

FULL PAPER

Open Access



Adjoint-based direct data assimilation of GNSS time series for optimizing frictional parameters and predicting postseismic deformation following the 2003 Tokachi-oki earthquake

Masayuki Kano^{1*} , Shin'ichi Miyazaki², Yoichi Ishikawa³ and Kazuro Hirahara^{4,5}

Abstract

Postseismic Global Navigation Satellite System (GNSS) time series followed by megathrust earthquakes can be interpreted as a result of afterslip on the plate interface, especially in its early phase. Afterslip is a stress release process accumulated by adjacent coseismic slip and can be considered a recovery process for future events during earthquake cycles. Spatio-temporal evolution of afterslip often triggers subsequent earthquakes through stress perturbation. Therefore, it is important to quantitatively capture the spatio-temporal evolution of afterslip and related postseismic crustal deformation and to predict their future evolution with a physics-based simulation. We developed an adjoint data assimilation method, which directly assimilates GNSS time series into a physics-based model to optimize the frictional parameters that control the slip behavior on the fault. The developed method was validated with synthetic data. Through the optimization of frictional parameters, the spatial distributions of afterslip could roughly (but not in detail) be reproduced if the observation noise was included. The optimization of frictional parameters reproduced not only the postseismic displacements used for the assimilation, but also improved the prediction skill of the following time series. Then, we applied the developed method to the observed GNSS time series for the first 15 days following the 2003 Tokachi-oki earthquake. The frictional parameters in the afterslip regions were optimized to $A-B \sim O(10 \text{ kPa})$, $A \sim O(100 \text{ kPa})$, and $L \sim O(10 \text{ mm})$. A large afterslip is inferred on the shallower side of the coseismic slip area. The optimized frictional parameters quantitatively predicted the postseismic GNSS time series for the following 15 days. These characteristics can also be detected if the simulation variables can be simultaneously optimized. The developed data assimilation method, which can be directly applied to GNSS time series following megathrust earthquakes, is an effective quantitative evaluation method for assessing risks of subsequent earthquakes and for monitoring the recovery process of megathrust earthquakes.

Keywords: Data assimilation, Afterslip, The 2003 Tokachi-oki earthquake, GNSS, Frictional parameters, Postseismic deformation prediction

Main Text

Introduction

Large megathrust earthquakes have been recurrently observed in many subduction zones (Ando 1975). During such seismic cycles, afterslip can often be observed as a slow transient fault slip at plate boundaries following

*Correspondence: masayuki.kano.a3@tohoku.ac.jp

¹ Graduate School of Science, Tohoku University, 6-3, Aramaki-aza-aoba, Aoba-ku, Sendai 980-8578, Japan

Full list of author information is available at the end of the article

large earthquakes. The occurrence of afterslip results from a stress change due to the coseismic slip, and their slip areas are usually adjacent to the coseismic slip area. Such transient slips last for months to years in the case of magnitude (M) 7.0 or larger interplate earthquakes (Miyazaki et al. 2004; Hsu et al. 2006; Chlieh et al. 2007; Ozawa et al. 2011). While afterslip itself is a stress release process, it plays various tectonic roles during seismic cycles.

Stress release due to afterslip further yields a stress perturbation in the surrounding areas that can be inferred as spatio-temporal evolution or propagation of afterslip. Such stress perturbations due to afterslip sometimes trigger subsequent earthquakes in different time scales (Murakami et al. 2006; Pollitz et al. 2006; Miyazaki and Larson 2008; Uchida et al. 2009; Ohta et al. 2012). Miyazaki and Larson (2008) reported the afterslip distribution 1 h after the 2003 M 8.0 Tokachi-oki earthquake by analyzing the high-rate Global Navigation Satellite System (GNSS) data. The resulting afterslip was distributed between the mainshock and the largest M 7.4 aftershock that occurred ~ 1 h after the mainshock. Hence, the largest aftershock was possibly triggered by the initial afterslip. Another example includes the daily scale triggering related to the 2011 M 9 Tohoku-oki earthquake. Ohta et al. (2012) reported that the 2-day-long afterslip due to the M 7.3 earthquake occurred on March 9, 2011, 2 days before the Tohoku-oki earthquake, and it was located on the northern side of the hypocenter of the M 9.0 mainshock. This implies that the Tohoku-oki earthquake can be more or less loaded by the afterslip of the foreshock. For year-long triggering, Uchida et al. (2009) showed that the 2004 M 7.1 Kushiro-oki earthquake, located ~ 100 km east of the 2003 Tokachi-oki earthquake, could be advanced by the eastward propagation of the afterslip of the Tokachi-oki earthquake. These examples show that the triggering of subsequent earthquakes due to the spatio-temporal evolution of afterslip can often occur, although whether the stress perturbations of afterslip can directly trigger subsequent earthquakes highly depends on the amplitude of the perturbation and the level of accumulated stress on possible source regions.

Afterslip and the viscoelastic relaxation effect of coseismic slip reflect the recovery process following large earthquakes (Diao et al. 2014; Sun and Wang 2015; Klein et al. 2016; Itoh et al. 2019). Itoh et al. (2019) modeled the spatio-temporal evolution of afterslip for 7.5 years by considering the viscoelastic effect following the 2003 Tokachi-oki earthquake. Based on the spatial distribution of afterslip, they concluded that the interplate locking of the afterslip area had not fully recovered to the pre-seismic state after 7.5 years. Hence, investigating the afterslip

evolution both in time and space provides information on the probability of triggered earthquakes and on the relocking process following the megathrust earthquakes. To quantify this information, it is important to assess the previous and current states of afterslip based on the observations and evaluate the future evolution based on physics-based simulations.

Spatio-temporal evolution of fault slip in seismic cycles along a subducting plate can often be simulated based on a physics-based model (Hori et al. 2004; Liu and Rice 2005; Kato 2008), which consists of the quasi-dynamic equation of motion in an elastic medium (Rice 1993) and the rate- and state-dependent friction (RSF) law combined with the evolution law for the state variable (Dieterich 1979; Ruina 1983). In this model, frictional parameters in the frame of the RSF law qualitatively determine the behaviors of various fault slip styles including fast coseismic slip, slow transient aseismic slip such as afterslip and slow slip events, and steady slips. By assigning these parameters with trial and error, a wide variety of slip phenomena can be qualitatively reproduced (Nakata et al. 2012, 2014; Ohtani et al. 2014).

For a more realistic simulation and quantitative reproduction of observed data, many data assimilation techniques have been developed and widely adopted in various fields of geophysics (e.g., Lewis et al. 2006). The data assimilation incorporates observations into physics-based models to optimize simulation variables and/or parameters in the system of interest and forecast their evolutions. They have been originally developed and expanded in the fields of meteorology and oceanology and now are applied to practical problems such as weather forecasts. In the field of seismology, data assimilation methods have been recently introduced to estimate and forecast tsunamis and seismic waves (Hoshiba and Aoki 2015; Maeda et al. 2015; Wang et al. 2017; Furumura et al. 2019; Oba et al. 2020).

To estimate and predict the fault state, such as slip velocities and stress states, van Dinther et al. (2019) and Hirahara and Nishikiori (2019) have introduced the ensemble Kalman filter (EnKF), one of the sequential data assimilation methods. The EnKF sequentially updates the simulation variables and/or physical parameters of interest every time the observations are acquired. Hirahara and Nishikiori (2019) verified that the EnKF can accurately recover the synthesized daily GNSS surface velocity by sequentially optimizing the slip velocities and frictional parameters in the case of slow slip events. On the contrary, Kano et al. (2013, 2015) adopted a non-sequential approach, an adjoint data assimilation method (Lewis et al. 2006), to optimize frictional parameters. The adjoint method optimizes the simulation variables and/or physical

parameters, such that the simulated time series quantitatively explains all the observations obtained during the study period. Kano et al. (2015) applied the adjoint method to the real afterslip data following the 2003 Tokachi-oki earthquake and successfully reproduced and forecasted the slip velocities by optimizing spatially heterogeneous frictional parameters. However, they assimilated indirect observations of slip velocities along the plate interface inferred by Miyazaki et al. (2004) using a kinematic inversion approach as if the slip velocities were directly observed. In this study, we update the adjoint method developed by Kano et al. (2015) to assimilate direct observations, i.e., observed GNSS time series. Then, we apply the developed method to the postseismic deformation following the 2003 Tokachi-oki earthquake to optimize the frictional parameters and simulation variables and to predict future evolutions of postseismic deformation.

The outline of the paper is as follows. First, we summarize the adjoint data assimilation method for optimizing frictional parameters and simulation variables together with relevant observations and a physics-based model. Next, we present the numerical experiments conducted to optimize the frictional parameters with synthetic postseismic data to test the feasibility of the developed method, followed by the application to the real GNSS time series after the 2003 Tokachi-oki earthquake. Finally, we present the additional assimilation conducted for the simultaneous optimization of both frictional parameters and simulation variables, the comparison of the results when assimilating the indirect observations of slip velocities, and a discussion of the future improvements of the developed method.

Adjoint data assimilation method

An adjoint data assimilation method incorporates the observation data to a physics-based numerical simulation to optimize the initial conditions and/or model parameters in the simulation. In this study, an adjoint method was employed to predict the postseismic time series following the 2003 Tokachi-oki earthquake by optimizing the frictional parameters in the physics-based simulation. In this section, after summarizing the 2003 Tokachi-oki earthquake and the following postseismic observations, we describe the physics-based numerical model. Then, we briefly introduce the formulation of the adjoint method for our problem and describe the data assimilation settings used throughout this study following Kano et al. (2015). A more general explanation of the adjoint data assimilation method is summarized in Lewis et al. (2006).

2003 Tokachi-oki earthquake and the postseismic GNSS observations

This study focuses on the postseismic deformation following the 2003 Tokachi-oki earthquake, which was a megathrust earthquake of M 8.0 that occurred on September 25, 2003 (UT) off the Tokachi region in north-eastern Japan. The earthquake region is located in the southwestern part of the Kurile subduction zone where many megathrust events have repeatedly occurred due to a subduction of the Pacific plate including 1843, 1952, and 2003 Tokachi-oki earthquakes (Hatori 1984; Hirata et al. 2003; Tanioka et al. 2004). The main rupture zone of the 2003 Tokachi-oki earthquake coincides with that in the previous M 8.2 earthquake in 1952 (Yamanaka and Kikuchi 2003), which we hereafter refer to as the Tokachi-oki asperity (TA).

A significant postseismic deformation following the 2003 Tokachi-oki earthquake was continuously detected by the GNSS Earth Observation Network System (GEONET) in Japan, operated by the Geospatial Information Authority of Japan (GSI) (Miyazaki et al. 2004; Ozawa et al. 2004; Baba et al. 2006). Based on such a dense network, Miyazaki et al. (2004) inferred the afterslip distributions for the first one month, which were complementarily distributed adjacent to the TA, exhibiting that frictional properties on both slip areas were spatially different. The spatio-temporal evolution of the afterslip likely triggered the 2004 M 7.1 Koshiro-oki earthquake, located ~100–150 km northeast of the TA (Murakami et al. 2006; Uchida et al. 2009), through a stress perturbation due to the propagation of afterslip. The source area of the Koshiro-oki earthquake is referred to as KA in this study. We modeled these two asperities, TA and KA, as velocity-weakening zones in the next subsection.

We utilized the daily GNSS time coordinates recorded at 54 stations of GEONET in Hokkaido and northern Tohoku areas (Fig. 1), which are analyzed and offered as the GEONET F3 solutions by the GSI (Nakagawa et al. 2009). Data period is 1 month following the mainshock, i.e., September 26 to October 26. We calculated the differences in the coordinates from September 26, defined as day 00, for the next 30 days as cumulative displacements. The time series in two horizontal components are rotated to the trench-parallel (X-axis) and trench-perpendicular (Y-axis) directions. We subtracted the coseismic steps due to two aftershocks that occurred on September 29 and October 8 with magnitudes greater than 6.2 listed in the hypocenter catalog provided by the Japan Meteorological Agency. The coseismic steps were theoretically calculated by assuming a single uniform slip fault model with earthquake focal mechanisms determined by the broadband seismic networks [F-net,

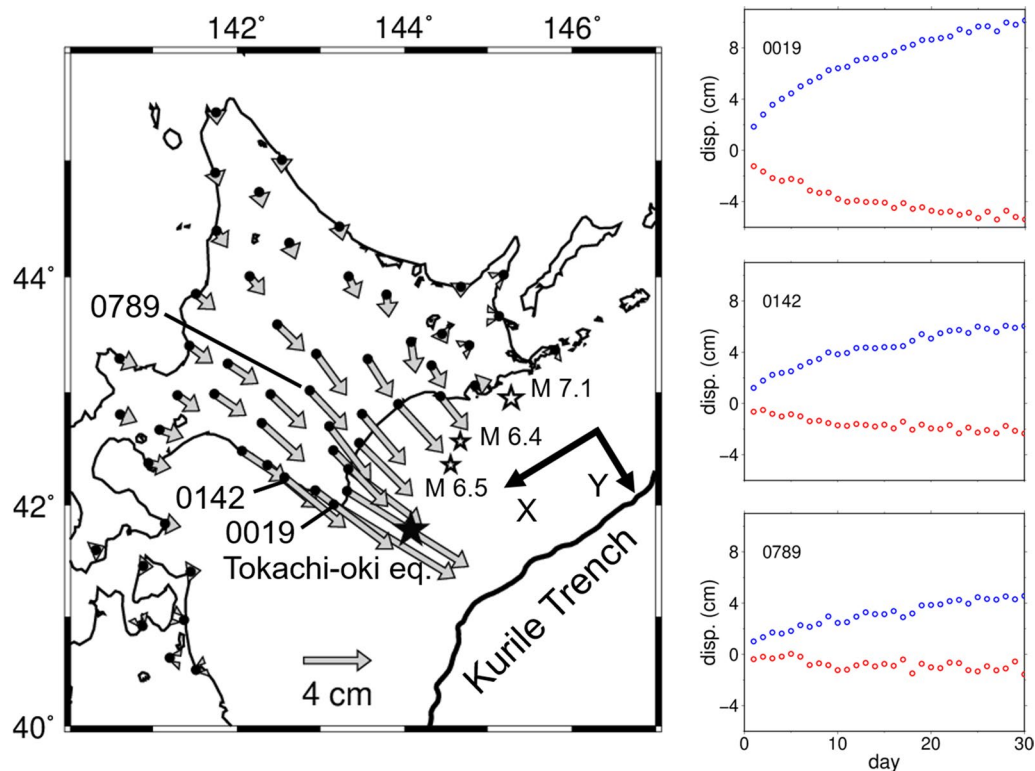


Fig. 1 Observed displacement vectors and GNSS time series. The left panel shows the observed cumulative displacement for 15 days following the 2003 Tokachi-oki earthquake. Black and white stars indicate the epicenters of the mainshock and three aftershocks, which are the M 7.1 2004 Kushiro-oki earthquake and the two M 6 class earthquakes occurred within 15 days following the mainshock. The coseismic steps for the two M 6 class earthquakes have been removed. The right panels show the examples of the GNSS time series at three stations for 30 days. The time series for the first 15 days are used for assimilation, while those for the later 15 days are used for testing the prediction skill of the developed method. Red and blue circles indicate the cumulative displacement in trench-parallel (X) and trench-perpendicular (Y) directions, respectively. The black dots show the locations of the GEONET stations used in this study

National Research Institute for Earth Science and Disaster Resilience (NIED) 2019] using the elastic homogeneous half-space dislocation theory of Okada (1992). We divided the entire time series into two periods: the observations for the first 15 days were assimilated into numerical simulations for optimizing frictional parameters (assimilation period), and those for the later 15 days were adopted for verifying the predictability of our assimilation results (prediction period).

Figure 1 shows the examples of the GNSS time series, which exhibits a logarithmically decaying curve. The cumulative displacement vectors for the assimilation period exhibit a clear coherent southeastward movement with a maximum amplitude of ~ 8.5 cm at the 0019 station, and decay with the distance from the epicenter. Because we analyzed 30 days of data following the mainshock, these signals are mainly attributed to the afterslip on the plate interface rather than viscoelastic effect. This assumption is reasonable in view of Itoh et al. (2019), which showed that the calculated surface displacement

due to the viscoelastic effect of the 2003 Tokachi-oki earthquake at an inland station located at northern Hokkaido was, at most, a few mm in the first month following the earthquake, even in the case of a low viscosity (Additional file 1: Figure S13 in Itoh et al. 2019).

Fault modeling and governing equations

We modeled the plate boundary along the southwestern Kurile subduction zone as a planar rectangular fault (Fig. 2). The entire fault is 330 km long and 220 km wide with a dip angle of 20° and consists of two asperities, TA and KA, which correspond to the source region of the Tokachi-oki and Kushiro-oki earthquakes, respectively, and the surrounding region. TA is 60 km long and 60 km wide, and KA is 40 km long and 60 km wide. The surrounding region is divided into 666 subfaults of 10×10 km, where aseismic fault slip is expected to occur. This fault model is exactly the same as that in Kano et al. (2015).

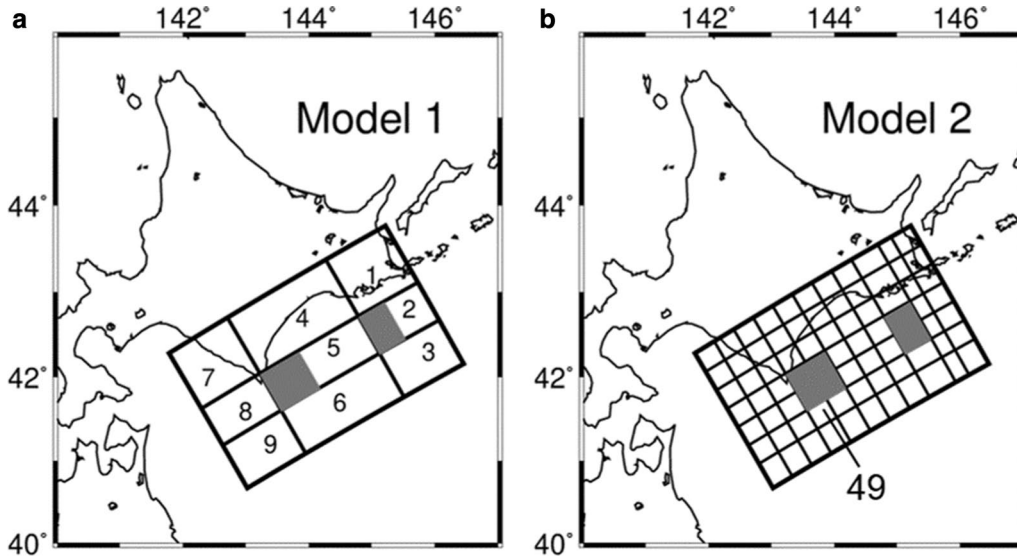


Fig. 2 Fault models. Thick rectangle is the modeled region, which is divided into 10 km \times 10 km subfaults. Two gray regions represent the Tokachi-oki and the Kushiro-oki asperities, respectively. The entire fault is divided into 9 and 71 regions in **a** model 1 and **b** model 2, respectively. Frictional parameters are assumed to be uniform within each region. Model 1 is used in the numerical experiments, while model 2 in the application to real data

Temporal evolutions of fault slips are simulated by solving the physics-based governing equations at each subfault, which consist of a quasi-dynamic equation of motion in an elastic medium (Rice 1993) and the RSF law with an aging law as the evolution of state variables (Dietrich 1979; Ruina 1983):

$$\sigma_i \mu_i(t) = \sum_j k_{ij} (v_{pl} t - s_j(t)) - \frac{G}{2v_s} v_i(t), \quad (1)$$

$$\sigma_i \mu_i(t) = \sigma_i \mu_0 + A_i \ln \left(\frac{v_i(t)}{v_0} \right) + B_i \ln \left(\frac{v_0 \theta_i(t)}{L_i} \right), \quad (2)$$

$$\frac{d\theta_i(t)}{dt} = 1 - \frac{v_i(t) \theta_i(t)}{L_i}, \quad (3)$$

$$\frac{ds_i(t)}{dt} = v_i(t). \quad (4)$$

In this set of equations, simulation variables at subfault i where temporal evolutions are calculated are the slip velocity $v_i(t)$, state variable $\theta_i(t)$, frictional coefficient $\mu_i(t)$, and slip $s_i(t)$. In the simulation, $v_i(t)$ and $\theta_i(t)$ were independently calculated and then $\mu_i(t)$ and $s_i(t)$ were obtained using Eqs. (2) and (4). v_{pl} is the plate convergence velocity, G is the rigidity, v_s is the shear wave velocity, and μ_0 is the reference frictional coefficient while sliding at a reference velocity v_0 at the

steady state. These constants were set as $v_{pl} = 9.0$ cm/year (DeMets et al. 1990), $G = 30$ GPa, $v_s = 3.0$ km/s. The constants μ_0 and v_0 were not explicitly given in the simulation because these values are necessary only to calculate $\mu_i(t)$, which was not calculated in this study. The slip response function k_{ij} represents the shear stress change at subfault i due to a unit slip at subfault j . Thus, the first term on the right-hand side of Eq. (1) indicates the contribution of the stress change for all subfaults. We calculated k_{ij} assuming a homogeneous elastic half-space based on Okada (1992). The second term represents the radiation damping, which approximates the inertial term (Rice 1993) and is often adopted in quasi-dynamic simulations rather than treating the fully dynamic effect. A_i , B_i , and L_i are the frictional parameters at subfault i that control the slip behavior of the subfault. A corresponds to the instantaneous change due to a change in the slip velocity, while B is related to the evolution of the state variable. At steady state, i.e., $d\theta/dt = 0$ in Eq. (3), the change in the frictional coefficient at the instantaneous velocity change is proportional to $A - B$. Thus, when $A - B$ is negative, the frictional coefficient decreases, showing a velocity-weakening frictional property leading to an unstable slip; when $A - B$ is positive, it yields a velocity-strengthening frictional property, resulting in a stable fault slip. L is the characteristic slip distance. In this study, we set two asperities as the velocity-weakening regions and fix the frictional parameters as $A = 40$ kPa,

$A-B = -40$ kPa, $L = 40$ mm in TA, $A = 40$ kPa, $A-B = -30$ kPa, $L = 40$ mm in KA by referring to Kano et al. (2015).

This study focuses on optimizing the frictional parameters in the surrounding region. Since it is difficult to spatially resolve frictional parameters in all subfaults, we assume spatial homogeneity of frictional parameters with two different spatial scales based on Kano et al. (2015) (Fig. 2): we divide the modeled region into 9 regions in model 1 in the synthetic test to evaluate the feasibility of the developed method (Fig. 2a), and 71 regions in model 2 to apply the method to real GNSS data (Fig. 2b). The frictional parameters are assumed to be uniform within each region. Therefore, the numbers of parameters to be optimized in models 1 and 2 are 27 and 213, respectively. Note that the spatial distributions of the frictional parameters were obtained based on model 2 in Kano et al. (2015).

By taking the time derivatives of Eqs. (1) and (2) and combining them with Eq. (3), the governing equations can be rewritten as the following differential equations:

$$\begin{aligned} \frac{dv_i(t)}{dt} &= \frac{\sum_j k_{ij}(v_{pl} - v_j(t)) - \frac{A_i - (A_i - B_i)}{\theta_i(t)} \left(1 - \frac{v_i(t)\theta_i(t)}{L_i}\right)}{\frac{A_i}{v_i} + \frac{G}{2v_s}} \\ &\equiv \frac{P_i(v_1, \dots, v_{N_f}, \theta_i, A_i, A_i - B_i, L_i)}{Q_i(v_i, A_i)}, \end{aligned} \quad (5)$$

$$\frac{d\theta_i(t)}{dt} = 1 - \frac{v_i(t)\theta_i(t)}{L_i} \equiv R_i(v_i, \theta_i, L_i), \quad (6)$$

where N_f is the number of all subfaults. We here define the functions P_i , Q_i , and R_i of simulation variables and frictional parameters for clarity and hereafter simply represent without arguments of each function. We solve Eqs. (5) and (6) using the adaptive time-step Runge–Kutta method (Press et al. 1996) to obtain the temporal evolution of simulation variables.

optimum value is conducted by iteratively updating the simulation variables and parameters based on the gradients of the cost function with respect to the simulation variables and parameters calculated using adjoint equations. This section briefly summarizes the outline of the adjoint data assimilation method for optimizing frictional parameters. The derivation of the adjoint equations for our problem is based on the study of Kano et al. (2015).

The steps of the adjoint data assimilation method used in this study can be summarized as follows:

Step 1: Assign the values of simulation variables \mathbf{x}_f and frictional parameters \mathbf{c}_f as starting values for parameter search as well as their error variance–covariance matrices \mathbf{B}_x and \mathbf{B}_c . These values, hereafter referred to as first-guess values, must be assigned in advance based on the prior knowledge of simulation variables and frictional parameters. In this study, we adopt the slip velocities estimated by the kinematic inversion analysis (Miyazaki et al. 2004) as the first-guess values. We set the first-guess values of state variables as $\theta_i = v_i/L_i$ for each subfault i under steady-state conditions. The corresponding error variance–covariance matrix \mathbf{B}_x is assumed to be diagonal, in which each diagonal element is set to be the square of the standard deviation of $\sigma_v = 1.0 \times 10^{-8}$ m/s for slip velocities (Miyazaki et al. 2004) and infinity for state variables because little information is available for state variables. As for the frictional parameters, we adopt the same first-guess values as Kano et al. (2015) in the numerical experiments summarized in Table 1 and spatially uniform values of $A-B = 40$ kPa, $A = 160$ kPa, and $L = 10$ mm for the application to the real data, which are roughly searched through a grid search. We do not explicitly consider \mathbf{B}_c as explained in Step 2.

Step 2: Define a cost function:

$$J(x_0, c) = \tilde{J}(x_0, x_1, \dots, x_{N_t}, c) = \frac{1}{2} \sum_{t=1}^{N_t} (Hx_t - d_t)^T R_t^{-1} (Hx_t - d_t) + \frac{1}{2} (x_0 - x_f)^T B_x^{-1} (x_0 - x_f), \quad (7)$$

Adjoint data assimilation method for our problem

The adjoint method searches an optimum value in a high-dimensional space of simulation variables and parameters of interest where the cost function, an indicator of the misfit between the observed and theoretical values, takes its minimum value. The search for the

where vector \mathbf{x}_t contains the simulation variables on day t , and \mathbf{c} is a frictional parameter vector. An observation vector \mathbf{d}_t consists of a cumulative displacement on day t , \mathbf{H} is an observation matrix connecting the simulation variables \mathbf{x}_t and the observed quantity \mathbf{d}_t and is calculated by Okada (1992). Mathematically, the only difference between Kano et al. (2015) and this study is

Table 1 True, first-guess, and optimized values of frictional parameters in each subregion (Fig. 2a) in the numerical experiments, and the values of cost function calculated by using Eq. (7)

Region	True values	First-guess values	Optimized values	
			Exp. 1	Exp. 2
1				
$A-B$ (kPa)	8.00	6.00	8.13	6.98
A (kPa)	60.0	60.0	37.0	50.2
L (mm)	30.0	50.0	47.3	49.0
2				
$A-B$ (kPa)	10.0	8.00	7.18	9.00
A (kPa)	40.0	40.0	37.3	49.7
L (mm)	10.0	10.0	9.73	9.03
3				
$A-B$ (kPa)	6.00	7.00	7.03	7.95
A (kPa)	20.0	30.0	28.0	21.0
L (mm)	20.0	18.0	17.3	17.0
4				
$A-B$ (kPa)	6.00	8.00	6.19	8.95
A (kPa)	20.0	40.0	20.5	30.6
L (mm)	20.0	20.0	20.3	19.0
5				
$A-B$ (kPa)	8.00	6.00	7.61	6.99
A (kPa)	60.0	60.0	68.8	69.7
L (mm)	10.0	10.0	9.96	9.04
6				
$A-B$ (kPa)	10.0	6.00	8.29	4.99
A (kPa)	40.0	60.0	63.2	50.0
L (mm)	10.0	10.0	10.5	11.0
7				
$A-B$ (kPa)	10.0	12.0	11.4	11.0
A (kPa)	40.0	50.0	56.1	60.0
L (mm)	20.0	17.0	18.3	18.0
8				
$A-B$ (kPa)	6.00	8.00	6.77	7.00
A (kPa)	20.0	20.0	29.5	30.0
L (mm)	30.0	20.0	21.4	21.0
9				
$A-B$ (kPa)	8.00	8.00	11.5	9.00
A (kPa)	60.0	40.0	33.7	30.0
L (mm)	10.0	8.00	9.17	9.00
Cost function (assimilation period)				
Exp. 1	0	3.80×10^1	2.41×10^{-2}	
Exp. 2	2.04×10^3	2.10×10^3		2.03×10^3
Cost function (prediction period)				
Exp. 1	0	2.82×10^2	4.58×10^{-1}	
Exp. 2	2.95×10^3	3.20×10^3		3.10×10^3

related to the observation matrix \mathbf{H} because we assimilate directly observed surface displacements, while Kato et al. (2015) utilized slip velocities on the plate interface. N_t is the number of assimilated epochs, i.e., 15 in this study. Hence, the first term on the right-hand side of Eq. (7) indicates the misfit between the theoretical and observed displacements at GNSS stations in the assimilation period, and is referred to as the misfit term. We assume that the variance–covariance matrix for observations, \mathbf{R}_t , is a time-invariant and diagonal matrix, where each element is set to be the square of the standard deviation of the sum of the observation error, 1 mm, and the modeling error defined as 20% of the cumulative displacement on day 15 in each component. We introduce the modeling error considering the assumption of homogeneous elastic media (Sato et al. 2007, 2010). The second term on the right-hand side of Eq. (7) provides a constraint for the parameters not to be optimized distantly from the first-guess values, and hereafter, is referred to the first-guess term. However, this term is zero when we focus only on the optimization of frictional parameters, except for the first subsection in the “Discussion” section where we attempt the simultaneous optimization for both frictional parameters and initial values of simulation variables. In addition, we consider the first-guess term only for the simulation variables as described in Step 1, but not for frictional parameters. This is equivalent to the assumption that elements of the error variance–covariance matrix for frictional parameters are infinite; thus, the frictional parameter vector \mathbf{c} can take any value. Note that we only impose non-negative constraints for A , B , and L in updating the parameters in Step 6. Thus, the parameters are optimized within a physically plausible range. Because the theoretical displacements at all epochs are uniquely calculated by setting the simulation variables at the initial time, or initial values \mathbf{x}_0 and frictional parameters \mathbf{c} , the cost function is a function of $(\mathbf{x}_0, \mathbf{c})$. Hereafter, we use J when referring to the cost function as a function of $(\mathbf{x}_0, \mathbf{c})$, while \tilde{J} is used as a function of the simulation variables at all days \mathbf{x}_t and parameters \mathbf{c} .

Step 3: Simulate the temporal evolution of simulation variables using the first-guess values.

Step 4: Calculate the value of the cost function in Eq. (7) using the theoretical displacements.

Step 5: Solve the following adjoint equations to obtain the gradients of the cost function with respect to the simulation variables $\partial J / \partial \mathbf{x}_0$ and frictional parameters $\partial J / \partial \mathbf{c}$:

$$-\frac{d\lambda_{v_i}(t)}{dt} = \frac{1}{Q_i} \left(\frac{B_i}{L_i} + \frac{A_i P_i}{Q_i V_i^2} \right) \lambda_{v_i}(t) - \sum_j \frac{k_{ji}}{Q_j} \lambda_{v_j}(t) - \frac{\theta_i}{L_i} \lambda_{\theta_i}(t) + \lambda_{s_i}(t), \quad (8)$$

$$-\frac{d\lambda_{\theta_i}(t)}{dt} = \frac{B_i}{Q_i \theta_i^2} \lambda_{v_i}(t) - \frac{v_i}{L_i} \lambda_{\theta_i}(t), \quad (9)$$

$$-\frac{d\lambda_{s_i}(t)}{dt} = \frac{\partial \tilde{J}}{\partial s_i}, \quad (10)$$

$$-\frac{d\lambda_{A_i}(t)}{dt} = -\frac{1}{Q_i} \left(\frac{R_i}{\theta_i} + \frac{P_i}{Q_i v_i} \right) \lambda_{v_i}(t), \quad (11)$$

$$-\frac{d\lambda_{A_i-B_i}(t)}{dt} = \frac{R_i}{Q_i \theta_i} \lambda_{v_i}(t), \quad (12)$$

$$-\frac{d\lambda_{L_i}(t)}{dt} = -\frac{B_i v_i}{Q_i L_i^2} \lambda_{v_i}(t) + \frac{v_i \theta_i}{L_i^2} \lambda_{\theta_i}(t), \quad (13)$$

where λ represents the adjoint variables. The subscripts of λ indicate the simulation variables or frictional parameters. These adjoint equations are numerically integrated backwards with time using $\lambda=0$ at the final time-step. In the backward integration, since only the term on the right-hand side of Eq. (10) is not multiplied by the adjoint variables, the adjoint variables are kept at zero until this term is added for the first time. This term is a derivative of the slip with respect to the cost function (Eq. (7)) and, thus, corresponds to the misfit between the theoretical and observed displacements when the observation is acquired. Thus, the adjoint equations integrate the information of the misfit through backward propagation. The resulting adjoint variable at initial time $\lambda(0)$ is equal to the gradient of the cost function with respect to the corresponding simulation variable or frictional parameter, i.e., the elements of $\partial J / \partial \mathbf{x}_0$ and $\partial J / \partial \mathbf{c}$.

Step 6: Update the simulation variables and frictional parameters of interest using the gradients derived in Step 5 based on the following steepest descent method:

$$\mathbf{x}_0^{(k+1)} = \mathbf{x}_0^{(k)} - \alpha_{\mathbf{x}} \frac{\partial J}{\partial \mathbf{x}_0^{(k)}}, \quad (14)$$

$$\mathbf{c}^{(k+1)} = \mathbf{c}^{(k)} - \alpha_{\mathbf{c}} \frac{\partial J}{\partial \mathbf{c}^{(k)}}, \quad (15)$$

where $\alpha_{\mathbf{x}}$ and $\alpha_{\mathbf{c}}$ control the magnitudes of the updates of the simulation values and frictional parameters at the k th iteration, respectively, which are set as:

$$\alpha_{v_{0i}} = \eta_{v_0} \left/ \frac{\partial J}{\partial v_{0i}^{(1)}} \right., \quad (16)$$

$$\alpha_{\theta_{0i}} = \eta_{\theta_0} \left/ \frac{\partial J}{\partial \theta_{0i}^{(1)}} \right., \quad (17)$$

$$\alpha_{A_i} = \eta_A \left/ \frac{\partial J}{\partial A_i^{(1)}} \right., \quad (18)$$

$$\alpha_{A_i-B_i} = \eta_{A-B} \left/ \frac{\partial J}{\partial (A_i^{(1)} - B_i^{(1)})} \right., \quad (19)$$

$$\alpha_L = \eta_L \left/ \frac{\partial J}{\partial L_i^{(1)}} \right.. \quad (20)$$

The coefficients η are determined as $\eta_A = 1.0 \times 10^4$ Pa, $\eta_{A-B} = 1.0 \times 10^3$ Pa, and $\eta_L = 1.0 \times 10^{-3}$ m in the numerical experiments following Kano et al. (2015) and $\eta_A = 5.0 \times 10^3$ Pa, $\eta_{A-B} = 5.0 \times 10^3$ Pa, and $\eta_L = 5.0 \times 10^{-2}$ m in applications to real data. In addition, we set $\eta_{v_0} = 1.0 \times 10^{-9}$ m/s and $\eta_{\theta_0} = 2.5 \times 10^3$ s when we additionally consider the optimization of the simulation variables. If the frictional parameters are updated to be negative A , B , and L , we reduce the corresponding coefficient η to $\eta/2$. Note that we allow negative $A-B$. Since we first only focus on the frictional parameters, the initial slip velocities and state variables are not updated but fixed to the first-guess values. However, this strategy is appropriate only if we have good estimates for the first-guess values of the simulation variables. Therefore, in the “Discussion”, we attempt to estimate both parameters and initial simulation variables. Because $\partial J / \partial \mathbf{x}_0$ are automatically calculated by the adjoint equations, Eqs. (8)–(13), in calculating $\partial J / \partial \mathbf{c}$, we can easily include the simulation variables as targets to be optimized.

Step 7: Simulate the temporal evolution of the simulation variables again.

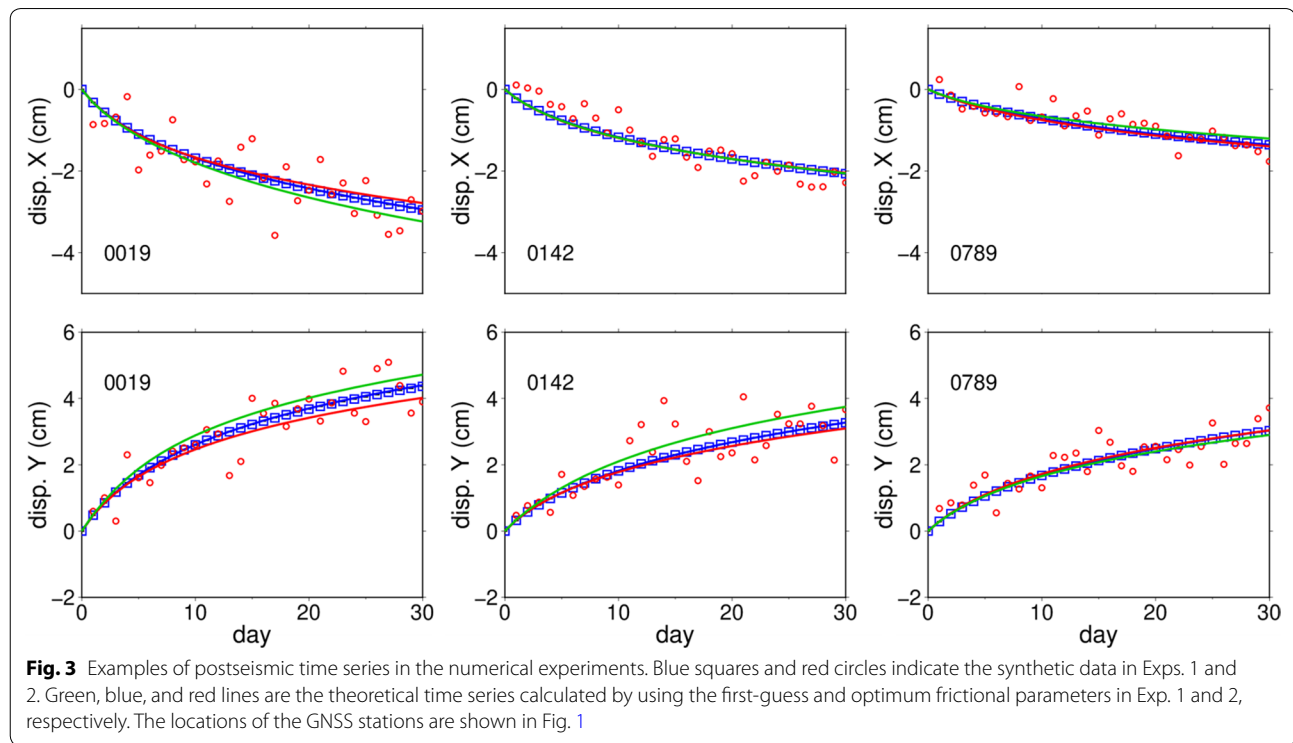
Step 8: Repeat Steps 4–7 iteratively until the cost function converges or until 10,000 iterations are conducted. We adopt the criterion of convergence as follows:

$$\left| \frac{J^{(k)} - J^{(k+1)}}{J^{(k+1)}} \right| < 1.0 \times 10^{-6}. \quad (21)$$

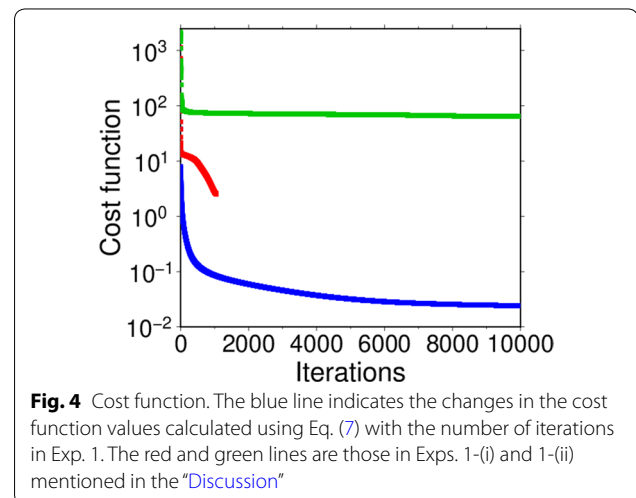
Numerical experiments based on synthetic data

Synthetic data

The developed adjoint data assimilation method are validated based on two numerical experiments, Exps. 1 and



2, using synthetic postseismic time series. This subsection describes how to generate the synthetic data used in numerical experiments using physics-based simulations. The procedure is the same as in Kano et al. (2015), except for the observed physical quantity used for data assimilation. We assume the spatial heterogeneity of frictional parameters (Fig. 2a) and set their values as summarized in Table 1, which are referred to as true values. As for the initial values of the slip velocities, we assume the slip velocities estimated by Miyazaki et al. (2004) on day 00, the day after the Tokachi-oki earthquake. The initial values of the state variables are defined as $\theta_i = v_i / L_i$. By assigning these initial values and true frictional parameters, we simulate the spatio-temporal evolution of the slip velocity and the corresponding surface displacement in both X and Y directions at all the GNSS stations for 30 days. The surface displacements are sampled daily and utilized as synthetic data in Exp. 1. In addition, we add the observation noise that obeys the Gaussian distribution with an average of 0 and the standard deviation of the sum of the observation error and the modeling error, defined as 1 mm and 20% of the cumulative displacement on day 15 in each component. The degree of the observation noise can be considered as the variance–covariance matrix for observations \mathbf{R}_t in Step 2 mentioned in the previous subsection. These synthetic data with noise are utilized in Exp. 2. For both experiments, we assimilate the synthetic data for the first 15 days in the assimilation



period and use those for the later 15 days in the prediction period for testing the prediction skill of our method. The synthetic data are summarized in Fig. 3 and Additional file 1: Figure S1, showing a logarithmic transient curve with a maximum displacement of ~ 3.0 cm in X and ~ 4.5 cm in Y directions for 30 days.

Results

The adjoint data assimilation method iteratively reduced the cost function for both experiments

(Table 1 and Fig. 4). In Exp. 1, the number of iterations reached its maximum value of 10,000 iterations in our experimental setting (Fig. 4). The values of the cost function significantly decreased from its initial value of 3.80×10^1 calculated by using the first-guess frictional parameters and did not largely change after ~ 1000 iterations. Thus, we consider that the cost function converged and define the frictional parameters after 10,000 iterations as the optimized values in Exp. 1. The optimization finished only in two iterations in Exp. 2. The value of the cost function decreased from 2.10×10^3 to 2.03×10^3 after the optimization, which is smaller than 2.04×10^3 calculated by using true values. Because synthetic data in Exp. 2 include the observation noise, the global minimum of the cost function does not necessarily exist at the true parameter values, and the optimized values indicate a location close to at least one of the local minimums of the cost function.

Table 1 presents the comparison of the optimized frictional parameters with the true values as well as the first-guess values, showing that some parameters are optimized close to the true values, while some are not. To investigate the reliability of the optimized parameters, we calculated the cost functions by varying one frictional parameter with the other ones fixed to their optimized values (Fig. 5). Because we search for a 27-dimensional parameter space, this calculation cuts a one-dimensional cross-section of such a

high-dimensional space, which roughly reflects the parameter sensitivity with respect to the cost function. If values of the cost function largely vary depending on a value of one parameter, then the theoretical surface displacement is sensitive to the parameter change. If not, a small change of the parameter does not affect the cost function; in other words, the observed data rarely constrain the parameter. Furthermore, because the theoretical displacement is a sum of the product of fault slips and the observation matrix \mathbf{H} , the parameter sensitivity to the cost function includes the contributions of both the parameter sensitivity to the slip velocities in each region and the amplitude of the corresponding element of \mathbf{H} , which are determined by the plate geometry and the locations of the observation sites, irrespective of the parameters.

In Exp. 1, the shapes of the cost function for most parameters can be approximated by parabolic functions with vortices close to their optimized values (stars in Fig. 5). In regions 4–9 where significant afterslip occurs, the values of the cost function largely change with the parameter values. The large curvature in these regions may reflect high sensitivity to the observations. Such variations are smaller in regions 1 and 3 because there is little afterslip in these regions, and the parameters rarely contribute to the observed displacement. Despite the differences in sensitivity, the frictional parameters are optimized such that each cross-section of the cost

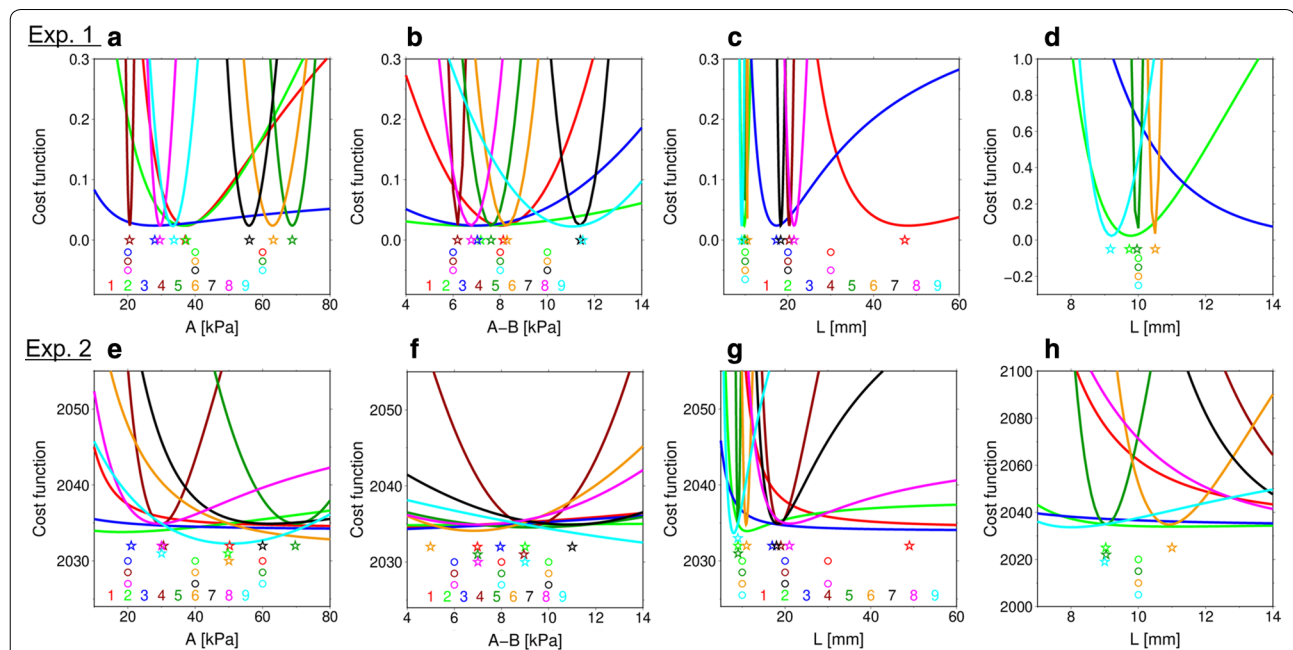
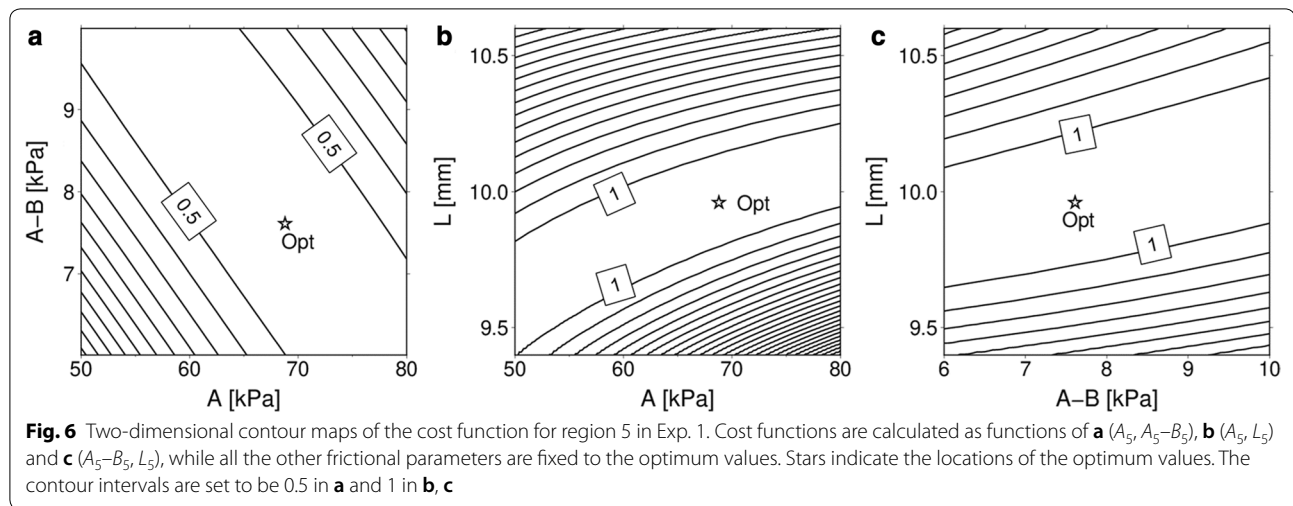


Fig. 5 Cost function with respect to parameter changes in the numerical experiments. **a–d** Cost functions in Exp. 1 as a function of each frictional parameter of **a** A , **b** $A-B$, and **c** L , and **d** the enlarged view of the horizontal axis of **c**. The color of each line corresponds to the number of region shown in Fig. 2a. Stars and circles indicate the optimum and true frictional parameters with corresponding colors. **e–h** Same as **a–d**, but for Exp. 2



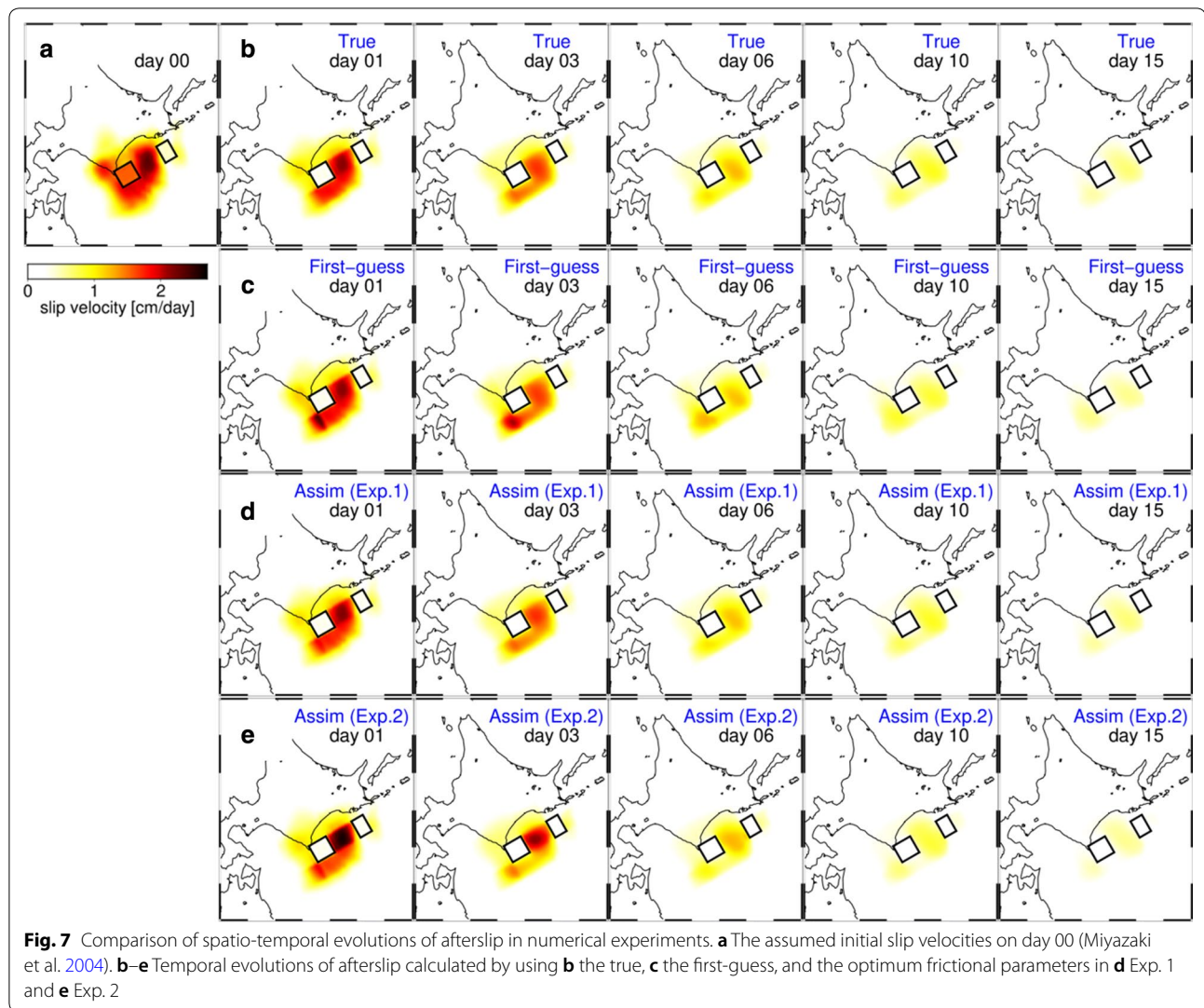
function takes its minimum value. Even in such a situation, the optimized values are not always close to the true values. Figure 6 shows the two-dimensional contour maps of the cost function by, respectively, changing two frictional parameters in region 5. In this calculation, all the other parameters are fixed to the optimized ones. These contour maps indicate the trade-offs of each two-parameter set, and the optimized parameters are located at the trough of the contours. It is not possible to search the parameters in the directions where the cost function is almost flat, which may result in a discrepancy between the true and optimized parameters.

If we include the observation error in the synthetic data (Exp. 2), the cross-sections of the cost function indicate the parabolic forms as in Exp. 1 (Fig. 5a–d). The difference is that the shapes of the parabolic functions become broader for all the parameters in Exp. 2 than those in Exp. 1; i.e., the variation of a frictional parameter changes the cost function values slightly in Exp. 2 (vertical axis for Exp. 2 in Fig. 5e–h is larger than that for Exp. 1 in Fig. 5a–d). Thus, the observation noise results in larger estimation errors. The characteristics of the relative sensitivity among regions are similar to the case of noise-free synthetic data. For example, A in region 4 is the most sensitive to the cost function, while those in regions 1–3 are less sensitive (Fig. 5a and e). Therefore, the inclusion of the observation noise does not change the relative sensitivity of the parameters among regions with respect to the cost function, but such relative sensitivity can be determined by the physics of the governing equations and the observation matrix. In addition, the cost function for some parameters such as L in regions 1 and 3 does not show its minimum at the optimized values but monotonically decreases, while those for most parameters become minimum close to their optimized values as

in Exp. 1. If some parameters are much more sensitive to the cost function than the others, the parameters cannot update when the highly sensitive parameters are close to the minimum even though it is possible for less sensitive parameters to update. Figure 5g represents the situation that L in regions 1 and 3 with few afterslip cannot update any longer. A similar situation can be observed in the case of assimilating real data, which will be presented in the next section.

Although the optimized values are not always close to the true ones, both experiments revealed good estimates in terms of the cost function, i.e., good fit to the data. Figures 3 and Additional file 1: Figure S1 display the comparison of the theoretical displacements and the synthetic observation data. Note that these results are for the assimilation period (day 1–15); those for the prediction period (day 16–30) are presented later in this section. The theoretical displacements (blue lines) in Exp. 1 fit the observed data without noise (blue squares). The green lines in Fig. 3 and Additional file 1: Figure S1 represent the theoretical displacement using the first-guess values. By updating the frictional parameters through data assimilation, the theoretical displacements were modified from green to blue lines, resulting in a quantitatively better fitting to the observations. This is also true for Exp. 2: the green line was updated to the red one by assimilating the synthetic data (red circles).

Figure 7 compares the snapshots of the spatial distribution of afterslip using the true, first-guess, and optimized frictional parameters. In the case of the first-guess parameters (Fig. 7c), afterslip continues to accelerate for the first few days in region 9, south of TA, which is not inferred in the case of the true parameters (Fig. 7b). This slip acceleration disappeared after assimilating noise-free synthetic data in Exp. 1 (Fig. 7d), and the temporal

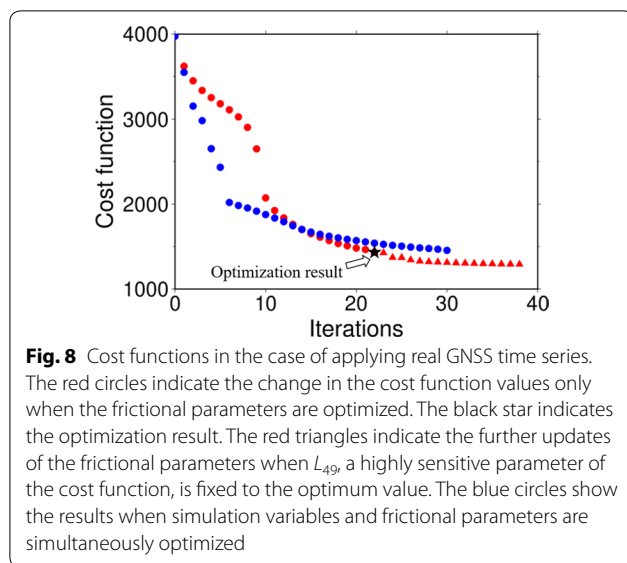


changes in slip velocity were modified close to the true value in region 9. Moreover, the slip velocity calculated using the optimized parameters in Exp. 1 closely resembles the true ones in all regions. When synthetic data with observation error are assimilated in Exp. 2 (Fig. 7e), the estimated afterslip distributions are mostly consistent with the true ones, with some differences in detail. For example, the acceleration of afterslip is inferred for day 1 in region 5, to the east of the TA (Fig. 7e). Therefore, even though the observed data can be reproduced, the completely true recovery of the slip distribution is not always possible from the assimilation of crustal deformation on the ground surface with observation noise, but a consistent spatial pattern can be roughly inferred.

To investigate the prediction skill of the data assimilation method, we compare crustal deformation for the prediction period (day 16–30) (Fig. 3 and Additional file 1: Figure S1, Table 1). Figure 3 and Additional file 1:

Figure S1 indicate that, for both experiments, the theoretical displacements with the optimized values better explain the synthetic data for the prediction period than those with the first-guess values, even though the data for the prediction period are not assimilated. The cost function calculated only for the prediction period is smaller for the optimized values than for the first-guess values (Table 1). Therefore, the assimilation of the crustal deformation data for the first 15 days leads to the more accurate prediction of postseismic deformation for the following 15 days.

In summary, the adjoint data assimilation was adopted for the synthetic afterslip data with and without the observation noise to optimize the frictional parameters. Frictional parameters sensitive to the data were optimized so that the cost function could be minimized along the cross-section of each parameter. The optimized parameters well reproduced the synthetic data in the



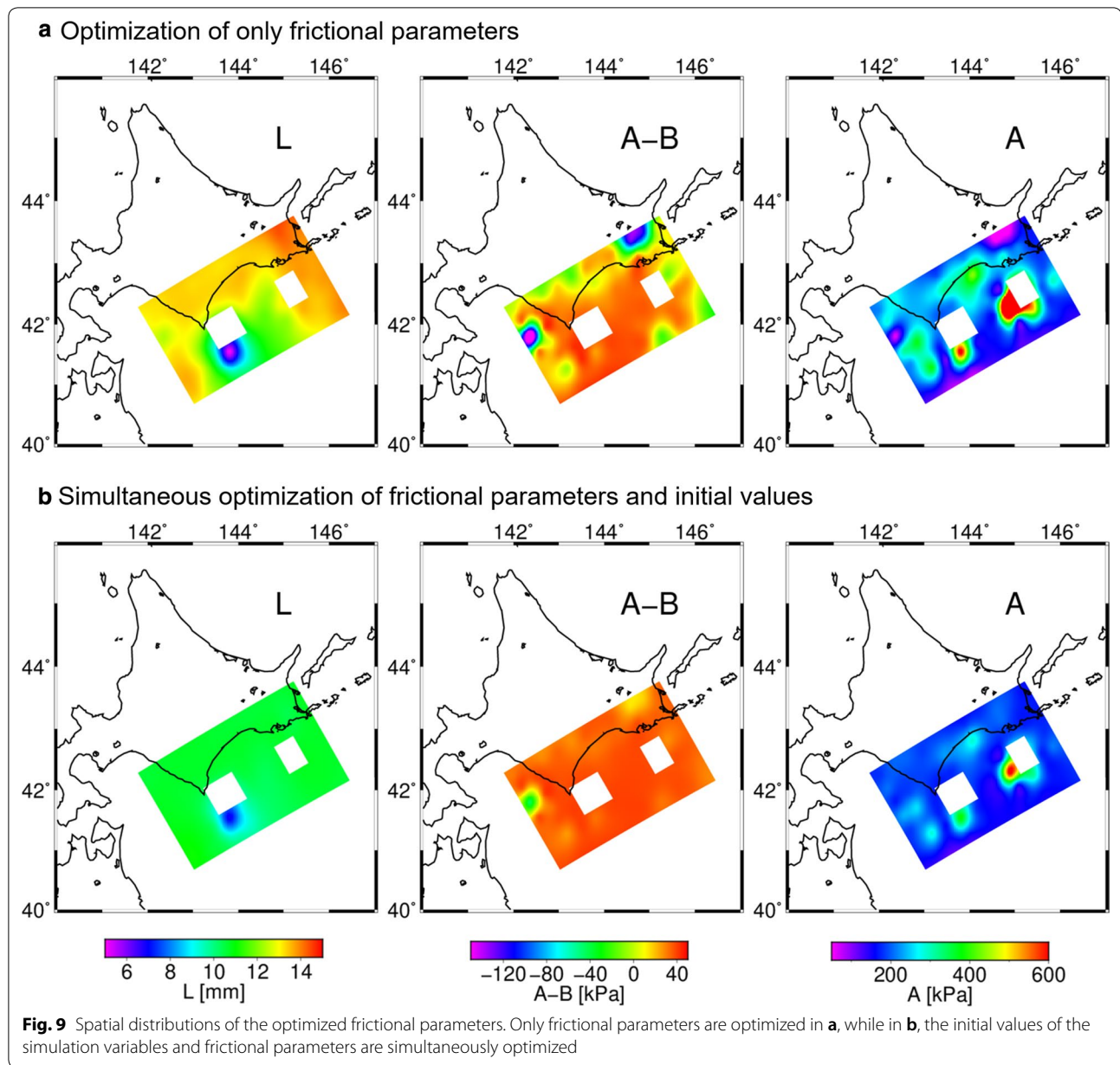
assimilation period, although there was a difference for the slip velocity on the plate interface calculated by using the true and optimized parameters, reflecting a limitation for resolving a fault slip when using crustal deformation data on the ground surface with observation noise. Nonetheless, the prediction of the time series of crustal deformation was improved through data assimilation.

Application to the GNSS data following the 2003 Tokachi-oki earthquake

Figure 8 (red circles) illustrates the updates of the cost function values starting from the first-guess values. The cost function value decreased from 3.97×10^3 before data assimilation to 1.43×10^3 after 22 iterations. The frictional parameters, $A-B$, A , and L , were optimized as $O(10 \text{ kPa})$, $O(10^2 \text{ kPa})$, and $O(10 \text{ mm})$, respectively (Fig. 9a). To investigate the robustness of the parameters, we calculated the values of the cost function by changing one parameter, while all the other parameters are fixed to their optimum values. As the number of parameters is large, we cannot calculate the values of the cost function for many cases as in Fig. 5; instead, we simply calculate the cost function with a small perturbation of $\delta(A-B)$, δA , and δL added to or subtracted from the optimized value and, respectively, obtain the difference in the cost function values in the optimized case. Here, we assume $\delta(A-B) = 1 \text{ kPa}$, $\delta A = 10 \text{ kPa}$, and $\delta L = 1 \text{ mm}$. If the cost function becomes larger for both adding and subtracting small perturbations for a certain parameter, the shape of the cross-section of the cost function is roughly approximated as a parabolic function with a vortex of optimized values. Such a parabolic function is rarely indicated, but most parameters

monotonically decrease or increase around the optimum values: the cost function becomes smaller (larger) when the perturbation is added and larger (smaller) when it is subtracted because L in region 49 (Fig. 2b), where a large afterslip is inferred (Fig. 10a), is highly sensitive to the data. In such a case, the other parameters are rarely updated after the highly sensitive parameter is optimized. Therefore, we conduct additional data assimilation starting from optimized values under the same conditions, but the parameters in region 49 are fixed to the optimized values. As expected, the cost function further decreases (red triangles in Fig. 8), and the other parameters are further updated. In this procedure, it is implicitly assumed that the parameters in region 49 are orthogonal to the other parameters, but further investigation is necessary to validate this assumption. Even though further optimization may occur, we define the first optimization results (indicated by the black star in Fig. 8) as “optimized parameters” as more subjectively determined parameters. Figure 10a shows the spatio-temporal evolution of the slip velocities. The large afterslip area is located on the northeastern side of the TA just after the mainshock on day 00 (Miyazaki et al. 2004). The main slip location soon moved to the shallower side of the TA and continued to slip, exceeding 2 cm/year in the next 6 days, followed by a gradual decay until the end of the assimilation period. The time constant for decaying afterslip on the shallower side of the TA is longer than that on the northeastern side of the TA, which is consistent with the results of Itoh et al. (2019) although they analyzed a much longer time series of 7.5 year of postseismic deformation.

Figure 11 and Additional file 1: Figure S2 present the comparison of the theoretical displacements calculated by using the optimum parameters (red and blue lines) and those by using the first-guess parameters (pink and light blue lines) with the observed GNSS time series (circles). The theoretical displacements calculated using the optimized parameters reproduce the observed postseismic time series, especially at stations where large postseismic signals were identified. Station 0019 shows the largest observed displacement of ~ 4.1 and ~ 7.4 cm in X and Y directions, respectively, for the first 15 days following the mainshock, and the theoretical displacement of ~ 3.5 and ~ 7.9 cm in X and Y directions, almost similar to the observations (Fig. 11). These theoretical displacements better fit to the observations than those by the first-guess parameters (pink and light blue lines). The daily misfit (red circles in Fig. 12), which is defined as same as the first term on the right-hand side of Eq. (7) but before summing up for index t , decreased



by assimilating the observations compared to the case without data assimilation (pink circles).

The optimum parameters have the prediction skill of the GNSS time series. In most stations, the predicted time series with the optimum values show a higher degree of agreement with the observations than the first-guess parameters in the prediction period (day 16–30) as well as the assimilation period (day 1–15). When the misfit values of all stations are summed for each day (Fig. 12), they decrease in the prediction period by assimilating the observations in the previous 15 days even though the data in the prediction period

are not assimilated. Thus, the prediction skill of the GNSS time series can be improved through the adjoint data assimilation in the applications to the real GNSS time series.

Discussion

Simultaneous estimation of the initial values of the simulation variables

The theoretical GNSS time series can be calculated by adopting the frictional parameters and the initial values of the simulation variables, i.e., slip velocity and state variables. In previous sections, we focused only on

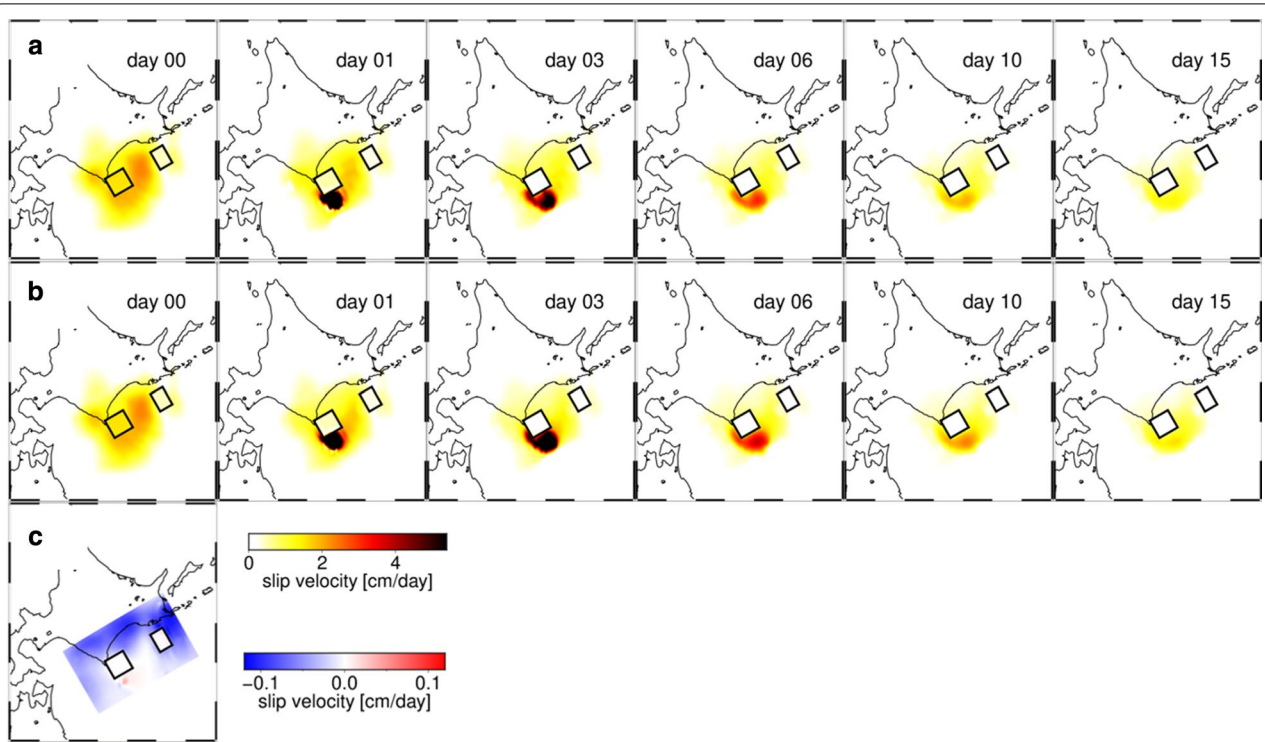


Fig. 10 History of afterslip following the 2003 Tokachi-oki earthquake. The spatio-temporal evolutions are calculated by using the optimized frictional parameters **a** without and **b** with the optimization of the initial velocities on day 00. **c** The difference in the initial slip velocities between **a** and **b**

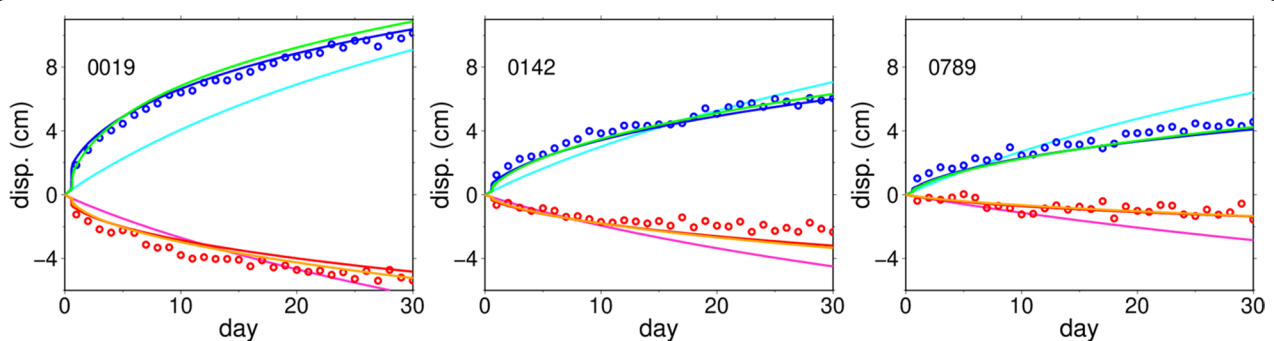
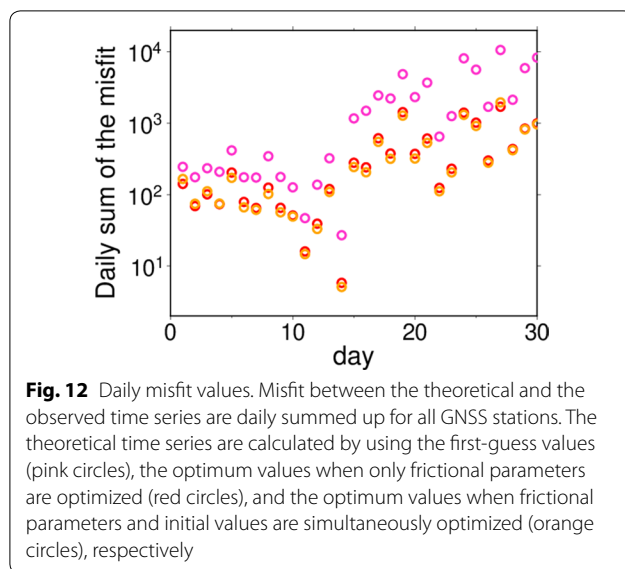


Fig. 11 Examples of postseismic time series in the application to real GNSS time series. The red and blue circles indicate the GNSS time series in X and Y directions. The pink, red, and orange lines are the theoretical time series in the X direction calculated by using the first-guess values, the optimum values when only frictional parameters are optimized and the optimum values when frictional parameters and initial values are simultaneously optimized, respectively. The light blue, blue, and green lines indicate the corresponding theoretical time series in the Y-direction. The locations of the GNSS stations are indicated in Fig. 1

the optimization of frictional parameters, but fixed the pre-determined first-guess values for the initial values. In adjoint data assimilation, gradients of the cost function to the initial values and to the frictional parameters are simultaneously computed. Therefore, we conducted data assimilation for optimizing the initial values as well as frictional parameters. In this case, the number of

variables to be optimized increased from 213 (frictional parameters) to 1545 (frictional parameters, initial values for slip velocity and state variable for all subfaults).

As a result, the cost function gradually decreased and finally satisfied the convergence condition after 30 iterations (blue circles in Fig. 8). The value of the resulting cost function is 1.45×10^3 , which is slightly larger than



that in the case with no optimization of the initial values (Table 2). When the initial values were optimized, the first-guess term, i.e., the second term on the right-hand side of Eq. (7), had a finite value (1.17×10^2), which was zero when only the frictional parameters were optimized. Therefore, the simultaneous optimization of the initial values resulted in slightly better estimates if only the misfit terms were evaluated.

The history of the afterslip distribution is summarized in Fig. 10b. The optimized initial velocity on day 00 is not significantly different from that in the case without optimization (Fig. 10a). The calculation of the difference between the optimized and first-guess initial velocities (Fig. 10c) showed that the initial velocity became smaller for most subfaults, by ~ 0.1 cm/day on deeper subfaults, but increased slightly on the shallower side of the TA. The temporal evolution of afterslip was slightly different: the decay time for afterslip was longer on the shallower side of the TA when the initial values were optimized.

Similarly, the theoretical displacements decayed more slowly (Fig. 11 and Additional file 1: Figure S2), which resulted in a larger displacement after a long time period (e.g., day 30). The simultaneous optimization yielded a

slightly better reproduction of the observed displacement in the assimilation period. This also holds for the prediction period, during which the cost function and the daily misfit show slightly smaller values (Table 2, Fig. 12).

The spatial distribution of the frictional parameters was also different (Fig. 9b). Most frictional parameters did not largely change from the first-guess values. Instead, some parameters, which are spatially isolated such as those in the shallower side of the TA were updated. The spatial characteristics of these updated parameters (e.g., smaller L and larger A on the shallower side of TA) were estimated even when only frictional parameters were optimized. This result indicates that the cost function is smaller when the initial values are updated than most frictional parameters because the observed crustal deformation is more sensitive to the initial values, especially slip velocity, compared to most frictional parameters.

We assume that the error of the first-guess values for slip velocity is $\sigma_v = 1.0 \times 10^{-8}$ m/s ($= 8.64 \times 10^{-2}$ cm/day) based on the inversion result of Miyazaki et al. (2004). If we impose less constraint on the first-guess values of the initial slip velocity as $\sigma_v = 1.0 \times 10^{-7}$ m/s, the resulting cost function becomes 1.35×10^3 , consisting of the first-guess term of 5.03 and the misfit term of 1.30×10^3 . At a larger error value σ_v , the first-guess term is smaller, while the misfit term is almost the same as in the case of smaller σ_v . Then, the optimized initial velocities are largely updated from the first-guess values (Additional file 1: Figure S3), resulting in a faster slip on the shallower side of the TA by up to 0.40 cm/day. The following afterslip and the theoretical displacement become large as well (Additional file 1: Figure S4). The optimized frictional parameters in Additional file 1: Figure S5 largely resemble those in the case of smaller σ_v (Fig. 9b), exhibiting slightly smaller updates. Thus, if we allow less constraint on the first-guess initial velocities, the cost function decreases by the update of the initial values rather than the frictional parameters as in the previous case (Figs. 9b and 10c).

The assimilation results mentioned in this subsection do not have significant impacts on the estimation of the afterslip evolutions and the fitting to the observations, even when the initial values are simultaneously

Table 2 Values of cost functions calculated by using Eq. (7) in the case of real GNSS time series

	Optimization		Cost function	
	Frictional parameters	Initial values	Assimilation period	Prediction period
First-guess values			3.97×10^3	6.12×10^4
Optimized values	Optimized	Fixed	1.43×10^3	1.07×10^4
	Optimized	Optimized	1.45×10^3 (including first-guess term of 1.17×10^2)	1.03×10^4

estimated (Figs. 10, 11, 12). If we have reasonably good estimates of the simulation variables, the optimization of the initial simulation variables is not always necessary. However, in practical applications of adjoint data assimilation, prior knowledge of the initial values is usually limited. Thus, it may be reasonable to adopt an inverted slip velocity obtained from the kinematic analysis as the first-guess values for the initial slip velocities. Nonetheless, inversion results are sometimes biased when prior constraints such as the spatial and/or temporal smoothing of the slip are assumed. In such a case, the amount of error of the first-guess initial velocities would be set larger than that obtained in the kinematic analysis so that the initial velocity can be largely updated with less constraint.

Did the fault system reach steady state on the next day of the mainshock in the Tokachi-oki region?

The first-guess initial state variables were assigned by assuming steady state, i.e., $v\theta/L=1$, and using the first-guess frictional parameter L . In the case of optimizing only the frictional parameters, the initial values of v and θ were fixed to the first-guess values, while L was optimized in the data assimilation. Since the optimized L was different from the first-guess values (Fig. 9a), the steady-state assumption is not valid as in the spatial distribution of $v\theta/L$ on day 00 (Fig. 13a). On the contrary, when both initial values and frictional parameters were optimized (Fig. 13b), the spatial pattern of $v\theta/L$ was similar to Fig. 13a, showing that the steady-state assumption does not hold even if the initial slip velocities and state variables are included as the optimized variables. Figure 13c and d shows examples of the temporal evolution of $v\theta/L$

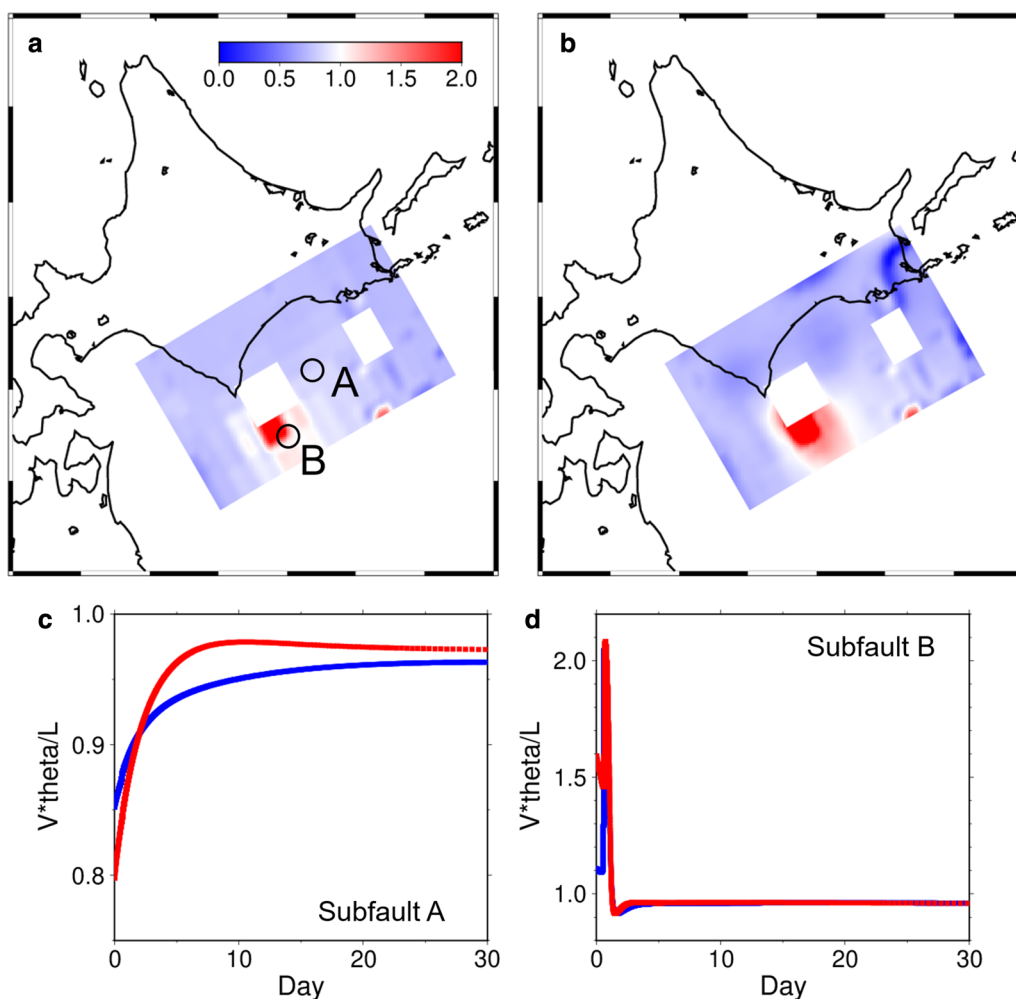


Fig. 13 Spatial distributions of $v\theta/L$ and examples of the temporal evolution. The spatial distributions of $v\theta/L$ of **a** only when the frictional parameters are optimized, and **b** both initial values and frictional parameters are optimized. The temporal evolutions of $v\theta/L$ at subfaults **c** A and **d** B are indicated by blue and red lines, which correspond to the results of **a** and **b**, respectively. The locations of subfaults A and B are indicated in **a**

at subfaults A and B, located on the northeastern and shallower side of TA. These figures indicate that it took at least a few days to reach the steady state. Based on these results, we consider that the physical state on the next day of the Tokachi-oki earthquake did not reach a steady state. This fact is different from the steady-state assumption that was introduced to impose coseismic stress change to investigate the postseismic behavior (Fukuda et al. 2009), which should be further investigated.

Comparison with the frictional parameters estimated from the Tokachi-oki afterslip region

The frictional parameters are optimized as $A-B \sim O(10 \text{ kPa})$, $A \sim O(10^2 \text{ kPa})$, and $L \sim O(10 \text{ mm})$. In the Tokachi-oki region, frictional parameters were estimated using the postseismic GNSS time series of the 2003 Tokachi-oki earthquake (Miyazaki et al. 2004; Fukuda et al. 2009). Miyazaki et al. (2004) estimated $A-B$ to be on the order of $O(10^2 \text{ kPa})$, which is one-order larger than our result, based on the relation between the kinematically inverted slip velocity and the stress change assuming a steady state. Since the analysis period in Miyazaki et al. (2004) was the same as in this study, the difference of the parameters may result from the difference of the forward modeling and the assumed fault model, that is, Miyazaki et al. (2004) utilized a kinematic forward model in a single spring–slider system, while we adopted the quasi-dynamic model with an elastic continuum model. The other difference is that Miyazaki et al. (2004) assumed a steady state. As we discussed in the previous subsection, this assumption did not hold at least within a few days following the mainshock in a continuum model. Nonetheless, to investigate the impact of parameter optimization under the assumption of a steady state, we conducted a parameter search assuming the spatial homogeneity of $A-B$ by a one-dimensional grid search using the same GNSS data and the first-guess initial velocities. The parameter search was conducted within the range of $A-B$ from 0 to 1,600 kPa with an interval of 4 kPa. As a result, an $A-B$ of 8 kPa attained the minimum value of the cost function of $J = 2.98 \times 10^3$, which was almost twice as high as that in our assimilation result (Table 2). If we assume $A-B = 100 \text{ kPa}$, the cost function became much larger value of $J = 6.40 \times 10^3$, indicating worse fitting to the observation. This result of a smaller $A-B$ compared to that without a steady-state assumption suggests that the assumption in our model would further enlarge the difference between the parameter estimations between Miyazaki et al. (2004) and this study.

Fukuda et al. (2009) focused on the early afterslip in 5 h following the 2003 Tokachi-oki mainshock and estimated $A-B \sim O(100 \text{ kPa})$, $A \sim O(100 \text{ kPa})$, and $L \sim O(1 \text{ mm})$ using a single spring–slider system. The parameter A was on

the same order as our results, while $A-B$ and L were estimated to be one-order larger and smaller, respectively. Although these differences may be partly caused by the single spring–slider model as discussed above in the case of Miyazaki et al. (2004), the main reason is likely the difference in the data period: Fukuda et al. (2009) focused on the 5 h directly after the main shock, while we assimilated the data for the following 15 days.

We showed that there are trade-offs among the parameters in region 5, even in the simple numerical experiment (Exp. 1) (Fig. 6), indicating a negative trade-off between $A-B$ and A and positive trade-offs between A and L , and $A-B$ and A . Trade-offs among parameters were also investigated in Fukuda et al. (2009), which showed the negative trade-off between A and L and no clear trade-offs between $A-B$ and A or L . These patterns are different from our results as well. We guess that these differences also arose from the assumed model. For this difference, the number of the frictional parameters to be optimized was 27 in this study, larger than that of 3 in Fukuda et al. (2009), and thus, this resulted in the complex trade-off relations among this large number of parameters.

Comparison with the results of assimilating slip velocity data

The numerical experiment Exp. 1 revealed that the frictional parameters were not optimized close to the true values because of the trade-offs among parameters (Fig. 6) even though the observation noise was not considered. To investigate the origin of these trade-offs, we conducted an additional numerical experiment using the synthetic slip velocity data without the observation noise under the same experimental settings as in Exp. 1. Then, we conducted a two-dimensional grid calculation by changing two of the frictional parameters in region 5 by fixing all the other frictional parameters to the optimized ones. Figure 14a shows the contour maps of the cost function, which is defined by the sum of the misfits between the theoretical and observed slip velocities at all subfaults. All the contour maps indicate similar trade-offs to those found when the surface displacement was assimilated (Fig. 6). Furthermore, Fig. 14b shows the contour maps of the pairs of frictional parameters in region 5 in the case of fixing all the other parameters to their true values, showing similar trade-offs to those in Fig. 14a. These results indicate that the trade-offs among frictional parameters shown in Figs. 6 and 14 are originated not by the difference in the assimilated physical quantities, i.e., the difference of the observation matrix \mathbf{H} , but by the nature of our data assimilation system including the physics-based governing equations and the observed phenomena. Therefore, the developed data assimilation system has low resolution along the troughs of the contour maps.

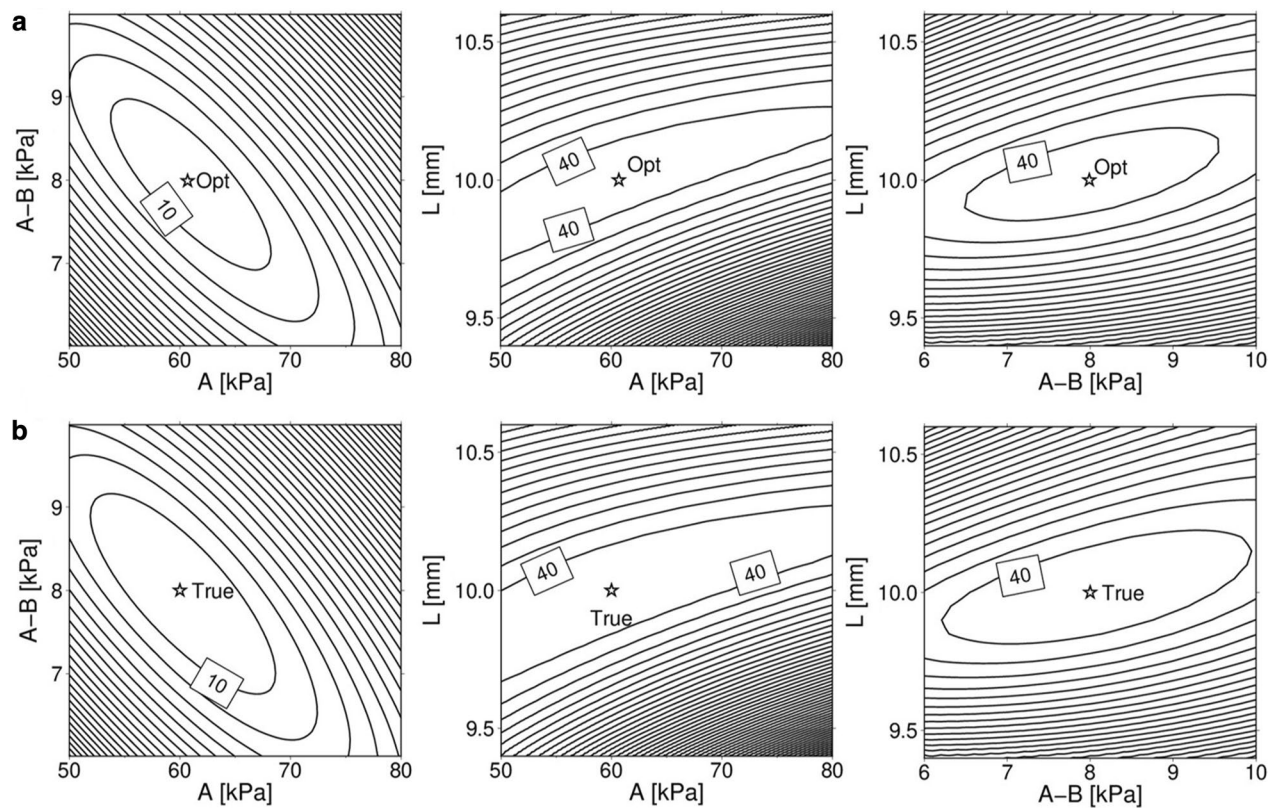


Fig. 14 Two-dimensional contour maps of the cost function when slip velocities are assimilated. Cost functions are calculated as functions of (A_5, A_5-B_5) , (A_5, L_5) , and (A_5-B_5, L_5) , while all the other frictional parameters are fixed to **a** their optimum values and **b** true values. Stars indicate the locations of **a** the optimum values and **b** the true values. The contour intervals are set to be 10 for (A_5, A_5-B_5) and 40 for (A_5, L_5) and (A_5-B_5, L_5)

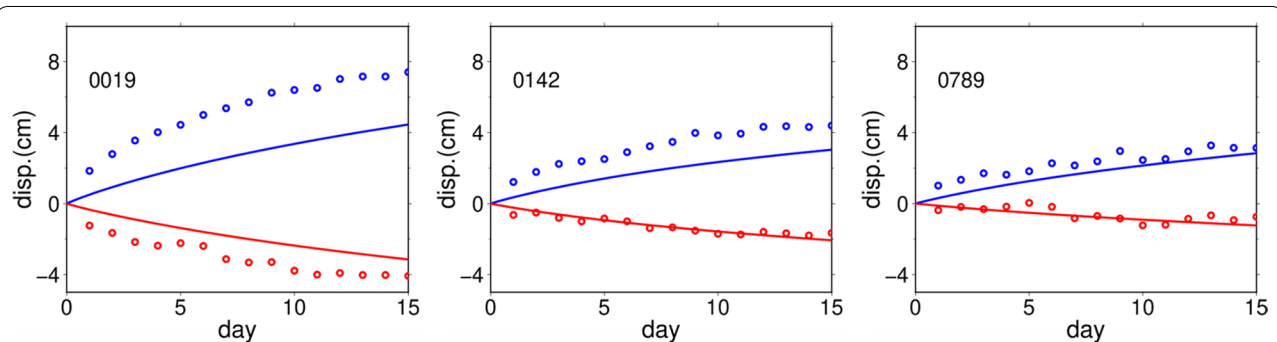


Fig. 15 Examples of the postseismic time series using the results of Kano et al. (2015). Red and blue circles indicate the GNSS time series in the X and Y directions. The red and blue lines are the corresponding theoretical time series calculated by using the optimum values obtained in Kano et al. (2015), which assimilated the slip velocities estimated by Miyazaki et al. (2004) as observations

Kano et al. (2015) estimated the spatial distribution of the frictional parameters in the afterslip region following the 2003 Tokachi-oki earthquake using an adjoint data assimilation method. The only difference between this study and Kano et al. (2015) is the assimilated observation data: Kano et al. (2015) adopted the slip velocity on

the plate interface inferred from the GNSS data using the kinematic inversion analysis (Miyazaki et al. 2004) as “observation” as if they were in situ observations. Although the spatial variations of the frictional parameters in Kano et al. (2015) are not similar to those in this study, the frictional parameters are optimized in

the same order of $L \sim O(10 \text{ mm})$, $A \sim B \sim O(10 \text{ kPa})$, and $A \sim O(100 \text{ kPa})$. Figure 15 and Additional file 1: Figure S6 display the comparison of the theoretical surface displacement calculated by using the optimized frictional parameters in Kano et al. (2015) with the GNSS time series used in this study. The theoretical time series underestimates the displacement in the Y-direction at stations where large postseismic deformation was observed. The resulting cost function is 3.71×10^3 , which is much larger than our result of 1.43×10^3 . This difference may be caused by the difference in the assimilated data. The slip velocity data in Kano et al. (2015) were obtained in the kinematic inversion analysis where the inverted daily slip velocity indicates the daily averaged slip velocity. Therefore, it may not always be reasonable to fit the daily averaged slip velocity by the continuously varying slip velocity calculated using physics-based equations. The kinematic analysis often imposes the smoothing of the slip velocity as a prior constraint resulting in strong correlations between slip velocities in neighboring subfaults. Nonetheless, Kano et al. (2015) did not consider the covariance matrix of the observed assimilated slip velocity, assuming that slip velocities were independently observed. Inclusion of such a covariance effect, i.e., the consideration of the covariance matrix in the cost function, may improve the optimization of in situ observations, which is beyond the scope of this study.

Effect of the first-guess values of frictional parameters

In the numerical experiments, we assumed the first-guess frictional parameters adopted in Kano et al. (2015), which were not so far from the true values (Table 1). Kano et al. (2015) roughly determined these values based on a two-dimensional grid calculation by changing two of the frictional parameters in each region. Therefore, these first-guess values were suitable to some degree. To investigate the effect of the choice of the first-guess values, we additionally conducted two numerical experiments using the noise-free synthetic data adopted in Exp. 1 by setting two different first-guess values: Exp. 1-i, which used the first-guess values with 100% perturbation added to the true values and Exp. 1-ii, where 900% perturbation were added to (or one-order larger from) the true values. The optimization results are summarized in Table 3 with the first-guess frictional parameters. The cost function calculated using these first-guess values are 1.14×10^3 and 2.16×10^3 in cases Exp. 1-i and Exp. 1-ii, respectively, which are much larger than 3.80×10^1 for Exp. 1 (Table 1). The adjoint data assimilation reduced these values to 2.51×10^0 and 6.45×10^1 after 1,046 and 10,000 iterations (Fig. 4), respectively, which were larger than the result (2.41×10^{-2}) in Exp. 1. The theoretical surface displacements are compared in Fig. 16, indicating that

the ability to explain the observations becomes worse as the time from the mainshock elapsed when larger perturbations are added in Exp. 1-ii. Especially in the prediction period (day 16–30), the misfit terms in the cost function in Exps. 1-i and 1-ii become four-orders of magnitude larger than in Exp. 1. These results indicate the necessity to estimate good first-guess frictional parameters or starting points for the parameter search. This will be discussed further in the next subsection, in terms of future improvements.

Future improvements

Our assimilation results can quantitatively reproduce the daily postseismic deformation both in the assimilation and prediction periods. When we focus on the short time-scale deformation of the theoretical displacement, there is a step-like movement between day 0 and 1 regardless of the initial values being optimized or not (Fig. 11). This step-like movement is caused by the sudden acceleration of afterslip close to the shallower edge of the TA (Fig. 9a). Since we assimilated the daily GNSS time series, the temporal resolution was not shorter than a day, and thus, it is difficult to resolve if such slip acceleration truly occurred from the data used in this study. In addition, the daily coordinates represent an average position for 24 h, and consequently would not accurately reflect the exact position of the representative time, i.e., 12:00 PM in the case of GEONET F3 solutions, especially when significant deformation occurs during one day. Furthermore, Twardzik et al. (2019) reported that ~64% of the postseismic signal detected in the first 36 h following the mainshock occur within the first 12 h in the case of three large earthquakes along the South American subduction zone. Therefore, to clarify the temporal evolution of the slip in detail, especially for the initial stage of the afterslip, e.g., within one day following the mainshock, high-sampling GNSS data should be assimilated, which will be the topic of a future study. Another possible update related to the observation data is an inclusion of the vertical component of GNSS time series. The vertical component may be more sensitive to the location of a fault slip, and thus, will contribute to better constraint of the spatial distribution of the afterslip area.

We demonstrated that the adjoint data assimilation is capable of predicting weekly postseismic deformation by optimizing the frictional parameters and initial values of simulation variables. This study fixed the data period of the data assimilation for the first 15 days following the mainshock and predicted the following time series. In a practical application of predicting the postseismic deformation, we can sequentially update the predictions in real time similar to weather forecasts: every time new data are derived, we can assimilate the new data as well

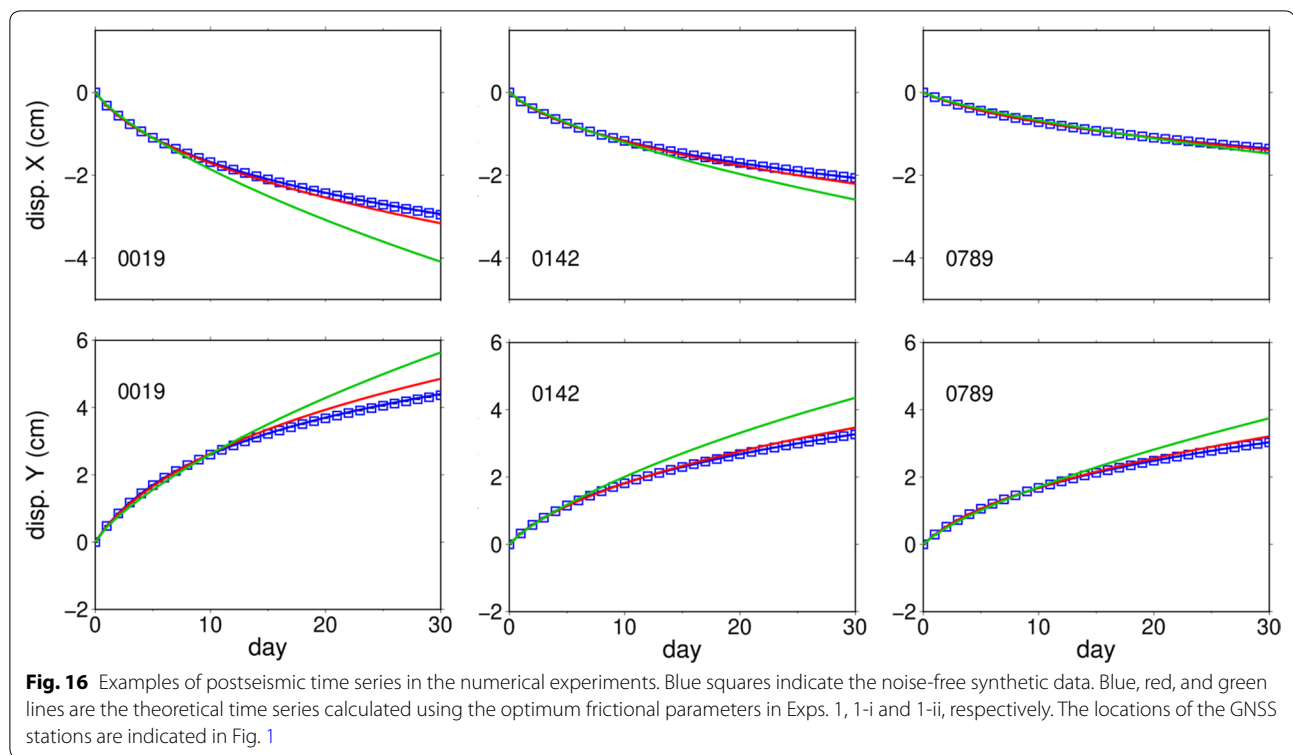
Table 3 True, first-guess and optimized values of frictional parameters in each subregion in the numerical experiments in Exps. 1-i, 1-ii, and 1, and values of the cost function calculated by using Eq. (7)

Region	True values	First-guess values		Optimized values		
		Exp. 1-i	Exp. 1-ii	Exp. 1-i	Exp. 1-ii	Exp. 1
1						
A–B (kPa)	8.00	16.0	80.0	19.7	109	8.13
A (kPa)	60.0	120	600	74.4	233	37.0
L (mm)	30.0	60.0	300	54.7	281	47.3
2						
A–B (kPa)	10.0	20.0	100	20.0	112	7.18
A (kPa)	40.0	80.0	400	51.6	279	37.3
L (mm)	10.0	20.0	100	14.9	60.3	9.73
3						
A–B (kPa)	6.00	12.0	60.0	15.9	68.8	7.03
A (kPa)	20.0	40.0	200	16.6	70.8	28.0
L (mm)	20.0	40.0	200	41.6	224	17.3
4						
A–B (kPa)	6.00	12.0	60.0	13.1	63.9	6.19
A (kPa)	20.0	40.0	200	27.9	108	20.5
L (mm)	20.0	40.0	200	38.5	219	20.3
5						
A–B (kPa)	8.00	16.0	80.0	2.96	142	7.61
A (kPa)	60.0	120	600	42.5	234	68.8
L (mm)	10.0	20.0	100	9.42	149	9.96
6						
A–B (kPa)	10.0	20.0	100	23.9	117	8.29
A (kPa)	40.0	80.0	400	25.9	119	63.2
L (mm)	10.0	20.0	100	18.0	157	10.5
7						
A–B (kPa)	10.0	20.0	100	21.9	106	11.4
A (kPa)	40.0	80.0	400	61.2	422	56.1
L (mm)	20.0	40.0	200	36.0	149	18.3
8						
A–B (kPa)	6.00	12.0	60.0	12.8	68.9	6.77
A (kPa)	20.0	40.0	200	28.0	71.0	29.5
L (mm)	30.0	60.0	300	60.6	422	21.4
9						
A–B (kPa)	8.00	16.0	80.0	10.8	— 117	11.5
A (kPa)	60.0	120	600	124	294	33.7
L (mm)	10.0	20.0	100	11.2	192	9.17
Cost function (assimilation period)	0	1.14×10^3	2.16×10^3	2.51×10^0	6.45×10^1	2.41×10^{-2}
Cost function (prediction period)	0	1.33×10^4	2.08×10^4	5.01×10^3	6.55×10^3	4.58×10^{-1}

as the previous observations and update the prediction of postseismic time series through a further update of frictional parameters, which have been optimized using the previous observations. Such sequential updates will improve the prediction skill.

Further improvement of the adjoint data assimilation method and relevant numerical models is possible to

attain more accurate predictions. In general, the optimization of the adjoint method depends on the first-guess values because the method always searches for the direction in the parameter space at which the cost function becomes smaller. This often results in a local minimum of the cost function. In addition, as discussed in the previous subsection, if the adjoint method is not



started based on suitable first-guess values, it results in a worse fitting to and prediction of the observation. This is partly because the adjoint method linearizes the non-linear forward simulation. To avoid such an effect, it is possible to develop a hybrid data assimilation approach by utilizing an ensemble-based approach, such as EnKF (Hirahara and Nishikiori 2019; van Dinter et al. 2019) or the simple grid calculation (Kano et al. 2015), which do not require the linearization of the forward equation, for example, to obtain better first-guess values then optimize them using the adjoint method. Another disadvantage of the adjoint method is that it is difficult to directly obtain the uncertainties of the optimized model parameters. The uncertainties are principally evaluated by calculating the second-order derivatives of the cost function around the optimized parameters, which is computationally expensive to be implemented. Ito et al. (2016) recently developed a second-order adjoint method that is applicable even to massive systems within reasonable computational time and resources. Following Ito et al. (2016), Ito et al. (2017) proposed a data assimilation method to predict the time series together with the uncertainties by considering the uncertainties of the optimized model parameters. Implementation of their methods to our adjoint system will lead to the evaluation of the uncertainties of the optimized frictional parameters and the prediction of postseismic deformation with the uncertainties. Finally,

this study focused on the short time-scale (\sim one month) evolution of postseismic GNSS time series that is mostly attributed to afterslip. When the year-long postseismic time series is considered, the introduction of the viscoelastic effect to numerical simulations will be essential for achieving more accurate predictions.

Conclusions

We developed an adjoint method for optimizing frictional parameters by directly assimilating the GNSS time series. The developed method was validated through numerical experiments using synthetic postseismic data, showing that the data could be well reproduced. Although the spatial distribution of afterslip was not completely recovered, the estimated slip velocities were consistent with the true ones. Using the optimized frictional parameters, the prediction skill of the following postseismic time series was significantly improved. The developed method was applied to the GNSS time series following the 2003 Tokachi-oki earthquake. The frictional parameters were optimized to $A \sim B \sim O(10 \text{ kPa})$, $A \sim O(100 \text{ kPa})$, and $L \sim O(10 \text{ mm})$. The optimized parameters reproduced the observed time series and predicted the following GNSS time series. The largest afterslip was inferred on the shallower side of the coseismic slip area. All these characteristics were inferred even if the simulation variables as well as the frictional parameters were

additionally optimized. Although future improvements are needed, the developed data assimilation method will provide a more quantitative evaluation for assessing risks of subsequent earthquakes and for monitoring the recovery process of megathrust earthquakes.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s40623-020-01293-0>.

Additional file 1: Figure S1. Postseismic time series at all stations in numerical experiments. The blue squares and red circles indicate the synthetic data in Exps. 1 and 2. The green, blue, and red lines are the theoretical time series calculated by using the first-guess and optimum frictional parameters in Exp. 1 and 2, respectively. **Figure S2.** Postseismic time series at all stations in the application to real data. The red and blue circles indicate the GNSS time series in X and Y directions. The pink, red, and orange lines are the theoretical time series in the X direction calculated by using the first-guess values, the optimum values when only frictional parameters are optimized and the optimum values when frictional parameters and initial values are simultaneously optimized, respectively. The light blue, blue, and green lines indicate the corresponding theoretical time series in the Y direction. **Figure S3.** (a) Spatio-temporal evolution of slip velocities when the diagonal elements of error variance-covariance matrix for slip velocities are set to be the squared of the larger standard deviation of 1.0×10^{-7} m/s compared to the results shown in Fig. 10b. (b) Difference between the spatial distribution of the first-guess initial slip velocity and that of the optimized initial slip velocity. **Figure S4.** Postseismic time series at all stations when the diagonal elements of error variance-covariance matrix for slip velocities are set to be the squared of the larger standard deviation of 1.0×10^{-7} m/s compared to the results shown in Figure S2. The red and blue circles indicate the GNSS time series in X and Y directions. The red, and blue lines are the corresponding theoretical time series in X and Y directions calculated by using the optimum values. **Figure S5.** Spatial distribution of the optimized frictional parameters when the diagonal elements of error variance-covariance matrix for slip velocities are set to be the squared of the larger standard deviation of 1.0×10^{-7} m/s compared to the results shown in Fig. 9b. **Figure S6.** Postseismic time series at all stations. The red and blue circles indicate the GNSS time series in X and Y directions. The red and blue lines are the corresponding theoretical time series calculated by using the optimum values obtained in Kano et al. (2015).

Abbreviations

EnKF: Ensemble Kalman filter; GEONET: GNSS Earth Observation Network System; GNSS: Global Navigation Satellite System; GSI: Geospatial Information Authority of Japan; KA: Kushiro-oki asperity; NIED: National Research Institute for Earth Science and Disaster Resilience; RSF: Rate and state-dependent friction law; TA: Tokachi-oki asperity.

Acknowledgements

We used the GNSS data provided by the GSI and the focal mechanisms of two aftershocks by F-net. We used the EIC computer system of the Earthquake Research Institute, the University of Tokyo. Generic Mapping Tools by Wessel and Smith (1998) were used to generate figures. The fruitful comments from Paul Segall, an anonymous reviewer, and the editor Takuya Nishimura helped a lot in improving the manuscript.

Authors' contributions

MK designed this study, developed the adjoint data assimilation method, applied it to both synthetic and real postseismic time series and prepared the manuscript. YI supported the development of the data assimilation method through technical advice and interpretation of the obtained results in terms of methodology. SM and KH contributed to the discussion of the obtained results. All authors read and approved the final manuscript.

Funding

This study was supported by the JSPS KAKENHI Grant Number JP15K17743 in Grant-in-Aid for Young Scientists (B), JP18K03779, JP18K03796 in Grant-in-Aid for Scientific Research (C), JP16H06473, JP16H06474, JP19H04620 in Scientific Research on Innovative Areas "Science of Slow Earthquakes", JST CREST Grant Number JPMJCR1763 and the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan, under its Earthquake and Volcano Hazards Observation and Research Program.

Availability of data and materials

The GEONET F3 solutions are provided by the GSI and available at https://datahouse1.gsi.go.jp/terras/terras_english.html. The focal mechanisms of two aftershocks are from the NIED F-net website (<https://www.fnet.bosai.go.jp/top.php?LANG=en>).

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ Graduate School of Science, Tohoku University, 6-3, Aramaki-aza-aoba, Aoba-ku, Sendai 980-8578, Japan. ² Faculty of Science, Kyoto University, Kitashirakawa-oiwake-cho, Sakyo-ku, Kyoto 606-8502, Japan. ³ Japan Agency for Marine-Earth Science and Technology, 3173-25, Showa-machi, Kanazawa-ku, Yokohama 236-0001, Japan. ⁴ Kagawa University, 1-1, Saiwai-cho, Takamatsu 760-0016, Japan. ⁵ Riken, AIP center, 2-2-2, Seika-cho, Souraku, Kyoto 619-0288, Japan.

Received: 29 June 2020 Accepted: 13 October 2020

Published online: 22 October 2020

References

- Ando M (1975) Source mechanisms and tectonic significance of historical earthquakes along the Nankai trough. *Japan Tectonophysics* 27(2):119–140. [https://doi.org/10.1016/0040-1951\(75\)90102-X](https://doi.org/10.1016/0040-1951(75)90102-X)
- Baba T, Hirata K, Hori T, Sakaguchi H (2006) Offshore geodetic data conducive to the estimation of the afterslip distribution following the 2003 Tokachi-oki earthquake. *Earth Planet Sci Lett* 241(1–2):281–292. <https://doi.org/10.1016/j.epsl.2005.10.019>
- Chlieh M et al (2007) Coseismic slip and afterslip of the great (M_w 9.15) Sumatra-Andaman earthquake of 2004. *Bull Seism Soc Am* 97:152–173. <https://doi.org/10.1785/0120050631>
- DeMets C, Gordon RG, Argus DF, Stein S (1990) Current plate motions. *Geophys J Int* 101(2):425–478. <https://doi.org/10.1111/j.1365-246X.1990.tb06579.x>
- Diao F, Xiong X, Wang R, Zheng Y, Walter TR, Weng H, Li J (2014) Overlapping post-seismic deformation processes: afterslip and viscoelastic relaxation following the 2011 Mw 9.0 Tohoku (Japan) earthquake. *Geophys J Int* 196:218–229. <https://doi.org/10.1093/gji/ggt376>
- Dieterich JH (1979) Modeling of rock friction 1. Experimental results and constitutive equations. *J Geophys Res* 84:2161–2168. <https://doi.org/10.1029/JB084iB05p02161>
- Fukuda J, Johnson KM, Larson KM, Miyazaki S (2009) Fault friction parameters inferred from the early stages of afterslip following the 2003 Tokachi-oki earthquake. *J Geophys Res* 114:B04412. <https://doi.org/10.1029/2008JB006166>
- Furumura T, Maeda T, Oba A (2019) Early forecast of long-period ground motions via data assimilation of observed ground motions and wave propagation simulations. *Geophys Res Lett* 46:138–147. <https://doi.org/10.1029/2018GL081163>
- Hatori T (1984) Source area of the east Hokkaido tsunami generated in April, 1843. *Bull Earthq Res Inst Univ Tokyo* 59:423–431 (in Japanese with English abstract)

- Hirahara K, Nishikiori K (2019) Estimation of frictional properties and slip evolution on a long-term slow slip event fault with the ensemble Kalman filter: numerical experiments. *Geophys J Int* 219:2074–2096. <https://doi.org/10.1093/gji/ggz2415>
- Hirata K, Geist E, Satake K, Tanioka Y, Yamaki S (2003) Slip distribution of the 1952 Tokachi-Oki earthquake (M 8.1) along the Kurile Trench deduced from tsunami waveform inversion. *J Geophys Res* 108:2196. <https://doi.org/10.1029/2002JB001976>
- Hori T, Kato N, Hirahara K, Baba T, Kaneda Y (2004) A numerical simulation of earthquake cycles along the Nankai Trough in southwest Japan: lateral variation in frictional property due to the slab geometry controls the nucleation position. *Earth Planet Sci Lett* 228(3–4):215–226. <https://doi.org/10.1016/j.epsl.2004.09.033>
- Hoshiya M, Aoki S (2015) Numerical shake prediction for earthquake early warning: data assimilation, real-time shake mapping, and simulation of wave propagation. *Bull Seismol Soc Am* 105(3):1324–1338. <https://doi.org/10.1785/0120140280>
- Hsu YJ, Simons M, Avouac JP, Galetzka J, Sieh K, Chlieh M, Natawidjaja D, Prawirodirdjo L, Bock Y (2006) Frictional afterslip following the 2005 Nias-Simulue earthquake. *Sumatra Science* 63:L10304. <https://doi.org/10.1126/science.1126960>
- 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. *Phys Rev E* 94(4):043307. <https://doi.org/10.1103/PhysRevE.94.043307>
- Ito S, Nagao H, Kasuya T, Inoue J (2017) Grain growth prediction based on data assimilation by implementing 4DVar on multi-phase-field model. *Sci Technol Adv Mater* 18(1):857–868
- Itoh Y, Nishimura T, Ariyoshi K, Matsumoto H (2019) Interplate slip following the 2003 Tokachi-oki earthquake from ocean bottom pressure gauge and land GNSS data. *J Geophys Res* 124:4205–4230. <https://doi.org/10.1029/2018JB016328>
- Kano M, Miyazaki S, Ito K, Hirahara K (2013) An adjoint data assimilation method for optimizing frictional parameters on the afterslip area. *Earth Planets Space* 65:1575–1580. <https://doi.org/10.5047/eps.2013.08.002>
- Kano M, Miyazaki S, Ishikawa Y, Hiyoshi Y, Ito K, Hirahara K (2015) Real data assimilation for optimization of frictional parameters and prediction of afterslip in the 2003 Tokachi-oki earthquake inferred from slip velocity by an adjoint method. *Geophys J Int* 203:646–663. <https://doi.org/10.1093/gji/ggv289>
- Kato N (2008) Numerical simulation of recurrence of asperity rupture in the Sanriku region, northeastern Japan. *J Geophys Res* 113:B06302. <https://doi.org/10.1029/2007JB005515>
- Klein E, Fleitout L, Vigny C, Garau JD (2016) Afterslip and viscoelastic relaxation model inferred from the large-scale postseismic deformation following the 2010 Mw 8.8 Maule earthquake (Chile). *Geophys J Int* 205(3):1455–1472. <https://doi.org/10.1093/gji/ggv086>
- Lewis J, Lakshminarayanan S, Dhall S (2006) Dynamic data assimilation: a least squares approach. Cambridge University Press, Cambridge
- Maeda T, Obara K, Shinohara M, Kanazawa T, Uehira K (2015) Successive estimation of a tsunami wavefield without earthquake source data: a data assimilation approach toward real-time tsunami forecasting. *Geophys Res Lett* 42:7923–7932. <https://doi.org/10.1002/2015GL065588>
- Miyazaki S, Larson KM (2008) Coseismic and early postseismic slip for the 2003 Tokachi-oki earthquake sequence inferred from GPS data. *Geophys Res Lett* 35:L04302. <https://doi.org/10.1029/2007GL032309>
- Miyazaki S, Segall P, Fukuda J, Kato T (2004) Space time distribution of afterslip following the 2003 Tokachi-oki earthquake: Implications for variations in fault zone frictional properties. *Geophys Res Lett* 31:L06623. <https://doi.org/10.1029/2003GL019410>
- Murakami M, Suito H, Ozawa S, Kaidzu M (2006) Earthquake triggering by migrating slow slip initiated by M8 earthquake along Kuril Trench. *Japan Geophys Res Lett* 33:L09306. <https://doi.org/10.1029/2006GL025967>
- Nakagawa H et al (2009) Development and validation of GEONET new analysis strategy (Version 4). *J Geogr Surv Inst* 118:1–8 (in Japanese)
- Nakata R, Hyodo M, Hori T (2012) Numerical simulation of afterslips and slow slip events that occurred in the same area in Hyuga-nada of southwest Japan. *Geophys J Int* 190(2):1213–1220. <https://doi.org/10.1111/j.1365-246X.2012.05552.x>
- National Research Institute for Earth Science and Disaster Resilience (2019). NIED F-net. doi:<https://doi.org/10.17598/nied.0005>.
- Oba A, Furumura T, Maeda T (2020) Data-assimilation-based early forecasting of long-period ground motions for large earthquakes along the Nankai Trough. *J Geophys Res*. <https://doi.org/10.1029/2019JB019047>
- Ohta Y et al (2012) Geodetic constraints on afterslip characteristics following the March 9, 2011, Sanriku-oki earthquake. *Japan Geophys Res Lett* 39:L16304. <https://doi.org/10.1029/2012GL052430>
- Ohtani M, Hirahara K, Hori T, Hyodo M (2014) Observed change in plate coupling close to the rupture initiation area before the occurrence of the 2011 Tohoku earthquake: implications from an earthquake cycle model. *Geophys Res Lett* 41(6):1899–1906. <https://doi.org/10.1002/2013GL058751>
- Okada Y (1992) Internal deformation due to shear and tensile faults in a half-space. *Bull Seismol Soc Am* 82(2):1018–1040
- Ozawa S, Kaidzu M, Murakami M, Imakiire T, Hatanaka Y (2004) Coseismic and postseismic crustal deformation after the Mw 8 Tokachi-oki earthquake in Japan. *Earth Planets Space* 56(7):675–680. <https://doi.org/10.1186/BF03352530>
- Ozawa S, Nishimura T, Suito H, Kobayashi T, Tobita M, Imakiire T (2011) Coseismic and postseismic slip of the 2011 magnitude-9 Tohoku-Oki earthquake. *Nature* 475:373–376. <https://doi.org/10.1038/nature10227>
- Pollitz F, Banerjee P, Burgmann R (2006) Post-seismic relaxation following the great 2004 Sumatra-Andaman earthquake on a compressible self-gravitating Earth. *Geophys J Int* 267:397–420. <https://doi.org/10.1111/j.1365-246X.2006.03018.x>
- Press WH, Teukolsky SA, Vetterling WT, Flannery BP (1996) Numerical Recipes in Fortran 77: The Art of Scientific Computing, 2nd edn. Cambridge University Press, New York, p 963
- Rice JR (1993) Spatio-temporal complexity of slip on a fault. *J Geophys Res* 98:9885–9907. <https://doi.org/10.1029/93JB00191>
- Ruina A (1983) Slip instability and state variable friction laws. *J Geophys Res* 88(B12):10359–10370. <https://doi.org/10.1029/JB088B12p10359>
- Sato K, Minagawa N, Hyodo M, Baba T, Hori T, Kaneda Y (2007) Effect of elastic inhomogeneity on the surface displacements in the northeastern Japan: Based on three-dimensional numerical modeling. *Earth Planets Space* 59(10):1083–1093. <https://doi.org/10.1186/BF03352051>
- Sato K, Baba T, Hori T, Hyodo M, Kaneda Y (2010) Afterslip distribution following the 2003 Tokachi-oki earthquake: an estimation based on the Green's functions for an inhomogeneous elastic space with subsurface structure. *Earth Planets Space* 62(12):923–932. <https://doi.org/10.5047/eps.2010.11.007>
- Sun T, Wang K (2015) Viscoelastic relaxation following subduction earthquakes and its effects on afterslip determination. *J Geophys Res* 120:1329–1344. <https://doi.org/10.1002/2014JB011707>
- Tanioka Y, Hirata K, Hino R, Kanazawa T (2004) Slip distribution of the 2003 Tokachi-oki earthquake estimated from tsunami waveform inversion. *Earth Planets Space* 56:373–376. <https://doi.org/10.1186/BF03353067>
- Twardzik C, Vergnolle M, Sladen A, Avallone A (2019). Unravelling the contribution of early postseismic deformation using sub-daily GNSS positioning. *Sci Rep* 9(1):1–12. <https://www.nature.com/articles/s41598-019-39038-z>
- Uchida N, Yui S, Miura S, Matsuzawa T, Hasegawa A, Motoya Y, Kasahara M (2009) Quasi-static slip on the plate boundary associated with the 2003 M8.0 Tokachi-oki and 2004 M7.1 off-Kuroshio earthquakes, Japan. *Gondwana Res* 16:527–533. <https://doi.org/10.1016/j.jgr.2009.04.002>
- van Dinther Y, Künsch HR, Fichtner A (2019) Ensemble data assimilation for earthquake sequences: probabilistic estimation and forecasting of fault stresses. *Geophys J Int* 217:1453–1478. <https://doi.org/10.1093/gji/ggz063>
- Wang Y, Satake K, Maeda T, Gusman AR (2017) Green's function-based tsunami data assimilation: a fast data assimilation approach toward tsunami early warning. *Geophys Res Lett* 44:10282–10289. <https://doi.org/10.1002/2017GL075307>
- Wessel P, Smith WHF (1998) New, improved version of Generic Mapping Tools released. *Eos Trans AGU* 79:579
- Yamanaka Y, Kikuchi M (2003) Source process of the recurrent Tokachi-oki earthquake on September 26, 2003, inferred from teleseismic body waves. *Earth Planets Space* 55:e21–e24. <https://doi.org/10.1186/BF03352479>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.