



Exploration of nonlinear parallel heterogeneous reaction pathways through Bayesian variable selection

Ryosuke X. Oyanagi^{1,a} , Tatsu Kuwatani^{1,2} , and Toshiaki Omori^{3,4}

¹ Research Institute for Marine Geodynamics, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokosuka 236-0061, Japan

² PRESTO, Japan Science and Technology Agency (JST), Kawaguchi 332-0012, Japan

³ Graduate School of Engineering, Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe 657-8501, Japan

⁴ Center for Mathematical and Data Sciences, Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe 657-8501, Japan

Received 19 August 2020 / Accepted 15 January 2021 / Published online 2 February 2021
© The Author(s) 2021

Abstract. Inversion is a key method for extracting nonlinear dynamics governed by heterogeneous reaction that occur in parallel in the natural sciences. Therefore, in this study, we propose a Bayesian statistical framework to determine the active reaction pathways using only the noisy observable spatial distribution of the solid phase. In this method, active reaction pathways were explored using a Widely Applicable Bayesian Information Criterion (WBIC), which is used to select models within the framework of Bayesian inference. Plausible reaction mechanisms were determined by maximizing the posterior distribution. This conditional probability is obtained through Markov chain Monte Carlo simulations. The efficiency of the proposed method is then determined using simulated spatial data of the solid phase. The results show that active reaction pathways can be identified from the redundant candidates of reaction pathways. After these redundant reaction pathways were excluded, the controlling factor of the reaction dynamics was estimated with high accuracy.

1 Introduction

Nonlinear reaction dynamics in natural sciences are governed by reaction mechanism variables, such as reaction rate constants, diffusion coefficients, and reaction pathways [1, 2]. Therefore, understanding the nonlinear dynamics of reaction and mechanism are important for scientific theory and ensuring precise simulations in Earth sciences. Many of the important reactions in natural sciences are heterogeneous, i.e., the reaction occurs at the interface between two or more phases, which are typically minerals and water [1, 3]. Understanding the mechanism of heterogeneous reactions is important for multiple disciplines, including environmental science, geochemistry, hydrogeology, oceanography, nuclear waste disposal, and soil science. Inversion is a common way of understanding the mechanism of heterogeneous reactions; however, the complexity and nonlinearity of heterogeneous reactions, which typically involve sequences of several reactions occurring in parallel, makes inversion difficult.

Many laboratory experiments have attempted to understand the mechanism of heterogeneous reactions. These experiments typically involve quantitative observations of either (1) the temporal evolution of the concentrations of aqueous species [4–7], or (2) the spatial distribution of minerals or concentrations of aqueous

species at a single time [8–11]. The reaction rate constant in a reaction pathway can be determined inversely from the former by analyzing the temporal evolution of the solution chemistry and equilibration. However, it is impossible to determine the reaction rate and reaction pathways from the latter because the stability of a mineral cannot be determined without fluid composition data. Therefore, it is important to develop a versatile method that can extract nonlinear dynamics directly from a spatial dataset.

The number of model parameters to be estimated is an important factor for inversion analysis. In heterogeneous parallel reactions, many possible reaction pathways are present, and many reaction rate constants need to be estimated. However, in general, the extent of fitting increases with the number of model parameters used for fitting. This phenomenon is well known as overfitting, which is a very common concept in the field of machine learning [12–14]. When overfitting occurs, the estimated model parameters are substantially affected by noise, which can reduce the predictive ability of the model. Thus, to estimate reliable reaction rate constants by inversion, reaction rate constants of unimportant reaction pathways must be identified and excluded from redundant candidates of reaction pathways.

In this study, we propose a framework for exploring the reaction pathways of nonlinear and parallel heterogeneous reactions from the spatial distribution of minerals alone. We used the concept of Bayesian estima-

^a e-mail: royanagi@jamstec.go.jp (corresponding author)

tion from machine learning [12, 13]. The concept has been applied to various fields including physics [15–20], brain science [21], astronomy [22], and Earth sciences [23]. Notably, the Bayesian estimation has been applied to identify reaction pathways in the fields of biology and chemical engineering [24–30]. In these fields, the estimation of reaction rates from noisy observations is the technically important and challenging topic. Biochemical reaction networks are characterized by nonlinear systems that can be changed by spatio-temporal scale. These characteristics are similar to reactions in the Earth sciences where heterogeneous reactions between rock and solution are coupled with diffusive and/or advective transport of solution. Therefore, the Bayesian estimation can be an effective way to identify heterogeneous reaction pathways in the Earth sciences.

Under Bayesian estimation, a statistical model can be selected by evaluating the Bayesian free energy, and the important variables can be identified among redundant variables from a set of candidate models and given data [12]. Typically, huge computational costs are required to estimate the Bayesian free energy. Therefore, the Schwarz information criterion [31] that approximates Bayesian free energy has been used, but its usage is validated only if the statistical model is regular (i.e., the posterior distribution can be approximated by normal distribution). Recently, an alternative approach that gives an approximate value of the Bayesian free energy and can be used for both regular and singular statistical models has been proposed [32]. This approach has made it significantly easier to apply Bayes' theorem for nonlinear model selection problems. Using the framework proposed in this study, we successfully extract important reaction pathways of heterogeneous reactions using only the observable spatial distribution of minerals after the reaction.

This paper is organized as follows. In Sect. 2, we introduce the method for estimating the reaction rate constants from the observed spatial mineral distribution. We first present a forward modeling approach for an observed spatial distribution of solids. Bayesian inference is employed to construct the posterior distribution, i.e., the conditional probability of the model parameters for an observed spatial distribution of phases. In Sect. 3, we validate the proposed estimation method by determining the reaction mechanism from a synthetic spatial distribution of solid phases. Section 4 presents the discussion and summary.

2 Method for estimating Model parameters

In this study, we analyze the spatio-temporal dynamics of a nonlinear heterogeneous reaction. To reflect prospected experimental constraints, we suppose a situation in which temporal changes are not observable and spatial changes at a single time are observable. Figure 1 shows a conceptual model of parallel heterogeneous reactions. In this conceptual model, three minerals (M_a , M_b , and M_c) are present at the initial time

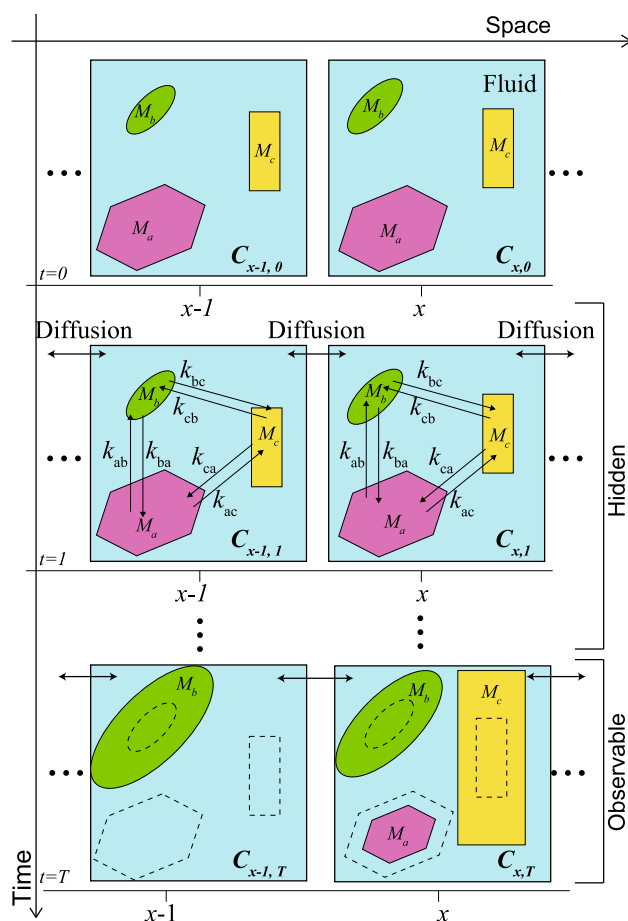
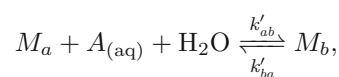


Fig. 1 Conceptual model of parallel heterogeneous reactions in the presence of three minerals (M_a , M_b , and M_c) in discretized time t and space x for a finite difference approach. At each grid point, a set of partial differential equations conserving the masses of the chemical components in the system is solved, considering diffusion and heterogeneous reactions. Heterogeneous mineral-water reactions are included in the model by using reaction rates that depend on the solution composition at each x and t ($C_{x,t}$). Active reaction pathways are determined from the local variable $C_{x,t}$. At the observable time $t = T$, the type and amount of mineral present varies with x because the reaction pathways differ according to $C_{x,t}$

$t = 0$. These minerals can react and either be consumed or produced by each other. For example, one reaction can be written in the following generalized form:



where $A_{(aq)}$ is an aqueous species in the solution, and k'_{ab} and k'_{ba} (s^{-1}) are the rate coefficients of the reaction pathway for M_b after M_a and M_a after M_b , respectively. Other reactions can be written similarly; in this heterogeneous reaction network, the total number of reaction pathways L is given by $L = N_m(N_m - 1)$ where N_m is the total number of minerals. Notably, the value of

k'_l is non-zero when the reaction pathway l is active but zero when the reaction pathway l is inactive. This concept is used for identifying important reaction pathways, as described later. Whether the reaction pathways are active depends on the local variable of concentration of $A_{(aq)}$ (C (mol cm⁻³); Fig. 1). The concentration C varies in space x' and time t' according to the reaction and the diffusion of $A_{(aq)}$. The amount of mineral i ($m^{(i)}$ (mol cm⁻³)), varies with x' and t' because it is a function of C .

To reflect realistic experimental environments, C is set as an unobservable value. The observable values in this study are the spatial variation of the amount of minerals and the volume fraction of the solution (porosity) at the start ($t' = 0$ in Fig. 1) and end of the reaction ($t' = T$ in Fig. 1). Note that discretized space and time step (x, t) are used in Fig. 1 rather than continuous space and time (x', t')

2.1 Forward modeling of mineral-fluid interactions using a reaction-diffusion model

The spatio-temporal evolution of $C_{x',t'}$ can be modeled using the mass-conservation equation of aqueous species [1];

$$\frac{\partial(\varphi C)}{\partial t'} = D \frac{\partial}{\partial x'} \left(\varphi^2 \frac{\partial C}{\partial x'} \right) - \sum_{l=1}^L \frac{\partial \Xi^l(C)}{\partial t'}, \quad (1)$$

where $\varphi(-)$ is porosity, D (cm²/s) is the constant of diffusivity for aqueous species, and $\Xi^l(C)$ (mol cm⁻³) is the loss or gain density of aqueous species according to the reaction pathway l . The first term on the right side of Eq. (1) expresses the concentration change due to diffusion of the aqueous species, whereas the second term expresses the change due to heterogeneous reactions. The rate of the heterogeneous reaction in a reaction pathway l , $\partial \Xi^l(C)/\partial t'$ (mol cm⁻³ s⁻¹), can be written as a linear equation as a function of concentration, as follows:

$$\frac{\partial \Xi^l(C)}{\partial t'} = k'_l(C_{(eq)}^l - C), \quad (2)$$

where $C_{(eq)}^l$ (mol cm⁻³) and k'_l (s⁻¹) are the constant of equilibrium concentration and the rate constant for a reaction pathway l , respectively. From the rate of heterogeneous reactions, the rate of increase or decrease in the amount of mineral i can be written as

$$\frac{\partial m^{(i)}}{\partial t'} = \varphi \sum_{l=1}^L v_i^l \frac{\partial \Xi^l(C)}{\partial t'}, \quad (3)$$

where $m^{(i)}$ (mol cm⁻³) and v_i^l are the amount of mineral i and stoichiometric coefficient of mineral i relative to aqueous species in the reaction pathway l , respectively. φ can be calculated from one minus the sum of the volume fraction of all minerals, as follows:

$$\varphi = 1 - \sum_{i=1}^{N_m} \bar{V}_i m^{(i)}, \quad (4)$$

where \bar{V}_i (cm³ mol⁻¹) is a constant representing the molar volume of mineral i . When Eqs. (1)–(3) are solved, x' and t' need to be discretized. The discretization of Eq. (1) is described in the Appendix.

As a result of discretization, the concentration ($C_{x,t}$), amount of mineral i ($m_{x,t}^{(i)}$), and porosity ($\varphi_{x,t}$) for a discretized space x ($\in \{1, \dots, X\}$) and time t ($\in \{1, \dots, T\}$) are obtained through forward modeling by inputting the following unknown parameter vector Θ' :

$$\Theta' = \{k'_1, k'_2, \dots, k'_L\}. \quad (5)$$

For a given parameter set Θ' , the theoretical amount of all observation series (i.e., minerals and porosity) at the observable target time T with dimensions of $J = N_m + 1$ is

$$\begin{aligned} \mathbf{y}_x &= \{y_x^{(1)}, \dots, y_x^{(J-1)}, y_x^{(J)}\} \\ &= \{m_{x,T}^{(1)}, \dots, m_{x,T}^{(N_m)}, \varphi_{x,T}\}. \end{aligned} \quad (6)$$

The output for observation series j , $y_x^{\text{ex}(j)}$, is the sum of the response function for the corresponding input $y_x^{(j)}$ and the observation noise $\epsilon_x^{(j)}$; that is,

$$y_x^{\text{ex}(j)} = y_x^{(j)} + \epsilon_x^{(j)}. \quad (7)$$

At this point, we assume that the observation noise follows a Gaussian distribution with a mean of zero and a standard deviation of σ_j :

$$p(y_x^{\text{ex}(j)} | y_x^{(j)}) = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(y_x^{\text{ex}(j)} - y_x^{(j)})^2}{2\sigma_j^2}\right), \quad (8)$$

for $x \in X_{\text{obs}}$, where X_{obs} is a set of observable spatial points. Thus, in the forward modeling, the conditional probability of the observed spatial distribution of the mineral and porosity using the model parameters Θ' is written as

$$\begin{aligned} p(\mathbf{y}_x^{\text{ex}} | \mathbf{y}_x) &= p(\mathbf{y}_x^{\text{ex}} | \Theta') \\ &= \prod_{j=1}^J \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(y_x^{\text{ex}(j)} - y_x^{(j)})^2}{2\sigma_j^2}\right), \end{aligned} \quad (9)$$

where $\mathbf{y}_x^{\text{ex}} = \{y_x^{\text{ex}(1)}, \dots, y_x^{\text{ex}(J)}\}$ is the set of observed series at position x . This formula can be used to obtain the probabilistic estimation of the mineral distribution after reaction as the conditional probability for the given model parameters. This can then be compared with experimental data.

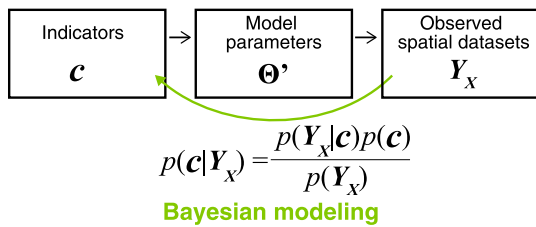


Fig. 2 Schematic of the Bayesian modeling procedure. The model parameter ($\Theta' = \{k'_1, \dots, k'_L\}$) is generated from the indicator vector \mathbf{c} . The observed data set is generated from the model parameters. To identify the active reaction pathways, $p(\mathbf{c}|\mathbf{Y}_X)$ is calculated using Bayes' theorem

2.2 Exhaustive search method for exploring reaction pathways

This section describes the method for achieving the first goal of this study, i.e., to identify the active reaction pathways. After the active reaction pathways are identified, the value of the rate constants are estimated (goal 2) in Section II C.

We can determine whether a reaction pathway l is active or inactive using the k'_l value of a reaction pathway l . Thus, Θ' can be formulated as

$$\begin{aligned} \Theta' &= \mathbf{c} \circ \Theta \\ &= \{c_1 k_1, c_2 k_2, \dots, c_L k_L\}, \end{aligned} \quad (10)$$

where k_l is the intrinsic rate constant of the reaction pathways, the symbol \circ represents the Hadamard product, and $\Theta = \{k_1, \dots, k_L\}$. \mathbf{c} is an L dimensional binary vector, defined as

$$\mathbf{c} = (c_1, c_2, \dots, c_L) \in \{0, 1\}^L. \quad (11)$$

Each variable c_l is 0 or 1; $c_l = 1$ if the l th variable belongs to the combination and is used for inversion. In contrast, $c_l = 0$ if it does not. The term \mathbf{c} is used to represent the indicators (Fig. 2).

This formulation is key for clarifying one aspect of the program setting; i.e., identifying the active reaction pathways. The best \mathbf{c} for modeling an objective variable is determined by minimizing the Bayesian free energy \mathcal{F} using the exhaustive search (ES) method, which is the simplest variable selection method [33, 34]. In the ES method, all combinations of used and unused variables are exhaustively searched, which requires estimation of 2^L combinations of variable. With the ES method, the $\mathcal{F}(\mathbf{c})$ is calculated for all combinations of explanatory variables; the combination that minimizes the $\mathcal{F}(\mathbf{c})$ is determined as the optimal combination.

Using Bayes' theorem, the posterior probability can be expressed as proportional to the likelihood function and the prior probability, as follows:

$$p(\mathbf{c}|\mathbf{Y}_X) = \frac{p(\mathbf{Y}_X|\mathbf{c})p(\mathbf{c})}{p(\mathbf{Y}_X)} \propto p(\mathbf{Y}_X|\mathbf{c}). \quad (12)$$

Here, $\mathbf{Y}_X = \{\mathbf{y}_1^{\text{ex}}, \dots, \mathbf{y}_X^{\text{ex}}\}$ and X is the number of spatial data points in the single observation series. We assume a continuous uniform prior probability $p(\mathbf{c})$, where the posterior distribution is proportional to a marginalized likelihood function defined as follows:

$$p(\mathbf{Y}_X|\mathbf{c}) = \int p(\mathbf{Y}_X|\mathbf{c}, \Theta)p(\Theta|\mathbf{c})d\Theta. \quad (13)$$

In addition, we assume that previous quantitative measurements of the minerals and porosity do not affect the present measurement; thus,

$$\begin{aligned} p(\mathbf{Y}_X|\mathbf{c}, \Theta) &= \prod_{x=1}^X p(\mathbf{y}_x^{\text{ex}}|\mathbf{c}, \Theta) \\ &= \prod_{x=1}^X \prod_{j=1}^J \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(y_x^{\text{ex}(j)} - y_x^{(j)})^2}{2\sigma_j^2}\right). \end{aligned} \quad (14)$$

Because reaction rate constants must be positive and may vary by orders of magnitude, as for $p(\Theta|\mathbf{c})$, we assume that the case of $c_l = 1$, $p(k_l|c_l = 1)$ has a continuous uniform distribution in logarithmic space, as follows:

$$p(k_l|c_l = 1) = \begin{cases} \frac{1}{k_l(\ln(k_l^{(\max)})) - \ln(k_l^{(\min)})} & \text{for } x \in [k_l^{(\min)}, k_l^{(\max)}] \\ 0 & \text{otherwise} \end{cases} \quad (15)$$

where $k_l^{(\max)}$ and $k_l^{(\min)}$ are the maximum and minimum values of the rate coefficient for the reaction pathway l within the range of searched values. In the case of $c_l = 0$, $p(k_l|c_l = 0)$ is assumed to be the delta function, as follows:

$$p(k_l|c_l = 0) = \delta(k_l) \quad (l = 1, \dots, L). \quad (16)$$

The negative logarithm of the marginalized likelihood is termed the Bayesian free energy \mathcal{F} , which is defined as

$$\mathcal{F} = -\ln p(\mathbf{Y}_X|\mathbf{c}). \quad (17)$$

The minimization of \mathcal{F} is identical to the posterior probability maximization. However, it is difficult to analytically calculate the integral over the parameter Θ in the marginal likelihood. A well-established solution is thermodynamic integration [16] that numerically calculates \mathcal{F} using the Markov chain Monte Carlo (MCMC) method [35]; however, this method generates huge computational costs because the expectations over several distributions with different pseudo-temperatures must be calculated. In this study, we numerically calculate the WBIC for the Bayesian model selection of reaction pathways. The WBIC value gives an approximate value of \mathcal{F} [32]. This approach is expected to substantially reduce the computational costs of numerical calculation. The WBIC is defined as

$$\text{WBIC} = \frac{\int nL_n p(\mathbf{Y}_X | \mathbf{c}, \Theta)^\beta p(\Theta | \mathbf{c}) d\Theta}{\int p(\mathbf{Y}_X | \mathbf{c}, \Theta)^\beta p(\Theta | \mathbf{c}) d\Theta}, \text{ where}$$

$$\beta = \frac{1}{\ln(n)}, \tag{18}$$

where n is the total number of data points ($n = XJ$) and β is the inverse pseudo-temperature ($\beta > 0$). L_n is a negative logarithmic likelihood function, defined as

$$L_n = -\frac{1}{n} \ln p(\mathbf{Y}_X | \mathbf{c}, \Theta). \tag{19}$$

As defined in Eq. (18), WBIC is identical to the average nL_n over the posterior distribution with $\beta = 1/\ln(n)$, whose β is different from the standard Bayesian estimation of the posterior ($\beta = 1$). The MCMC method [36] is employed to obtain WBIC values. In the MCMC simulations, the candidate is generated using a Gaussian proposal density ($\mathcal{N}(0, 0.01)$). The same proposal is used for all models. The candidate is accepted with the Metropolis-type probability from the transition from Θ to Θ^* as follows:

$$w(\Theta^* | \Theta) = \min\{1, \exp[-\beta(E(\Theta^*) - E(\Theta))]\}, \tag{20}$$

where E is an energy function,

$$E(\Theta) = nL_n - \frac{1}{\beta} \ln p(\Theta | \mathbf{c}). \tag{21}$$

Using these equations, WBIC values for 2^7 combinations of \mathbf{c} were calculated using the MCMC method with the setting of $\beta = 1/\ln(n)$. The vector of active reaction pathways, i.e., the reaction pathways with non-zero rate constants, can be given by the indicator vector that minimizes the WBIC value.

$$\hat{\mathbf{c}} = \arg \min_{\mathbf{c}} \text{WBIC}(\mathbf{c}). \tag{22}$$

2.3 Estimation of model parameters

After $\hat{\mathbf{c}}$ is identified, the parameter in Θ can be estimated using Bayes' theorem, as follows:

$$p(\Theta | \mathbf{Y}_X, \hat{\mathbf{c}}) \propto p(\mathbf{Y}_X | \Theta, \hat{\mathbf{c}}) p(\Theta | \hat{\mathbf{c}}). \tag{23}$$

The likelihood function $p(\mathbf{Y}_X | \Theta, \hat{\mathbf{c}})$ is then given by

$$p(\mathbf{Y}_X | \Theta, \hat{\mathbf{c}}) = \prod_{x=1}^X p(y_x^{\text{ex}} | \Theta, \hat{\mathbf{c}})$$

$$= \prod_{x=1}^X \prod_{j=1}^J \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(y_x^{\text{ex}(j)} - y_x^{(j)})^2}{2\sigma_j^2}\right).$$

$$\tag{24}$$

Thus, the posterior distribution can be expressed as

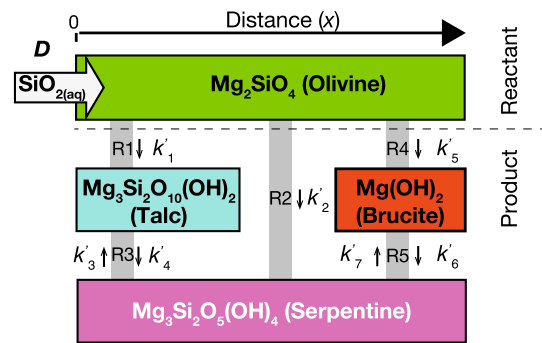


Fig. 3 Schematic illustration of the reactive transport model considered in this study [37]. At $x = 0$, $\text{SiO}_{2(\text{aq})}$ is transported by diffusion at a rate of D through the porous media of powdered olivine, and secondary minerals (talc, serpentine, and brucite) are formed. The gray bold line shows the overall reaction between the two minerals. k'_1 is the effective rate constant of the reaction pathway l

$$p(\Theta | \mathbf{Y}_X, \hat{\mathbf{c}}) \propto \exp\left[-\frac{1}{2} \sum_{j=1}^J \frac{1}{\sigma_j^2} \sum_{x=1}^X (y_x^{\text{ex}(j)} - y_x^{(j)})^2\right] p(\Theta | \hat{\mathbf{c}}), \tag{25}$$

which is obtained by the MCMC method.

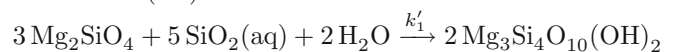
3 Validation

3.1 An example of parallel heterogeneous reactions

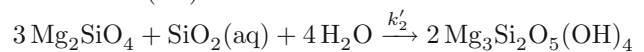
In this section, we validate the proposed method using simulated data. As an example of the potential applications of the proposed methodology, we consider coupled diffusion and reaction in the presence of olivine, quartz, and H_2O (Fig. 3). This is chosen for the following reasons: (1) The diffusional metasomatic zoning of talc and serpentine zones between quartz and olivine has been classically modeled [38–40]; (2) this heterogeneous reaction network is relatively simple, but involves the dissolution/precipitation processes of several minerals as well as element diffusion [8, 41].

In this example, heterogeneous reactions between four types of minerals ($m_{x,t}^{(i)}$, $i = 1, \dots, 4$) at 300°C and 10 MPa are considered ($N_m = 4$). At $t = 0$, the porous media of the reactant mineral, Mg_2SiO_4 (olivine), are initially present (Fig. 3). The minerals predicted to be produced from the reactant are $\text{Mg}_3\text{Si}_2\text{O}_5(\text{OH})_4$ (serpentine), $\text{Mg}_3\text{Si}_4\text{O}_{10}(\text{OH})_2$ (talc), and $\text{Mg}(\text{OH})_2$ (brucite). At $x = 0$, the concentration of $\text{SiO}_2(\text{aq})$, $C_{x,t}$, is externally buffered at a constant value. All reaction pathway candidates can be given as

Reaction 1 (R1):



Reaction 2 (R2):



Reaction 3 (R3):

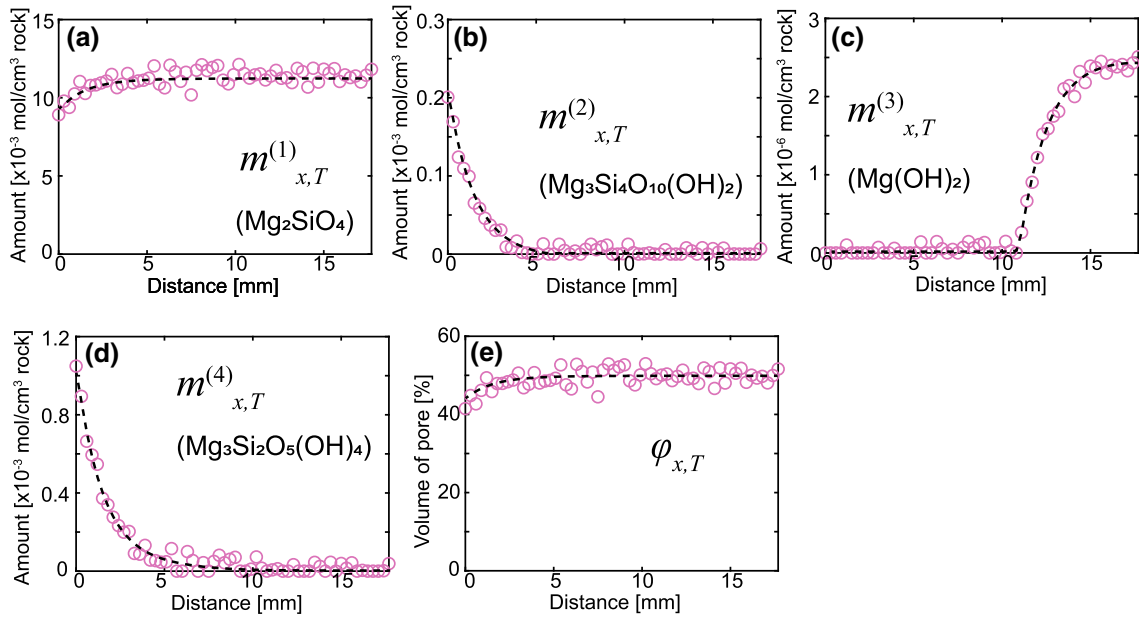
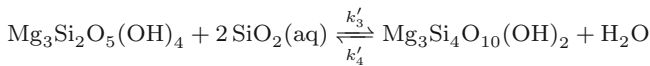
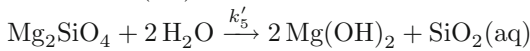


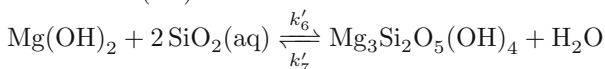
Fig. 4 Synthesized spatial distribution of minerals ($m_{x,T}^{(i)}$, $i = 1, \dots, 4$), and porosity ($\varphi_{x,T}$). The black dotted line is the theoretical spatial distribution of minerals ($m_{x,t}^{(i)}$) with artificial parameters at a certain time step T . The pink circle shows the observed spatial distribution of minerals ($y_x^{\text{ex}(i)}$) with added Gaussian noise



Reaction 4 (R4):



Reaction 5 (R5):



Here, we assume that R1, R2, and R4 are irreversible reactions (Fig. 3). Based on the stoichiometric relationship in these reaction equations, the v_i^j for Mg_2SiO_4 are given as $\{v_1^1, \dots, v_7^1\} = \{3/5, 3, 0, 0, -1, 0, 0\}$. Similarly, for $\text{Mg}_3\text{Si}_4\text{O}_{10}(\text{OH})_2$, $\{v_1^2, \dots, v_7^2\} = \{-2/5, 0, -1/2, -1/2, 0, 0, 0\}$; for $\text{Mg}(\text{OH})_2$, $\{v_1^3, \dots, v_7^3\} = \{0, 0, 0, 0, 2, 1/2, 1/2\}$; and for $\text{Mg}_3\text{Si}_2\text{O}_5(\text{OH})_4$, $\{v_1^4, \dots, v_7^4\} = \{0, -2, 1/2, 1/2, 0, -1/2, -1/2\}$. The model parameters (i.e., rate constants of each reaction pathway) can be expressed by $\Theta' = \{k'_1, k'_2, \dots, k'_7\}$ with the dimension $L = 7$.

3.2 Validation dataset

The black dotted line in Fig. 4 (a–e) shows the mineral spatial distribution, which is obtained by $\Theta' = \{10^{-3.301}, 10^{-3.301}, 0, 0, 10^{-3.301}, 0, 0\}$; i.e., k'_1, k'_2 , and k'_5 are non-zero. The equilibrium concentrations are set as $\{C_{(eq)}^1, \dots, C_{(eq)}^7\} = \{10^{-6.57}, 10^{-9.41}, 10^{-5.86}, 10^{-5.86}, 10^{-7.54}, 10^{-8.01}, 10^{-8.01}\}$ and D is set to $10^{-4.2}$ [41]. The temperature-pressure variable \bar{V}_i at 300 °C and 10 MPa was obtained from the petrological software *Perple_X* [42], which can determine physical properties of minerals thermodynamically.

We assume that the mineral spatial distribution can be observed at steps of 0.03 cm and is represented by adding Gaussian noise. To impose the same level of Gaussian noise on each observable series, σ_j was assumed to be $\sigma_j = \sigma' \times y_{\text{max}}^{(j)}$ ($j = 1, \dots, 5$), whereas σ' is a noise magnitude, and $y_{\text{max}}^{(j)}$ is maximum value of a single observation series ($y_{\text{max}}^{(j)} = \max(y_1^{(j)}, \dots, y_X^{(j)})$). The σ' is set to $10^{-1.5}$, which corresponds to 6.3% deviation, and thus $\{\sigma_1, \dots, \sigma_5\} = \{10^{-3.44}, 10^{-5.17}, 10^{-4.45}, 10^{-7.11}, 10^{0.20}\}$. After adding Gaussian noise, we set negative values of the synthesized spatial data to zero, because the amount of mineral cannot be negative. The circle in Fig. 4(a–e) shows the synthesized mineral spatial distribution. Each observed series contains 60 data points ($X = 60$) and there are five observation series ($J = 5$); thus, the total number of data points is 300 ($n = 300$). The parameters for prior information $k_i^{(\text{min})}$ and $k_i^{(\text{max})}$ are set to 10^{-8} and 10^0 , respectively. The noisy dataset is then analyzed using the proposed method to determine its effectiveness for extracting and estimating the model parameters Θ .

3.3 Validation result

10^4 samples were obtained using the MCMC method with the setting of $\beta = 1/\ln(n)$, 1000 burn-in time, and 10 thinning intervals, which were used to calculate WBIC values for 2^L candidates of \mathbf{c} . As shown in a later section (Sect. 3.5), autocorrelation of chains is generally low in each MCMC run, suggesting that chains are well mixed. Figure 5 shows the calculated value of WBIC for

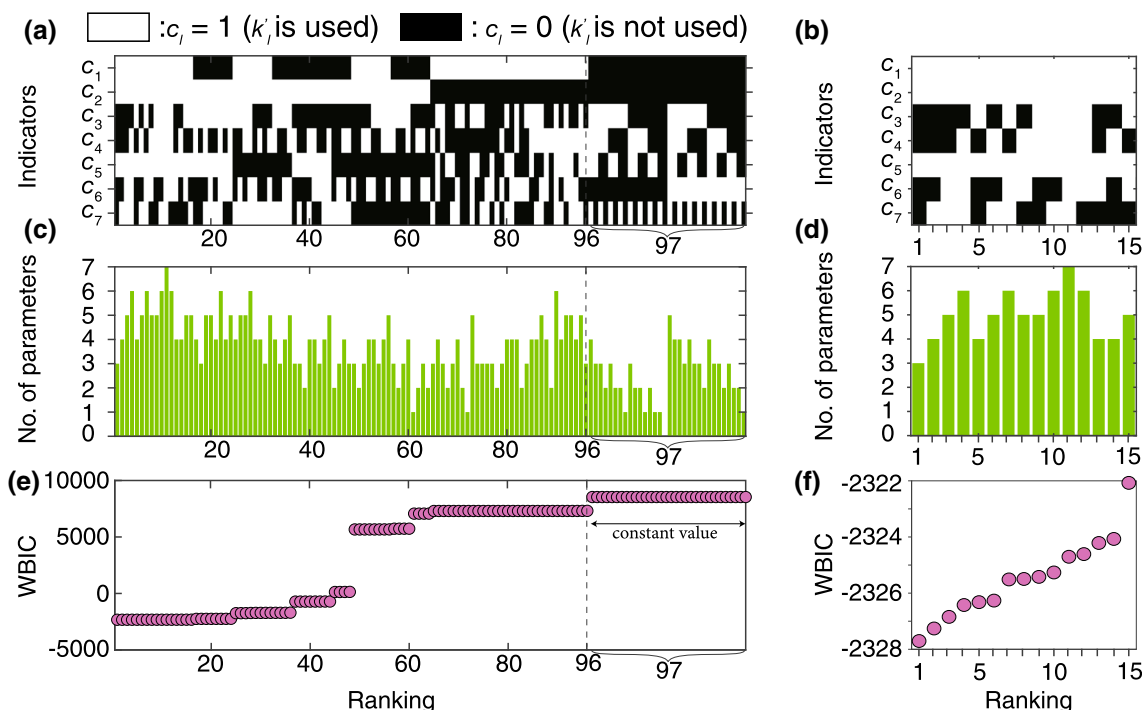


Fig. 5 **a** $2^7 (= 128)$ combinations of used ($c_l = 1$) and unused ($c_l = 0$) parameters k'_1, \dots, k'_7 . **b** A magnified image of **(a)**. **c** Number of kinetic parameters used. **d** A magnified image of **(c)**. **e** WBIC values for each combination of used and unused parameters. **f** A magnified image of **(e)**. In **(a–e)**, the horizontal axes show the ranking of WBIC values in increasing order. Among 128 combinations, 32 combinations have uniformly highest WBIC values, which are ranked at 97

2^7 combinations of **c**. Among 2^7 combinations, 96 models have different WBIC values, and WBIC increases as ranking increases from 1 to 96 (Fig. 5a–d). In contrast, another 32 models have the highest WBIC values, and these worst models are uniformly ranked at 97 (Fig. 5a–d). The ranking result shows that c_1, c_2 , and $c_5 = 1$ appeared frequently in combinations ranked at < 15 (Fig. 5a, b). In contrast, c_3, c_4, c_6 , and $c_7 = 0$ frequently ranked at < 15 (Fig. 5a, b), indicating that k'_3, k'_4, k'_6 , and k'_7 are comparably unimportant parameters. When all parameters are used ($c_l = 1, l = 1, \dots, 7$), the WBIC value is -2324.7 and ranked at 11 (Fig. 5a–d). WBIC is lowest (WBIC = -2327.4) when c_1, c_2 , and $c_5 = 1$, which is the same combination used in the input observation series. This demonstrates that the proposed method can identify the active reaction pathways from redundant candidates of reaction pathways.

For the combinations of used model parameters that minimize WBIC values (i.e., k'_1, k'_2 , and k'_5), another MCMC run was conducted to estimate the model parameters. The first 10^3 trials of MCMC sampling were not used. The auto correlation (AC) between MCMC sampling and lag time showed steep decay with increasing lag length, suggesting that MCMC sampling is less correlated and independent (Fig. 6a–c). The posterior distribution of these parameters after 10^4 Monte Carlo steps is similar to the shape of a normal distribution with a single peak (Fig. 6). The mean values of the posterior distribution for parameters $\log_{10}(k'_1), \log_{10}(k'_2)$, and $\log_{10}(k'_5)$ are $-3.299, -3.301$,

and -3.291 , respectively (Fig. 6a–c), which are close to their true values.

3.4 Dependence of model selection accuracy on spatial resolution and observation noise

Here we investigated how robustly the true model can be selected for different numbers of total observations and different levels of observation noise. With a set of n and σ' , the validations are repeatedly conducted for 10 trials. Different datasets are used for each trial.

Figure 7a–i show the result of model selection for all ten trials at each set of n and σ' . Combination of indicators ranked within 5th are shown. At $n = 300$, true models are definitely selected as the best (rank 1) models four times when noise levels are low ($\sigma' = 10^{-2.0}$; Fig. 7a). The number of times the true models were selected are decreased at high noise levels, and are five and three times among ten trials at $\sigma' = 10^{-1.5}$ (Fig. 7b) and $10^{-1.0}$ (Fig. 7c), respectively. Similarly, at $n = 100$, the number of times the true models were selected as the best model are 2, 3, and 2 times among 10 trials for $\sigma' = 10^{-2.0}$ (Fig. 7d), $10^{-1.5}$ (Fig. 7e), and $10^{-1.0}$ (Fig. 7f), respectively; At $n = 30$, the number of times the true models were selected as the best model are 2, 1, and 0 times among 10 trials for $\sigma' = 10^{-2.0}$ (Fig. 7g), $10^{-1.5}$ (Fig. 7h), and $10^{-1.0}$ (Fig. 7i), respectively.

We note that the true combination of the indicator (c_1, c_2 , and $c_5 = 1$) appeared frequently in the top 5

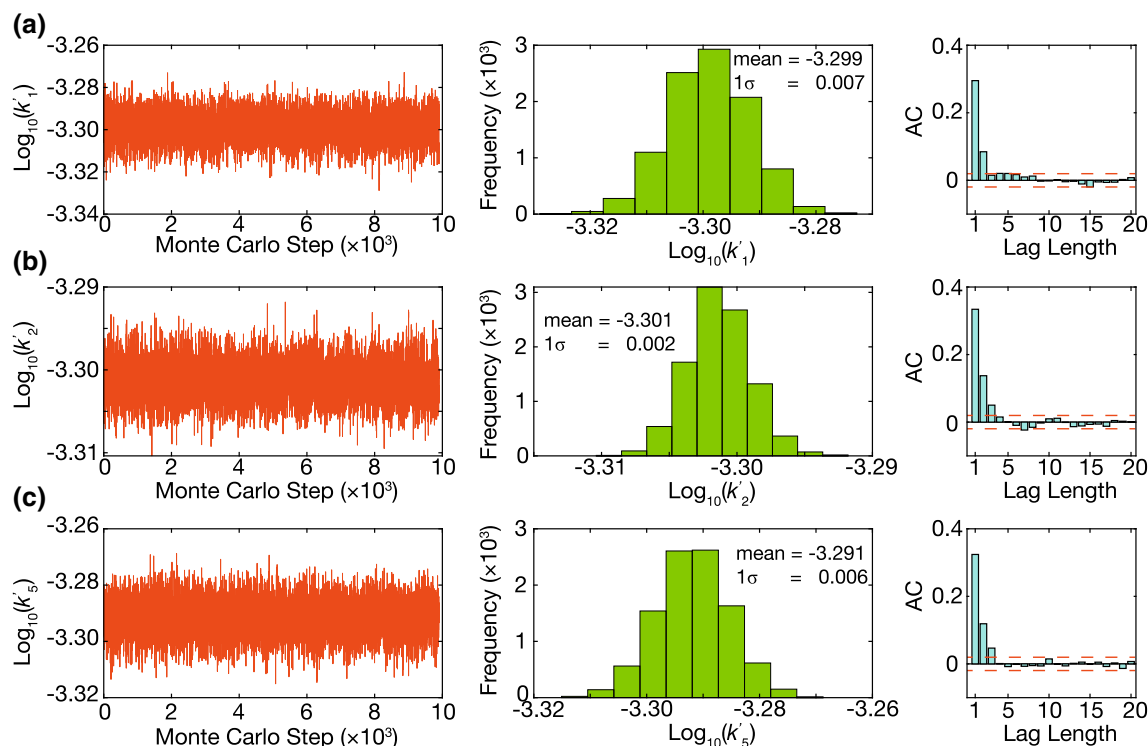


Fig. 6 Result of MCMC sampling, showing MCMC chains for 10^4 iterations, posterior distributions, and autocorrelation (AC) of MCMC sampling for parameter **a** k_1' , **b** k_2' , and **c** k_5' . The red dotted line in the AC plot is the 95 % confidence interval

models during 10 trials at any set of n and σ' (Fig. 7a–i). Moreover, the true model is generally selected in the top 5 models. For example, at $n = 300$, the true models are ranked in top 5 for 10, 8, and 7 times among 10 trials for $\sigma' = 10^{-2.0}$ (Fig. 7a–c), $10^{-1.5}$, and $10^{-1.0}$ (Fig. 7a–c), respectively.

3.5 Dependence of parameter estimation accuracy on spatial resolution and observation noise

Here we investigated how robustly the non-zero variables (k_1' , k_2' , and k_5') can be estimated for different numbers of total observations and different levels of observation noise. We evaluated the discrepancy between estimated value and true value as the normalized root mean square error (NRMSE) by

$$NRMSE = \sqrt{\frac{1}{3} \sum_{l=1,2,5} \left(\frac{k_l'^{(\text{est})} - k_l'^{(\text{true})}}{k_l'^{(\text{true})}} \right)^2},$$

where $k_l'^{(\text{est})}$ and $k_l'^{(\text{true})}$ are the posterior mean and true value of the rate constant for reaction pathway l , respectively. For each setting of the number of observations (n) and observation noise (σ'), we evaluated the NRMSE values for several trials because the NRMSE values can be affected by a random number. Then, we used one of the trials as a typical result because the calculated NRMSE results for the several trials were consistent.

Figure 8 shows that the calculated NRMSE on n and σ' . We obtained NRMSE in the range of $n = 15 - 300$ ($\log_{10}(n) = 1.17 - 2.48$) and $\sigma' = 10^{-2.0} - 10^{-1.0}$ (corresponding to 2 – 20% deviation relative to the maximum value of each observation series). Results presented in Fig. 8 suggest that the NRMSE increases with (1) decreasing total number of observation and (2) increasing noise intensity. The NRMSE shown in Fig. 6 ($\log_{10}(n) = 2.48$ ($n = 300$) and $\sigma' = 10^{-1.5}$) was $10^{-1.7}$, which corresponds to 2% error on average. Even when the number of observation was limited ($\log_{10}(n) = 1.30$ ($n = 20$)), rate constants were estimated with similar NRMSE ($10^{-1.7}$) for moderate noise levels ($\sigma' = 10^{-1.75}$), whereas it was estimated with slightly higher NRMSE ($10^{-1.2}$) for intense noise levels ($\sigma' = 10^{-1.25}$). For most regions except for $\log_{10}(n) < 1.30$ ($n < 20$) and $\sigma' > 10^{-1.50}$, the NRMSE was lower than $10^{-1.0}$, which corresponds to 10% estimation error on average (Fig. 8). The estimates were very accurate from the viewpoint of heterogeneous reaction in the Earth sciences, suggesting that our proposed method is effective against both observational sparseness and noise intensities and can be used for estimating kinetic parameters.

4 Discussion and summary

In this study, we developed a method for identifying active reaction pathways from redundant reaction pathway candidates using observed spatial data of the solid

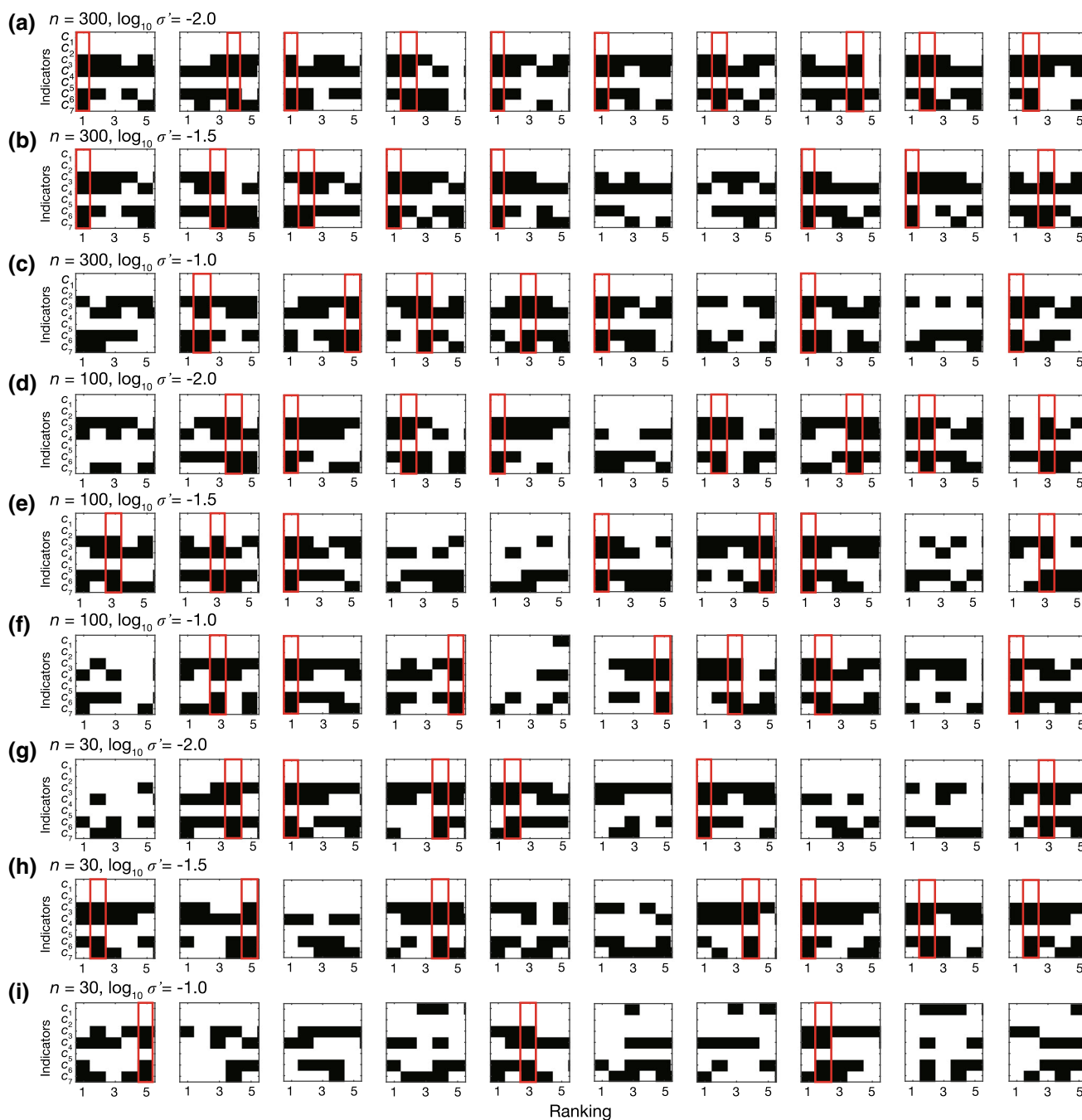


Fig. 7 Top 5 combinations of used ($c_l = 1$, white) and unused ($c_l = 0$, black) parameters among 10 trials of validation tests while changing the number of observation (n) and observation noise (σ'). The true model is marked using a red square

phase and Bayesian modeling. Verification using simulated data showed that the proposed method provides effective estimation for a given noisy dataset.

In general, it is difficult to determine the reaction pathways in nonlinear parallel heterogeneous reactions without measuring the solution chemistry. Even when the solution chemistry dataset is provided, the exploration and identification of reaction pathways involves multiple considerations due to the complexity and non-linearity of the reactions. As demonstrated during the

estimation of known kinetic parameters from the noisy spatial distribution of minerals, our method objectively identifies both the active reaction pathways and unimportant reaction pathways, instead of arbitrarily ignoring them. The proposed method can also be employed to evaluate the estimation accuracy for each unknown variable, which is important for scientific research.

In this study, the lowest WBIC value was not attained using all seven variables (-2324.7 , corresponding to a ranking of 11; Fig. 5a). This suggests that

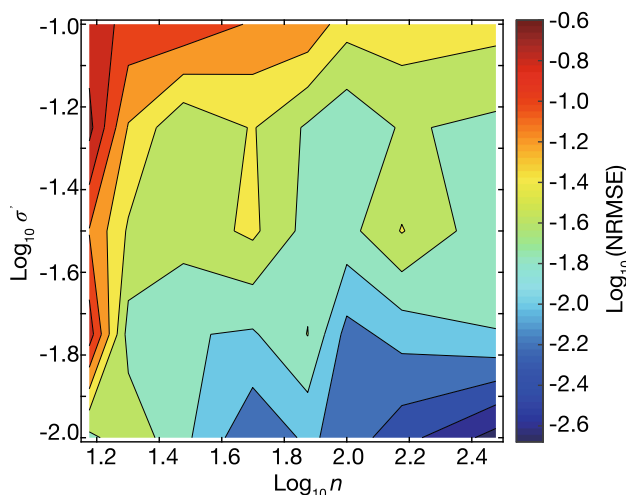


Fig. 8 Dependence of discrepancy between true and estimated parameters in the total number of observations (n) and the magnitude of observation noise (σ').

overfitting can be avoided thorough Bayesian inversion, and the estimated model parameters are not affected by noise.

For variable combinations ranked at < 15 , the indicator suggests that non-zero parameters k'_1 , k'_2 , and k'_5 always appeared in the combinations (Fig. 5a, b); however, the number of unimportant parameters (i.e., with a zero value) included in each combinations differed. For the model ranked first, no zero-value parameters were included. For the model ranked second, the indicator suggests that k'_1 , k'_2 , k'_5 , and k'_6 were included, whereas k'_1 , k'_2 , k'_3 , k'_5 , and k'_6 were included in the model ranked third (Fig. 5a, b). Because unimportant parameters were included in the models ranked second and third, the WBIC values were larger than those for the parameter combination ranked first.

Large variability was observed in WBIC values, especially between rank 48 (WBIC = 138.2) and 49 (WBIC = 5665.7; Fig. 5e). The indicators suggest that the former used three variables: k'_2 , k'_6 , and k'_7 (Fig. 5a), whereas the latter used two variables: k'_1 and k'_2 (Fig. 5). When the variable combinations ranked 48 and 49 were compared, the number of variables was larger for the lower-ranked model. This suggests that the extent of fitting led to large differences in the WBIC values between the models ranked at 48 and 49.

It is important to note the characteristics of the 32 worst models identified by highest WBIC value (Fig. 5a–e). Among these worst models, all of the models without any parameters (i.e., all rate parameters set to zero) and with less than five parameter are similarly selected as worst models. This is because reactions that are necessary for another reaction to take place are restricted by these models. Common features in these worst models is that k'_1 and k'_2 are not used (Fig. 5a). Without these reaction pathways, the combination of variables become worst case, even when the important reaction k'_5 is used. These results suggests that reac-

tion pathway k'_1 and k'_2 are necessary to drive another reaction, such as k'_5 . We suggest that identifying the common features of variables in the worst models may also be helpful to understand the sequences of reactions.

Although it is difficult to analytically derive Bayesian free energy \mathcal{F} in general, there are several methods to calculate \mathcal{F} [16], such as thermodynamic integration [43], nested sampling [44], and the non-equilibrium Monte Carlo method [45]. In this study, we used an information criterion, WBIC, that may provide a computationally efficient approximation of the \mathcal{F} . The \mathcal{F} can be approximated by a Bayesian information criteria (BIC) [31] only when the statistical model is regular, i.e., the posterior distribution can be approximated as the normal distribution [31, 32]. In the example described in this study, the shape of the posterior distributions obtained by MCMC method are similar to those of the normal distribution (Fig. 6). To investigate the effectiveness of BIC, the same validation numerical experiments were conducted. As a result, we found that the BIC also identified active reaction pathways in the case of the validation test shown in the present study. However, the usage of BIC cannot be validated without sampling because the shape of the posterior distribution cannot be constrained prior to sampling. In contrast, the WBIC, which is a generalized version of the BIC, can be used in both regular and singular statistical models and regardless of the shape of the posterior distribution [32]. Analyzing real heterogeneous reactions could require a solution for a singular model because heterogeneous reactions between minerals and fluids can be complex. Thus, WBIC as a generalized version of BIC would be the preferred information criterion for selecting models via Bayesian inference.

In this study, the noise variances σ_j for each observation series are treated as known constants. In the realistic situation, noise variance can be estimated as a hyperparameter by Bayesian estimation. However, such hyperparameter estimation is not necessary because approximate values for noise variance are known for the heterogeneous reaction between minerals and fluid.

In this study, spatial data points at a single time are used as observable data (n). As we showed in Fig. 7, the true model is likely difficult to be selected when $n = 30$ with a corresponding space interval of 0.3 cm (Fig. 7g–i). Because the spatial observation step can be 0.01 cm at minimum, n may not be less than 100, and an extreme situation in which only 30 observation data points can be used is unrealistic. However, even in the extreme situation, the parameters used (c_1 , c_2 , and $c_5 = 1$) in the true model frequently appeared in the top five models, suggesting that the true models can be interpreted by checking common indicators in the high ranked models.

The Bayesian approach has been applied to infer kinetic reaction pathways in the fields of chemical engineering and biochemistry [26, 29]. In these works, the Bayes factor, which quantifies the support for one model over another, was used to select models. The Bayes factor between two hypothesized models is obtained by calculating the difference of \mathcal{F} between the two models.

Such a calculation of \mathcal{F} can be substituted by WBIC to reduce computational cost.

Although the effectiveness of our approach was demonstrated by analyzing a synthesized dataset, the limits of its applicability must be noted. It is easy to imagine that the exhaustive search method used in this study will become intractable for a large number of variables (L). Therefore, to reduce the computational load, a relaxation approach such as the least absolute shrinkage and selection operator (LASSO) method [21, 46] using an l_1 -norm regularization term could be used when the ES method is computationally infeasible. However, exhaustive search (ES) is necessary for the strict selection of efficient variables [33, 34, 47]. This computational explosion problem associated with the ES method can be improved through the development of another relaxation method.

The proposed approach serves as a basic inversion framework for extracting active reaction pathways during parallel heterogeneous reactions. The applications of our methods to real data will be considered separately. The development of analytical methods for heterogeneous reactions and applications to a real dataset will provide a better understanding of important and dynamic Earth processes.

We acknowledge valuable comments from Kenji Nagata, Kenta Yoshida, Atsushi Okamoto, and members of the study group “Analyses of high-dimensional data in Earth science using machine learning” held in the Earthquake Research Institute at the University of Tokyo. We appreciate Takuya Ishibashi and Kazuki Yoshida for their help to improve the efficiency of calculations. This work was supported by the Japan Society for the Promotion of Science KAKENHI (JP18J01649, JP19K14827, and JP15KK0010), the Cooperative Research Program of the Earthquake Research Institute, University of Tokyo (ERI JURP 2018-B-01), the Japan Science and Technology Agency (JST) PRESTO (JPMJPR1676), and JST CREST (JPMJCR1761 and JPMJCR1914).

Author contributions

RXO conducted design of study, experiments, and drafting the manuscript. TK and TO contributed to the research idea development and revised the manuscript critically for important intellectual content. All authors gave their final approval of the manuscript version to be submitted the authors’ original work, has not received prior publication and is not under consideration for publication elsewhere.

Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Authors’ comment: Data are available from the corresponding authors upon request.]

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits

use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

A Discretization of partial differential equation

The partial differential equation (Eq. 1) must be discretized to solve. The left side of Eq. (1) can be discretized with respect to time to obtain the difference equation

$$\frac{\partial(\varphi C)}{\partial t'} = \frac{\varphi_{x,t+1}C_{x,t+1} - \varphi_{x,t}C_{x,t}}{\Delta t}, \tag{26}$$

where t denotes a time step and Δt is a time interval used for time discretization. The first term on right side of Eq. (1) can be discretized with respect to space to obtain the difference equation

$$D \frac{\partial}{\partial x'} \left(\varphi^2 \frac{\partial C}{\partial x'} \right) = D \left[\Phi_e \frac{C_{x+1,t} - C_{x,t}}{\Delta x^2} - \Phi_w \frac{C_{x,t} - C_{x-1,t}}{\Delta x^2} \right], \tag{27}$$

where

$$\Phi_e = \frac{2\varphi_{x,t} \varphi_{x+1,t}}{\varphi_{x,t} + \varphi_{x+1,t}}, \quad \Phi_w = \frac{2\varphi_{x-1,t} \varphi_{x,t}}{\varphi_{x-1,t} + \varphi_{x,t}}, \tag{28}$$

and Δx is a space interval used for space discretization. Combining the previous discretized equations, we obtain the following difference equation:

$$\begin{aligned} & \frac{\varphi_{x,t+1}C_{x,t+1} - \varphi_{x,t}C_{x,t}}{\Delta t} \\ & = D \left[\Phi_e \frac{C_{x+1,t} - C_{x,t}}{\Delta x^2} - \Phi_w \frac{C_{x,t} - C_{x-1,t}}{\Delta x^2} \right] \\ & \quad - \sum_l^L k_l'(C_{(eq)}^l - C_{x,t}). \end{aligned} \tag{29}$$

We note that for Eq. (29) to be solved, the parameter $\varphi_{x,t}$ must be known. Therefore, Eqs. (2)–(4) were first calculated to obtain $\varphi_{x,t+1}$. Using Eq. (29) and $\varphi_{x,t+1}$, the parameter $C_{x,t+1}$ was constrained.

References

1. A.C. Lasaga, *Kinetic Theory in the Earth Sciences* (Princeton University Press, 1998), ISBN 0-691-03748-5
2. C.I. Steefel, *In kinetics of water-rock interaction* (Springer, New York, 2008), pp. 545–589

3. J.D. Rimstidt, *Geochemical Rate Models* (Cambridge University Press, Cambridge, 2013), ISBN 9781139342773, <http://ebooks.cambridge.org/ref/id/CBO9781139342773>
4. J.D. Rimstidt, H.L. Barnes, *Geochim. Cosmochim. Acta* **44**, 1683 (1980)
5. T. Omori, T. Kuwatani, A. Okamoto, K. Hukushima, *Phys. Rev. E* **94**, 033305 (2016)
6. A. Okamoto, Y. Ogasawara, Y. Ogawa, N. Tsuchiya, *Chem. Geol.* **289**, 245 (2011)
7. C. Zhu, P. Lu, Z. Zheng, J. Ganor, *Geochim. Cosmochim. Acta* **74**, 3963 (2010)
8. R. Oyanagi, A. Okamoto, N. Hirano, N. Tsuchiya, *Earth Planet. Sci. Lett.* **425**, 44 (2015)
9. K. Maher, C.I. Steefel, D.J. DePaolo, B.E. Viani, *Geochim. Cosmochim. Acta* **70**, 337 (2006)
10. K. Maher, C.I. Steefel, A.F. White, D.A. Stonestrom, *Geochim. Cosmochim. Acta* **73**, 2804 (2009)
11. R. Abart, E. Petrishcheva, F.D. Fischer, J. Svoboda, *Am. J. Sci.* **309**, 114 (2009)
12. C. Bishop, *Pattern recognition and machine learning* (Springer, Verlag New York, 2006)
13. I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning* (The MIT press, 2016), <http://www.deeplearningbook.org>
14. J. Lever, M. Krzywinski, N. Altman, *Nat. Methods* **13**, 703 (2016)
15. Q. Yang, C. Sing-Long, E. Reed, *Chaos: an interdisciplinary. Journal of Nonlinear Science* **30**, 53122 (2020)
16. U. Von Toussaint, *Rev. Modern Phys.* **83**, 943 (2011)
17. R. Tamura, K. Hukushima, *Phys. Rev. B* **95**, 1 (2017)
18. M. Meier, R. Preuss, V. Dose, *New J. Phys.* **5**, 133 (2003)
19. U.V. Toussaint, R. Fischer, K. Krieger, V. Dose, *New J. Phys.* **1**, 11 (1999)
20. IEEE Nuclear Science Symposium And Medical Imaging Conference **1**, 52425 (2008)
21. S. Otsuka, T. Omori, *Neural Networks* **109**, 137 (2019)
22. K. Nagata, S. Sugita, M. Okada, *Neural Networks* **28**, 82 (2012)
23. T. Kuwatani, H. Nagao, S.I. Ito, A. Okamoto, K. Yoshida, T. Okudaira, *Phys. Rev. E* **98**, 043311 (2018)
24. S. Matera, W.F. Schneider, A. Heyden, A. Savara, *ACS Catal.* **9**, 6624 (2019). <https://doi.org/10.1021/acscatal.9b01234>
25. N. Pullen, R.J. Morris, *PLoS One* **9**, 1 (2014)
26. V. Vyshemirsky, M.A. Girolami, *Bioinformatics* **24**, 833 (2008)
27. P. Loskot, K. Atitey, L. Mihaylova, *Front. Genetics* **10**, 549 (2019)
28. N. Galagali, Y.M. Marzouk, *Chem. Eng. Sci.* **123**, 170 (2015)
29. T.R. Xu, V. Vyshemirsky, A. Gormand, A. von Kriegsheim, M. Girolami, G.S. Baillie, D. Ketley, A.J. Dunlop, G. Milligan, M.D. Houslay et al., *Sci. Signal.* **3**, ra20 (2010)
30. D. Schnoerr, G. Sanguinetti, R. Grima, *J. Phys. A: Math. Theor.* **50**, 093001 (2017)
31. G. Schwarz, *Annal. Stat.* **6**, 461 (1978)
32. S. Watanabe, *J. Mach. Learn. Res.* **14**, 867 (2013)
33. K. Nagata, J. Kitazono, S. Nakajima, S. Eifuku, R. Tamura, M. Okada, *IPJS Online Trans.* **8**, 25 (2015)
34. Y. Igarashi, K. Nagata, T. Kuwatani, T. Omori, Y. Nakanishi-Ohno, M. Okada, *J. Phys.: Conf. Ser.* **699**, (2016)
35. Y. Ogata, *Ann. Inst. Stat. Math.* **42**, 403 (1990)
36. N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, *J. Chem. Phys.* **21**, 1087 (1953), 5744249209
37. R. Oyanagi, A. Okamoto, N. Tsuchiya, *Minerals* **8**, 579 (2018)
38. P.C. Lichtner, *Geochim. Cosmochim. Acta* **52**, 143 (1988)
39. J.D. Frantz, H.K. Mao, *Am. J. Sci.* **276**, 817 (1976)
40. D.S. Korzhinskii, *Miner. Deposita* **3**, 222 (1968)
41. R. Oyanagi, A. Okamoto, N. Tsuchiya, *Geochim. Cosmochim. Acta* **270**, 21 (2020)
42. J. Connolly, *Geochem. Geophys. Geosyst.* **10** (2009)
43. Y. Ogata, *Numer. Math.* **55**, 137 (1989)
44. J. Skilling et al., *Bayesian Anal.* **1**, 833 (2006)
45. H. Ahlers, A. Engel, *Eur. Phys. J. B* **62**, 357 (2008)
46. R. Tibshirani, *J. Roy. Stat. Soc.: Ser. B (Methodol.)* **58**, 267 (1996)
47. T. Nakayama, Y. Igarashi, K. Sodeyama, M. Okada, *Chem. Phys. Lett.* **731**, 136622 (2019)