AIAA AVIATION Forum
17-21 June 2019, Dallas, Texas
AIAA Aviation 2019 Forum

10.2514/6.2019-2888

Multi-response Gaussian Process for Multidisciplinary Design Optimization

Jangho Park *, Anuj Arora †, and Seongim Choi ‡
Virginia Polytechnic Institute and State University, Blacksburg, VA, 24060

In the practical design of aircraft, multiple disciplines need to be considered for accurate design result; such as aerodynamics, structures, and thermodynamics. In the manner, several design objectives which are generally tightly coupled in physics should be handled in the multidisciplinary design. In this paper, for the tight coupling of different principles, a multi-response Gaussian Process Regression (GPR) method is introduced and implemented into the Efficient Global Optimization (EGO). Using the method, multiple responses, which are physically related in practical design, can be analytically correlated through the surrogate modeling and provides more accurate estimation compared to the traditional GPR methods. The validation of the method is shown and the applications of aeroelasticity are provided in the paper.

Nomenclature

Cov	=	covariance
$\boldsymbol{B}$	=	the matrix of covariance between responses
$G(a, b)$	=	Gaussian process with follows mean a and variance b
n	=	the number of observed samples
p	=	the dimension of input covariates
$\boldsymbol{Q}$	=	the matrix of covariance between samples
q	=	the dimension of output covariates
$\boldsymbol{S}$	=	the matrix of correlation between sample noise
$\mathbf{x}$	=	the vector of input covariates
$\mathbf{y}$	=	the vector of response covariates

I. Introduction

The military and commercial aircraft sectors both show a dramatic escalation in cost across a unit's life cycle. The unit cost of a U.S. combat aircraft has grown by a factor of four every decade [1]. Affordability is a very significant challenge for aircraft manufacturers.

Fig. 1 shows the life cycle cost (LCC) of an aircraft at different stages of its life [2]. It has been recognized from aircraft design experience that reducing the total LCC is the right path to affordability. At the early design stages, namely the pre-conceptual design stage, conceptual design stage, and preliminary design stage, the cumulative percentage of LCC is less than 10%. However, the impact on cost (IOC) is enormous, accounting for more than 80% of the total LCC. Fig. 1 clearly shows that the relation between cost and the impact on cost is disproportionate. Design changes at later design stages become increasingly expensive. Therefore, obtaining accurate design results at the early design stages provides huge cost savings, along with improved efficiency and effectiveness of the whole design process.

In the practical design of aircraft, at the early design stage, such as conceptual design and preliminary design, multiple disciplines, such as aerodynamics, structures, and thermodynamics, need to be considered for the accurate design results. For example, since the wing structure is not a rigid body, the wing can be deformed under the aerodynamic-loading [**?**]. Therefore, the aerodynamic performance and structural characteristics of wing are tightly coupled, and for the accurate design analysis and optimization, aeroelastic analysis should be considered.

*Ph.D, Aerospace and Ocean Engineering, astrox@vt.edu, and AIAA Member
†Master, Aerospace and Ocean Engineering Department, aarora23@vt.edu, and AIAA Member
‡Professor, Aerospace and Ocean Engineering Department, schoi1@vt.edu, and AIAA Member

In the multidisciplinary design process, several design objectives, which stand for the performance of different disciplines, can be considered. For example, if the aspect ratio of the wing, which is one of the design variable, increases, lift to drag ratio of the wing increases due to the reduction of induced drag. However, since the span length increases, high aspect ratio wing has higher bending stress for a given load which leads to larger deformation. Since both lift to drag ratio and deformation are the function of aspect ratio, it can be said that they are correlated. Therefore, for effective multidisciplinary design, the relation between objectives and design variables and the relation between different objectives should be considered.

In the gradient-based optimization, multidisciplinary design has been carried out using Adjoint method. Prasad et al. derived the coupled adjoint equation for aerodynamics and structural analysis [3]. Since two principle are tightly coupled by adjoint variables, the multidisciplinary design analysis and optimization can be done accurately and cost-efficiently. However, if the design problem is changed, such as the coupling of aerodynamics and thermodynamics, the adjoint equations should be re-derived for the corresponding problem.

In the gradient-free optimization, Simpson et al. used Kriging models for multidisciplinary design optimization of aerospike nozzle [4]. The thrust, weight of engine, and gross lift-off weight off (GLOW) of aircraft were set as objective functions, and Kriging models were constructed for each objective separately. Therefore, the estimated value of objectives were calculated in every design iterations. Park et al. performed multidisciplinary design optimization for the control surface of tailless aircraft [5]. In this research, the control surface of tailless aircraft was optimized considering control-ability of aircraft as a constraint. The response surface of time-to-bank-30-degrees was constructed, and the objective function was set to minimize the area of control surface to minimized the drag and hinge moment.

However, in these studies, the objective functions were not coupled in surrogate models because they were built separately without any mathematical relation between the models even they are coupled in the physical phenomena. To couple the objectives from different principles, Bowcutt used a new objective, which included the effect of different disciplines. For the multidisciplinary optimization of air breathing hypersonic vehicle, which considers aerodynamics, propulsion, stability, control-ability, and mass property of the aircraft, he used the objective function as maximizing flight range [6]. Constructing the surrogate model of flight range, he could mildly couple the effect of multiple disciplines in one surrogate model.

For the tight coupling of different principles, Wang et al. and Kleijnen et al. suggested Gaussian Process Regression (GPR) of multiple response variables [7, 8]. Traditional GPR models only a single response variable which only considers the correlation between the location of the given data points and the trial data points. However, they added correlation terms which considers the coupling between different responses. As the correlation of objectives are considered in the mathematical model, the physical interaction between objectives, such as the relation between lift to drag ratio and maximum deformation of the wing, can be modeled as well.

In this paper, multi-response GRP models provided by Wang et al. and Kleijnen et al. are introduced and advanced using iterative Maximum Likelihood Estimation (MLE) process. (Sec. II) The advanced multi-response GPR models are validated with test functions and implemented into Efficient Global Optimization (EGO) process. (Sec. III) Improved EGO process will be compared with EGO process with single response GRP model as pros and cons. In Sec. IV, the

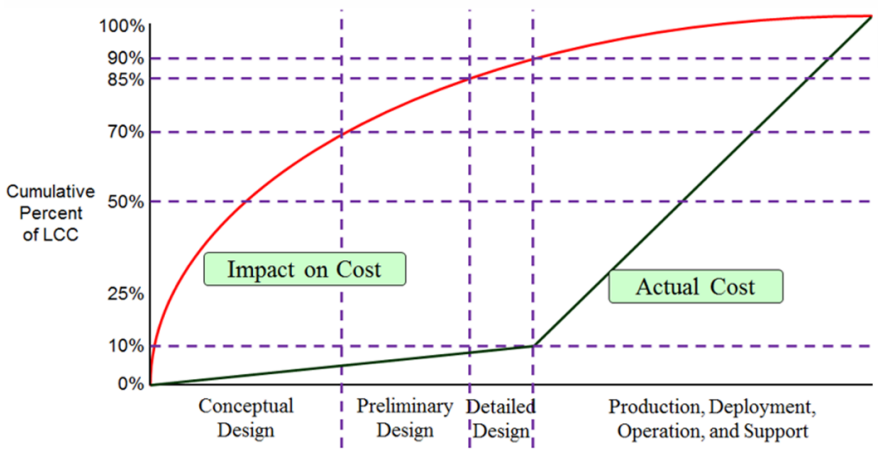


Fig. 1 Life cycle cost and impact on cost for each design stage [2]

application of the optimization method will be suggested.

II. Literature Survey

To provide research background of the study, literature on the traditional Kriging model and Co-Kriging model is shown. Previous studies about Gaussian Process Regression models with multiple output covariates are also introduced.

In the practical multidisciplinary design cases, we have several types of outputs and each type is a specific transformation of the same input combination. For example, in aeroelasticity analysis, for an ordinary design sample, we can obtain both aerodynamic analysis result and elasticity analysis result. To make more than two surrogate models of correlated outputs, CoKriging gives good prediction of both outputs. Using CoKriging, the multi-variate outputs can be predicted by a linear combination of all the available data values [9]. In CoKriging, the covariance between responses can be expressed into closed-form equation, which is a function of observation of outputs [10].

If the gradient value at the sample location can be obtained, it will be superb information to improve the accuracy of surrogate model. Since the gradient is tightly related with the function value, CoKriging model is ideal to include gradient data into the model. Chung et al. developed the CoKriging model which can incorporate gradient information in addition to the function values of the sample points for constructing approximation models, and approached improves on the accuracy and efficiency of using the Kriging method for high dimensional design problems [11]. Forrester et al. and Yamazaki et al suggested Gradient-enhanced model and Hessian-enhanced model, which incorporate gradient or Hessian information as a second set of basis functions in Kriging model [12, 13]. Laurenceau used Gradient-enhanced Kriging model to build the response surface of aerodynamic functions, and the improvement in global accuracy of the model can be observed compared to ordinary Kriging model [14].

In variable-fidelity analysis, different simulation models with different degrees of realism are correlated to each other models since they are explaining same phenomena. Therefore, CoKriging has been extended to variable-fidelity CoKriging. Han et al develped variable-fidelity model combined direct GEK and bridge function. Using the model, a global surrogate model for the aerodynamic coefficient and drag polar of RAE2822 airfoil was constructed [15–17]. Jo et al developed variable-fidelity Kriging model where gradient data from Adjoint solver are directly utilized to improve the accuracy of the Kriging. The model is implemented to Efficient Global Optimization (EGO), and UCAV, unmanned combat air vehicle which is tailless aircraft, has been optimized to maximize lift coefficient and minimize drag coefficient [18].

While CoKriging calculates the covariance between responses as a function of observation of outputs, there are another attempts to set the covariance between responses as unknown hyper-parameters in Maximum Likelihood Estimation (MLE). Neal suggested use a single hyper-parameter, which is the noise variance, for all covariance between outputs, however this method may not be enough to differentiate the effect of multiple outputs [19]. Wang et al [20] developed a multiple response model to jointly model the quantile with a correlated and less-noisy expectation to improve the fit and predictions from the quantile metamodel. Conti et al suggested separable multiple response Gaussian Process(GP) model which is extended model of single-output Bayesian model [21, 22]. Ver Hoef et al developed multi response GP model with the convolution method model [23], and Goulard et at proposed the model with the linear model of coregionalization(LMC) [24]. Alvarez et at suggested multiple output GP model with dependent output Gaussian processes constructed through the convolution formalism [25].

Wang et al developed Gaussian process regression with multiple response variables by assuming that the covariance function for multi-response is the nominal uni-output covariance multiplied by the covariances between different outputs [7]. This approach can be extended to the case of different input covariates for different outputs covariates. Hong et al extended the method to have heteroscedastic and correlated noises terms [26, 27].

Kleijnen et al proposed multivariate Kriging metamodels for multi-response model with nonseparable dependence covariance structure [8]. This approach is not able to extended to the case of different design location for different responses, but the covariance between responses can be set in form of correlation function. In this research, we focused on Wang and Kleijnen's multi-response Kriging model. With Kleijnen's multivariate Kriging metamodels, Sadoughi et al proposed multivariate system reliability analysis (MSRA), for complex engineered systems with highly nonlinear and dependent components that are connected in series, parallel, and mixed configurations [28].

In this paper, Wang and Kleijnen's methods are advanced with iterative Maximum Likelihood Estimation (MLE) process and applied to Efficient Global Optimization (EGO) process.

III. Mathematical formulation of Gaussian Process Regression Models

In this section, the mathematical formulations of different Gaussian Process Regression(GPR), Kriging, models are introduced. First, the mathematics of the ordinary GPR model which solves single response problems are discussed. Second, the formulation of the ordinary GPR model is extended to develop a multiple response Kriging (MRK) model of a reduced-covariance MRK and a fully expended covariance MRK, respectively. The discrepancy of the two MRK models lies in the structure of the covariance matrix. The former is similar to that by Wang et. al. [], and the latter by Kleiginen et. al. []. The procedure of iterative optimization for the Maximum Likelihood Estimation (MLE) is also developed to improve the accuracy of the fully expanded covariance MRK.

A. A single response GPR model

A Gaussian process (GP) is a realization process of random variables, a finite number of which have a joint Gaussian distribution [29]. This process can be specified by the mean function $m(\boldsymbol{x})$ and covariance function $c(\boldsymbol{x}, \boldsymbol{x'})$ of a real process $f(\boldsymbol{x})$ as below,

$$\begin{aligned} m(\boldsymbol{x}) &= E[f(\boldsymbol{x})] \\ c(\boldsymbol{x}, \boldsymbol{x'}) &= E[(f(\boldsymbol{x}) - m(\boldsymbol{x}))(f(\boldsymbol{x'}) - m(\boldsymbol{x'}))] \end{aligned} \tag{1}$$

where $\boldsymbol{x}$ is made up of p-dimensional input covariates, $\boldsymbol{x} = [x_1, x_2, \cdots, x_p]^T$. The Gaussian process can be written as

$$f(\boldsymbol{x}) \sim G(m(\boldsymbol{x}), c(\boldsymbol{x}, \boldsymbol{x'})) \tag{2}$$

The GPR or Kriging model is a meta model for which a function values are estimated by a Gaussian process [30]. Using linear regression, the response variable y can be expressed as a function of covariates and a regression parameter $\boldsymbol{w} = [w_1, w_2, \cdots, w_p]^T$.

$$y_i = \boldsymbol{x}_i^T \boldsymbol{w} + \epsilon_i, \; i = 1, 2, \cdots, n, \tag{3}$$

where n is the number of observed data points, and ϵ is additive Gaussian noise with zero mean and unknown variance σ_ϵ^2. The ordinary GPR model sets the mean of response variable y as zero. Under the Bayesian update approach, a prior distribution of the model needs to be assumed. In the regression model, the regression parameter $\boldsymbol{w}$ will be assumed to follow a Gaussian distribution with zero mean and covariance σ_w^2 which is independent of ϵ_i as below.

$$\boldsymbol{w} \sim G(\boldsymbol{0}, \sigma_w^2 \boldsymbol{I}_p) \tag{4}$$

where $\boldsymbol{I}_p$ is $p \times p$ identity matrix.

Now, for a trial location $\hat{\boldsymbol{x}} \in \mathbb{R}^p$ which has the estimated function value $\hat{y}$, the joint probability distribution of observed data augmented by $\hat{\boldsymbol{x}}$ and $\hat{y}$ can be used to evaluate the Gaussian process posterior distribution at $\hat{\boldsymbol{x}}$ [31].

$$\begin{bmatrix} \boldsymbol{y} \\ \hat{y} \end{bmatrix} \sim G\left(\boldsymbol{0}, \begin{bmatrix} \boldsymbol{C}(\boldsymbol{x}, \boldsymbol{x}) & \boldsymbol{C}(\boldsymbol{x}, \hat{\boldsymbol{x}}) \\ \boldsymbol{C}(\hat{\boldsymbol{x}}, \boldsymbol{x}) & \boldsymbol{C}(\hat{\boldsymbol{x}}, \hat{\boldsymbol{x}}) \end{bmatrix} \right) \tag{5}$$

where $\boldsymbol{y} = [y_1, y_2, \cdots, y_n]^T$ is a vector of the observed response variable, and C is a covariance matrix. Since there are n number of $\boldsymbol{x}$ vectors corresponding to n number of observed data points, the size of $C(\boldsymbol{x}, \boldsymbol{x})$ matrix is $(n \times n)$. The size of $\boldsymbol{C}(\boldsymbol{x}, \hat{\boldsymbol{x}})$ is $(n \times 1)$, and $\boldsymbol{C}(\hat{\boldsymbol{x}}, \boldsymbol{x})$ is transpose matrix of $\boldsymbol{C}(\boldsymbol{x}, \hat{\boldsymbol{x}})$. The size of $\boldsymbol{C}(\hat{\boldsymbol{x}}, \hat{\boldsymbol{x}})$ matrix is (1×1). The elements of the $\boldsymbol{C}$ matrix can be obtained by calculating the expected covariance. For two observed data points, $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ $(i, j = 1, 2, \cdots, n)$, the covariance between two data can be calculated as below,

$$\begin{aligned} C_{ij} = \boldsymbol{C}(\boldsymbol{x}_i, \boldsymbol{x}_j) &= cov[y_i, y_j] = E\left[(y_i - E[y_i]) \left(y_j - E[y_j] \right) \right] \\ &= \sigma_w^2 \boldsymbol{x}_i^T \boldsymbol{x}_j + \delta_{ij} \sigma_\epsilon^2 \end{aligned} \tag{6}$$

where δ_{ij} is the delta function; if $i = j$, $\delta_{ij} = 0$, and if $i \neq j$, $delta_{ij} = 1$. In Eq. 6, the covariance matrix is a function of σ_w^2, the variance of prior probability. Since the observation of $\boldsymbol{y}$ and the estimated function value of y^* are connected by the joint probability in Eq. 5, using Gaussian identities, the posterior probability, which is the conditional probability of y^* for given data $\boldsymbol{y}$ can be calculated as below,

$$
\begin{aligned}
p\left(\hat{y}|\boldsymbol{y}\right) &= G\left(\mu(\hat{y}), \sigma^2(\hat{y})\right) \\
\mu(\hat{y}) &= \boldsymbol{C}(\hat{\boldsymbol{x}}, \boldsymbol{x})\boldsymbol{C}^{-1}(\hat{\boldsymbol{x}}, \boldsymbol{x})\boldsymbol{y} \\
\sigma^2(\hat{y}) &= \boldsymbol{C}(\hat{\boldsymbol{x}}, \hat{\boldsymbol{x}}) - \boldsymbol{C}(\hat{\boldsymbol{x}}, \boldsymbol{x})\boldsymbol{C}^{-1}(\hat{\boldsymbol{x}}, \boldsymbol{x})\boldsymbol{C}(\boldsymbol{x}, \hat{\boldsymbol{x}})
\end{aligned} \tag{7}
$$

where $\mu(\hat{y})$ is the mean of $\hat{y}$ which is the estimated function value of a trial location $\hat{\boldsymbol{x}}$, and $\sigma^2(\hat{y})$ is estimation uncertainty.

Instead of inferring parameter $\boldsymbol{w}$, the regression model can be summarized by several forms of the covariance function, which are related to $\boldsymbol{C}$ matrix. The covariance function below is one of an example form which is gamma exponential function [32].

$$
\begin{aligned}
C_{ij} &= \prod_{k=1}^{p} C_k\left(\theta, \gamma, x_{ik} - x_{jk}\right) \\
C_k &= \exp\left(-\theta_k \left|d_k\right|^{\gamma}\right), 0 < \gamma \leq 2
\end{aligned} \tag{8}
$$

Since k varies from 1 to p, there are $(p + 1)$ unknown variables, which are θ_k and $gamma$. These parameters are called hyperparameters. Assuming a Gaussian process, the optimal hyperparameters of the covariance function can be determined using optimization process, which called Maximum Likelihood Estimation (MLE) [33]. The MLE problem is defined below.

$$
\min_{\theta,\gamma}\left\{\psi(\theta, \gamma) \equiv \left|\boldsymbol{C}(\boldsymbol{x}, \boldsymbol{x})\right|^{1/n} \sigma_w^2\right\} \tag{9}
$$

where $|\boldsymbol{C}(\boldsymbol{x}, \boldsymbol{x})|$ is the determinant of $\boldsymbol{C}(\boldsymbol{x}, \boldsymbol{x})$.

B. Reduced-Covariance Multi-Response Kriging (RC-MRK)

In next two sections, the single response GPR will extended to encorporate multiple correlated response. The first method is reduced covariance multi-response Kriging (RC-MRK) model. The model is developed by Wang et al [7].

For the q-dimensional responses, $\boldsymbol{y} = \left[y_1, y_2, \cdots, y_q\right]^T$, the multi-response linear regression model can be expressed as below,

$$
\boldsymbol{y}_i = \boldsymbol{K}_i\boldsymbol{W} + \boldsymbol{\epsilon}_i, i = 1, 2, \cdots, n \tag{10}
$$

where n is the number of observed data points and $\boldsymbol{K}_i = \boldsymbol{I}_q \otimes \boldsymbol{x}_i^T$ is the Kronecker product of the $q \times q$ identity matrix and p-dimensional input covariates vector. Therefore, the size of $\boldsymbol{K}_i$ is $q \times pq$. $\boldsymbol{W}$ is the concatenation of the regression vectors for the q response variables, which can be expressed as $\boldsymbol{W} = \left[\boldsymbol{w_1^T}, \boldsymbol{w_2^T}, \cdots, \boldsymbol{w_q^T}\right]^T$. The size of $\boldsymbol{W}$ is $pq \times 1$. $\boldsymbol{\epsilon}_i$ is a q-dimensional vector of noise given by a zero mean Gaussian distribution as below,

$$
\boldsymbol{\epsilon}_i \sim G(\boldsymbol{0}, \boldsymbol{S}) \tag{11}
$$

where $\boldsymbol{S}$ is $q \times q$ correlation matrix between the elements of $\boldsymbol{\epsilon}_i$ vector. Since the noise is generally assumed to be independently and identically distributed noise, $\boldsymbol{S}$ becomes a diagonal matrix.

If the prior probability for the vectors of regression parameter $\boldsymbol{w}_g$ and $\boldsymbol{w}_h$, where g and h varies from 1 to q, is assumed to follow a Gaussian process, the covariance matrix between regression vectors can be defined as below,

$$
\operatorname{cov}\left(\boldsymbol{w}_g, \boldsymbol{w}_h\right) = \boldsymbol{\Sigma}^{gh} = b_{gh}\sigma^2\boldsymbol{I}_p \tag{12}
$$

where b_{gh} forms a $q \times q$ symmetric matrix B, which corresponds the covariance between outputs.

Now, for trial locations $\hat{\boldsymbol{x}} \in \mathbb{R}^p$ which has vectors of estimated function value $\hat{\boldsymbol{y}}$, the joint probability distribution of observed data augmented by $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$ can be considered to evaluate the Gaussian process posterior distribution at $\hat{\boldsymbol{x}}$.

$$
\begin{bmatrix} \boldsymbol{y}_{aug,RC-MRK} \\ \hat{\boldsymbol{y}} \end{bmatrix} \sim G\left(\boldsymbol{0}, \begin{bmatrix} \boldsymbol{C}(\boldsymbol{x}, \boldsymbol{x}) & \boldsymbol{C}(\boldsymbol{x}, \hat{\boldsymbol{x}}) \\ \boldsymbol{C}(\hat{\boldsymbol{x}}, \boldsymbol{x}) & \boldsymbol{C}(\hat{\boldsymbol{x}}, \hat{\boldsymbol{x}}) \end{bmatrix}\right) \tag{13}
$$

where $\boldsymbol{y}_{RC-MRK} = [\boldsymbol{y}_1^T, \boldsymbol{y}_2^T, \cdots, \boldsymbol{y}_q^T]^T = [y_{11}, \cdots, y_{1n}, y_{21}, \cdots, y_{2n}, \cdots, y_{q1}, \cdots, y_{qn}]^T$ is a augmented vector of observed response variables, and C is a covariance matrix between sample points. In Ref [7], the covariance function is suggested as below,

$$\begin{aligned} C_{ij}^{gh} &= Q_{ij} b_{gh} + \delta_{ij}\delta_{gh} S_{gh} \\ Q_{ij} &= a_0 + a_1 \sum_{d=1}^{p} x_{id} x_{jd} + \nu_0 \exp\left(- \sum_{d=1}^{p} \eta_d \left(x_{id} - x_{jd} \right)^2 \right) \end{aligned} \tag{14}$$

where a_0, a_1, ν_0, and η_d are hyper-parameters. Each hyper-parameter correspond bias term, the weight of linear correlation function, the weight of Gaussian correlation function, and the weight on distance in Gaussian correlation function, respectively. From Eq. 14, the covariance matrix of $\boldsymbol{C}$ can be derived as below.

$$\begin{aligned} \boldsymbol{C} &= \begin{bmatrix} b_{11}\boldsymbol{Q} + S_{11}\boldsymbol{I}_n & b_{12}\boldsymbol{Q} & \cdots & b_{1q}\boldsymbol{Q} \\ b_{21}\boldsymbol{Q} & b_{22}\boldsymbol{Q} + S_{22}\boldsymbol{I}_n & \cdots & b_{2q}\boldsymbol{Q} \\ \vdots & \vdots & \ddots & \vdots \\ b_{q1}\boldsymbol{Q} & b_{q2}\boldsymbol{Q} & \cdots & b_{qq}\boldsymbol{Q} + S_{qq}\boldsymbol{I}_n \end{bmatrix} \\ &= \boldsymbol{B} \otimes \boldsymbol{Q} + \boldsymbol{S} \otimes \boldsymbol{I}_n \end{aligned} \tag{15}$$

where $\otimes$ denotes the Kronecker product. Since $\boldsymbol{C}$ is a function of samples with multiple response covariates, the size is $qn \times qn$. $\boldsymbol{S}$ matrix is q-dimensional diagonal matrix of noise which is related to ϵ_i, above. Since these terms are unknown as well, we have $S_{11}, S_{22}, \cdots, S_{qq}$ unknown variables, which are the elements of $\boldsymbol{S}$ matrix and another hyper-parameters. In the covariance matrix, $\boldsymbol{B}$ is the matrix of covariance between responses, which is a symmetric matrix. Therefore, the matrix can be re-organized as shown below.

$$\boldsymbol{B} = \boldsymbol{\phi}\boldsymbol{\phi}^T, \boldsymbol{\phi} = \begin{bmatrix} \phi_{11} & 0 & \cdots & 0 \\ \phi_{21} & \phi_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{q1} & \phi_{q1} & \cdots & \phi_{q1} \end{bmatrix} \tag{16}$$

where ϕ_{11},ϕ_{21},ϕ_{22}, $\cdots$ ϕ_{qq} are the hyper-parameters for the covariance between responses.

Therefore, in multiple-response GPR model, we have $\left(p + \frac{q(q+3)}{2} + 3 \right)$ number of unknown hyper-parameters, which should be determined. These hyper-parameters can be determined by Maximum Likelihood Estimation (MLE) process. The equation below shows the equation of likelihood.

$$L = -\frac{1}{2}\log[\det \boldsymbol{C}] - \frac{1}{2}\boldsymbol{y}_{aug,RC-MRK}^T \boldsymbol{C}^{-1} \boldsymbol{y}_{aug,RC-MRK} - \frac{nq}{2}\log[2\pi] \tag{17}$$

For this optimization process, non-linear Gaussian-Seidel method can be used. While single-response Gaussian process regression has only $(p + 4)$ number of hyper-parameters, multi-response Gaussian process regression has much more number of variables since we need to consider the covariance between responses. For example, if we have 4 variables and 3 responses, Gaussian process regression will have 16 variables, and performing the optimization of 16 variables will be difficult. Therefore, the optimization process is divided into two sub-optimization process. First of all, from the initial guess of hyper-parameters, the variables related to the covariance between responses, which are in $\boldsymbol{B}$ matrix, are fixed as constants, and rest of all parameters are optimized. Then, the variables related to the covariance between responses, which are the elements of $\boldsymbol{S}$ matrix, a_0, a_1, ν_0, and η_d, are optimized with fixed parameters from the first step.

From the equation, using Gaussian identities, the posterior probability, which is the conditional probability of $\boldsymbol{y}^*$ for given data $\boldsymbol{y}$ can be calculated as below,

$$
\begin{aligned}
p\left(\hat{\boldsymbol{y}}|\boldsymbol{y}_{aug,RC-MRK}\right) &= G\left(\boldsymbol{\mu}(\hat{\boldsymbol{y}}), \boldsymbol{\sigma}^2(\hat{\boldsymbol{y}})\right) \\
\boldsymbol{\mu}(\hat{\boldsymbol{y}}) &= \left[\boldsymbol{B} \otimes \boldsymbol{Q}^*\right]^T \boldsymbol{C}^{-1} \boldsymbol{y}_{aug,RC-MRK} \\
\boldsymbol{\sigma}^2(\hat{\boldsymbol{y}}) &= \boldsymbol{B} \otimes Q\left(\hat{\boldsymbol{x}}, \hat{\boldsymbol{x}}\right) + \boldsymbol{S} - \left[\boldsymbol{B} \otimes \boldsymbol{Q}^*\right]^T \boldsymbol{C}^{-1} \left[\boldsymbol{B} \otimes \boldsymbol{Q}^*\right]
\end{aligned}
\tag{18}
$$

where $\mu(\hat{\boldsymbol{y}})$ is the mean of $\hat{\boldsymbol{y}}$, which is a vector of estimated multi-reponse function values at a trial location $\hat{\boldsymbol{x}}$, and $\sigma^2(\hat{\boldsymbol{y}})$ is a matrix of estimation uncertainty for each response. The estimation uncertainty matrix includes the variance of the prediction at the trial location and the covariance of the prediction between different responses as shown in Eq.19. $\boldsymbol{Q}^* = [Q(\boldsymbol{x}_1, \hat{\boldsymbol{x}}), Q(\boldsymbol{x}_2, \hat{\boldsymbol{x}}), \cdots, Q(\boldsymbol{x}_n, \hat{\boldsymbol{x}})]$ is the covariances between the test data point $\hat{\boldsymbol{x}}$ and the training sample sets. The detailed derivation is introduced in Ref [7].

$$
\boldsymbol{\sigma}^2(\hat{\boldsymbol{y}}) = \begin{bmatrix}
Var(\hat{y}_1) & Cov(\hat{y}_1, \hat{y}_2) & \cdots & Cov(\hat{y}_1, \hat{y}_q) \\
Cov(\hat{y}_2, \hat{y}_1) & Var(\hat{y}_2) & \cdots & Cov(\hat{y}_2, \hat{y}_q) \\
\vdots & \vdots & \ddots & \vdots \\
Cov(\hat{y}_q, \hat{y}_1) & Cov(\hat{y}_q, \hat{y}_2) & \cdots & Var(\hat{y}_q)
\end{bmatrix}
\tag{19}
$$

C. Fully-Expanded-Covariance Multi-Response Kriging (FEC-MRK)

The second method is fully-expanded-covariance multi-response Kriging(FEC-MRK),proposed by Kleijnen et al [8]. For n number of observed training samples, the covariance matrix C is defined as shown in Eq. 20. While Wang et al.

While Wang et al. used Kronecker product for covariance across reponses and the spatial covariance, Kleijnen et al. included the effect of both covariance into one matrix, $\boldsymbol{Cov}\left(\boldsymbol{y}(\boldsymbol{x}_i), \boldsymbol{y}(\boldsymbol{x}_j)\right)(i, j = 1, 2, \cdots, n)$.

While Wang et al. used Kronecker product to separate the covariance between responses and the covariance between samples, Kleijnen et al. included the effect of both covariance into one matrix, $\boldsymbol{Cov}\left(\boldsymbol{y}(\boldsymbol{x}_i), \boldsymbol{y}(\boldsymbol{x}_j)\right)(i, j = 1, 2, \cdots, n)$.

For $i = j$, which is diagonal elements of the matrix, the covariance between samples are equal to 1, and only the covariance between responses are considered. Therefore, $\boldsymbol{B}$ indicates the covariance between responses, and it does not vary with the input combination $\boldsymbol{x}$. For q-dimensional response covariates, the size of $\boldsymbol{B}$ is $q \times q$, and it is symmetric matrix.

$$
\begin{aligned}
\boldsymbol{C} &= \begin{bmatrix}
\boldsymbol{B} & \boldsymbol{Cov}\left(\boldsymbol{y}(\boldsymbol{x}_1), \boldsymbol{y}(\boldsymbol{x}_2)\right) & \cdots & \boldsymbol{Cov}\left(\boldsymbol{y}(\boldsymbol{x}_1), \boldsymbol{y}(\boldsymbol{x}_n)\right) \\
\boldsymbol{Cov}\left(\boldsymbol{y}(\boldsymbol{x}_2), \boldsymbol{y}(\boldsymbol{x}_1)\right) & \boldsymbol{B} & \cdots & \boldsymbol{Cov}\left(\boldsymbol{y}(\boldsymbol{x}_2), \boldsymbol{y}(\boldsymbol{x}_n)\right) \\
\vdots & \vdots & \ddots & \vdots \\
\boldsymbol{Cov}\left(\boldsymbol{y}(\boldsymbol{x}_n), \boldsymbol{y}(\boldsymbol{x}_1)\right) & \boldsymbol{Cov}\left(\boldsymbol{y}(\boldsymbol{x}_n), \boldsymbol{y}(\boldsymbol{x}_2)\right) & \cdots & \boldsymbol{B}
\end{bmatrix} \\
\boldsymbol{B} &= \begin{bmatrix}
\sigma_1^2 & \sigma_{1,2} & \cdots & \sigma_{1,q}^2 \\
\sigma_{2,1}^2 & \sigma_2^2 & \cdots & \sigma_{2,q}^2 \\
\vdots & \vdots & \ddots & \vdots \\
\sigma_{q,1}^2 & \sigma_{q,2}^2 & \cdots & \sigma_{q,q}^2
\end{bmatrix} = \begin{bmatrix}
b_{11} & b_{12} & \cdots & b_{1q} \\
b_{21} & b_{22} & \cdots & b_{2q} \\
\vdots & \vdots & \ddots & \vdots \\
b_{q1} & b_{q2} & \cdots & b_{qq}
\end{bmatrix}
\end{aligned}
\tag{20}
$$

where $\sigma_{g,h}$ is the covariance between the response, $cov(y_g, y_h)$. $\boldsymbol{B}$ can be divided to $\boldsymbol{B} = \boldsymbol{A}\boldsymbol{A}^T$, where $\boldsymbol{A}$ is the eigen-decomposition of $\boldsymbol{B}$ while guaranteeing that $\boldsymbol{A}$ is positive-definite matrix [34]. Using $\boldsymbol{A}$ matrix, $\boldsymbol{Cov}\left(\boldsymbol{y}(\boldsymbol{x}_i), \boldsymbol{y}(\boldsymbol{x}_j)\right)$ can be expressed as below,

$$
\begin{aligned}
\boldsymbol{Cov}\left(\boldsymbol{y}(\boldsymbol{x}_i), \boldsymbol{y}(\boldsymbol{x}_j)\right) &= \boldsymbol{A}\boldsymbol{R}\boldsymbol{A}^T \\
\boldsymbol{R} &= \begin{bmatrix}
R(d_{i,j}; \theta^{(1)}) & 0 & \cdots & 0 \\
0 & R(d_{i,j}; \theta^{(2)}) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & R(d_{i,j}; \theta^{(q)})
\end{bmatrix}
\end{aligned}
\tag{21}
$$

where $R(d_{i,j};\theta^{(g)})$ is Gaussian correlation function $\exp\left[-\theta^{(g)}\left(\boldsymbol{x}_i-\boldsymbol{x}_j\right)^2\right]$. In the paper, for the comparison with RC-MRK method, the correlation function in Eq. 14 is used for REC-MRK. Using the correlation function in Eq. 14, $\boldsymbol{R}$ can be written as below.

$$\boldsymbol{R}=\begin{bmatrix} Q_{ij}^{(1)} & 0 & \cdots & 0 \\ 0 & Q_{ij}^{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & Q_{ij}^{(q)} \end{bmatrix}$$
$$Q_{ij}^{(g)}=a_0^{(g)}+\nu_0^{(g)}\exp\left(-\sum_{d=1}^{p}\eta_d\left(x_{id}-x_{jd}\right)^2\right) \tag{22}$$

where $a_0^{(g)}$, $\nu_0^{(g)}$, and η_d are hyper-parameters. Each hyper-parameter correspond bias term, the weight of Gaussian correlation function, and the weight on distance in Gaussian correlation function, respectively. Therefore, in $\boldsymbol{R}$ matrix, we have $p+2q$ number of hyper-parameters.

Since $\boldsymbol{A}$ should be positive definite matrix, using the Cholesky transformation, $\boldsymbol{A}$ can be divided into $\boldsymbol{L}\boldsymbol{L}^T$, where $\boldsymbol{L}$ is lower triangular matrix and the diagonal elements are non-negative. Now, unknown $\boldsymbol{B}$ matrix is a function of $\boldsymbol{L}$ matrix. The elements of $\boldsymbol{L}$ should be set as hyper-parameters, and we have $\frac{q(q+1)}{2}$ number of hyper-parameters related to the covariance between multiple responses.

Now, for an arbitrary hyper-parameters, the covariance matrix $\boldsymbol{C}$ can be calculated. The size of $\boldsymbol{C}$ is $qn\times qn$. For $\frac{q(q+5)}{2}+p$ number of unknown hyper-parameters, Maximum Likelihood Estimation (MLE) is performed. The equation of likelihood L is shown in Eq. 23 [33].

$$L=-\log\left[\det\boldsymbol{C}\right]-\log\left[\det\left[\boldsymbol{F}^T\boldsymbol{C}^{-1}\boldsymbol{F}\right]\right]-\left(\boldsymbol{y}_{aug,FEC-MRK}-\boldsymbol{F}\hat{\boldsymbol{\mu}}\right)^T\boldsymbol{C}^{-1}\left(\boldsymbol{y}_{aug,FEC-MRK}-\boldsymbol{F}\hat{\boldsymbol{\mu}}\right)$$
$$\hat{\boldsymbol{\mu}}=\left(\boldsymbol{F}^T\boldsymbol{C}^{-1}\boldsymbol{F}\right)^{-1}\boldsymbol{F}^T\boldsymbol{C}^{-1}\boldsymbol{y}_{aug,FEC-MRK} \tag{23}$$

with $\boldsymbol{F}=\boldsymbol{1}_n\otimes\boldsymbol{I}_q$ where $\boldsymbol{1}_n$ denotes an n-dimensional vector with ones. $\hat{\boldsymbol{\mu}}$ denotes the GLS estimator and $\boldsymbol{y}_{aug,FEC-MRK}=[\boldsymbol{y}_1^T,\boldsymbol{y}_2^T,\cdots,\boldsymbol{y}_n^T]^T=[y_{11},\cdots,y_{1q},y_{21},\cdots,y_{2q},\cdots,y_{n1},\cdots,y_{nq}]^T$ is an augmented vector of observed response variables [35].

Using the optimized hyper-parameters, the posterior probability can be calculated as below,

$$p\left(\hat{\boldsymbol{y}}|\boldsymbol{y}_{aug,FEC-MRK}\right)=G\left(\boldsymbol{\mu}(\hat{\boldsymbol{y}}),\boldsymbol{\sigma}^2(\hat{\boldsymbol{y}})\right)$$
$$\boldsymbol{\mu}(\hat{\boldsymbol{y}})=\hat{\boldsymbol{\mu}}+\hat{\boldsymbol{C}}\boldsymbol{C}^{-1}\left(\boldsymbol{y}_{aug,FEC-MRK}-\boldsymbol{F}\hat{\boldsymbol{\mu}}\right) \tag{24}$$
$$\boldsymbol{\sigma}^2(\hat{\boldsymbol{y}})=\boldsymbol{B}-\hat{\boldsymbol{C}}\boldsymbol{C}^{-1}\hat{\boldsymbol{C}}^T+\boldsymbol{U}\left(\boldsymbol{F}^T\boldsymbol{C}^{-1}\boldsymbol{F}\right)^{-1}\boldsymbol{U}^T$$

where $\boldsymbol{U}=\boldsymbol{I}_q-\hat{\boldsymbol{C}}\boldsymbol{C}^{-1}\boldsymbol{F}$ and $\hat{\boldsymbol{C}}=(Cov(\boldsymbol{y}(\hat{\boldsymbol{x}}),\boldsymbol{y}(\boldsymbol{x}_1)),Cov(\boldsymbol{y}(\hat{\boldsymbol{x}}),\boldsymbol{y}(\boldsymbol{x}_2)),\cdots,Cov(\boldsymbol{y}(\hat{\boldsymbol{x}}),\boldsymbol{y}(\boldsymbol{x}_n)))$. The size of $\hat{\boldsymbol{C}}$ matrix is $q\times qn$.

D. Comparison between RC-MRK and FEC-MRK

In the previous sub-sections, RC-MRK and FEC-MRK methods are introduced. The key difference between two methods are how to handle the covariance matrix between responses and the number of hyper-parameters.

In RC-MRK, the covariance matrix between multivariate responses, B, can be separated to the spatial covariance matrix Q by Kronecker product as shown in Eq.15. To put it another way, each covariance between responses $(b_{11},b_{12},b22,\cdots,b_{qq})$ is expressed as a single unknown constant, and these values are multiplied to the spatial covariance. Since B and Q matrices are separated, the method can be extended to the case of different sample number and locations for different responses.

Table 1 Comparison of the number of hyper-parameters

	Single response GRP	RC-MRK	FEC-MRK
Hyper-parameters for the covariance between samples	$q(p+3)$	$p+3$	$p+3q$
Hyper-parameters for the covariance between responses	0	$\frac{q(q+3)}{2}$	$\frac{q(q+1)}{2}$
Total number of Hyper-parameters	$q(p+3)$	$p+\frac{q(q+1)}{2}+3$	$p+\frac{q(q+5)}{2}$

In FEC-MRK, the covariance matrix between different responses is tightly combined with the spatial covariance matrix R, which is nonseparable dependence covariance structure, as shown in Eq.21. Since the the covariance matrix of response and spatial covariance matrix are coupled to each other, the meta-model is not able to be used for different sample number and locations in different responses. However, the method has much more flexibility in the form of covariance matrix of responses. While the correlation function between responses in RC-MRK is set as a single constant, which is not a function of any variate, FEC-MRK can use the correlation function as a function of response variate. For example, using exponential correlation function, $\sigma_{g,h}$ component in Eq.20 which will be multiplied to the spatial covariance matrix between $\boldsymbol{x_i}$ and $\boldsymbol{x_j}$ can be expressed as $b_{gh} \times \exp\left[\left|y_{g,i} - y_{h,j}\right|^2\right]$.

Since the form of covariance matrix $\boldsymbol{C}$ is different in two meta-model, the number of hyper-parameter in MLE process is different. Table 1 and Fig.2 show the comparison of the number of hyper-parameters for single-response GPR model, RC-MRK, and FEC-MRK. It is assumed that the correlation function in Eq.14 is used for the covariance matrix for all the cases. As shown in the table and graph, the number of hyper-parameters for single-response GRP is always larger than multi-response Kriging since we need separate responses for each output. FEC-MRK has $2q-3$ more hyper-parameters compared to RC-MRK. Since we will deal with $q > 1$, FEC-MRK always has more hyper-parameters than RC-MRK.

For example, if we are constructing a GPR model with two input covariates and three response covariates, the number of hyper-paramters in RC-MRK is 14, and the number of hyper-parameters in FEC-MRK is 17, while single-response GPR model uses only 5. However, since three single-response GRP models should be constructed separately, total number of hyper-parameters for single-responses GPR model will be 15.

E. Multi-response Kriging with iterative Maximum Likelihood Estimation

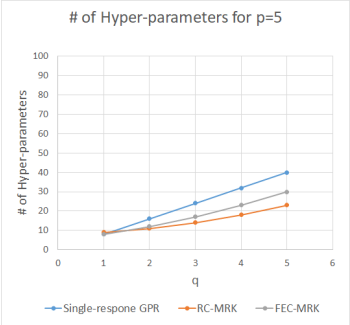
Even the number of hyper-parameters are the largest for single-response GPR model, since they are obtained in separate models, the number of hyper-parameters which should be calculated in a single model remains small, $p+3$. However, in multi-response GPR model, the number of hyper-parameters is remarkably increases in a single model since the covariance between responses are set as unknown variable.

As the number of variables in the optimization of MLE process increases, the stability of optimization decreases, and the chance to find optimum hyper-parameter is reduced. To guarantee the accuracy of MLE process, two multi-response GPR models are advanced using iterative Maximum Likelihood Estimation process.

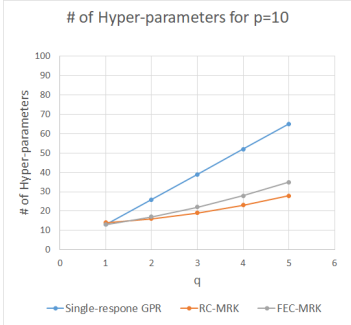
The schematics of iterative Maximum Likelihood Estimation process is shown in Fig. 3. The iterative Maximum Likelihood Estimation process separates hyper-parameters into two groups; hyper-parameters for the covariance between samples and hyper-parameters for the covariance between responses. Each group of hyper-parameters are optimized through two separated sub-optimization in order.

First of all, hyper-parameters are initialized with constant numbers. In the next step, which is sub-optimization 1 in Fig.3, the hyper-parameters for the covariance between samples are optimized to have maximum likelihood. To avoid the numerical instability of matrix inversion during the optimization process, the range of design variable is restricted with step size Δ.

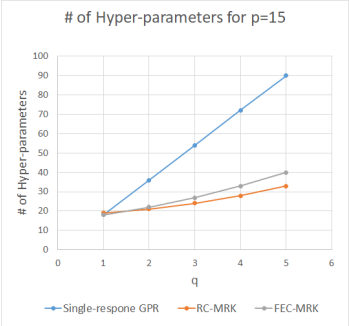
For example, in RC-MRK model, the initial values of hyper-parameters are set as $a_0^{(0)}$, $a_1^{(0)}$, $\nu_0^{(0)}$, $\eta_d^{(0)}$, $S_{11}^{(0)}$, $S_{22}^{(0)}$, $\cdots$, $S_{qq}^{(0)}$, $\phi_{11}^{(0)}$,$\phi_{21}^{(0)}$,$\phi_{22}^{(0)}$, $\cdots$ $\phi_{qq}^{(0)}$. Then, the maximum likelihood L_1 is optimized w.r.t a_0, a_1, ν_0, η_d, and S_{11}, S_{22}, $\cdots$, S_{qq}. The range of design variables are set as $\pm\Delta$ of initial values. The hyper-parameters for the covariance between response are fixed as same as initial values. The detailed definition of the optimization problem of this step is shown below.



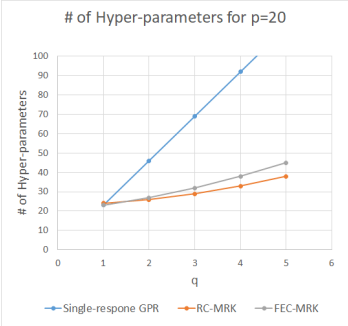
(a) The number of input variates = 5



(b) The number of input variates = 10



(c) The number of input variates = 15



(d) The number of input variates = 20

Fig. 2 The comparison of the number of hyper-parameter for different input and output variate

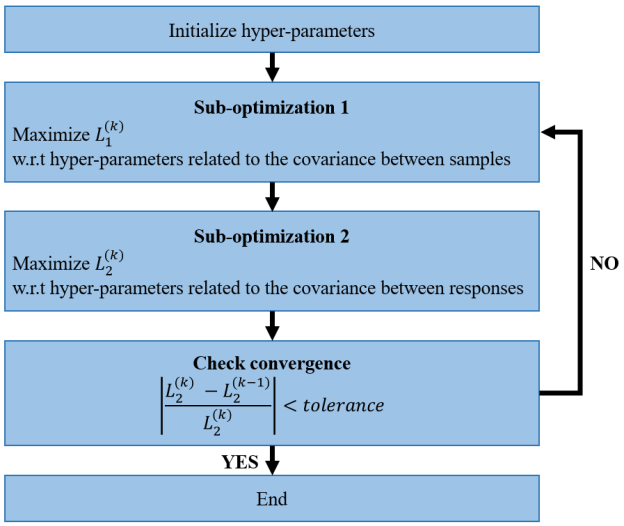


Fig. 3 The schematics of iterative Maximum Likelihood Estimation (k = iteration number)

$$
\begin{aligned}
\text{maximize} \quad & Likelihood\ L_1^{(1)} \\
\text{w.r.t} \quad & a_0, a_1, \nu_0, \eta_d, S_{11}, S_{22}, \cdots, S_{qq}, \\
\text{subject to} \quad & a_0^{(0)} - \Delta \le a_0 \le a_0^{(0)} + \Delta, \\
& a_1^{(0)} - \Delta \le a_1 \le a_1^{(0)} + \Delta, \\
& \nu_0^{(0)} - \Delta \le \nu_0 \le \nu_0^{(0)} + \Delta, \\
& \eta_d^{(0)} - \Delta \le \eta_d \le \eta_d^{(0)} + \Delta, \\
& S_{11}^{(0)} - \Delta \le S_{11} \le S_{11}^{(0)} + \Delta, \\
& \vdots \\
& S_{qq}^{(0)} - \Delta \le S_{qq} \le S_{qq}^{(0)} + \Delta,
\end{aligned}
\tag{25}
$$

In the third step, the hyper-parameters for the covariance between responses are optimized to have maximum likelihood. In this optimization process, as same as the previous sub-optimization, the range of design variables are set as $\pm\Delta$ of initial values. The hyper-parameters related to the spatial covariance are set as the value which is obtained in the previous step. In RC-MRK model, the optimization problem can be written as below.

$$\begin{aligned}
&\text{maximize} && Likelihood\ L_2^{(1)} \\
&\text{w.r.t} && \phi_{11}, \phi_{21}, \phi_{22}, \cdots, \phi_{qq}, \\
&\text{subject to} && \phi_{11}^{(0)} - \Delta \le \phi_{11} \le \phi_{11}^{(0)} + \Delta, \\
& && \phi_{21}^{(0)} - \Delta \le \phi_{21} \le \phi_{21}^{(0)} + \Delta, \\
& && \phi_{22}^{(0)} - \Delta \le \phi_{22} \le \phi_{22}^{(0)} + \Delta, \\
& && \qquad \vdots \\
& && \phi_{qq}^{(0)} - \Delta \le \phi_{qq} \le \phi_{qq}^{(0)} + \Delta, \\
& && a_0 = a_0^{(1)}, a_1 = a_1^{(1)}, \nu_0 = \nu_0^{(1)}, \eta_d = \eta_d^{(1)}, S_{11} = S_{11}^{(1)}, S_{22} = S_{22}^{(1)}, \cdots S_{qq} = S_{qq}^{(1)}
\end{aligned} \tag{26}$$

where $a_0^{(1)}, a_1^{(1)}, \nu_0^{(1)}, \eta_d^{(1)}, S_{11}^{(1)}, S_{22}^{(1)}, \cdots S_{qq}^{(1)}$ are the hyper-parameters for the spatial covariates which are obtained from Eq.26. The process is iterate until the relative error of L_2 is less than the tolerance. (The fourth step in Fig.3) In this paper, strong termination criteria is applied; the error should be less than $1.0E-10$.

F. Implementation into Efficient Global Optimization (EGO)

Efficient Global Optimization (EGO) is widely using optimization technique for the computationally expensive black-box functions [36]. In EGO process, we can set five design steps as shown in Fig. 4. First of all, small number of initial samples are randomly selected. In this research, Latin Hypercube Sampling (LHS) is applied to construct sets of sample points [37]. In the second step, the function value of the samples are evaluated with computation or experiment. In third, surrogate model is constructed for the observed sample. For single objective problem, Kriging surrogate model is generally used in this step. Using the surrogate model, the optimum is explored in the design space. Since we have surrogate model, the function value at the trial location can be calculated with very small computational cost. Therefore, we can use gradient-free optimization such as Genetic Algorithm(GA) [38].

If the error between the predicted optimum by the surrogate model, and the value calculated from the function analysis is lower than the termination criteria, the design process is terminated. Otherwise, the design process continues to the next step, which is adaptive sampling with infill sampling criteria.(The fifth step in Fig. 4) Since we begin with small number of initial samples, the global and local accuracy of the surrogate should not be accurate at the early design iteration, and this step is significant in overall EGO design process.In this paper, to improve both exploration of design space to improve the global accuracy of the model and exploitation to improve the local accuracy of the fit around the global optimum, Multi-Point Multi-Objective Infill Sampling Criteria (MPMO-ISC) by Yi et al. is used [39]. In MPMO-ISC, multi-objective optimization is performed to provide a Pareto front with local and global accuracy which are represented by optimum value and variance from Kriging model, respectively. Using the Parato front, multiple number of additional samples can be selected to improve both local and global accuracy of the surrogate model. The details of MPMO-ISC is introduced in Ref [39].

In traditional multidisciplinary design optimization, if there are more than one discipline, there should be same number of individual surrogate models which are totally separated. However, using multi-response GPR model, only one GPR model is constructed, and it gives estimated function valued which are correlated by covariance terms; ϕ_{11},ϕ_{21},ϕ_{22}, $\cdots$ ϕ_{qq} in RC-MRK and L matrix in FEC-MRK. (Fig. 5) Therefore, more accurate function estimation is available, and fewer number of samples are required for the convergence of optimization. The validation of improvement in global accuracy and local accuracy of optimization process using multi-response GPR model will be provided in the next section.

IV. Validation of Multi-Response GPR model

In this section, the global accuracy of multi-response GPR methods will be validated with analytic test functions. Moreover, the meta-models are implemented into Efficicent Global Optimization (EGO) process, and the local accuracy

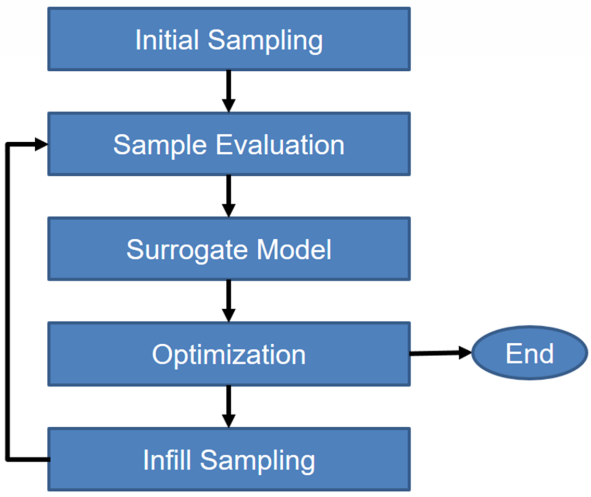


Fig. 4 Schematic of EGO process

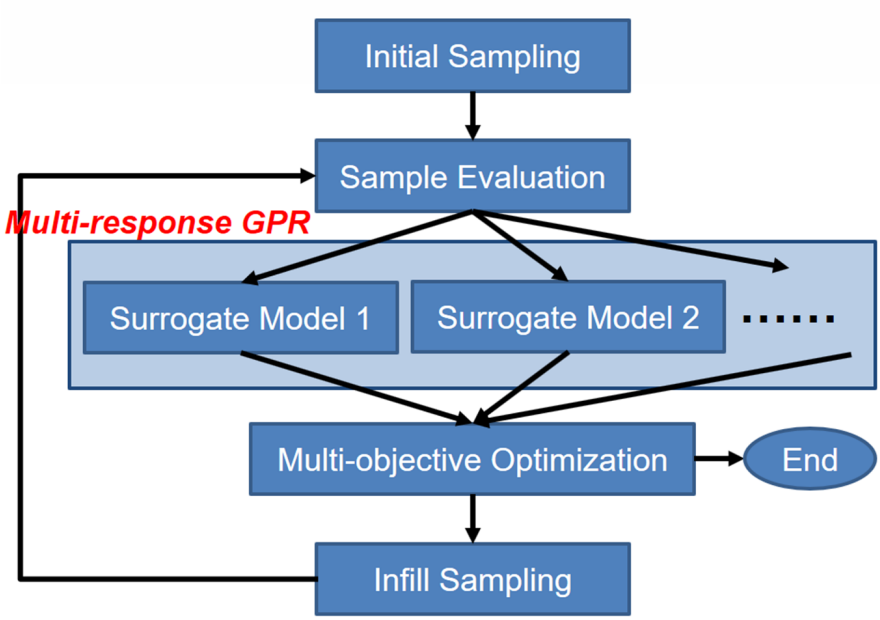


Fig. 5 Schematic of EGO process using multi-response GPR model

and efficiency as a surrogate model in the optimization process is verified. The results are compared with the result with single-response GPR model.

A. Global accuracy test

To validate the improvement in global accuracy using multi-response GPR model, the global accuracy test is performed. In this test, single- and multi-response GPR models are created for the test analytic functions with multiple discipline and root mean square errors(RMSE) are compared. RMSE is computed by comparing with the exact analytic function with constructed GPR model. The observed samples are selected using Latin Hypercube Sampling(LHS) [37]. To consider the randomness of Latin Hypercube Sampling, the result of 20 sample sets are averaged.

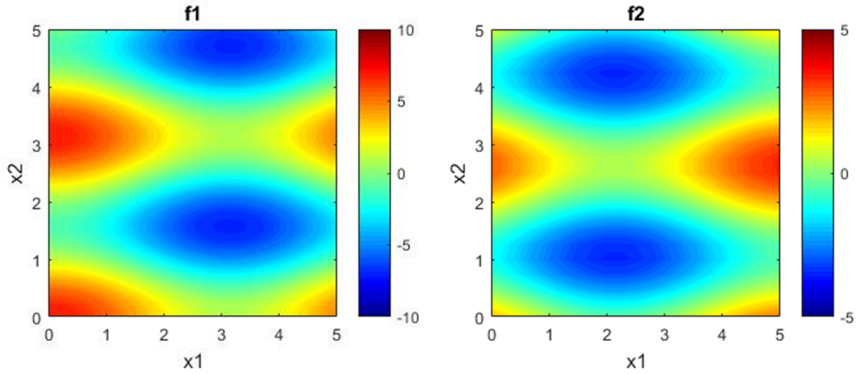


Fig. 6 The contour of weakly coupled functions

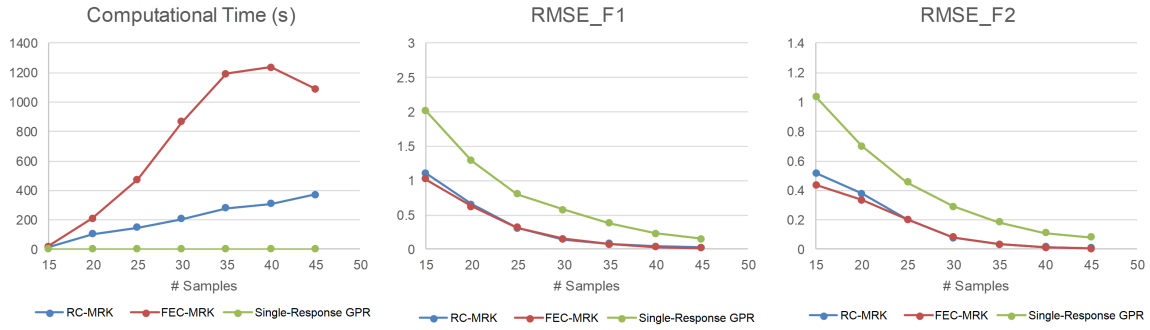


Fig. 7 The result of global accuracy test for weakly coupled function

1. Weakly coupled function

As the weakly coupled function, the test function by Wang et al is used [7]. The test function has two design variables, and two responses, which are correlated. To enhance the coupling between two responses, the functions are modified as shown below. Both design variables x_1 and x_2 varies from 0.0 to 5.0. The contours of exact function value are shown in Fig. 6.

$$\begin{aligned} f_1 &= 3\cos(x_1) + 4\cos(2x_2) \\ f_2 &= 3/2\cos(x_1 + 1) + 2\cos(2x_2 + 1) \end{aligned} \tag{27}$$

To compare the RMSE and computational time to find the hyper-parameters of each GPR model, Monte Carlo simulation [40] has been performed. From 15 number of samples, the number of sample is increased by 5. For each different number of sample, 20 sample sets are randomly picked using Latin Hypercube Sampling (LHS). The RMSE and computational time is averaged for these 20 sample sets. The RMSE is calculated with 21×21 evenly spread test samples.

Fig. 7 shows the result of global accuracy test for weakly coupled function. For the comparison, the single response GPR model is constructed by DACE, which is Matlab Kriging code [32]. Since RC-MRK and FEC-MRK methods uses huge matrix inversion and iterative MLE process, the computational time is much larger. While single response GPR model takes less than a second, multi-response GRP method takes a minutes to 25 minutes by different number of samples. However, since the correlation between responses are considered, the average RMSE is remarkably lower in both multi-response GPR models.

2. Strongly coupled function

As the strongly coupled function, the analytic function which is previously solved by Sellar et al is used [41]. The test function has three design variables, and two responses, which are strongly correlated by each other response. The functions are as shown below. The contours of exact function value are shown in Fig. 8.

$$
\begin{aligned}
f_1 &= z_1^2 + x_1 + x_2 - 0.2 f_2 \\
f_2 &= \sqrt{f_1} + z_1 + x_2 \\
0 &\le x_1 \le 10 \\
0 &\le x_2 \le 10 \\
-10 &\le z_1 \le 10
\end{aligned}
\tag{28}
$$

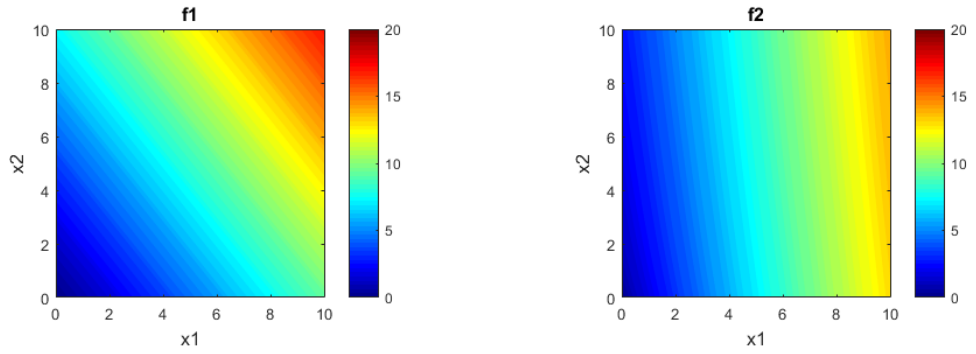


Fig. 8 The iso-contour of strongly coupled functions ($z_1 = 0$)

As shown in the previous test, the computational time to find the hyperparameters of each GPR model and RMSE is compared. From 15 number of samples, the number of sample is increased by 5. For each different number of sample, 20 sample sets are randomly picked using Latin Hypercube Sampling. The RMSE and computational time is averaged for these 20 sample sets. The RMSE is calculated with $11 \times 11 \times 11$ evenly spread test samples.

Fig. 9 shows the result of global accuracy test for strongly coupled function. For the comparison, the single response GPR model is constructed by DACE, which is Matlab Kriging code. As shown in the previous global accuracy test with weakly coupled function, the computational time of multi-response GPR method is much larger than single response GRP model. However, the average RMSE is generally lower in both multi-response GPR models for all the functions.

B. Optimization problems with analytic functions

To verify the efficiency and local accuracy of Efficient Global Optimization(EGO) process using multi-response GPR model in multidisciplinary optimization problems, the optimization problems are solved with analytic functions. The number of iterations, the number of samples for the convergence, and RMSE of converged responses are compared for weakly and strongly coupled problems.

1. Weakly coupled function

The first test comes with weakly coupled function. The weakly coupled functions in the previous subsection is used as two different disciplines, (Eq. 27) but the high-level objective functions are added to consider the optimization problem. The problem is described as below. The exact contour of function F is shown in Fig. 10. The function has one global minimum at (x_1, x_2) =(2.6416,1.2847), and one local minimum.

$$
\begin{aligned}
\underset{x_1,x_2}{\text{minimize}} \quad & F = f_1 + 0.5 f_2 + x_2 \\
\text{subject to} \quad & 0 \le x_1 \le 5, \\
& 0 \le x_2 \le 5
\end{aligned}
\tag{29}
$$

For this optimization problem, the number of initial samples are set as 10, and randomly picked with LHS. For the comparison, DACE, RC-MRK and FEC-MRK are used as the surrogate model, and the results are compared. As the adaptive sampling process, one sample at the optimum of each EGO iteration is added to improve the local accuracy of the surrogate model, and samples at the maximum prediction uncertainty in each surrogate model is added to improve

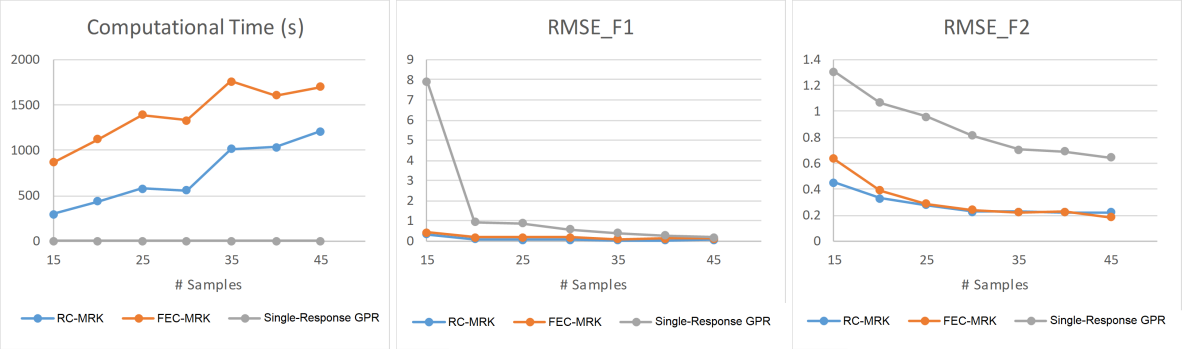


Fig. 9 The result of global accuracy test for strongly coupled function

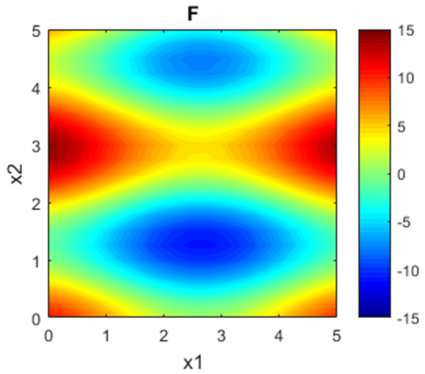


Fig. 10 The contour of objective function for weakly coupled case

the global accuracy. The design iteration is terminated if the relative error of optimum function value is lower than 10^{-4}. The relative error of optimum function value is calculated as below.

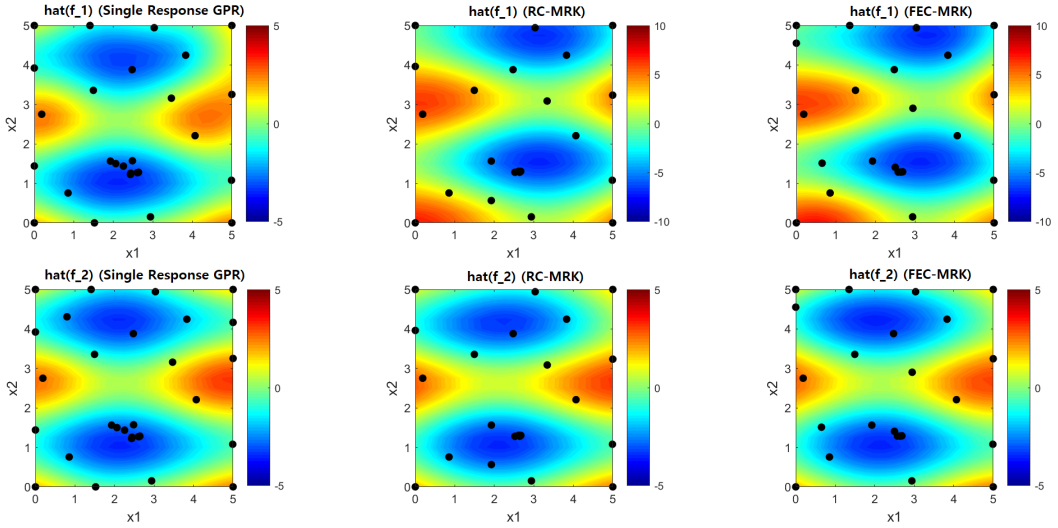
$$error = \left| \frac{F_{opt}^{k} - F_{opt}^{k-1}}{F_{opt}^{k}} \right| \tag{30}$$

where k is the iteration number. To consider the randomness of the initial sampling, the test has been performed 20 times, and the results are averaged.

Fig. 11 shows the comparison of F contour at the last iteration for one test case. Fig. 12 shows the result of EGO test for weakly coupled functions. Compared to the single response GPR, both multi-response GRP metamodel used less number of design iterations and samples for the convergence. Especially, the number of samples to find the optimum is decreased about 20%. Moreover, comparing the average RMSE at the last design iteration, both multi-response GPR models have lower RMSE even fewer number of samples are used. Therefore, the global accuracy of the surrogate model during the EGO process is promising.

2. Strongly coupled function

The second test comes with strongly coupled function. The strongly coupled functions in the previous subsection is used as two different disciplines (Eq. 28), but the high-level objective functions are added to consider the optimization problem. The problem is described as below. The exact iso-contour of function F at $z_1 = 0$ is shown in Fig. 13. The objective function has one global minimum at (x_1, x_2, z_1) =(0,0,1.9776).



(a) EGO with Single response GRP
 # samples = 31
 # iteration = 10
RMSE of f_1 = 0.4201
RMSE of f_2 = 0.1769

(b) EGO with RC-MRK
 # samples = 23
 # iteration = 7
RMSE of f_1 = 0.3108
RMSE of f_2 = 0.1648

(c) EGO with FEC-MRK
 # samples = 24
 # iteration = 7
RMSE of f_1 = 0.3778
RMSE of f_2 = 0.1111

Fig. 11 The comparison of f_1 and f_2 contour at the last iteration (Weakly coupled function)

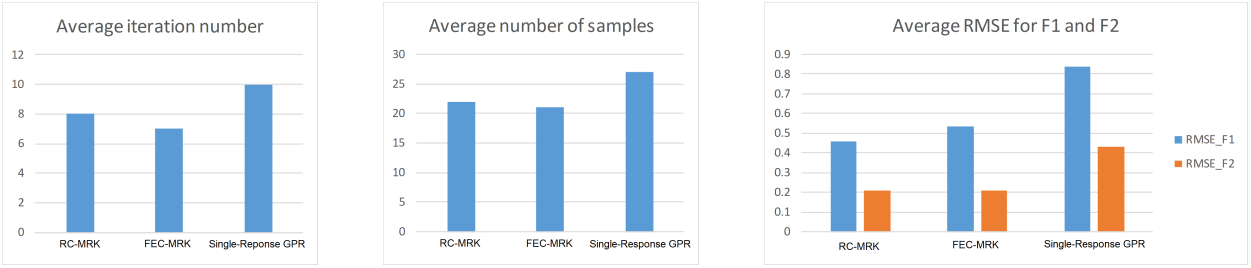


Fig. 12 The result of EGO test for weakly coupled functions

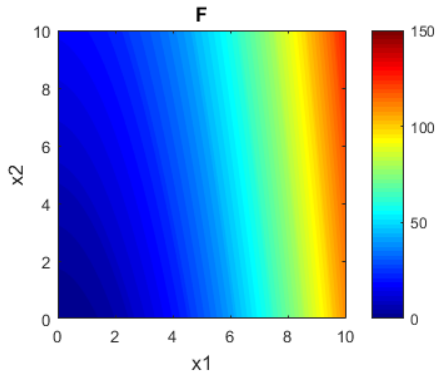


Fig. 13 The contour of objective function for strongly coupled case($z_1 = 0$)

$$\begin{aligned} \underset{x_1,x_2,z_1}{\text{minimize}} \quad & F = x_1^2 + x_2 + f_1 + \exp\left[-f_2\right] \\ \text{subject to} \quad & 0 \le x_1 \le 10, \\ & 0 \le x_2 \le 10, \\ & -10 \le z_1 \le 10, \end{aligned} \tag{31}$$

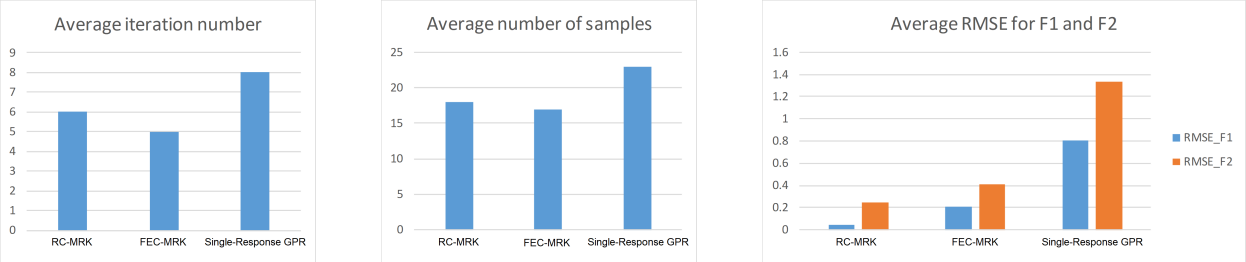


Fig. 14 The result of EGO test for strongly coupled functions

For this optimization problem, the number of initial samples are set as 15, and randomly picked with LHS. For the comparison, DACE, RC-MRK and FEC-MRK are used as the surrogate model, and the results are compared. As the adaptive sampling process, one sample at the optimum of each EGO iteration is added to improve the local accuracy of the surrogate model, and samples at the maximum prediction uncertainty in each surrogate model is added to improve the global accuracy. The design iteration is terminated if the relative error of optimum function value is lower than 10^{-4}. To consider the randomness of the initial sampling, the test has been performed 20 times, and the results are averaged.

Fig.14 shows the result of EGO test for strongly coupled functions. Compared to the single response GPR, both multi-response GRP metamodel used less number of design iterations and samples for the convergence; the number of samples to find the optimum is decreased up to 25%. Moreover, the average RMSE at the last design iteration is dramatically lower with both multi-response GPR models.

From the EGO test, it is verified that for the multidisciplinary design problem, multi-response GPR model can decrease the number of samples which is required for the design convergence. Moreover, since the global accuracy of the surrogate model is better due to the correlation information between the responses, the global optimum of the complex problems can be guaranteed.

V. Applications

In the previous chapter, the accuracy and the cost efficiency of multi-reponse GPR model is validated. In this chapter, the model will be used in the design application.

A. Prediction of Aeroelastic torsion

Multi-response GPR model is a useful tool for most aerospace applications, as the majority of the aerodynamic/structural/thermal etc. since the outputs are highly coupled. A simple example where one can use multi-response GPR is the prediction of aeroelastic torsion of an airfoil due to aerodynamic loading.

In this application, the spring model is used for the analytic model for the prediction of the torsion [42]. Fig. 15 shows the model description. For the lift L and aerodynamic moment M_0 on the airfoil, the aerodynamic torsion can be calculated as shown in Eq. 32.

$$\begin{aligned} M_0 + Le &= K\theta \\ M_0 &= \frac{1}{2}\rho V^2 c^2 C_{M,0} \\ L &= \frac{1}{2}\rho V^2 c C_L \\ C_L &= C_{L,0} + \frac{\partial C_L}{\partial \alpha}(\alpha + \theta) \end{aligned} \tag{32}$$

where e is the distance between aerodynamic center and flexural center, K is spring constant and c is the chord length of airfoil. Solving the equation, aerodynamic torsion θ can be calculated. Aerodynamic analysis is performed with Thin airfoil thoery.

The analytic model outputs the coefficient of lift (C_L) and the angle of twist (θ) with respect to the flexural axis, which are strongly correlated responses. In this application, these two outputs are trained with FEC-MRK meta-model, and the predicted C_L and θ are compared with actual value. For the problem, four design variables are set to control the

Table 2 The boundary of design variables

Parameter	Symbol	Lower bound	Upper bound
Taper ratio	λ	0.5	1
Wing span (m)	b	15	20
Root chord (m)	C_r	2	2.5
Angle of Attack (deg.)	α	0	0.18

shape of a wing planform; taper ratio(λ), span(b), root chord(c_r) and angle of attack(α). The boundary of each design variable is shown in Table 2. For the training data, each of these design variables were randomly sampled using Latin Hypercube Sampling.

A visual depiction of the results is shown in Fig. 16. In Fig. 16a, 20 training samples are used, and in Fig. 16b, 45 training samples are used. For the validation of the model, 50 test samples are selected and at the design location of test samples, the actual C_L and θ are compared with the predicted values. If the algorithm was a perfect interpolator, each red circle would form the circumference of each blue point i.e. perfect overlap. As shown in Fig. 16, we have reasonable accuracy in the both cases, whose RMSE is about 10^{-3}.

B. Airfoil optimization

The application shown in the previous subsection can be extended to an airfoil optimization problem with aeroelastic analysis. In this application, NACA0012 airfoil will be set as a base airfoil, and the airfoil will be designed to maximize the lift coefficient while minimizing the twist associated with the lift. The objective function for the minimization problem is shown in Eq. 33.

$$OBJ = \left(\frac{C_l}{C_l'}\right)^{-1} + \frac{\theta}{\theta'} \tag{33}$$

where C_l is the lift coefficient of the airfoil, and θ is the wing twist. The variables with superscript denote the value from baseline airfoil, which is NACA0012. To perturb the airfoil, the radial basis function (RBF) as a few control points with a specified deflection [43]. Eq.34 shows RBF equation.

$$f(\xi) = \exp\left(\frac{-\xi^2}{2\sigma^2}\right) \tag{34}$$

where ξ is the distance of mesh node to control point, and σ is the radius of influence. In this paper, two control points near leading edge and trailing edge of the airfoil are used. For two control points, we have six design variables, which are horizontal and vertical location and σ for each control point. Fig. 17 shows NACA0012 diagram with control point and design variables.

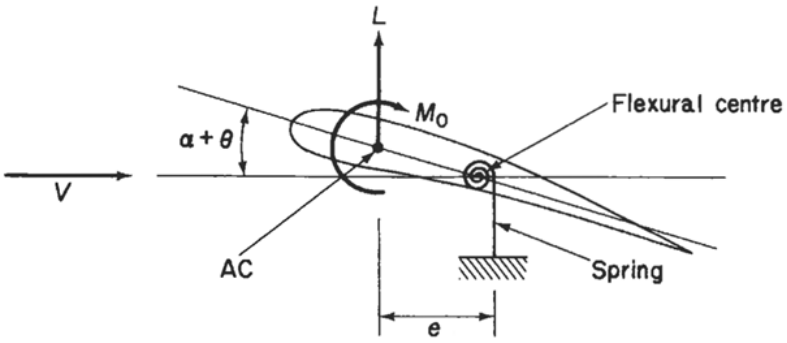
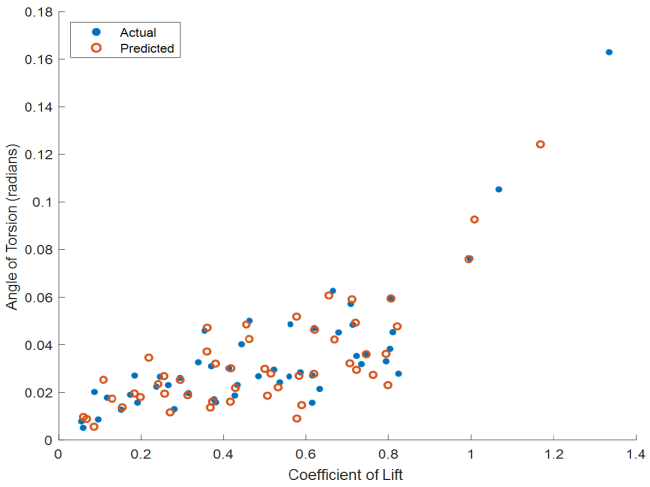
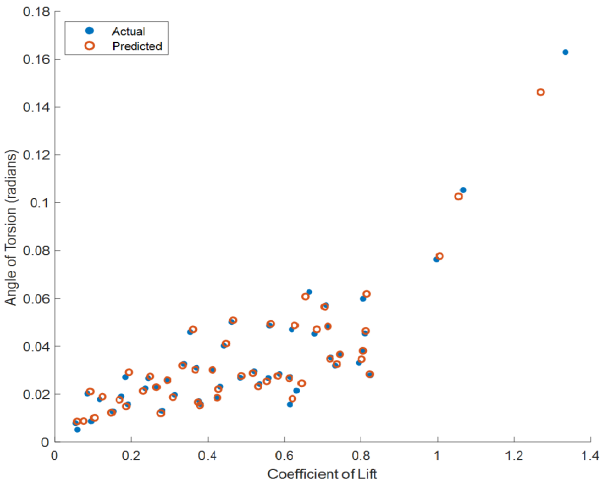


Fig. 15 Aerodynamic force and aeroelastic torsion of the wing



(a) Number of training samples = 20, RMSE $\sim 10^{-3}$



(b) Number of training samples = 45, RMSE $\sim 10^{-4}$

Fig. 16 The result of the prediction of aeroelastic torsion

The aerodynamic condition of the problem is set as $Re = 10^6$ and $Mach = 0.2$. For the aerodynamic analysis, incompressible RANS equations were solved to determine the lift coefficient at zero angle of attack and the lift curve slope using QuickerSim [44]. Fig. 18 shows aerodynamic analysis result of the baseline airfoil (NACA 0012). For the elastic model, spring model in the previous application is used.

Fig. 19 shows the schematic of airfoil optimization process. First of all, the samples are randomly generated using Latin Hypercube Sampling. In this study, 50 training samples are selected. These sample airfoils are analyzed with aerodynamic solver and aeroelasticity solver. From these two solvers, the lift coefficient and wing twist can be calculated,

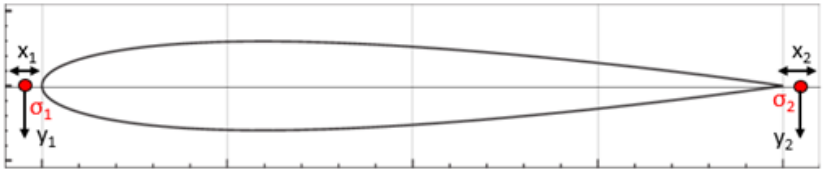


Fig. 17 NACA 0012 diagram with design variables

Table 3 The airfoil optimization result

	C_l	θ	Computational time
Opt. using Multi response GPR model	0.8439	0.0732 rad	1.9407×10^3
Opt. using aero. and aeroelastic solver	0.8548	0.0750 rad	1.1251×10^5

and FEC-MRK model is generated for two responses. Then, the optimization is performed using genetic algorithm [45]. In the genetic algorithm, 50 generations with 100 population size is used.

Table 3 show the optimization result. For the comparison, the optimization which does not us multi-response GRP model, but all the analysis for the optimization has runs using CFD and aeroelastic model has been performed. As shown in Table 3, the optimization using FEC-MRK yields results that are almost as good as the optimum from the full solver. Howver, FEC-MRK based optimization takes approximately 33 minutes whereas the full solver takes approximately 31 hours for the same.

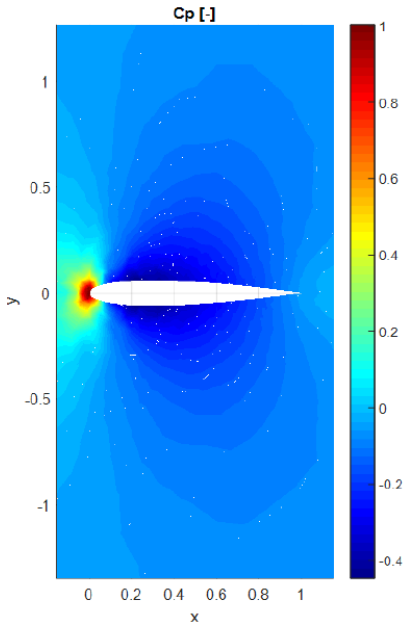


Fig. 18 Aerodynamic analysis result of NACA0012 baseline airfoil

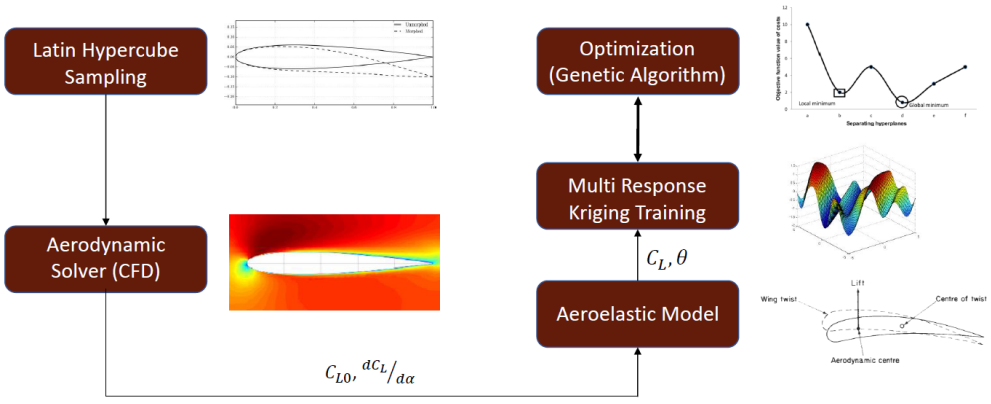


Fig. 19 Schematic of airfoil optimization process

VI. Conclusion

In this paper, a multi-response Gaussian Process Regression model is suggested for multidisciplinary design analysis and optimization. Two different approaches of multi-response GPR model are introduced, and the methods are advanced using iterative Maximum Likelihood Estimation (MLE) process. Since the model can consider the correlation between responses, which are physically related in nature, the accuracy of function estimation is improved. The advance methods are validated with analytic test function, and it is shown that the meta-model is cost efficient and accurate, especially in global accuracy. Moreover, the method is implemented into the Efficient Global Optimization (EGO). Using the method, aeroelasticity design problem, which is one of the famous multidisciplinary design optimization problem, is solved, and the efficiency and the accuracy of design result is compared.

As the future work, the method will be extended to the case whose sample points are different by different responses. In this manner, we can use multi-response GPR model even we do not have all the responses for an arbitrary sample point. This approach can be applied to variable fidelity analysis and optimization.

References

[1] Augustine, N. R., *Augustine's Laws*, 6th ed., American Institute of Aeronautics and Astronautics, Reston, Va, 1997.

[2] *Defense Systems Management College*, Dec. 1991.

[3] Prasad, R., Kim, H. S., Im, D., Choi, S., and Yi, S., "Analysis and Sensitivity Calculation using High Fidelity Spectral Formulation-Based FSI and Coupled Adjoint Method," *AIAA AVIATION Forum, 17th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, AIAA, Washington, DC, 2016. doi:10.2514/6.2016-3993.

[4] Simpson, T. W., Mauery, T. M., Korte, J. J., and Mistree, F., "Kriging Models for Global Approximation in Simulation-Based Multidisciplinary Design Optimization," *AIAA Journal*, Vol. 39, No. 12, 2001, pp. 2233–2241. doi:10.2514/2.1234.

[5] Park, J., Jo, Y., Yi, S., Choi, J.-Y., Raj, P., and Choi, S., "Variable-Fidelity Multidisciplinary Design Optimization for Innovative Control Surface of Tailless Aircraft," *AIAA AVIATION Forum, 34th AIAA Applied Aerodynamics Conference*, AIAA, Washington, DC, 2016. doi:10.2514/6.2016-4038.

[6] Bowcutt, K. G., "Multidisciplinary Optimization of Airbreathing Hypersonic Vehicles," *Journal of Propulsion and Power*, Vol. 17, No. 6, 2001, pp. 1184–1190. doi:10.2514/2.5893.

[7] Wang, B., and Chen, T., "Gaussian process regression with multiple response variables," *Chemometrics and Intelligent Laboratory Systems*, Vol. 142, 2015, pp. 159 – 165. doi:https://doi.org/10.1016/j.chemolab.2015.01.016.

[8] Kleijnen, J. P., and Mehdad, E., "Multivariate versus univariate Kriging metamodels for multi-response simulation models," *European Journal of Operational Research*, Vol. 236, No. 2, 2014, pp. 573 – 582. doi:https://doi.org/10.1016/j.ejor.2014.02.001.

[9] Cressie, N., "Statistics for Spatial Data," *Terra Nova*, Vol. 4, No. 5, 1992. doi:10.1111/j.1365-3121.1992.tb00605.x.

[10] Myers, D. E., "Matrix formulation of co-kriging," *Journal of the International Association for Mathematical Geology*, Vol. 14, No. 3, 1982, pp. 249–257. doi:10.1007/BF01032887.

[11] Chung, H.-S., and Alonso, J., "Using gradients to construct cokriging approximation models for high-dimensional design optimization problems," *40th AIAA Aerospace Sciences Meeting Exhibit*, AIAA, Reno,NV, 2002. doi:10.2514/6.2002-317.

[12] Forrester, A. I. J., Sobester, A., and Keane, A. J., *Engineering Design via Surrogate Modelling - A Practical Guide.*, Wiley, 2008.

[13] Yamazaki, W., Rumpfkeil, M., and Mavriplis, D., "Design Optimization Utilizing Gradient/Hessian Enhanced Surrogate Model," *28th AIAA Applied Aerodynamics Conference, Fluid Dynamics and Co-located Conferences*, AIAA, Chicago, IL, 2010. doi:10.2514/6.2010-4363.

[14] Laurenceau, J., and Sagaut, P., "Building Efficient Response Surfaces of Aerodynamic Functions with Kriging and Cokriging," *AIAA Journal*, Vol. 46, No. 2, 2008, pp. 498–507. doi:10.2514/1.32308.

[15] Han, Z.-H., Zimmermann, R., and Goretz, S., "A New Cokriging Method for Variable-Fidelity Surrogate Modeling of Aerodynamic Data," *48th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition, Aerospace Sciences Meetings*, AIAA, 2010. doi:10.2514/6.2010-1225.

[16] Han, Z., Zimmerman, R., and GÃ¶rtz, S., "Alternative Cokriging Method for Variable-Fidelity Surrogate Modeling," *AIAA Journal*, Vol. 50, No. 5, 2012, pp. 1205–1210. doi:10.2514/1.J051243.

[17] Han, Z.-H., Görtz, S., and Zimmermann, R., "Improving variable-fidelity surrogate modeling via gradient-enhanced kriging and a generalized hybrid bridge function," *Aerospace Science and Technology*, Vol. 25, No. 1, 2013, pp. 177 – 189.

[18] Jo, Y., Choi, S., and Lee, D., "Variable-Fidelity Design Method Using Gradient-Enhanced Kriging Surrogate Model with Regression," *14th AIAA Aviation Technology, Integration, and Operations Conference, AIAA AVIATION Forum*, AIAA, 2014. doi:10.2514/6.2014-2867, 0.

[19] Neal, R. M., *Bayesian Learning for Neural Networks*, Springer-Verlag, Berlin, Heidelberg, 1996.

[20] Wang, S., and Ng, S. H., "A joint Gaussian process metamodel to improve quantile predictions," *2017 Winter Simulation Conference (WSC)*, 2017, pp. 1891–1902.

[21] Conti, S., and OâEUR(tm)Hagan, A., "Bayesian emulation of complex multi-output and dynamic computer models," *Journal of Statistical Planning and Inference*, Vol. 140, No. 3, 2010, pp. 640–651.

[22] O'Hagan, A., *Some Bayesian Numerical Analysis. In : Bayesian statistics*, Vol. 4, Oxford University Press, New York, 1992.

[23] Ver Hoef, J. M., and Barry, R. P., "Constructing and fitting models for cokriging and multivariable spatial prediction," *Journal of Statistical Planning and Inference*, Vol. 69, No. 2, 1998, pp. 275–294.

[24] Goulard, M., and Voltz, M., "Linear coregionalization model: Tools for estimation and choice of cross-variogram matrix," *Mathematical Geology*, Vol. 24, No. 3, 1992, pp. 269–286. doi:10.1007/BF00893750.

[25] Álvarez, M. A., and Lawrence, N. D., "Computationally Efficient Convolved Multiple Output Gaussian Processes," *J. Mach. Learn. Res.*, Vol. 12, 2011, pp. 1459–1500.

[26] Hong, X., Ren, L., Chen, L., Guo, F., Ding, Y., and Huang, B., "A weighted Gaussian process regression for multivariate modelling," *2017 6th International Symposium on Advanced Control of Industrial Processes (AdCONIP)*, 2017, pp. 195–200.

[27] Hong, X., Huang, B., Ding, Y., Guo, F., Chen, L., and Ren, L., "Multi-model multivariate Gaussian process modelling with correlated noises," *Journal of Process Control*, Vol. 58, 2017, pp. 11–22.

[28] Sadoughi, M., Li, M., and Hu, C., "Multivariate system reliability analysis considering highly nonlinear and dependent safety events," *Reliability Engineering & System Safety*, Vol. 180, 2018, pp. 189–200.

[29] Rasmussen, C., and Williams, C., *Gaussian Processes for Machine Learning*, Adaptative computation and machine learning series, University Press Group Limited, 2006.

[30] Stein, M. L., *Interpolation of Spatial Data: Some Theory for Kriging (Springer Series in Statistics)*, 1st ed., Springer, 1999.

[31] Roberts, S., Osborne, M., Ebden, M., Reece, S., Gibson, N., and Aigrain, S., "Gaussian processes for time-series modelling," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, Vol. 371, 2013. doi:10.1098/rsta.2011.0550.

[32] Nielsen, H. B., Lophaven, S. N., and Søndergaard, J., "DACE - A Matlab Kriging Toolbox," *Informatics and Mathematical Modeling, Technical Report, IMM-TR-2002-12*, 2002.

[33] Santner, T. J., Williams, B. J., Notz, W., and Williams, B. J., *The design and analysis of computer experiments*, Vol. 1, Springer, 2003.

[34] Fricker, T. E., Oakley, J. E., and Urban, N. M., "Multivariate Gaussian Process Emulators With Nonseparable Covariance Structures," *Technometrics*, Vol. 55, No. 1, 2013, pp. 47–56. doi:10.1080/00401706.2012.715835.

[35] Svenson, J., and Santner, T., "Multiobjective optimization of expensive-to-evaluate deterministic computer simulator models," Vol. 94, 2015.

[36] Jones, D. R., Schonlau, M., and Welch, W. J., "Efficient Global Optimization of Expensive Black-Box Functions," *Journal of Global Optimization*, Vol. 13, No. 4, 1998, pp. 455–492. doi:10.1023/A:1008306431147.

[37] Tang, B., "Orthogonal Array-Based Latin Hypercubes," Vol. 88, No. 424, 1993, pp. 1392–1397. doi:10.2307/2291282.

[38] Davis, L., "Handbook of genetic algorithms," 1991.

[39] YI, S., Kwon, H. I., and Choi, S., "Efficient Global Optimization using a Multi-point and Multi-objective Infill Sampling Criteria," *52nd Aerospace Sciences Meeting, AIAA SciTech Forum*, AIAA, 2014. doi:10.2514/6.2014-0898.

[40] Mooney, C. Z., *Monte Carlo simulation*, Thousand Oaks, Calif. : Sage Publications, 1997.

[41] Sellar, R., Batill, S., and Renaud, J., "Response surface based, concurrent subspace optimization for multidisciplinary system design," *34th Aerospace Sciences Meeting and Exhibit, Aerospace Sciences Meetings*, AIAA, Reno,NV, 1996. doi:10.2514/6.1996-714.

[42] Hodges, D. H., and Pierce, G. A., *Introduction to Structural Dynamics and Aeroelasticity*, 2nd ed., Cambridge Aerospace Series, Cambridge University Press, 2011. doi:10.1017/CBO9780511997112.

[43] Morris, A. M., Allen, C. B., and Rendall, T. C. S., "CFD-based optimization of aerofoils using radial basis functions for domain element parameterization and mesh deformation," *International Journal for Numerical Methods in Fluids*, Vol. 58, No. 8, 2008, pp. 827–860. doi:10.1002/fld.1769.

[44] "QuickerSim," https://quickersim.com/, ????

[45] Mitchell, M., *An Introduction to Genetic Algorithms*, MIT Press, Cambridge, MA, USA, 1998.