated to provide the mean and variance of the predictor, as shown in Figure 16. Just after three samplings ('x' mark), EGO searched for the global minimum with all criteria (EI, [A],[B]) satisfied. The PR was improved from $0.9735 \rightarrow 0.9880$ (+1.49%) with accompanying DC60 $0.6505 \rightarrow 0.5109$, as listed in Table 3. A total of 12 ($=9+3$) CFD runs were performed to find the global minimum.

In order to validate the EGO solution, the design space was contoured with 41 crowded sample points and this noisy response was assumed to be the true design space, as shown in Figure 17. By comparing Figure 16 to Figure 17 the global minimum searched for by EGO can be called a true global minimum. Also, the EI function and function value (pressure recovery $x - 1.0$) were converged in Figure 18.

Figure 19 compares the designed S-duct shape to the original S-duct. The duct diameter after the inlet became larger and the duct diameter in front of the outlet became smaller. The large duct diameter after the inlet prevented the static pressure drop which occurred at the back of the first bent by the centrifugal force, so the reverse flow region was removed after the design as shown on the surface streamline in Figure 20. The negative x-velocity region passing the first bent was almost removed after the design, as shown in Figure 21. That means there is only rare flow separation on the bottom of the S-duct.

Conclusively, the high total pressure area at the outlet of the S-duct became larger than the original one, and the low total pressure area became smaller

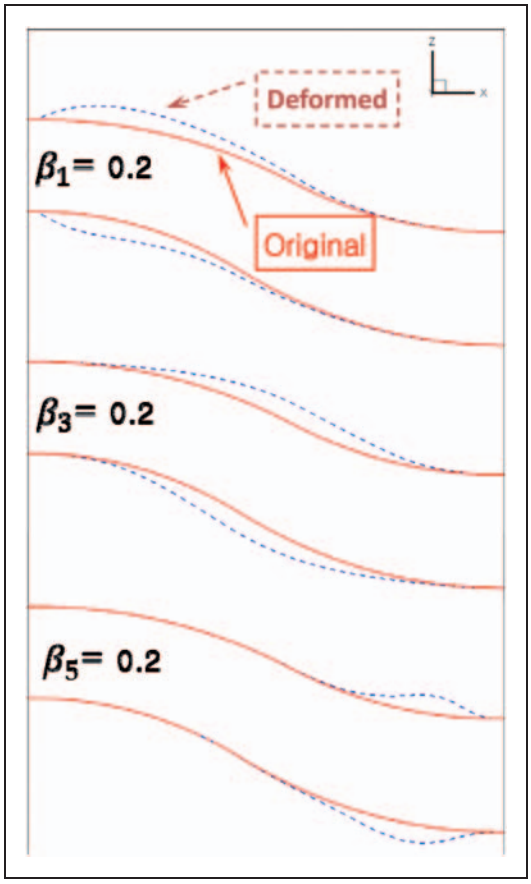

Figure 15. S-duct shape change due to design variables.

Table 2. Pressure recovery sensitivity.

	PR	Sensitivity
Baseline	0.9734	–
DV1 = 0.05 only	0.9875	0.282
DV2 = 0.05 only	0.9845	0.222
DV3 = 0.05 only	0.9789	0.110
DV4 = 0.05 only	0.9727	−0.014
DV5 = 0.05 only	0.9673	−0.122

PR: total pressure recovery.

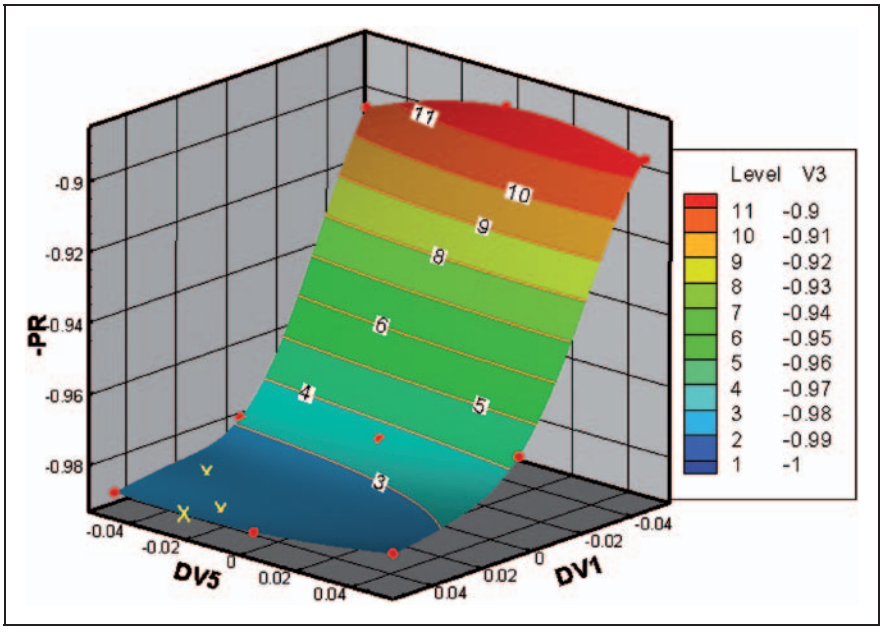

Figure 16. EGO solution of S-duct with 2DV. EGO: efficient global optimization.

Table 3. Summary of S-duct shape design results.

	2DV		
	DOT	EGO	Criteria
Original PR	0.9735		
Original DC60	0.6505		
Designed PR	0.9880	0.9880	
Designed DC60	0.5123	0.5109	
Difference of PR (%)	1.49	1.49	
Difference of DC (%)	−21.25	−21.46	
No. of initial DOE	–	9	
No. of sampling	–	3	
Total CFD runs	22	12	
EI	–	0.000089	<1% of current best min[a]
[A]	–	0.000145	<0.01
[B]	–	0.000670	<0.05
DV1	0.050000	0.050000	
DV5	−0.028627	−0.025447	

DOT: Design Optimization Tools; EGO: efficient global optimization; PR: total pressure recovery; DC: distortion coefficient; DOE: design of experiment; CFD: computational fluid dynamics; EI: expected improvement.
[A] Relative error between the optimum based on sampled points and the optimum based on the kriging model; abs((Samp_opt − Krig_opt)/Samp_opt). [B] Relative error of the optimum based on sampled points.
[a]'Current best min' is the minimum of all sample points.

than the original, as shown in Figure 22 because the energy loss due to the flow separation was reduced.

To check the EGO efficiency, local optimization was conducted using the gradient-based optimizer Design Optimization Tools (DOT), version 4.00 (www.vrand.com). The sensitivity was calculated by the finite difference method (FDM) and the initial step for the FDM was 0.0001. The Broyden–Fletcher–Goldfarb–Shanno method was used for the search direction. In Figure 23, the convergence histories of

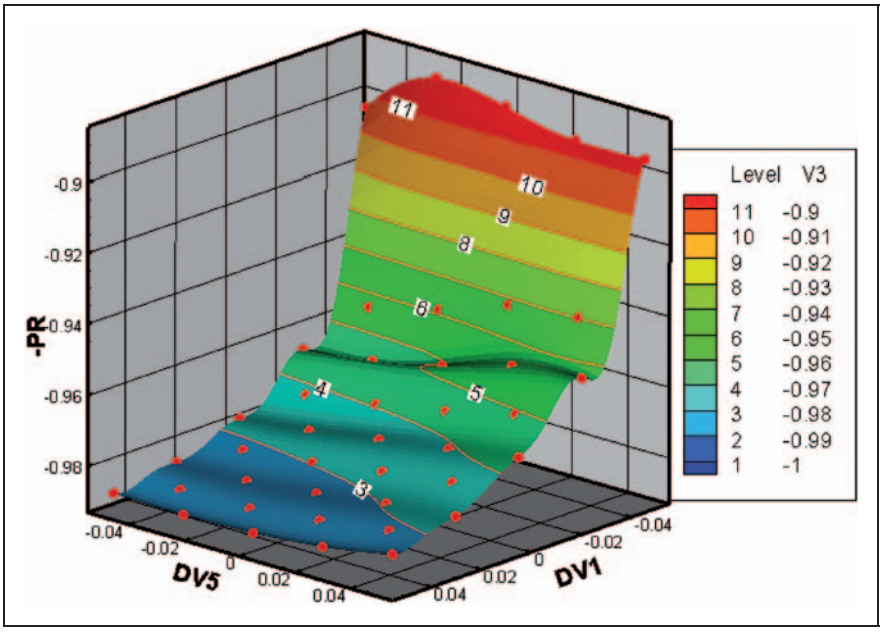

Figure 17. S-duct true design space with 2DV.

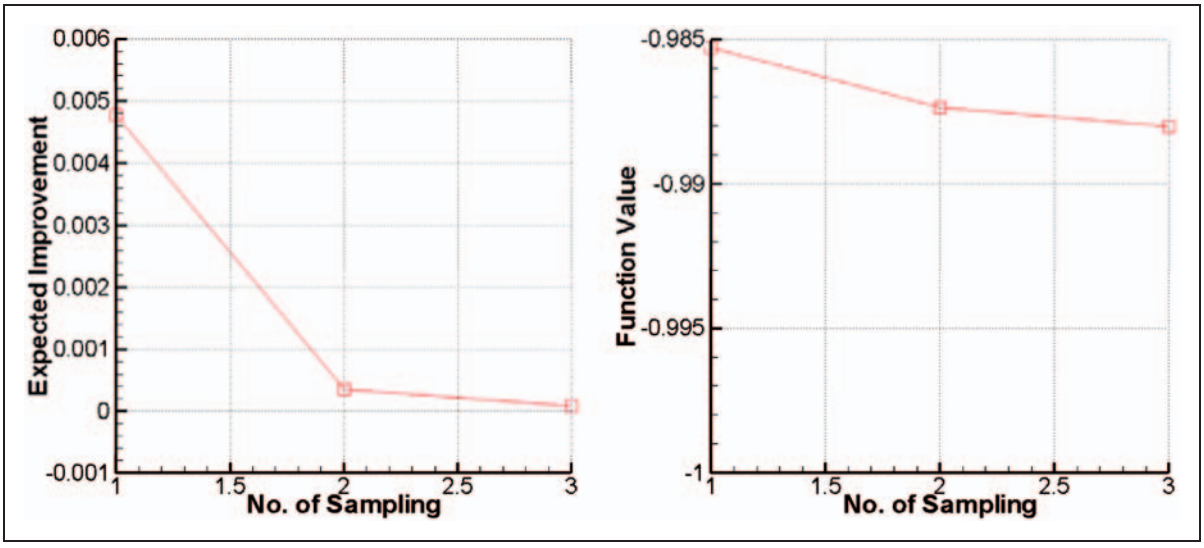

Figure 18. EI and PR of EGO with two design variables.

pressure recovery from both EGO and DOT methods were plotted according to the number of CFD runs. The unexpected peak for DOT is due to the 1D line search, and finally, the solution of DOT was converged within tolerance (relative error 1.E−3). Up to nine CFD runs, they were evaluated to prepare the initial sample data. From the 10th CFD run, EGO selected the next sample points. For DOT method, a total of 22 CFD runs were performed to converge the solution. This shows that EGO was able to search for the

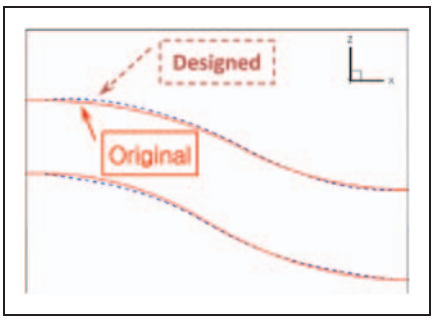

Figure 19. Designed S-duct surface shape.

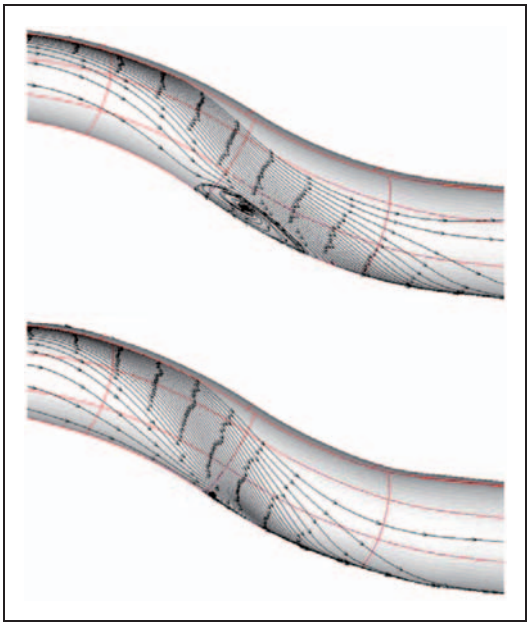

Figure 20. Original (up) and designed (down) S-duct surface: streamlines.

global minimum efficiently as the analogous iteration level of the local minimum search. The PR improved from $0.9735 \rightarrow 0.9880$ (+1.49%) with an accompanying DC60 $0.6505 \rightarrow 0.5123$, as listed in Table 3. Additionally, the final design variables (DV1, DV5) = (0.05, −0.028627) were close to those of EGO. That means the local minimum by DOT can be regarded as the global minimum in this case.

Conclusions

The EGO method was validated with several test functions (2D, 3D, 6D). Although EGO becomes ineffective to search a global minimum as the number of design variables increases, EGO could overcome the local search of gradient-based optimization and the design space realization difficulty of surrogate-based methods.

When the EGO method was applied to the S-duct shape design with two design variables, the EGO method searched for the minimum globally within the given design space just after three samplings with no need to concern about the local minima. Also, we could

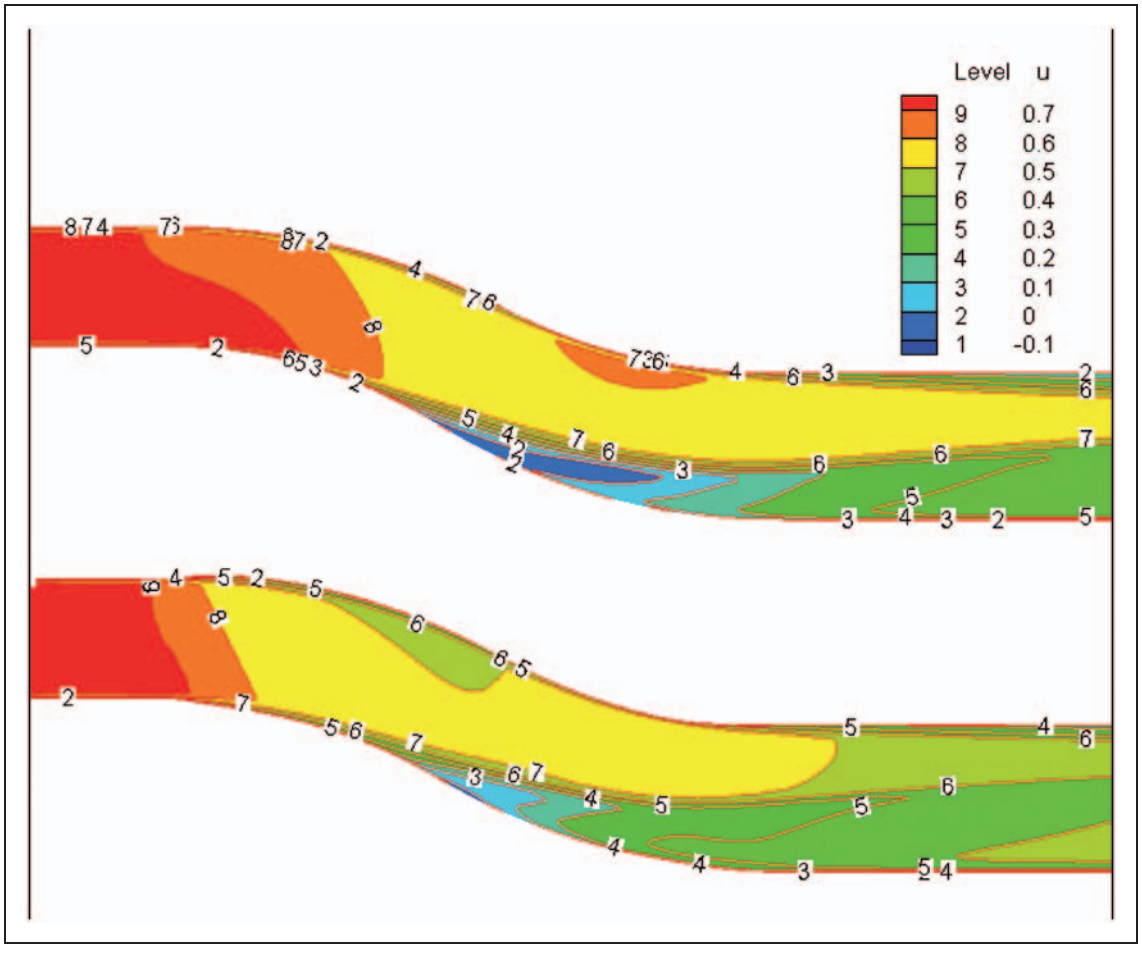

Figure 21. Original (up) and designed (down) S-duct: *U*-velocity.

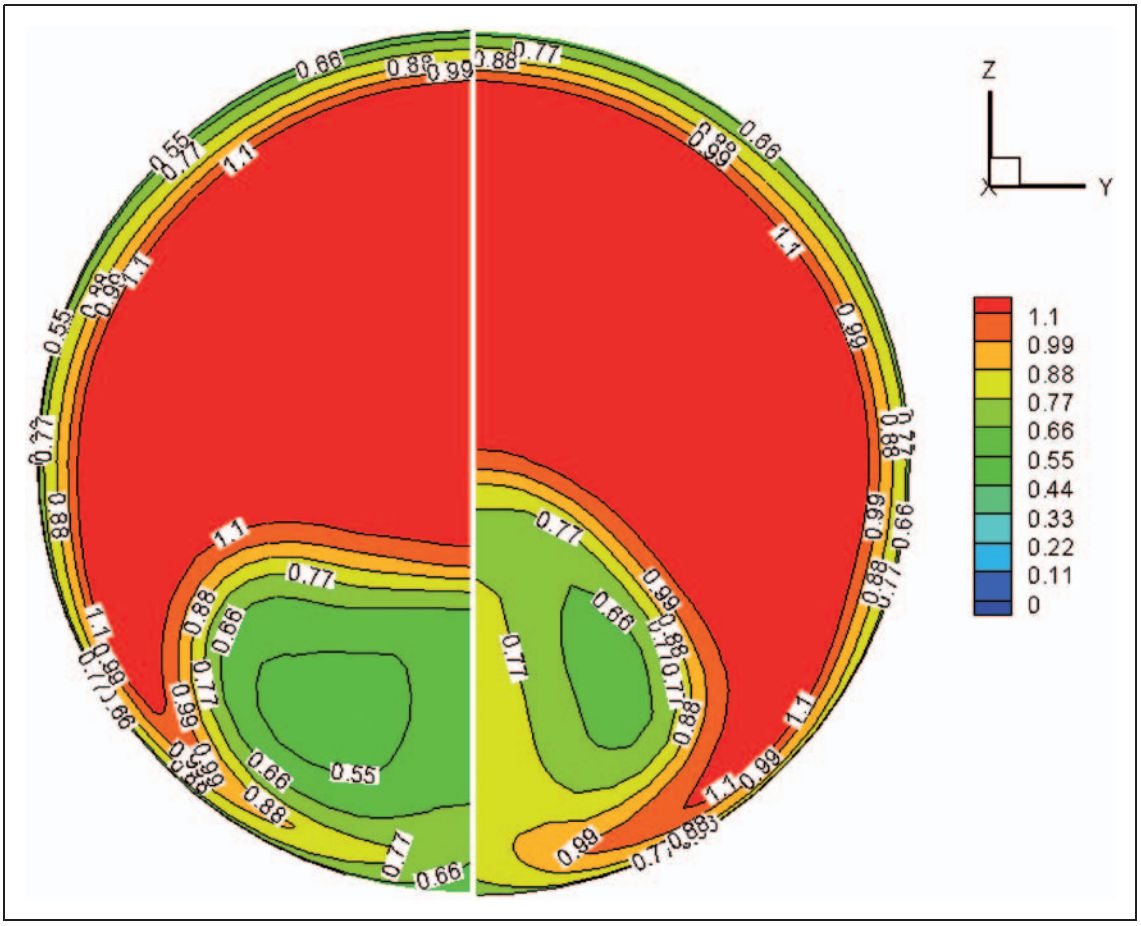

Figure 22. Original (left) and designed (right) S-duct outlet: total pressure coefficient.

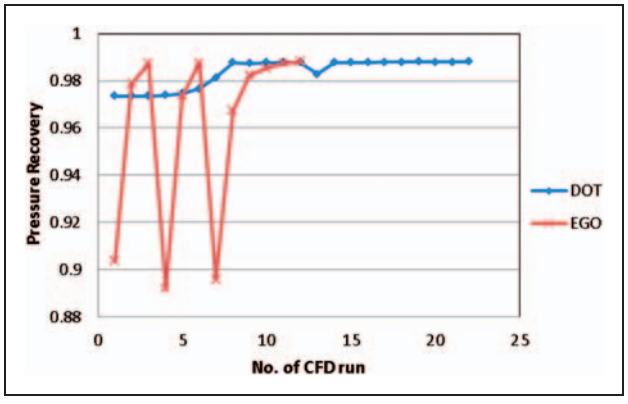

Figure 23. Pressure recovery history of EGO and DOT with 2DV. CFD: computational fluid dynamics; EGO: efficient global optimization; DOT: Design Optimization Tools.

look at the design space after EGO and assured that the searched solution is a real global minimum. The results showed that the EGO method is effective to find the solution globally even for complicated 3D internal flow design problems.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

1. Wellborn SR, Reichert BA and Okiishi TH. An experimental investigation of the flow in a diffusing S-duct. In: *AIAA/SAE/ASME/ASEE 28th joint propulsion conference and exhibit,* Nashville, TN, 6–8 July 1992, AIAA-92-3622.
2. Harloff GJ, Reichert BA and Wellborn SR. Navier-Stokes analysis and experimental data comparison of compressible flow in a diffusing S-duct. AIAA paper 92-2699, 1992.
3. Cooper GK. The PARC code: theory and usage. Arnold Engineering Development Center TR-87-24, Tullahoma, TN, October 1987.
4. Zhang W, Knight DD and Smith D. Automated design of a three-dimensional subsonic diffuser. *J Propul Power* 2000; 16(6): 1132–1140.
5. Zhang W-L. *Three-dimensional subsonic diffuser design optimization and analysis*. PhD Dissertation, Department of Mechanical and Aeronautic Engineering, Rutgers University, New Brunswick, NJ, May 1999.

6. Lee BJ and Kim C. Automated design methodology of turbulent internal flow using discrete adjoint formulation. *Aerosp Sci Technol* 2007; 11: 163–173.
7. Lefantzi S and Knight DD. Automated design optimization of a three-dimensional S-shaped subsonic diffuser. *J Propul Power* 2002; 18(4): 913–921.
8. Aranake A, Lee JG, Knight D, et al. Automated design optimization of a three-dimensional subsonic diffuser. *J Propul Power* 2001; 27(4): 838–846.
9. Morgans RC, Zander AC, Hansen CH, et al. EGO shape optimization of horn loaded loudspeakers. *Optim Eng* 2008; 9: 361–374.
10. Sasena MJ, Parkinson M, Reed M, et al. Adaptive experimental design applied to an ergonomics testing procedure. In: *ASME DETC conference*, Montreal, Canada, 29 September–2 October 2002, ASME paper DETC2002/DAC-34091.
11. Jones DR, Schonlau M and Welch W. Efficient global optimization of expensive black-box functions. *J Global Optim* 1998; 13: 455–492.
12. Lee TH and Jung JJ. Kriging metamodel based optimization, Chapter 16. In: JS Arora (ed.) *Optimization of structural and mechanical systems*. Singapore: World Scientific, 2007, pp.445–485.
13. Sacks J, Welch WJ, Mitchel TJ, et al. Design and analysis of computer experiments. *Stat Sci* 1989; 4(4): 409–423.
14. Simpson TW. *A concept exploration method for product family design: Appendix A*. PhD Dissertation, Department of Mechanical Engineering, Georgia Institute of Technology, 1998.
15. Soren NL, Hans BN and Jacob S. DACE, a Matlab Kriging Toolbox version 2.0, Technical Report IMM-TR-2002-12, Technical University of Denmark, August 2002.
16. Sasena MJ. *Flexibility and efficiency enhancements for constrained global design optimization with kriging approximations*, Chapter 4.1.2, PhD Dissertation, Department of Mechanical Engineering, University of Michigan, 2002.
17. Forrester AIJ, Sobester A and Keane AJ. *Engineering design via surrogate modelling: a practical guide*. Hoboken, NJ: Wiley, 2008.
18. Sung CH, Park SH and Kwon JH. Multigrid diagonalized-ADI method for compressible flows. In: *31st AIAA CFD conference*, 11–14 June 2001, Anaheim, CA, AIAA paper 2001-2556.
19. Park SH and Kwon JH. Implementation of $k-\omega$ turbulence models in an implicit multigrid method. *AIAA J* 2004; 42(7): 1348–1357.
20. FLUENT. FLUENT 6.3 user's guide, chapter 7.8.4.

Appendix

Notation

Cp	static pressure coefficient
DV	design variable
$f_{\min}$	minimum value of all sample points
R	correlation function
$\mathbf{R}$	correlation matrix
$\mathbf{s}$	sample point
$\mathbf{x}$	untried point
$\hat{\mathbf{Y}}$	prediction value
$\mathbf{Y}$	response vector of sample points
θ	correlation parameter
$\boldsymbol{\theta}$	correlation parameter vector
$\hat{\sigma}$	prediction standard deviation
ϕ	PDF
Φ	CDF