ADAPTIVE DESIGN OF EXPERIMENT BASED ON GAUSSIAN PROCESS MODELS WITH NONSTATIONARY KERNEL APPLIED TO WIND TUNNEL TESTING

Uihwan Choi*, Junhong Kim*, Yeongbin Lee, Jaemyung Ahn***
***KAIST, **Agency for Defense Development**

Keywords: *Design of Experiment, Wind Tunnel Testing, Gaussian Process, Nonstationary Kernel, Adaptive Sampling*

Abstract

Wind tunnel testing often requires many costly experiment runs, thereby making it one of the most cost-intensive experiments known. To effectively reduce the required number of experiments while maintaining statistical accuracy of the results, many researchers have extensively studied Adaptive Design of Experiment (ADoE). Efficiency of the ADoE is highly dependent on an underlying metamodel, which is crucial for predicting responses on factors where the experiments are not yet conducted. However, most studies make use of a conventional ADoE using well-known stationary GP (ADoE-GP) as its metamodel. Although efficient, conventional ADoE has limitations in predicting responses when the experiment shows highly nonlinear traits. Therefore, in this paper, we propose an ADoE with Nonstationary Gaussian Process (ADoE-NGP) model instead of a conventional ADoE-GP. Using actual wind tunnel experiment data in a high angle-of-attack region where the responses are highly nonlinear, performance of the proposed ADoE-NGP is measured against conventional ADoE-GP in terms of the required number of experiments while satisfying the same statistical accuracy. Results indicate that the proposed ADoE-NGP show better performance results compared to conventional ADoE-GP in both number of experiment runs and accuracy.

1 Introduction

Design of Experiment (DOE) refers to methodologies that utilize statistics to select an experiment sequence that guarantees the accuracy of the whole experiment while minimizing the resources required for such experiments. In fields where available resources are limited, such as wind tunnel testing, the DoE can economically improve the efficiency and accuracy of the experiments. The study on the application of DoE in the field of wind tunnel testing has been carried out by NASA Langley Research Center, starting with the preliminary experiments led by DeLoach in the late 1990s [1]. While many traditional DoE approaches selects a set of experiment points *prior to* conducting experiments, Adaptive Design of Experiment (ADoE) selects the next experiment points sequentially *during* the experiments based on the current best metamodel constructed from the currently obtained experiment data. A prior study on the application of ADoE to wind tunnel testing [2] shows that it can effectively reduce the experiment runs while maintaining required statistical accuracy.

When applying the ADoE to wind tunnel testing, it is important to select a proper metamodel that captures the key features of the underlying aerodynamics of the aircraft. However, for many wind tunnel tests, experimenters commonly encounter difficulty gaining even a slight hint on the aircraft aerodynamics, especially in cases where the aircraft has a novel configuration or prior knowledge on the aircraft aerodynamics is unpredictable. Since wind tunnel tests tend to have such little prior knowledge, the Gaussian Process (GP) model becomes an attractive option for the underlying metamodel. The GP model is a data-driven model that does not require any fixed form of function that maps given inputs to desired outputs. However, the GP model requires prior selection of covariance function, or kernel, before applying it. Although a common choice is a stationary kernel due to its generality and simplicity, it sometimes results in prediction performance degradation for nonstationary datasets. In wind tunnel testing, aerodynamics at high angle-of-attack region is considered nonstationary, whereas low angle-of-attack region shows relatively linear traits.

In this paper, we suggest an ADoE methodology that shows improvements on wind tunnel testing, especially under nonstationary experiment regions. The proposed ADoE uses GP model as its metamodel, whose kernel is a nonstationary kernel. An algorithm to obtain the nonstationary kernel for wind tunnel testing with rapid change of response surface slope is suggested based on the gradient analysis of the traditional stationary GP model. Finally, prediction performance comparison between the proposed ADoE with nonstationary GP and that with conventional stationary GP are discussed.

2 Methodology

2.1 Overview on Three-phase ADoE

In [2], the authors suggested a three-phase ADoE composed of Initial Phase, Adaptive Phase and Confirmation Phase. In the Initial Phase, a preliminary regression model for the aerodynamic coefficients are obtained by the experiments at the N_{IP} points sampled randomly or sampled by a traditional DoE technique. In the Adaptive Phase, additional N_{AP} points are selected sequentially from current experimental points using any metamodel such as GP model to update the current regression model. Then, the reliability of the constructed models is verified in the Confirmation Phase through randomly selected N_{CP} points that has not yet been conducted in Initial Phase and Adaptive Phase.

Overall procedure for the three-phase ADoE is summarized in Figure 1. In this paper, we mainly focus on the Adaptive Phase, because the next point selection algorithm during this phase highly influences the performance of the ADoE methodology.

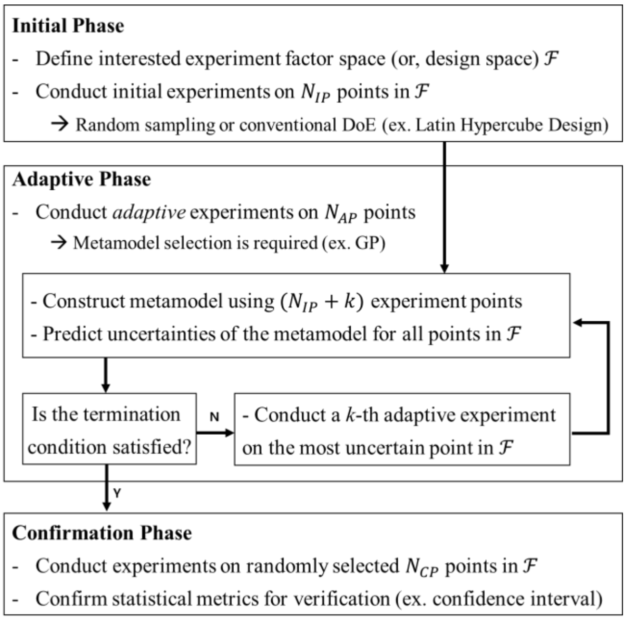

Figure 1. Overall workflow for the three-phase Adaptive Design of Experiment

2.2 Gaussian Process Regression Model

Gaussian Process Regression is a Bayesian nonparametric nonlinear regression model based on Gaussian Process (GP). The GP is a generalization of a multivariate Gaussian distribution to infinitely many variables [3]. The objective of the Gaussian Process is to learn a function $f\colon \mathbf{x} \to y$ from a given data $\mathcal{D} = \{(\mathbf{x}_1, \mathrm{y}_1), \dots, (\mathbf{x}_\mathrm{N}, \mathrm{y}_\mathrm{N})\} = \{\mathbf{X}, \mathbf{y}\}$.
In many regression problems, the generated data includes measurement noise and GP has an advantage when the noise is Gaussian white noise. With the assumption that the measurement noise has Gaussian distribution of zero mean and standard deviation σ_n, prediction of GP model (zero mean function) at any test input $\mathbf{x}_T$ based on the given data $\mathcal{D}$ is obtained as follows

$$\begin{aligned} \mathbb{E}(\mathbf{x}_T) &= K_{TN}[K_{NN} + \sigma_n^2 I]^{-1}\mathbf{y} \\ \mathrm{Var}(\mathbf{x}_T) &= K_{TT} \\ &\quad -K_{TN}[K_{NN} + \sigma_n^2 I]^{-1}K_{NT} + \sigma_n^2 I \end{aligned} \tag{1}$$

where K_{**} is a covariance matrix with its element $K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$. Covariance function $k(\mathbf{x}_i, \mathbf{x}_j)$, or kernel, characterizes the correlation between different points in the input space, that is, $k(\mathbf{x}_i, \mathbf{x}_j) \equiv \mathbb{E}[f(\mathbf{x}_i) f(\mathbf{x}_j)]$. In order to fully define a GP model, selection of the kernel type and determination of kernel hyperparameters are required. Kernel types are divided into two categories: stationary kernel and nonstationary kernel.

2.3 Kernel of Gaussian Process

2.3.1 Stationary Kernel

Stationary kernel is designed to be a function of $\boldsymbol{r} = \mathbf{x}_i - \mathbf{x}_j$ which implies that the kernel value does not change with respect to translations in the input space. Most of the known GP kernel types such as squared exponential (SE), exponential, Matérn or rational quadratic kernel are stationary kernels. Among them, the SE kernel is the most utilized stationary kernel due to its simplicity and smoothness. For *D*-dimensional input, the Automatic Relevance Determination Squared Exponential (ARD-SE) kernel shown below is widely used as an extended form of the SE kernel in a non-isotropic input space.

$$k_S(\mathbf{x}_i, \mathbf{x}_j) = \sigma_f^2 \exp\left(-\frac{1}{2}\sum_{k=1}^{D}\frac{(x_{i,k} - x_{j,k})^2}{l_k^2}\right) \qquad (2)$$

The kernel has $D + 1$ hyperparameters to be optimized: σ_f and l_k $(k = 1,2, \dots, D)$. It is known that σ_f affects the amplitude of the correlation between two inputs and l_k reveals different length-scale for each input dimension. In the remainder of this paper, a GP model with the ARD-SE kernel is considered as a reference stationary GP model due to its wide applications in many engineering fields including terrain modeling, aerodynamics and robotics.

2.3.2 Nonstationary Kernel

In some nonlinear regression problem domains, the stationary GP model shows prediction performance degradation that cannot be ignored. This happens when the regression problem which produces a dataset for the GP model construction involves highly nonlinear, rough, or irregular outputs on a certain subset of the desired input space. These kind of problem characteristics are against the basic assumption for the stationary GP model: *stationarity* on the entire input space. Nonstationary Gaussian Process (NGP) model has been introduced in order to overcome the shortcomings by adopting nonstationary kernels as its covariance function. A typical example of nonlinear regression problem using the NGP model is terrain modeling, where traditional stationary GP model (or, Kriging model) has limitations on modeling extremely rough and irregular terrain surfaces covering a wide target area with several local mountains or cliffs [4]. Aerodynamics models obtained through wind tunnel testing may also reveal nonlinearity and irregularity on a set of certain experiment conditions and experiment factors. For example, lift coefficient at high angle-of-attack region for most of aerial vehicles is known to be highly nonlinear compared to that at low angle-of-attack region. In order to effectively capture the non-linearity, we apply the following form of the nonstationary kernel [5] which is a nonstationary version of the stationary ARD-SE kernel.

$$k_{NS}(\mathbf{x}_i, \mathbf{x}_j) = \sigma_f^2 |\Sigma_i|^{\frac{1}{4}} |\Sigma_j|^{\frac{1}{4}} \left|\frac{\Sigma_i + \Sigma_j}{2}\right|^{-\frac{1}{2}} \cdot \exp\left(-(\mathbf{x}_i - \mathbf{x}_j)^T \left(\frac{\Sigma_i + \Sigma_j}{2}\right)^{-1} (\mathbf{x}_i - \mathbf{x}_j)\right) \qquad (3)$$

Unlike the ARD-SE kernel, hyperparameters of the introduced nonstationary kernel are a set of matrix $\Sigma_i = \Sigma_i(\mathbf{x}_i)$, or local kernel matrix, for all input points under interest. Thus, to apply the nonstationary kernel appropriately, it is important to develop algorithms to define those local kernel matrices Σ_i that fully captures the characteristics of the nonlinear regression problem to be solved.

2.4 Adaptive Design of Experiment with GP

ADoE with the stationary GP (ADoE-GP) uses only stationary GP model for its metamodel and the pseudo code is presented below in Algorithm 1. The ADoE-GP is the reference ADoE algorithm to be compared with the proposed

Algorithm 1 Adaptive Design of Experiment with Stationary Gaussian Process (ADoE-GP)

inputs : Initial dataset $(\boldsymbol{X}_0, \boldsymbol{y}_0)$, measurement noise s.d. σ_n, a set of target inputs $\mathcal{X}_q = \{\mathbf{x}_{q,1}, \dots, \mathbf{x}_{q,Q}\}$
returns : mean & standard deviation for the set of target inputs $\mu_{\mathcal{X}_q}, \sigma_{\mathcal{X}_q}$

1: **for** $k = 1{:}N_{AP}$ **do**
2: $\Theta^* \leftarrow$ optimize hyperparameters of stationary kernel k_S with $\{\mathbf{X}^{(k-1)}, \boldsymbol{y}^{(k-1)}, \sigma_n\}$
3: **for** $i = 1{:}Q$ **do**
4: $\left\{\mu_{q,i}^{(k)}, \sigma_{q,i}^{(k)}\right\} \leftarrow$ prediction at $\mathbf{x}_{q,i}$ using GP with $\{\mathbf{X}^{(k-1)}, \boldsymbol{y}^{(k-1)}, k_S|\Theta^*, \sigma_n\}$
5: **end**
6: $i_{MostInfo} \leftarrow \mathrm{argmax}_i\, \sigma_{q,i}{}^{(k)}$
7: $\mathbf{x}_{next} \leftarrow \mathbf{x}_{q,i_{MostInfo}}$
8: $y_{next} \leftarrow$ **Measurement**$(\mathbf{x}_{next})$
9: $\mathbf{X}^{(k)} \leftarrow \left[\mathbf{X}^{(k-1)}, \mathbf{x}_{next}\right], \boldsymbol{y}^{(k)} \leftarrow \left[\boldsymbol{y}^{(k-1)}, y_{next}\right]$
10: **end**
11: $\mu_{\mathcal{X}_q} \leftarrow \left\{\mu_{q,1}^{(N)}, \dots, \mu_{q,Q}^{(N)}\right\}, \sigma_{\mathcal{X}_q} \leftarrow \left\{\sigma_{q,1}^{(N)}, \dots, \sigma_{q,Q}^{(N)}\right\}$

ADoE with Nonstationary GP (ADoE-NGP), whose pseudo code is given in Algorithm 2.

Main features of the ADoE-NGP are 1) local kernel matrix adaptation part (line 3-8) and 2) prediction with a set of nonstationary kernel matrices part (line 9-11). The theoretic base for the first local kernel matrix adaptation part originates from [4]. However, to apply the original idea in the context of ADoE where initially there are only a few and sparse data points, normalized gradient vector calculation (line 5) and learning rate calculation with prediction uncertainty (line 6) are newly modified and kernel matrix update (line 7) is done only one time for the algorithm to possess advantage in computation time. The second part is identical to line 3-5 in Algorithm 1 except the part where the required kernel $k_{NS}|\left(\Sigma_1, \dots, \Sigma_Q\right)$ are a set of local kernel matrices representing nonstationary smoothness along the input space instead of $k_S|\Theta^*$.

Algorithm 2 Adaptive Design of Experiment with Nonstationary Gaussian Process (ADoE-NGP)

inputs : experiment dataset $(\mathbf{X}, \boldsymbol{y})$, measurement noise standard deviation σ_n, a set of target inputs $\mathcal{X}_q = \{\mathbf{x}_{q,1}, \dots, \mathbf{x}_{q,Q}\}$
returns : mean & standard deviation for the set of target inputs $\mu_{\mathcal{X}_q}, \sigma_{\mathcal{X}_q}$

1: **for** $k = 1{:}N_{AP}$ **do**
2: $\Theta^* \leftarrow$ optimize hyperparameters of stationary kernel k_S with $\{\mathbf{X}^{(k-1)}, \boldsymbol{y}^{(k-1)}, \sigma_n\}$
3: **for** $i = 1{:}Q$ **do**
4: $\left\{\mu_{q,i}^{(k)}, \nabla\mu_{q,i}^{(k)}, \sigma_{q,i}^{(k)}\right\} \leftarrow$ prediction at $\mathbf{x}_{q,i}$ using GP with $\left\{\mathbf{X}^{(k-1)}, \boldsymbol{y}^{(k-1)}, k_S|\Theta^*, \sigma_n\right\}$
5: $\nabla\hat{\mu}_{q,i}^{(k)} \leftarrow \left[\frac{\partial_1 \mu_{q,i}^{(k)}}{(\sigma_f^*/\lambda_{x_1}^*)}, \frac{\partial_D \mu_{q,i}^{(k)}}{(\sigma_f^*/\lambda_{x_2}^*)}, \dots, \frac{\partial_D \mu_{q,i}^{(k)}}{(\sigma_f^*/\lambda_{x_D}^*)}\right]^T$, where $\partial_j \mu_{q,i} \equiv \left.\frac{\partial\mu}{\partial x_j}\right|_{\mathbf{x}_{q,i}}, j = 1, \dots, D$
6: $\epsilon \leftarrow \epsilon_{min} + (\epsilon_{max} - \epsilon_{min}) \exp\left(-\sigma_q/\left(\frac{\sigma_n+\sigma_f^*}{2}\right)\right)$
7: $\Sigma_i \leftarrow (1-\epsilon)\,\mathrm{diag}\left(\lambda_{x_1}^{*\,2}, \dots, \lambda_{x_D}^{*\,2}\right) + \epsilon\,\mathrm{diag}\left(\lambda_{x_1}^{*\,2} \exp\left(-\left|\nabla\hat{\mu}_{q,i}^{(k)}\right|\right), \dots, \lambda_{x_D}^{*\,2} \exp\left(-\left|\nabla\hat{\mu}_{q,i}^{(k)}\right|\right)\right)$
8: **end**
9: **for** $i = 1{:}Q$ **do**
10: $\left\{\mu_{q,NS,i}^{(k)}, \sigma_{q,NS,i}^{(k)}\right\} \leftarrow$ prediction at $\mathbf{x}_{q,i}$ using GP with $\left\{\mathbf{X}^{(k-1)}, \boldsymbol{y}^{(k-1)}, k_{NS}|\left(\Sigma_1, \dots, \Sigma_Q\right), \sigma_n\right\}$
11: **end**
12: $i_{MostInfo} \leftarrow \mathrm{argmax}_i\, \sigma_{q,NS,i}^{(k)}$
13: $\mathbf{x}_{next} \leftarrow \mathbf{x}_{q,i_{MostInfo}}$
14: $y_{next} \leftarrow$ **Measurement**$(\mathbf{x}_{next})$
15: $\mathbf{X}^{(k)} \leftarrow \left[\mathbf{X}^{(k-1)}, \mathbf{x}_{next}\right], \boldsymbol{y}^{(k)} \leftarrow \left[\boldsymbol{y}^{(k-1)}, y_{next}\right]$
16: **end**
17: $\mu_{\mathcal{X}_q} \leftarrow \left\{\mu_{q,NS,1}^{(N)}, \dots, \mu_{q,NS,Q}^{(N)}\right\}, \sigma_{\mathcal{X}_q} \leftarrow \left\{\sigma_{q,NS,1}^{(N)}, \dots, \sigma_{q,NS,Q}^{(N)}\right\}$

3 Results

3.1 Reference Wind Tunnel Experiment Data

In order to verify the metamodeling performance of the ADoE with the proposed nonstationary GP, wind tunnel experiment data of the variant of unmanned combat aerial vehicle (UCAV) 1303 [6] is used as a reference One-Factor-at-a-Time (OFAT) data. UCAV 1303 has a tailless delta-wing, and it is known that such a vehicle configuration usually reveals very complex aerodynamic phenomena especially when it encounters flow at a high angle-of-attack. We focus on the roll moment coefficient C_l with respect to both angle-of-attack α and angle-of-sideslip β. Figure 2 is the OFAT data points of $C_l = C_l(\alpha, \beta)$ for $\alpha \in \{0deg, 1deg, \dots, 24deg\}$ and $\beta = \{0deg, 2deg, \dots, 20deg\}$. These 275 data points are considered as a true reference model for the stationary and nonstationary GP model.

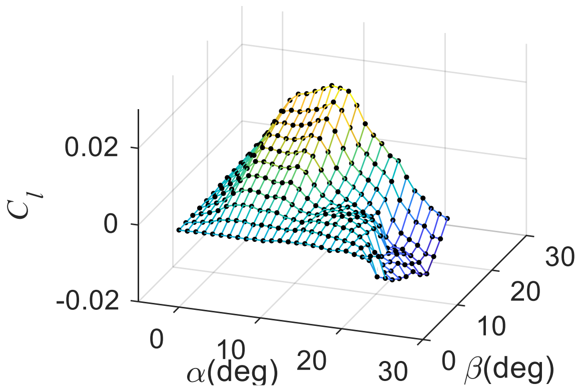

Figure 2. OFAT data (275 points) of roll moment coefficient

3.2 Comparison of ADoE Performance

To compare the stationary GP model and the nonstationary GP model in the perspective of ADoE, the ADoE-GP (Algorithm 1) and ADoE-NGP (Algorithm 2) are applied to randomly sampled $N_{IP} = 15$ initial data points (Initial Phase) to obtain $N_{AP} = 150$ adaptive data points (Adaptive Phase). For each k-th run in the Adaptive Phase, prediction on mean of $C_l = C_l(\alpha, \beta)$ obtained from the stationary GP model and the nonstationary GP model are compared with the OFAT data by calculating their root-mean-square-error (RMSE) with the 275 OFAT data points. To account for the effect of measurement noise during the wind tunnel experiment, every initial and adaptive include a uniform noise $n_{C_l} \sim \mathcal{U}(-0.001, 0.001)$. These whole procedures are repeated for 500 times to obtain meaningful statistical results considering measurement noise error and several random combinations of initial points. The simulation results are shown through Figures 3-4.

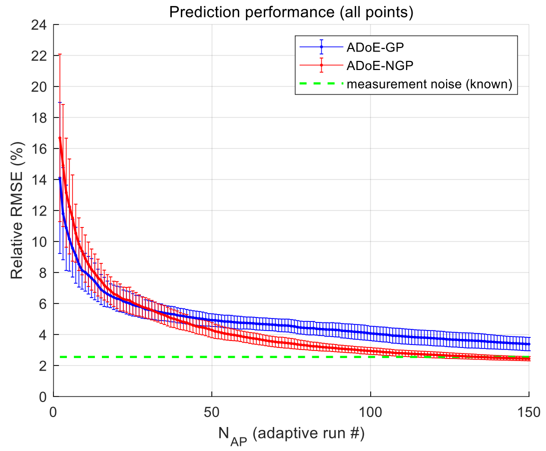

Figure 3. Comparison of ADoE metamodel prediction accuracy for all OFAT data points

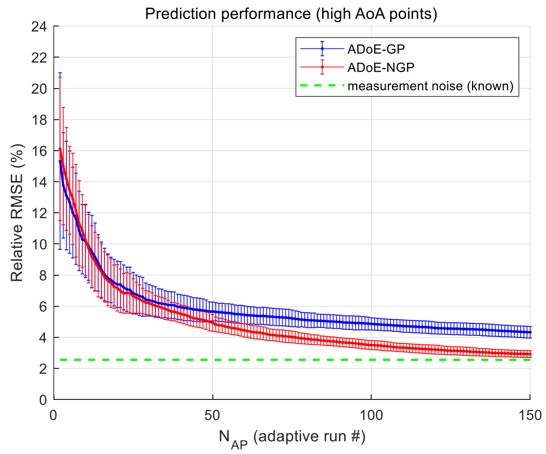

Figure 4. Comparison of ADoE metamodel prediction accuracy for high angle-of-attack data points

In Fig. 3, early runs of the Adaptive Phase show little difference between the ADoE-GP and the ADoE-NGP. However, after approximately 30 runs in the Adaptive Phase, ADoE-NGP shows better performance than ADoE-GP in terms of the RMSE with the original OFAT data. Also, when the Adaptive Phase runs are conducted more than 100 times, it turns out that the RMSE of the ADoE-NGP gradually approaches the standard deviation of the known measurement noise which can be considered as the rough lower bound of the metamodel prediction performance. This trend appears more strongly when the metamodel predictions of both ADoE-GP and ADoE-NGP are compared with only the OFAT points with $\alpha > 15\ deg$. In the high angle-of-attack region, the aerodynamic coefficients are very nonlinear and irregular. Therefore, it can be seen in Fig. 4 that the ADoE-NGP outperforms ADoE-GP even in the early runs of the Adaptive Phase.

4 Conclusions

In order to improve the stationary GP model used for the underlying metamodel of the ADoE in wind tunnel testing experiments, nonstationary GP model with nonstationary kernel is proposed. Then, overall nonlinear regression performance improvement is observed based on the Monte-Carlo simulations with realistic wind tunnel data. The proposed ADoE process shows better performance especially for high angle-of-attack region data where aerodynamic responses are expected to be nonlinear and irregular. However, the proposed ADoE method has some limitations such as instability in the early runs of the Adaptive Phase. Our future works will cover the drawbacks of the proposed ADoE with nonstationary GP model.

References

[1] DeLoach R., et al., “The modern design of experiments - A technical and marketing framework”, 21st AIAA Advanced Measurement Technology and Ground Testing Conference, Denver, CO, USA, 2000.

[2] H. Jin, D. Lee, J. Lee, J. Ahn, “Hyperparameters Optimization for Adaptive Design of Experiment Applied to Wind Tunnel Testing”, 30th Congress of the International Council of the Aeronautical Sciences, DCC, Daejeon, Korea, September 25-30, 2016

[3] Rasmussen, C. E., and Williams, C. K. I., Gaussian Processes for M achine Learning. Cambridge:The MIT Press, 2006

[4] Lang, Tobias, Christian Plagemann, and Wolfram Burgard, "Adaptive Non-Stationary Kernel Regression for Terrain Modeling." Robotics: Science and Systems. Vol. 6. 2007.

[5] Paciorek, C. J., and Schervish, M. J. (2004). “Nonstationary covariance functions for Gaussian process regression”, Advances in Neural Information Processing Systems (pp. 273-280).

[6] Byeongmin Jeong, Dongoo Lee, Hojoon Shim, Jaemyung Ahn, Han-Lim Choi, Seung-O Park, and Se-Yoon Oh. "Yaw-Control Spoiler Design Using Design of Experiments Based Wind Tunnel Testing", Journal of Aircraft, Vol. 52, No. 2 (2015), pp. 713-718.

Acknowledgement

This work was conducted at High-Speed Vehicle Research Center of KAIST with the support of the Defense Acquisition Program Administration and the Agency for Defense Development under Contract UD170018CD.

Contact Author Email Address

mailto:jaemyung.ahn@kaist.ac.kr

Copyright Statement