Proceedings of the 21st ICSMGE, Vienna, Austria, 14 – 19 June 2026. Pistrol, Adam & Schweiger (eds.)
Published by: ÖGG, Austrian Society for Geomechanics, Salzburg, Austria, ISBN 978-3-9503898-4-5
https://doi.org/10.53243/ICSMGE2026-2392

The influence of prior on the simultaneous estimation of geometric features and physical property fields using Bayesian inference

Tatsuya Shibata, Kazunori Fujisawa
Graduate School of Agriculture, Kyoto University, Japan, shibata.tatsuya.84z@st.kyoto-u.ac.jp

Michael C. Koch
Research School of Earth Sciences, Australian National University, Australia

ABSTRACT: When estimating subsurface structures, a priori information such as geological maps, borehole records, and design drawings can be incorporated into the estimation model as definitive information on geometric features like the locations of interfaces and buried objects. However, this a priori information rarely matches the actual situations exactly and generally includes a certain degree of uncertainty. As a result, the inverse analysis is performed using a model that differs from the actual situation, which introduces uncertainty in the estimation results. Our proposed approach, which simultaneously estimates geometric features and the physical property distribution (referred to as the simultaneous estimation method), can effectively address this problem. In our method, geometric features are treated as estimation targets rather than as definitive information. Known geometric features are explicitly parameterized and estimated simultaneously with the surrounding physical property distribution. Furthermore, since the simultaneous estimation method is based on Bayesian inference, a priori information about the geometric features is naturally incorporated into the estimation process as a prior distribution, and the estimation results are obtained as a posterior distribution. However, even when geometric features are estimated, if the a priori information differs significantly from the actual situations, the prior distribution may be set inappropriately, which can decrease estimation accuracy. Therefore, in this study, we aim to investigate the influence of a priori information on geometric features in the results of simultaneous estimation by comparing the results for different hyperparameter sets. For this investigation, we addressed an inverse problem to estimate both the locations of layer interfaces and the hydraulic conductivity distribution within each layer from seepage flow observation data. The results demonstrated that when appropriate hyperparameters are selected for the hydraulic conductivity distribution, the interface locations can be estimated accurately, even if the a priori information regarding the layer interface locations is somewhat inaccurate.

KEYWORDS: Inverse problem, interface detection, hyperparameters.

1 INTRODUCTION

Visualization of subsurface structures using non-destructive testing is typically achieved by estimating the spatial fields of physical properties, such as hydraulic conductivity, electrical resistivity, elastic modulus, and P- and S-wave velocities, based on relevant observation data (Koch *et al.*, 2020; Yang *et al.*, 2019). However, in most cases, these data are noisy and limited, which introduces uncertainty in the estimation. This uncertainty makes it difficult to apply non-destructive testing in areas where high accuracy is required, such as structural design.

Bayesian inference provides an effective framework for dealing with such uncertainties. In this framework, the spatial field of physical properties is modeled as a random field. In terms of dimensionality reduction, the random field is efficiently parameterized using the Karhunen–Loève (KL) expansion (Ghanem & Spanos, 2003). For mathematical convenience, a Gaussian random field is often used as the prior. However, Gaussian priors tend to produce spatially smooth fields (Uribe *et al.*, 2020), which makes it difficult to detect sharp interfaces, such as geological boundaries or embedded objects, where physical properties change abruptly. As a result, it becomes challenging to achieve both dimensionality reduction and accurate interface detection simultaneously.

To address this issue, we previously proposed a method that explicitly introduces geometry parameters related to geometric features—such as buried anomalies and interfaces of soil layers—and estimates them simultaneously with the background spatial field (Koch *et al.*, 2021; Shibata *et al.*, 2025). This simultaneous estimation approach makes it possible to preserve the advantages of dimensionality reduction using the KL expansion while also allowing for sharp detection of interfaces.

In practical settings, prior geometric information is often available, such as geological maps, cone penetration data, borehole data, and structural design documents. However, such information rarely matches the actual subsurface structures exactly and may contain some uncertainty, making it difficult to utilize effectively in conventional deterministic estimation. Since the simultaneous estimation method is based on Bayesian inference, such geometric information can be treated not as definitive information but as probabilistic variables to be inferred. For example, parameters such as the depth of an interface or the radius of a buried object can be estimated simultaneously with the physical property field. Thus, prior knowledge about geometric features can be naturally incorporated into the estimation process as prior distributions.

However, if the prior distributions assigned to the geometric parameters differ significantly from the true conditions, the accuracy of the estimation may deteriorate. Therefore, handling prior information is a crucial issue when applying the simultaneous estimation method. In this study, we investigate the influence of prior information on geometric features in the estimation results.

The remainder of this paper is organized as follows. In Section 2, the modeling of the inverse problem is explained. Section 3 briefly describes the framework of Bayesian inference. In Section 4, hyperparameters of the simultaneous estimation method are explained. Section 5 presents the numerical study of inverse analysis of one-dimensional seepage flow for different hyperparameter choices.

2 MODELING OF THE INVERSE PROBLEM

2.1 *Karhunen–Loève expansion*

In Bayesian inference, the computational cost increases exponentially with the dimensionality of the parameters (curse of dimensionality). Therefore, when dealing with spatial fields, which are essentially infinite-dimensional objects, it is crucial to represent them using a small number of parameters. To address this issue, the simultaneous estimation approach

employs the Karhunen–Loève (KL) expansion to reduce the dimensionality of the parameter space.

Let $u(\mathbf{z})$ be a Gaussian random field defined over a domain $\mathcal{D}$, with mean $\bar{u}(\mathbf{z})$ and covariance kernel $C(\mathbf{z}, \mathbf{z}^*)$. The KL expansion of $u(\mathbf{z})$ is given by

$$\begin{aligned} u(\mathbf{z}, \omega) &= \bar{u}(\mathbf{z}) + \sum_{i=1}^{\infty} \sqrt{\lambda_i}\, \phi_i(\mathbf{z})\, {}^1\theta_i(\omega) \\ &\approx \bar{u}(\mathbf{z}) + \sum_{i=1}^{M} \sqrt{\lambda_i}\, \phi_i(\mathbf{z})\, {}^1\theta_i(\omega), \end{aligned} \tag{1}$$

where ${}^1\theta_i \sim \mathcal{N}(0,1)$, and λ_i and ϕ_i are the eigenvalues and eigenfunctions, respectively, of the following integral equation:

$$\int_{\mathcal{D}} C(\mathbf{z}, \mathbf{z}^*)\phi_i(\mathbf{z}^*)\mathrm{d}\mathbf{z}^* = \lambda_i \phi_i(\mathbf{z}). \tag{2}$$

By truncating the infinite series in Equation (1) and retaining only the first M terms with the largest eigenvalues, the random field $u(\mathbf{z})$ can be approximated using a finite-dimensional vector ${}^1\boldsymbol{\theta} = [\, {}^1\theta_1, \ldots, {}^1\theta_M]^\top$. This yields a significant reduction in the dimensionality of the random field.

In this study, we use a Gaussian kernel for the covariance function $C(\mathbf{z}, \mathbf{z}^*)$, given by

$$C(\mathbf{z}, \mathbf{z}^*) = \sigma^2 \exp\left(-\frac{|\mathbf{z} - \mathbf{z}^*|^2}{2l^2}\right), \tag{3}$$

where σ^2 is the variance (scale), and l is the correlation length. Both σ^2 and l are treated as hyperparameters. The correlation length l characterizes the spatial range over which significant correlation is expected. A larger l implies that the correlation extends over longer distances, allowing the KL expansion to be accurately approximated using fewer terms. However, this also leads to smoother realizations of the field, potentially suppressing local variations. In addition to σ^2 and l, the mean function $\bar{u}(\mathbf{z})$ is also treated as a hyperparameter in this framework.

2.2 *Parameterization of the inverse problem*

Consider a domain $\mathcal{D} = \bigcup \mathcal{D}_i$ consisting of multiple soil layers $\mathcal{D}_i$ $(i = 1,2, \ldots)$. The steady seepage flow through $\mathcal{D}$, governed by the continuity equation and Darcy's law, is represented as

$$\nabla \cdot \mathbf{q}(\mathbf{z}) = Q(\mathbf{z}), \tag{4}$$

$$\mathbf{q}(\mathbf{z}) = -k(\mathbf{z})\nabla h(\mathbf{z}), \tag{5}$$

where $\mathbf{z} \in \mathcal{D}$, $h(\mathbf{z})$ denotes the hydraulic head, $\mathbf{q}(\mathbf{z})$ is seepage flow velocity, $Q(\mathbf{z})$ is the source term, and $k(\mathbf{z})$ is the hydraulic conductivity field. In this study, Dirichlet boundary conditions are prescribed on the entire boundary of the domain.

For this steady seepage flow problem, an inverse problem is set up to estimate the hydraulic conductivity $k(\mathbf{z})$ within the domain $\mathcal{D}$, based on measurements of the hydraulic head and flow rate. The interfaces between soil layers are explicitly parameterized by the geometry parameter ${}^2\boldsymbol{\theta}$, which allows the domain $\mathcal{D}$ to be divided into multiple layers $\mathcal{D}_i$. The hydraulic conductivity $k_i(\mathbf{z})$ for each layer $\mathcal{D}_i$ is the target to be estimated. To ensure non-negativity, $k_i(\mathbf{z})$ is modeled through the KL expansion of the logarithmic hydraulic conductivity $u_i(\mathbf{z}) = \log_{10} k_i(\mathbf{z})$. Each $u_i(\mathbf{z})$ is represented by a random vector ${}^1\boldsymbol{\theta}_i$, and all such vectors are collectively written as ${}^1\boldsymbol{\theta} = \left[\, {}^1\boldsymbol{\theta}_1,\ {}^1\boldsymbol{\theta}_2, \ldots\right]^\top$. The complete set of parameters to be inferred is thus given by $\boldsymbol{\theta} = \left[\, {}^1\boldsymbol{\theta},\ {}^2\boldsymbol{\theta}\right]^\top$.

If the geometry parameter ${}^2\boldsymbol{\theta}$ is not introduced, the domain $\mathcal{D}$ cannot be decomposed into layers, and the KL expansion must be applied over the entire domain $\mathcal{D}$. Due to the smoothness of KL expansions with a Gaussian prior, this approach leads to overly smoothed estimates and obscures the interfaces between regions (Shibata *et al.*, 2025). The introduction of the geometry parameter ${}^2\boldsymbol{\theta}$ enables flexible modeling by reducing the dimensionality of the KL expansion while, at the same time, complementing the drawback of the KL expansion, which cannot represent abrupt spatial variations in physical properties.

3 BAYESIAN PARAMETER ESTIMATION

Bayes' theorem is expressed as follows:

$$p(\boldsymbol{\theta}|\mathbf{y}) = \frac{p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{\int p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}}, \tag{6}$$

where $\mathbf{y}$ denotes the observation, assumed to follow the observation model

$$\mathbf{y} = \mathbf{x}(\boldsymbol{\theta}) + \mathbf{r}. \tag{7}$$

Here, $\mathbf{x}(\boldsymbol{\theta})$ represents the state vector (i.e., the output of the forward model given parameters $\boldsymbol{\theta}$), which in this study corresponds to hydraulic head and flow rate. In problems governed by partial differential equations, such as seepage flow, $\mathbf{x}(\boldsymbol{\theta})$ is generally not available in closed form. The term $\mathbf{r}$ denotes observational noise, which is assumed to follow a Gaussian distribution $\mathcal{N}(\mathbf{0}, \mathbf{R})$. Based on Equation (7), the likelihood $p(\mathbf{y}|\boldsymbol{\theta})$ can be written as

$$p(\mathbf{y}|\boldsymbol{\theta}) = \mathcal{N}(\mathbf{x}(\boldsymbol{\theta}), \mathbf{R}). \tag{8}$$

The prior distribution $p(\boldsymbol{\theta})$ is assumed to be Gaussian. Specifically, the prior of ${}^1\boldsymbol{\theta}$, associated with the KL expansion, is $p({}^1\boldsymbol{\theta}) = \mathcal{N}(\mathbf{0}, \mathbf{I}_M)$ as a natural choice. For the geometry parameter ${}^2\boldsymbol{\theta}$, we adopt the prior

$$p({}^2\boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2), \tag{9}$$

where $\boldsymbol{\mu}_2$ and $\boldsymbol{\Sigma}_2$ are hyperparameters. If physical constraints restrict the possible values of ${}^2\boldsymbol{\theta}$, a truncated Gaussian distribution should be used instead. The posterior $p(\boldsymbol{\theta}|\mathbf{y})$ represents the distribution of $\boldsymbol{\theta}$ conditioned on the observation $\mathbf{y}$, and is the target to be estimated in the inverse problem. Substituting the prior $p(\boldsymbol{\theta})$ and likelihood $p(\mathbf{y}|\boldsymbol{\theta})$ into Equation (6), the posterior can be written as

$$p(\boldsymbol{\theta}|\mathbf{y}) \propto \exp\big(-\varphi(\boldsymbol{\theta})\big), \tag{10}$$

where $\varphi(\boldsymbol{\theta})$ ig given as follows:

$$\begin{aligned} \varphi(\boldsymbol{\theta}) \equiv{}& \frac{1}{2}\big(\mathbf{y} - \mathbf{x}(\boldsymbol{\theta})\big)^\top \mathbf{R}^{-1}\big(\mathbf{y} - \mathbf{x}(\boldsymbol{\theta})\big) \\ &+ \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}_\theta^{-1}(\boldsymbol{\theta} - \boldsymbol{\mu}). \end{aligned} \tag{11}$$

Since $\varphi(\boldsymbol{\theta})$ includes the forward model output $\mathbf{x}(\boldsymbol{\theta})$, it is generally not analytically tractable with respect to $\boldsymbol{\theta}$, implying that $p(\boldsymbol{\theta}|\mathbf{y})$ cannot be obtained in closed form. In such cases, sampling approaches, such as Markov chain Monte Carlo (MCMC) methods, are typically used to numerically approximate the posterior distribution. Since all MCMC methods cancel out the normalization constant, it suffices to consider $\exp\big(-\varphi(\boldsymbol{\theta})\big)$.

In this study, we adopt a simultaneous estimation approach based on the Hamiltonian Monte Carlo (HMC) method (Neal, 2011), a variant of MCMC. By exploiting the gradient with

respect to the posterior $\partial\varphi(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$, HMC significantly improves sampling efficiency compared to conventional methods such as the Metropolis-Hastings algorithm, which relies on random-walk proposals.

4 HYPERPARAMETERS

In the simultaneous estimation, the hyperparameters include the mean vector $\boldsymbol{\mu}_2$ and covariance matrix $\boldsymbol{\Sigma}_2$ for the geometric parameters, and the mean $\bar{u}$, scale σ^2, and correlation length l for the logarithmic hydraulic conductivity u, which is represented using the KL expansion. The effect of the choice of the correlation length l on the estimation results has been investigated by Shibata *et al.* (2025). In this study, we further examine how the choice of the mean $\boldsymbol{\mu}_2$ and covariance $\boldsymbol{\Sigma}_2$ for the geometry parameters influences the estimation results, in addition to the correlation length l. It is desirable to set $\boldsymbol{\mu}_2$ and $\boldsymbol{\Sigma}_2$ such that the prior distribution adequately covers the true values. In cases where prior information is limited, an unbiased prior can be constructed by choosing the mean $\boldsymbol{\mu}_2$ as the midpoint of the plausible range of the parameters and setting the covariance $\boldsymbol{\Sigma}_2$ to a relatively large value. On the other hand, if prior knowledge, such as geological maps or columnar sections, indicates the potential extent of geometric features, this information should be reflected in the choice of $\boldsymbol{\mu}_2$ and $\boldsymbol{\Sigma}_2$. However, it should be noted that if the prior information is inaccurate, it may degrade the estimation accuracy. In this study, we assume that appropriate values for the mean $\bar{u}$ and the scale σ^2 of the KL expansion can be selected. This assumption is justified by the availability of prior information on soil properties obtained from sources such as cone penetration tests or borehole exploration, which can provide a basis for setting reasonable values.

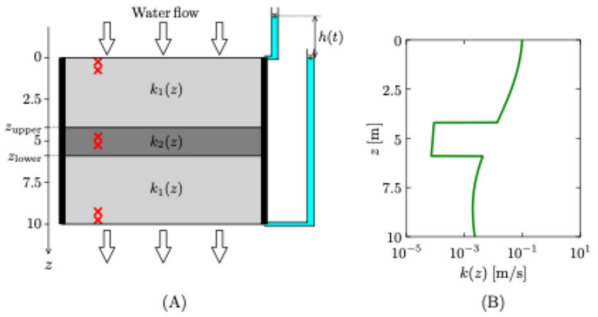

Figure 1. (A) Target domain consists of two subdomains $\mathcal{D}_1$ and $\mathcal{D}_2$. (B) True hydraulic conductivity field

5 NUMERICAL STUDY

We consider an inverse problem of estimating the spatial field of hydraulic conductivity for a one-dimensional seepage flow problem, as illustrated in Figure 1(A). The domain consists of two subdomains: $\mathcal{D}_1 = \{z: 0 \le z < 4.2, 5.9 < z \le 10\}$ and $\mathcal{D}_2 = \{z: 4.2 \le z \le 5.9\}$. The hydraulic conductivity field is shown in Figure 1(B). Let $k_i(z)$ and $u_i(z)$ denote the hydraulic conductivity and its logarithm, respectively, for subdomain $\mathcal{D}_i$ $(i = 1,2)$. The true fields $u_1(z)$ and $u_2(z)$ are generated as realizations from the KL expansions with $(\bar{u}_1, \sigma_1^2, l_1) = (-2, 1, 5)$ and $(\bar{u}_2, \sigma_2^2, l_2) = (-4, 1, 10)$, respectively. Dirichlet boundary conditions are imposed at the top and bottom of the domain. The hydraulic head at the top boundary is given by $h(t) = 0.1 + 0.01(t-1)$ $(t = 1, \ldots, 21)$, while the bottom boundary is kept fixed at zero. Observation data consist of the hydraulic heads at six points (marked as red crosses in Figure 1(A)) and the water inflow rate at the top under the 21 boundary conditions. Using these observations, we simultaneously estimate the domain interface parameters ${}^2\boldsymbol{\theta} = [z_{\text{upper}}, z_{\text{lower}}]^\top = [{}^2\theta_1, {}^2\theta_2]^\top$ and the spatial field of the logarithmic hydraulic conductivity $u_i(z)$ in each subdomain $\mathcal{D}_i$. Here, $\mathcal{D}_i$ is represented as follows:

$$\begin{aligned} \mathcal{D}_1({}^2\boldsymbol{\theta}) &= \{z: 0 \le z < {}^2\theta_1, {}^2\theta_2 < z \le 10\} \\ \mathcal{D}_2({}^2\boldsymbol{\theta}) &= \{z: {}^2\theta_1 \le z \le {}^2\theta_2\} \end{aligned}. \tag{12}$$

Both $u_1(z)$ and $u_2(z)$ are parameterized using KL expansions with ${}^1\boldsymbol{\theta}_1$ and ${}^1\boldsymbol{\theta}_2$, respectively, and the combined vector is denoted as ${}^1\boldsymbol{\theta} = [{}^1\boldsymbol{\theta}_1, {}^1\boldsymbol{\theta}_2]^\top$. To avoid realizations of a non-physical domain, constraints $1.05 < {}^2\theta_1 < 4.7$ and $5.3 < {}^2\theta_2 < 8.95$ are imposed on ${}^2\boldsymbol{\theta}$. Consequently, the prior distribution for ${}^2\boldsymbol{\theta}$ is modeled as a truncated Gaussian distribution.

The hyperparameters include the means $\bar{u}_i$, variances σ_i^2, and correlation lengths l_i of the KL expansions for $u_i(z)$ $(i = 1,2)$, as well as the mean vector $\boldsymbol{\mu}_2$ and covariance matrix $\boldsymbol{\Sigma}_2$ of the geometry parameters ${}^2\boldsymbol{\theta}$. In all numerical experiments, the values $(\bar{u}_1, \bar{u}_2, \sigma_1^2, \sigma_2^2) = (-2, -4, 1, 1)$ are used, matching those used to generate the true field. Other choices of hyperparameters are summarized in Table 1.

Table 1. Hyperparameter choices.

	l_1	l_2	$\boldsymbol{\mu}_2$	$\boldsymbol{\Sigma}_2$
Case 1	5.0	10.0	$[3.5, 6.5]$	$(10^{-5})^2 \cdot \mathbf{I}_2$
Case 2	5.0	10.0	$[2.5, 7.5]$	$0.4^2 \cdot \mathbf{I}_2$
Case 3	2.5	5.0	$[2.5, 7.5]$	$0.8^2 \cdot \mathbf{I}_2$
Case 4	10.0	10.0	$[4.2, 5.9]$	$0.1^2 \cdot \mathbf{I}_2$

Figure 2 shows results of hydraulic conductivity fields. In Case 1, with extremely small variances $\boldsymbol{\Sigma}_2 = (10^{-5})^2 \cdot \mathbf{I}_2$, effectively fixing the geometric parameters to the deterministic value ${}^2\boldsymbol{\theta} = \boldsymbol{\mu}_2 = [3.5, 6.5]^\top$. This corresponds to the conventional approach, where prior knowledge of the interface is used as a fixed input without estimating the interface locations. As shown in Figure 3, the posterior distribution $p({}^2\boldsymbol{\theta}|\mathbf{y})$ becomes sharply peaked, resembling a delta function. Since this prior knowledge is inaccurate (the true interface is at $(z_{\text{upper}}, z_{\text{lower}}) = (4.1, 5.8)$), the estimated 95% credible intervals (CIs) fail to include the true conductivity distribution, even though the proper correlation lengths $(l_1, l_2) = (5.0, 10.0)$ are used. The hydraulic conductivity is overestimated in $\mathcal{D}_1$ and $\mathcal{D}_2$, highlighting the adverse effect of using incorrect prior information as definitive values. In Case 2, while the correlation lengths are set correctly as $(l_1, l_2) = (5.0, 10.0)$, a relatively small covariance $\boldsymbol{\Sigma}_2 = 0.4^2 \cdot \mathbf{I}_2$ is used. Since approximately 95% of values from a Gaussian distribution $\mathcal{N}(m, s)$ lie within $m \pm 2s$, and the mean is set to $\boldsymbol{\mu}_2 = [2.5, 7.5]^\top$, this prior implies that the interface is likely within $1.7 \le {}^2\theta_1 \le 3.3$ and $6.7 \le {}^2\theta_2 \le 8.3$. These ranges are inconsistent with the true geometry. Nevertheless, the 95% CI captures the sharp transition near the interface (see Figure 2, Case 2), and the posterior $p({}^2\boldsymbol{\theta}|\mathbf{y})$ clearly identifies the correct interface positions (Figure 3). This improvement is attributed to the appropriate modeling of the spatial field via the KL expansion, which increases the dominance of the likelihood relative to the prior. However, if $\boldsymbol{\Sigma}_2$ is further reduced, the behavior becomes similar to Case 1, effectively fixing the geometry and deteriorating the estimation accuracy. In Case 3, shorter correlation lengths $(l_1, l_2) = (2.5, 5.0)$ than those used in the true field are adopted. As for the geometry parameters, $\boldsymbol{\mu}_2 = [2.5, 7.5]^\top$ and $\boldsymbol{\Sigma}_2 = 0.8^2 \cdot \mathbf{I}_2$ are set. Although the prior

implies the true interfaces lie within $0.9 \leq {}^2\theta_1 \leq 4.1$ and $5.9 \leq {}^2\theta_2 \leq 9.1$, which are slightly inconsistent with the true geometry, the 95% CIs still successfully capture both the conductivity field and the interface locations (Figure 2 and Figure 3). Compared to Case 2, the posterior exhibits larger variances. This is because smaller correlation lengths generate higher-frequency fields, which have greater expressiveness, resulting in many hydraulic conductivity fields that differ from the true field but produce a forward response that matches observations. In Case 4, a larger correlation length $l_1 = 10$ than the true value used to generate the true field is employed, while $\boldsymbol{\mu}_2 = [4.2,\ 5.9]^\mathsf{T}$ and $\boldsymbol{\Sigma}_2 = 0.1^2 \cdot \mathbf{I}_2$ are chosen. This implies prior knowledge that the true interfaces lie within $4.1 \leq {}^2\theta_1 \leq 4.3$ and $5.8 \leq {}^2\theta_2 \leq 6.0$, closely matching the actual values. As a result, the estimated interface locations are highly accurate (Figure 2 and Figure 3). However, the 95% CI for the hydraulic conductivity fails to include the true field, leading to poor estimation accuracy. This is because the overly large correlation length $l_1 = 10.0$ produces extremely smooth fields compared to the true spatial field, which has $l_1 = 5.0$.

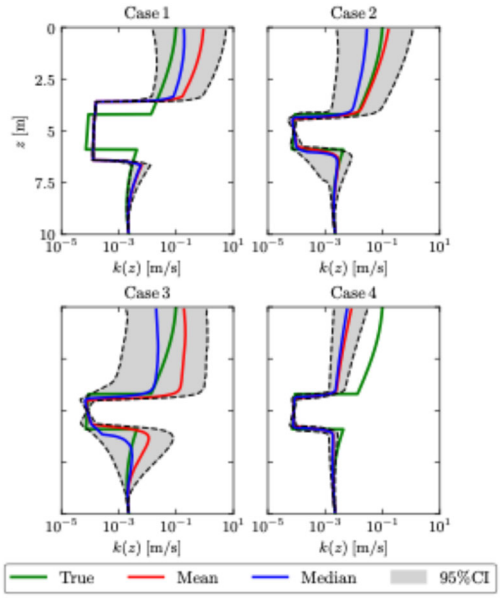

Figure 2. Estimated hydraulic conductivity fields for 4 different hyperparameter choices.

6 CONCLUSION

In this study, we investigated the influence of correlation length and prior information about geometric features on the results of simultaneous estimation by using a one-dimensional seepage flow problem. It was demonstrated that when using a correlation length that is equal to or smaller than the true value, the interface location can be accurately estimated, even though the prior information about the geometry is somewhat inaccurate. Conversely, when a correlation length larger than the true value is used, the estimated hydraulic conductivity field becomes overly smooth, making accurate estimation difficult, even with precise prior information about the geometric features. These results suggest that in simultaneous estimation, the correlation length, a hyperparameter of the spatial field, has a greater influence on the accuracy of the estimation than the hyperparameters of the geometry parameters. Therefore, a future direction for improving the simultaneous estimation approach would be to extend it into a hierarchical Bayesian framework that enables the systematic estimation of an appropriate correlation length.

Moreover, the finding that reasonably accurate estimations can be achieved even with inaccurate prior information about geometric features supports the proactive use of prior information. If the prior information is accurate, it can improve the estimation accuracy; and even if it is not, there is no strong reason to avoid using it as long as acceptable accuracy can still be achieved. However, since the influence of prior information may vary depending on factors such as the type of inverse problem and the available observation data, further detailed investigations across a broader range of problems are necessary to guide the appropriate selection of hyperparameters.

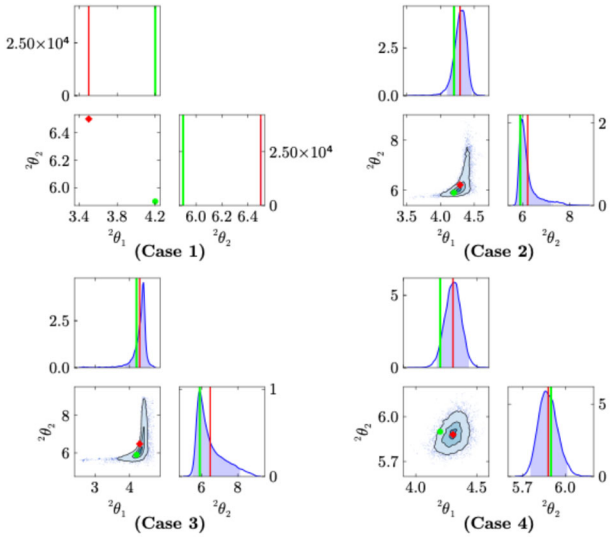

Figure 3. Posterior distributions $p({}^2\boldsymbol{\theta}|\mathbf{y})$. (The red diamond and line represent the mean. The green circle and line represent the true values).

7 ACKNOWLEDGEMENTS

This work was supported by JSPS KAKENHI Grant Numbers JP22K18352 and JP22K20601.

8 REFERENCES

Ghanem, R.G. & Spanos, P.D., 2003. *Stochastic finite elements: a spectral approach*. North Chelmsford, MA: Courier Corporation.

Koch, M.C., Fujisawa, K. & Murakami, A., 2020. Adjoint Hamiltonian Monte Carlo algorithm for the estimation of elastic modulus through the inversion of elastic wave propagation data. *International Journal for Numerical Methods in Engineering*, 121(6), pp.1037–1067.

Koch, M.C., Osugi, M., Fujisawa, K. & Murakami, A., 2021. Hamiltonian Monte Carlo for Simultaneous Interface and Spatial Field Detection (HMCSISFD) and its application to a piping zone interface detection problem. *International Journal for Numerical and Analytical Methods in Geomechanics*, 45(17), pp.2602–2626.

Neal, R.M., 2011. MCMC using Hamiltonian dynamics. In: S. Brooks, A. Gelman, G.L. Jones & X.-L. Meng, eds. *Handbook of Markov Chain Monte Carlo*. Boca Raton, FL: Chapman and Hall/CRC, pp.113–162.

Shibata, T., Koch, M.C., Papaioannou, I. & Fujisawa, K., 2025. Efficient Bayesian inversion for simultaneous estimation of geometry and spatial field using the Karhunen–Loève expansion. *Computer Methods in Applied Mechanics and Engineering*, 441, p.117960.

Uribe, F., Papaioannou, I., Betz, W. & Straub, D., 2020. Bayesian inference of random fields represented with the Karhunen–Loève expansion. *Computer Methods in Applied Mechanics and Engineering*, 358, p.112632.

Yang, H.-Q., Zhang, L., Xue, J., Zhang, J. & Li, X., 2019. Unsaturated soil slope characterization with Karhunen–Loève and polynomial chaos via Bayesian approach. *Engineering with Computers*, 35, pp.337–350.