공학석사 학위논문

Efficient Global Sensitivity Analysis Using Adaptively-learned Gaussian Process

적응형 가우시안 프로세스를 이용한 효율적인 전역 민감도 해석

2023년 8월

서울대학교 대학원

건설환경공학부

김 유 진

Efficient Global Sensitivity Analysis Using Adaptively-learned Gaussian Process

적응형 가우시안 프로세스를 이용한 효율적인 전역 민감도 해석

지도교수 송 준 호

이 논문을 공학석사 학위논문으로 제출함
2023년 8월

서울대학교 대학원
건설환경공학부
김 유 진

김유진의 공학석사 학위논문을 인준함
2023년 8월

Chair _____(Seal)
Vice Chair _____(Seal)
Examiner _____(Seal)

Abstract

Sensitivity analysis (SA), which is the study of how uncertainty in the 'output' of a model can be apportioned to different sources of uncertainty in the model 'input', have been extensively performed in fields of engineering. Especially, many studies on global sensitivity analysis (GSA) methodologies have been conducted based on the Sobol' method to take advantage of its applicability to various systems including nonlinear ones.

However, in real engineering problems described by complex system models, i.e., high-dimensional and highly-nonlinear systems, sampling-based SA method has limitations in that the analysis time increases with the input dimension and it is also challenging to apply the surrogate modeling-based SA method to highly nonlinear structures.

In this study, a new framework of GSA using actively-learned gaussian process is introduced to efficiently perform GSA for computationally expensive system models. The adaptively-learned gaussian process is intended to accurately build a surrogate model even with much fewer model evaluations. Additionally, to pursue both exploration and exploitation in the adaptive learning, Leave-One-Out Cross-Validation (LOOCV) error is applied as a weight to the objective function of the existing design-of-experiment (DoE) method, which only considers the prediction variance. Furthermore, we propose modified objective function for the adaptively-learned GP, which allows the proposed algorithm to adaptively suggest an appropriate number of experiments to perform effective GSA.

The validity and efficiency of the proposed GSA framework are demonstrated

through two structural numerical examples: 7-story shear building and 9-story frame structure. The results confirm that the method facilitate convergence to the accurate sensitivity index with a significantly reduced number of structural model evaluations.

Keywords: global sensitivity analysis, adaptive learning, gaussian process, design of experiments

Student Number: 2021-29502

Table of Contents

LIST OF FIGURES ..... **v**

LIST OF TABLES ..... **v**

Chapter 1. Introduction **1**

1.1. Research Background 1

1.2. Research Objectives 5

1.3. Outline 6

Chapter 2. Theoretical Background **7**

2.1. Variance-based Global Sensitivity Analysis (GSA) 7

2.2. Gaussian Process (GP) Formulation 10

2.3. Adaptive DoE formulation 14

Chapter 3. Proposed Algorithm of GSA using Adaptive GP **21**

3.1. GSA using Adaptive GP 21

3.2. GSA using Modified Adaptive GP 24

Chapter 4. Verification of Proposed Method **27**

4.1. Example 1: 7-story Shear Building 27

4.2. Example 2: 9-story Frame Building 34

Chapter 5. Conclusions **41**

REFERENCES **42**

국문 초록 **44**

감사의 글 **46**

List of Figures

Figure 3.1 Flow chart of GSA using adaptive GP 23
Figure 3.2 Flow chart of GSA using modified adaptive GP 24
Figure 4.1 Configuration of the 7-story shear building 29
Figure 4.2 Material used for the 7-story shear building 29
Figure 4.3 Ground acceleration time history of El Centro earthquake 30
Figure 4.4 First order Sobol' index of example 1, case 1 31
Figure 4.5 First order Sobol' index of example 1, case 2 32
Figure 4.6 First order Sobol' index of example 1, case 3 32
Figure 4.7 Convergence graph along the number of evaluations, example 1 33
Figure 4.8 Configuration of the 9-story frame 35
Figure 4.9 Material used for the 9-story frame 35
Figure 4.10 Pushover curve of the 9-story frame 36
Figure 4.11 Quasi-Cyclic test curve of the 9-story frame 36
Figure 4.12 First order Sobol' index of example 2 37
Figure 4.13 Convergence graph along the number of evaluations, example 2 38
Figure 4.14 Convergence graph with confidence level, example 2 39
Figure 4.12 NRMSE, example 2 40

List of Tables

Table 4.1 Input variables of the shear building 30
Table 4.2 The number of experiments and computation cost 31
Table 4.3 Input variables of 9-story frame 36
Table 4.4 The number of experiments and computation cost of 9-story frame 37

Chapter 1. Introduction

1.1. Research Background

Sensitivity analysis (SA) is the study of how uncertainty in the 'output' of a model can be apportioned to different sources of uncertainty in the model 'input' (Saltelli et al., 2008). SA has been extensively used in engineering and scientific fields to address several purposes of system analysis and modeling: (a) *dimensionality reduction* to identify uninfluential factors in a system that may be redundant and fixed or removed in subsequent analyses (e.g., Sobol' et al., 2007); (b) *data worth assessment* to identify processes, parameters and scales that dominantly control a system, for which new data acquisition reduces targeted uncertainty the most (e.g., Guillaume et al., 2019; Partington et al., 2020); and (c) *decision support* to quantify the sensitivity of an expected outcome to different decision options, constraints, assumptions and/or uncertainties (Tarantola et al., 2002).

There are two main categories in SA: local sensitivity analysis (LSA) and global sensitivity analysis (GSA). LSA focuses on sensitivity around a 'nominal point' in the problem space. However, LSA has limitations as it provides only a localized view of the problem space at the nominal point, especially when investigating parameter importance in mathematical modeling (Saltelli and Annoni, 2010; Saltelli et al., 2019). To address this limitation, GSA, which offers a comprehensive representation of how different factors interact across the entire problem space to influence the system's output, has gained its prominence.

Many methodologies have been conducted to perform GSA, such as the Fourier

amplitude sensitivity test (FAST), methods based on correlation ration, Kullback-Leibler divergence based approaches, and Sobol' indices related methods. Among these methods, the Sobol' sensitivity indices method based on variance-decomposition is a prominent one owing to its applicability to a wide range of models, including both linear and nonlinear models.

The simplest way to compute the Sobol' indices is to implement double-loop Monte Carlo simulation (MCS), which requires a large number of evaluations of the prediction model and is unaffordable if the prediction model is computationally expensive. To reduce the computational cost of the Sobol' indices, various algorithms have been proposed roughly in the following two ways: *metamodel-based* methods and *sample-based* methods. As explained earlier, the basic sample-based method for GSA is the double-loop MCS based purely on model evaluations and the double-loop MCS has expensive computational cost which is proportional to both dimension of the input space and the number of a sample points. Thus, most of various sample-based methods are developed to reduce the computational cost of SA. For example, Sobol' (1990) discussed how to efficiently estimate the Sobol' indices using MCS and the scheme was more accurately developed by Homma and Saltelli (1996). In this scheme, the required number of samples in the double-loop procedure is reduced to a number which is proportional to the dimension of the input variables. Similarly, Glen and Issac (2012) developed an approach to compute the Sobol' indices by switching the columns of two separately generated MCS sample matrices, which is usually called quasi-MCS.

In the metamodel-based methods, the original system model, which is computationally expensive, is replaced with a cheaper surrogate model such as a

gaussian process model, polynomial chaos expansion model or regression model. After building the surrogate model, the Sobol' indices can be obtained using either analytical or direct MCS-based methods. Sudret proposed that if the original model is approximated by a polynomial chaos expansion (PCE), the Sobol' index can be calculated by post-processing the PCE coefficients. Chen et al. proposed another analytical method for commonly used surrogate models such as the linear regression model, Gaussian process model, Gaussian radial basis model, and MARS model and analytical solution of the index is available if the inputs are normally or uniformly distributed. These analytical methods reduce the number of model evaluations significantly, but may require: 1) extra approximations and assumptions, and 2) extra computational cost in building the surrogate model.

In practical engineering applications, the true analytical form of the given system may be unknown especially for the case of the computationally expensive systems. In that situation, sampling-based methods cannot be easily adopted due to the fact that the computational time rapidly increases with the required number of sample points. Metamodel-based methods are applicable in this situation since they approximate expensive model to a cheaper one, but limitations exist in the aspect of computational time due to the fact that they still need training points obtained by evaluating the true model, which is very time-consuming.

Therefore, the main objective of this thesis is to develop a more efficient surrogate-based method whose training points for building a surrogate are obtained by adaptive gaussian process in order to reduce the number of model evaluations.

1.2. Research Objectives

This study has two main objectives. The first one is to suggest how to build a surrogate model using adaptive gaussian process for the accurate and effective SA of the complex systems, which requires expensive computational cost. Because surrogate modeling needs several experiments to train the surrogate model, it takes time to evaluate the true system to obtain the experiments. Therefore, this study proposes using adaptive learning to build the surrogate, and then, once the surrogate has been built, one would produce a number of test points with cheap evaluation cost to obtain Sobol' indices.

Second, this study proposes an adaptive gaussian process method with modified convergence criteria, which is more appropriate in calculating the Sobol' indices. The existing one using Leave-One-Out Cross Validation (LOOCV) is also applicable and valid for SA but the proposed algorithm automatically decides the proper number of needed experiments to the user-defined confidence level. Therefore, this study proposes a new framework to perform more adaptive GP specialized for the purpose of SA.

1.3. Outlines

Chapter 1 is the introductions of the thesis, which present the research background, objects, scopes and outlines. Chapter 2 provides the theoretical background about global sensitivity analysis, gaussian process, and gaussian process using adaptive learning, which are the basis for the proposed method in this thesis. Chapter 3 presents the proposed GSA methodology based on the methods presented in Chapter 2. Chapter 4 presents two structural model examples with seismic excitation to validate the proposed method in Chapter 3. The first example deals with a shear building to test the proposed algorithm's validity. The second example investigates a 9-story frame building to expand the algorithm's applicability to more high-dimensional and computationally expensive system. In this study, the structural models are built by OpenSees with actual seismic data. Chapter 5 summarizes the study and provides academic and practical implications, study limitations, and suggestions for future research.

Chapter 2. Theoretical Backgrounds

2.1. Variance-based Global Sensitivity Analysis (GSA)

Assuming that $y = f(\boldsymbol{x})$ is a real integrable function for a physics model or system, where $\mathbf{X} = \{X_1, \dots, X_n\} \in \mathbb{R}^{n_x}$ is a vector of random input variables, the variance $\mathrm{Var}(Y)$ of Y can be decomposed as follows (Saltelli et al., 2008):

$$\mathrm{Var}(Y) = \Sigma_{i=1}^{n} V_i + \Sigma_{1 \le i < j}^{n} V_{ij} + \cdots + V_{12\dots n_x} \tag{2.1}$$

where $V_i = \mathrm{Var}_{X_i}\left(\mathbf{E}_{\mathbf{X}_{\sim i}}(Y|X_i)\right)$ is the variance of Y caused by X_i without considering its interactions with other input variables (i.e., $\mathbf{X}_{\sim i}$), $\mathbf{E}(\cdot)$ is the expectation operator, and $V_{1\dots k}$, $\forall k = 2, \dots, n_x$, represents the proportion of $\mathrm{Var}(Y)$ caused by variables $\{X_1, \dots, X_k\}$.

Based on the above variance decomposition, the Sobol' indices are defined as (Saltelli et al., 2008)

$$S_i = \frac{V_i}{Var(Y)}, S_{ij} = \frac{V_{ij}}{Var(Y)}, S_{1\dots k} = \frac{V_{1\dots k}}{Var(Y)}, \forall k = 2, \dots, n_x \tag{2.2}$$

where n_x is dimension of the input space, S_i is the first-order index, S_{ij} is the second-order index, and $S_{1\dots k}$ is the higher-order index corresponding to input variables $\{X_1, \dots, X_k\}$.

The number of indices will grow dramatically if the higher-order indices are used. For this reason, the first-order and total-effect Sobol' indices are commonly used and are given by

$$S_i = \frac{\mathrm{Var}_{X_i}\left(\mathbf{E}_{\mathbf{X}_{\sim i}}(Y|X_i)\right)}{\mathrm{Var}(Y)}, \forall i = i, \dots, n_x \tag{2.3}$$

$$S_{T_i} = 1 - \frac{\mathrm{Var}_{\mathbf{X}_{\sim \mathrm{i}}}\left(\mathbf{E}_{X_i}(Y|\mathbf{X}_{\sim i})\right)}{\mathrm{Var}(Y)}, \forall i = i, \dots, n_x \tag{2.4}$$

or

$$S_{T_i} = \frac{\mathbf{E}_{\mathbf{X}_{\sim \mathrm{i}}}\left(\mathrm{Var}_{X_i}(Y|\mathbf{X}_{\sim i})\right)}{\mathrm{Var}(Y)}, \forall i = i, \dots, n_x \tag{2.5}$$

where S_i and S_{T_i} are the first-order and total-effect Sobol' indices of X_i, respectively.

It should be noted that the above variance decomposition is derived based on the independence assumption of the input variables. When the input variables are correlated, $\mathrm{Var}(Y)$ cannot be decomposed as in Eq. (2.1). However, as Saltelli and Tarantola (2002) has pointed out, S_i and S_{T_i} computed using the above formulas are still informative for the importance measure of dependent input variables. In addition, sensitivity indices are defined in two types, namely full sensitivity index and independent sensitivity index, to perform GSA of model output with dependent random variables (Mara TA & Tarantola S, 2012). The full sensitivity index includes the effects of the dependence of a VoI with other inputs while the independent sensitivity indices represent the effects of a VoI that are not due to its dependence with other variables (Mara et al., 2015). Mara et al. defined that the indices given in Eqs. (2.2) and (2.3) are, respectively, the full first-order sensitivity index and the independent total-effect index when they are applied to GSA of model output with dependent random variables. In this thesis, we therefore focus on how to compute Eqs. (2.2) or (2.3) for generalized problems with or without dependent input variables.

As discussed in Chapter 1, directly solving Eqs. (2.2) or (2.3) requires a double-

loop MCS. Assume that there is a data matrix $\mathbf{X} \in \mathbb{R}^{n \times n_x}$ given as follows:

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}^{(1)} \\ \vdots \\ \mathbf{X}^{(n)} \end{bmatrix} = \begin{bmatrix} x_1^{(1)} & \cdots & x_{n_x}^{(1)} \\ \vdots & \ddots & \vdots \\ x_1^{(n)} & \cdots & x_{n_x}^{(n)} \end{bmatrix} \tag{2.6}$$

in which $\mathbf{X}^{(q)} = \left[x_1^{(q)}, \ldots, x_{n_x}^{(q)}\right], \forall q = 1, \ldots, n$ is the q^{th} sample of X and n is the number of samples in the data matrix.

Then, the computational cost of the double-loop MCS is proportion to n^{n_x}, which is expensive in case of high-dimensional and highly nonlinear systems. The computation can be reduced via existing short cuts using the instrument proposed by Saltelli (2008). Assume that two data matrices $\mathbf{A} \in \mathbb{R}^{n \times n_x}$ and $\mathbf{B} \in \mathbb{R}^{n \times n_x}$ are defined as follows:

$$\mathbf{A} = \begin{bmatrix} x_1^{(1)} & \cdots & x_{n_x}^{(1)} \\ \vdots & \ddots & \vdots \\ x_1^{(n)} & \cdots & x_{n_x}^{(n)} \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} x_{n_x+1}^{(1)} & \cdots & x_{2n_x}^{(1)} \\ \vdots & \ddots & \vdots \\ x_{n_x+1}^{(n)} & \cdots & x_{2n_x}^{(n)} \end{bmatrix} \tag{2.7}$$

and define a matrix $\mathbf{C}_i$ formed by all columns of $\mathbf{B}$ except the ith column, which is taken from $\mathbf{A}$:

$$\mathbf{C}_i = \begin{bmatrix} x_{n_x+1}^{(1)} & \cdots & x_i^{(1)} & \cdots & x_{2n_x}^{(1)} \\ & \vdots & \ddots & \vdots & \\ x_{n_x+1}^{(n)} & \cdots & x_i^{(n)} & \cdots & x_{2n_x}^{(n)} \end{bmatrix} \tag{2.8}$$

Then compute the model output for all the input values in the sample matrices $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}_i$, obtaining three vectors of model outputs:

$$y_{\mathbf{A}} = f(\mathbf{A}), \ y_{\boldsymbol{B}} = f(\mathbf{B}) \ \text{ and } \ y_{\mathbf{C}_i} = f(\mathbf{C}_i) \tag{2.9}$$

Using the above output vectors, the Sobol' indices can be obtained as follows:

$$S_i = \frac{\mathrm{Var}_{X_i}\left(\mathbf{E}_{\mathbf{X}_{\sim i}}(Y|X_i)\right)}{\mathrm{Var}(Y)} = \frac{y_A \cdot y_{C_i} - f_0^2}{y_A \cdot y_A - f_0^2} = \frac{(1/n)\Sigma_{j=1}^{n} y_A^j y_{C_i}^j - f_0^2}{(1/n)\Sigma_{j=1}^{n}\left(y_A^j\right)^2 - f_0^2} \tag{2.10}$$

$$S_{T_i} = 1 - \frac{\mathrm{Var}_{\mathbf{X}_{\sim i}}\left(\mathbf{E}_{X_i}(Y|\mathbf{X}_{\sim i})\right)}{\mathrm{Var}(Y)} = 1 - \frac{(1/n)\Sigma_{j=1}^{n} y_B^j y_{C_i}^j - f_0^2}{(1/n)\Sigma_{j=1}^{n}\left(y_A^j\right)^2 - f_0^2} \tag{2.11}$$

where $f_0^2 = \left(\frac{1}{n}\Sigma_{j=1}^{n} y_A^j\right)^2$ is the mean, and the symbol $(\cdot)$ denotes the scalar product of two vectors.

2.2. Gaussian Process (GP) formulation

Kriging combines two components: a global regression to capture underlying trends in the high-fidelity model behavior and a GP that performs a local fitting to the regression residuals. To formulate the metamodel, consider a high-fidelity simulation model with an input vector $\mathbf{x} \in \mathbb{R}^{n_x}$ and a scalar output $z \in \mathbb{R}$. The fundamental hypothesis of Kriging is that the true model response is one realization of a stochastic process of the following form:

$$\tilde{z}(\mathbf{x}) = \mathbf{f}(\mathbf{x})^T \boldsymbol{\beta} + h(\mathbf{x}) \tag{2.12}$$

where the first term is the global regression component, comprised by an $n_p \times 1$ vector of n_p dimensional basis functions $\mathbf{f}(\mathbf{x}) = [f_1(\mathbf{x}), \dots, f_{n_p}(\mathbf{x})]$ (typically $\mathbf{f}(\mathbf{x})$ is a lower order polynomial), and a $n_p \times 1$ vector of regression coefficients $\boldsymbol{\beta}$ and the second term fits the regression residuals with a GP that has a zero mean, and a covariance function of the form:

$$\mathrm{cov}\left(\mathbf{x}^l, \mathbf{x}^m\right) = \tilde{\sigma}^2 R(\mathbf{x}^l, \mathbf{x}^m | \mathbf{s}) \tag{2.13}$$

where $\tilde{\sigma}^2$ is the process variance and $R(\mathbf{x}^l, \mathbf{x}^m | \mathbf{s})$ the selected correlation function, dependent on hyper-parameters. The correlation function between experiments $\mathbf{x}^l$ and $\mathbf{x}^m$ is defined through some distance measure $d(\mathbf{x}^l, \mathbf{x}^m | \mathbf{s})$. Common choices for this correlation function include the generalized exponential, Gaussian, cubic, or Matérn correlation functions (Williams and Rasmussen 2006; Kleijnen 2009). In this thesis, the generalized exponential will be used for its flexibility, leading to:

$$d(\mathbf{x}^l, \mathbf{x}^m) = \sum_{k=1}^{n_x} s_k \left| x_k^l - x_k^m \right|^{s_{n_x}+1}; \mathbf{s} = [s_1, \dots, s_{n_x}, s_{n_x} + 1] \quad (2.14)$$

$$R(\mathbf{x}^l, \mathbf{x}^m) = \exp\left(-d(\mathbf{x}^l, \mathbf{x}^m)\right) = \Pi_{k=1}^{n_x} \exp(-s_k \left| x_k^l - x_k^m \right|^{s_{n_x}+1}) \quad (2.15)$$

Let $\{\mathbf{x}^j; j = 1, \dots, n\}$ represent the available simulation experiments and denote by $\mathbf{X} = [\mathbf{x}^1, \dots, \mathbf{x}^n]^T \in \mathbb{R}^{n \times n_x}$ the input matrix, and by $\mathbf{Z} = [z(\mathbf{x}^1), \dots, z(\mathbf{x}^n)] \in \mathbb{R}^n$, the corresponding output vector. We also denote by $\mathbf{F}(\mathbf{X}) = [\mathbf{f}(\mathbf{x}^1), \dots, \mathbf{f}(\mathbf{x}^n)]^T \in \mathbb{R}^{n \times n_p}$ the matrix of the basis functions and the correlation matrix by $\mathbf{R}(\mathbf{X}) \in \mathbb{R}^{n \times n}$ defined as $R(\mathbf{x}^l, \mathbf{x}^m | \mathbf{s})$ for $l, m = 1, \dots, n$. Also, for every new point $\mathbf{x}$, we denote by $\mathbf{r}(\mathbf{x}) = [R(\mathbf{x}, \mathbf{x}^l | \mathbf{s}), \dots, R(\mathbf{x}, \mathbf{x}^n | \mathbf{s})$ the correlation vector between the new input and each element of $\mathbf{X}$. Based on the given set of experiments, the Kriging predictive mean is given by (Sacks et al. 1989):

$$\hat{z}(\mathbf{x}|\mathbf{X}) = \mathbf{f}(\mathbf{x})^{\boldsymbol{T}} \boldsymbol{\beta}^* + \mathbf{r}(\mathbf{x}|\mathbf{X})^T \mathbf{R}(\mathbf{X})^{-1} (\mathbf{Z} - \mathbf{F}(\mathbf{X}) \boldsymbol{\beta}^*) \quad (2.16)$$

where $\boldsymbol{\beta}^* = \left(\mathbf{F}(\mathbf{X})^T \mathbf{R}(\mathbf{X})^{-1} \mathbf{F}(\mathbf{X})\right)^{-1} \mathbf{F}(\mathbf{X})^T \mathbf{R}(\mathbf{X})^{-1} \mathbf{Z}$ corresponds to the generalized least squares estimate of $\boldsymbol{\beta}$. Kriging also provides an estimate of the predictive mean's variability, ultimately an estimate of the metamodel error (Jin et al. 2002), quantified through the predictive variance:

$$\sigma^2(\mathbf{x}|\mathbf{X}) = \tilde{\sigma}^2(\mathbf{X}) \sigma_n^2(\mathbf{x}|\mathbf{X})$$

$$\text{with } \sigma_n^2(\mathbf{x}|\mathbf{X}) = [1 - \mathbf{r}(\mathbf{x}|\mathbf{X})^T \mathbf{R}(\mathbf{X})^{-1} \mathbf{r}(\mathbf{x}|\mathbf{X})$$

$$+ \mathbf{u}(\mathbf{x}|\mathbf{X})^T \{\mathbf{F}(\mathbf{X})^T \mathbf{R}(\mathbf{X})^{-1} \mathbf{F}(\mathbf{X})\}^{-1} \mathbf{u}(\mathbf{x}|\mathbf{X})] \quad (2.17)$$

where $\mathbf{u}(\mathbf{x}|\mathbf{X}) = \mathbf{F}(\mathbf{X})^T\mathbf{R}(\mathbf{X})^{-1}\mathbf{r}(\mathbf{x}|\mathbf{X}) - \mathbf{f}(\mathbf{x})$.

This normalized variance describes the variation of the metamodel prediction error in the $\mathbf{x}$ domain and is independent of observations $\mathbf{Z}$, while the process variance corresponds to an independent of $\mathbf{x}$ scaling constant, and is a function of the observations $\mathbf{Z}$. The maximum likelihood estimate of the process variance is:

$$\tilde{\sigma}^2 = \frac{(\mathbf{Z}-\mathbf{F}(\mathbf{X})\boldsymbol{\beta}^*)^T(\mathbf{Z}-\mathbf{F}(\mathbf{X})\boldsymbol{\beta}^*)}{n} \tag{2.18}$$

Through proper tuning of the hyperparameter vector $\mathbf{s}$, Kriging has been proven efficient in approximating even highly complex functions (Simpson et al. 2001b). In this work, this tuning is performed using maximum likelihood estimation (Lophaven et al. 2002). The predictive capability of Kriging can be evaluated using LOOCV, where each experiment $\mathbf{x}^i$ is sequentially removed from the sample set, and the remaining ones are utilized to provide prediction $\hat{z}(\mathbf{x}^i|\mathbf{X}_{-i})$, with $\mathbf{X}_{-i}$ denoting the original dataset excluding $\mathbf{x}^i$. Closed-form solutions exists for the LOOCV (Dubrule 1983; Sundararajan and Keerthi 2001) statistics with no need to explicitly evaluate the metamodels that correspond to observations $\hat{z}(\mathbf{x}^i|\mathbf{X}_{-i})$. The predictive mean and variance are given, respectively, by:

$$\hat{z}(\mathbf{x}^i|\mathbf{X}_{-i}) = z(\mathbf{x}^i) - \frac{[\mathbf{R}(\mathbf{X})^{-1}(\mathbf{Z}-\mathbf{F}(\mathbf{X})\boldsymbol{\beta}^*)]_i}{[\mathbf{R}(\mathbf{X})^{-1}]_{ii}} \tag{2.19}$$

$$\sigma^2(\mathbf{x}^i|\mathbf{X}_{-i}) = \frac{1}{[\mathbf{B}]_{ii}} \tag{2.20}$$

where $[\cdot]_{pq}$ is used to denote the entry on the pth row and qth column in a matrix, $[\cdot]_p$ is the pth element of a vector, and matrix $\mathbf{B}$ is given by:

$$\mathbf{B} = \begin{bmatrix} \tilde{\sigma}^2 \mathbf{R}(\mathbf{X}) & \mathbf{F}(\mathbf{X}) \\ \mathbf{F}(\mathbf{X})^T & \mathbf{0} \end{bmatrix}^{-1} \tag{2.21}$$

The estimate for the LOOCV error, characterizing ultimately the metamodel bias, is:

$$e_i = \frac{\left[\mathbf{R}(\mathbf{X})^{-1}(\mathbf{Z}-\mathbf{F}(\mathbf{X})\boldsymbol{\beta}^*)\right]_i}{\left[\mathbf{R}(\mathbf{X})^{-1}\right]_{ii}} \tag{2.22}$$

2.3. Adaptive DoE Incorporating Bias Information

Let X^d denote the domain of interest within which the metamodel will be eventually used to provide predictions. The general objective here is to establish a metamodel that is globally accurate within X^d. Kriging's ability to provide a local estimate of the prediction error variability according to Eq. (2.17), is the basis for many adaptive DoE approaches to satisfy this objective. Given the observation set $\mathbf{X}$ these approaches search for the next experiment $\mathbf{x}_{\text{new}}$ that will provide the greatest anticipated improvement according to some score functions. The most popular among such score functions is the integrated mean squared error given by (Sacks et al. 1989):

$$IMSE(\mathbf{X}, \mathbf{x}_{\text{new}}) = \int_{X^d} \sigma_n^2(\mathbf{x}|\mathbf{X}, \mathbf{x}_{\text{new}})\, d\mathbf{x} \tag{2.20}$$

where $\sigma_n^2(\mathbf{x}|\mathbf{X}, \mathbf{x}_{\text{new}})$ is the normalized predictive variance considering the new experiment $\mathbf{x}_{\text{new}}$, given by:

$$\begin{aligned}&\sigma_n^2(\mathbf{x}|\mathbf{X}, \mathbf{x}_{\text{new}})\\ &= 1 + \mathbf{u}(\mathbf{x}|\mathbf{X}, \mathbf{x}_{\text{new}})^T\{\mathbf{F}(\mathbf{X}, \mathbf{x}_{\text{new}})^T\mathbf{R}(\mathbf{X}, \mathbf{x}_{\text{new}})^{-1}\mathbf{F}(\mathbf{X}, \mathbf{x}_{\text{new}})\}^{-1}\mathbf{u}(\mathbf{x}|\mathbf{X}, \mathbf{x}_{\text{new}})\\ &\quad -\mathbf{r}(\mathbf{x}|\mathbf{X}, \mathbf{x}_{\text{new}})^T\mathbf{R}(\mathbf{X}, \mathbf{x}_{\text{new}})^{-1}\mathbf{r}(\mathbf{x}|\mathbf{X}, \mathbf{x}_{\text{new}})\end{aligned} \tag{2.21}$$

Note that this normalized variance ignores the process variance $\tilde{\sigma}^2$, since the latter has no impact on the optimization, acting merely as a scaling constant (A. P. Kyprioti et al., 2020). The term $IMSE(\mathrm{X}, \mathrm{x}_{\text{new}})$ corresponds to the average error of the metamodel established by using the dataset $\mathbf{X}$ and the new point $\mathbf{x}_{\text{new}}$ as support

points, assuming no modification in the hyperparameters after the addition of $\mathbf{x}_{\text{new}}$. The above equation requires the evaluation of the augmented correlation matrix $\mathbf{R}(\mathbf{X}, \mathbf{x}_{\text{new}})$, the new basis function vector $\mathbf{F}(\mathbf{X}, \mathbf{x}_{\text{new}})$, and the correlation vector $\mathbf{r}(\mathbf{x}|\mathbf{X}, \mathbf{x}_{\text{new}})$, all considering augmenting the existing experiments $\mathbf{X}$ by the new one $\mathbf{x}_{\text{new}}$. For adding a single new experiment $\mathbf{x}_{\text{new}}$, the potentially computational demanding inversion of the updated correlation matrix, $\mathrm{R}(\mathrm{X}, \mathrm{x}_{\text{new}})$ required for Eq. (2.21), can be simplified using the 1-rank update:

$$\mathbf{R}(\mathbf{X}, \mathbf{x}_{\text{new}})^{-1} = \begin{bmatrix} \mathbf{R}(\mathbf{X})^{-1} + \frac{\mathbf{R}(\mathbf{X})^{-1}\mathbf{r}(\mathbf{x}_{\text{new}}|\mathbf{X})\mathbf{r}(\mathbf{x}_{\text{new}}|\mathbf{X})^T\mathbf{R}(\mathbf{X})^{-1}}{\eta_{new}} & -\frac{\mathbf{R}(\mathbf{X})^{-1}\mathbf{r}(\mathbf{x}_{\text{new}}|\mathbf{X})}{\eta_{new}} \\ -\frac{\mathbf{r}(\mathbf{x}_{\text{new}}|\mathbf{X})^T\mathbf{R}(\mathbf{X})^{-1}}{\eta_{new}} & \frac{1}{\eta_{new}} \end{bmatrix} \tag{2.23}$$

where $\eta_{new} = 1 - \mathbf{r}(\mathbf{x}_{\text{new}}|\mathbf{X})^T\mathbf{R}(\mathbf{X})^{-1}\mathbf{r}(\mathbf{x}_{\text{new}}|\mathbf{X})$. This expression does not entail any new matrix inversions beyond scalar ones, since the partitioned inversion $\mathbf{R}(\mathbf{X})^{-1}$ is known after developing the metamodel for the observation set $\mathbf{X}$.

The optimal experiment(s) can be then selected to minimize the $IMSE$ over the domain of interest:

$$\mathbf{x}_{\text{new}}^{*} = \arg \min_{\mathbf{x}_{\text{new}} \in X^d} IMSE(\mathbf{X}, \mathbf{x}_{\text{new}}) \tag{2.24}$$

Note that this corresponds to a challenging optimization problem, with multiple local minima (Picheny et al. 2010) and a score function that requires a cumbersome integration over X^d with respect to the updated variance for each $\mathbf{x}_{\text{new}}$ examined. An efficient optimization scheme will be discussed in the next section. An alternative score function can be formulated by considering the maximum error over domain X^d instead of the integrated one (Sacks et al. 1989):

$$MMSE_{\mathrm{u}}(\mathbf{X}, \mathbf{x}_{\mathrm{new}}) = \max_{\mathbf{x} \in X^d} \sigma_n^2(\mathbf{x}|\mathbf{X}, \mathbf{x}_{\mathrm{new}}) \tag{2.25}$$

Rather than using the updated predictive variance considering the new experiment(s), given by Eq. (2.21), it is common in such instances (McKay et al. 1979; Jin et al. 2002) to use the variance for the current metamodel as it leads to a computationally simpler DoE procedure. The corresponding maximum mean squared error approach selects the new experiment based on the largest current estimation of the prediction error, leading to a score function that considers strictly the existing training set:

$$MMSE(\mathbf{x}) = \sigma_n^2(\mathbf{x}|\mathbf{X}) \tag{2.26}$$

and to an optimal new experiment identification:

$$\mathbf{x}_{\mathrm{new}}^* = \arg\max_{\mathbf{x} \in X^d} MMSE(\mathbf{x}) \tag{2.27}$$

In this paper, MMSE has been used as the objective function since it is efficient approximation of IMSE, i.e., more cheap and similar accuracy comparing to IMSE.

The previous discussion focuses on adding one new experiment at a time. When multiple simulations for experiments can be simultaneously performed, it is often preferred to add a batch of n_q new experiments at the same time according to either $IMSE$ or $MMSE$ score functions ($MMSE$ in this paper). This can be done by sequentially identifying new experiments one-at-a-time as described above, adding each of them to $\mathbf{X}$ and updating $\mathbf{R}, \mathbf{F}$ and $\mathbf{r}$, the essential components of the normalized variance, without updating the Kriging hyperparameter structure as the latter would require the output for the new experiments, and then proceeding to the

identification of the next experiment until the desired batch size has been reached.

The score function variants discussed in the above focus entirely on the metamodel prediction variability, quantified through the normalized predictive variance, and thus, ignore any possible bias in these predictions. To incorporate information about the latter, an adjustment of these score functions is proposed here utilizing a weighting term quantified through the LOOCV error. This leads to the modification of the $MMSE$ presented in Eq. (2.26) to a weighted $MMSE$, $MMSE_w$ defined as follows:

$$MMSE_w(\mathbf{x}) = \varphi(\mathbf{x})^\rho \sigma_n^2(\mathbf{x}|\mathbf{X}) \tag{2.28}$$

where weights $\varphi(\mathbf{x})$ use information regarding both the LOOCV mean and variance as:

$$\varphi(\mathbf{x}) = \frac{\Sigma_{i=1}^n w_i(\mathbf{x})[1+e_i^2/\sigma^2(\mathbf{x}^i|\mathbf{X}_{-i})]}{\Sigma_{i=1}^n w_i(\mathbf{x})} \tag{2.29}$$

with $w_i(\mathbf{x})$ obtained using Voronoi cells as explained later in Eq. (2.31)

Contrary to the predictive variance, the LOOCV error is known only for the input locations of the initial training set. In order to approximate the LOOCV error across the whole domain X^d, the weighted average interpolation over the set X is adopted, leading to the following $\varphi(\mathbf{x})$ definition:

$$\varphi(\mathbf{x}) = \frac{\Sigma_{i=1}^n w_i(\mathbf{x}) e_i^2}{\Sigma_{i=1}^n w_i(\mathbf{x})} \tag{2.30}$$

For the weights, $w_i(\mathbf{x})$ sets to 1 only for the closest experiment to $\mathbf{x}$ and 0 otherwise, corresponding to the natural neighbor (NN) interpolation:

$$w_i(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in V_i = \left\{\mathbf{x} \in X^d, \left\|\mathbf{x} - \mathbf{x}^i\right\| \leq \left\|\mathbf{x} - \mathbf{x}^j\right\|, \forall j \neq i\right\}, i, j = 1, \ldots, n \\ 0 & \text{else} \end{cases} \quad (2.30)$$

where V_i is the Voronoi cell associated with $\mathbf{x}^i$ for set $\mathbf{X}$, and the difference between vectors $\|\cdot\|$ is chosen as the distance used for the Kriging correlation function $\left\|\mathbf{x} - \mathbf{x}^i\right\| = d(\mathbf{x}, \mathbf{x}^i|\mathbf{s})$. Note that this choice for $w_i(\mathbf{x})$, using Voronoi cells, has been adopted in similar studies that relied on interpolation for LOOCV error (Le Gratiet and Cannamela 2015; Liu et al. 2016). Thus, the two main components in Eq. (2.28), bias and variance, show two fundamental strategies for the identification of the new experiments

1. The LOOCV error weight contains information about the metamodel bias. A larger weight in certain regions suggests a smaller local metamodel accuracy, which is due to the inadequate capturing of the local nonlinearities of the original function, as will be shown later in the thesis. By introducing this weight, a local exploitation is encouraged in such problematic domains.

2. Metamodel variance $\sigma_n^2(\mathbf{x}|\mathbf{X}, \mathbf{x}_{\text{new}})$, ultimately quantifies the closeness between the input $\mathbf{x}$ to all current experiments $\mathbf{X}$, and to the new experiment $\mathbf{x}_{\text{new}}$ as well. A large variance suggests that the input region is "under-sampled" considering the distribution of existing experiments $\mathbf{X}$ and the added one $\mathbf{x}_{\text{new}}$ in X^d. By aiming to reduce such variance within the DoE, a global exploration of the input domain is promoted.

Both components should be used within the DoE process. By incorporating information about the metamodel bias, the LOOCV weight $\varphi(\mathbf{x})$ can be particularly impactful for capturing local nonlinearities of the original function. On

the other hand, since it is fully formulated based on LOOCV information, it inherits the well-known LOOCV limitation, i.e., ignoring the existence of a specific experiment at each localized domain (Kleijnen and Beers, 2004). Therefore, the proposed LOOCV weight can only offer incomplete details about the true bias given $\mathbf{X}$. Its drawback is partially compensated by using the updated predictive variance $\sigma_n^2(\mathbf{x}|\mathbf{X}, \mathbf{x}_{\text{new}})$. As discussed, this quantity provides information about the influence of adding the new experiment $\mathbf{x}_{\text{new}}$ (through the variance updating), but more importantly, incorporates the distribution of all existing experiments $\mathbf{X}$ in X^d into the DoE, and ultimately avoids putting new experiments within close proximity to existing ones.

In summary, the identification of the new experiment(s) using the proposed adaptive framework is performed through Eq. (2.27) simply using the weighted $MMSE_w$ of Eq. (2.28) as the objective function.

Algorithm 1. Adaptive DoE with MMSE

Step 1 Candidate experiments: generate n_c candidate experiments $\{\mathbf{x}_{\text{new}}^c; c = 1, \dots, n_c\}$ following a uniform distribution in X^d.

Step 2 Ranking and prescreening of experiments: evaluate $\{\sigma_n^2(\mathbf{x}_{\text{new}}^c|\mathbf{X}), \varphi(\mathbf{x}_{\text{new}}^c); c = 1, \dots, n_c\}$ and retain only the $a_r n_c$ candidate experiments that correspond to the highest values of $\varphi(\mathbf{x}_{\text{new}}^c)^\rho \sigma_n^2(\mathbf{x}_{\text{new}}^c|\mathbf{X})$, with a_r being the desired percentage of candidate experiments that have larger weighted-variance values.

Step 3 Final selection: select as new experiment the one that provides the

minimum value for $\{MMSE(\mathbf{x}_{\text{new}}^{c}); c = 1, \dots, a_r a_c n_c\}$.

The prescreening in step 2 offers a substantial reduction of the computational burden, as it avoids the estimation of the $MMSE$ for candidate experiments that are not expected to correspond to the optimum; proportion $(1 - a_r)$ of candidate experiments in subdomains of X^d with low current prediction variability are ultimately ignored. Note that the current variance $\sigma_n^2(\mathbf{x}|\mathbf{X})$ used in the prescreening step can be estimated with small computational cost, so the overall computational burden for this step is negligible.

The remaining characteristics of the numerical optimization scheme that need to be discussed are the selection of N_s and n_c. The choice of N_s impacts the estimation error. It is still important to set n_c large enough to promote an adequate exploration of X^d at each iteration, and the inclusion of the step 2 in the optimization algorithm has been introduced to accommodate such a selection.

Chapter 3. Proposed Algorithm of GSA using Adaptively-learned Gaussian Process

3.1. GSA using Adaptive GP

As explained in Chapter 1, the motivation of the proposed method is to develop a computationally efficient algorithm for SA. Here, computational efficiency has two meanings – (1) taking less computational time for an evaluation of the model and (2) reducing a number of the needed experiments by adaptively selecting the location of them. To achieve these goals, we introduce a surrogate modeling method by GP to replace the system model with a cheaper model and an adaptive GP method by considering both predictive variance of GP and the LOOCV error, which is explained in algorithm 1 of Chapter 2. As shown in Figure 3.1, GSA using adaptive GP can be performed with the proposed algorithm explained as follows:

Algorithm 2. GSA using Adaptive GP

Step 1 Initial experiments and corresponding outputs: generate n_{init} initial experiments $\{\mathbf{x}_{\text{init}}^{i} \in \mathbb{R}_{x}^{n}; i = 1, \dots, n_{init}\}$ following a uniform distribution in X^{d} and run simulation to obtain n_{init} initial outputs $\{y_{\text{init}}^{i} \in \mathbb{R}; i = 1, \dots, n_{init}\}$.

Step 2 Calibrate GP's parameters: evaluate initial parameters of GP using the initial data and calibrate the GP parameters.

Step 3 Train GP using adaptive learning: repeat **Algorithm 1** n_{train} times which is user-defined parameter.

Step 4 Perform SA: obtain Sobol' indices using the GP model built on **Step 3.**

The most computationally demanding part of the algorithm above is calibrating the GP model with a new experiment, which requires updating of the GP's parameters. However, it shows very similar degree of accuracy with much smaller number of experiments compared to the original GP. Therefore, it is able to obtain the Sobol' indices accurately and more efficiently by using the proposed algorithm.

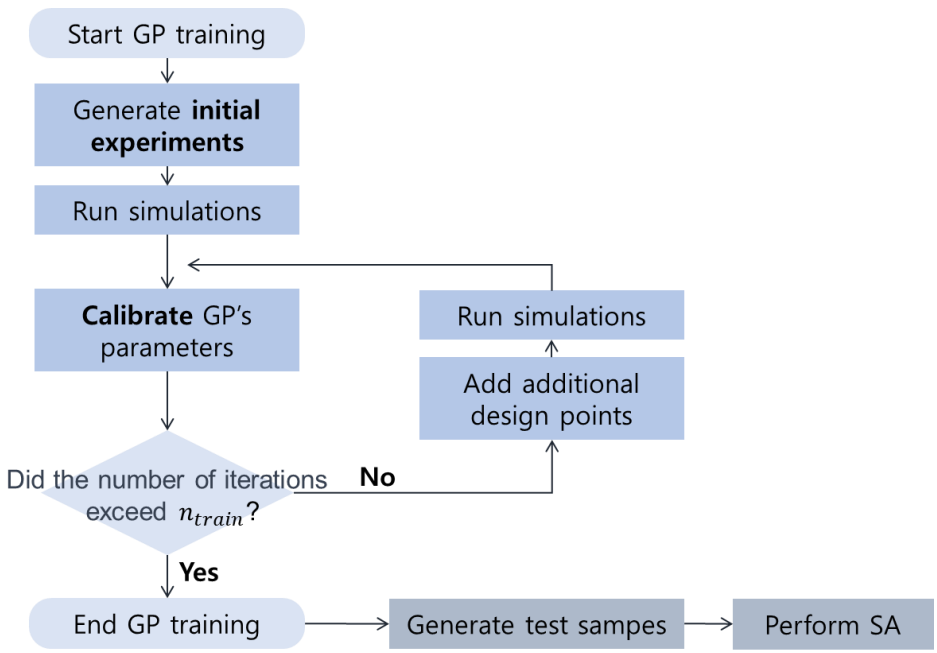

Figure 3.1 Flow chart of GSA using adaptive GP

3.2. GSA using Modified Adaptive GP

The basic idea of adaptive learning is to find an optimal point by minimizing or maximizing its objective function. $MMSE$ minimize the maximum error, which is defined as a multiplication of bias and variance while $IMSE$ minimize the average error expressed as the integration of the error over the whole input space as shown in Eq. (2.28).

However, as mentioned earlier in Section 1.2, the objective function of the adaptive learning proposed in Section 2.3 is not specialized for performing SA. To solve this problem, this study proposes the modified adaptive gaussian process, which has more appropriate objective function to obtain the Sobol' indices and its own convergence criteria in order to help users make decisions by suggesting a proper number of experiments that are needed to guarantee an enough level of confidence.

In the original methods, the algorithm converges at the user-defined number of training experiments, n_{train}, since it does not have its own convergence criteria. Thus, there are some over-trained cases, which means that the training process of GP goes far even if it already converged enough confidence level. In contrast, the modified method has its own convergence criteria which incorporates the confidence information of the Sobol' indices. The overall flow of it is as follows (Figure 3.2):

Algorithm 3. GSA using Modified Adaptive GP

Step 1 Initial experiments and corresponding outputs: generate n_{init} initial experiments $\{\mathbf{x}_{\text{init}}^{i} \in \mathbb{R}_{x}^{n}; i = 1, \dots, n_{init}\}$ following a uniform distribution in X^d

and run simulation to obtain n_{init} initial outputs $\{y_{\text{init}}^{i} \in \mathbb{R}; i = 1, \dots, n_{init}\}$.

Step 2 Calibrate GP's parameters: evaluate initial parameters of GP using the initial data and calibrate the GP parameters.

Step 3 Train GP using modified adaptive learning: the algorithm repeats **Algorithm 1.** and stops GP training automatically depend on its own convergence criteria defined by the confidence level, which is

Step 4 Perform SA: obtain Sobol' indices using the GP model built on **Step 3.**

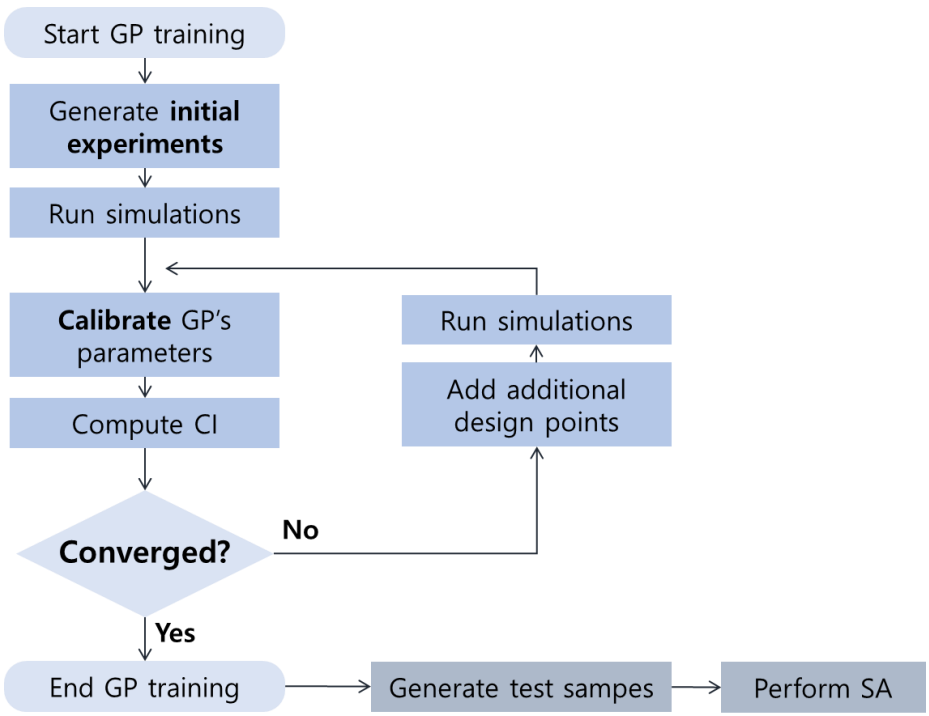

Figure 3.2 Flow chart of GSA using modified adaptive GP

Chapter 4. Verification of Proposed Method

In this chapter, two numerical examples are presented to confirm the validity of the proposed algorithm in Chapter 3. The target structures are including a 7-story shear building and a 9-story frame structure. Both structures are excitated by the earthquake using the actual record data of ground acceleration, El Centro, as shown in Figure 4.3.

Structural modeling and analysis are performed in OpenSees. Numerical simulations were performed in MATLAB 2021b. Calibration of the GP parameters was performed utilizing DACE toolbox developed by Lophaven et al (2002).

4.1. Example 1: 7-story Shear Building

The target structure is a 7-story of shear building as shown in Figure 4.1. As the material property, uniaxial steel object with isotropic strain hardening is used (Filippou, 1983) as shown in Figure 4.2.

Three input variables are chosen as Table 4.1; E_0 is young's modulus, F_y is yielding stress, and α is post-ratio. They are the interest variables and the Sobol' indices are obtained for each input variables. GSA was performed by modified adaptive GP. It is tested by changing the number of initial experiments and the proposed algorithm automatically decided the number of new experiments which are adaptively found by the algorithm. The results of the 7-story shear building using the proposed algorithm for each case; GSA using modified adaptive GP are shown in

Figure 4.4, Figure 4.5 and Figure 4.6. Comparing the results from Figure 4.4, 4.5 and 4.6, one can find that there is not significant difference between them, which means that the proposed algorithm is even valid for smaller number of initial points. The convergence graph comparing the proposed algorithm and the Quasi-MCS explained in Eq. (2.10) is shown in Figure 4.7 and the proposed one shows similar accuracy compared with the Quasi-MCS.

Figure 4.1 Configuration of the 7-story shear building

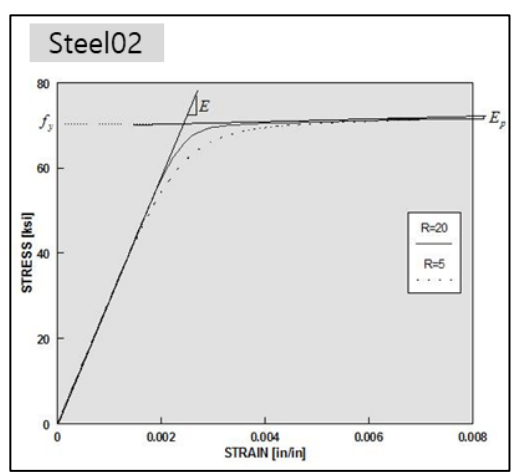

Figure 4.2 Material used for the 7-story shear building

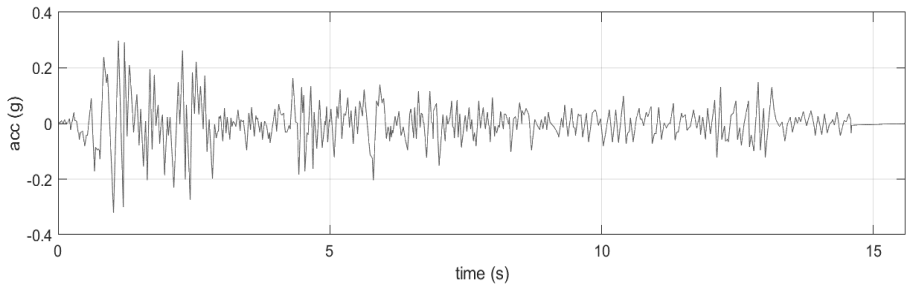

Figure 4.3 Ground acceleration time history of El Centro earthquake

(ksi unit)	$\boldsymbol{\mu}$	$\boldsymbol{\delta}$
$\boldsymbol{E_0}$	326	0.1
$\boldsymbol{F_y}$	50	0.1
$\boldsymbol{\alpha}$	0.2	0.1

Table 4.1 Input variables of the 7-story shear building

	Case 1	Case 2	Case 3
#. of Initial points	50	25	10
#. of new DoE	93	102	118
Time(sec)	131.53	145.81	146.53

Table 4.2 The number of experiments and computation cost

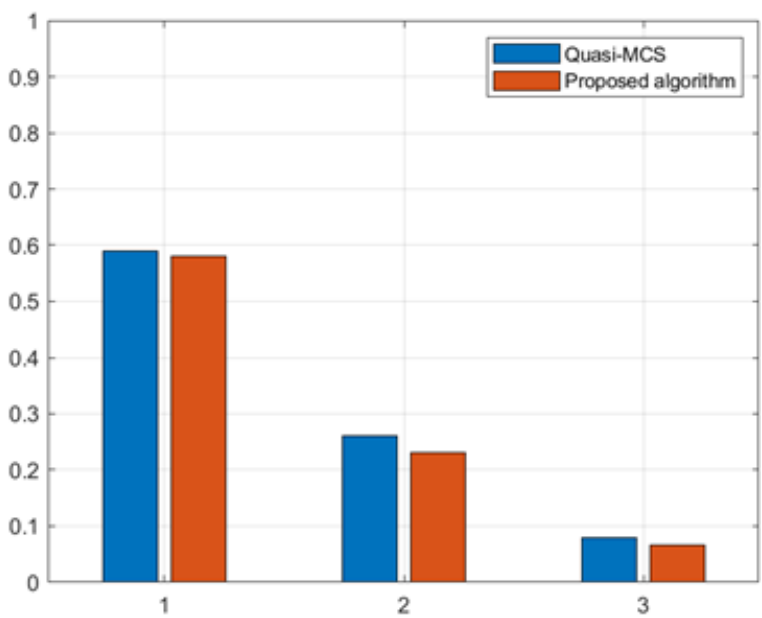

Figure 4.4 First order Sobol' index of example 1, case 1

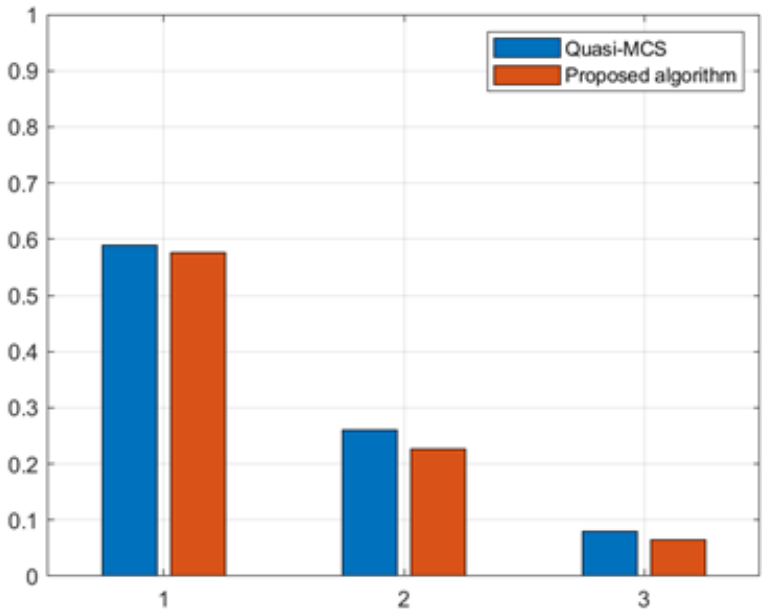

Figure 4.5 First order Sobol' index of example 1, case 2

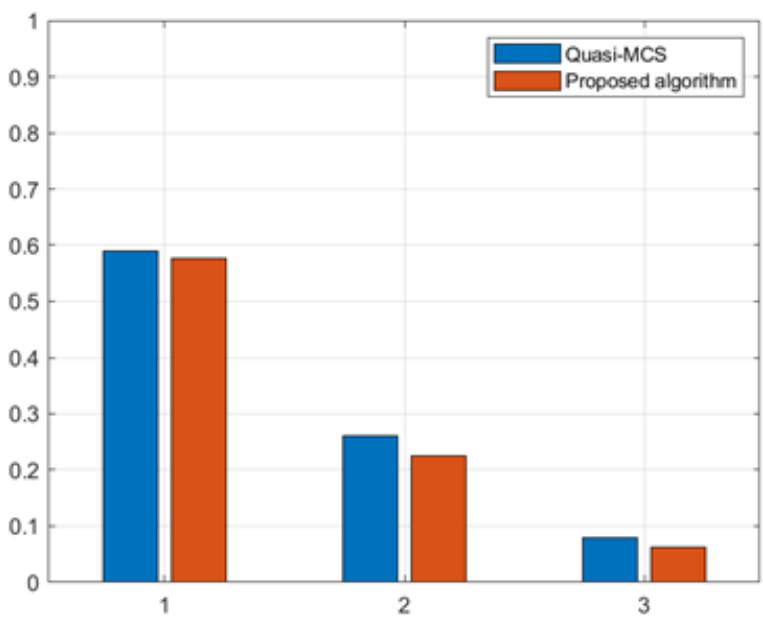

Figure 4.6 First order Sobol' index of example 1, case 3

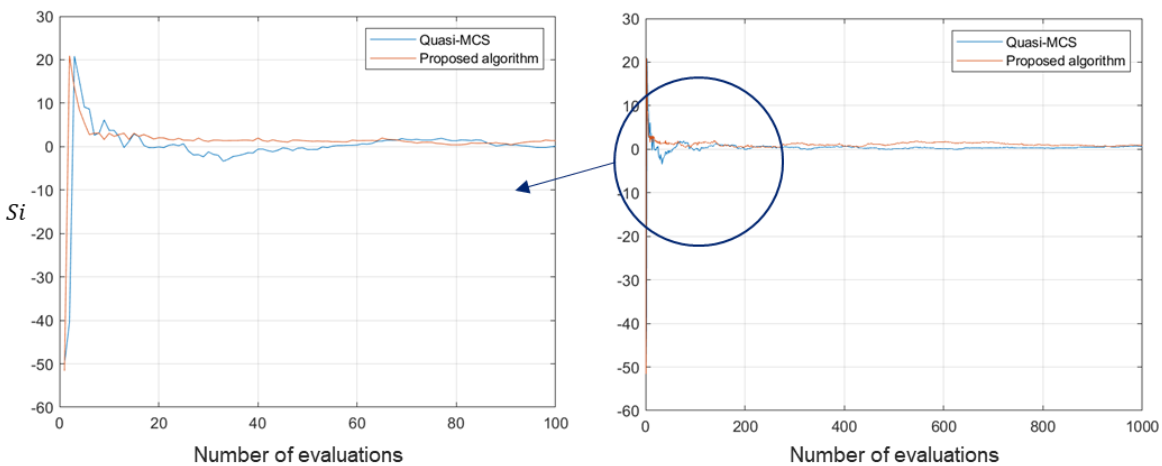

Figure 4.7 Convergence graph along the number of evaluations

4.2. Example 2: 9-story Frame

As shown in Figure 4.8, the second target structure is 9-story frame building. It is excitated by the same earthquake with Example 1, i.e., El Centro. The Steel01 material of OpenSees in Figure 4.9, which is a uniaxial bilinear steel material object with kinematic hardening and isotropic hardening described by a nonlinear evolution equation, is used for the structure.

To test the validity of the proposed algorithm in the higher dimension than the example 1, six input parameters are chosen as shown in Table 4.2. The OpenSees model is tested its validity with pushover and quasi-cyclic test and the results are shown in Figure 4.10 and 4.11. Four input variables are chosen as Table 4.3; E_0 is young's modulus, F_y is yielding stress, α is post-ratio, and PGA is factor for peak ground acceleration. They are the interest variables and the Sobol' indices are obtained for story 1 and story 9 and the results of the 9-story frame using the proposed algorithm are shown in Figure 4.12. One can find that the impact of PGA becomes larger in story 9. As shown in Figure 4.14, the proposed algorithm's advantage is that one can obtain both the confidence level and prediction about the model.

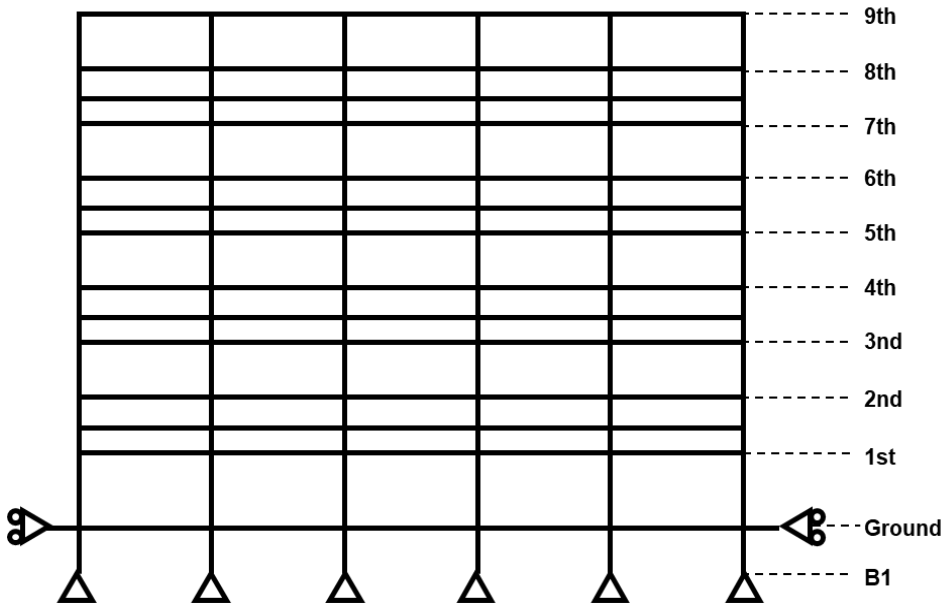

Figure 4.8 Configuration of the 9-story frame

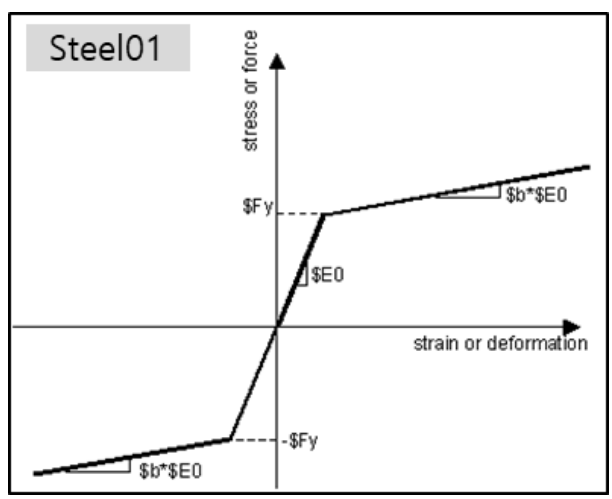

Figure 4.9 Material used for the 9-story frame

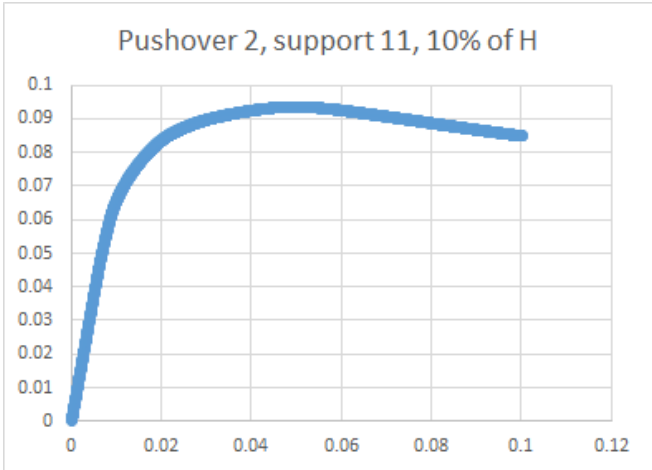

Figure 4.10 Pushover curve of the 9-story frame

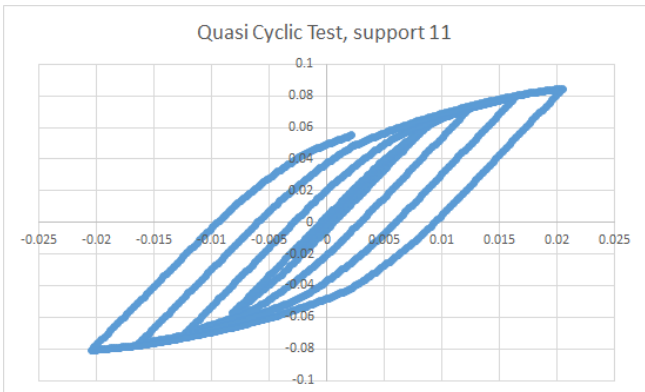

Figure 4.11 Quasi-Cyclic test curve of the 9-story frame

(SI Unit)	μ	δ
$E_{0,column}$	200,000	0.1
$F_{y,\,column}$	345	0.1
α_{column}	0.01	0.1
PGA Weight factor	0.01	0.1

Table 4.3 The input variables of 9-story frame

	Case 1	Case 2	Case 3
#. of Initial points	50	25	10
#. of new DoE	213	242	258
Time(sec)	259.53	260.81	263.53

Table 4.4 The number of experiments and computation cost of 9-story frame

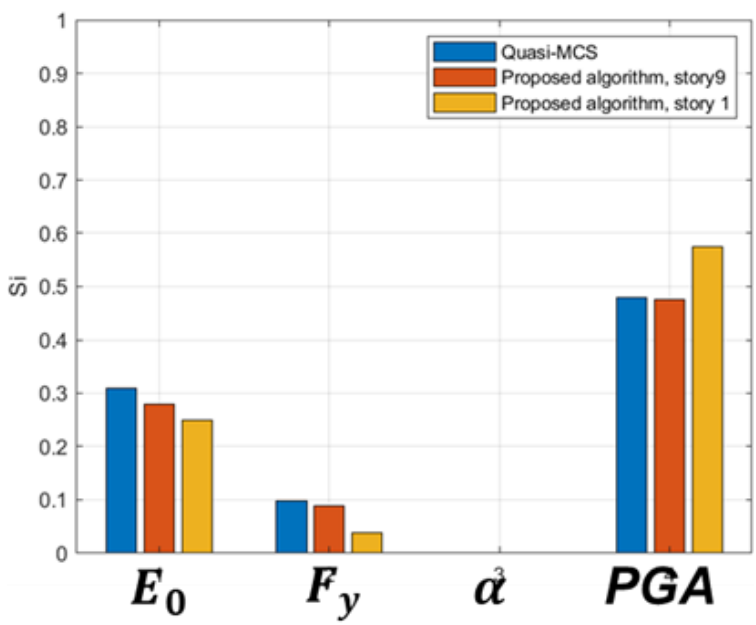

Figure 4.12 First order Sobol' index of example 2

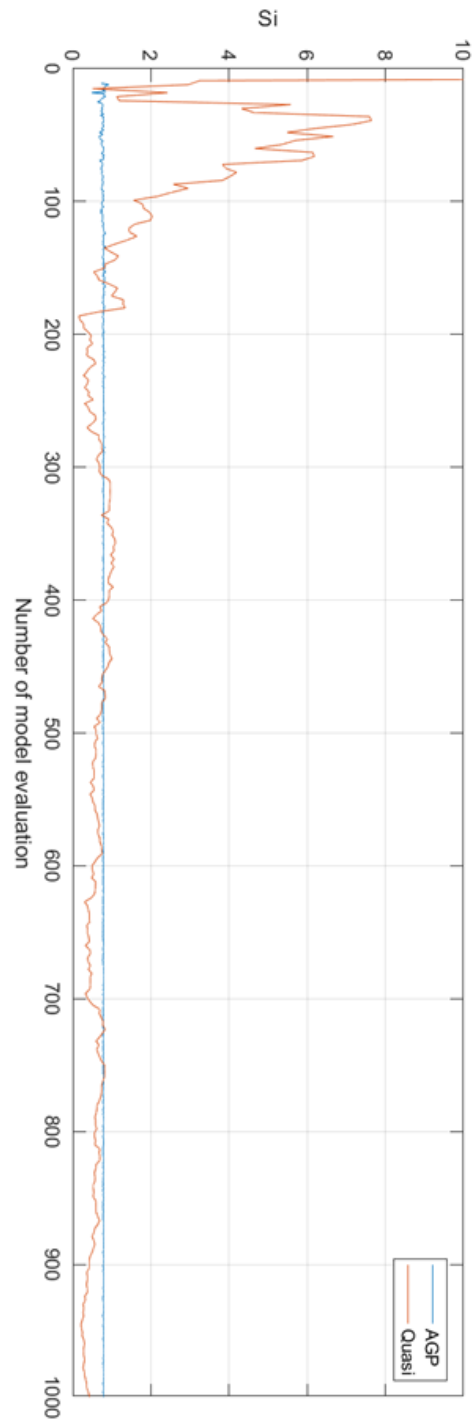

Figure 4.13 Convergence graph along the number of evaluations, example 2

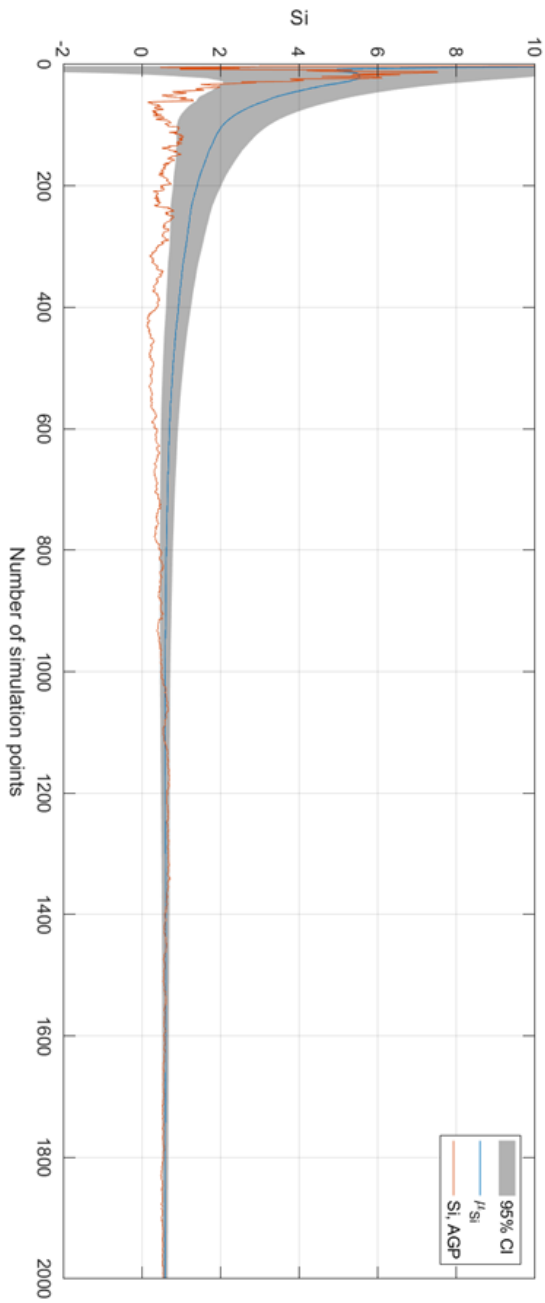

Figure 4.14 Convergence graph with confidence level, Example 2

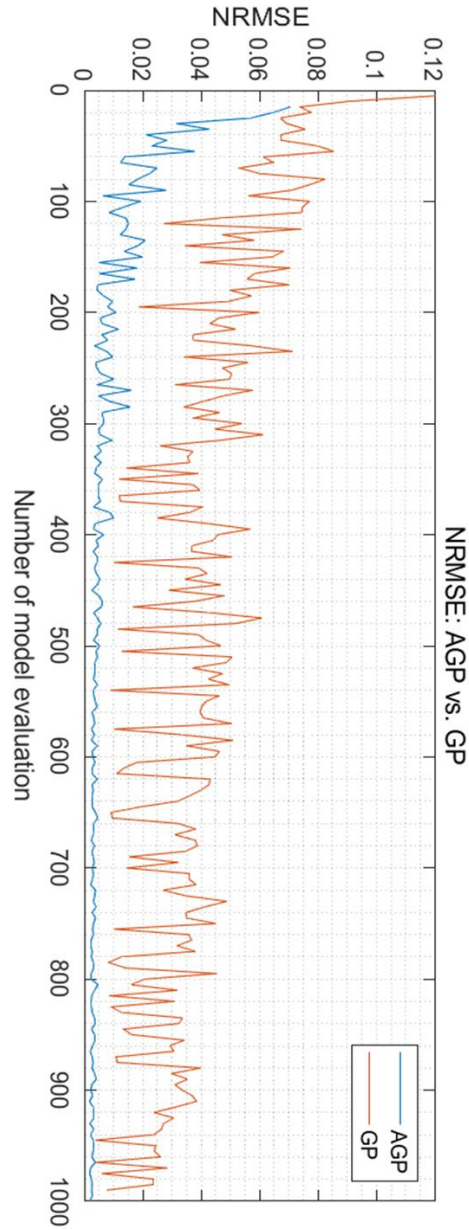

Figure 4.15 NRMSE, example 2

Chapter 5. Conclusions

In this thesis, GSA algorithm using adaptive GP that can be effectively used for complex systems is introduced. It is confirmed that the proposed algorithm enables the accurate estimation of the Sobol' indices. Furthermore, the algorithm can identify the variability of the Sobol' indices in order to help user's decision making and it has its own convergence criteria to stop the algorithm automatically.

There are two major further studies based on this study. First, expanded algorithm for total-effect Sobol' indices can be developed. It would give a comprehensive understanding of the given system if the total-effect Sobol' indices could be found even in the complex systems.

Second is to develop more adaptivity that even can identify the number of initial experiments. Even though the proposed one decides the number of new experiments, it could be more helpful for the user if the algorithm can determine the number of initial experiments.

Further studies based on results of this study are expected to enhance the effective GSA for the more complex systems.

References

Dubrule, O. (1983). Cross validation of kriging in a unique neighborhood. J Int Assoc Math Geol 15(6):687–699

Filippou, F. C., Popov, E. P., Bertero, V. V. (1983). "Effects of Bond Deterioration on Hysteretic Behavior of Reinforced Concrete Joints". Report EERC 83-19, Earthquake Engineering Research Center, University of California, Berkeley.

Glen G, Isaacs K. (2012) Estimating Sobol sensitivity indices using correlations. Environ Model Softw;37:157–66

Guillaume, J.H.A., Jakeman, J.D., Marsili-Libelli, S., Asher, M., Brunner, P., Croke, B., Hill, M.C., Jakeman, A.J., Keesman, K.J., Razavi, S., Stigter, J.D., (2019). Introductory overview of identifiability analysis: a guide to evaluating whether you have the right type of data for your modeling purpose. Environ. Model. Software 119, 418–432.

Jin R., Chen W., Sudjianto A. (2002) On sequential sampling for global metamodeling in engineering design. In: ASME 2002 international design engineering technical conferences and computers and information in engineering conference. American Society of Mechanical Engineers, pp 539-548

Kleijnen JP (2009) Kriging metamodeling in simulation: a review. Eur J Oper Res 192(3):707–716

Lophaven SN, Nielsen HB, Søndergaard J. (2002). DACE-A Matlab Kriging toolbox, version 2.0

Mara TA, Tarantola S. (2012). Variance-based sensitivity indices for models with dependent inputs. Reliability Engineering System Safety;107:115–21.

Mara TA, Tarantola S, Annoni P. (2015) Non-parametric methods for global sensitivity analysis of model output with dependent inputs. Environ Model Software;72:173–83.

Sacks J, Welch WJ, Mitchell TJ, Wynn HP (1989) Design and analysis of computer experiments. Stat Sci 4(4):409–435

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. and Tarantola, S. (2008). Global Sensitivity Analysis, John Wiley & Sons, Ltd.

Saltelli A, Tarantola S. (2002). On the relative importance of input factors in mathematical models: safety assessment for nuclear waste disposal. J Am Stat Assoc; 97(459):702–9.

Simpson TW, Poplinski J, Koch PN, Allen JK (2001) Metamodels for computer-based engineering design: survey and recommendations. Eng Comput 17(2):129–150

Sobol' IM. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Math Comput Simul, 55, 271-80.

Sobol', I.M., Tarantola, S., Gatelli, D., Kucherenko, S.S., Mauntz, W., (2007). Estimating the approximation error when fixing unessential factors in global sensitivity analysis. Reliab. Eng. Syst. Saf. 92, 957–960.

Sundararajan S, Keerthi SS (2001) Predictive approaches for choosing hyperparameters in Gaussian processes. Neural Comput 13(5): 1103–1118

Tarantola, S., Giglioli, N., Jesinghaus, J., Saltelli, A., (2002). Can global sensitivity analysis steer the implementation of models for environmental assessments and decisionmaking? Stoch. Environ. Res. Risk Assess. 16, 63–76.

Williams CK, Rasmussen CE (2006) Gaussian processes for machine learning, vol 2. vol 3. MIT Press Cambridge, MA

국문 초록

민감도 해석은 시스템의 출력값이 가지는 불확실성에 대해 입력값들 각각의 불확실성이 차지하는 중요도를 정량화하는 연구로, 공학적 시스템의 외부환경에 대한 응답에서 입력변수들의 중요도, 구조물 최적 설계의 변수 설정 등과 같이 다양한 공학 분야에서 사용되고 있다. 특히, 소볼의 전역 민감도 해석 방법론이 비선형 시스템에도 적용될 수 있는 이의 범용성으로 인해 주로 연구되어 왔다.

그러나, 출력값의 표본을 필요로 하는 샘플링 기반 및 대리모형 기반 방법론과 같은 기존 방법론들은 긴 해석 시간을 필요로 하는 실제 공학적 문제의 고도화된 시스템에 적용하기 어렵다는 한계를 가진다.

따라서 이 연구에서는 계산이 어려운 복잡 시스템에 효율적으로 적용될 수 있는 적응형 가우시안 프로세스를 이용한 전역 민감도 해석 방법론을 제안한다. 훨씬 적은 수의 모델 해석으로도 실제 시스템과 매우 유사한 대리 모형을 수립할 수 있도록 적응형 학습 방법론을 가우시안 프로세스에 도입했다. 특히, 예측 분산만을 사용하는 기존 방법론들과 달리 Leave-One-Out Cross-Validation을 이용하여 편향 정보 또한 고려하는 적응형 가우시안 프로세스를 사용하여 탐색과 탐험 모두를 수행할 수 있는 방법론을 제안한다.

마지막으로, 해석 대상 시스템에 필요한 출력값 표본의 적절한 수를 알고리즘이 구할 수 있도록 민감도 해석에 맞는 적응형 가우시안 프로세스의 목적 함수를 제안한다.

전단 빌딩 및 9층 골조 구조물로 구성된 두 가지 구조물 예제를 통해 제안한 전역 민감도 해석 알고리즘을 검증하였고, 본 연구에서 제안한 방법을 통해 더 적은 수의 시스템 해석으로 상당히 정확한 민감도 지표 값에 수렴하는 효과적인 민감도 해석을 수행할 수 있을 것으로 기대된다.

주요어 : 전역 민감도 해석, 적응형 학습, 가우시안 프로세스, 실험계획법

학번 : 2021-29502

감사의 글

많이 방황하고 흔들렸던 학위기간 동안 저를 이끌어주신 송준호 교수님과 이상리 박사님께, 진심으로 감사드립니다. 기다려주시고 격려해주신 두 분의 가르침 덕분에 부족하기만한 제가 졸업까지의 여정을 무사히 마칠 수 있었습니다.

표현하지는 못했지만, 저의 미숙함을 뒤에서 각자의 방식으로 배려해주시고 도와주신 연구실 분들께 죄송하고, 감사드립니다.

각자 몫의 힘듦을 감내하는 것만으로 벅찼을텐데도 버팀목이 되어 준 도현이와 수민이, 의연함을 배우게 해 준 유현이, 벼랑 끝의 못난 투정도 진지하게 들어주었던 영민이, 저보다 더 저를 믿어준 성한 오빠, 존재만으로 힘이 되어주는 은서와 윤영이. 제 주변에 머물러 주신 모든 분들께, 고맙고, 사랑합니다.

이따금 힘들기도 했지만, 사랑하고 사랑해주신 분들의 소중함을 깨닫고 저의 부족함을 돌아볼 수 있는 시간이었기에, 그런 성장의 기회를 가질 수 있었음에 깊이 감사드립니다.

그리고 무엇보다도, 토목공학과 연구를 사랑하는 분들과 함께 할 수 있어 행복한 시간이었습니다. 모자란 스스로가 부끄럽기도 했지만 저와 같은 것들을 열망하는 사람들과 있으면서 깊이 즐거웠습니다.

받은 배려와 가르침이 헛되지 않도록, 공학도로써 쓰임받을 수 있도록 정진하겠습니다.