Adjoint-based marker-in-cell data assimilation for constraining thermal and flow processes from Lagrangian particle records

Atsushi Nakao[1], Tatsu Kuwatani[2], Shin-ichi Ito[3], and Hiromichi Nagao[4]

[1]Akita University
[2]Japan Agency for Marine-Earth Science and Technology
[3]The University of Tokyo
[4]Earthquake Research Institute, The University of Tokyo

May 28, 2024

Abstract

Geophysical problems often involve Lagrangian particles that follow surrounding flows and record information about the system, such as the pressure and temperature path recorded in metamorphic rocks. These Lagrangian particles can be useful for constraining unknown parameters, such as their sources and the thermal and flow processes of the surrounding fluid. To use information about Lagrangian particles to constrain unknown parameters about the surrounding fluid in an inverse manner, we have developed a 4D-Var (four-dimensional variational) data assimilation for thermal convection in a particle-grid coupled system. Here we consider particles advected in a thermally convecting, highly viscous fluid that mimics geochemical tracers in the Earth's mantle, and estimate time series of thermal and velocity fields only from the particle records, focusing on their high traceability in the laminar flow. We present preliminary 4D-Var results using a sufficient amount of synthetic particle position and velocity data. The 4D-Var run achieves a 60-Myr time reversal of thermal convection with a horizontal wavelength of 6,000 km, without using any temperature data. For smaller scale convection, the cost function tends not to decrease well, but with a shorter retrospective time domain or a large weight on early stage information, the reconstruction improves. While this work focuses on mantle dynamics, our framework has the potential to constrain thermal, flow, and mixing processes in any other laminar flow containing Lagrangian particles that record useful information.

Adjoint-based marker-in-cell data assimilation for constraining thermal and flow processes from Lagrangian particle records

Atsushi Nakao[1,2], Tatsu Kuwatani[2], Shin-ichi Ito[3,4], and Hiromichi Nagao[3,4]

[1]Graduate School of Engineering Science, Akita University, Akita, Japan

[2]Research Institute for Marine Geodynamics, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokosuka, Japan

[3]Earthquake Research Institute, The University of Tokyo, Tokyo, Japan

[4]Graduate School of Information Science and Technology, The University of Tokyo, Tokyo, Japan

Key Points:

- We have developed a new inversion scheme for a grid-particle system for retrospective estimation of thermal and flow processes
- The method has successfully reconstructed thermal convection from synthetic particle positions and velocities without using temperature data
- Our method highlights the advantages of particle traceability in laminar flow in the context of data assimilation

Corresponding author: Atsushi Nakao, atsushi.nakao@gmail.com or a-nakao@gipc.akita-u.ac.jp

Abstract

Geophysical problems often involve Lagrangian particles that follow surrounding flows and record information about the system, such as the pressure and temperature path recorded in metamorphic rocks. These Lagrangian particles can be useful for constraining unknown parameters, such as their sources and the thermal and flow processes of the surrounding fluid. To use information about Lagrangian particles to constrain unknown parameters about the surrounding fluid in an inverse manner, we have developed a 4D-Var (four-dimensional variational) data assimilation for thermal convection in a particle-grid coupled system. Here we consider particles advected in a thermally convecting, highly viscous fluid that mimics geochemical tracers in the Earth's mantle, and estimate time series of thermal and velocity fields only from the particle records, focusing on their high traceability in the laminar flow. We present preliminary 4D-Var results using a sufficient amount of synthetic particle position and velocity data. The 4D-Var run achieves a 60-Myr time reversal of thermal convection with a horizontal wavelength of 6,000 km, without using any temperature data. For smaller scale convection, the cost function tends not to decrease well, but with a shorter retrospective time domain or a large weight on early stage information, the reconstruction improves. While this work focuses on mantle dynamics, our framework has the potential to constrain thermal, flow, and mixing processes in any other laminar flow containing Lagrangian particles that record useful information.

Plain Language Summary

We study how tiny particles carried by moving fluids deep inside the Earth can tell us a lot about processes that we cannot observe directly. These particles, like pieces of rock from deep underground, travel with the flow and collect valuable information about their environment, such as temperature and pressure. We have developed a method to better understand the hidden motions and properties of these fluids, such as how heat circulates in the Earth, by looking backward in time. Our approach uses computer simulations to track how these particles have moved, allowing us to reconstruct the fluid's movements and properties. We tested our method on a simple model and were able to map the flow of heat and matter inside the Earth back 60 million years. This approach could help scientists better understand not only the Earth's dynamics, but also those of other fluid systems that carry similar particles.

1 Introduction

Problems in the geosciences often involve "Lagrangian" particles that passively follow the "Eulerian" flow of the surrounding fluid and sometimes record essential information about the thermal, flow, and chemical processes of the surrounding fluid. For example, metamorphic rocks that we can sample today have been flowing as part of the Earth's mantle convection for tens of millions of years, and their mineral and chemical compositions reflect the P–T (pressure and temperature) of the surrounding environment (Kuwatani, Nagata, et al., 2018; Kuwatani, Nagao, et al., 2018). In addition, rock structures may reflect the surrounding flow and stress (Iwamori, 2003). Similarly, the chemical and mineral composition of volcanic ash can reflect the cooling process of the plume that carried the ash (Matsumoto & Nakamura, 2017), and the distribution of volcanic ash deposits reflects the surrounding wind speed and the strength of the eruption (Shimizu et al., 2019). Similarly, pumice floating on the ocean erupted from a submarine volcano is a key to estimating ocean temperature and flow processes by analyzing the growth conditions of shells on the pumice (Watanabe et al., in revision). These examples underscore the central role of Lagrangian particles in deciphering Earth's past environmental conditions.

Although it is very difficult to track individual Lagrangian particles in high-Re (Reynolds number) fluids, such as volcanic ash in a volcanic plume and pumice floating on the ocean, particle tracking can be easier for low-Re fluids, such as rocks derived from the mantle. This is due to the laminar flow of the mantle (Turcotte & Schubert, 2002) and the very small diffusion coefficient of the elements that make up the mantle and crustal minerals under high P–T conditions (e.g., Fei et al., 2013). Based on this background, a so-called marker-in-cell (or marker-and-cell) method has been widely used in computational fluid dynamics for geodynamic modeling (e.g., Gerya & Yuen, 2003). The method computes both a grid and particles simultaneously; the velocity fields are solved in the grid system, and non-diffusive particles passively advect following the velocity field of the surrounding fluid. Thus, this method can compute advection of compositional heterogeneity smaller than the resolution of the grid system without numerical diffusion. Previous studies have used marker-in-cell methods to simulate the transport of crustal and volatile components during plate subduction processes (e.g., Nakao et al., 2016, 2018), and to track the P–T paths of crustal rocks in response to tectonic events such as continental collision and plate subduction (e.g. Warren et al., 2008; Chowdhury et al., 2021).

In previous studies, particle tracking using the marker-in-cell method has been dominated by forward simulations; therefore, even though rock data may contain information about the temperature, pressure, and flow of the surrounding environment, there has been no framework to date that uses data from mantle-derived rocks as a quantitative constraint on the surrounding environment or convection processes as an inverse problem. In other words, there are few data assimilation frameworks in the geodynamics community that consider both a grid and particles in a system to exploit information from mantle-derived rocks, as will be explained below.

Let us look at data assimilation schemes and their application to geodynamics. There are two main approaches to data assimilation: sequential methods and 4D-Var (four-dimensional variational) methods (e.g., Kalnay, 2003). On the one hand, sequential data assimilation, including the Kalman filter, is a mathematical technique used in forecasting to incorporate new observations into a model's current state estimate. This process updates predictions by combining observed data with prior knowledge, optimizing accuracy in real-time applications such as weather forecasting and navigation systems. On the other hand, 4D-Var, also known as adjoint-based data assimilation, is used to refine model predictions by optimizing initial conditions. In this method, the gradient of a cost function representing the difference between model output and observations is calculated by backward integration. For problems in geodynamics, the unknown parameters are the thermal, chemical, and flow structures of the Earth's deep interior over billions of years, none of which can be observed in real time. Geophysical observations such as seismic, electromagnetic, and gravity are useful for estimating unknown parameters of the present-day crust and mantle (e.g., Iwamori et al., 2021; Kuwatani et al., 2023), but not over geologic time. Therefore, it is difficult to apply sequential data assimilation to geodynamic phenomena, except for a short time scale such as post-seismic crustal deformation (Fukuda & Johnson, 2021). Thus, 4D-Var is dominant in the geodynamic community (Bunge et al., 2003; Li et al., 2017; Price & Davies, 2018; Ghelichkhan et al., 2021; Nakao et al., 2024) compared to sequential methods (Bocher et al., 2016), and this study also focuses on 4D-Var.

In each of the previous geodynamic 4D-Var data assimilations, only "Eulerian" field data are assimilated to reconstruct thermal and flow processes. In detail, previous studies often use the present-day thermal structure estimated from seismic tomography, along with time series of the Earth's surface velocity field estimated from the plate reconstruc-

tion model (e.g., Müller et al., 2016), so that a forward model develops consistently with these two Eulerian observations. Although some studies include Lagrangian particles in 4D-Var, their scheme does not use any "Lagrangian" information to constrain thermal and flow processes, but the path of Lagrangian particles is estimated based on the velocity fields optimized by Eulerian observations (Ghelichkhan et al., 2021; Nakao et al., 2024). Some studies in fluid dynamical communities use Lagrangian information to reconstruct Rayleigh-Bénard convection by using a sequential data assimilation (Bauer et al., 2022), but to our knowledge, a 4D-Var scheme that considers markers and grids in a system is very rare even in communities outside geodynamics. Thus, there are probably no previous studies that use Lagrangian information to estimate their origin or the time evolution of the surrounding Eulerian field. This study is the first to propose a theoretical framework for the 4D-Var marker-in-cell data assimilation for this purpose, i.e., retrospective estimation of thermal and flow processes based on Lagrangian information.

This paper is organized as follows. Section 2 describes the 4D-Var data assimilation scheme in order to use Lagrangian particle data based on the concept described above. Section 3 describes the generation of synthetic Lagrangian particle data to apply the 4D-Var marker-in-cell method for its validation. Section 4 describes the results of the 4D-Var simulations assimilating the synthetic data generated in section 3. Section 5 highlights the advantages of our 4D-Var methods for traceable particles and discusses the future application of the method to the real petrological data. Section 6 concludes the paper.

2 Methods

2.1 Forward equations

We considered a highly viscous, incompressible fluid in which laminar flow occurs. The fluid convects in a two-dimensional rectangular domain with a horizontal domain $x \in [x_0, x_1]$ and a vertical domain $z \in [z_0, z_1]$, where $\boldsymbol{x} = (x, z)^\top$ denotes the Cartesian coordinates (Figure 1). The fluid motion develops as a function of time $t \in [t_0, t_1]$. The driving force of the fluid is thermal buoyancy due to cooling along the upper boundary and heating along the lower boundary and, as such, the fluid flows spontaneously. The fluid contains tracer particles which are passively advected following the surrounding fluid flow. As such, we considered three governing equations: equations of motion,

energy conservation, and advection of tracer particles. The forward governing equations are written in a non-dimensional form.

The equation of motion of a highly viscous fluid flowing due to thermal buoyancy, which combines the conservation equations of mass and momentum, can be written as follows (Turcotte & Schubert, 2002):

$$\eta \nabla^4 \psi = \mathrm{Ra} \frac{\partial T}{\partial x}, \tag{1}$$

where η is the constant effective viscosity of the fluid, $\psi(\boldsymbol{x}, t)$ is the stream function, Ra is the thermal Rayleigh number (i.e., a nondimensional number representing the intensity of thermal convection), and $T(\boldsymbol{x}, t)$ is the fluid temperature. The stream function ψ is defined as follows:

$$\boldsymbol{u} = (u, w)^\top = \left(\frac{\partial \psi}{\partial z}, -\frac{\partial \psi}{\partial x} \right)^\top, \tag{2}$$

where $\boldsymbol{u}(\boldsymbol{x}, t)$ is the velocity field of the fluid. The nabla, Laplacian, and biharmonic operators are defined as follows:

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial z} \right), \tag{3}$$

$$\nabla^2 = \nabla \cdot \nabla^\top = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}, \tag{4}$$

and

$$\nabla^4 = \nabla^2 \nabla^2 = \frac{\partial^4}{\partial x^4} + 2 \frac{\partial^4}{\partial x^2 \partial z^2} + \frac{\partial^4}{\partial z^4}, \tag{5}$$

respectively. The thermal Rayleigh number Ra is defined as follows:

$$\mathrm{Ra} = \frac{\rho_0 \alpha_T \Delta T g h^3}{\eta_0 \kappa}, \tag{6}$$

where ρ_0 is the reference fluid density, α_T is the thermal expansivity, $\Delta T (= T_1 - T_0 > 0)$ is the temperature difference between the lower and upper boundaries of the fluid, g is acceleration due to gravity, h is the fluid thickness, η_0 is the reference viscosity, and κ is the thermal diffusivity. The parameters that make up Ra are constant. Equation (1) is valid for a two-dimensional domain, the Boussinesq approximation, and a highly viscous fluid with an infinite Prandtl number and constant viscosity. A free slip condition was imposed on all four boundaries of the forward equation of motion (equation 1) (Figure 1a).

We considered heat transport due to advection and conduction such that the energy conservation equation can be written as

$$\frac{\partial T}{\partial t} + \boldsymbol{u}^\top \cdot (\nabla T)^\top = k\nabla^2 T, \tag{7}$$

where k is the thermal conductivity. The fluid is cooled at T_0 along the upper boundary, heated at T_1 along the lower boundary, and insulated along the side boundaries as the boundary conditions of equation (7) (Figure 1a).

Tracer particles passively follow the surrounding velocity field as

$$\frac{d\boldsymbol{x}_i}{dt} = \boldsymbol{u}_i \tag{8}$$

where $\boldsymbol{x}_i(t) = (x_i(t), z_i(t))^\top$ is the position of particle i, and $\boldsymbol{u}_i(t) = \boldsymbol{u}(\boldsymbol{x}_i(t), t)$ is the velocity of particle i ($i = 1, ..., N$). Each particle has composition C_i that is constant throughout the model time domain:

$$C_i = \text{const.} \tag{9}$$

This assumption is valid only when the diffusion coefficient of the particles is so small that it can be ignored, as in the Earth's mantle.

2.2 Cost function

We now address the problem of how to reconstruct fluid convection in the forward model on the basis of Lagrangian particle information. To do this, we define a cost function that quantifies the misfit between the observed data and the corresponding modeled variable values as follows:

$$J = J_1 + J_2. \tag{10}$$

J_1 is the time-integrated summation of the squared error between observed and modeled particle positions for all particles, and J_2 is similarly the squared velocity error:

$$J_1 = \int_{t_0}^{t_1} dt \sum_{i=1}^{N} \frac{1}{2}\alpha \left(\boldsymbol{x}_i - \boldsymbol{x}_i^{\text{obs}}\right)^\top \left(\boldsymbol{x}_i - \boldsymbol{x}_i^{\text{obs}}\right) \tag{11}$$

$$J_2 = \int_{t_0}^{t_1} dt \sum_{i=1}^{N} \frac{1}{2}\beta \left(\boldsymbol{u}_i - \boldsymbol{u}_i^{\text{obs}}\right)^\top \left(\boldsymbol{u}_i - \boldsymbol{u}_i^{\text{obs}}\right) \tag{12}$$

where α and β are the data assimilation weights. The time dependence of the weights is considered in section 4.2. $\boldsymbol{x}_i^{\text{obs}}(t) = (x_i^{\text{obs}}(t), z_i^{\text{obs}}(t))^\top$ and $\boldsymbol{u}_i^{\text{obs}}(t) = (u_i^{\text{obs}}(t), w_i^{\text{obs}}(t))^\top$ denote the observed value of the position and velocity of particle i, respectively, i.e., $d\boldsymbol{x}_i^{\text{obs}}/dt =$

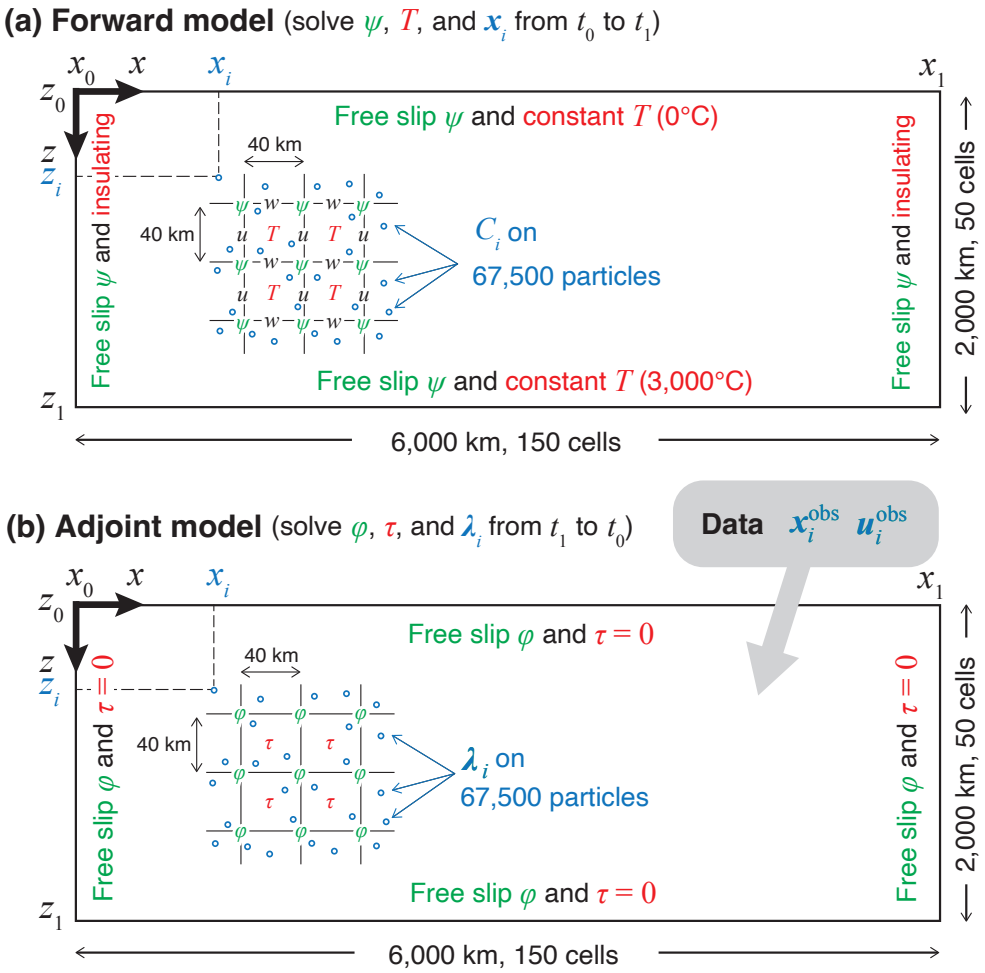

Figure 1. Model setup for solving (a) forward and (b) adjoint equations of the 4D-Var simulation. Green characters are related to the equations of motion, red characters to the equations of energy conservation, and blue characters to the tracer particles. ψ = stream function; $\boldsymbol{u} = (u, w)^\top = (\partial\psi/\partial z, -\partial\psi/\partial x)^\top$ = velocity; T = temperature; $\boldsymbol{x}_i = (x_i, z_i)$ = position of particle i; φ = adjoint stream function; τ = adjoint temperature; $\boldsymbol{\lambda}_i = (\lambda_i, \mu_i)$ = adjoint position of particle i; $\boldsymbol{x}_i^{\mathrm{obs}}$ = observed position of particle i; $\boldsymbol{u}_i^{\mathrm{obs}}$ = observed velocity of particle i; $x \in [x_0, x_1]$ = horizontal distance; $z \in [z_0, z_1]$ = vertical distance; $t \in [t_0, t_1]$ = time.

$\boldsymbol{u}_i^{\mathrm{obs}}$. If an observation of particle i is missing at time t, we replace the observed values with those in the model, i.e., $x_i^{\mathrm{obs}}(t) = x_i(t)$, $z_i^{\mathrm{obs}}(t) = z_i(t)$, $u_i^{\mathrm{obs}}(t) = u_i(t)$, or $w_i^{\mathrm{obs}}(t) = w_i(t)$, so that the missing observation is not evaluated in the cost function. We also apply this rule in the adjoint equations in the following subsection so that the adjoint model skips the variable correction when an observation is missing.

To reproduce the position and the velocity of each particle in the forward model that are consistent with the observations, we minimized J subject to the forward governing equations (i.e., equations 1, 7, and 8). In this case, the following Lagrangian function L can be defined using the Lagrange multipliers φ, τ, and $\boldsymbol{\lambda}_i$:

$$\begin{aligned} L &= J \\ &+ \int_{t_0}^{t_1} dt \int_{x_0}^{x_1} dx \int_{z_0}^{z_1} dz\, \varphi \left(\eta \nabla^4 \psi - \mathrm{Ra} \frac{\partial T}{\partial x} \right) \\ &+ \int_{t_0}^{t_1} dt \int_{x_0}^{x_1} dx \int_{z_0}^{z_1} dz\, \tau \left(\frac{\partial T}{\partial t} + \boldsymbol{u}^\top \cdot (\nabla T)^\top - k \nabla^2 T \right) \\ &+ \int_{t_0}^{t_1} dt \sum_{i=1}^{N} \boldsymbol{\lambda}_i^\top \left(\frac{d\boldsymbol{x}_i}{dt} - \boldsymbol{u}_i \right). \end{aligned} \tag{13}$$

Hereafter, we refer to $\varphi(\boldsymbol{x}, t)$ as the adjoint stream function, $\tau(\boldsymbol{x}, t)$ as the adjoint temperature, and $\boldsymbol{\lambda}_i(t) = (\lambda_i(t), \mu_i(t))^\top$ as the adjoint position for particle i. Although the adjoint temperature is widely used in previous 4D-Var geodynamical models (e.g., Bunge et al., 2003), the adjoint stream function is introduced firstly in Nakao et al. (2024) as a proxy of an adjoint velocity vector in the previous studies (e.g., Bunge et al., 2003), and the adjoint particle position is the original of this paper.

2.3 Adjoint equations

Taking variation of the Lagrangian function defined in equation (13) and comparing coefficients, we derived the following non-dimensional adjoint equations.

The coefficients of variation of ψ yields the adjoint equation of motion:

$$\eta \nabla^4 \varphi + (\nabla \tau) \times (\nabla T) = \nabla \times \sum_{i \in dS} \beta (\boldsymbol{u}_i - \boldsymbol{u}_i^{\mathrm{obs}}). \tag{14}$$

Note that the cross product of two-dimensional vectors $\boldsymbol{A} = (A_x, A_z)^\top$ and $\boldsymbol{B} = (B_x, B_z)^\top$ becomes the scalar defined as $\boldsymbol{A} \times \boldsymbol{B} = A_x B_z - B_x A_z$. $dS(\boldsymbol{x})$ in the right-hand side of equation (14) denotes a sufficiently small area around the position on the grid at which this equation is considered. This operation is required because the velocity error can only be defined at the particle positions in the present framework, i.e. it is the conversion of

the residual velocity values from the particles to the field. dS corresponds to a grid cell in the numerical calculation. The velocity correction only takes place if the cell contains the particle.

The variation of T yields the adjoint equation of energy conservation:

$$-\frac{\partial \tau}{\partial t} - \boldsymbol{u}^{\top} \cdot (\nabla \tau)^{\top} - k\nabla^2 \tau + \mathrm{Ra}\frac{\partial \varphi}{\partial x} = 0. \tag{15}$$

The variation of $\boldsymbol{x}_i$ yields the adjoint equation for advection of particle i:

$$\frac{d\boldsymbol{\lambda}_i}{dt} = \alpha \left(\boldsymbol{x}_i - \boldsymbol{x}_i^{\mathrm{obs}}\right) - \boldsymbol{\lambda}_i^{\top}(\nabla_i \boldsymbol{u}), \tag{16}$$

where $\nabla_i \boldsymbol{u} = (\partial u/\partial x_i, \partial w/\partial z_i)$.

The variation of $T(t_0)$ and $\boldsymbol{x}_i(t_0)$ yields gradient of the cost function with respect to the forward variable values at the initial condition:

$$\frac{\partial J}{\partial T(\boldsymbol{x}, t_0)} = -\tau(\boldsymbol{x}, t_0) \tag{17}$$

$$\frac{\partial J}{\partial \boldsymbol{x}_i(t_0)} = -\boldsymbol{\lambda}_i(t_0). \tag{18}$$

The variation of $T(t_1)$ and $\boldsymbol{x}_i(t_1)$ yields the terminal condition for the backward calculation:

$$\tau(\boldsymbol{x}, t_1) = 0 \tag{19}$$

$$\boldsymbol{\lambda}_i(t_1) = \boldsymbol{0}. \tag{20}$$

Note that equations (16), (18), and (20) appear for the first time in this paper in connection with the introduction of the adjoint particle position $\boldsymbol{\lambda}_i$, while equations (15), (17), and (19) are the same as in Nakao et al. (2024), which does not correct the particle positions. On the other hand, the variational calculation yields a free slip boundary condition for equation (14) and $\tau = 0$ for all four boundaries (Figure 1b), in the agreement with Nakao et al. (2024).

Equations (17) and (18) allow the temperature field and the position of each particle at the initial condition to be updated based on the values of the adjoint variables at the initial condition so that the cost function decreases. The adjoint variables at the initial condition can be obtained by solving equations (14), (15), and (16) backwards in time, starting from the terminal conditions (equations 19 and 20).

Table 1. Physical parameters used to calculate the thermal Rayleigh number and to dimensionalize the simulation results.

Symbol	Explanation	Value	Unit
c_p	Isobaric specific heat	1.2×10^3	J kg^{-1} K^{-1}
g	Gravitational acceleration	10	m s^{-2}
h	Thickness of the convective layer	2×10^6	m
ΔT	Temperature difference between upper and lower boundaries	3×10^3	K
α_T	Thermal expansivity	2.5×10^{-5}	K^{-1}
η_0	Reference viscosity	10^{22}	Pa s
κ	Thermal diffusivity	10^{-6}	m^2 s^{-1}
ρ_0	Reference density	3.9×10^3	kg m^{-3}

Table 2. Non-dimensional physical variables used in the governing equations and the dimensionalization. Physical parameters used for the dimensionalization are listed in Table 1.

Symbol	Explanation	Dimensional value	Unit of dimensional value
C_i	Composition of particle i	–	–
k	Thermal conductivity	$k \times \rho_0 c_p \kappa$	W m^{-1} K^{-1}
T	Temperature	$T \times \Delta T$	K
$\boldsymbol{u} = (u, w)^\top$	Velocity field	$\boldsymbol{u} \times \kappa / h$	m s^{-1}
$\boldsymbol{u}_i = (u_i, w_i)^\top$	Velocity of particle i	$\boldsymbol{u}_i \times \kappa / h$	m s^{-1}
$\boldsymbol{x}_i = (x_i, z_i)^\top$	Position of particle i	$\boldsymbol{x}_i \times h$	m
η	Viscosity	$\eta \times \eta_0$	Pa s
$\boldsymbol{\lambda}_i = (\lambda_i, \mu_i)^\top$	Adjoint position of particle i	–	–
τ	Adjoint temperature	–	–
φ	Adjoint stream function	–	–
ψ	Stream function	$\psi \times \kappa$	m^2 s^{-1}
$\boldsymbol{x} = (x, z)^\top$	Cartesian coordinates	$\boldsymbol{x} \times h$	m
t	Time	$t \times h^2 / \kappa$	s

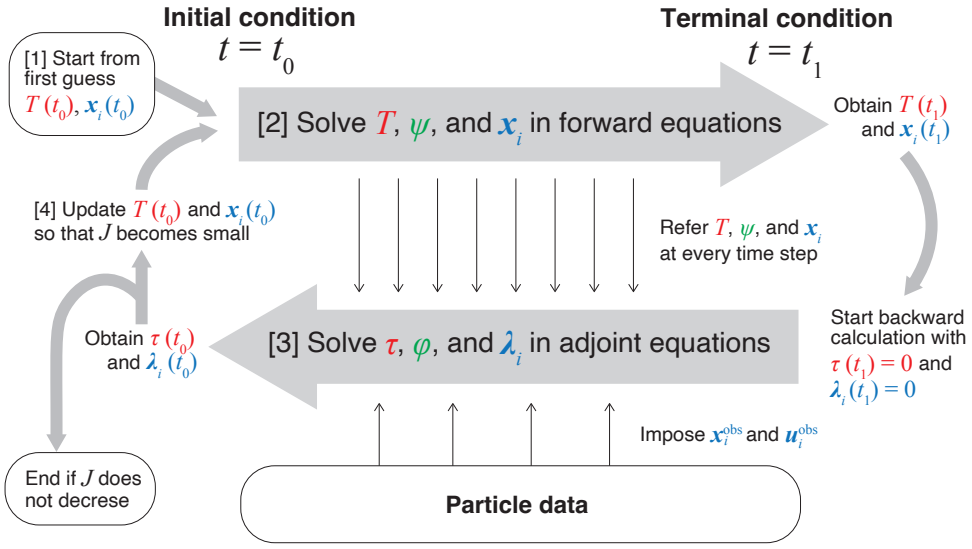

Figure 2. Flowchart of the 4D-Var simulation. J = cost function; T = temperature; ψ = stream function; $\boldsymbol{x}_i$ = position of particle i; τ = adjoint temperature; φ = adjoint stream function; $\boldsymbol{\lambda}_i$ = adjoint position of particle i; $\boldsymbol{x}_i^{\mathrm{obs}}$ = observed position of particle i; $\boldsymbol{u}_i^{\mathrm{obs}}$ = observed velocity of particle i; $t \in [t_0, t_1]$ = time. Details are provided in section 2.4.

2.4 Algorithm for data assimilation

We constructed a data assimilation algorithm as shown in Figure 2. The flow chat is based on Nakao et al. (2024), and we newly add the process to assimilate particle information to update particle positions, as well as temperature field, at the initial condition. [Process 1] Set the first guess of the variables to optimize, that is, $T(\boldsymbol{x}, t_0)$ and $\boldsymbol{x}_i(t_0)$. [Process 2] Solve the forward equations (1), (7), and (8) until $T(\boldsymbol{x}, t_1)$ and $\boldsymbol{x}_i(t_1)$ are obtained. [Process 3] Solve the adjoint equations (14), (15), and (16) backwards in time starting from t_1 to t_0. This process refers to the particle data and variables solved in process 2 at each step. [Process 4] Obtain $\tau(\boldsymbol{x}, t_0)$ and $\boldsymbol{\lambda}_i(t_0)$, and update $T(\boldsymbol{x}, t_0)$ and $\boldsymbol{x}_i(t_0)$. [Process 5] Solve the forward equations again from the updated initial conditions. These steps are repeated until J becomes sufficiently small.

2.5 Parameter values and dimensionalization

Although we have described the governing equations for the general fluid in sections 2.1 and 2.3, we now consider the fluid to be Earth's mantle by substituting the parameter values listed in Table 1, which yield $\mathrm{Ra} = 2.34\times10^6$. Using these parameters, the variables in the governing equations are converted to dimensional values as shown in Table 2. Furthermore, as shown in Figure 1, we considered a geological spatiotemporal scale such that $x_0 = 0$ and $x_1 = 3$ (0 to 6,000 km horizontally in the dimensional form), $z_0 = 0$ and $z_1 = 1$ (0 to 2,000 km vertically), $t_0 = 0$, and $t_1 h^2/\kappa = 30$, 60, 90 Myr. The temperatures of the upper and lower boundaries were set to $T_0\Delta T = 273$ K (= 0°C) and $T_1\Delta T = 3{,}273$ K (= 3,000°C), respectively. We used a constant viscosity of $\eta = 1$ (10^{22} Pa s in the dimensional form) and a constant thermal conductivity of $k = 1$ (4.68 W/m/K in the dimensional form).

2.6 Numerical schemes

The forward and adjoint equations in sections 2.1 and 2.3 are discretized in space and time using a finite volume method and a staggered grid (Figure 1). The grid is rectangular and contains 150 uniform 40-km-width cells along the horizontal axis and 50 uniform 40-km-height cells along the vertical axis. This resolution is comparable to or better than global seismic tomography (e.g., Gu et al., 2001). The variables were also discretized along the time axis with uniform 2×10^4-yr intervals. It was confirmed in advance that the discretization of the forward equations satisfies the Courant–Friedrichs–Lewy condition with sufficient buffer at each time step in the case of the parameter sets in Table 1. Although it is not obvious whether the discretized adjoint equations satisfy the condition, we have confirmed that the numerical simulation of the adjoint part does not break down. We used the solvers in Nakao et al. (2024), and an additional solver was introduced in the present model to solve the adjoint particle advection. The model domain contains 67,500 particles and their both forward and adjoint positions are updated at each time step by using the discretized equations in Appendix A.

A simple gradient descent method was used to update the temperature fields and the particle positions at $t = t_0$ in the forward model:

$$T^{n+1}(\boldsymbol{x}, t_0) = T^n(\boldsymbol{x}, t_0) + \varepsilon_T^n \tau^n(\boldsymbol{x}, t_0) \tag{21}$$

$$\boldsymbol{x}_i^{n+1}(t_0) = \boldsymbol{x}_i^n(t_0) + \varepsilon_x^n \boldsymbol{\lambda}_i^n(t_0) \tag{22}$$

Table 3. Details of reference forward experimental runs for data generation.

Run ID	Horizontal wavelength	Time window (Myr)
R1	$\sim 3h$ (6,000 km)	60
R2	$\sim 1.5h$ (3,000 km)	30, 60, 90
R3	$\sim h$ (2,000 km)	60

where index n indicates the number of iteration loops. ε_T^n and ε_x^n are the learning rates for updating the temperature field and particle positions, respectively. If the cost function becomes larger after these operations, smaller learning rates are given: $\varepsilon_T^{n+1} = 0.8\varepsilon_T^n$ and $\varepsilon_x^{n+1} = 0.8\varepsilon_x^n$, and then, the optimization restarts from the previous iteration loop. The 4D-Var optimization loop stops when J does not decrease even after updating the larning rates 10 times, or when the iteration step reaches $n = 10{,}001$ (Figure 2).

In the first iteration loop ($n = 1$), we assume a homogeneous thermal structure $T(\boldsymbol{x}, t_0) = 0.5$ and particle positions $\boldsymbol{x}_i(t_0) = \boldsymbol{x}_i^{\mathrm{obs}}(t_1)$ as the first guess. The latter setting is based on the experimental or petrological background that the chemical composition of the Lagrangian samples can be measured after the flow process.

3 Reference forward models and data sampling

To demonstrate the effectiveness of the 4D-Var data assimilation scheme described in the previous section, we generate synthetic particle data sets that can be obtained from reference forward models that mimic Earth's mantle convection. The model setup of the reference forward models is the same as that of the forward part of the 4D-Var model (Figure 1), except for the thermal structure and particle positions under the initial condition. The initial thermal structure and particle positions of the reference models are the unknown variables to be solved by the 4D-Var simulations.

Spherical harmonic analysis of the seismic velocity of the Earth's interior indicates that degree-2 structure dominates in the lowermost mantle, and degree 2 and 5 dominate in the uppermost mantle (Gu et al., 2001). This suggests that the dominant wavelength of convecting cells is ∼4,000 km (degree-5) to ∼20,000 km (degree-2) at the top of the mantle. To generate such a convective structure, we assume a mixed Gaussian func-

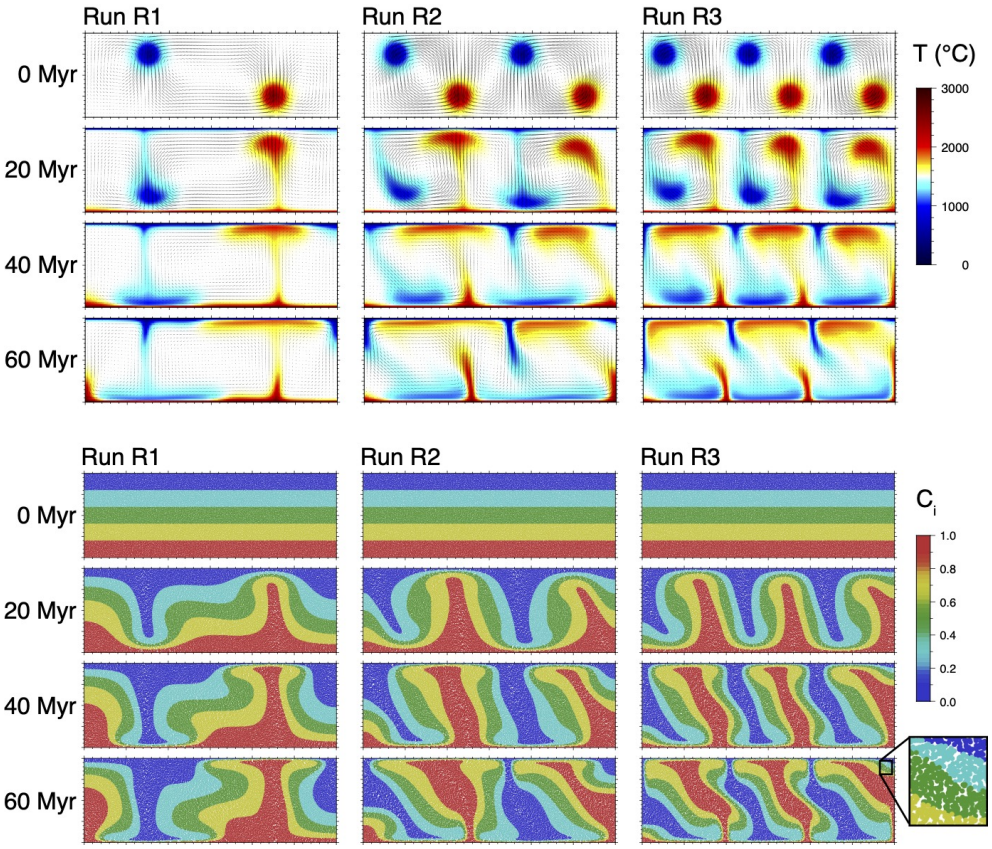

Figure 3. Snapshots of the three reference forward models R1, R2, and R3 at four time steps. Upper half panels show temperature (colored contours) and velocity (arrows). Lower half panels show the position and chemical composition of all particles.

tion in the initial thermal structure such that

$$T = 0.5 + 0.1 \sum_j \delta T_j \tag{23}$$

where δT_j is the j-th thermal anomaly

$$\delta T_j = \frac{p_j}{2\pi\sigma} \exp\left(-\frac{(\boldsymbol{x} - \bar{\boldsymbol{x}}_j)^\top (\boldsymbol{x} - \bar{\boldsymbol{x}}_j)}{2\sigma^2} \right). \tag{24}$$

$\bar{\boldsymbol{x}}_j$ is the center of j-th thermal anomaly, e.g., $\bar{\boldsymbol{x}}_1 = (0.75, 0.25)$ and $\bar{\boldsymbol{x}}_2 = (2.25, 0.75)$ for run R1. $\sigma = 0.1$ is the standard deviation of the mixed Gaussian function. p_j is the temperature polarity; $p_j = 1$ $(j = 1, 3, ...)$ or $p_j = -1$ $(j = 2, 4, ...)$. We generated three reference forward models R1, R2, and R3 with horizontal spacings of 2,000, 3,000, and 6,000 km, respectively (Table 3), which are comparable to the horizontal wavelength of the seismic velocity structure. The upper half of Figure 3 shows that the placed thermal anomalies induce the convection at the corresponding horizontal wavelength.

In addition to the thermal anomalies, nine tracer particles are placed in each grid cell at the initial condition. The positions of the particles in a cell are randomly determined. The chemical compositions of each particle, or particle IDs, are set as a function of depth so that they form a stratified structure at the initial condition for each reference model. The lower half of Figure 3 shows that the placed particles are advected following the velocity fields induced by the corresponding thermal anomalies placed at the initial conditions.

We sample the position $\boldsymbol{x}_i^{\text{obs}}$ and velocity $\boldsymbol{u}_i^{\text{obs}}$ of all particles at each time step from the reference forward runs. This is unrealistic for the geological context because the rock samples have little information about the horizontal position and are only available when they are near the Earth's surface. Rather, our sampling is similar to observing tracer particles in a convecting transparent fluid in laboratory experiments to simulate the dynamics of subducting slabs, where the particles are used to visualize the flow of the fluid (Schellart, 2004; Guillaume et al., 2021). The ideal particle dataset is used to demonstrate the validity of data assimilation for a marker-cell coupled system first proposed in this study. Future application to petrological data will be discussed in section 5.2.

4 Results

This section shows how the 4D-Var experiments reconstruct the reference forward runs in the previous section using only particle data $\boldsymbol{x}_i^{\text{obs}}$ and $\boldsymbol{u}_i^{\text{obs}}$.

Figure 4 shows the updating process of the initial temperature and particle positions when assimilating data sampled from reference run R1. A homogeneous thermal structure and the particle distributions observed in the final state of the reference forward model are used as the initial guess (first row of Figure 4). As the iteration progresses in Figure 4, the positive and negative initial temperature anomalies appear (1st column), the residual temperature and velocity fields between the 4D-Var and reference forward models become small (2nd column), the layered structure of the initial particle positions is reconstructed (3rd column), and the error distances of the particle positions between the 4D-Var and reference forward models become small (4th column).

Figure 5a shows that the cost function J of run 1 (equation 10), as well as the learning rate for updating the initial particle positions ε_x (equation 22), decreases as the iteration progresses. This shows that the temperature field and the particle positions are properly corrected so that the forward part of the 4D-Var run explains the time series of the sampled particle positions and velocity.

Figure 6 shows the evolution of the forward part of run 1 after sufficient optimization until the cost function is no longer decreasing (n = 1,333). Although the temperatures in the central parts of the thermal anomalies are not fully reproduced in the initial condition, the residual temperature and velocity are much smaller in the later stages, e.g., $\sim$100°C around the cold downwelling and hot upwelling parts and much smaller in the thermally homogeneous parts in the terminal condition (2nd column of Figure 6). As the flow structure is correctly reconstructed, the distribution of the passive particles also has good much with those in the reference model at each step, e.g., 100–200 km residual around the cold downwelling and hot upwelling parts, and several km around in the thermally heterogeneous parts (4th column of Figure 6). Thus, the 4D-Var run well reconstructs the thermal and flow fields and their time series only from the Lagrangian particle positions and velocities.

The following subsections investigate the accuracy of the reconstruction of the thermal convection as a function of the horizontal wavelength of the convective cells of the reference forward runs (section 4.1), the time dependence of the assimilation weights α and β (section 4.2), and the time window of the assimilation (section 4.3), by varying the parameters listed in Table 4.

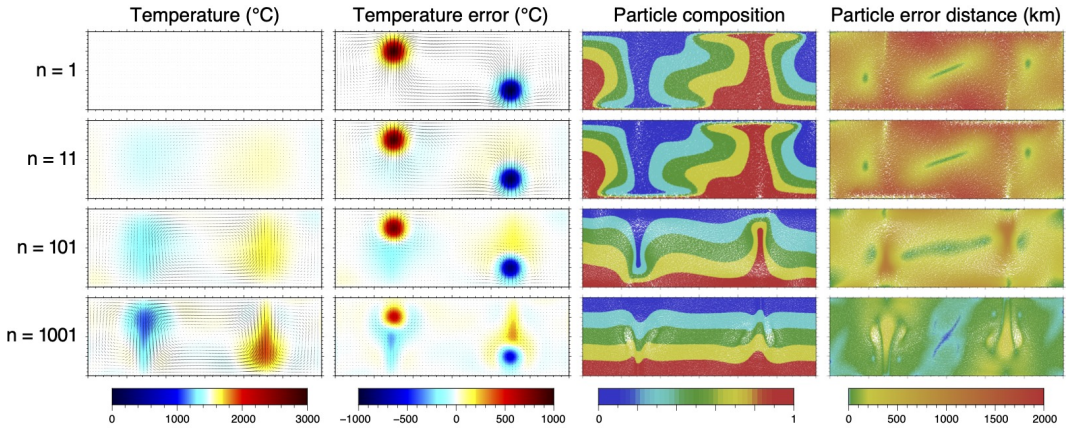

Figure 4. Values of the variables at the initial condition ($t = t_0$) of run 1 at four different iteration steps n; the lower rows of the figure represent more advanced optimization. First columns show the optimized temperature (colored contours) and velocity fields (arrows). Second columns show errors of the optimized temperature (colored contours) and velocity (arrows) fields between the 4D-Var model (run 1) and its reference model (run R1). Third columns show the optimized particle positions and compositions. Fourth columns show particle positions, and the color of the particles indicates the distance between the modeled particle locations in run 1 and the true particle positions in run R1.

Table 4. Details of 4D-Var experimental runs. [1]Assimilation weight of particle positions. [2]Assimilation weight of particle velocities.

Run ID	Data	α^1	β^2	Time window (Myr)
1	Run R1	10^7	1	60
2	Run R2	10^7	1	60
3	Run R3	10^7	1	60
4	Run R2	$10^7(t-t_0)/(t_1-t_0)$	$(t-t_0)/(t_1-t_0)$	60
5	Run R2	$10^7(t_1-t)/(t_1-t_0)$	$(t_1-t)/(t_1-t_0)$	60
6	Run R2	10^7	1	30
7	Run R2	10^7	1	90

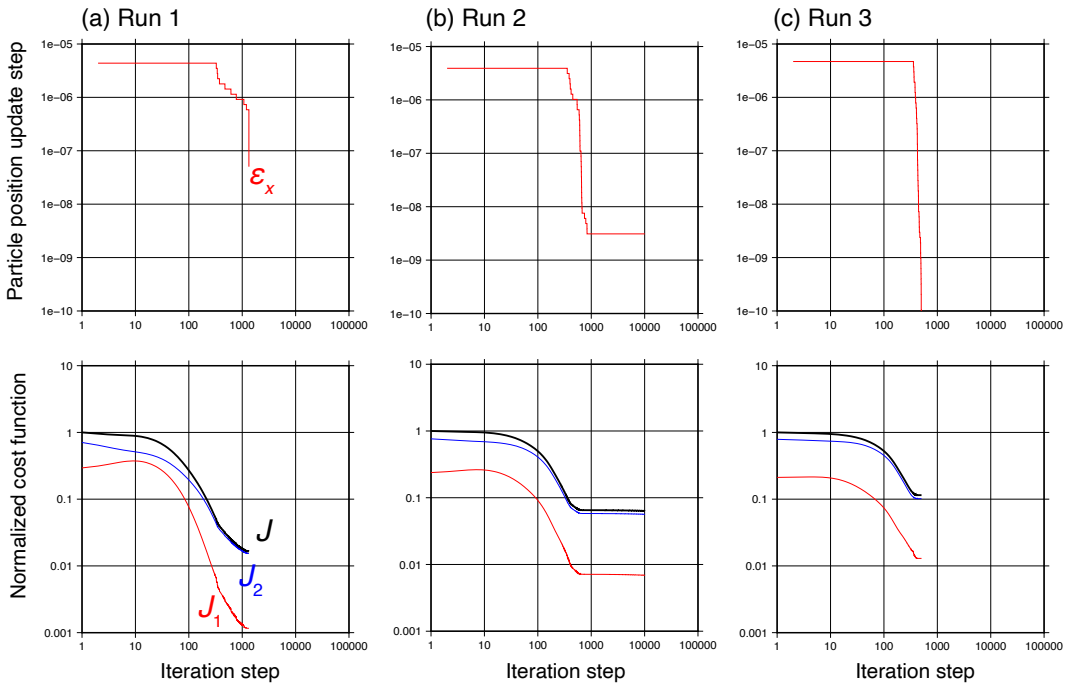

Figure 5. Graphs of cost functions as a function of optimization step n for runs 1, 2, and 3. Upper three figures show the learning rate of particle positions ε_x in equation (22). Lower three figures show cost functions J_1 (squared error of particle positions as in equation 11; thin red lines), and J_2 (squared error of particle velocities as in equation 12; thin blue lines), and J ($J_1 + J_2$ as in equation 10; thick black lines).

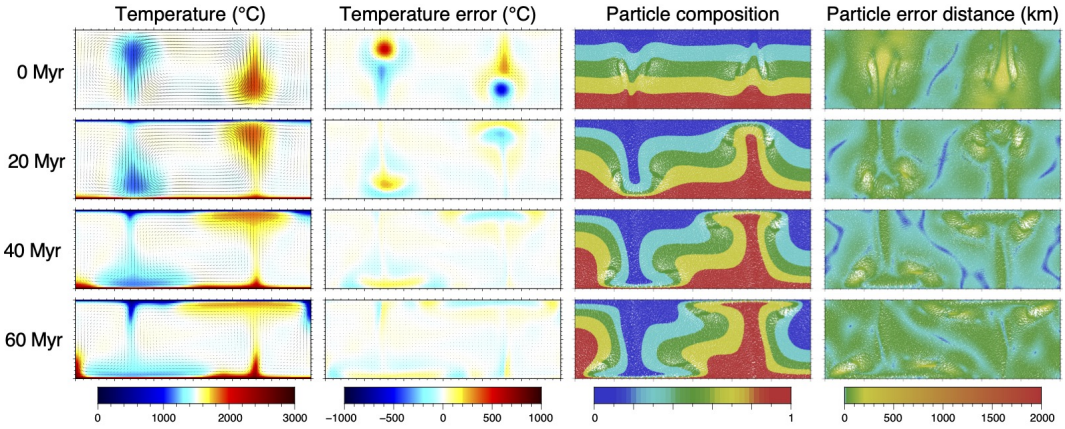

Figure 6. Results of the forward part of run 1 at iteration step n=1,333 at four different time steps; the lower rows of the figure represent more advanced time steps. First columns show the optimized temperature (colored contours) and velocity (arrows). Second columns indicate errors of the optimized temperature (colored contours) and velocity (arrows) fields between the 4D-Var model (run 1) and its reference model (run R1). Third columns show optimized particle positions and compositions. Fourth columns show particle positions, and the color of the particles indicates the distance between the modelled particle positions in run 1 and the true particle positions in run R1.

4.1 Effects of the convection pattern

This subsection investigates the effects of the horizontal wavelength of the convective cells of the reference forward models on the reconstruction of the thermal convection by comparing runs 1–3 after sufficient optimization. Similar to run 1, large temperature residuals remain between the 4D-Var and reference models around the initial thermal anomalies and the corresponding downwelling and upwelling parts in runs 2 and 3 (Figures 7 and 8). The temperature errors cause velocity errors around the corresponding downwelling and upwelling flows, resulting in the larger particle error distances around the vertical flows; e.g., Figures 7 and 8 show particle error distances of several hundred km around the thermal anomalies at the initial conditions in runs 2 and 3. Thus, as the number of the thermal anomalies placed at the initial conditions increases, the cost function tends not to decrease (Figure 5). These results suggest that the accuracy of the 4D-Var scheme decreases when reconstructing complex thermal and flow patterns with small convective cells.

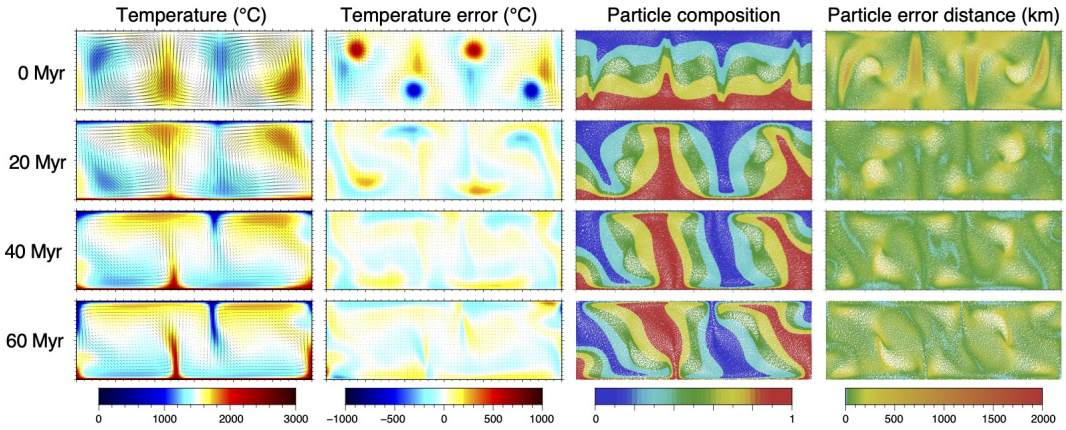

Figure 7. Same as Figure 6 but for run 2 at iteration step $n = 10,001$.

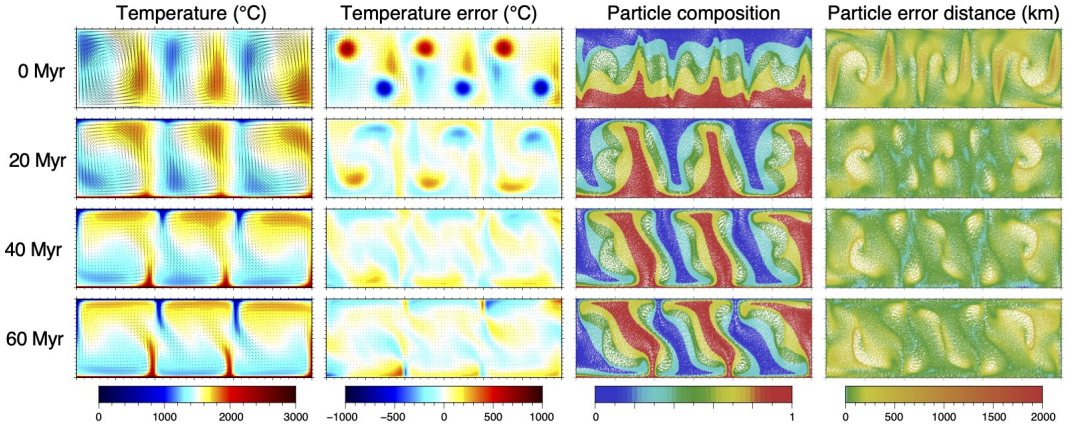

Figure 8. Same as Figure 6 but forrun 3 at iteration step $n = 505$.

4.2 Effects of time dependence of the assimilation weight

Here we examine cases where the 4D-Var scheme successfully reconstructs complex thermal and flow structures and processes due to small convective cells. This subsection examines the time dependence of the assimilation weights α and β in the cost function J (equation 10) by comparing runs 4, 2, and 5, each of which assimilates data sampled from run R2. Run 4 uses larger weights in the later stages, run 2 uses constant weights, and run 5 uses larger weights in the early stages (Table 4). The ratio α/β is constant in each run. Figure 9 shows that, among the three experiments, run 5 has the best reconstruction of thermal and flow structures and particle positions, especially in the early stages. Figure 10 shows that the cost function decreases more when larger weights are set in the early stages.

4.3 Effects of the assimilation time window

This subsection examines the time window for data assimilation (i.e., $t_1 - t_0$) by comparing runs 6, 2, and 7, which use the dataset sampled from the period of 30, 60, and 90 Myr evolution of the reference forward model R2, respectively (Table 4). The 30 Myr time window accurately reconstructs thermal convection at a wavelength of 4,000 km (run 6 of Figure 11). With the 90 Myr time window, the reconstruction of the thermal convection is poor in the early stages (run 7 of Figure 11). The thermal structure and particle distribution in run 7 are consistent with the reference run in the later stage, because the particle positions at $t = t_1$ (90 Myr) are imposed as a first guess, from which the optimization proceeds. For time windows longer than 100 Myr, numerical calculations tend to break down. Figure 12 shows that the cost function decreases more when the time window is shorter.

Here is the summary of the results. The 4D-Var run achieves a 60-Myr time reversal of the convective structure with a horizontal wavelength of 6,000 km (run 1). For thermal convection with a wavelength less than 6,000 km, the cost function tends not to decrease well (runs 2 and 3). Nevertheless, the reconstruction improves with a large weight on early-stage information (run 5) or a shorter retrospective time domain (e.g., 30 Myr for a wavelength of 4,000 km, run 6).

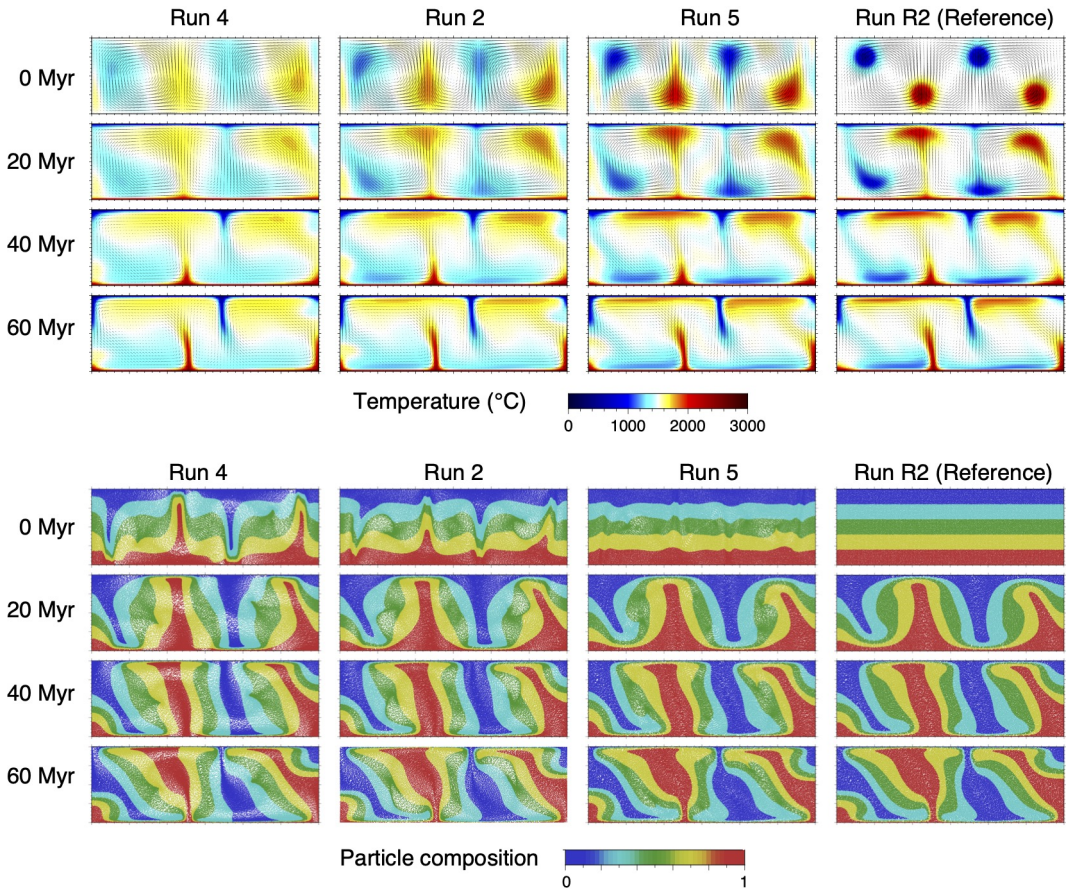

Figure 9. Snapshots of the three 4D-Var models run 4 (n = 10,001), run 2 (n = 10,001), and run 5 (n = 10,001) with different hyperparameter time dependence, at four time steps, compared to their reference forward model run 2. Upper panels show temperature (colored contours) and velocity (arrows). Lower panels show the position and chemical composition of the particles.

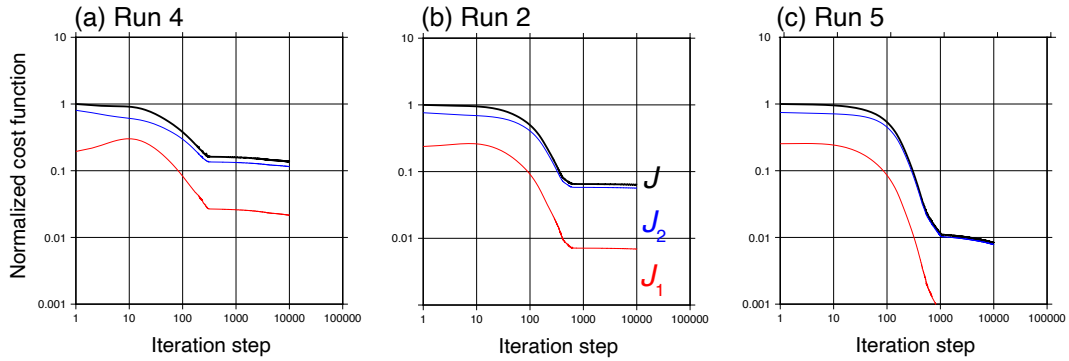

Figure 10. Same as the lower half of Figure 5 but for runs 4, 2, and 5 with different hyperparameter time dependency.

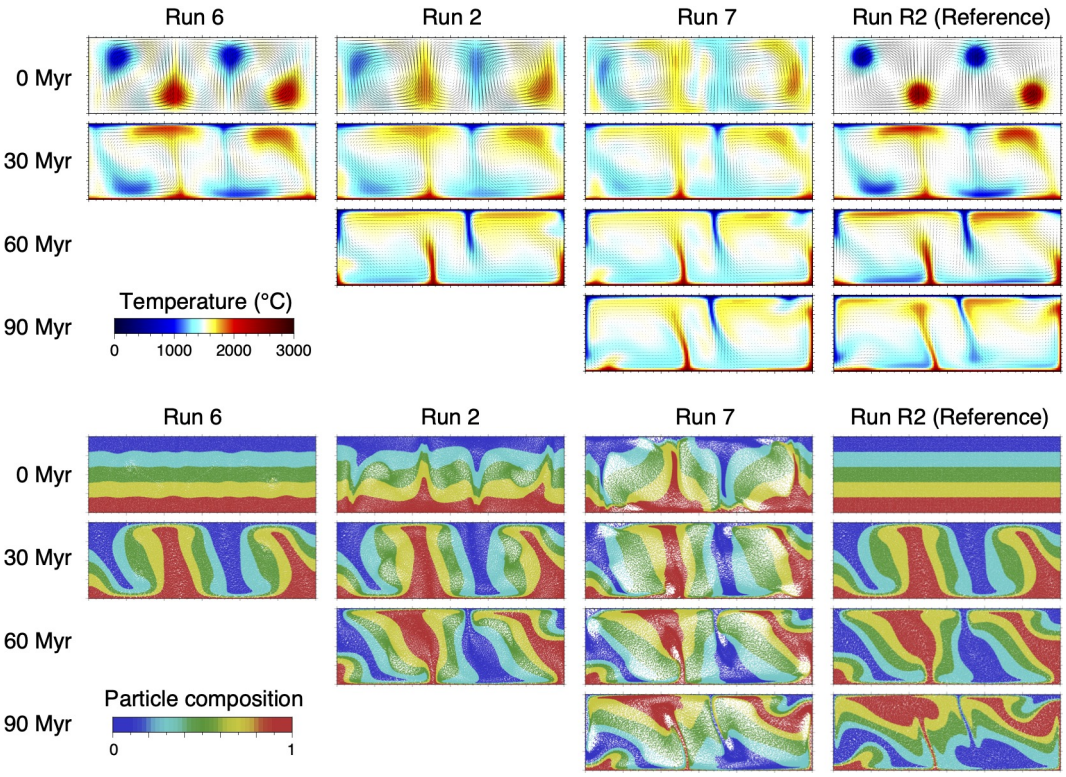

Figure 11. Same as Figure 9 but for run 6 (n = 3,356), run 2 (n = 10,001), and run 7 (n = 3,001) with different assimilation time windows.

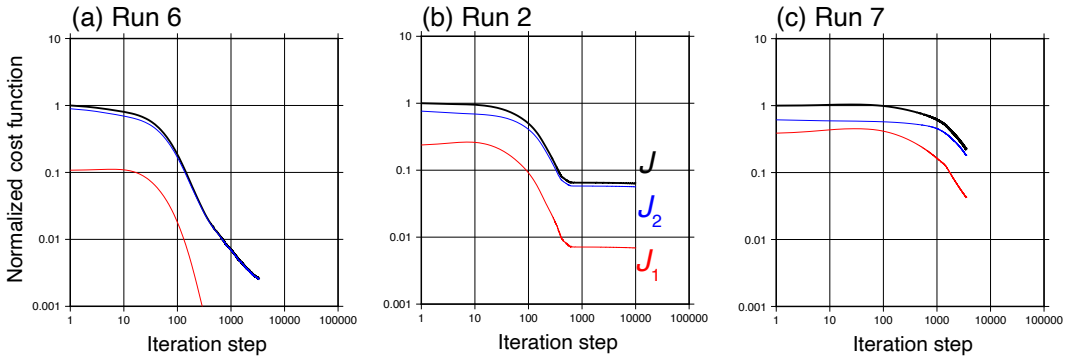

Figure 12. Same as the lower half of Figure 5 but for runs 6, 2, and 7, with different assimilation time windows.

5 Discussion

5.1 Application to fluids containing traceable particles

The 4D-Var marker-in-cell data assimilation achieves a tens of Myr time reversal of the convective structure at a horizontal wavelength of thousands of km, comparable to the target scales of geodynamics (Gu et al., 2001; Müller et al., 2016). The method has successfully estimated the thermal structures of the reference forward models without using any temperature data, because the temperature is optimized as the driving force of the particle motions.

This study provides new values for a marker-in-cell method. Despite the much higher computational cost, a marker-in-cell method is necessary to accurately compute the advection of chemical heterogeneity at resolutions smaller than the grid resolution in geodynamic forward models. Rather, the present results highlight the advantages of traceability of identifiable particles for constraining the convection process of the surrounding fluid in the context of inverse problems.

The 4D-Var marker-in-cell method is applicable to models that do not use a stream function as long as the velocity field can be calculated or is known. For example, the way to implement particle position correction in a three-dimensional model is simply to add a dimension to $\boldsymbol{x}_i$, $\boldsymbol{u}_i$, and $\boldsymbol{\lambda}_i$ in the forward and adjoint equations of the particle advection. It is also possible to implement the method in the system containing source and sink of the fluid flow, such as fluid addition from fractures to fluid flows in an underground porous medium (Nagao et al., 2013). Thus, our method can be applied not only to the mantle dynamics but also to a wide range of the fluid containing traceable particles.

It should be noted that our 4D-Var marker-in-cell method may not be applicable to particles in turbulent flows. This is because it is difficult to track individual particles for long periods of time in their complex flows. Due to the difficulty of particle tracking, previous data assimilation studies for turbulent flows have converted Lagrangian velocity into Eulerian velocity at each time step using a particle image velocimetry (PIV) imaging (Bauer et al., 2022). In fact, forward numerical simulations to calculate the transport of chemical species in turbulent flows usually do not consider identical particles, but define the number density per unit cell, which is solved by using an advection-diffusion equation (e.g., volcanic ash in a pyroclastic flow; Shimizu et al., 2019). Such modeling

is justified because the random behavior of individual particles can be represented macroscopically as diffusion. Thus, untraceable particles can be treated like heat, which can be expressed as an advection-diffusion equation (equation 7) and solved by a grid. Therefore, the source and flow process of the untraceable particles in the turbulent flows can be constrained in the same way that the thermal structure of the past mantle is solved based on the present mantle temperature estimated from seismic tomography (e.g., Bunge et al., 2003; Li et al., 2017; Nakao et al., 2024).

5.2 Application of petrological data

It is necessary to discuss how to apply real petrological data reflecting mantle convection to the 4D-Var method, because this study used an ideal dataset mimicking markers in transparent fluid convecting in a water tank (e.g., Schellart, 2004). In the petrological context, the vertical particle position z_i^{obs} and the vertical particle velocity w_i^{obs} assimilated in the 4D-Var models correspond to the pressure estimated by geobarometers and its time derivative, respectively. Meanwhile, the horizontal position x_i^{obs} and the horizontal velocity u_i^{obs} would be estimated based on paleomagnetism. Previous studies have converted paleomagnetic observations into surface velocity fields using plate reconstruction models (e.g., Bunge et al., 2003; Price & Davies, 2018; Ghelichkhan et al., 2021), while future studies based on the 4D-Var marker-in-cell method can use such observations directly as Lagrangian data. The fundamental problem with rock sampling is that rock samples are available only after they are exposed on the Earth's surface. Therefore, it may be difficult for our method to constrain the velocity and temperature of downwelling flows due to the unavailability of the downwelling samples. To constrain them as much as possible, it will be necessary in the future study based on real datasets to include the velocity data around converging boundaries, which is useful to estimate temperature of the downwelling velocity and temperature via a force balance between negative thermal buoyancy and shear stress around the negative thermal anomaly (Nakao et al., 2024).

We have considered only the particle positions and velocities in the cost function and the adjoint equations for simplicity, but it is possible to consider other additional particle information, such as the temperature estimated by geothermometers. In this case, a temperature error term will appear in the adjoint energy conservation equation (equation 15). The Lagrangian temperature records will provide the new insight, since seis-

mic tomography is the only data reflecting the thermal structure in the previous geodynamic data assimilation studies (e.g., Bunge et al., 2003; Price & Davies, 2018; Ghelichkhan et al., 2021). For example, if there are factors other than temperature that induce fluid flow, such as buoyancy due to chemical heterogeneity, there would be a trade-off between temperature and chemical composition estimates; temperature estimated from rock samples would be useful to resolve such a trade-off.

Our numerical experiments show that the reconstruction of the thermal convection tends to be successful when large assimilation weights are given for particle data sampled in the early stages (section 4.2), and when the data assimilation time window for reversal is short (section 4.3). This result seems reasonable because a 4D-Var algorithm generally optimizes variables at the initial conditions through which the thermal convection develops deterministically. However, the weighting of old petrological data may require some careful consideration, because the older the rocks, the larger the expected errors in the estimates of P, T, age, and other parameters due to weathering. To clarify appropriate assimilation weights for given petrological data, a second-order adjoint method that assesses the uncertainty of data assimilation solutions would be useful (Ito et al., 2016). It should be noted, of course, that when applying the 4D-Var marker-in-cell method, sufficient rock records reflecting the past mantle are required to infer the mantle convection process.

6 Conclusions

This study has developed a 4D-Var marker-in-cell data assimilation framework that uses particle data to estimate time series of thermal and flow structures within the surrounding fluid. The 4D-Var marker-in-cell method has been applied to positions and velocities of synthetic particles, sampled from a convection model that emulates Earth's mantle convection. This application aimed to simulate the reconstruction process based on the lateral wavelength of the convecting cells, the data assimilation time window, and the time dependence of the hyperparameters. The 4D-Var method has successfully achieved a 60-Myr reversal of the convective structure with a horizontal wavelength of 6,000 km, without using any temperature data. For convection at smaller scales, the cost function tends not to decrease effectively. However, improvements in reconstruction have been observed with either a shorter retrospective time domain or an increased emphasis on early-stage information.

Our method highlights the advantages of using highly traceable particles in laminar flow, especially in the context of data assimilation for reconstructing fluid convection processes. Lagrangian data, derived from geobarometers and geothermometers, provide insights into the past of the deep mantle. In contrast, Eulerian data, derived from geophysical surveys, provide a snapshot of the present or shallow Earth dynamics. Consequently, merging these complementary datasets and integrating them into a future data assimilation framework holds significant potential for quantitatively reconstructing the evolution of mantle dynamics over historical timescales.

Appendix A Interpolation of velocity for particle advection equations

The forward advection equation of particle i (equation 8) is approximated by a first-order Lagrangian interpolation as

$$\frac{dx_i}{dt} = \frac{x_i - x_W}{x_E - x_W} u_E + \frac{x_E - x_i}{x_E - x_W} u_W \tag{A1}$$

$$\frac{dz_i}{dt} = \frac{z_i - z_N}{z_S - z_N} w_S + \frac{z_S - z_i}{z_S - z_N} w_N. \tag{A2}$$

Indices W, E, N, and S indicate left, right, top, and bottom boundaries of the cell at which tracer particle i locates, respectively (i.e., $x_W \leq x_i \leq x_E$ and $z_N \leq z_i \leq z_S$), as described in Figure A1.

A simple cost function is defined here as

$$J = \int_{t_0}^{t_1} dt \sum_{i=1}^{N} \frac{1}{2} \alpha \left[\left(x_i - x_i^{\text{obs}} \right)^2 + \left(z_i - z_i^{\text{obs}} \right)^2 \right]. \tag{A3}$$

Given equations (A1) to (A3), we can derive the adjoint equation for the advection of particle i as

$$\frac{d\lambda_i}{dt} = \alpha \left(x_i - x_i^{\text{obs}} \right) - \lambda_i \frac{u_E - u_W}{x_E - x_W} \tag{A4}$$

$$\frac{d\mu_i}{dt} = \alpha \left(z_i - z_i^{\text{obs}} \right) - \mu_i \frac{w_S - w_N}{z_S - z_N}. \tag{A5}$$

Open Research Section

All data for this work can be downloaded from Zenodo (https://doi.org/10.5281/zenodo.11207943).

Acknowledgments

Simulation results were visualized using Generic Mapping Tools (Wessel et al., 2019).

This research was supported by the Japan Science and Technology Agency (CREST; Grants

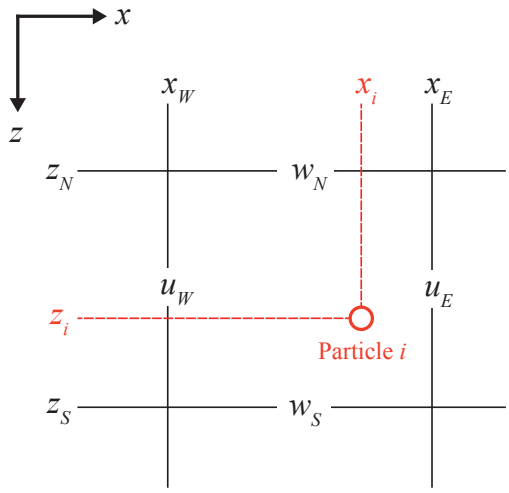

Figure A1. Configuration of physical variables on the staggered grid and tracer particles for solving the forward and adjoint marker advection equations (equations A1, A2, A4, and A5).

JPMJCR1761 and JPMJCR1763), MEXT Project for Seismology towards Research Innovation with Data of Earthquakes (STAR-E; Grant JPJ010217), JSPS KAKENHI Grants (20K21785, 21H04750, 22K03542, 22K14131, 23H00466, and 24K00744), and Joint Research Programs of the Earthquake Research Institute, University of Tokyo (2021-B-01, 2022-B-06, and 2024-B-01).

References

Bauer, C., Schiepel, D., & Wagner, C. (2022). Assimilation and extension of particle image velocimetry data of turbulent Rayleigh–Bénard convection using direct numerical simulations. *Experiments in Fluids*, *63*, 22. doi: https://doi.org/10.1007/s00348-021-03369-3

Bocher, M., Coltice, N., Fournier, A., & Tackley, P. J. (2016). A sequential data assimilation approach for the joint reconstruction of mantle convection and surface tectonics. *Geophysical Journal International*, *204*(1), 200–214. doi: https://doi.org/10.1093/gji/ggv427

Bunge, H.-P., Hagelberg, C., & Travis, B. (2003). Mantle circulation models with variational data assimilation: inferring past mantle flow and structure from plate motion histories and seismic tomography. *Geophysical Journal International*, *152*(2), 280–301. doi: https://doi.org/10.1046/j.1365-246X.2003.01823

Chowdhury, P., Chakraborty, S., & Gerya, T. V. (2021). Time will tell: Secular change in metamorphic timescales and the tectonic implications. *Gondwana Research*, *93*, 291–310. doi: https://doi.org/10.1016/j.gr.2021.02.003

Fei, H., Wiedenbeck, M., Yamazaki, D., & Katsura, T. (2013). Small effect of water on upper-mantle rheology based on silicon self-diffusion coefficients. *Nature*, *498*(7453), 213–215. doi: https://doi.org/10.1038/nature12193

Fukuda, J., & Johnson, K. M. (2021). Bayesian inversion for a stress-driven model of afterslip and viscoelastic relaxation: Method and application to postseismic deformation following the 2011 MW 9.0 Tohoku-Oki earthquake. *Journal of Geophysical Research: Solid Earth*, *126*(5), e2020JB021620. doi: https://doi.org/10.1029/2020JB021620

Gerya, T. V., & Yuen, D. A. (2003). Characteristics-based marker-in-cell method with conservative finite-differences schemes for modeling geological flows with strongly variable transport properties. *Physics of the Earth and Planetary Interiors*, *140*(4), 293–318. doi: https://doi.org/10.1016/j.pepi.2003.09.006

Ghelichkhan, S., Bunge, H.-P., & Oeser, J. (2021). Global mantle flow retrodictions for the early Cenozoic using an adjoint method: evolving dynamic topographies, deep mantle structures, flow trajectories and sublithospheric stresses. *Geophysical Journal International*, *226*(2), 1432–1460. doi: https://doi.org/10.1093/gji/ggab108

Gu, Y. J., Dziewonski, A. M., Su, W., & Ekström, G. (2001). Models of the mantle shear velocity and discontinuities in the pattern of lateral heterogeneities. *Journal of Geophysical Research: Solid Earth*, *106*(B6), 11169–11199. doi: https://doi.org/10.1029/2001JB000340

Guillaume, B., Funiciello, F., & Faccenna, C. (2021). Interplays between mantle flow and slab pull at subduction zones in 3D. *Journal of Geophysical Research: Solid Earth*, *126*(5), e2020JB021574. doi: https://doi.org/10.1029/2020JB021574

Ito, S., Nagao, H., Yamanaka, A., Tsukada, Y., Koyama, T., Kano, M., & Inoue, J. (2016). Data assimilation for massive autonomous systems based on a second-order adjoint method. *Physical Review E*, *94*(4), 043307.

Iwamori, H. (2003). Viscous flow and deformation of regional metamorphic belts at convergent plate boundaries. *Journal of Geophysical Research: Solid Earth*,

108(B6). doi: https://doi.org/10.1029/2002JB001808

Iwamori, H., Ueki, K., Hoshide, T., Sakuma, H., Ichiki, M., Watanabe, T., ... others (2021). Simultaneous analysis of seismic velocity and electrical conductivity in the crust and the uppermost mantle: a forward model and inversion test based on grid search. *Journal of Geophysical Research: Solid Earth*, *126*(9), e2021JB022307. doi: https://doi.org/10.1029/2021JB022307

Kalnay, E. (2003). *Atmospheric modeling, data assimilation and predictability.* Cambridge University Press.

Kuwatani, T., Nagao, H., Ito, S., Okamoto, A., Yoshida, K., & Okudaira, T. (2018). Recovering the past history of natural recording media by Bayesian inversion. *Physical Review E*, *98*(4), 043311. doi: https://doi.org/10.1103/PhysRevE.98.043311

Kuwatani, T., Nagata, K., Sakai, T., & Iwamori, H. (2023). Bayesian inversion of lithology and liquid phase parameters from seismic velocity and electrical conductivity in the crust and uppermost mantle. *Journal of Geophysical Research: Solid Earth*, *128*(12), e2023JB026836. doi: https://doi.org/10.1029/2023JB026836

Kuwatani, T., Nagata, K., Yoshida, K., Okada, M., & Toriumi, M. (2018). Bayesian probabilistic reconstruction of metamorphic p–t paths using inclusion geothermobarometry. *Journal of Mineralogical and Petrological Sciences*, *113*(2), 82–95. doi: https://doi.org/10.2465/jmps.170923

Li, D., Gurnis, M., & Stadler, G. (2017). Towards adjoint-based inversion of time-dependent mantle convection with nonlinear viscosity. *Geophysical Journal International*, *209*(1), 86–105. doi: https://doi.org/10.1093/gji/ggw493

Matsumoto, K., & Nakamura, M. (2017). Syn-eruptive breakdown of pyrrhotite: a record of magma fragmentation, air entrainment, and oxidation. *Contributions to Mineralogy and Petrology*, *172*, 1–19. doi: https://doi.org/10.1007/s00410-017-1403-8

Müller, R. D., Seton, M., Zahirovic, S., Williams, S. E., Matthews, K. J., Wright, N. M., ... others (2016). Ocean basin evolution and global-scale plate reorganization events since Pangea breakup. *Annual Review of Earth and Planetary Sciences*, *44*, 107–138. doi: https://doi.org/10.1146/annurev-earth-060115-012211

Nagao, H., Sato, K., & Higuchi, T. (2013). MCMC algorithm to estimate physical parameters of a fracture from tracer test. *Journal of the Japanese Association for Petroleum Technology*, *78*(2), 197–209. doi: https://doi.org/10.3720/japt.78.197

Nakao, A., Iwamori, H., & Nakakuki, T. (2016). Effects of water transportation on subduction dynamics: Roles of viscosity and density reduction. *Earth and Planetary Science Letters*, *454*, 178–191. doi: https://doi.org/10.1016/j.epsl.2016.08.016

Nakao, A., Iwamori, H., Nakakuki, T., Suzuki, Y. J., & Nakamura, H. (2018). Roles of hydrous lithospheric mantle in deep water transportation and subduction dynamics. *Geophysical Research Letters*, *45*(11), 5336–5343. doi: https://doi.org/10.1029/2017GL076953

Nakao, A., Kuwatani, T., Ito, S., & Nagao, H. (2024). Adjoint-based data assimilation for reconstruction of thermal convection in a highly viscous fluid from surface velocity and temperature snapshots. *Geophysical Journal International*, *236*(1), 379–394. doi: https://doi.org/10.1093/gji/ggad417

Price, M., & Davies, J. (2018). Profiling the robustness, efficiency and limits of the forward-adjoint method for 3-D mantle convection modelling. *Geophysical Journal International*, *212*(2), 1450–1462. doi: https://doi.org/10.1093/gji/ggx489

Schellart, W. (2004). Kinematics of subduction and subduction-induced flow in the upper mantle. *Journal of Geophysical Research: Solid Earth*, *109*(B7). doi: https://doi:10.1029/2004JB002970

Shimizu, H. A., Koyaguchi, T., & Suzuki, Y. J. (2019). The run-out distance of large-scale pyroclastic density currents: a two-layer depth-averaged model. *Journal of Volcanology and Geothermal Research*, *381*, 168–184. doi: https://doi.org/10.1016/j.jvolgeores.2019.03.013

Turcotte, D. L., & Schubert, G. (2002). *Geodynamics*. Cambridge University Press.

Warren, C., Beaumont, C., & Jamieson, R. A. (2008). Modelling tectonic styles and ultra-high pressure (UHP) rock exhumation during the transition from oceanic subduction to continental collision. *Earth and Planetary Science Letters*, *267*(1-2), 129–145. doi: https://doi.org/10.1016/j.epsl.2007.11.025

Watanabe, H. K., Nagai, Y., Sakai, S., Kobayashi, G., Yamamori, L., Tada, N., ...

Yusa, Y. (in revision). Heterogeneous shell growth of the neustonic goose barnacle lepas anserifera: Its potential application for tracking floating materials. *Research Square*. doi: https://doi.org/10.21203/rs.3.rs-3819539/v1

Wessel, P., Luis, J. F., Uieda, L. a., Scharroo, R., Wobbe, F., Smith, W. H., & Tian, D. (2019). The generic mapping tools version 6. *Geochemistry, Geophysics, Geosystems*, *20*(11), 5556–5564. doi: https://doi.org/10.1029/2019GC008515