Proc. of the 8th International Symposium on Geotechnical Safety and Risk (ISGSR)
Edited by Jinsong Huang, D.V. Griffiths, Shui-Hua Jiang, Anna Giacomini and Richard Kelly

Published by Research Publishing, Singapore.
doi: 10.3850/978-981-18-5182-7_17-005-cd

ISGSR 2022

Observation Updating of Model Parameters for Consolidation Settlement Using Adaptive Surrogate Model

Tomoka Nakamura[1] and Ikumasa Yoshida[2]

[1]School of Engineering, Tokyo City University, Tamazutsumi 1-28-1, Setagaya-ku, Tokyo, Japan.
E-mail: thomoka1226@gmail.com
[2]Department of Urban and Civil Engineering, Tokyo City University,
Tamazutsumi 1-28-1, Setagaya-ku, Tokyo, Japan.
E-mail: iyoshida@tcu.ac.jp

Abstract: Data assimilation or observation updating is a method for improving prediction by using observation data to reduce the uncertainty. Particle filter (PF) is one of data assimilation methods based on Bayesian inference and Monte Carlo approach. In PF, posterior probability distributions (PDFs) are expressed by many sample realizations. PF has a problem known as degeneracy, where weights tend to concentrate into only a few particles, which causes poor computational performance. If many particles are used, the degeneracy can be avoided, however it requires high computation cost. Adaptive surrogate model approach for reliability analysis have attracted attentions for their low computation cost recently. We introduce a method to estimate posterior probability distribution directly by using surrogate model constructed by adaptive Gaussian Process Regression (GPR) with multiple random fields. The learning function is also very important in active learning, and has two roles in estimation of posterior PDF: a) finding the location of large values of posterior PDF, b) avoiding the location where posterior PDF is once calculated. The proposed method is demonstrated through a consolidation settlement problem. The settlement is predicted by using soil/water coupling FEM analysis. Model parameters of FEM are updated by synthesized observation data of settlements.

Keywords: Data assimilation; Gaussian process regression; Observation update; Particle filter.

1 Introduction

Bayesian model updating provides a useful probabilistic framework for assimilating the information contained in observation data into models, where a prior probability distribution of model parameters is updated to a posterior probability distribution. Except for special cases, the posterior distribution may not be amendable to analytical derivation. Sample-based approaches have been developed to evaluate the posterior probability distribution numerically. One popular method is particle filter (PF), in which probability distributions are approximated by their sample realizations, called 'particles' in PF algorithm (Ristic et al. 2003). Straub and Papaioannou (2015) proposes Bayesian Updating with Structural Reliability Methods (BUS). Subset simulation (Au and Beck 2003; Au and Wang 2014) is often combined with the idea of BUS for efficiently generating samples from the posterior distribution. Another useful method for Bayesian model updating is Transitional Markov chain Monte Carlo (TMCMC) proposed by Ching and Chen (2007). These methods basically require many function calls (full model evaluation, e.g., calculation by finite element method), consequently large computation cost.

One of attractive way to reduce the total computation cost is to use surrogate model (meta-model, response surface) in order to reduce the number of function call which generally requires large computational cost in real world problems. Echard et al. (2011) proposed an active learning with surrogate model by Kriging with MCS for reliability analysis. Since then, many modified methods have been proposed for reliability assessment, especially low failure probability, e.g., Echard et al. (2013), Huang et al. (2016), Kim et al. (2020).

This surrogate model approach can be also applied to Bayesian updating of model parameters. For example, as the surrogate model in BUS with subset simulation, Giovanis et al. (2017) incorporates artificial neural networks, Wang and Shafieezadeh (2020), and Kitahara et al. (2021) incorporate adaptive Kriging. Ni et al. (2021) proposed a variational Bayesian inference approach to estimate posterior probability distributions by using a surrogate model of an adaptive Gaussian process modeling technique and a Gaussian mixture model.

In this paper, PF with a surrogate model is adopted to estimate posterior probability distribution of model parameters for consolidation settlement prediction. The adaptive surrogate model is constructed by Gaussian process regression (GPR) with multiple random fields (Yoshida et al. 2021).

2 Algorithm of Bayesian update by PF with adaptive GPR

The process of updating $\boldsymbol{x}$ with $\boldsymbol{z}$ is referred as the observation updating process. The updated PDF of $\boldsymbol{x}$ given the information $\boldsymbol{z}$, i.e., the posterior distribution, is obtained using the Bayes' theorem:

$$p(\boldsymbol{x} \mid z) = \frac{p(\boldsymbol{x}, z)}{p(z)} = \frac{p(z \mid \boldsymbol{x}) p(\boldsymbol{x})}{\int p(z \mid \boldsymbol{x}) p(\boldsymbol{x}) d\boldsymbol{x}} \tag{1}$$

The PF algorithm starts by generating random samples, often referred as 'particles', from the prior PDF $p(\boldsymbol{x})$:

$$\boldsymbol{x}^{(j)} \sim p(\boldsymbol{x}), \qquad j = 1, \cdots, n \tag{2}$$

where the superscript (j) denotes the j-th particle. The prior PDF $p(\boldsymbol{x})$ then has the following approximate representation, called an 'empirical PDF':

$$p(\boldsymbol{x}) \cong \frac{1}{n} \sum_{j=1}^{n} \delta(\boldsymbol{x} - \boldsymbol{x}^{(j)}) \tag{3}$$

where δ (▪) denotes the Dirac delta function. Substituting Eq. (3) into Eq. (1) yields:

$$p(\boldsymbol{x} \mid z) = \frac{p(z \mid \boldsymbol{x}) \frac{1}{n} \sum_{j=1}^{n} \delta(\boldsymbol{x} - \boldsymbol{x}^{(j)})}{\int p(z \mid \boldsymbol{x}) \frac{1}{n} \sum_{i=1}^{n} \delta(\boldsymbol{x} - \boldsymbol{x}^{(i)}) d\boldsymbol{x}} \tag{4}$$

where,

$$a^{(j)} = \frac{p(z \mid \boldsymbol{x}^{(j)})}{\sum_{i=1}^{n} p(z \mid \boldsymbol{x}^{(i)})} \tag{5}$$

The term $a^{(j)}$ represents the 'weight' (likelihood ratio) of particle j after updating. PF evaluates denominator of Eq. (1), i.e., evidence C_E, simply by:

$$p(z) \cong \int p(z \mid \boldsymbol{x}) \frac{1}{n} \sum_{i=1}^{n} \delta(\boldsymbol{x} - \boldsymbol{x}^{(i)}) d\boldsymbol{x} = \frac{1}{n} \sum_{i=1}^{n} p(z \mid \boldsymbol{x}^{(i)}) \tag{6}$$

The algorithm of PF is combined with the idea of adaptive surrogate model. GPR with two random Gaussian random fields is used as surrogate model to express the distribution of posterior PDF in this study. The procedure for the Bayesian update PF with adaptive GPR can be summarized as follows.

Step 1 Generate model parameter vector $\boldsymbol{x}_{\mathrm{cal}}^{(j)}$ and calculate the numerator of posterior PDF $\boldsymbol{n}_{\mathrm{cal}}=\{\log[p(z|\boldsymbol{x}_{\mathrm{cal}}^{(i)})p(\boldsymbol{x}_{\mathrm{cal}}^{(i)})]$, i=1, $j\}$. Use these samples to construct a surrogate model of the function $\log[p(z|\boldsymbol{x})p(\boldsymbol{x})]$ by GPR. Generate a large number of model parameter vectors $\boldsymbol{x}_{\mathrm{gpr}}^{(i)}$. Use these to calculate the evidence $C_{E,j}$ and posterior PDF based on the constructed surrogate model.

Step 2 Assume (or estimate) a standard deviation (SD) σ_k and a scale of fluctuation (SOF) $\delta_{k,i}$. Calculate the covariance matrix $\boldsymbol{M}$ involved in the GPR.

Step 3 Determine the next sample for the calculation of the numerator as the one that maximizes the learning function $U(\boldsymbol{x})$, i.e., $\boldsymbol{x}_{\mathrm{cal}}^{(j+1)}$= argmax($U(\boldsymbol{x})$). $U(\boldsymbol{x})$ is explained below

Step 4 Calculate the logarithm of the numerator of posterior PDF, i.e., $n_{\mathrm{cal}}^{(j+1)}$, at $\boldsymbol{x}_{\mathrm{cal}}^{(j+1)}$. Expand the set of samples by including $n_{\mathrm{cal}}^{(j+1)}$ in the vector $\boldsymbol{n}_{\mathrm{cal}}$ and $\boldsymbol{x}_{\mathrm{cal}}^{(j+1)}$ in $\boldsymbol{X}_{\mathrm{cal}}=[\boldsymbol{x}_{\mathrm{cal}}^{(1)}, \boldsymbol{x}_{\mathrm{cal}}^{(2)}, \ldots, \boldsymbol{x}_{\mathrm{cal}}^{(j)}, \boldsymbol{x}_{\mathrm{cal}}^{(j+1)}]$. Advance the sample size index j by 1.

Step 5 Calculate the evidence $C_{E,j}$ and posterior PDF by using the particles $\boldsymbol{x}_{\mathrm{gpr}}^{(i)}$ and surrogate model GPR ($\boldsymbol{x}$| $\boldsymbol{n}_{\mathrm{cal}}$**,** $\boldsymbol{X}_{\mathrm{cal}}$).

Step 6 If j is greater than the required sample size, stop. Otherwise, go to Step 2 to generate a new sample

The learning function $U(\boldsymbol{x})$ in Step 3 is based on the constructed surrogate model in Step 2 and is given by

$$U(\boldsymbol{x}) = r(\boldsymbol{x})(\mu(\boldsymbol{x}) - \min(\boldsymbol{n}_{\mathrm{cal}})) \tag{7}$$

$$r(\boldsymbol{x}) = \sigma_{\mathrm{poeterior}}(\boldsymbol{x}) / \sigma_{\mathrm{prior}}(\boldsymbol{x}) \tag{8}$$

where $\sigma_{\mathrm{poeterior}}(\boldsymbol{x})$ and $\sigma_{\mathrm{prior}}(\boldsymbol{x})$ are posterior and prior SD at $\boldsymbol{x}$ of the surrogate model by GPR. The learning function has two roles: a) finding the location with large values of $p(z|\boldsymbol{x})p(\boldsymbol{x})$, and b) avoiding the location where $p(z|\boldsymbol{x})p(\boldsymbol{x})$ has already been calculated so far.

The proposed method is illustrated using the single-variable example. It involves updating a scalar random variable x whose prior PDF is assumed to be N(0,1), i.e., Gaussian with zero mean and unit SD. The observation z has a mean value of 4 and its observation error is Gaussian with a SD of 0.6. This problem has an analytical solution: the posterior distribution is Gaussian with a mean of 2.94 and a SD of 0.51, i.e., N(2.94, 0.51^2). In Step 1, two samples $\boldsymbol{x}_{\mathrm{cal}}^{(i)}$, i=1, j (j=2) are initially generated and the logarithm of numerator $\boldsymbol{n}_{\mathrm{cal}}$ is calculated. In Step 2, the SD σ and SOF δ for estimating $\mu_n(\boldsymbol{x})$ are assumed as follows: σ_1=2 and $\delta_{1,1}$ = 3 for the first random field; σ_2=1 and $\delta_{2,1}$ =0.3 for the second random field. In Step 3, the learning function $U(\boldsymbol{x})$ is maximized to give a new sample $\boldsymbol{x}_{\mathrm{cal}}^{(j+1)}$ = argmax($U(\boldsymbol{x})$) for the next function call. The SOF δ_r for $r(\boldsymbol{x})$ of the learning function is assumed to be equal to that of the second random filed, i.e., $\delta_r = \delta_{2,1}$ =0.3.

Figure 1 shows the true numerator of the posterior PDF (black line), the value of numerator (black dot) of the two samples $\boldsymbol{x}_{\mathrm{cal}}^{(j)}$, j=1,2 generated in Step 1, surrogate model estimated by GPR (red line) based on these two samples, the learning function $U(\boldsymbol{x})$ (blue line) and the next sample $\boldsymbol{x}_{\mathrm{cal}}^{(j+1)}$ (red dot) to calculate the numerator. Figure 1(a) shows the surrogate distribution (red line) does not match the true one (black line) well because there are only two points for the estimation. The learning function is the same as the surrogate distribution except near the two sampling points. The next sample $\boldsymbol{x}_{\mathrm{cal}}^{(j+1)}$ is determined such that the learning function is maximized, i.e., at the peak of the blue line. Step 2 to 6 are repeated until the number of function calls j reaches 16. Figure 1(b) shows the results when j=16. There are more samples near the peak, and the estimated distribution is in good agreement with the true one.

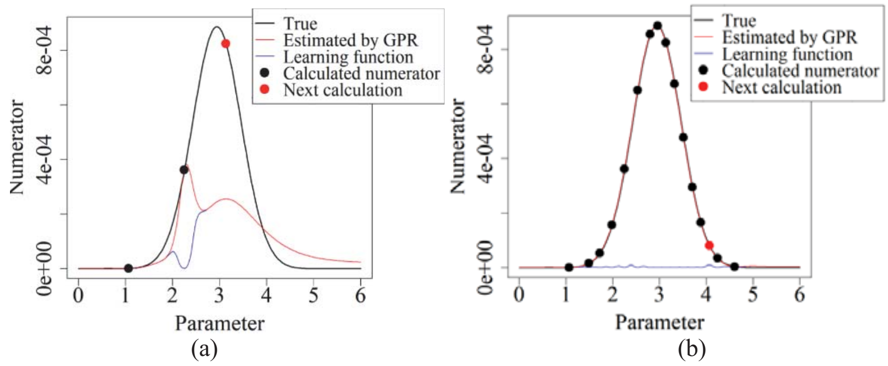
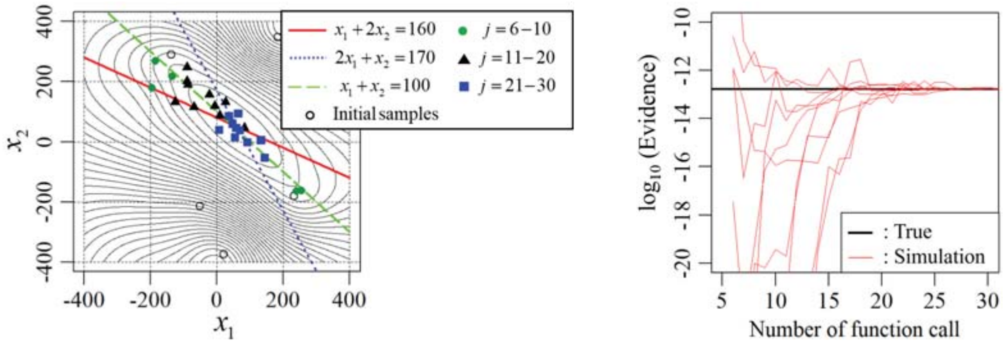


Figure 1. True and estimated numerator of posterior PDF, and learning function of single variable illustrative example, (a)the number of function call is two (j=2) and (b) the number of function call is 16 (j=16).

The proposed method is also demonstrated by the simple linear inverse problem with three observations and two unknown variables.

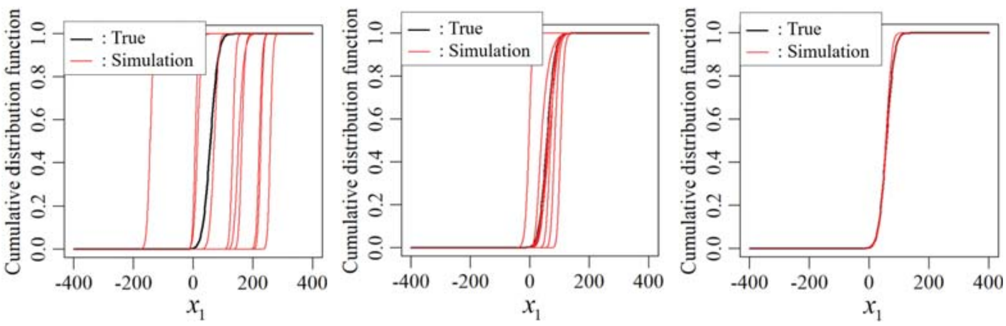
$$\boldsymbol{z}=\begin{pmatrix}160\\170\\100\end{pmatrix},\boldsymbol{x}=\begin{pmatrix}x_1\\x_2\end{pmatrix},\boldsymbol{H}=\begin{bmatrix}1&2\\2&1\\1&1\end{bmatrix},\boldsymbol{R}=\begin{bmatrix}30^2&0&0\\0&30^2&0\\0&0&30^2\end{bmatrix} \quad (9)$$

where, observation z is given by z=Hx+v, R indicates covariance matrix of observation error v. The parameter vector x whose prior PDF is assumed to be N(0,$100^2 I$), where 0 and I indicates zero vector and an identity matrix. In Step 1, five samples $x_{\mathrm{cal}}(i)$, i=1, j (j=5) are initially generated and the logarithm of numerator n_{cal} is calculated. In Step 2, the SD σ and SOF δ for surrogate model by GPR are assumed as follows: σ_1 = 2, $\delta_{1,1}$ = 8.0×10^2 for the first random field; σ_2 = 1, $\delta_{2,1}$ =1.6×10^2 for the second random field. These parameters were determined empirically. In Step 3, the learning function $U(x)$ is maximized by particle swarm optimization (PSO) to give a new sample $x_{\mathrm{cal}}^{(j+1)}$ for the next function call. In the implementation of PSO, the number of population and generation (iteration) are taken to be 100 and 50, respectively. Figure 2(a) shows the estimated posterior PDF and thirty function call points. The three lines indicate the observation equation, z=Hx. Figure 2(b) shows the convergence process of evidence $C_{E,j}$. The evidence gradually converges and becomes almost constant after 25 function calls. The estimated CDFs of parameter 1 by 10 independent runs are shown in Figure 3, where the number of function call is 10, 20, and 30. The estimated CDFs by 30 function calls are in good agreement with the true one.



(a) (b)

Figure 2. Estimated posterior distribution after 30 function calls of two-dimensional linear problem, (a) Function call points and contour by surrogate model, (b) log10(Evidence $C_{E,j}$).



(a) j=10 (b) j =20 (c) j =30

Figure 3. Estimated CDFs of parameter 1 by 10 runs for the two-dimensional linear problem, (a) after 10 function calls, (b) after 20 function calls, (c) after 30 function calls.

3 Model for settlement prediction

Figure 4 shows the model ground and observation locations for settlement prediction. The embankment is constructed on the ground consisting of five layers: the first and fifth layers are sand, and the second, third and fourth layers are clay. The first and fifth layers are modelled by linear elastic model, and second, third, and fourth layers are modelled by Cam-clay model. The observations of settlement are given at location S1 under the embankment.

Figure 5 shows the FE mesh used in the analysis. The half of the ground area is modelled because of the symmetry. The parameters of the second to fourth layers used in the analysis are shown in Table 1. The elastic modulus, Poisson's ratio and coefficient of permeability of layer 1 and 5, which are assumed to be linear elastic and deterministic, are 1.28×10^5 (kPa), 0.33, 1.00×10^{-3} (m/s) respectively. Coefficient of permeability k and compression index λ of second clay layers are estimated by the proposed method based on the observation data. In addition to these two parameters of the clay layers, parameters α and β for loading process are also estimated. The total number of parameters for estimation is four.

Table 1 Deterministic model parameters for settlement prediction.

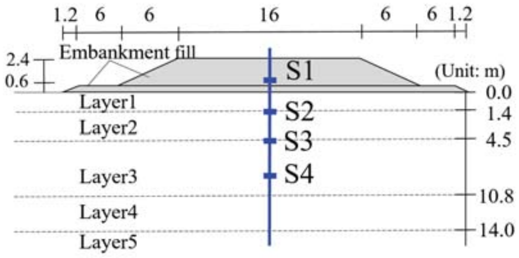
	v	κ	e	M	K_0	K_i	λ_K	σ_v' (kPa)	OCR	λ	k (m/s)
Layer 2	0.348	0.123	1.837	1.023	0.563	0.563	0.833	14.70	1.00	-	-
Layer 3	0.348	0.123	1.525	1.023	0.563	0.563	0.833	42.35	1.00	0.568	0.006
Layer 4	0.348	0.123	1.278	1.023	0.563	0.563	0.833	97.65	1.00	0.512	0.003

v: Poisson's ratio, κ: Swelling index, e: Void ratio, M: Critical state parameter, K_0: Coefficient of earth pressure at rest, K_i: Coefficient of earth pressure at rest (overconsolidation), λ_K: Permeability change index, σ_v': Overburden pressure (kPa), OCR: Over consolidation ratio, λ: Compression index, k: Coefficient of permeability (m/s)

Table 2 Model parameters for estimation.

	Parameter	Range
Loading condition	$\log_{10}\alpha$	-0.1 – 0.1
	$\log_{10}\beta$	-1.7 – 0.2
Layer 2	$\log_{10}\lambda$	-0.4 – 0.17
	$\log_{10}k$ (m/s)	-3.0 – 1.0

λ: Compression index, k: Coefficient of permeability (m/s)



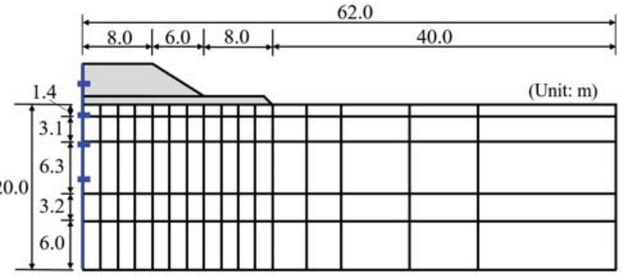
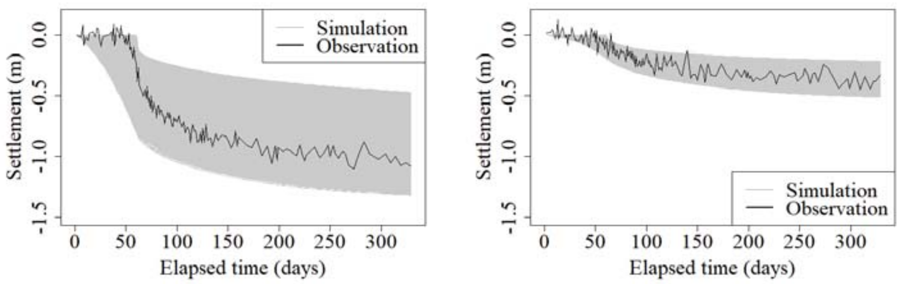


Figure 4. The model ground and observation locations S1-S4. **Figure 5.** Finite element mesh for settlement prediction.

The loading process by embankment construction is often treated deterministically in simulations. However, it may contain uncertainties especially when the simulation model for settlement is two-dimensional because the real world construction process is three-dimensional. The loading process is defined by Eq. (10), and the parameters α and β are also estimated based on the observation data.

$$f_t = \begin{cases} \alpha F_0 \{(\exp(\beta t)-1)/(\exp(\beta T)-1)\} & t \leq T \\ \alpha F_0 & t > T \end{cases} \tag{10}$$

where, f_t is the embankment load at time t, F_0 is final embankment load, volume multiplied by unit weight of soil, T is construction period. The four parameters listed in the Table 2 are estimated based on the synthesized time-series of settlement shown in Figure 6 (black line). The logarithm of model parameters are normalized to the standard space [0, 1] by the range indicated in Table 2.



(a)S1 (b)S4

Figure 6. MCS with respect to the model parameters in Table.1 and observation data for updating at S1 and S4.

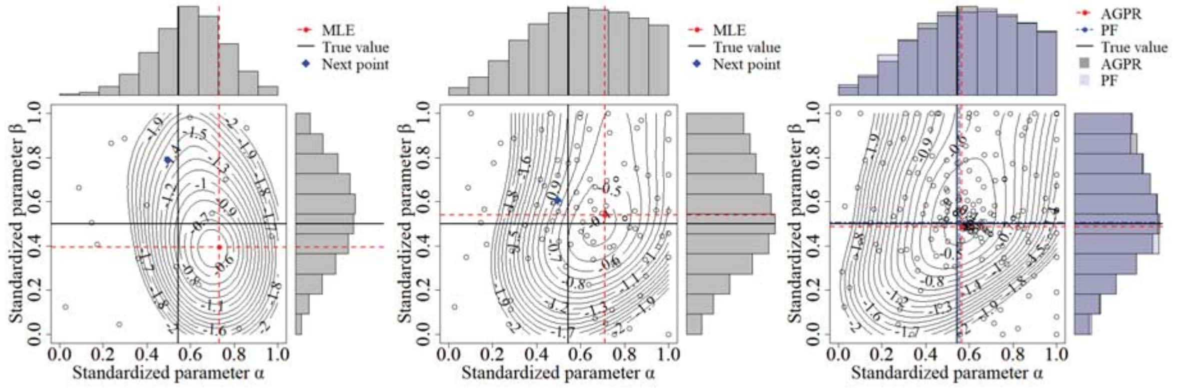
4 Updating of model parameters and settlement prediction

4.1 Procedure of updating by PF with adaptive GPR

In Step 1, twenty initial samples ($\boldsymbol{x}_{\text{cal}}^{(i)}$, i=1, 20) are generated by Sobol method, at which the logarithm of numerator $\boldsymbol{n}_{\text{cal}}$ is calculated. In Step 2, the initial values of SD σ and SOF δ for estimating $\mu_n(\boldsymbol{x})$ are assumed to be $\sigma_1 = 10$, $\sigma_2 = 5$, $\delta_{1,i}$=1.0, and $\delta_{2,i} = 0.2$, i=1, 4 in standardized space. These 10 parameters are estimated by ML after 70, 120 and 170 function calls. It is assumed that the SOF δ_r in $r(\boldsymbol{x})$ of the learning function follow uniform distribution of U(0.1, 0.4) up to 100 function calls, and then U(0.001, 0.4) thereafter. This is intended to capture the overall trend of the target distribution in the early stage, and then the details later. In Step 3, PSO is used to search a next sample $\boldsymbol{x}_{\text{cal}}^{(j+1)}$ for function call in the same way as two-dimensional problems. The number of population and generation (iteration) of PSO are 100 and 50. Steps 2 to 6 are repeated until the number of function calls j reaches 200.

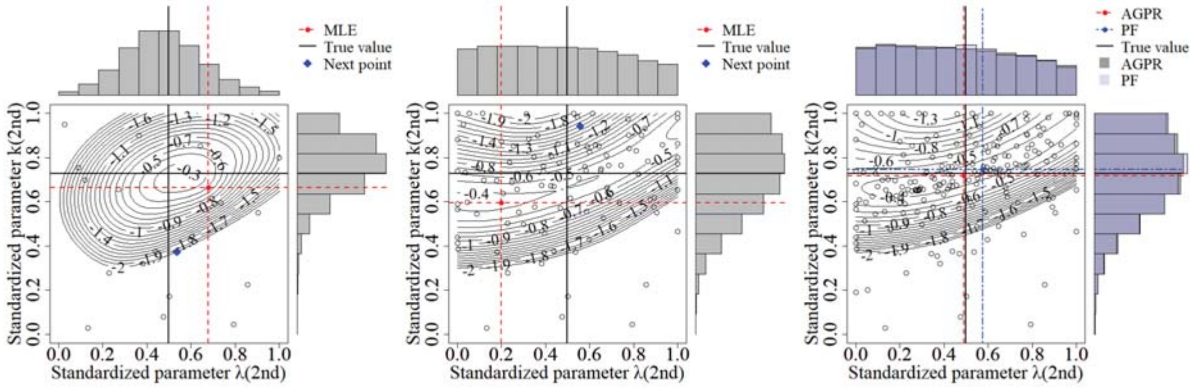
4.2 Updated parameters for settlement prediction

Figure 7 and 8 show the estimated parameters by direct PF and PF with adaptive surrogate model, the histograms of the loading parameter α and β, the coefficient of permeability k, and compression indices λ of the second layer.



(a) initial(j=20) (b) j=100 (c) j=200

Figure 7. Histogram and contour map of cross-sectional joint PDF of estimated model parameters α and β(standardized), (a) number of function call is 20 (j=20), (b) number of function call is 100 (j=100) and (c) final result (j=200).



(a) initial(j=20) (b) j=100 (c) j=200

Figure 8. Histogram and contour map of cross-sectional joint PDF of estimated model parameters compression index λ and permeability k of 2nd layer (standardized), (a) number of function call is 20 (j=20), (b) number of function call is 100 (j=100) and (c) final result (j=200).

Compression indices λ are not sensitive because their histogram are almost flat. The other parameters vary significantly, indicating that they are sensitive/influential parameters. The cross lines in Figure 7(c) and 8(c) indicate the estimates with the largest likelihood value among the 200 function calls by the proposed method, and 20,000 function calls by direct PF. For sensitive parameters, there is a good agreement with the true parameter values. Based on the results presented here, the estimates with the maximum likelihood values among samples by the proposed method are better than those by PF. The estimate of posterior distribution by the proposed method with only 200 function calls can be considered acceptable if fine detail is not required.

5 Conclusions

PF with adaptive GPR for Bayesian updating method has been applied to update settlement predictions simulated by soil-water coupled FEM with Cam-clay model. The posterior distributions of four model parameters have been estimated by 200 function calls and compared with those estimated by direct PF with 20,000 function calls. Three parameters except the compression indices, which is not sensitive, show good agreement with true values. In this study, the four-dimensional problem is solved and the good result is obtained. The applicability to a higher dimensional problem and stopping criteria are future research topics.

References

Au, S.K. and Beck, J.L. (2003). Subset simulation and its application to seismic risk based on dynamic analysis. *Journal of Engineering Mechanics,* 129(8):901-917.

Au, S.K. and Wang, Y. (2014). Engineering risk assessment with subset simulation. *John Wiley & Sons Inc.*

Ching, J. and Chen, YC. (2007). Transitional Markov chain Monte Carlo Method for Bayesian model updating, model class selection, and model averaging. *J Eng Mech*, 133(7):816-832.

Echard, B., Gayton, N. and Lemaire, M. (2011). AK-MCS: an active learning reliability method combining Kriging and Monte Carlo simulation. *Struct Saf*, 33(2):145-54.

Echard, B., Gayton, N., Lemaire, M. and Relun, N. (2013). A combined importance sampling and Kriging reliability method for small failure probabilities with time-demanding numerical models. *Reliab Eng Syst Safety*, 111:232-240.

Giovanis, D.G., Papaioannou, I., Straub, D. and Papadopoulos, V. (2017). Bayesian updating with subset simulation using artificial neural networks. *Computer Methods in Applied Mechanics and Engineering*, 319, 124-145.

Huang, X., Chen, J. and Zhu, H. (2016). Assessing small failure probabilities by AK-SS: An active learning method combing Kriging and Subset Simulation. *Struct Saf*, 59:86-95.

Kim, J. and Song, J.(2020). Probability-Adaptive Kriging in n-Ball (PAK-Bn) for reliability analysis. *Struct Saf*, 85:101924.

Kitahara, M., Bi, S., Broggi, M. and Beer, M. (2021). Bayesian Model Updating in Time Domain with Metamodel-Based Reliability Method. *ASCE-ASME J. Risk Uncertainty Eng. Syst., Part A: Civ. Eng.*,7(3): 04021030.

Ni, P., Li, J., Hao, H., Han, Q. and Du, X. (2021). Probabilistic model updating via variational Bayesian inference and adaptive Gaussian process modeling. *Computer Methods in Applied Mechanics and Engineering*, 383, 113915.

Ristic, B., Arulampalam. S., and Gordon, N. (2003). Beyond the Kalman filter: particle filters for tracking applications. *Artech House, U.K.*

Straub, D. and Papaioannou, I. (2015). Bayesian updating with structural reliability methods. *J Eng Mech*;141(3):04014134.

Wang, Z. and Shafieezadeh, A. (2020). Highly efficient Bayesian updating using metamodels: An adaptive Krigingbased approach. *Struct Saf*, 84, 101915.

Yoshida, I., Tomizawa, Y., and Otake, Y. (2021). Estimation of trend and random components of conditional random field using Gaussian process regression. *Computers and Geotechnics*, 136.