Mathematical Geosciences (2024) 56:333–360
https://doi.org/10.1007/s11004-023-10084-8

Joint Clarification of Contaminant Plume and Hydraulic Transmissivity via a Geostatistical Approach Using Hydraulic Head and Contaminant Concentration Data

Shizuka Takai[1,2] · Taro Shimada[1] · Seiji Takeda[1] · Katsuaki Koike[2]

Received: 10 November 2022 / Accepted: 25 June 2023 / Published online: 31 July 2023

Abstract

To enable proper remediation of accidental groundwater contamination, the contaminant plume evolution needs to be accurately estimated. In the estimation, uncertainties in both the contaminant source and hydrogeological structure should be considered, especially the temporal release history and hydraulic transmissivity. Although the release history can be estimated using geostatistical approaches, previous studies use the deterministic hydraulic property field. Geostatistical approaches can also effectively estimate an unknown heterogeneous transmissivity field via the use of joint data, such as a combination of hydraulic head and tracer data. However, tracer tests implemented over a contaminated area necessarily disturb the in situ condition of the contamination. Conversely, measurements of the transient concentration data over an area are possible and can preserve the conditions. Accordingly, this study develops a geostatistical method for the joint clarification of contaminant plume and transmissivity distributions using both head and contaminant concentration data. The applicability and effectiveness of the proposed method are demonstrated through two numerical experiments assuming a two-dimensional heterogeneous confined aquifer. The use of contaminant concentration data is key to accurate estimation of the transmissivity. The accuracy of the proposed method using both head and concentration data was verified achieving a high linear correlation coefficient of 0.97 between the true and estimated concentrations for both experiments, which was 0.67 or more than the results using only the head data. Furthermore, the uncertainty of the contaminant plume evolution was successfully evaluated by considering the uncertainties of both the initial plume and the transmissivity distributions, based on their conditional realizations.

✉ Shizuka Takai
takai.shizuka@jaea.go.jp

1 Nuclear Safety Research Center, Japan Atomic Energy Agency, 2-4 Shirakata, Tokai, Ibaraki 319-1195, Japan

2 Department of Urban Management, Graduate School of Engineering, Kyoto University, Katsura, Kyoto 615-8540, Japan

Keywords Conditional realization · Contaminant plume distribution · Geostatistical inversion · Hydraulic transmissivity · Water-soluble contaminant

1 Introduction

Groundwater contamination is a serious environmental issue that has been occurring in various areas worldwide for many years. Contamination is caused by various contaminants such as anthropogenic chemicals and radioactive and microbiological substances. Contamination by soluble and mobile contaminants tends to spread inconspicuously into extensive areas because its extension is not noticed unless the water quality is systematically monitored. To suppress the extension of contamination and form a proper remediation plan via pumping and chemical processing, correct estimations of the contaminant plume evolution and its uncertainty given the water quality data at wells are essential. To achieve this, the construction of an accurate hydrogeological model to simulate the groundwater flow and contaminant transport is indispensable.

A contaminant source (i.e., the source location or release history) and hydraulic properties are two predominant uncertain factors in the estimation of the contaminant plume distribution. Uncertainty in the contaminant source arises in accidental contamination events, because the source is not identified or recorded except laboratory experiments. In real-world events, release history records are particularly rare, as in past cases of underground contamination at nuclear facilities (OECD 2014); although source locations have been detected by preliminary surveys or historical site assessments in drain lines, sumps, pipes, and tanks, the temporal release histories have not been traced in all cases. Among the relevant hydraulic properties, the permeability expressed by the hydraulic conductivity or transmissivity is the most essential parameter for simulating groundwater flow and contaminant transport. Despite the significance and heterogeneity of the spatial distribution, the amount and location of the measured permeability data at wells are usually limited by practical constraints such as time and cost. Therefore, there are large uncertainties in permeability distributions estimated by simply interpolating and extrapolating the measured permeability or indirectly using hydraulic head or other data. The correctness and uncertainty of the estimated permeability distribution necessarily control the estimation accuracy of the contaminant plume evolution.

The joint identification of the contaminant release history and hydraulic properties has been studied using several approaches. Examples of hydraulic property identification approaches include non-linear maximum likelihood estimation (Wagner 1992) and trained artificial neutral network (Singh and Datta 2004) for homogeneous fields and restart normal-score ensemble Kalman filter (EnKF) (Sanchez-Leon et al. 2016; Chen et al. 2018; Xu and Gomez-Hernandez 2018) and ensemble smoother (ES) with multiple data assimilation (Xu et al. 2021) for heterogeneous fields. The high accuracy of EnKF methods has been confirmed in a sandbox experiment (Chen et al. 2021). However, there are two problems with the above studies:

1. The release patterns of contaminants are assumed to be known. Typically, the contaminant source parameters, such as the source location, initial release time, release duration, and mass-loading rate, are determined by assuming constant release (e.g., Xu et al. 2021). However, the release pattern in actual contamination events is usually unknown and uncertain. Therefore, a random function should be applied to the release pattern (Snodgrass and Kitanidis 1997).
2. Although EnKF methods have the advantage of enabling production of non-Gaussian distributions without considering the spatial correlation of the hydraulic conductivity, they require an impractically large amount of spatiotemporal measurement data of the head and concentration for usual contaminant cases.

To overcome the first issue, the quasi-linear geostatistical approach (GA; e.g., Snodgrass and Kitanidis 1997; Gyzl et al. 2004; Shlomi and Michalak 2007) is applicable by introducing the prior information of release history with geostatistical trend and covariance. The applicability of GA has been verified at real-world sites contaminated by water-soluble contaminants: 1,4-dioxane (Woodbury et al. 1998; Michalak and Kitanidis 2002), tetrachloroethene and trichloroethne (Michalak and Kitanidis 2003), and hexachlorocyclohexane (Gyzl et al. 2014). However, most of these studies only incorporated the uncertainty of the contaminant source using a deterministic hydraulic property model.

GA can also solve the second issue if extended to hydraulic tomography studies (e.g., Li et al. 2007, 2008; Cardiff et al. 2009; Cardiff and Barrash 2011; Pouladi et al. 2021), and its applicability using head data has been verified by field tests (e.g., Illuman et al. 2009; Wang et al. 2017; Zha et al. 2018; Luo et al. 2022). However, GA tends to generate a spatially smoother best estimate than the true distribution, which is its main drawback. This smoothing effect is caused by modeling the hydraulic conductivity as a multivariate Gaussian, which is usually inadequate for the estimation of heterogeneous fields such as aquifers in fluvial deposits, where several strata with highly different permeabilities coexist (Mo et al. 2020). However, the assumption of a Gaussian field is applicable to cases of groundwater contamination that occur in a single aquifer. The smoothing effect has been improved via joint inversion of the head and temperature data (Jiang and Woodbury 2006) and the head and tracer data (e.g., Harvey and Gorelick 1995; Cirpka and Kitanidis 2000; Xu and Kitanidis 2014), as well as in combination with a convolution neural network (Vu and Jardani 2022).

Although the tracer test data can indeed improve the performance and accuracy of GA, the implementation of many tests over a contaminated area necessarily disturbs the contamination situation, renders situation assessments difficult, and possibly further extends the contamination. In contrast to such impractical testing, measuring the transient concentration data in groundwater at wells over an area is possible and preserves the situation. Therefore, through the joint use of head and transient concentration data, the estimation accuracies of both the contaminant plume distribution and the hydraulic conductivity are expected to be effectively improved. To achieve this, an estimation of the unknown initial plume distribution is indispensable.

Given the above background, this study aims to accurately estimate the contaminant plume evolution by considering uncertainties in both the temporal release history and the heterogeneous transmissivity fields. Accordingly, the GA method is further

developed for a joint clarification of the contaminant plume and transmissivity distributions using both the head and contaminant concentration data. The joint clarification is achieved by combining previous estimation methods for each component: a contaminant plume with an unknown release history is estimated using the method of Shlomi and Michalak (2007) and the hydraulic transmissivity is estimated using the method of Kitanidis and Lee (2014). This paper begins with a review of the previous estimation methods and then, proposes a combined method. This method consists of the following three steps: separate initial estimations of the transmissivity and the initial plume distributions using the head and concentration data, respectively; an iterative update of their distributions via joint use of the data; and an estimation of the contaminant plume evolution and its uncertainties based on their conditional realizations. The proposed method is verified by two numerical experiments assuming groundwater contamination in a two-dimensional aquifer and the results are discussed finally.

2 Methods

2.1 Iterative Estimation of Contaminant Plume and Hydraulic Transmissivity

Previous geostatistical approaches for contaminant plume estimation with unknown release histories (e.g., Shlomi and Michalak 2007) cannot consider the uncertainty of the hydraulic transmissivity. To address this problem, this study developed a GA method to estimate the contaminant plume evolution $\boldsymbol{z}(\boldsymbol{x},t)$ ($\boldsymbol{x}$: space and t: time) and its uncertainty by combining previous estimation methods for the contaminant plume and transmissivity, reviewed in Sects. 2.2 and 2.3, for the uncertainties of the release history $\boldsymbol{s}(t)$ and log-transmissivity $\boldsymbol{r}(\boldsymbol{x})$, respectively. Because both $\boldsymbol{z}$ and $\boldsymbol{r}$ are necessary for each estimation, an iterative approach using both the head and concentration data is proposed as shown in Fig. 1.

The first step is the initial estimation of $\boldsymbol{r}$ and the initial contaminant plume $\boldsymbol{z}_0 = \boldsymbol{z}(t_0)$ (t_0: initial measurement time), using the head $\boldsymbol{\varphi}$ and initial concentration data $\boldsymbol{z}_0^*$, separately. The next step is to update $\boldsymbol{r}$ based on the estimated $\boldsymbol{z}_0$ using both $\boldsymbol{\varphi}$ and the transient concentration data $\boldsymbol{z}^*(t)$. The posterior pdfs of $\boldsymbol{r}$ and $\boldsymbol{z}_0$ are iteratively calculated until the posterior pdf of $\boldsymbol{r}$ reaches its maximum. In this step, the mutual uncertainties of $\boldsymbol{r}$ and $\boldsymbol{z}_0$ are not considered (i.e., the uncertainty of $\boldsymbol{r}$ is not considered in the $\boldsymbol{z}_0$ estimation, and vice versa). Finally, the best estimate of $\boldsymbol{z}(t)$ is obtained using the best estimates of $\boldsymbol{r}$ and the corresponding $\boldsymbol{z}_0$. To consider both the uncertainties of $\boldsymbol{r}$ and $\boldsymbol{z}_0$, the estimation method for the uncertainty of $\boldsymbol{z}(t)$ based on $N_r \times N_{z0}$ conditional realizations of $\boldsymbol{r}$ and $\boldsymbol{z}_0$ (N: number of realizations) is developed as described in Sect. 2.4.

2.2 Geostatistical Inversion for Initial Contaminant Plume Estimation

This section reviews preceding studies of the quasi-linear GA for the estimation of the contaminant plume distribution from a known source with an unknown release history (e.g., Kitanidis 1995; Snodgrass and Kitanidis 1997; Shlomi and Michalak

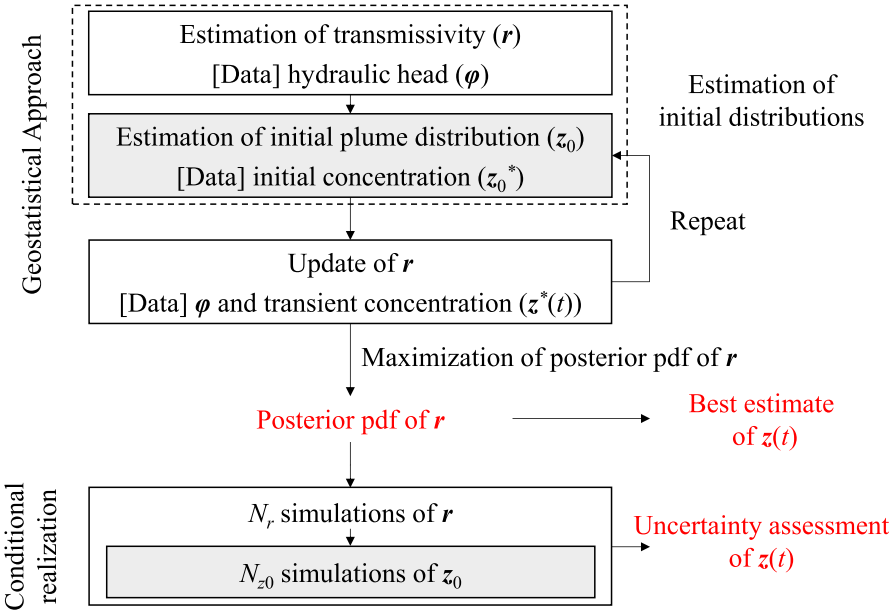

Fig. 1 Flowchart of the iterative estimation of the contaminant plume and the hydraulic transmissivity distributions, $\boldsymbol{z}(\boldsymbol{x},t)$ and $\boldsymbol{r}(\boldsymbol{x})$, respectively

2007). Under a steady state flow, $\boldsymbol{z}_0^* \in \mathbb{R}^{n_z \times 1}$ (n_z-dimensional real space) is related linearly to the release history $\boldsymbol{s} \in \mathbb{R}^{m_t \times 1}$ at each time $t_j (j = 1, \ldots, m_t)$ such that

$$\boldsymbol{z}_0^* = \boldsymbol{H}_s^* s + \boldsymbol{v}_z, \tag{1}$$

where $\boldsymbol{H}_s^* \in \mathbb{R}^{n_z \times m_t}$ and $\boldsymbol{v}_z \in \mathbb{R}^{n_z \times 1}$ stand for the Jacobian matrix and the model mismatch error at the measurement points, respectively. $\boldsymbol{H}_s^*$ expresses the sensitivity of the concentrations at each measurement point and time and can be calculated in advance by a flow and transport simulation for the release of a unit concentration pulse. Therefore, the unknown $\boldsymbol{s}$ can be obtained by solving Eq. (1) inversely.

The geostatistical inversion incorporates the temporal correlation of $\boldsymbol{s}$ and assumes that $\boldsymbol{s}$ and $\boldsymbol{v}_z$ are random vectors following the multivariate Gaussian distributions $\boldsymbol{s} \sim N\left(\boldsymbol{X}_s \boldsymbol{\beta}_s, \boldsymbol{Q}_s(\theta_s)\right)$ and $\boldsymbol{v}_z \sim N(\boldsymbol{0}, \boldsymbol{R}_z)$, where $\boldsymbol{X}_s \in \mathbb{R}^{m_t \times p_s}$ is a known matrix of basis functions; $\boldsymbol{\beta}_s \in \mathbb{R}^{p_s \times 1}$ are p_s unknown drift coefficients; $\boldsymbol{Q}_s(\theta_s) \in \mathbb{R}^{m_t \times m_t}$ is the generalized covariance matrix of $\boldsymbol{s}$; θ_s is the structural parameter of $\boldsymbol{Q}_s$; and $\boldsymbol{R}_z$ is the error covariance matrix of $\boldsymbol{z}_0^*$. This study assumes an uncorrelated error of $\boldsymbol{R}_z = \sigma_{R_z}^2 \boldsymbol{I}$, where $\sigma_{R_z}^2$ is the variance of the error and $\boldsymbol{I} \in \mathbb{R}^{n_z \times n_z}$ is the identity matrix. The unknown $\boldsymbol{s}$ can be estimated from $\boldsymbol{z}_0^*$ by maximizing the posterior pdf $p''\left(\boldsymbol{s}, \boldsymbol{\beta}_s\right)$ obtained via Bayes' rule as

$$-\ln p''\left(\boldsymbol{s}, \boldsymbol{\beta}_s\right) = \tfrac{1}{2}\left(\boldsymbol{z}_0^* - \boldsymbol{H}_s^* \boldsymbol{s}\right)^T \boldsymbol{R}_z^{-1}\left(\boldsymbol{z}_0^* - \boldsymbol{H}_s^* \boldsymbol{s}\right) + \tfrac{1}{2}\left(\boldsymbol{s} - \boldsymbol{X}_s \boldsymbol{\beta}_s\right)^T \boldsymbol{Q}_s^{-1}\left(\boldsymbol{s} - \boldsymbol{X}_s \boldsymbol{\beta}_s\right). \tag{2}$$

The structural parameters $\boldsymbol{\theta} = \left(\theta_s, \sigma_{R_z}\right)^T$ can be iteratively estimated using a restricted maximum likelihood approach that minimizes the objective function $L(\boldsymbol{\theta})$ (Kitanidis 1995)

$$L(\boldsymbol{\theta}) = \tfrac{1}{2}\ln|\boldsymbol{\Sigma}_s| + \tfrac{1}{2}\ln\left|\boldsymbol{X}_s^T\boldsymbol{H}_s^{*T}\boldsymbol{\Sigma}_s^{-1}\boldsymbol{H}_s^*\boldsymbol{X}_s\right| + \tfrac{1}{2}\boldsymbol{z}_0^{*T}\boldsymbol{\Xi}_s^{-1}\boldsymbol{z}_0^*, \tag{3}$$

$$\boldsymbol{\Sigma}_s = \boldsymbol{H}_s^*\boldsymbol{Q}_s\boldsymbol{H}_s^{*T} + \boldsymbol{R}_z, \tag{4}$$

$$\boldsymbol{\Xi}_s = \boldsymbol{\Sigma}_s^{-1} - \boldsymbol{\Sigma}_s^{-1}\boldsymbol{H}_s^*\boldsymbol{X}_s\left(\boldsymbol{X}_s^T\boldsymbol{H}_s^{*T}\boldsymbol{\Sigma}_s^{-1}\boldsymbol{H}_s^*\boldsymbol{X}_s\right)^{-1}\boldsymbol{X}_s^T\boldsymbol{H}_s^{*T}\boldsymbol{\Sigma}_s^{-1}. \tag{5}$$

Then, the best estimate $\hat{\boldsymbol{s}}$ and its posterior covariance $\boldsymbol{V}_{\hat{s}}$ are derived by solving the following equation system

$$\begin{pmatrix} \boldsymbol{\Sigma}_s & \boldsymbol{H}_s^*\boldsymbol{X}_s \\ \left(\boldsymbol{H}_s^*\boldsymbol{X}_s\right)^T & 0 \end{pmatrix}\begin{pmatrix} \boldsymbol{\Lambda}_s^T \\ \boldsymbol{M}_s \end{pmatrix} = \begin{pmatrix} \boldsymbol{H}_s^*\boldsymbol{Q}_s \\ \boldsymbol{X}_s^T \end{pmatrix}, \tag{6}$$

$$\hat{\boldsymbol{s}} = \boldsymbol{\Lambda}_s\boldsymbol{z}_0^*, \tag{7}$$

$$\boldsymbol{V}_{\hat{s}} = \boldsymbol{Q}_s - \boldsymbol{Q}_s\boldsymbol{H}_s^{*T}\boldsymbol{\Lambda}_s^T - \boldsymbol{X}_s\boldsymbol{M}_s, \tag{8}$$

where $\boldsymbol{\Lambda}_s \in \mathbb{R}^{m_t \times n_z}$ and $\boldsymbol{M}_s \in \mathbb{R}^{p_s \times m_t}$ are the weight matrix and the Lagrange multiplier, respectively. To enforce concentration non-negativity, a power transformation (Box and Cox 1964) is applied, such that

$$\tilde{\boldsymbol{s}} = \alpha\left(\boldsymbol{s}^{1/\alpha} - 1\right), \tag{9}$$

where α is a positive number. Because Eq. (1) is not linear in the transformed space, $\tilde{\boldsymbol{s}}$ and $\boldsymbol{\theta}$ are solved iteratively using the quasi-linear approach (Snodgrass and Kitanidis 1997) in which α is chosen to be as small as possible while ensuring that $\tilde{\boldsymbol{s}} > \alpha$. After obtaining the best estimate and its covariance, the solutions are back-transformed into the original space by

$$\boldsymbol{s} = \left(\frac{\tilde{\boldsymbol{s}}+\alpha}{\alpha}\right)^{\alpha}. \tag{10}$$

Once $\hat{\boldsymbol{s}}$ and $\boldsymbol{V}_{\hat{s}}$ are determined, the best estimate of $\widehat{\boldsymbol{z}_0} \in \mathbb{R}^{m\times 1}$ and its posterior covariance $\boldsymbol{V}_{\hat{z}}$ can be solved as

$$\widehat{\boldsymbol{z}_0} = \boldsymbol{H}_s\hat{\boldsymbol{s}}, \tag{11}$$

$$\boldsymbol{V}_{\hat{z}} = \boldsymbol{H}_s\boldsymbol{V}_{\hat{s}}\boldsymbol{H}_s^T, \tag{12}$$

where m is the number of estimation points and $\boldsymbol{H}_s \in \mathbb{R}^{m\times m_t}$ is the Jacobian matrix at all estimation points.

2.3 Principle Component Geostatistical Approach for Hydraulic Transmissivity Estimation

This study adopts the principal component geostatistical approach (PCGA: Kitanidis and Lee 2014; Lee and Kitanidis 2014), as reviewed below, to estimate the hydraulic transmissivity distribution. The observation $\boldsymbol{y} \in \mathbb{R}^{n\times 1}$ can be expressed by the forward model h with $\boldsymbol{r} \in \mathbb{R}^{m\times 1}$ and the observation error $\boldsymbol{v} \in \mathbb{R}^{n\times 1}$ as

$$\boldsymbol{y} = h(\boldsymbol{r}) + \boldsymbol{v}. \tag{13}$$

For the present case, $\boldsymbol{y}$ corresponds to only head data or to head and concentration data. $\boldsymbol{r}$ and $\boldsymbol{v}$ are assumed to follow the multivariate Gaussian distributions $\boldsymbol{r} \sim N(\boldsymbol{X\beta}, \boldsymbol{Q}(\theta_r))$ and $\boldsymbol{v} \sim N(0, \boldsymbol{R})$, where $\boldsymbol{X} \in \mathbb{R}^{m\times p}$ is a known matrix of basis functions; $\boldsymbol{\beta} \in \mathbb{R}^{p\times 1}$ represents p unknown drift coefficients; $\boldsymbol{Q}(\theta_r) \in \mathbb{R}^{m\times m}$ is a generalized covariance matrix of $\boldsymbol{r}$; θ_r is the structural parameter of $\boldsymbol{Q}$; and $\boldsymbol{R}$ is the error covariance matrix of $\boldsymbol{y}$. As in the above release history, the best estimate $\hat{\boldsymbol{r}}$ is obtained by maximizing $p''(\boldsymbol{r}, \boldsymbol{\beta})$. Because Eq. (13) is not linear, the quasi-linear approach (Kitanidis 1995) is applied to approximate the true $\hat{\boldsymbol{r}}$ with the latest estimate $\overline{\boldsymbol{r}}$, such that

$$h(\hat{\boldsymbol{r}}) = h(\overline{\boldsymbol{r}}) + \boldsymbol{H}(\hat{\boldsymbol{r}} - \overline{\boldsymbol{r}}), \boldsymbol{H} = \left.\frac{\partial h}{\partial \boldsymbol{r}}\right|_{\boldsymbol{r}=\overline{\boldsymbol{r}}}. \tag{14}$$

The following equation system is solved to update $\overline{\boldsymbol{r}}$ until it converges

$$\begin{pmatrix} \boldsymbol{\Sigma} & \boldsymbol{HX} \\ (\boldsymbol{HX})^T & 0 \end{pmatrix} \begin{pmatrix} \overline{\boldsymbol{\xi}} \\ \overline{\boldsymbol{\beta}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{y} - h(\overline{\boldsymbol{r}}) + \boldsymbol{H}\overline{\boldsymbol{r}} \\ 0 \end{pmatrix}, \tag{15}$$

$$\boldsymbol{\Sigma} = \boldsymbol{HQH}^T + \boldsymbol{R}, \tag{16}$$

$$\overline{\boldsymbol{r}} = \boldsymbol{X}\overline{\boldsymbol{\beta}} + \boldsymbol{QH}^T\overline{\boldsymbol{\xi}}. \tag{17}$$

Once the optimal solution $\hat{\boldsymbol{r}}$ is obtained, its posterior covariance $\boldsymbol{V}_{\hat{r}}$ can be calculated in the same way as in Eqs. (6) and (8).

PCGA was proposed to obtain $\hat{\boldsymbol{r}}$ efficiently with small computation cost by improving the conventional GA through two approaches. The first approach is the use of Taylor expansion for the indirect expression of $\boldsymbol{H}$

$$\boldsymbol{Ha} \approx \frac{1}{\delta_r}[h(\boldsymbol{a} + \delta_r\boldsymbol{a}) - h(\boldsymbol{a})], \tag{18}$$

where $\boldsymbol{a}$ is the target vector, such as $\overline{\boldsymbol{r}}$, and δ_r is a finite difference interval that can be optimized as (Lee et al. 2016)

$$\widehat{\delta_r} = \frac{\sqrt{\varepsilon_r}}{\boldsymbol{a}_2^2} \max\left(\left|\boldsymbol{r}^T\boldsymbol{a}\right|, |\boldsymbol{r}|^T|\boldsymbol{a}|\right)\text{sign}\left(\boldsymbol{r}^T\boldsymbol{a}\right), \tag{19}$$

where ε_r is the relative machine precision depending on the precision of the forward model; $|\boldsymbol{a}| = (|a_1|, \ldots, |a_m|)^T$; and sign() indicates the sign of a value. The second approach is a low-rank approximation of $\boldsymbol{Q}$ as

$$\boldsymbol{Q} \approx \boldsymbol{Z}_Q \boldsymbol{Z}_Q^T = \sum_{i=1}^{K} \boldsymbol{\zeta}_i \boldsymbol{\zeta}_i^T, \boldsymbol{\zeta}_i = \sqrt{\lambda_i} \boldsymbol{V}_i, \tag{20}$$

where λ_i and $\boldsymbol{V}_i \in \mathbb{R}^{m \times 1}$ are the ith eigenvalue and eigenvector of $\boldsymbol{Q}$ in the descending order. The order K can be defined such that the relative error of the low-rank approximation, λ_{K+1}/λ_1, is sufficiently small. All of the above calculations are implemented by normalizing the drift and covariance of the prior model, as explained in the Appendix, following Kitanidis and Lee (2014).

2.4 Conditional Realizations of the Transmissivity and Initial Plume Distributions

The uncertainty of $z(\boldsymbol{x},t)$ is assessed considering the uncertainties of both the initial contaminant plume and the transmissivity distributions by generating their conditional realizations. The conditional realization can be drawn from the posterior pdf using either the Cholesky decomposition of the posterior covariance (Harvey and Gorelick 1995; Nowak 2009; Troldborg et al. 2012) or the parametric bootstrapping sampling method (Kitanidis 1995; Kitanidis and Lee 2014). Because of its simplicity and smallness of calculation, the Cholesky approach is adopted here where the ith conditional realization of the transmissivity distribution $\widehat{\boldsymbol{r}_{ci}}$ is

$$\widehat{\boldsymbol{r}_{ci}} = \widehat{\boldsymbol{r}_{ui}} + (\boldsymbol{H}\,\boldsymbol{Q})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{y} - h(\widehat{\boldsymbol{r}_{ui}}) + \boldsymbol{v}_i), \tag{21}$$

where $\widehat{\boldsymbol{r}_{ui}} \sim N(\hat{\boldsymbol{r}}, V_{\hat{r}})$ is the ith unconditional realization of the transmissivity randomly sampled from the posterior pdf and $\boldsymbol{v}_i \sim N(0, \boldsymbol{R})$ is the ith random measurement error of the head and concentration. In the same way, the ith realization of the initial contaminant plume distribution $\widehat{z_{0c_i}}$ can be written as

$$\widehat{z_{0c_i}} = \boldsymbol{H}_s \widehat{\boldsymbol{s}_{ci}}, \widehat{\boldsymbol{s}_{ci}} = \widehat{\boldsymbol{s}_{ui}} + \left(\boldsymbol{H}_s^* \boldsymbol{Q}_s\right)^T \boldsymbol{\Sigma}_s^{-1} \left(\boldsymbol{z}_0^* - \boldsymbol{H}_s^* \widehat{\boldsymbol{s}_{ui}} + \boldsymbol{v}_{z_i}\right), \tag{22}$$

where $\widehat{\boldsymbol{s}_{ui}}$ is the ith unconditional realization of the release history and $\boldsymbol{v}_{z_i} \sim N(0, \boldsymbol{R}_z)$ is the ith random measurement error of the initial concentration. $\widehat{\boldsymbol{s}_{ui}}$ can be inversely calculated from the ith realization of the release history in the transformed space, $\widehat{\tilde{\boldsymbol{s}}_{ui}} \sim N\left(\widehat{\tilde{\boldsymbol{s}}}, V_{\widehat{\tilde{s}}}\right)$.

3 Numerical Experiment

3.1 Physical Model

The above proposed geostatistical approach was tested via numerical experiments of two-dimensional steady state groundwater flow and contaminant transport in the

steady state. Let a transmissivity field T in a confined aquifer be spatially variable but locally isotropic. The governing equation for groundwater flow in a saturated porous media is expressed as

$$\nabla \cdot \boldsymbol{u} = Q_f \delta(\boldsymbol{x} - \boldsymbol{x}_f), \quad \boldsymbol{u} = -T\nabla\varphi, \tag{23}$$

where $\boldsymbol{u}$ is the groundwater flow velocity; T is the transmissivity; φ is the hydraulic head; $\delta(\boldsymbol{x})$ is the Dirac delta function; and Q_f is the pumping rate at a well location $\boldsymbol{x}_f$. Under this state, the contaminant transport is expressed by the advection–dispersion equation as

$$\nabla \cdot (\boldsymbol{D}\nabla c - \boldsymbol{V}c) = R_f \frac{\partial c}{\partial t} + \lambda_f c, \tag{24}$$

where c is the dimensionless concentration; $\boldsymbol{V} = \boldsymbol{u}/\varepsilon$ is the actual groundwater velocity; $\boldsymbol{D}$ is the dispersion tensor; ε is the porosity; R_f is the retardation factor; and λ_f is the radioactive or first-order biochemical decay constant. Each component of $\boldsymbol{D}$ is formulated as

$$D_{ij} = \alpha_T |\boldsymbol{V}| \delta_{ij} + (\alpha_L - \alpha_T) \frac{V_i V_j}{|\boldsymbol{V}|} + D_m \tau \delta_{ij} (i, j = 1, 2), \tag{25}$$

where α_L and α_T are the longitudinal and transverse dispersivities, respectively; D_m is the molecular diffusion coefficient; and τ is the tortuosity.

3.2 Settings of Two Cases

On the basis of reviews of groundwater contamination events caused by water-soluble pollutants, the contamination extent generally ranges from scales of 100 m to 1 km (e.g., for over 2000 sites in California, the median plume length was 270 m for 1,4-dioxane, 115 m for 1,1,1-trichloroethne, 95 m for trichloroethene, and 123 m for 1,1-dichloroethene; Adamson et al. 2014). At such a scale, contaminated water can be pumped from wells set at several locations. To simplify the present experiment, only one pumping well was set at $(x, y) = (25, 0)$ m in a model domain of 100 m along the x-axis (the flow direction) $\times$ 50 m along the y-axis (Fig. 2).

Two cases of transmissivity fields with different degrees of heterogeneity were prepared by referring to the experimental model of Lee and Kitanidis (2014): case 1 had a smooth spatial change in the transmissivity and case 2 had a highly heterogeneous field with local changes in the transmissivity. The mean log-transmissivities (m^2/d) of 2.4 were the same for both cases; this is the product of the assigned aquifer thickness, 10 m, with a typical hydraulic conductivity for porous sand, 10^{-5} m/s (Zanini and Kitanidis 2009). The difference between the two cases is expressed by the spatial correlation range of the field, case 1 is long and case 2 is short, as shown by the covariance function in Table 1. The covariance functions defined were a generalized cubic covariance with a linear drift following Zanini and Kitanidis (2009), which is continuously differentiable and smooth (Kitanidis and Lee 2014), for case 1 and an isotropic exponential covariance with a constant drift for case 2. Constant-head

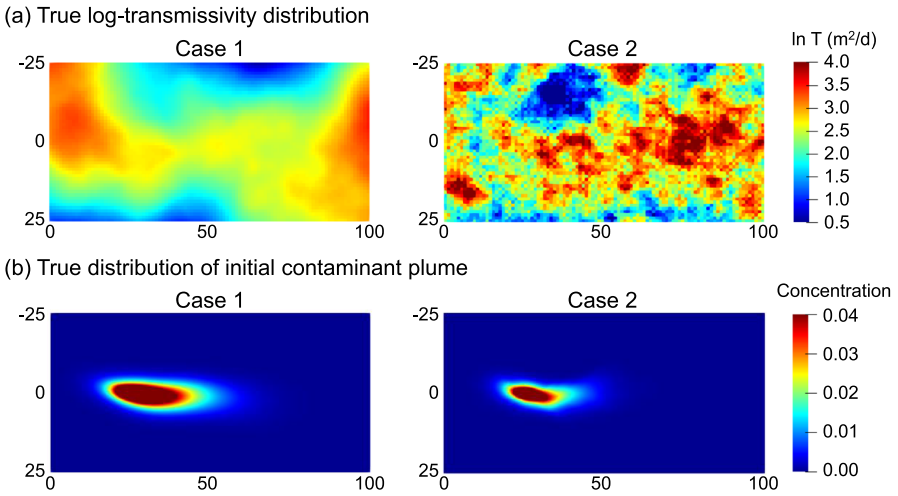

Fig. 2 True value distributions of (**a**) the log-transmissivity and (**b**) the initial contaminant plume for cases 1 (left) and 2 (right)

Table 1 Parameter settings for the two cases with smooth and heterogeneous transmissivity fields (case 1 and 2, respectively)

	Case 1	Case 2
Geometric parameters		
Domain length and width (m)	100, 50	
Cell size (m)	1	
Pumping rates (m^2/d)	25	
Porosity (–)	0.25	
Dispersivity (m)	5.0 (longitudinal), 0.5 (transverse)	
Diffusion coefficient (m^2/s)	1×10^{-9}	
Geostatistical parameters		
Mean log-transmissivity (ln m^2/d)	2.4	
Covariance kernel	$\theta_r \lvert \boldsymbol{x} - \boldsymbol{x}' \rvert^3$, $\theta_r = 8.0 \times 10^{-6}$	$\exp(-\lvert \boldsymbol{x} - \boldsymbol{x}' \rvert / \theta_r)$, $\theta_r = 15$
Drift	$[\mathbf{1} \times \boldsymbol{y}]$	$[\mathbf{1}]$
Measurement error (standard deviation)		
Head (m)	0.05	
Concentration (–) or travel time (month)	10%	

"–" stands for unitless

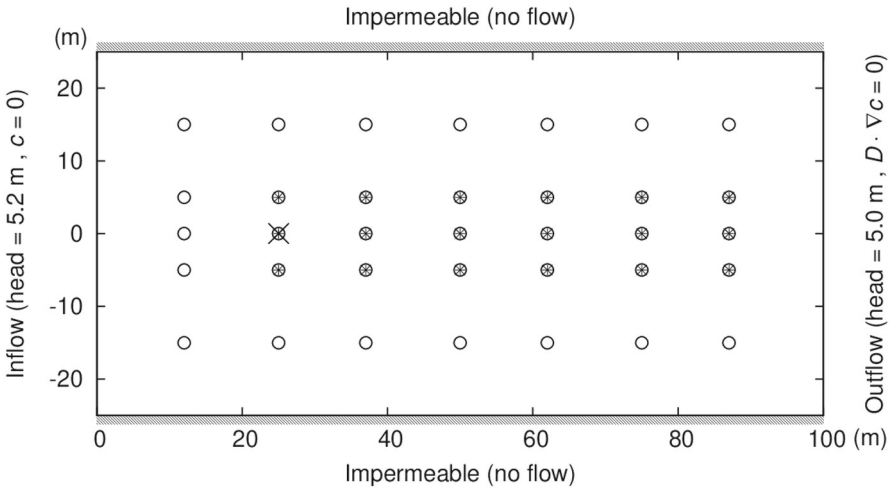

Fig. 3 Setting of the boundary conditions and well locations in the model domain. Open circles and asterisks indicate the monitoring wells of the hydraulic head (all 35 wells) and the contamination (18 wells), respectively. The cross mark indicates the location of the pumping well and contaminant source. The meaning of the circles and the cross mark are the same in Figs. 5, 6, 10, 11, and 13

boundaries were set at $x = 0$ (inflow) and 100 m (outflow) with a head difference of 0.2 m, and impermeable boundary conditions were set at both y edges ($y = \pm$ 25 m) (Fig. 3). The contaminant concentration at $x = 0$ m and the dispersive flux at $x =$ 100 m were both zero. The longitudinal and transverse dispersivities were defined as 5.0 and 0.5, respectively, considering that $\alpha_L \sim 0.1 L_p$ (Lallemand-Barres and Peaudecerf 1978; Pickens and Grisak 1981; Spitz and Moreno 1996) or $\alpha_L = 0.83 [\log_{10}(L_p)]^{2.414}$, where L_p is the plume length from the source [m] (Xu and Eckstein 1995) and α_T is approximately $0.1\alpha_L$ (Gelhar et al. 1992; Wiedemeier et al. 1999). R_f and λ_f were not considered, and τ was set to 1.

The contaminant plume distribution originated from a known source at $(x, y) = (25, 0)$ m. The release of the contaminant starts at $t = -300$ days (case 1) and -150 days (case 2) before the initial measurement time ($t = 0$) and ends at $t = 0$. The source intensity was assumed to increase linearly from 0 to 1. Hydraulic heads were measured at 35 monitoring wells under steady state for one pumping well at the source location (Fig. 4), and contaminant concentrations were measured monthly over 1 year ($t = 0$ to 1 year) at 18 monitoring wells located uniformly on the downstream side. In the calculation, the mean travel time was used instead of the transient concentration data, as suggested for tracer data (Harvey and Gorelick 1995; Ezzedine and Rubin 1996; Cirpka and Kitanidis 2000; Lee and Kitanidis 2014)

$$\overline{t_{\boldsymbol{x}_i}} = \frac{\sum_{t=t_0}^{t_{end}} t z(\boldsymbol{x}_i, t) \Delta t}{\sum_{t=t_0}^{t_{end}} z(\boldsymbol{x}_i, t) \Delta t}, \tag{26}$$

where $\overline{t_{\boldsymbol{x}_i}}$ is the mean travel time at position $\boldsymbol{x}_i$; $t = (t_0, \ldots, t_{end})$ is the measurement time; and Δt is the measurement interval.

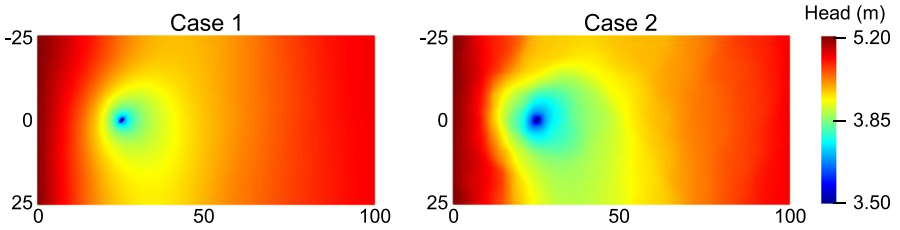

Fig. 4 Hydraulic head distribution under steady state when pumping at one well for cases 1 (left) and 2 (right). The location of the pumping well is shown in Fig. 3

Using the 35 head and 18 travel time data, the log-transmissivities at the 5,000 (100 $\times$ 50) cells at intervals of 1 m along the x- and y-axes were estimated. Assuming that the transmissivities at the 35 wells were known, the unknown structural parameters θ_r were estimated to be 2.0×10^{-5} for case 1 and 12.5 for case 2. In the contaminant plume estimation, the optimal values of both σ_R and θ_s were determined simultaneously. The standard deviation of the measurement errors of the head and concentration (or the mean travel time as mentioned above) were set to 0.05 m (approximately 5% of the maximum head change as a result of pumping) and 10%, respectively, following Lee and Kitanidis (2014). A Gaussian random error with zero mean and a corresponding standard deviation was added to all of the measurement data. The unknown release histories were recovered at ten-day intervals over the 1,350 days prior to the start of measurement, which is sufficiently long to express the Jacobian matrix for the contaminant plume evolution from the source to the model boundary.

3.3 Calculation Execution Conditions

The forward simulation of the groundwater flow and transport was executed using 3D-SEEP (Kimura and Muraoka 1986), based on the three-dimensional Garlerkin finite element method. Singular value decompositions for the low-rank approximations were computed in parallel using the ScaLAPACK package (Blackford et al. 1997). The linear systems of Eqs. (6) and (15) were solved using the generalized minimal residual method with a criterion for the relative residual error of $\leq 1 \times 10^{-8}$. A PC with an Intel Core i9-11900 K (3.50 GHz) CPU and 64-GB memory was used for the numerical experiments.

For both cases, the initial transmissivity field was set to be uniform with a log mean of -9.0 m^2/s. Following Lee and Kitanidis (2014), the optimum number of the low-rank approximation of $\boldsymbol{Q}$ was set to $K = 96$, in which the relative error of the approximation λ_{K+1}/λ_1 was 3.1×10^{-4}% for case 1 and 1.2% for case 2. Only for the joint inversion of case 2, which is a strongly nonlinear problem, was K changed to 350 with $\lambda_{K+1}/\lambda_1 = 0.18\%$. To ensure the monotonic convergence of the nonlinear transmissivity estimation problem [Eq. (13)], the optimal solution was identified using a line search (Zanini and Kitanidis 2009)

$$\hat{\boldsymbol{r}} = \boldsymbol{r}_i \delta_{ls} + \boldsymbol{r}_{i+1}(1 - \delta_{ls}), \tag{27}$$

where $\boldsymbol{r}_i$ is the previous estimate and $\boldsymbol{r}_{i+1}$ is the updated estimate found using the Gauss–Newton procedure [Eq. (17)] and δ_{ls} is a scalar. The range of δ_{ls} was set to $-0.1 \leq \delta_{ls} \leq 1.1$ following Zanini and Kitanidis (2009). Finally, the calculations of the transmissivity estimation converged entirely within 18 iterations for all cases with $\varepsilon_r = 5 \times 10^{-6}$.

To obtain the final solution, the estimated transmissivity distributions were updated two and three times for cases 1 and 2, respectively. The uncertainty of the contaminant plume distributions was evaluated using the results of 10,000 ($N_r = 100 \times N_{z0} = 100$) realizations. Because of rounding errors, the eigenvalues of the posterior covariance included small negative values (approximately -10^{-7}); all the negative eigenvalues were therefore changed to 1×10^{-10}.

4 Results

4.1 Hydraulic Transmissivity

The best estimates and estimation variances of the log-transmissivity distributions are shown in Figs. 5 and 6, respectively. Even for the results using only the head data, sufficient accuracy of the best estimates can be confirmed by the near agreement between the simulated and measured heads having small root mean square errors (RMSE) of 0.047 (case 1) and 0.054 (case 2) (Fig. 7a). However, because of the measurement error, the results are spatially much smoother than the true fields in both cases. In particular, large underestimates occurred at the relatively high transmissivity portions; these are continuously distributed from the upstream to the downstream (case

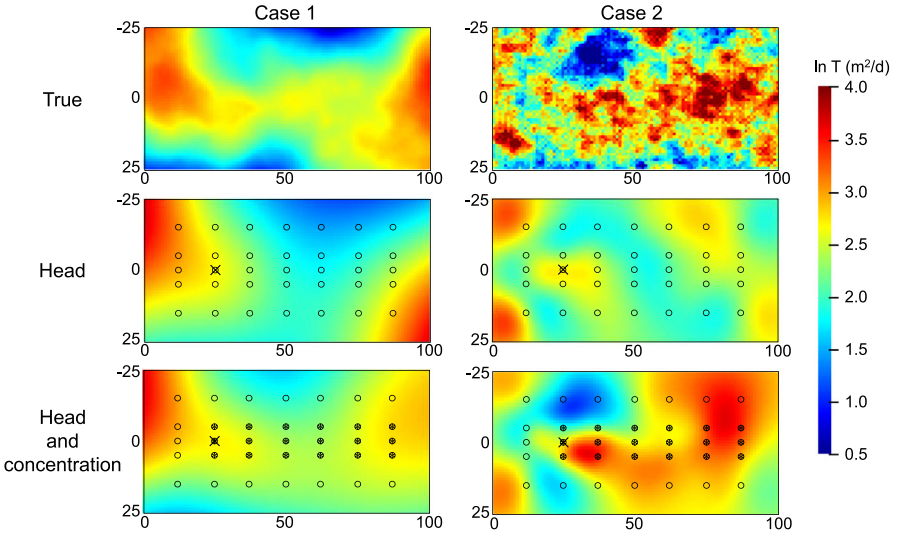

Fig. 5 True (top) and best estimates of the log-transmissivity distributions for cases 1 and 2 using only the head data (middle) and using both the head and concentration data (bottom)

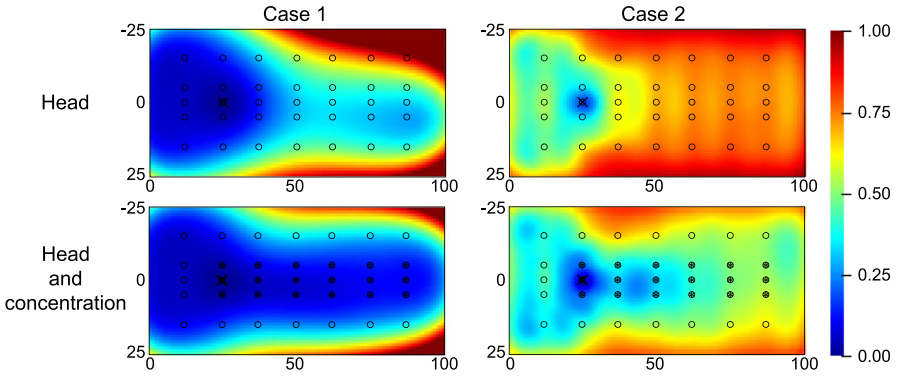

Fig. 6 Estimation variances of the log-transmissivity distributions for cases 1 and 2 using only the head data (top) and using both the head and concentration data (bottom)

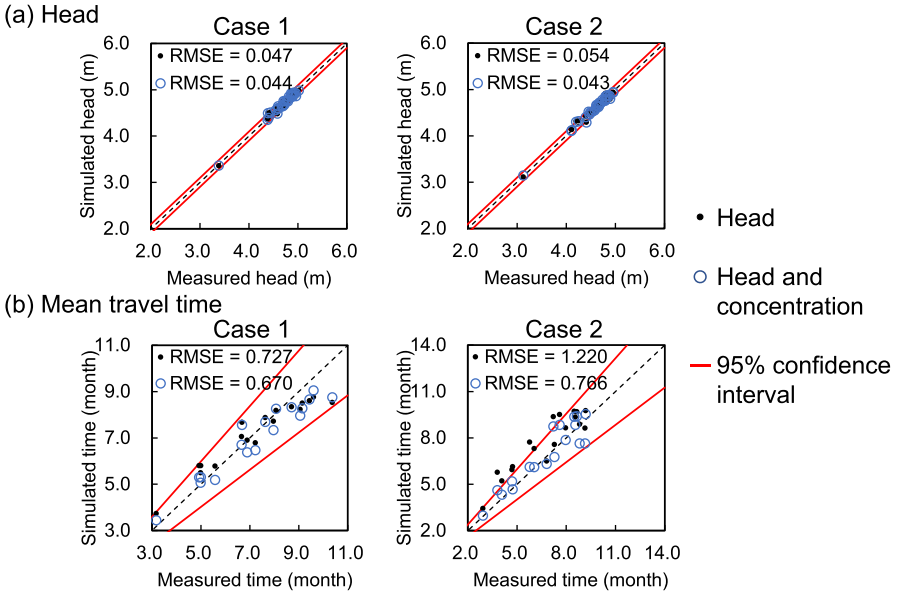

Fig. 7 Simulated versus measured values for (a) the head and (b) the mean travel time using only the head data (closed circle) and using both the head and concentration data (open circles)

1) and are heterogeneously distributed on the downstream side (case 2). This smoothing effect was remarkable in case 2 with the highly heterogeneous field; furthermore, the estimation uncertainty significantly increased with distance from the pumping well.

Conversely, through the joint use of the head and concentration data, large improvements in the estimation accuracy were confirmed for both cases, for example, the relatively high transmissivity portions were well reproduced. Both the measurement data of the head and the mean travel time were adequately reproduced within their

95% confidence intervals (Fig. 7b). Even though the estimation accuracy was low in the portions outside the contaminant plume transport, such as in the vicinity of the domain boundary, the joint data use obviously decreased the estimation uncertainties along the contaminant plume evolutions for both cases compared with the results using only the head data (Fig. 6).

The estimation accuracies of the log-transmissivity distributions were assessed in the measurement area, $x = [12, 87]$ (m), $y = [-15, 15]$ (m), using the descriptive statistics: the coefficient of determination (R^2) between the estimated and true values (Fig. 8), mean (μ_0) and variance (ν_0) of the log-transmissivity field; and mean ($\mu_2 = \text{mean}\left[\left(\boldsymbol{r}_{\text{true}} - \hat{\boldsymbol{r}}\right)^2\right]$) and variance ($\nu_2 = \text{var}\left[\left(\boldsymbol{r}_{\text{true}} - \hat{\boldsymbol{r}}\right)^2\right]$) of the square differences (Table 2). These statistical parameters demonstrate the large improvement of the estimation accuracy via the joint use of the head and concentration data for both cases (e.g., μ_2 was decreased by 76% and 50% for cases 1 and 2, respectively).

Table 2 Summary of statistics of the best estimates of the log-transmissivity distributions (m^2/d) for cases 1 and 2

Case		μ_0	ν_0	μ_2	ν_2
1	True	2.460	0.075	–	–
	Head	2.170	0.137	0.176	0.034
	Head and concentration	2.470	0.037	0.041	0.004
2	True	2.560	0.601	–	–
	Head	2.250	0.070	0.616	0.560
	Head and concentration	2.660	0.362	0.311	0.177

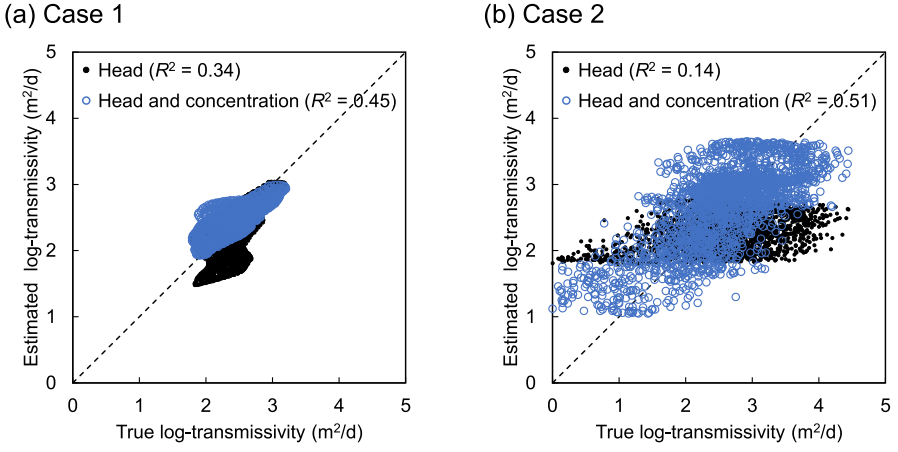

Fig. 8 Estimated versus true log-transmissivities in the measurement area for (a) case 1 and (b) case 2 using only the head data (closed circles) and both the head and concentration data (open circles)

Table 3 Optimal values of the structure parameters for the best estimates of the release history in the two cases

Parameter value	Case 1		Case 2	
	Head	Head and concentration	Head	Head and concentration
θ_s	4.2×10^{-7}	6.3×10^{-7}	1.4×10^{-6}	2.5×10^{-6}
σ_{R_z}	1.0×10^{-2}	1.0×10^{-2}	1.0×10^{-2}	1.0×10^{-2}

4.2 Contaminant Plume Distribution

Using the optimal values of θ_s and σ_{R_z} (Table 3), the release histories for the best estimates of the log-transmissivity distributions were estimated for the two cases, as shown in Fig. 9. When using both the head and concentration data, the recovered release peaks were slightly closer to the true peaks for both cases compared with the results when using only the head data. This is due to the improvement in the estimated transmissivity distribution, in particular the decrease in the underestimation induced when using only the head data. The magnitude of the estimated variance of the release history depends on both the concentration measurement error and the heterogeneity of the transmissivity distribution. As reported by Butera and Tanda (2003), the model mismatch error of the hydraulic field can increase the estimation variance for two main reasons: errors in the flow direction and the magnitude of the dispersion. The large confidence intervals in Fig. 9 are caused by the former factor in case 1 using only the head data; and by the latter factor in case 2 because of the large heterogeneous transmissivity.

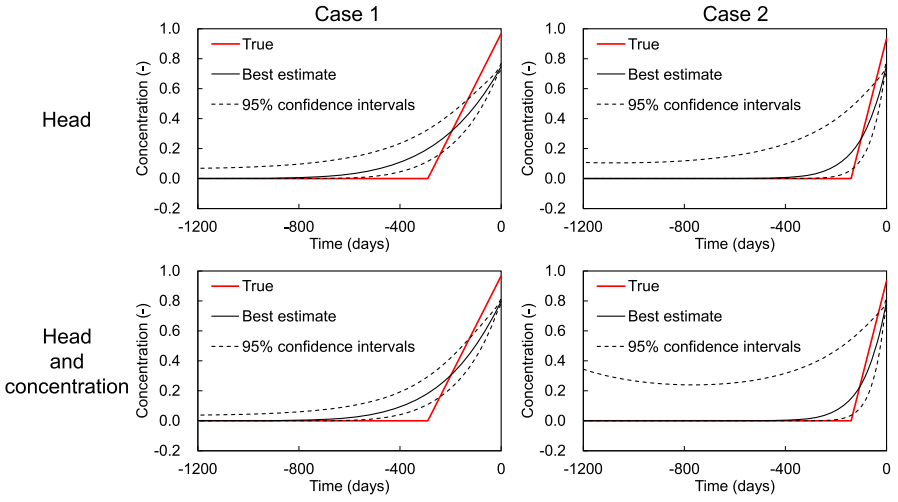

Fig. 9 Estimated release histories using only the head data (dashed line) and using both the head and concentration data (solid line) for cases 1 (left) and 2 (right). The red line indicates the true release history

The best estimates of the transmissivity distributions were applied to the prediction of the best estimate of the contaminant plume evolution, as drawn in Figs. 10 and 11 for the results of the plume distribution at $t = 0$, 1, and 2 years for cases 1 and 2. When using only the head data, the contaminant transport velocity was underestimated in both cases. This underestimation resulted in the maximum concentrations at $t = 1$ and 2 years being 1.4 (case 1) and 1.7 (case 2) times as large as the true concentrations. In addition, the center of plume distribution differs from the true position. At $t =$ 2 year, the true maximum concentrations are located at 85 m (case 1) and outside the model domain (case 2). However, the estimated positions when using only the head data are 69 m (case 1) and 66 m (case 2). Conversely, given the joint use of the

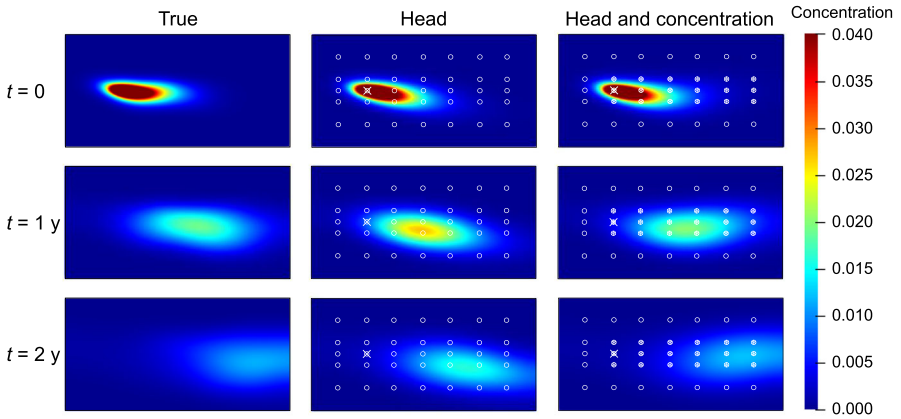

Fig. 10 True (left) and best estimates of the contaminant plume distributions at $t = 0$ (the initial measurement time), 1, and 2 years later for case 1, obtained using only the head data (middle) and using both the head and concentration data (right)

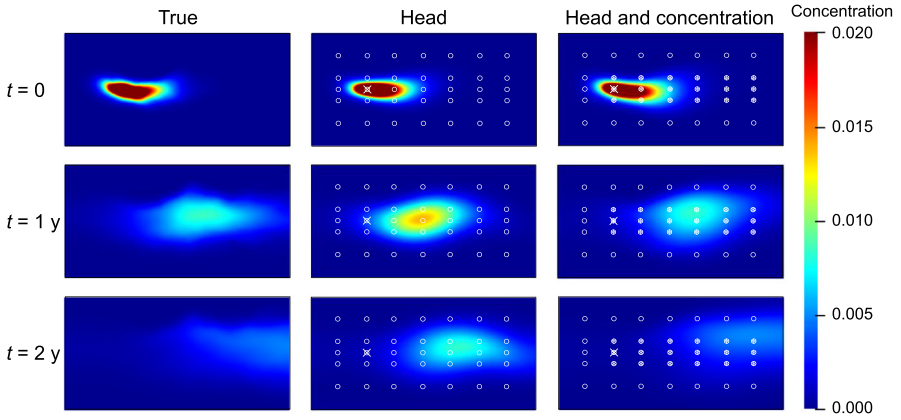

Fig. 11 True (left) and best estimates of contaminant plume distributions at $t = 0$ (the initial measurement time), 1, and 2 years later for case 2, obtained using only the head data (middle) and using both the head and concentration data (right)

head and concentration data, the plume evolution is well reproduced in both cases. The advantage of using both forms of data was proved by the significant increases in the mean of the linear correlation coefficients at one month interval from $t = 0$ to 2 years between the true and best estimated concentrations at the 18 measurement points (Fig. 12): the time-averaged correlation coefficient increased from 0.72 to 0.97 for case 1 and from 0.67 to 0.97 for case 2.

The uncertainties in the contaminant plume transport were quantified for the results through the joint data use (Fig. 13). Obviously, the resultant uncertainties were sufficiently small compared with the best estimate values in Figs. 9 and 10. At $t = 0$, the uncertainties were relatively high as a result of the estimated variance of the release histories (Fig. 9); this was more conspicuous in case 2. However, these fluctuations decreased with time.

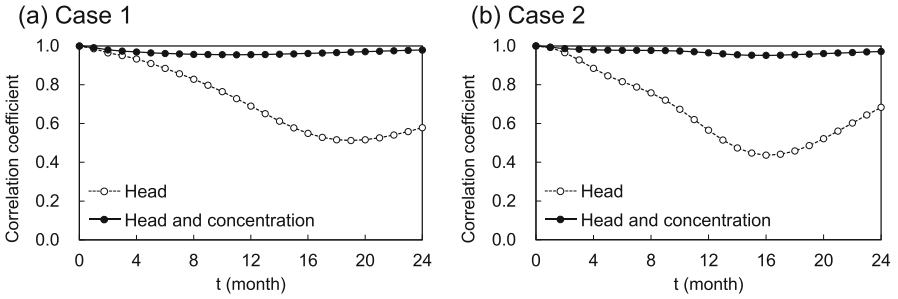

Fig. 12 Time evolution of the correlation coefficients between the true and best estimated concentration at measurement points for (a) case 1 and (b) case 2 obtained using only the head data (open circles) and using both the head and concentration data (closed circles)

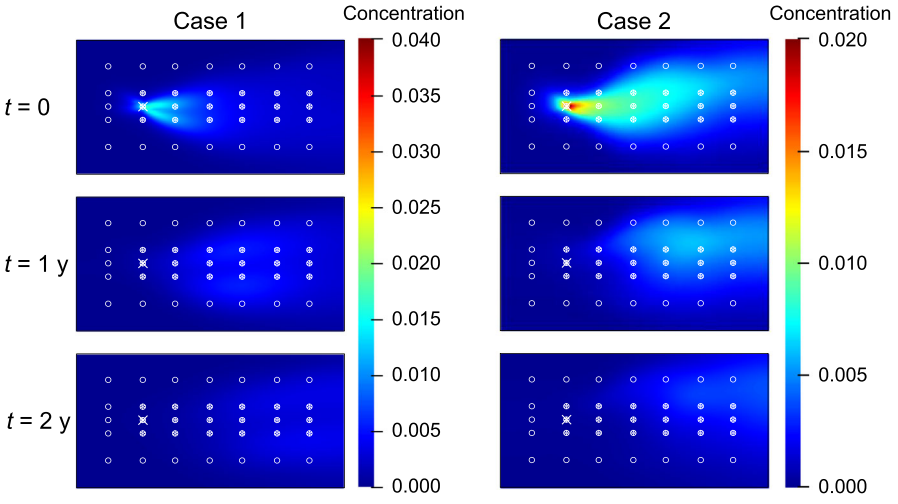

Fig. 13 Standard deviations of the contaminant plume distributions at $t = 0$, 1, and 2 years later for cases 1 (left) and 2 (right) using both the head and concentration data

5 Discussion

The effectiveness and high accuracy of the joint clarification of the initial contaminant plume and transmissivity distributions using both the head and concentration data were demonstrated for the prediction of the contaminant plume evolution. The proposed GA method is applicable for any water-soluble contaminant with or without retardation and/or radioactive or biochemical decay. However, the flow and transport simulation was simply implemented under steady state flow conditions. This assumption cannot be satisfied for cases that need to consider unsteady flow caused by typically periodic pumping. The next step is to incorporate the unsteady flow in both unconfined and confined aquifers.

As mentioned above, the two main novel points of this study differing from previous studies were the non-use of tracer data and the consideration of the uncertainty of the initial contaminant plume distribution. While a smaller model domain (40 m $\times$ 20 m) and a smaller head measurement error (5.0×10^{-4} m) than that used in this study was targeted, the effectiveness of the joint use of the head and tracer data with a defined anisotropic exponential covariance of the transmissivity was, in previous studies, demonstrated to improve the estimation accuracy of log-transmissivity inside the tracer paths (Cirpka and Kitanidis 2000; Lee and Kitanidis 2014). For a similar exponential covariance model (case 2), the joint use of the head and tracer data was demonstrated to improve the estimation accuracy based on the RMSE ($\sqrt{\mu_2}$) of the log-transmissivity in the entire domain by 10% compared with that obtained using only the head data (Cirpka and Kitanidis 2000). Although this study did not use tracer data, the RMSE in the measurement area was improved by 29% via the joint use of the head and concentration data. This high accuracy contributes to accurately reproducing the contaminant plume evolution even if the release history is unknown. Another noteworthy advantage of the present method is its capability to evaluate the uncertainty of the contaminant plume evolution, considering the uncertainties of both the initial contaminant plume and the transmissivity distributions. This feature resulted in the sufficiently small estimation uncertainty of the contaminant plume evolution even for the highly heterogeneous transmissivity field (case 2) compared with the best estimate values.

As an extension of the proposed method to practical applications, the following three points need to be considered. The first point is the need to estimate the three-dimensional permeability (i.e., the hydraulic conductivity). Although this is possibly straightforward given the proposed method, a problem is the large cost of acquiring a sufficient amount of multi-depth data of the head and contaminant concentration for the three-dimensional estimation. Therefore, a suitable data amount and the location of the measurement data should be specified depending on the hydrogeological features of the target area. The second point is the reduction of the smoothing effect associated essentially with the spatial estimation of geostatistical methods. As shown in this study, even though the estimated transmissivity field is spatially smoother than the true field, the contaminant plume evolution can be reproduced well because the contaminant plume spreads over time primarily as a result of mechanical dispersion. However, reducing the smoothing effect is indispensable when a hotspot-shaped concentration anomaly much higher than the surroundings needs to be reproduced. This reduction

may be possible by applying geostatistical simulations to conditional realizations, typically via sequential Gaussian simulations (e.g., Deutsch and Journel 1998), turning band simulations (e.g., de Sá et al. 2021a, b), and the incorporation of discontinuous geological structures such as lithological contacts and unconformities (Fienen et al. 2008, 2009; Koike et al. 2022) and fracture/fault distribution (Zha et al. 2017). The third point is to incorporate the constraints on release history, except for the non-negativity, for example, the upper limit of concentration as a result of the contaminant solubility. This can be achieved via Gibbs sampling (Michalak 2008), a representative Markov chain Monte Carlo method whose effectiveness for contaminant plume estimation was demonstrated by Takai et al. (2022) using a set of field data.

6 Conclusions

This study developed a geostatistical method to achieve accurate estimation of contaminant plume evolutions via a joint clarification of the contaminant plume and hydraulic transmissivity distributions. One of the novelties of this method is the use of the contaminant concentration data in the consideration of the uncertainty of a contaminant plume distribution originating from an unknown release history. To verify the effectiveness and accuracy of the proposed method, two transmissivity fields with different spatial patterns were prepared: a high-contrast smooth field (case 1) and a highly heterogeneous field (case 2). The main obtained results are summarized as follows.

(1) Even though the contaminant plume distribution was unknown, a higher estimation accuracy of the hydraulic transmissivity distribution was achieved through the joint use of the head and concentration data than when using only the head data. For case 2, the estimation accuracy was improved similarly to the previous study (Cirpka and Kitanidis 2000) using tracer data.
(2) Using both the head and concentration data, the release peaks were recovered more accurately than when using only the head data. This superiority was due to the improvement of the estimated transmissivity distribution by decreasing the smoothing effect.
(3) Large difference between the results through the non-use and use of the concentration data with the head data was highlighted in the predicted contaminant plume evolution. Using only the head data, the velocity of the estimated plume transport was almost half that of the true value. Conversely, using both the head and concentration data, the plume evolutions were sufficiently predicted for both cases with high time-averaged correlation coefficients of 0.97, respectively, between the true and predicted concentrations. Furthermore, the uncertainties of the predicted plume distributions based on the conditional realizations of the initial plume and transmissivity distributions were sufficiently smaller than the magnitudes of the best estimates.

Consequently, the effectiveness and accuracy of the proposed approach were demonstrated even if the initial contaminant plume distribution is uncertain. Any forward modeling methods of the groundwater flow and the contaminant transport can be incorporated into the method. Accurate predictions of the contaminant plume

transport are helpful to effectively plan remediation, in particular, when choosing the number and location of pumping-up wells. Because the applicability was demonstrated only via numerical experiments in this study, our next step will be a practical application to actual contaminated fields using multi-depth head and concentration data, targeting three-dimensional space and considering an additional uncertainty such as the uncertainty in the boundary conditions of the flow and transport model.

Acknowledgements This study was funded by the Secretariat of Nuclear Regulation Authority, Nuclear Regulation Authority, Japan. Sincere thanks are extended to the two anonymous reviewers for their essential and constructive comments and suggestions that improved the clarity of this manuscript.

Author contributions All authors contributed to the study conception and design. Data analysis and modeling were planned and performed by ST and KK. The first draft of the manuscript was written by ST, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix: Normalization of the Prior Model

The geostatistical estimation of the transmissivity was implemented via the normalization of the prior model using the following procedure. More details are given in Kitanidis and Lee (2014).

List of symbols

Parameter		Dimension
All		
t	Time	
$\boldsymbol{x}$	Space	$m \times 1$
m_t	Number of unknowns (source intensity)	
m	Number of unknowns (contaminant plume/transmissivity)	
n_z	Number of observations (concentration)	
n_φ	Number of observations (head)	
$\delta(\boldsymbol{x})$	Dirac delta function	

List of symbols

Parameter		Dimension
Estimation of initial contaminant plume distribution		
$\boldsymbol{s}$	Discretized unknown source intensity	$m_t \times 1$
z_0	Discretized unknown initial contaminant plume	$m \times 1$
$\boldsymbol{z}_0^*$	Observation (initial concentration)	$n_z \times 1$
$\boldsymbol{H}_s, \boldsymbol{H}_s^*$	Jacobian matrix (for whole domain/observations)	$m \times m_t, n_z \times m_t$
$\boldsymbol{v}_z$	Observation error	$n_z \times 1$
σ_{R_z}	Standard deviation of error	
$\boldsymbol{R}_z$	Error covariance matrix	$n_z \times n_z$
$\boldsymbol{X}_s$	Known drift matrix	$m_t \times p_s$
$\boldsymbol{\beta}_s$	Unknown drift coefficients	$p_s \times 1$
$\boldsymbol{Q}_s$	Generalized prior covariance matrix of unknown function	$m_t \times m_t$
θ_s	Structure parameter of covariance $\boldsymbol{Q}_s$	
$\boldsymbol{V}_{\hat{s}}$	Covariance matrix of estimated source intensity	$m_t \times m_t$
$\boldsymbol{V}_{\hat{z}}$	Covariance matrix of estimated initial contaminant plume	$m \times m$
Estimation of hydraulic transmissivity distribution		
n	Number of observations	
$\boldsymbol{r}$	Discretized unknown log-transmissivity	$m \times 1$
$z(t)$	Discretized unknown contaminant plume at time t	$m \times 1$
$\boldsymbol{y}$	Observation	$n \times 1$
$\boldsymbol{\varphi}$	Observation (head)	$n_\varphi \times 1$
$\boldsymbol{z}^*$	Observation (transient concentration)	$n_z \times (t_0, \ldots, t_{end})$

List of symbols

Parameter		Dimension
$\bar{\boldsymbol{t}}$	Observation (mean travel time)	$n_z \times 1$
$\boldsymbol{H}$	Jacobian matrix	$n \times m$
$\boldsymbol{v}$	Observation error	$n \times 1$
σ_R	Standard deviation of error	
$\boldsymbol{R}$	Error covariance matrix	$n \times n$
$\boldsymbol{X}$	Known drift matrix	$m \times p$
$\boldsymbol{\beta}$	Unknown drift coefficients	$p \times 1$
$\boldsymbol{Q}$	Generalized prior covariance matrix of unknown function	$m \times m$
θ_r	Structure parameter of covariance $\boldsymbol{Q}$	
K	Rank of approximation of $\boldsymbol{Q}$	
$\boldsymbol{V}_{\hat{\boldsymbol{r}}}$	Covariance matrix of estimated transmissivity	$m \times m$
δ_r	Finite difference interval	
ε_r	relative machine precision	
λ_i	ith eigenvalue of $\boldsymbol{Q}$	
$\boldsymbol{V}_i$	ith eigenvector of $\boldsymbol{Q}$	$m \times 1$
δ_{ls}	Finite difference interval for line search	
Conditional realization		
N_{z_0}	Number of realizations (initial contaminant plume)	
N_r	Number of realizations (transmissivity)	
$\boldsymbol{s_u}, \boldsymbol{s_c}$	Unconditional/conditional realization of $\boldsymbol{s}$	$m_t \times 1$
$\boldsymbol{z}_{0_c}$	Conditional realization of $\boldsymbol{z}_0$	$m \times 1$
$\boldsymbol{r_u}, \boldsymbol{r_c}$	Unconditional/conditional realization of $\boldsymbol{r}$	$m \times 1$
Physical model		
$\boldsymbol{u}$	Groundwater velocity	2×1
T	Transmissivity	$m \times 1$
Q_f	Pumping rate	

List of symbols

Parameter		Dimension
$\boldsymbol{V}$	Actual groundwater velocity	2×1
R_f	Retardation factor	
λ_f	Decay constant	
$\boldsymbol{D}$	Dispersion tensor	2×2
ε	Porosity	
α_L, α_T	Dispersivity (longitudinal and transverse)	
D_m	Molecular diffusion coefficient	
τ	Tortuosity	
L_p	Plume length	

First, the drift matrix $\boldsymbol{X}$ is replaced with its normalized and isomorphic matrix $\boldsymbol{U}$ such that

$$\boldsymbol{U} = \begin{cases} \boldsymbol{X}/\sqrt{m} & (p = 1) \\ \boldsymbol{U}\boldsymbol{S}_X\boldsymbol{V}_X^T & (p > 1) \end{cases}, \tag{28}$$

where $\boldsymbol{S}_X \in \mathbb{R}^{p\times p}$ is a diagonal matrix of the singular values and the columns of $\boldsymbol{V}_X \in \mathbb{R}^{p\times p}$ are the orthonormal eigenvectors of $\boldsymbol{X}^T\boldsymbol{X}$. Then, $\boldsymbol{U}$ is used to compute the detrending matrix $\boldsymbol{P} \in \mathbb{R}^{m\times m}$

$$\boldsymbol{P} = \boldsymbol{I} - \boldsymbol{U}\boldsymbol{U}^T. \tag{29}$$

The next step is to detrend the low-rank covariance $\boldsymbol{Q}$ [Eq. (20)]. First, $\boldsymbol{Z}_Q$ is replaced with $\boldsymbol{P}\boldsymbol{Z}_Q$. Then, the singular value decomposition of $\boldsymbol{Z}_Q$ is calculated such that

$$\boldsymbol{Z}_Q = \boldsymbol{U}_Z\boldsymbol{S}_Z\boldsymbol{V}_Z^T, \tag{30}$$

where $\boldsymbol{U}_Z \in \mathbb{R}^{m\times K}$ is the unitary matrix; $\boldsymbol{S}_Z \in \mathbb{R}^{K\times K}$ is a diagonal matrix of the singular values; and the columns of $\boldsymbol{V}_Z \in \mathbb{R}^{K\times K}$ are the orthonormal eigenvectors of $\boldsymbol{Z}_Q^T\boldsymbol{Z}_Q$. Using $\boldsymbol{C} = \boldsymbol{U}_Z^T\boldsymbol{Q}\boldsymbol{U}_Z$, $\boldsymbol{Q}$ is replaced with its detrended and isomorphic matrix $\boldsymbol{PQP}$

$$\boldsymbol{P}\boldsymbol{Q}\boldsymbol{P} \approx \boldsymbol{U}_Z\boldsymbol{C}\boldsymbol{U}_Z^T. \tag{31}$$

Then, $\boldsymbol{HQ}$ and $\boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^T$ can be approximated as

$$\boldsymbol{H}\boldsymbol{Q} \approx (\boldsymbol{H}\boldsymbol{U}_Z)\boldsymbol{C}\boldsymbol{U}_Z^T \equiv \boldsymbol{B}\boldsymbol{C}\boldsymbol{U}_Z^T, \tag{32}$$

$$\boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^T \approx (\boldsymbol{H}\boldsymbol{U}_Z)\boldsymbol{C}(\boldsymbol{H}\boldsymbol{U}_Z)^T = \boldsymbol{B}\boldsymbol{C}\boldsymbol{B}^T. \tag{33}$$

Finally, $\overline{\boldsymbol{r}}$ is updated iteratively by the following linear equation system corresponding to Eq. (6)

$$\begin{pmatrix} \boldsymbol{\Sigma} & \boldsymbol{H}\boldsymbol{X} \\ (\boldsymbol{H}\boldsymbol{X})^T & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\Lambda}^T\boldsymbol{A}_p \\ \boldsymbol{M}\boldsymbol{A}_p \end{pmatrix} = \begin{pmatrix} \boldsymbol{H}\boldsymbol{Q}\boldsymbol{A}_p \\ \boldsymbol{X}^T\boldsymbol{A}_p \end{pmatrix}, \boldsymbol{A}_p = (\boldsymbol{U}_Z, \boldsymbol{U}), \tag{34}$$

$$\overline{\boldsymbol{r}} = \boldsymbol{A}_p\big(\boldsymbol{\Lambda}^T\boldsymbol{U}_Z\big)^T(\boldsymbol{y} - h(\overline{\boldsymbol{r}}) + \boldsymbol{H}\overline{\boldsymbol{r}}). \tag{35}$$

Once the optimal solution $\hat{\boldsymbol{r}}$ is obtained, the posterior covariance $\boldsymbol{V}_{\hat{r}}$ can be approximately calculated as

$$\boldsymbol{V}_{\hat{r}} = \boldsymbol{U}_Z\boldsymbol{C}\boldsymbol{U}_Z^T - \boldsymbol{X}\big(\boldsymbol{M}\boldsymbol{A}_p\big)\boldsymbol{A}_p^T - \boldsymbol{Q}\boldsymbol{H}^T\big(\boldsymbol{\Lambda}^T\boldsymbol{A}_p\big)\boldsymbol{A}_p^T. \tag{36}$$

References

Adamson DT, Mahendra S, Walker KL Jr, Rauch SR, Sengupta S, Newell CJ (2014) A multisite survey to identify the scale of the 1,4-dioxane problem at contaminated groundwater sites. Environ Sci Technol Lett 1:254–258. https://doi.org/10.1021/ez500092u

Blackford LS, Choi J, Cleary A, D'Azevodo EF, Demmel J, Dhillon IS, Dongarra J, Hammarling S, Henry G, Petitet A, Stanley K, Walker DW, Whaley RC (1997) ScaLAPACK: a linear algebra library for message-passing computers. In: Proceedings of the eighth SIAM conference on parallel processing for scientific computing (PPSC 1997), Minnesota, USA, March 1997. Society for Industrial and Applied Mathematics

Box GEP, Cox DR (1964) An analysis of transformations. J R Stat Soc Ser B 26:353–360

Butera I, Tanda MG (2003) A geostatistical approach to recover the release history of groundwater pollutants. Water Resour Res 39(12):1372. https://doi.org/10.1029/2003WR002314

Cardiff M, Barrash W (2011) 3-D transient hydraulic tomography in unconfined aquifers with fast drainage response. Water Resour Res 47(12):W12518. https://doi.org/10.1029/2010WR010367

Cardiff M, Barrash W, Kitanidis PK, Malama B, Revil A, Straface S, Rizzo E (2009) A potential-based inversion of unconfined steady-state hydraulic tomography. Ground Water 47(2):259–270. https://doi.org/10.1111/j.1745-6584.2008.00541.x

Chen Z, Gomez-Hernandez JJ, Xu T, Zanini A (2018) Joint identification of contaminant source and aquifer geometry in a sandbox experiment with the restart ensemble Kalman filter. J Hydrol 564:1074–1084. https://doi.org/10.1016/j.jhydrol.2018.07.073

Chen Z, Xu T, Gomez-Hernandez JJ, Zanini A (2021) Contaminant spill in a sandbox with non-Gaussian conductivities: simultaneous identification by the restart normal-score ensemble Kalman filter. Math Geosci 53:1587–1615. https://doi.org/10.1007/s11004-021-09928-y

Cirpka OA, Kitanidis PK (2000) Sensitivity of temporal moments calculated by the adjoint-state method and joint inversing of head and tracer data. Adv Water Resour 24(1):89–103. https://doi.org/10.1016/S0309-1708(00)00007-5

De Sá VR, Koike K, Goto T, Nozaki T, Takaya Y, Yamasaki T (2021a) A combination of geostatistical methods and principal components analysis for detection of mineralized zones in seafloor hydrothermal systems. Nat Resour Res 30:2875–2887. https://doi.org/10.1007/s11053-020-09705-4

De Sá VR, Koike K, Goto T, Nozaki T, Takaya Y, Yamasaki T (2021b) 3D geostatistical modeling of metal contents and lithofacies for mineralization mechanism determination of a seafloor hydrothermal deposit in the middle Okinawa Trough, Izena Hole. Ore Geol Rev 135:104194. https://doi.org/10.1016/j.oregeorev.2021.104194

Deutsch CV, Journel AG (1998) GSLIB: geostatistical software library and user's guide. Applied geostatistics series. Oxford University Press, New York

Ezzedine S, Rubin Y (1996) A geostatistical approach to the conditional estimation of spatially distributed solute concentration and notes on the use of tracer data in the inverse problem. Water Resour Res 32(4):853–861. https://doi.org/10.1029/95WR02285

Fienen MN, Clemo T, Kitanidis PK (2008) An interactive Bayesian geostatistical inverse protocol for hydraulic tomography. Water Resour Res 44(12):W00B01. https://doi.org/10.1029/2007WR006730

Fienen MN, Hunt R, Krabbenhoft D, Clemo T (2009) Obtaining parsimonious hydraulic conductivity fields using head and transport observations: a Bayesian geostatistical parameter estimation approach. Water Resour Res 45(8):W08405. https://doi.org/10.1029/2008WR007431

Gelhar LW, Welty C, Rehfeldt KR (1992) A critical review of data on field-scale dispersion in aquifers. Water Resour Res 28(7):1955–1974. https://doi.org/10.1029/92WR00607

Gyzl G, Zanini A, Fraczek R, Kura K (2014) Contaminant source and release history identification in groundwater: a multi-step approach. J Contam Hydrol 157:59–72. https://doi.org/10.1016/j.jconhyd.2013.11.006

Harvey CF, Gorelick SM (1995) Mapping hydraulic conductivity—sequential conditioning with measurements of solute arrival time, hydraulic-head, and local conductivity. Water Resour Res 31(7):1615–1626. https://doi.org/10.1029/95WR00547

Illuman WA, Liu X, Shinji T, Yeh TJ, Ando K, Saegusa H (2009) Hydraulic tomography in fractured granite: mizunami underground research site. Jpn Water Resour Res 45(1):W01406. https://doi.org/10.1029/2007WR006715

Jiang Y, Woodbury AD (2006) A full-Bayesian approach to the inverse problem for steady-state groundwater flow and heat transport. Geophys J Int 167:1501–1512. https://doi.org/10.1111/j.1365-246X.2006.03145.x

Kimura H, Muraoka S (1986) The 3D-SEEP computer code user's manual. Japan Atomic Energy Research Institute, JAERI-M 86-091. https://doi.org/10.11484/jaeri-m-86-091

Kitanidis PK (1995) Quasi-linear geostatistical theory for inversing. Water Resour Res 31(10):2411–2419. https://doi.org/10.1029/95WR01945

Kitanidis PK, Lee J (2014) Principal component geostatistical approach for large-dimensional inverse problems. Water Resour Res 50(7):5428–5443. https://doi.org/10.1002/2013WR014630

Koike K, Kiriyama T, Lu L, Kubo T, Heriawan MN, Yamada R (2022) Incorporation of geological constraints and semivariogram scaling law into geostatistical modeling of metal contents in hydrothermal deposits for improved accuracy. J Geochem Explor 233:106901. https://doi.org/10.1016/j.gexplo.2021.106901

Lallemand-Barres P, Peaudecerf P (1978) Recherche des relations entre les valeurs de la dispersivite macroscopique d'un milieu aquifere, ses autres caracteristiques et les conditions de measures. Etude Bibliographique Bull BRGM III(4):277–284

Lee J, Kitanidis PK (2014) Large-scale hydraulic tomography and joint inversion of head and tracer data using the Principal Component Geostatistical Approach (PCGA). Water Resour Res 50(7):5410–5427. https://doi.org/10.1002/2014WR015483

Lee J, Yoon H, Kitanidis PK, Werth CJ, Valocchi AJ (2016) Scalable subsurface inverse modeling of huge data sets with an application to tracer concentration breakthrough data from magnetic resonance imaging. Water Resour Res 52(7):5213–5231. https://doi.org/10.1002/2015WR018483

Li W, Englert A, Cirpka OA, Vanderborght J, Vereecken H (2007) Two-dimensional characterization of hydraulic heterogeneity by multiple pumping tests. Water Resour Res 43(4):W04433. https://doi.org/10.1029/2006WR005333

Li W, Englert A, Cirpka OA, Vereecken H (2008) Three-dimensional geostatistical inversion of flowmeter and pumping test data. Ground Water 46(2):193–201. https://doi.org/10.1111/j.1745-6584.2007.00419.x

Luo N, Illuman WA, Zha Y (2022) Large-scale three-dimensional hydraulic tomography analyses of long-term municipal wellfield operations. J Hydrol 610:127911. https://doi.org/10.1016/j.jhydrol.2022.127911

Michalak AM (2008) A Gibbs sampler for inequality-constrained geostatistical interpolation and inverse modeling. Water Resour Res 44(9):1–14. https://doi.org/10.1029/2007WR006645

Michalak AM, Kitanidis PK (2002) Application of Bayesian inference methods to inverse modeling for contaminant source identification at Gloucester Landfill, Canada. Comput Methods Water Resour XIV(2):259–1266

Michalak AM, Kitanidis PK (2003) A method for enforcing parameter nonnegativity in Bayesian inverse problems with an application to contaminant source identification. Water Resour Res 39(2):1033. https://doi.org/10.1029/2002WR001480

Nowak W (2009) Best unbiased ensemble linearization and the quasi-linear Kalman ensemble generator. Water Resour Res 45(4):W04431. https://doi.org/10.1029/2008WR007328
OECD (2014) Nuclear site remediation and restoration during decommissioning of nuclear installations: a report by the NEA co-operative programme on decommissioning. Radioactive Waste Management, NEA No. 7192. https://doi.org/10.1787/9789264222182-en
Pickens JF, Grisak GE (1981) Scale-dependent dispersion in a stratified granular aquifer. Water Resour Res 17(4):1191–1211. https://doi.org/10.1029/WR017i004p01191
Pouladi B, Linde N, Longuevergne L, Bour O (2021) Individual and joint inversion of head and flux data by geostatistical hydraulic tomography. Adv Water Resour 154:103960. https://doi.org/10.1016/j.advwatres.2021.103960
Sanchez-Leon E, Leven C, Haslauer CP, Cirpka OA (2016) Combining 3D hydraulic tomography with tracer tests for improved transport characterization. Ground Water 54(4):498–507. https://doi.org/10.1111/gwat.12381
Shlomi S, Michalak AM (2007) A geostatistical framework for incorporating transport information in estimating the distribution of a groundwater contaminant plume. Water Resour Res 43(3):1–12. https://doi.org/10.1029/2006WR005121
Singh RM, Datta B (2004) Groundwater pollution source identification and simultaneous parameter estimation using pattern matching by artificial neural network. Environ Forensics 5(3):143–153. https://doi.org/10.1080/15275920490495873
Snodgrass MF, Kitanidis PK (1997) A geostatistical approach to contaminant source identification. Water Resour Res 33(4):537–546. https://doi.org/10.1029/96WR03753
Spitz K, Moreno J (1996) A practical guide to groundwater and solute transport modeling, 1st edn. Wiley, New York
Takai S, Shimada T, Takeda S, Koike K (2022) Evaluating the effectiveness of a geostatistical approach with groundwater flow modeling for three-dimensional estimation of a contaminant plume. J Contam Hydrol 251:104097. https://doi.org/10.1016/j.jconhyd.2022.104097
Troldborg M, Nowak W, Lange IV, Santos MC, Binning PJ, Bjerg PL (2012) Application of Bayesian geostatistics for evaluation of mass discharge uncertainty at contaminated sites. Water Resour Res 48(9):W09535. https://doi.org/10.1029/2011WR011785
Vu MT, Jardani A (2022) Mapping of hydraulic transmissivity field from inversion of tracer test data using convolutional neural networks. CNN-2T. J Hydrol 606:127443. https://doi.org/10.1016/j.jhydrol.2022.127443
Wagner BJ (1992) Simultaneous parameter estimation and contaminant source characterization for coupled groundwater flow and contaminant transport modelling. J Hydrol 135:275–303. https://doi.org/10.1016/0022-1694(92)90092-A
Wang X, Jardani A, Jourde H (2017) A hybrid inverse method for hydraulic tomography in fractured and karstic media. J Hydrol 551:29–46. https://doi.org/10.1016/j.jhydrol.2017.05.051
Wiedemeier TH, Rifai HS, Newell CJ, Wilson JT (1999) Natural attenuation of fuels and chlorinated solvents in the subsurface. Wiley, New York
Woodbury A, Sudicky E, Ulrych TJ, Ludwig R (1998) Three-dimensional plume source reconstruction using minimum relative entropy inversion. J Contam Hydrol 32:131–158. https://doi.org/10.1016/S0169-7722(97)00088-0
Xu M, Eckstein Y (1995) Use of weighted least-squares method in evaluation of the relationship between dispersivity and field scale. Ground Water 33(6):905–908. https://doi.org/10.1111/j.1745-6584.1995.tb00035.x
Xu T, Gomez-Hernandez JJ (2018) Simultaneous identification of a contaminant source and hydraulic conductivity via the restart normal-score ensemble Kalman filter. Adv Water Resour 112:106–123. https://doi.org/10.1016/j.advwatres.2017.12.011
Xu T, Gomez-Hernandez JJ, Chen Z, Lu C (2021) A comparison between ES-MDA and restart EnKF for the purpose of the simultaneous identification of a contaminant source and hydraulic conductivity. J Hydrol 595:125681. https://doi.org/10.1016/j.jhydrol.2020.125681
Zanini A, Kitanidis PK (2009) Geostatistical inversing for large-contrast transmissivity fields. Stochas Environ Res Risk Assess 23(5):565–577. https://doi.org/10.1007/s00477-008-0241-7
Zha Y, Yeh TCJ, Illman WA, Onoe H, Mok CMW, Wen JC, Huang SY, Wang W (2017) Incorporating geologic information into hydraulic tomography: a general framework based on geostatistical approach. Water Resour Res 53(4):2850–2876. https://doi.org/10.1002/2016WR019185

Zha Y, Yeh TCJ, Illman WA, Zeng W, Zhang Y, Sun F, Shi L (2018) A reduced-order successive linear estimator for geostatistical inversion and its application in hydraulic tomography. Water Resour Res 54(3):1616–1632. https://doi.org/10.1002/2017WR021884