DESIGN EXPLORATION OF HIGH-LIFT AIRFOIL USING KRIGING MODEL AND DATA MINING TECHNIQUE

Masahiro Kanazaki*, Kentaro Tanaka[†], Shinkyu Jeong[‡], and Kazuomi Yamamoto[**]

* Japan Aerospace Exploration Agency,
7-44-1 Jindaiji-Higashi, Chofu, Tokyo, Japan
e-mail: kanazaki.masahiro@jaxa.jp
Web page: http://www.jaxa.jp/

[†] Ryoyu systems Co., Ltd.,
2-19-13, Takanawa, Minato-ku, Tokyo, Japan.
e-mail: kentaro@chofu.jaxa.jp

[‡] Institute of Fluid Science, Tohoku University
2-1-1 Katahira Aoba-ku, Sendai, Japan.
e-mail: jeong@edge.ifs.tohoku.ac.jp

[**]Japan Aerospace Exploration Agency,
7-44-1 Jindaiji-Higashi, Chofu, Tokyo, Japan
e-mail: yamamoto.kazuomi@jaxa.jp

Key words: High-lift Airfoil, Design Exploration, Data Mining

Abstract. *A multi-objective design exploration for a three-element airfoil consisted of a slat, a main wing, and a flap was carried out. The lift curve improvement is important to design high-lift system, thus design has to be performed with considered multi-angle. The objective functions considered here are to maximize the lift coefficient at landing and near stall conditions simultaneously. Kriging surrogate model which was constructed based on several sample designs is introduced. The solution space was explored based on the maximization of Expected Improvement (EI) value corresponding to objective functions on the kriging models. The improvement of the model and the exploration of the optimum can be advanced at the same time by maximizing EI value. In this study, a total of 90 sample points are evaluated using the Reynolds averaged Navier-Stokes simulation (RANS) for the construction of the kriging model. In order to obtain the information of the design space, two data mining techniques are applied to design result. One is functional Analysis of Variance (ANOVA) which can show quantitative information and the other is Self-Organizing Map (SOM) which can show qualitative information.*

1 INTRODUCTION

A civil aircraft wing is generally designed by considering about a cruise condition. On the contrary, when an aircraft lands or takes off, its wing should gain enough lift even at low-speed. In such condition, high-lift system which can increase the wing load at low-speed is required. Thus, high-lift system is one of the main interests in aircraft design due to its effect on landing/ take-off performances, and pay-load capacity of an aircraft.

One of a typical high-lift system is a multi-element wing. Flowfield around a multi-element wing has a complex physics caused by the interaction of each element. [1, 2] The interactions between the design and its physics have to be examined closely to design high efficient high-lift system.

In order to obtain the information of the relationship between the design space and the solution space for realistic design, high quality solutions have to be collected in the multi-objective design. In Ref. 3, Kriging surrogate model was introduced and perform the efficient global optimization. In Ref. 5, Analysis of Variance (ANOVA) and Self-Organizing Map (SOM) were applied to the aerodynamic design exploration. In Ref. 6, these data-mining techniques are coupled with Kriging model and high efficient design is performed. Moreover, these techniques are also applied to multi-disciplinary optimization (MDO), successfully. [7]

In authors study, Kriging surrogate model and MOGA (multi-objective GA) was applied to multi-objective design problem for a high-lift airfoil. The three-element airfoil as shown in Fig. 1 is used as a baseline setting. Generally, a slat increases the stall angle and a flap produces an upward shift in a lift curve as shown in Fig. 2 [1], thus multi-angle of attack should be considered. In this study, the multi-objective design of the three-element high lift system was defined, where objective functions are to maximize C_l at the angle of attack of 8 degree which corresponds to landing condition and 20 degree which corresponds to near stall angle and the design variables are element' settings. This study obtained many solutions which achieve higher solution than the baseline settings and Kriging surrogate models which correspond to each objective functions are constructed.

In this study, data mining techniques are applied to the sample designs which were collected previous study to obtain circumstantial information about the relation between the design space and the solution space. To obtain the quantitative information, ANOVA is applied and to obtain the qualitative information, SOM is applied. Using these results, the effect of the slat setting and the flap setting are investigated closely using RANS.
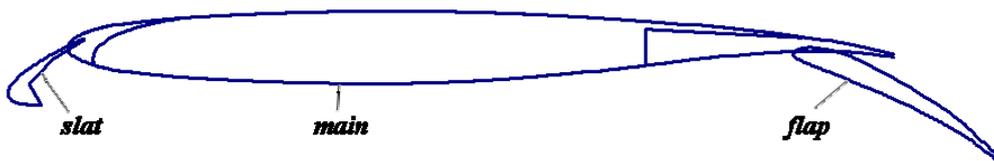
Figure 1 :Baseline airfoil and elements' settings.
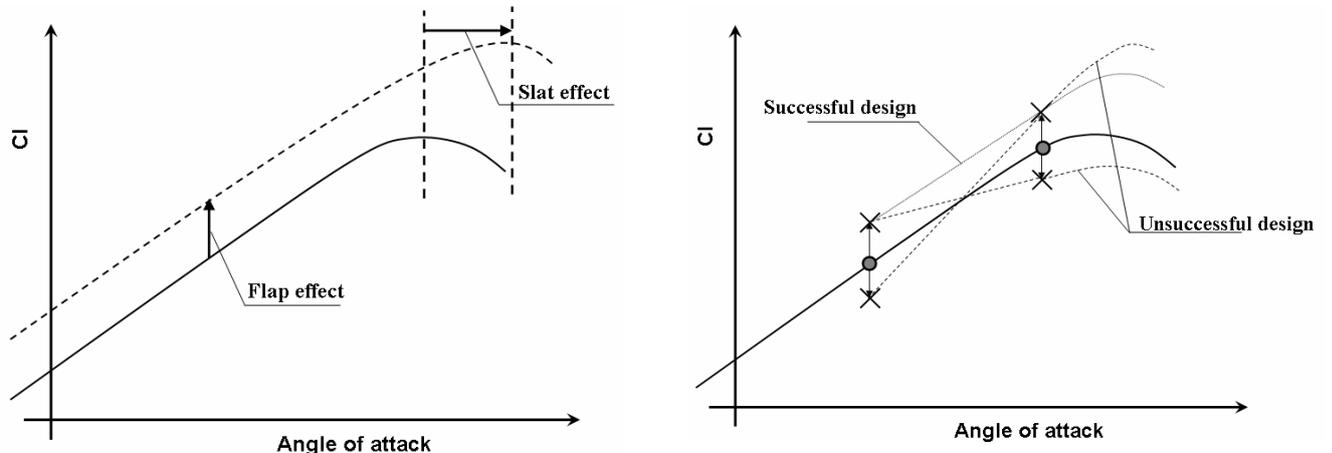
Figure2 :High-lift system effect on airfoil lift and ideal design.

2 FORMULATION

2.1 Flow Solver

Aerodynamic performances of sample designs for Kriging models are evaluated using a structured multi-block flow solver, UPACS (Unified Platform for Aerospace Computational Simulation) [9]. UPACS is developed at JAXA as a common-base code for aerodynamic researchers.

In this study, RANS is applied with Spalart-Allmaras turbulence model. Flux was evaluated by Roe's flux difference splitting with MUSCL for third-order spatial accuracy. The computational grid is decomposed into 35 sub-domains. Number of cells is about 10,000. To reduce mesh generation time, the deforming mesh method [10] is applied to deform the mesh around the baseline setting. Mach number is set to 0.2 and Reynolds number is set to 1.24×10^7.

2.2 Design Variables

As shown in Fig. 3, the overlap, the gap, and the deflection angle between elements are used as the design variables. Each design variable in limited as follows:

$$-0.01\ c \le overlap_{\text{slat}} \le 0.01\ c$$
$$0.01\ c \le gap_{\text{slat}} \le 0.04\ c$$
$$20.0 \le \theta_{\text{slat}} \le 30.0\ (\text{degree})$$
$$-0.01\ c \le overlap_{\text{flap}} \le 0.01\ c$$
$$0.01\ c \le gap_{\text{flap}} \le 0.03\ c$$
$$30.0 \le \theta_{\text{flap}} \le 40.0\ (\text{degree})$$

where c is the chord length of airfoil when flap and slat are retracted into the main element.
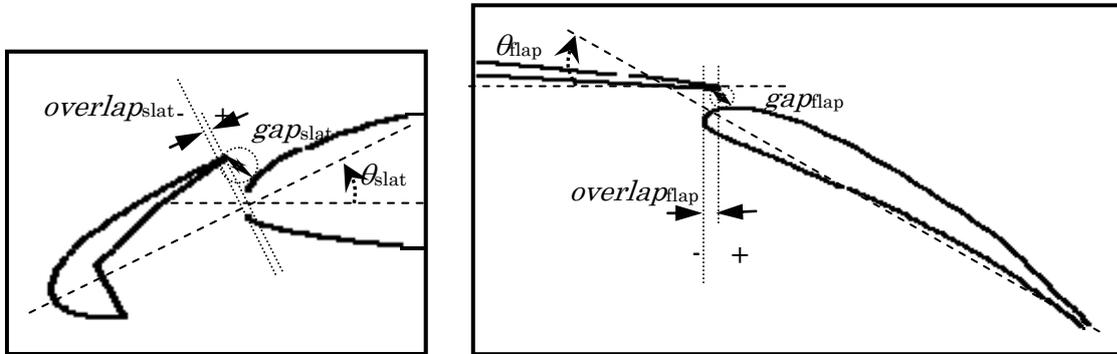
Figure. 3 Design parameters.

2.3 Objective functions

In this study, the design problem has two objective functions. The objective functions considered here are to maximize lift co-efficient at angle of attack of 8 degree ($C_{l\ 8}$) and 20 degree ($C_{l\ 20}$). Angle of attack of 8 degree is assumed the angle of attack at landing condition and 20 degree is assumed the stall angle decided from C_l -α of the baseline setting as discussed in Ref. 8.

2.4 Procedure of Multi-objective Design Exploration

The procedure of the present design (Fig. 4) is as follows: First, N samples which are decided by Latin hypercube sampling [11] which is one of the space filling methods are evaluated using RANS and Kriging surrogate models are constructed. Then, *m EI* maximum points are added as sample points, and model accuracy is improved by constructing Kriging models using $N+m$ samples. This process is iterated until improvement of objective functions becomes little. Finally, data mining technique can be applied to obtain the information of the design problem. The detail of each procedure is described in the following sections.

2.4.1   Kriging model

Kriging model [4] expresses the value $y(x_i)$ at the unknown design point x_i as:

$$y(x^i) = \mu + \varepsilon(x^i) \qquad\qquad (i = 1, 2, ..., m) \qquad\qquad (1)$$

where, m is the number of design variables, μ is a constant global model and $\varepsilon(x_i)$ represents a local deviation from the global model. The correlation between $\varepsilon(x_i)$ and $\varepsilon(x_j)$ is strongly related to the distance between the two corresponding point, x_i and x_j. In the model, the local deviation at an unknown point x is expressed using stochastic processes. Some design points are calculated as sample points and interpolated with Gaussian random function as the correlation function to estimate the trend of the stochastic process.

2.4.2   Improvement of Kriging model and selection of additional samples

Once the models are constructed, the optimum point can be explored using an arbitrary optimizer on the model. However, it is possible to miss the global optimum, because the surrogate model includes uncertainty at the predicted point. This study introduced EI values [3,4] as the criterion.

EI for present maximization problem can be calculated as follows:

$$E[I(\mathbf{x})] = (\hat{y} - f_{\max})\Phi\left(\frac{\hat{y} - f_{\max}}{s}\right) + s\phi\left(\frac{\hat{y} - f_{\max}}{s}\right)$$ (2)

where f_{max} is the maximum value among sample points and $\hat{y}$ is the value predicted by Eq. (1) at an unknown point x. Φ and are the standard distribution and normal density, respectively. EI consider the predicted function value and its uncertainty, simultaneously. Thus, the solution that has a large function value and a large uncertainty may be a promising solution. Therefore, by selecting the point where EI takes the maximum value, as the additional sample point, robust exploration of the global optimum and improvement of the model can be achieved simultaneously because this point has a somewhat large probability to become the global optimum. To apply multi-objective problem, this study considers two EI values based on two kriging models; EI_{Cl8} and EI_{Cl20}. Eq. (2) can be written for the present design problem as follows:

maximize: $$EI_{Cl8} = (\hat{y} - C_{l8\_\max})\Phi\left(\frac{\hat{y} - C_{l8\_\max}}{s}\right) + s\phi\left(\frac{\hat{y} - C_{l8\_\max}}{s}\right)$$

(3)

maximize: $$EI_{Cl20} = (\hat{y} - C_{l20\_\max})\Phi\left(\frac{\hat{y} - C_{l20\_\max}}{s}\right) + s\phi\left(\frac{\hat{y} - C_{l20\_\max}}{s}\right)$$

Maximizing these objective functions, non-dominated solutions between EI_{Cl8} and EI_{Cl20} can be obtained. Among these non-dominated solutions, three points are selected as additional sample points (Fig. 7): i) the point whose EI values of C_{l8} is maximum, ii) the mid point in the non-dominated solutions and iii) the point whose EI values of C_{l20} is maximum. Therefore, the value of m becomes 3 in this study.

2.4.3 Data mining technique

2.4.3.1 Analysis of Variance: ANOVA

An ANOVA [12] which is one of the data mining techniques is carried out to differentiate the contributions to the variance of the response from the model.

To evaluate the effect of each design variable, the total variance of the model is decomposed into that of each design variable and their interactions. The decomposition is accomplished by integrating variables out of the model $\hat{y}$. The main effect of design variable x_i is as follows:

$$\mu_i(x_i) \equiv \int \cdots \int \hat{y}(x_1, \ldots, x_n) dx_1, \ldots, dx_{i-1}, dx_{i+1}, \ldots, dx_n - \mu$$ (4)

Two-way interaction effect x_i and x_j is written as:

$$\mu_{i,j}(x_{i,j}) \equiv \int \cdots \int \hat{y}(x_1,\ldots,x_n)dx_1,\ldots,dx_{i-1},dx_{i+1}\ldots,dx_{j-1},dx_{j+1},\ldots,dx_n - \mu_i(x_i) - \mu_j(x_j) - \mu$$

(5)

where, total mean μ is as follows:

$$\mu \equiv \int \cdots \int \hat{y}(x_1,\ldots,x_n)dx_1,\ldots,dx_n$$

(6)

The variance due to the design variable x_i is

$$\varepsilon \equiv \int [\mu_i(x_i)]^2 dx_i$$

(7)

The proportion of the variance due to design variable x_i to total variance of model can be expressed as:

$$p \equiv \frac{\varepsilon}{\int \cdots \int [\hat{y}(x_1,\ldots,x_n) - \mu]^2 dx_1 \ldots dx_n}$$

(8)

The denominator of Eq. (8) means variance of the model. The value obtained by Eq. (8) indicates the sensitivity of the objective function to the variation of the design variable.

2.4.3.2 Self-organizing Map: SOM

SOM is an unsupervised learning, nonlinear projection algorithm [13] from high to low dimensional space. This projection is based on self-organization of a low-dimensional array of neurons. The weight between the input vector and the array of neurons are adjusted to represent features of the high dimensional data on low-dimensional map, in the projection algorithm. The closer two patterns are in the original space, the closer is the response of two neighboring neurons in the low-dimensional map. Thus, SOM reduces the dimension of input data while preserving their features. Using SOM, qualitative information can be obtained.

In this study, commercial software Viscovery® SOMine[14] Produced by Eudaptics GmbH is used. SOMine creates a map in a two dimensional hexagonal grid. Starting from multivariate data, the neurons on the grid gradually adapt to the intrinsic shape of the data distribution. Since the order on the grid reflects the neighborhood within the data, features of the data distribution can be read off from the emerging map on the grid. The trained SOM is systematically converted into visual information.

It is efficient to group all neurons by the similarity to facilitate SOM for the qualitative analysis, because number of neurons on the SOM is large is large as a whole. This process f grouping is called 'clustering'. Hierarchical agglomerative algorithm is used for the clustering here. First, ach node itself forms single cluster, and two clusters, which are adjacent in the map, are merged in each step. The distance between two clusters is calculated by using the SOM-ward distance. The number of clusters is determined by the hierarchical sequence of clustering. A relatively small number of clusters are used fir visualization, while a large number of clusters are used for the generation of weight vectors for respective design variables.
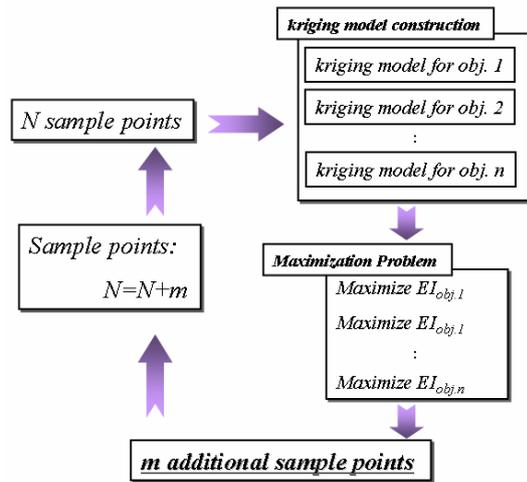
Figure4 Procedure of multi-objective global exploration.
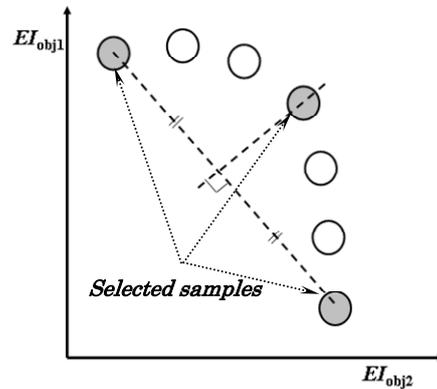
Figure5 Selection of additional samples based on EI maximization.

3   RESULTS

3.1 Design result

Figure6 shows the solutions obtained based on the present method. From this figure, the solutions obtained from the initial sampling distributed uniformly in the solution space, on the other hand, the solutions obtained from 15th-20th additional samplings achieve the better performance than that of the initial samplings. The non-dominated front gradually advances to the optimum direction as the improving process is preceded. These results show that the present method selects the additional samples properly.

3.2 Data Mining Result

3.2.1 Result of ANOVA

Total variances of models were decomposed into the variance due to each design variable. The proportion to the total variable of design variables and their interactions are shown in Fig. 16. According to Fig. 16(a), the flap setting gives over 70% effect on the $Cl8$. Moreover, according to this figure, the two-way interaction between $overlap_{flap}$ and gap_{flap} has a large effect on C_{l8}. This result suggests that $overlap_{flap}$ and gap_{flap} should be designed with considering their interaction carefully. Besides, θ_{flap} has a relative small effect because the maximum point of C_{l8} existent over the upper bound of θ_{flap} (See Fig. 8(b)). Generally the design space should be adapted in such case, however, the design space was determined based on practical use in this case. Therefore, elements' settings should design in this design space. According to Fig. 16(b), the slat and the flap setting both give effect on the C_{l20}. This result suggests that the proper setting of elements for C_{l20} is more difficult than that for C_{l8}. According to this figure, the gap of flap is also important design variable for each objective. Generally, a slat is set to increase stall angle, however, this result suggest that the flap setting has also important to the aerodynamic performance near stall condition. Not only slat but also flap should be designed carefully for near stall condition.

3.2.2   Result of SOM

To obtain quantitative information among the design space and the solution space from design results, SOM is employed. Once Kriging models are constructed, function's value at unknown points can be predicted. Using these Kriging models, the non-dominated solutions can be also obtained. Using sample points collected by the prediction of the non-dominated solutions, clustering is performed by SOM.

Figure9 (a) and (b) show SOM colored by each objective functions. In Fig (a), good C_{l8} performances are clustered in right hand side on the map and bad C_{l8} performance are clustered in left hand side. On the other hand, in Fig (a), good C_{l20} performances are clustered in left hand side on the map and bad C_{l8} performance are clustered in left hand side. This result suggests that two objective functions considered in this study have a strict trade-off.

Figure10 shows SOM colored by design variables. The SOM by θ_{flap} is thoroughly colored by red. It suggests that many solutions on the trade-off have to have highest value of θ_{flap} in the design space. The SOM by θ_{slat} is thoroughly colored by green. It suggest that θ_{slat} have to be mid-value (about 23 degree in this study) on the design space to obtain better solutions. The SOM by $overlap_{slat}$ is thoroughly colored by blue. It suggests that $overlap_{slat}$ have to be minimum value on the design space. Other maps are spotted patterns. It suggests that their design variables have interaction among other design variables.

3.2.3   Slat effect

Generally, they say that the slat has an influence on high angle of attack and the flap has an influence on low angle of attack. However, according to ANOVA result shown in 3.2.1, the interaction between the slat and the flap setting has an effect on C_{l20}. To invest the slat effect and its interaction with the flap, the slat only setting are designed by the procedure expressed in 2.4. Figure11 shows the comparison of Slat-Flap design and Slat only designs (7th samplings). According to this result, slat can only improve the lift at high angle of attack. It is agree with the general theory about high-lift airfoil. However, many solutions which obtained by Slat-Flap design achieve better C_{l20} than solutions which obtained by Slat only design. This result suggests that the flap can also improve lift at high angle of attack and they have interaction.

3.2.4   Flap effect

According to SOM result, flap deflection angle of many solutions achieving higher C_l is near upper bound (40 degree) in the design space. To invest the highest C_l obtained by flap deflection angle, the deflection angles, 40, 45, and 50 degree which out of design range are also calculated by RANS. Figure12 shows C_{l8}-flap deflection angle. According to this result, C_{l8} shows maximum value at flap deflection angle of 40-45 degree. This result suggests that the flap deflection angle should be less than 45 degree and the high-lift airfoil should stall if flap deflection angle becomes over 45 degree.
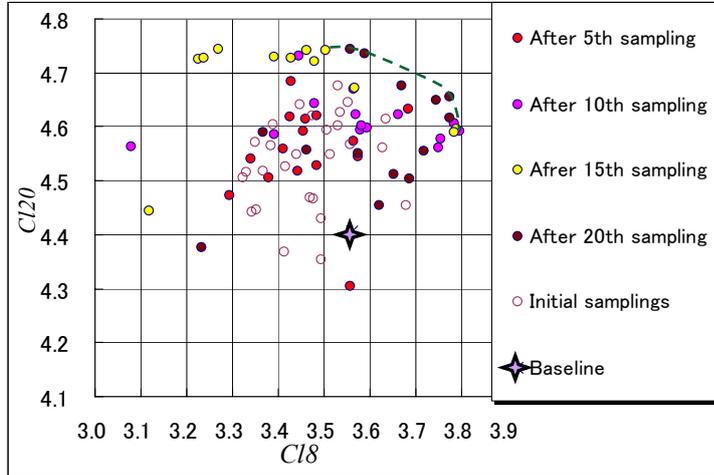
Figure6 Sample points obtained based on EI maximization.

(a)

(b)
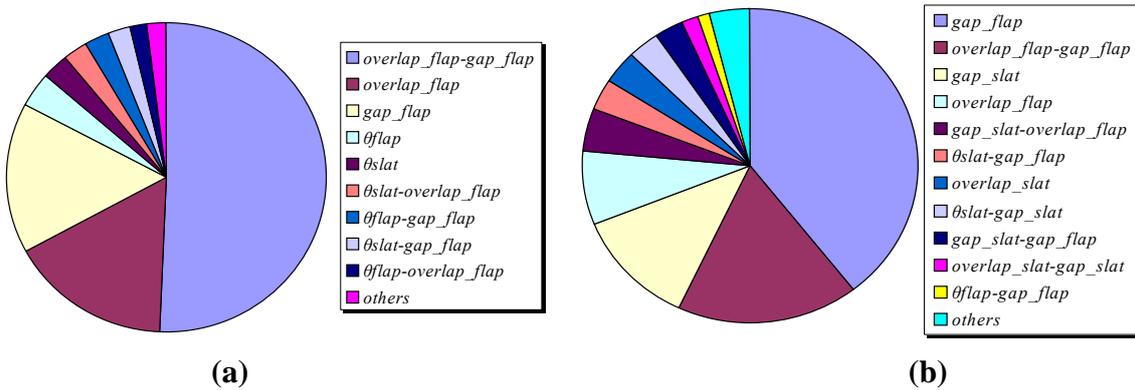
Figure8 Total proportion to the total variance of models: (a) C_{l8}, (b) C_{l20}.
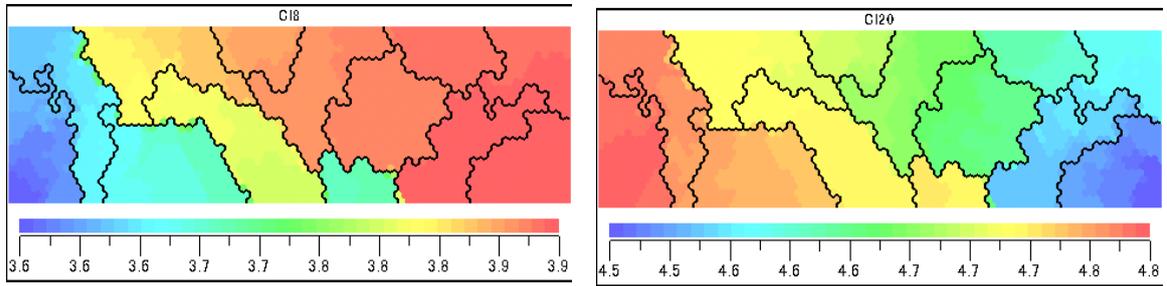
(a) **(b)**

Figure9 SOM: (a) colored by C_{l8}, (b) colored by C_{l20}.
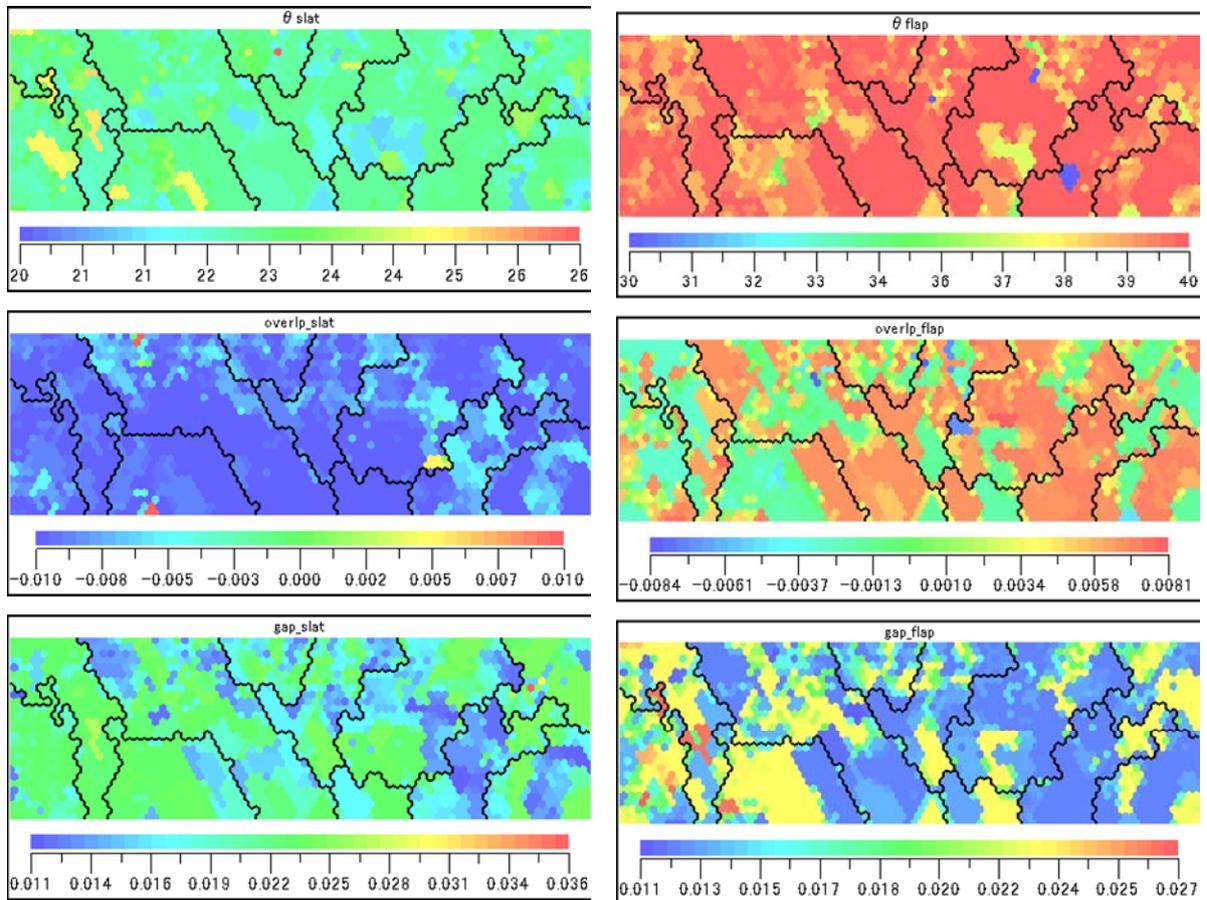
Figure10 SOM colored by design variables.
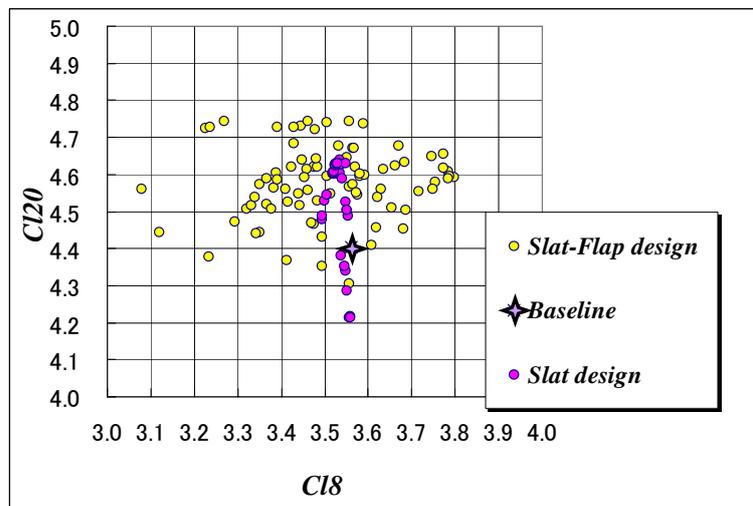
Figure11 The comparison of Slat-Flap design and Slat design.
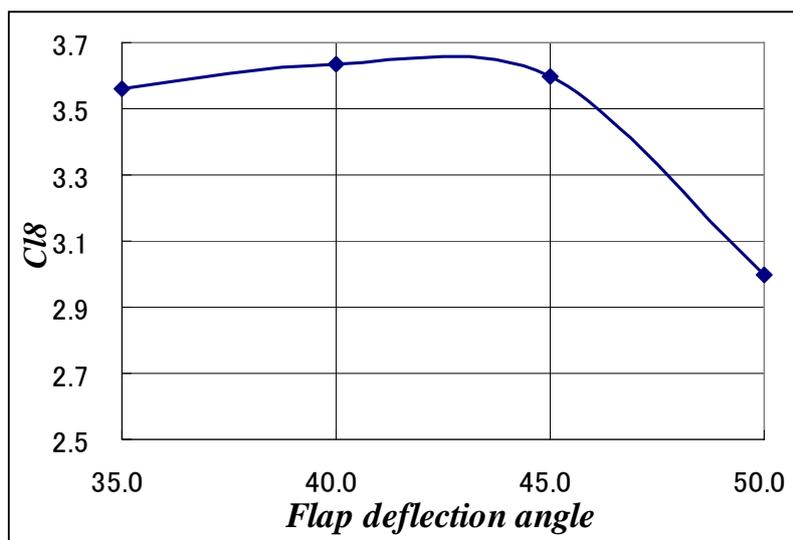
Figure12 Effect of flap deflection angle

4   CONCLUSIONS

Multi-objective design exploration for the elements' settings of the high-lift airfoil consisted of a slat, a main wing, and a flap was performed. There were two objective functions: maximizing lift coefficient at a landing condition (C_{l8}), maximizing lift coefficient near stall condition (C_{l20}). Flowfields were simulated by solving the Navier-Stocks equations with Spalart-Allmaras turbulent model using the multi-block structured grid method. The computational grids were deformed automatically for each design.

In this study, the objective functions, C_{l8} and C_{l20}, were transformed to the corresponding EI values on the kriging model and global optimization was performed based on maximizing their values. Using kriging surrogate model, the computational cost can be reduced and EI value permit to carry out high efficient design on the Kriging model. The resulting designs were also used as the additional samples to update the Kriging models.

Through the present method, the solutions based on the EI maximization advanced to the optimum direction in the solution space. As the result, element settings that give higher performance than that of baseline were successfully obtained. This result suggests that the present method can be applied to the multi-objective problem while reducing computational time drastically.

In order to obtain the information about design space, ANOVA which produces quantitative information and SOM which produces qualitative information by projecting the multi-dimensional data into two dimensional data are applied to the sampling result. This result shows the useful information for the design. From their data mining results, slat and flap effect are studied closely. According to these results, not only the slat but also the flap has to be designed carefully to obtain higher C_{l20}. To obtain higher C_{l8}, the flap deflection angle has to be decided with considering stall at the flap.

REFERENCES

[1] C. P. van Dam, "The aerodynamic design of multi-element high-lift systems for transport airplanes," *Progress in Aerospace Science*, Vol. 38, pp. 101-144, (2002).

[2] A. M. O. Smith, "High-Lift Aerodynamics," *Journal of Aircraft*, Vol. 12, No. 6, pp. 501-530, (1975).

[3] S. Jeong, M. Murayama, and K. Yamamoto,"Efficient Optimization Design Method Using Kriging Model," *Journal of Aircraft*, Vol. 42, pp.413-420, (2005).

[4] R. J.Donald, S. Matthias, and J.W. William ,"Efficient Global Optimization of Expensive Black-Box Function," *Journal of global optimization*, Vol. 13, pp. 455-192, (1998).

[5] K. Chiba, S. Obayashi, and K. Nakahashi, "Trade-off Analysis of Aerodynamic Wing Design for RLV," Proceedings of International Conference Parallel Computational Fluid Dynamics, to be appeared (2004).

[6] S. Jeong, and S. Obayashi, "Efficient Global Optimization (EGO) for Multi-Objective Problem and Data Mining," *Prooceedings of Congress on Evolutionary Computation 2005*, Vol. 3, pp. 2138-2145, (2005).

[7] T. Kumano, S. Jeong, S. Obayashi, Y. Ito, K. Hatanaka, and H., Morino,"Multidiciplinary Design Optimization of Wing Shape for a Small Jet Aicraft Using Kriging Model," AIAA 2006-932, (2006).

[8] M. Kanazaki, K. Tanaka, S. Jeong, and K. Yamamoto, "Multi-objective Aerodynamic Optimization of Elements' Setting for High-lift Airfoil Using Kriging Model," AIAA 2006-1471, (2006).

[9] R. Takaki, K. Yamamoto, T. Yamane, S. Enomoto, and J. Mukai,"The Development of

the UPACS CFD Environment," *High Performance Computing*, *Proceedings of ISHPC 2003*, *Springer*, pp. 307-319, (2003).

[10] Crumpton, P. I. and Giles, M. B.,"Implicit time accurate solutions on unstructured dynamic grids," AIAA Paper 95-1671-CP, pp. 284-294, (1995).

[11] Mckay, M. D., Beckman, R. J. and Conover, W. J., "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," *Technometric*, Vol. 21, No. 2, pp. 239-245, (1979).

[12] Sack, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P., "Design and Analysis of Computer Experiments (with Discussion)," *Statistical Science*, Vol. 4, pp. 409-435, (1989).

[13] J. C. Krzysztof, P. Witold, and W. S. Roman, *Data Mining Methods for Knowledge Discovery*, Kluwer Academic Publisher, (1998).

[14] Eudaptics software gmbh, http://www.eudaptics.com/somine/