Unraveling Processes and Rheology of the Tohoku Earthquake Cycle Using Bayesian Inference

Marsman, C. P.; Vossepoel, F. C.; D'Acquisto, M.; van Dinther, Y.; van de Wiel, L.; Govers, R.

DOI
10.1029/2024JB029665

Publication date
2025

Document Version
Final published version

Published in
Journal of Geophysical Research: Solid Earth

Citation (APA)
Marsman, C. P., Vossepoel, F. C., D'Acquisto, M., van Dinther, Y., van de Wiel, L., & Govers, R. (2025). Unraveling Processes and Rheology of the Tohoku Earthquake Cycle Using Bayesian Inference. *Journal of Geophysical Research: Solid Earth*, *130*(5), Article e2024JB029665. https://doi.org/10.1029/2024JB029665

JGR Solid Earth

RESEARCH ARTICLE

10.1029/2024JB029665

Unraveling Processes and Rheology of the Tohoku Earthquake Cycle Using Bayesian Inference

C. P. Marsman[1], F. C. Vossepoel[2], M. D'Acquisto[3], Y. van Dinther[1], L. van de Wiel[1], and R. Govers[1]

[1]Department of Earth Sciences, Utrecht University, Utrecht, the Netherlands, [2]Department of Geoscience and Engineering, Delft University of Technology, Delft, the Netherlands, [3]Department of Geodesy, GFZ German Research Centre for Geosciences, Telegrafenberg, Germany

Key Points:

- Rheological properties estimated from geodetic data before and after the 2011 Tohoku-Oki earthquake mostly align with laboratory tests
- Power-law viscosities are needed to disentangle afterslip and viscoelastic relaxation from geodetic observations
- Release of elastic stresses accumulated before the earthquake contribute to observed near-trench postseismic motion

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

C. P. Marsman,
c.p.marsman@uu.nl

Citation:

Marsman, C. P., Vossepoel, F. C., D'Acquisto, M., van Dinther, Y., van de Wiel, L., & Govers, R. (2025). Unraveling processes and rheology of the Tohoku earthquake cycle using bayesian inference. *Journal of Geophysical Research: Solid Earth*, *130*, e2024JB029665. https://doi.org/10.1029/2024JB029665

Received 6 JUN 2024
Accepted 16 APR 2025

Author Contributions:

Conceptualization: C. P. Marsman, F. C. Vossepoel, Y. van Dinther, R. Govers
Data curation: C. P. Marsman
Formal analysis: C. P. Marsman
Funding acquisition: R. Govers
Investigation: C. P. Marsman
Methodology: C. P. Marsman
Project administration: R. Govers
Resources: C. P. Marsman, L. van de Wiel, R. Govers
Software: C. P. Marsman, L. van de Wiel, R. Govers
Supervision: F. C. Vossepoel, R. Govers
Validation: C. P. Marsman
Visualization: C. P. Marsman

Abstract Geodetic data spanning different phases of the earthquake cycle offer insights into the spatiotemporal interplay between processes driving surface deformation, such as viscoelastic relaxation, afterslip, and (re)locking. However, quantifying their contributions and explaining pre- and post-earthquake displacements with a single set of rheological parameters is challenging. We set up a 2-D earthquake cycle finite element model that simulates the mantle and a thin low-viscosity shear zone with a temperature-dependent linear Maxwell or nonlinear power-law rheology. We use the ensemble smoother with multiple data assimilation to estimate ensembles of parameters describing the rheological makeup of the subduction zone. We assimilate onshore and offshore displacement time series acquired before and after the 2011 Tohoku-Oki earthquake. Our models provide a unique, robust solution using a temperature-dependent power-law rheology. The estimated creep parameters for the mantle wedge deeper than ∼50 km and sub-slab mantle align with laboratory experiments. However, different creep parameters are required for the shallow part of the mantle wedge than the deeper part to explain the observed postseismic response—highlighting the need for shallow viscoelastic relaxation. The trade-off between water fugacity and activation energy hinders their individual estimation but yields a well-constrained viscosity structure. The spatial distribution of vertical displacements as well as the temporal signature of early postseismic horizontal displacements are required to estimate individual parameters for afterslip and viscoelastic relaxation. Afterslip occurs downdip of the coseismic rupture. Near-trench landward motion during the early postseismic period is driven by elastic stress release beneath the oceanic plate and sub-slab asthenospheric flow.

Plain Language Summary We aim to understand the deep processes occurring around Japan's convergent plate boundary before and after the magnitude 9.1 Tohoku-Oki earthquake in 2011. We employ a mechanical model based on realistic subduction zone structures and temperatures. Model stresses can relax due to flow. During the earthquake cycle, flow may occur due to rapid movement along a deep shear zone above the subducting plate, or it may occur throughout the bulk of the asthenosphere. The speed of the flow is governed by viscosity in our model. We estimate the viscosities of the shear zone and asthenosphere from the observed surface deformation. A novel aspect of our study is that we can individually resolve the viscosities of the shear zone and the bulk asthenosphere from the observations during the first 2 years after the earthquake. Additionally, we find that the viscosities must be stress-dependent. Rapid flow naturally occurs beneath the subducting plate, as is needed to reproduce observations of enhanced surface motions near the trench after the earthquake. Relaxation of stresses that were inherited from before the earthquake results in slower flow than reported in other studies.

1. Introduction

Over the last few decades, dense networks of geodetic instruments have been established around several subduction zones. The rich variety of accumulated data from great subduction earthquakes, such as the 1960 Valdivia earthquake, the 1964 Alaska earthquake, the 2004 Sumatra earthquake, and the 2011 Tohoku-Oki earthquake, provides unique time windows into parts of the megathrust earthquake cycle. These time windows provide insights into different phases of the earthquake cycle. Our knowledge of the interplay between mechanisms that drive surface displacements throughout the earthquake cycle, such as megathrust (re)locking, coseismic slip, afterslip and viscoelastic relaxation, has greatly improved with the advent of space-based geodesy (Avouac, 2015; Govers et al., 2018; Wang et al., 2012). In this paper, we aim to capture the processes driving surface deformation

Writing – original draft: C. P. Marsman, F. C. Vossepoel, M. D'Acquisto, Y. van Dinther, L. van de Wiel, R. Govers
Writing – review & editing: C. P. Marsman, F. C. Vossepoel, M. D'Acquisto, Y. van Dinther, L. van de Wiel, R. Govers

during the interseismic and postseismic phases, using a consistent formulation of strain accumulation and release throughout the megathrust earthquake cycle. We use data assimilation to estimate ensembles of rheological parameters that explain both interseismic and postseismic observations of the 2011 Tohoku-Oki earthquake.

Recent modeling efforts highlight the role of megathrust geometry in strain accumulation and release, demonstrating how variations in the dip angle and curvature of the subducting slab affect the slip distribution (e.g., Moreno et al., 2009), afterslip and stress relaxation of the mantle (reviewed in Govers et al., 2018). Advances in computational power have allowed researchers to incorporate more realistic three-dimensional slab shapes into their models (e.g., Pollitz et al., 2008; Suito & Freymueller, 2009; Wang et al., 2001), which continue to improve with detailed mapping of the megathrust geometry (e.g., Hayes et al., 2018). Details in structural features such as a cold nose (Luo & Wang, 2021) and the volcanic arc (Hu et al., 2014; Itoh et al., 2019) significantly influence mantle fluid flow and the geodetic imprint. A better representation of mechanical properties and the geometry of the megathrust and margin is essential for accurately determining the regional rheological makeup and the spatiotemporal extent of subduction zone processes.

In numerical models the slab and overriding plate are frequently considered elastic throughout the earthquake cycle, and the mantle wedge (MW) and sub-slab mantle are considered viscoelastic. Several rheological models have been proposed for the mantle to describe different phases of the earthquake cycle. Typically, the rheological models used are a linear Maxwell, linear bi-viscous Burgers, or nonlinear power-law rheology. A linear Maxwell model is generally applied to conditions of steady-state deformation, whereas a Burgers model or a power-law model is used to describe both transient and steady-state behavior. A Burgers model is a convenient mathematical approach to approximate time-dependent mantle rheology, which consists of coupled Kelvin and Maxwell elements that represent a transient and steady-state rheology. In a power-law rheology, viscosity depends on stress (Melosh & Raefsky, 1980), in accordance with temperature-dependent laboratory-derived constitutive relations (e.g., Kirby & Kronenberg, 1987).

Recently, a nonlinear Burgers rheology for the mantle, which combines elements of steady-state power-law creep and transient power-law creep, has been inferred to explain postseismic deformation (e.g., Agata et al., 2019; Masuti et al., 2016; Muto et al., 2019). Synthetic tests on a strike-slip fault have shown that a linear Burgers and a power-law rheology may be indistinguishable from each other with geodetic data (Mallick et al., 2022). Following the 2004 Sumatra earthquake, onshore trenchward postseismic displacements have been explained using a Burgers or Maxwell rheology without afterslip (Pollitz et al., 2008; Tanaka et al., 2009), whereas Hoechner et al. (2011) found that either a Burgers rheology without afterslip or a Maxwell rheology with afterslip can explain the observed displacements. Hence, the choice of the mantle rheology affects the relative contributions of afterslip and viscoelastic relaxation to observed surface deformation.

Apparent discrepancies emerge when determining the relative contributions of afterslip and viscoelastic relaxation to postseismic crustal deformation and whether they can be uniquely constrained (e.g., Hu & Wang, 2012). Several postseismic studies published soon after the 2011 Tohoku-Oki earthquake have considered afterslip as the common driving mechanism for the observed transient deformation within the first few months (Ozawa et al., 2012; Perfettini & Avouac, 2014; Silverii et al., 2014). However, afterslip alone could not explain the near-trench landward motion recorded by several seafloor stations that came online later, which strongly suggested that viscoelastic relaxation is also required in the early postseismic phase (Sun et al., 2014; Watanabe et al., 2014). Subsequent studies incorporating both afterslip and viscoelastic relaxation into their models have concluded that these processes can be distinguished either spatially (Freed et al., 2017) or temporally (Fukuda & Johnson, 2021). As noted by Fukuda and Johnson (2021), models including stress-driven afterslip typically find patches of afterslip downdip of the coseismic rupture (Agata et al., 2019; Dhar et al., 2022; Fukuda & Johnson, 2021; Hu et al., 2016), whereas studies incorporating kinematic inversions of afterslip identify northern and southern patches of afterslip (Freed et al., 2017; Iinuma et al., 2016; Yamagiwa et al., 2015). These results indicate that the modeling approach to afterslip greatly influences the reconstruction of afterslip.

Afterslip can be estimated using both kinematic approaches and physics-based approaches that describe the mechanics governing slip. Solutions of kinematic inversions can result in a very good match with the observations. These solutions may yield unrealistic results because they may be inconsistent with stresses and material properties. More realistic results may be obtained by computing afterslip using a physics-based approach. Stress-driven afterslip of the 2011 Tohoku-Oki earthquake has been determined using frictional slip (e.g., Dhar et al., 2022; Fukuda & Johnson, 2021; Muto et al., 2019) or viscous deformation (e.g., Hu et al., 2014; Hu et al., 2016). It has been suggested that the evolution of afterslip follows a power-law relaxation law, of which linear creep and frictional slip are special cases (Montési, 2004). In the case of frictional slip, power-law relations with high power exponents can be used to describe frictional deformation (Barbot, 2023; Montési, 2004). For viscous deformation based on a power-law, the interplay between diffusion and dislocation creep becomes important (Bürgmann & Dresen, 2008). The specific physical conditions of a shear zone (SZ) at the subduction interface govern the relative importance of diffusion and dislocation creep and whether a particular type of rheology may appropriately describe deformation.

Geodetic data show enhanced landward offshore motion above the rupture zone (Watanabe et al., 2014) and enhanced landward onshore motion away from the rupture zone (Mitsui & Heki, 2013; Yuzariyadi & Heki, 2021) after the 2011 Tohoku-Oki earthquake. A low-viscosity sub-slab layer was introduced by Sun et al. (2014), and was shown by Suito (2017), to enhance landward near-trench motion above the rupture zone. Enhanced landward onshore observations away from the rupture zone have also been attributed to viscoelastic relaxation of the oceanic mantle (Sun et al., 2024). Postseismic power-law models have identified a temporary weak zone beneath the oceanic plate (Agata et al., 2019; Dhar et al., 2022; Muto et al., 2019). Observational support for the weak sub-slab layer comes from a reduction in seismic velocities beneath oceanic plates at approximately 80 km depth (Kawakatsu et al., 2009). However, the viscosities invoked in previous postseismic models are extremely low, sometimes lower than the viscosity in the hydrous mantle wedge. There is no consensus on the origin of this weak anomaly and the time span of postseismic observations is too short for conclusions about its persistence. Proposed explanations include a high partial melt fraction (Kawakatsu et al., 2009; Schmerr, 2012) or lithospheric weakening of the slab due to bending (Freed et al., 2017). The question remains whether landward offshore motion can be explained without the need of a persistent and very low-viscosity sub-slab layer.

Earthquake cycle models (Govers et al., 2018; Hirahara, 2002; Wang, 1995) provide the means to find a consistent framework describing the rheology over the entire earthquake cycle. Previous work has applied earthquake cycle models to explain observations at different phases of the earthquake cycle (e.g., Barbot, 2020; Hashima & Sato, 2017; Julve et al., 2024; Li et al., 2020; Sobolev & Muldashev, 2017). The total stress state in a subduction system continuously evolves throughout the earthquake cycle. Plate convergence at a locked megathrust fault results in interseismic stress accumulation. This is (partly) released during the coseismic phase as the megathrust unlocks. Additional stress on the megathrust interface is released in the form of afterslip. Coseismic slip and afterslip induced stresses are relaxed in the asthenosphere and in some cases in the lower crust (e.g., Cheng et al., 2022; Peña et al., 2020; van Dinther et al., 2019) postseismically. Processes on geological timescales, such as mantle convection, also contribute to the total stress state in a subduction system. "Single event" modeling approaches often fall short in describing the complex interplay of processes and mechanisms in earthquake cycles, particularly in terms of pre-stress conditions and their influence on rheology (especially in the power-law case) and postseismic behavior, which has not yet been investigated. Current earthquake cycle models can benefit from the growing understanding of megathrust geometry (e.g., Hayes et al., 2018), (thermo-)mechanical properties (e.g., van Keken & Wilson, 2023), and more accurate locking and coseismic slip distributions (e.g., Hashima et al., 2016; Itoh et al., 2021; Moreno et al., 2009). This allows for better interpretation of the processes and rheology that have a footprint in the observed surface deformation throughout the earthquake cycle.

In this paper we set up a two-dimensional (2-D) finite element earthquake cycle model of the Japan subduction zone with either a linear or a power-law rheology of the mantle, shear zone, and lower crust. A priori information, such as a realistic temperature field and a coseismic slip distribution, is integrated into the model. The power-law dependence on stress necessitates a consistent approach that includes the approximation of pre-stresses that cannot be given by models solely driven by coseismic slip. We follow a similar approach to Govers et al. (2018) in the sense that pre-stresses are initialized during repeated earthquake cycles. Interseismic and postseismic observations are assimilated using the ensemble smoother with multiple data assimilation (ESMDA). We aim to estimate a single set of parameters that describe the most appropriate rheological model for the mantle and shear zone downdip of the megathrust. Using ESMDA, we quantify probabilistic uncertainties and identify parameter trade-offs to investigate potential interactions between parameters representing afterslip and viscoelastic relaxation. Interestingly, our results indicate that significantly less viscoelastic relaxation from the sub-slab asthenosphere is required to match landward motion of the near-trench region during the early postseismic period. We propose another hypothesis to explain these observations, in which connectivity and cyclic loading of deeper parts of the elastic slab and oceanic plate near the trench drive this landward motion.

2. Analysis of Geodetic Observations

2.1. Parameterization of Interseismic and Postseismic Data

We use daily site coordinate solutions (Nakagawa, 2009) as part of the Global Navigation Satellite Systems (GNSS) Earth Observation Network System, abbreviated as GEONET. These solutions are computed in ITRF2005 (Altamimi et al., 2011) and are operated by the Geospatial Information Authority (GSI) of Japan. We collect interseismic coordinates observed between 1 January 1997 and 1 May 2000, that is, to avoid contamination by seismic swarms in mid-2000 (Japan Meteorological Agency, 2000). We fit the following parametric model for each interseismic time series solution at each GNSS site using linear least-squares:

$$x_{\mathrm{pre}}(t) = x_0 + v_0 t + a_1 \sin(2\pi t) + a_2 \cos(2\pi t) + b_1 \sin(4\pi t) + b_2 \cos(4\pi t) + \sum_{k=1}^{K} c_k \theta(t - t_k), \tag{1}$$

where x_{pre} is the solution of the interseismic time series vector composed of east, north, and vertical components; t represents the time since the start of the time series; x_0 is the displacement at the start of the interseismic time series; v_0 is the steady deformation rate; a_1, a_2, b_1, b_2 are annual and semi-annual seasonal coefficients; c is the change in deformation due to a sudden event; θ is the Heaviside function; and t_k is the time of a sudden event. We define the smoothed interseismic displacement vector as the steady deformation and subtract the deformation at the start of the time series.

The M_{w} 9.0 Tohoku-Oki earthquake occurred on March 11. We use the GEONET F3 solutions (i.e., the processed coordinates provided by the GSI of Japan) from 9 April 2011 to 9 April 2021. We do not include data from the first month after the earthquake as more stations became online again later and aim to reduce the effects from poroelastic rebound. We then fit the F3 solution of the postseismic time series, defined between 9 April 2011 and 9 April 2021, to the following equation which is modified from the fitting function of Tobita (2016) to include steps and seasonal effects:

$$\begin{aligned} x_{\mathrm{post}}(t') = {} & x'_0 + v_0 t' + p_1 \ln\left(1 + \frac{t'}{\tau_1}\right) + p_2 \ln\left(1 + \frac{t'}{\tau_2}\right) - p_3\left(1 - \exp\left(-\frac{t'}{\tau_3}\right)\right) \\ & + a'_1 \sin(2\pi t') + a'_2 \cos(2\pi t') + b'_1 \sin(4\pi t') + b'_2 \cos(4\pi t') + \sum_{k=1}^{K} c'_k \theta(t' - t'_k), \end{aligned} \tag{2}$$

where x_{post} is the F3 solution of the postseismic time series vector composed of east, north, and vertical components; t' represents the time since the earthquake, x'_0 is the deformation at the start of the postseismic time series; p_1, p_2, and p_3 are amplitudes; and τ_1, τ_2, and τ_3 are characteristic decay times. The parameters x'_0, p_1, p_2, p_3, a'_1, a'_2, b'_1, b'_2 and c'_k are determined via linear least-squares regression. For the decay times we use $\tau_1 = 2.12$ days, $\tau_2 = 287.5$ days and $\tau_3 = 450{,}005$ days for both the horizontal and vertical components (Fujiwara et al., 2022). The smoothed postseismic displacements are the sum of the linear trend, two logarithmic functions, and one exponential term, which represent contributions from locking, afterslip and viscoelastic relaxation. Subsequently, all displacements are referenced to the FUKUE GNSS station (station ID 950462, 128.8°E, 32.7°N), which is approximately 1,500 km away from the Tohoku-Oki earthquake epicenter, so that we can compare the data with our model results with respect to a fixed upper plate. Finally, the amplitude of displacement on 9 April 2011 is subtracted from later amplitudes to derive the smoothed 10-year postseismic time series of displacement since 9 April 2011.

Along with GEONET data, we compile seafloor time series from GNSS-Acoustic (GNSS-A) observations from campaign survey stations operated by the Japan Coast Guard (JCG) (Yokota et al., 2018) and Tohoku University (THK) (Honsho et al., 2019). For the THK data, only the horizontal component is available, and the data encompasses mostly September 2012 to May 2016. We estimate the postseismic deformation at the seafloor stations in a similar fashion to the GEONET data. However, we do not separately estimate the interseismic trend as little to no data is available before the earthquake. Instead, we estimate interseismic velocities from the linear trend in the postseismic time series using linear least-squares regression. In our estimation, we make a few assumptions as the data is temporally sparse. We do not consider seasonal variations and steps, that is, we set parameters p_3, a'_1, a'_2, b'_1, b'_2 and c'_k to zero in Equation 2. We assume the same decay times for the logarithmic terms (τ_1 and τ_2) as the

GEONET data. We disregard the exponential term as Fujiwara et al. (2022) found a large decay time (i.e., >1200 years) and assume that this term is captured by the linear trend. As the data time series of the THK data spans mostly between September 2012 and May 2016, we can ignore short-term afterslip, and thus also set p_1 in Equation 2 to zero for the THK data. The postseismic time series of the JCG data is taken as the sum of the linear component and two logarithm terms. Then, we subtract the displacement on 9 April 2011 from the postseismic time series to derive the smoothed postseismic time series of the JCG data between 9 April 2011 and 9 April 2021. The postseismic time series of the THK data is the sum of the linear trend and one logarithmic term. We subtract the displacement on 1 September 2012 to derive the smoothed postseismic time series of the THK data between 1 September 2012 and 30 April 2016.

2.2. Data Results for Interseismic and Postseismic Periods

To compare our two-dimensional model results to the data, we only select data within 50 km north and south of the transect (Figures 1a and 1b). Following Marsman et al. (2023), this transect was selected as it coincides with the center of the interseismic locked region inferred by Loveless and Meade (2011), as well as the region of most coseismic slip (Loveless & Meade, 2011) and postseismic motion (Nishimura, 2014; Ozawa et al., 2012). We define the division between the forearc and backarc at 300 km from the trench, that represents a simplified volcanic arc in 2-D (pink dashed lines, Figure 1a). We show the postseismic time series of select stations located near the trench, in the forearc, in the backarc and around the volcanic arc (Figures 1e–1h). The results of the parameterized time series for all stations along the swath profile can be found in Figure S7 in Supporting Information S1. For the assimilation, we compile a data vector **d** consisting of sampled observations of the smoothed interseismic and postseismic time series. We sample the 3.4-year interseismic time series every 4 months and sample the 10-year postseismic times series every 60 days during the first 2 years and every 120 days between 2 and 10 years after the earthquake.

The resultant GEONET interseismic velocities exhibit landward motion, with horizontal velocities of ~50 mm/yr on the eastern coast and ~20 mm/yr on the western coast within the transect (Figures 1a and 1c). Interseismic GEONET vertical velocities show subsidence in the forearc and uplift in the backarc within the transect (Figures 1a and 1c). These vertical GEONET velocities agree with rates derived from leveling data from the past 100 years (Sasajima et al., 2019). Although a general pattern can be found in the interseismic vertical velocities, they have a much lower signal-to-noise ratio than the interseismic horizontal data (Figures 1e–1g). Additionally, sources other than tectonics may be present in the data, such as groundwater extraction causing land subsidence (e.g., station 0556 at 319 km from the trench, Figures 1c, Sasajima et al. (2019)). The postseismic seafloor data mostly show landward motion (Figures 1b and 1d) at rates faster than the convergence rate (Watanabe et al., 2014), which indicates viscoelastic relaxation of the sub-slab asthenosphere (e.g., Govers et al., 2018). However, seafloor station MYGW and the GEONET data show trenchward motion, which is an imprint from viscoelastic relaxation of the mantle wedge and afterslip. The postseismic verticals show a more complex behavior with subsidence near the trench, coastal uplift on the eastern side and again subsidence in the backarc (Figures 1b and 1d).

3. Methods

3.1. Data Assimilation

3.1.1. Inverse Problem Formulation

As mentioned earlier, we intend to capture the processes driving surface displacements during the interseismic and postseismic phases. To do so, we aim to constrain a single set of model parameters (see Section 3.2) that control both interseismic and postseismic deformation, based on afterslip and bulk viscoelastic relaxation. We use data assimilation to estimate these parameters based on surface observations from before and after the 2011 Tohoku-Oki earthquake.

We adopt the ESMDA introduced by Emerick and Reynolds (2013) and use the ESMDA formulation by Evensen (2019) in our estimation problem. We denote the vector containing the unknown model parameters as $\boldsymbol{\Theta}$. The nonlinear forward model operator $\mathbf{m}$ translates the parameters into model outputs and the measurement operator $\mathbf{h}$ maps model outputs into the observation space. The vector function $\mathbf{g}$ represents both the application of

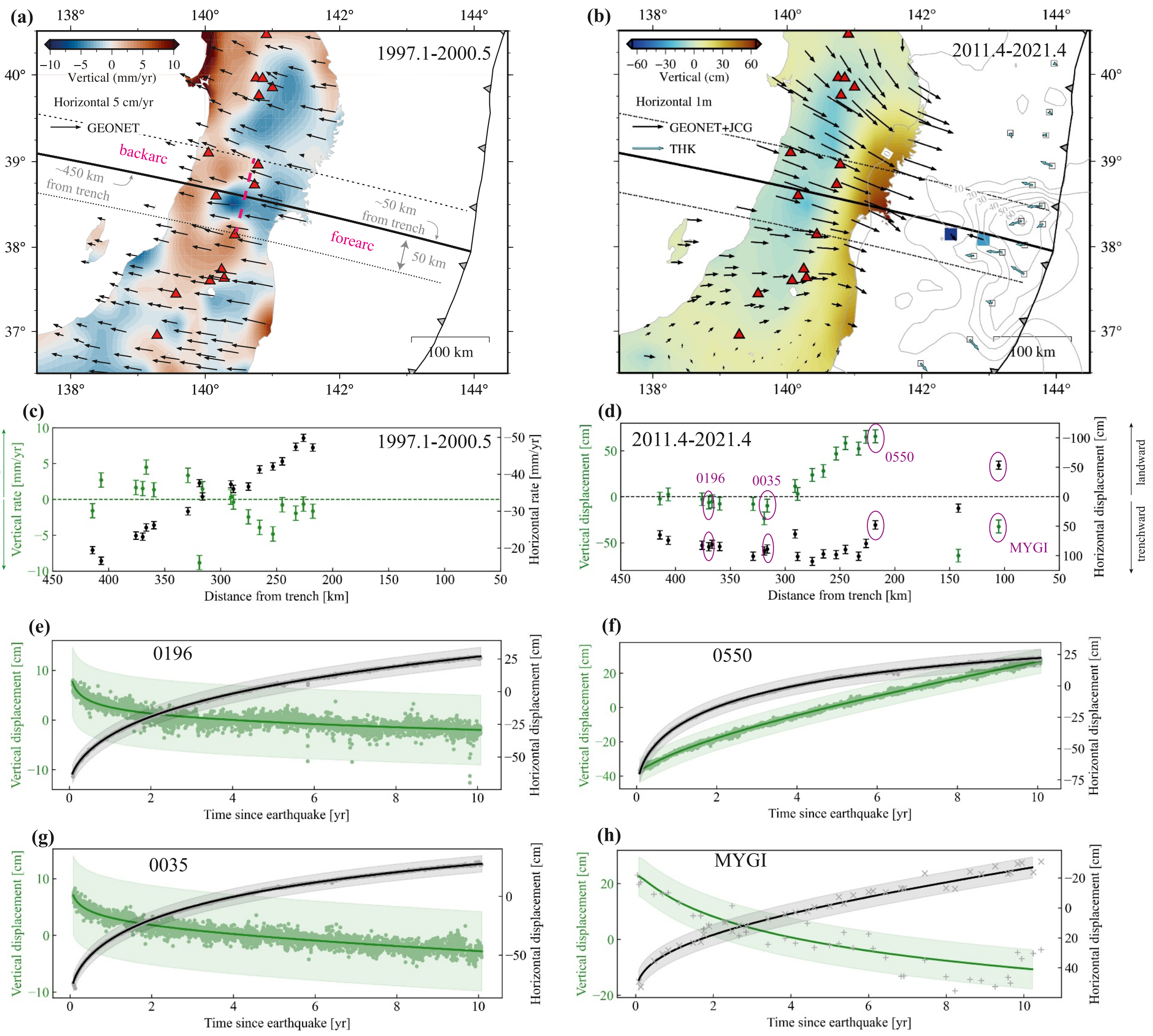


Figure 1. (a) Interseismic velocities. Red triangles indicate active volcanoes (Japan Meteorological Agency, 2013). The pink dashed line represents the simplified division between the forearc and backarc at 300 km from the trench. The thick black line shows the location of the middle of the swath profile, taken parallel to the relative plate motion direction, dashed lines indicate the width of the swath. (b) Cumulative 10-year postseismic displacements. Gray lines indicate the coseismic slip distribution from Hooper et al. (2013). (c) Interseismic velocities along the swath profile with a half-width of 50 km. Black and green colors represent the horizontal and vertical components. Error bars represent the total observation error of 1 cm (Section 3.3). (d) Cumulative 10-year postseismic displacements along the swath profile with a half-width of 50 km. Error bars represent the total observation error of 7 cm (Section 3.3). (e)–(h) Time series of GNSS displacements (offsets and seasonal components removed, indicated with dotted markers) and the smoothed times series (solid lines) at selected stations: 0196, 0550, 0035, and MYGI, respectively. The shaded bands represent the observation error of 7 cm. Sign convention: trenchward and upward motion are positive, whereas landward and downward motion are negative.

the forward model and measurement operator to generate model equivalent measurements, $\mathbf{g}(\cdot) = \mathbf{h}(\mathbf{m}(\cdot))$. We can relate the observed data $\mathbf{d}$ to the true model equivalent of the observations $\mathbf{y}^{\text{true}} = \mathbf{g}^{\text{true}}(\boldsymbol{\Theta}^{\text{true}})$ through

$$\mathbf{d} = \mathbf{g}^{\text{true}}(\boldsymbol{\Theta}^{\text{true}}) + \mathbf{e}, \tag{3}$$

in which $\mathbf{g}^{\text{true}}$ encompasses the perfect model (i.e., there are no model errors) and $\mathbf{e}$ represents the measurement errors. Measurement errors are assumed to be normally distributed, that is, $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{C}_{\text{dd}})$, where the error-covariance matrix $\boldsymbol{C}_{\text{dd}}$ describes the variance of the measurement errors. As with any other model, our model is imperfect (i.e., $\mathbf{g} \neq \mathbf{g}^{\text{true}}$) and thus model errors $\mathbf{q}$ arise, which are inherently unknown. The model equivalent of the observations $\mathbf{y}$ can be written as a function of the parameters $\boldsymbol{\Theta}$ and the model errors $\mathbf{q}$, such that

$$\mathbf{y} = \mathbf{g}(\boldsymbol{\Theta}, \mathbf{q}). \tag{4}$$

However, in our study we assume that model errors are ascribed to the observation error (e.g., Koohkan & Bocquet, 2012; Marsman et al., 2023). The observation error consists of representation errors, that is, discrepancies between model equivalent observations and actual observations, and instrumentation errors (e.g., Marsman et al., 2023). Thus, the model equivalent of observations are only a function of the parameters:

$$\mathbf{y} = \mathbf{g}(\boldsymbol{\Theta}). \tag{5}$$

The state vector contains the unknown quantities to be estimated. In our study, the state vector contains the unknown parameters and is defined as $\mathbf{z} = [\boldsymbol{\Theta}]^{\intercal}$. Through Bayes' theorem, we define the posterior probability density function (pdf) of the state vector $\mathbf{z}$ given the observed data $\mathbf{d}$ as

$$p(\mathbf{z}|\mathbf{d}) \propto p(\mathbf{z})\, p(\mathbf{d}|\mathbf{z}), \tag{6}$$

where $p(\mathbf{z})$ is the prior pdf of the parameters and model errors. We assume Gaussian distributions for the state vector and likelihood pdfs, that is, $p(\mathbf{z}) \sim \mathcal{N}\left(\mathbf{z}^f, \boldsymbol{C}_{\mathrm{zz}}\right)$, and $p(\mathbf{d}|\mathbf{z}) = p(\mathbf{e}) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{C}_{\mathrm{dd}})$ where $\mathbf{z}^f$ is the prior mean and $\boldsymbol{C}_{\mathrm{zz}}$ is the covariance matrix of the state vector elements. The posterior can be expanded as follows:

$$p(\boldsymbol{\Theta}|\mathbf{d}) = \exp\left\{-\frac{1}{2}\left(\boldsymbol{\Theta} - \boldsymbol{\Theta}^f\right)^{\intercal} \mathbf{C}_{\Theta\Theta}^{-1}\left(\boldsymbol{\Theta} - \boldsymbol{\Theta}^f\right)\right\} \times \exp\left\{-\frac{1}{2}(\mathbf{g}(\boldsymbol{\Theta}) - \mathbf{d})^{\intercal} \mathbf{C}_{\mathrm{dd}}^{-1}(\mathbf{g}(\boldsymbol{\Theta}) - \mathbf{d})\right\}, \tag{7}$$

where the superscript f stands for the prior mean and $\mathbf{C}_{\Theta\Theta}$ is the covariance matrix of the parameters. To find the set of parameters that best explain the observations, the posterior $p(\boldsymbol{\Theta}|\mathbf{d})$ needs to be maximized, which is equivalent to minimizing the cost function

$$\mathcal{J}(\boldsymbol{\Theta}) = \left(\boldsymbol{\Theta} - \boldsymbol{\Theta}^f\right)^{\intercal} \mathbf{C}_{\Theta\Theta}^{-1}\left(\boldsymbol{\Theta} - \boldsymbol{\Theta}^f\right) + (\mathbf{g}(\boldsymbol{\Theta}) - \mathbf{d})^{\intercal} \mathbf{C}_{dd}^{-1}(\mathbf{g}(\boldsymbol{\Theta}) - \mathbf{d}). \tag{8}$$

There are various methods available to minimize the above cost function to solve or approximate the posterior distribution $p(\boldsymbol{\Theta}|\mathbf{d})$, such as the ensemble smoother, ESMDA, or an iterative ensemble smoother. For an overview of these aforementioned methods we refer the reader to a book (Evensen et al., 2022) or the comparative study by Evensen (2018). A smoother is particularly well-suited for estimating model parameters in an earthquake cycle model, as smoothers estimate the state of a system from past, current and future observations, yielding a more accurate representation of earthquake cycle dynamics. We use ESMDA to approximately sample the posterior distribution $p(\boldsymbol{\Theta}|\mathbf{d})$, as explained in the next section. The method is computationally efficient, is easy to implement and can deal with nonlinear models to some extent.

3.1.2. The Ensemble Smoother With Multiple Data Assimilation

In ensemble methods, the prior, likelihood and posterior pdfs are approximated by an ensemble of parameters. By minimizing the solutions of an ensemble of cost functions, ensemble methods can provide approximate samples of the posterior pdf (Evensen, 2018). The main idea of ESMDA is to assimilate the same data set multiple times with an inflated covariance matrix of measurement errors to obtain an approximation of the posterior pdf. First, the likelihood function is adjusted as follows:

$$p(\mathbf{z}|\mathbf{d}) \propto p(\mathbf{z})\, p(\mathbf{d}|\mathbf{z}) = p(\mathbf{z})\, p(\mathbf{d}|\mathbf{z})^{\sum_{i=1}^{N_{\mathrm{mda}}} \frac{1}{\alpha_i}} \quad \text{with} \sum_{i=1}^{N_{\mathrm{mda}}} \frac{1}{\alpha_i} = 1, \tag{9}$$

where N_{mda} is the number of data assimilation steps and α_i is the inflation factor at data assimilation step i. Because multiple steps are introduced, the covariance matrix needs to be inflated to account for the fact that in each subsequent data assimilation step, the prior contains information from the observations. As a result, perturbed measurements are introduced $\mathbf{d}_{j,i} = \mathbf{d} + \mathbf{e}_{j,i}$, with sampled measurement perturbations $\mathbf{e}_{j,i} \sim \mathcal{N}\left(\mathbf{0}, \alpha_{i+1}\mathbf{C}_{dd}\right)$ for each ensemble member denoted with subscript j. The cost function $\mathcal{J}$ for each new step becomes

$$\begin{aligned}\mathcal{J}(\boldsymbol{\Theta}_{j,i+1}) = {} & (\boldsymbol{\Theta}_{j,i+1} - \boldsymbol{\Theta}_{j,i})^{\intercal}\mathbf{C}_{\Theta\Theta}^{-1}(\boldsymbol{\Theta}_{j,i+1} - \boldsymbol{\Theta}_{j,i}) \\ & + \left(\mathbf{g}(\boldsymbol{\Theta}_{j,i+1}) - \mathbf{d} - \sqrt{\alpha_{i+1}}\,\mathbf{e}_{j,i}\right)^{\intercal}(\alpha_{i+1}\mathbf{C}_{dd})^{-1}\left(\mathbf{g}(\boldsymbol{\Theta}_{j,i+1}) - \mathbf{d} - \sqrt{\alpha_{i+1}}\,\mathbf{e}_{j,i}\right).\end{aligned} \tag{10}$$

The procedure for ESMDA is summarized as follows. ESMDA starts with an initialization ($i = 0$) of the unknown parameters: $\boldsymbol{\Theta}_{j,i=0} \sim \mathcal{N}\left(\boldsymbol{\Theta}^{f}, \boldsymbol{C}_{\Theta\Theta}\right)$. The number of data assimilation steps needs to be established beforehand to initialize the set of inflation coefficients (Emerick & Reynolds, 2013). The inflation coefficients are defined according to the scheme by Evensen (2018):

$$\alpha_i = \alpha_i' \left(\sum_{i=1}^{N_{\mathrm{mda}}} \frac{1}{\alpha_i'}\right) \tag{11}$$

and

$$\alpha_{i+1}' = \frac{\alpha_i'}{\alpha_{\mathrm{geo}}}, \tag{12}$$

where we set $\alpha_{\mathrm{geo}} = \alpha_{i=0}' = 1.0$ such that $\alpha_{i+1}' = 1$. After initialization, a number of steps follow that each consist of a forecast step and an analysis step (see e.g., Marsman et al. (2023)).

In the forecast step, the forward model is run for each ensemble member, the model outputs are taken at the same space-time locations as the data, to obtain the prior ensemble of model equivalent observations

$$\mathbf{y}_{j,i} = \mathbf{g}(\boldsymbol{\Theta}_{j,i}). \tag{13}$$

In the update step, the model parameters are updated according to

$$\boldsymbol{\Theta}_{j,i+1} = \boldsymbol{\Theta}_{j,i} + \mathbf{C}_{\Theta \mathrm{y}}^{i}\left(\mathbf{C}_{\mathrm{yy}}^{i} + \alpha_{i+1}\boldsymbol{C}_{dd}\right)^{-1}\left(\mathbf{d} + \mathbf{e}_{j,i} - \mathbf{y}_{j,i}\right) \tag{14}$$

The cross-covariances matrices between the model parameters and the data, $\mathbf{C}_{\Theta \mathrm{y}}^{i}$, and the auto-covariance matrix of the model equivalent observations, $\mathbf{C}_{\mathrm{yy}}^{i}$, are computed at each data assimilation step as

$$\begin{aligned}\mathbf{C}_{\Theta \mathrm{y}}^{i} &= \frac{1}{N_e - 1}\sum_{j=1}^{N_e}\left(\boldsymbol{\Theta}_{j,i} - \overline{\boldsymbol{\Theta}}_i\right)\left(\mathbf{y}_{j,i} - \overline{\mathbf{y}}_i\right)^{\intercal} \\ \mathbf{C}_{\mathrm{yy}}^{i} &= \frac{1}{N_e - 1}\sum_{j=1}^{N_e}\left(\mathbf{y}_{j,i} - \overline{\mathbf{y}}_i\right)\left(\mathbf{y}_{j,i} - \overline{\mathbf{y}}_i\right)^{\intercal},\end{aligned} \tag{15}$$

where the overbars denote the ensemble mean. Then, the next data assimilation step takes place, and the forecast and update steps are repeated. The procedure is repeated $N_{\mathrm{mda}} - 1$ times until the final ensembles approximating the posterior pdfs of $\boldsymbol{\Theta}$ are obtained.

3.2. Forward Model Setup

We use a two-dimensional (2-D) earthquake cycle model for northeast Japan (Figure 2a). We choose a 2-D forward model for its computational efficiency, enabling us to explore a wide range of the parameters using data assimilation, as opposed to a more resource-intensive 3-D model. To reduce the 2-D effects on our interpretation of the results, we limit the scope of our results to the 100 km wide swath profile defined in Figure 1a. We use the finite element code GTECTON (Govers et al., 2018; Govers & Wortel, 1993) to solve the momentum equation for visco-elastic plane strain (Melosh & Raefsky, 1983). The model consists of an overriding plate (including an upper- and lower crust, each 15 km thick), a 60 km thick slab, an oceanic mantle, a mantle wedge, and a thin low-viscosity shear zone with a thickness of approximately 3 km to accommodate afterslip (Figure 2a). Afterslip has mostly been found in adjacent areas to the coseismic rupture zone (Avouac, 2015). For the 2011

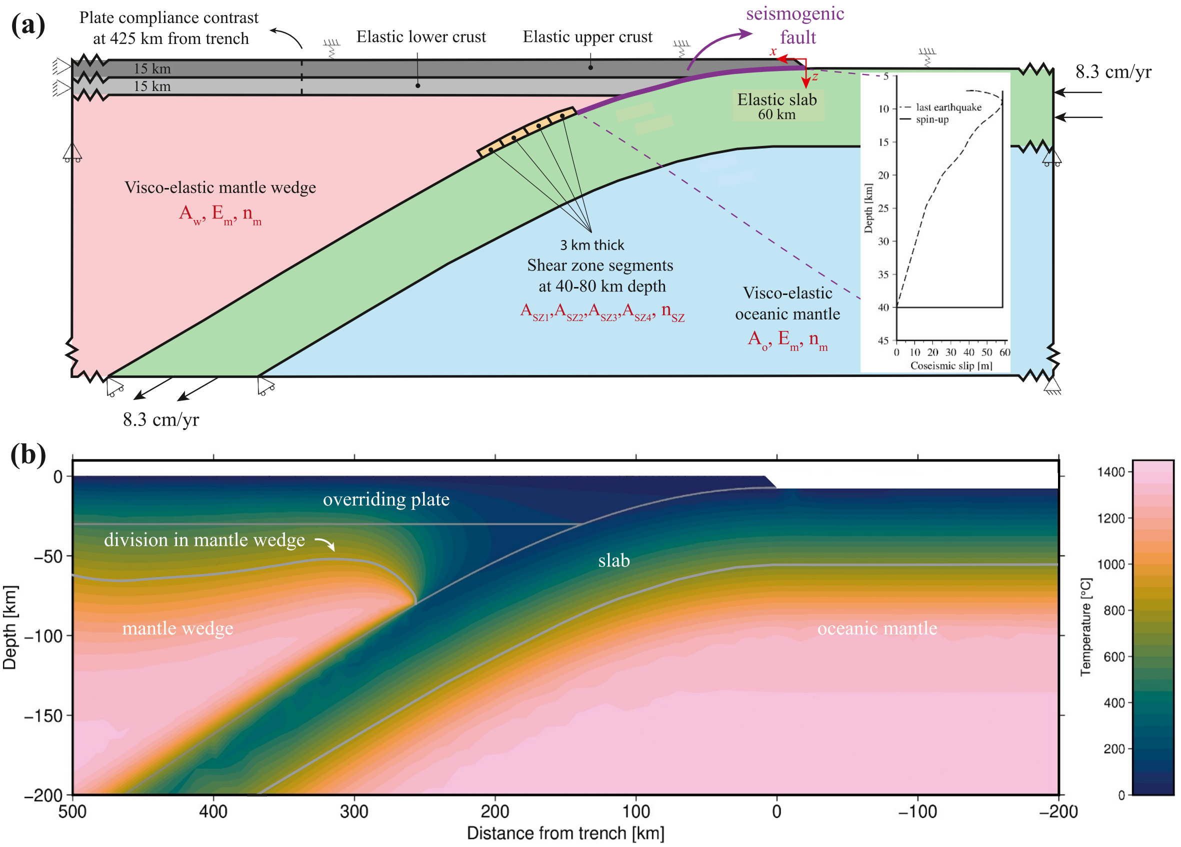


Figure 2. (a) 2-D model setup. The inset on the lower right corner shows the coseismic slip for the spin-up phase and the last earthquake. A, E, and n are the pre-exponent factor, activation energy, and stress power, respectively. Subscripts s, w, o, m, and SZ refer to the shallow mantle wedge, mantle wedge, oceanic mantle, all mantle bodies, and shear zone properties, respectively. $A_{SZ1}, A_{SZ2}, A_{SZ3}$, and A_{SZ4} refer to the shear zone pre-exponent factor between 40 and 50 km, 50–60 km, 60–70 km, 70–80 km depth, respectively. (b) The temperature model (Morishige, 2022), which is used as input to compute the effective viscosity in viscoelastic domains.

Tohoku-Oki earthquake, coseismic slip and afterslip areas have been shown to complement one another (Iinuma et al., 2016). However, some overlap between afterslip and the coseismic rupture has been identified (Johnson et al., 2012; Ozawa et al., 2012; Perfettini & Avouac, 2014; Silverii et al., 2014). Moreover, historically reoccurring earthquakes can illuminate areas of afterslip, which have been documented downdip of the 2011 Tohoku-Oki rupture zone (Uchida & Matsuzawa, 2013). Based on estimates of the depth beyond which the slab and mantle wedge are mechanically coupled (Wada & Wang, 2009), afterslip is commonly assumed to extend to 70–80 km depth (Dhar et al., 2022; Freed et al., 2017; Fukuda & Johnson, 2021; Ozawa et al., 2012; Sun et al., 2014). The narrative for this assumption is that afterslip is unlikely to persist below 80 km depth due to the increased temperature at these depths. Additionally, previously published afterslip inversions show little to no afterslip beyond 80 km depth (Iinuma et al., 2016; Yamagiwa et al., 2015) and the depth beyond which the slab and mantle wedge appear to be mechanically coupled has been constrained by heat flow data in Northeast Japan to 80 km depth (Morishige, 2022; Morishige & Kuwatani, 2020). We consider this depth as the maximum depth of afterslip in our model. We allow afterslip on a shear zone downdip of the rupture zone between 40 and 80 km depth. We test several configurations with one, two, or four shear zone segments and find that four segments are necessitated by the data (Text S1 in Supporting Information S1). We adopt the slab geometry of Zhao et al. (1997), which was also used to derive the temperature field in Morishige (2022) that serves as an input in our models (see below). We consider the geometry of the megathrust well-constrained, and thus do not investigate the role of the geometry on interseismic and postseismic deformation. To avoid boundary conditions at the model edges (except for the velocity boundary conditions around the slab) to affect our results, the model extends 1,000 km to the west and 400 km to the east (with respect to the trench) and to a 500 km depth.

Elastic heterogeneity affects coseismic deformation (Hashima et al., 2016). We adopt an elastic heterogeneous structure similar to Hashima et al. (2016) by assigning different Young's moduli to the overriding plate (87.5 and

875 GPa, near- and far-field, respectively), the slab (212.5 GPa) and the mantle (162.5 GPa). We assume a uniform Poisson ratio of 0.25. A compliance contrast in the overriding plate is incorporated at 425 km from the trench (D'Acquisto et al., 2023) to represent a stronger backarc region (Figure 2a). The far-field value of the Young's modulus for the overriding plate is 10 times higher than the near-field value, based on pilot tests to fit interseismic observations (Figure S8a in Supporting Information S1). We note that the coseismic model we adopt (see below) uses a different elastic structure, but we find an acceptable fit with the coseismic observations (Figure S8b in Supporting Information S1). The far-field value of the Young's modulus of the overriding plate does affect the postseismic results (Figure S8c in Supporting Information S1). However, investigating the influence of the plate compliance considering the imposed rheology, is beyond the scope of this work.

Deformation of the model is driven by velocity boundary conditions along the up- and downdip ends of the oceanic plate/slab. We impose a velocity of 83 mm/yr for the subducting Pacific plate relative to the North American plate at the Tohoku margin (DeMets et al., 2010). We expect that passive mantle convective flow is excited by slab motion with velocities that depend on the mechanical coupling of the slab and the surrounding asthenosphere. We fix the side of the overriding plate to represent the strong far-field lithosphere. The mantle edges are considered flow-open, but we add the following boundary conditions to decrease the spin-up time: we only allow horizontal motion in the shallow parts of the mantle wedge (<100 km depth) and oceanic mantle (<250 km depth) edge domains and slab-parallel motion around the downdip edge of the slab. We apply Winkler pressures (Govers & Wortel, 1999; Wu, 2004) on the upper surfaces of the oceanic and continental plates to simulate buoyancy representing isostatic forces resulting from vertical displacements of density interfaces (2,200 and 3,200 kg/m^3, for the oceanic and continental plates, respectively).

We define the seismogenic part of the megathrust between 0 and 40 km depth. We apply the split-node technique (Melosh & Raefsky, 1980) to simulate coseismic slip. The seismogenic zone is locked throughout the interseismic and postseismic phases and unlocked during the coseismic phase. It is important to address pre-stresses into our modeling as the pre-stress conditions may influence the rheology and thus postseismic behavior, especially when modeling stress-dependent viscosity. To represent the long-term background stress field, we build up pre-stresses in the model over multiple earthquake cycles until pre-stresses are considered steady, that is, the spin-up phase. We do that by kinematically imposing coseismic slip in our models. During the spin-up phase we release all accumulated slip deficit over the entire fault for each cycle, which is equal to the recurrence period times the plate convergence rate. Steady-state stresses are well achieved within the first 9 cycles. We tailor the last earthquake (the 10th cycle) to match the observed coseismic slip of the 2011 Tohoku-Oki earthquake. Note that the accumulated slip deficit prior to the Tohoku earthquake is consequently equal to the recurrence period times the plate convergence rate as we do not release slip deficit in some other form, such as foreshocks. We disregard the effects of this residual slip in our results and assume that this will have minimal consequences for the postseismic period we analyze. We adopt the coseismic model of Hooper et al. (2013), which is based on a joint inversion of GNSS, seafloor and tsunami data. We compute an average coseismic model based on 10 equidistant profiles defined within 50 km of the transect in Figure 1a. We taper the coseismic slip of the averaged model between 30 and 40 km to remove effects of afterslip as the data used in the inversion of the coseismic model contains some early postseismic data (see inset Figure 2a). Moreover, coseismic slip beyond 40 km depth results in a bad fit with coseismic observations, particularly onshore subsidence is overestimated by 1–2 m (Figure S9 in Supporting Information S1). The maximum slip we impose is 58.6 m at 9 km depth. The earthquake recurrence period is set to 706 years to match the maximum slip considering the plate convergence rate, which agrees with a recurrence period of 500–900 years of great earthquakes along the Japan trench (Usami et al., 2018). Similar to D'Acquisto et al. (2023) we identify a residual displacement accumulated over several earthquake cycles due to geometry changes of a curved slab from a model with free slip along the entire interface (i.e., no locking and no earthquakes). The residual does not affect horizontal deformation, whereas it is significant for interseismic vertical displacements (at most 3.4 mm/yr at the eastern coast) (Figure S10 in Supporting Information S1). We remove this residual from our model results for the last earthquake. Note that with this correction applied, the net surface displacement over one cycle (prior to the Tohoku earthquake) is zero.

Stress-driven afterslip is commonly modeled based on either stable frictional behavior (e.g., Marone, 1998; Marone et al., 1991) or viscous deformation (e.g., Bürgmann & Dresen, 2008; Montési, 2004; Montési & Hirth, 2003). In our study, we do not aim to constrain the frictional properties of the megathrust fault. We can model stress-dependent afterslip using a convenient method based on viscous deformation, which can produce similar results for postseismic deformation (Hearn et al., 2002). Instead, we model stress-driven afterslip using a ductile shear zone with either a linear or power-law creep rheology. The interplay between diffusion (i.e., linear) and dislocation creep (i.e., nonlinear) is relevant for viscous deformation; diffusion creep dominates in case of small grain sizes or small stresses, while dislocation creep dominates in the case of large grain sizes or high stresses (Bürgmann & Dresen, 2008). Brittle processes in low-temperature regions reduce the grain size and thus enhance diffusion creep (e.g., Montési & Hirth, 2003). The presence of water also enhances diffusion (Karato et al., 1986). We can expect high water content and small grain sizes along the subduction channel, especially downdip of the coseismic rupture, making diffusion creep important. However, large stresses and an increase in temperature may facilitate dislocation creep. Moreover, the impact of (coseismic) stress on grain size evolution can influence the relative importance of the two creep mechanisms (Montési & Hirth, 2003). In this study, we focus on modeling a single mechanism, assuming that either diffusion or dislocation creep will dominate, rather than modeling the interaction between the two.

We assume that the plate interface immediately undergoes a transition from brittle to ductile creep downdip of the seismogenic zone. In the power-law case, we describe the effective viscosity η_{SZ} within the shear zone with

$$\eta_{\mathrm{SZ}} = \frac{1}{2} A_{\mathrm{SZ}}^{-1} \sigma_E^{1-n_{\mathrm{SZ}}}, \tag{16}$$

where A_{SZ} is the pre-exponent factor of the shear zone, σ_E is the effective shear stress (defined as the second invariant of the deviatoric stress tensor) and n_{SZ} is the stress power of the shear zone. In our models, we assume that n_{SZ} is the same for all four shear zone segments, but we allow for different values of A_{SZ}. The reasoning for this is that we do not expect large variations in n_{SZ} along the shear zone, but the temperature, water content and shear zone width can significantly change along the shear zone which is represented by A_{SZ}. In the linear case, n_{SZ} is set to 1.0 and the term $\frac{1}{2}A_{\mathrm{SZ}}^{-1}$ becomes equivalent to the effective viscosity.

Laboratory tests suggest that mantle flow is driven by dislocation creep (Karato et al., 1986). Moreover, a power-law rheology successfully explained postseismic observations of the 1992 Landers and 1999 Hector Mine earthquakes (Freed & Bürgmann, 2004). Previous postseismic models of the 2011 Tohoku-Oki earthquake have successfully incorporated a power-law rheology with dislocation creep as the driving mechanism for viscoelastic relaxation (Agata et al., 2019; Dhar et al., 2022; Muto et al., 2019). We describe the rheology in the mantle using a Maxwell element with a steady-state stress-dependent power-law flow. The effective viscosity η_m in the mantle of the power-law flow is

$$\eta_m = \frac{1}{2} {A_m}^{-1} \sigma_E^{1-n_m} \exp\left(\frac{E_m}{\mathrm{RT}}\right), \tag{17}$$

where A_m is the pre-exponent factor in the mantle, E_m is the activation energy, R is the gas constant, and T is the temperature. We adopt the thermal structure of Morishige (2022), which is constrained by surface heat flow and bathymetry observations. In the thermal model, the depth where the slab appears to be decoupled from the overriding mantle wedge is defined at 80 km. This leads to the formation of a cold nose. The thermal structure of Morishige (2022) extends 100 km east from the trench, 560 km west from the trench and the bottom is at 200 km depth. We extend the thermal structure in the eastern and western directions to 500 and 1,000 km, respectively, using the given temperature at the boundaries of the thermal model. Additionally, we extend the temperature at the bottom boundary toward 500 km, assuming a constant slab angle for simplicity.

We use the geodetic data to constrain the pre-exponent factor and activation energy and make informed assumptions about their spatial variability as much as possible. First, we adopt a single value for the activation energy for the viscoelastic domains in the upper mantle in our model. The activation energy is a material property that depends on whether the rock deforms by dislocation creep or diffusion creep (e.g., Hirth & Kohlstedt, 2003). Shear wave splitting observations lend overwhelming support to the notion that the upper mantle is seismically anisotropic (Karato, 2008). Olivine is the most relevant mineral for the rheology of the upper mantle because it is the most abundant mineral and mechanically the weakest. Under both laboratory and field conditions, it develops a strongly anisotropic microstructure due to deformation by power-law creep, not diffusion creep (Karato & Wu, 1993). Consequently, shear wave splitting observations are often interpreted as evidence for power-law creep in the upper mantle. Reported variations in E_m (e.g., 410–600 kJ/mol; Hirth & Kohlstedt, 2003; Karato &

Jung, 2003) have limited impact on the viscosity structure in comparison to the variations in temperature (e.g., 400–1380°C (Figure 2b, Figure S11 in Supporting Information S1). Hence the temperature structure dictates the spatial variability in $\exp\left(\frac{E_m}{\mathrm{RT}}\right)$ in Equation 17. By adopting a single activation energy for the viscoelastic domains in the upper mantle, we simplify our model while ensuring it remains consistent with the dominant deformation mechanism and observational constraints.

Second, we assign three distinct pre-exponent factor values across the upper mantle. This spatial division is justified by significant spatial differences in water fugacity driven by the dewatering of subducted sediments and oceanic crust (Karato, 2008). To account for this, we introduce distinct pre-exponent factors for the mantle wedge and the (likely drier) sub-slab mantle (Figure 2a). We also include a division in the mantle wedge following the 800°C isotherm to distinguish between pre-exponent factors for the shallow and deeper parts of the mantle wedge, to achieve a better match with the observations (see results in Section 4.2). Other configurations, such as defining the division at the 900°C, 1,000°C or 1,100°C isotherms, have minimal effects on our results. The data assimilation process determines whether the pre-exponent factors need to vary between the shallow and deeper parts of the mantle wedge and the oceanic mantle. We thus define the parameters A_w, A_s, and A_o as the pre-exponent factor for the deeper mantle wedge, shallow mantle wedge, and oceanic mantle, respectively. For the nonlinear case, we set the stress power of the mantle n_m to 3.0, which is a typical value for olivine, the main constituent of mantle material (Ranalli, 1987).

3.3. Experimental Setup

One model run (using 6,621 time steps, including spin-up) takes about 3 hr on a single core of an AMD EPYC 7543 processor running at 2.8 GHz. Our models are thus computationally inexpensive. We run 100 simulations in parallel for each data assimilation step. Since we use four data assimilation steps, one entire data assimilation experiment takes about 12 hr when our simulations are run in parallel. This allows us to run several data assimilation experiments to investigate the effects of having different settings of the parameter space. We assimilate interseismic and postseismic geodetic observations described in Section 2.2. With these experiments, we aim to find (a) an appropriate rheology for the mantle and shear zone, and (b) whether afterslip and viscoelastic relaxation each have a unique signature on the surface displacements. An overview of all data assimilation experiments is presented in Table 1. The Gaussian approximation of the prior and posterior pdfs of each experiment can be found in Tables S1 and S2 in Supporting Information S1.

First, we present our best-fit results obtained with a power-law rheology of the mantle as well as the shear zone (Experiment 4, Section 4.1). The unknown parameters are the pre-exponent factor of the mantle wedge (A_w), the pre-exponent factor of the shallow mantle wedge (A_s), the pre-exponent factor of the oceanic mantle (A_o), the activation energy of the mantle (E_m), the pre-exponent factor of each of the shear zone segments (A_{SZ1}, A_{SZ2}, A_{SZ3}, A_{SZ4}) as shown in Figure 2a. Since the pre-exponent can vary across several orders of magnitude, we use the base 10 logarithm to estimate the pre-exponent factor. These unknown parameters are included in the state vector for Experiment 4:

$$\mathbf{z} = \left[\log_{10}(A_w), \log_{10}(A_s), \log_{10}(A_o), E_m, \log_{10}(A_{SZ1}), \log_{10}(A_{SZ2}), \log_{10}(A_{SZ3}), \log_{10}(A_{SZ4})\right]^{\intercal}. \tag{18}$$

Note that the unknown parameters included in the state vector may vary per experiment. Then, we investigate the effect of having lower viscosities in the shallow part of the mantle wedge by introducing a division in the mantle wedge (Experiments 4 and 5, Section 4.2). Subsequently, we focus on the rheology of the mantle and the shear zone in terms of a power-law or linear Maxwell creep (Experiments 1 to 4, Section 4.3). We fix the stress power of the mantle and shear zone. We purposely do not rely on a single experiment with a varying stress power, as this would require us to explore a wide range in parameter values corresponding to both linear and nonlinear creep and thus increasing the ensemble size. We thus run several experiments with realistic values of the stress power. In case of linear creep in the shear zone, we replace the creep parameters of the shear zone ($A_{\mathrm{SZ1}}, A_{\mathrm{SZ2}}, A_{\mathrm{SZ3}}, A_{\mathrm{SZ4}}$) with the Newtonian viscosities in each of the shear zone segments (η_{SZ1}, η_{SZ2}, η_{SZ3}, η_{SZ4}) in the state vector. For a linear Maxwell rheology of the mantle, we estimate the pre-exponent factor in each part of the mantle. This way, the viscosity structure in the mantle is still temperature-dependent and thus spatially heterogeneous.

Subsequently, we explore potential trade-offs between parameters representing afterslip and viscoelastic relaxation. We find that correlations between these parameters are better represented when we fix the activation energy

Table 1
Description of Data Assimilation Experiments

Exp.	ID	Description	Size state vector	Mean and st. dev. of cost function ensemble	Mean and st. dev. of data cost function ensemble
1	0043	Linear creep in the mantle and shear zone	8	1.6627 ± 0.91	0.1617 ± 0.04
2	0044	Linear creep in the mantle and power-law creep in shear zone	8	2.1975 ± 0.98	0.3086 ± 0.07
3	0001	Power-law creep in the mantle and linear creep in the shear zone	8	0.4733 ± 0.35	0.0692 ± 0.03
4	**0009**	**Power-law creep in the mantle and shear zone**	**8**	**0.4284 ± 0.39**	**0.0705 ± 0.04**
5	0012	Same as Exp. 4, but without a division in the mantle wedge	7	0.2969 ± 0.30	0.0728 ± 0.04
6	0013	Same as Exp. 4, but set $E_m = 420$ kJ/mol	7	0.1376 ± 0.07	0.0591 ± 0.02
7	0014	Same as Exp. 3, but set $E_m = 420$ kJ/mol	7	0.1152 ± 0.14	0.0498 ± 0.01
8	0016	Same as Exp. 4, but we include n_{SZ} in the state vector	9	0.4534 ± 0.29	0.0805 ± 0.03
9	0005	Same as Exp. 4, but set $n_{SZ} = 2.0$	8	1.1682 ± 0.66	0.1355 ± 0.07
10	0006	Same as Exp. 4, but set $n_{SZ} = 3.0$	8	0.6361 ± 0.42	0.0795 ± 0.04
11	0007	Same as Exp. 4, but set $n_{SZ} = 4.0$	8	0.3170 ± 0.28	0.1144 ± 0.02
12	0008	Same as Exp. 4, but set $n_{SZ} = 5.0$	8	1.1682 ± 0.66	0.1356 ± 0.07
13	0077	Same as Exp. 15, but the shear zone viscosity is temperature-dependent.	6	0.7614 ± 0.65	0.0878 ± 0.48
14	0078	Same as Exp. 13, but we include 4 shear zone segments	9	0.4371 ± 0.30	0.0826 ± 0.05
15[a]	0039	Same as Exp. 4, but we only include 1 shear zone segment	5	0.7387 ± 0.76	0.0692 ± 0.03
16[a]	0040	Same as Exp. 4, but we only include 2 shear zone segments	6	0.4838 ± 0.46	0.0674 ± 0.03
17[a]	0076	Same as Exp. 4, but A_{LC}, E_{LC}, n_{LC} are added to the state vector	11	0.3021 ± 0.08	0.1164 ± 0.01
18[a]	0071	Same as Exp. 17, but fix the lower crust stress power to $n_{LC} = 1.5$	10	0.3117 ± 0.24	0.1018 ± 0.02
19[a]	0072	Same as Exp. 17, but fix the lower crust stress power to $n_{LC} = 2.0$	10	0.3546 ± 0.32	0.1090 ± 0.02
20[a]	0073	Same as Exp. 17, but fix the lower crust stress power to $n_{LC} = 2.5$	10	0.2902 ± 0.25	0.1153 ± 0.02
21[a]	0074	Same as Exp. 17, but fix the lower crust stress power to $n_{LC} = 3.0$	10	0.3568 ± 0.18	0.1376 ± 0.02
22[a]	0075	Same as Exp. 17, but fix the lower crust stress power to $n_{LC} = 3.5$	10	0.3275 ± 0.10	0.1423 ± 0.01
23[a]	0017	Same as Exp. 4, but only assimilate horizontal observations	8	0.4636 ± 0.22[b]	0.4148 ± 0.14[b]
24[a]	0018	Same as Exp. 4, but only assimilate vertical observations	8	0.3557 ± 0.16[b]	0.1278 ± 0.09[b]
25[a]	0002	Same as Exp. 3, but only assimilate horizontal observations	8	0.5664 ± 0.29[b]	0.5525 ± 0.13[b]
26[a]	0003	Same as Exp. 3, but only assimilate vertical observations	8	0.3849 ± 0.19[b]	0.1947 ± 0.14[b]
27[a]	0083	Same as Exp. 4, but with a shallower shear zone segment	9	0.4486 ± 0.32	0.0705 ± 0.04

Note. The data cost function only entails the part of Equation 10 concerning the fit with the data, defined as $\mathcal{J}_{\mathrm{data}}(\boldsymbol{\Theta}_{j,i+1}) = \left(\boldsymbol{g}(\boldsymbol{\Theta}_{j,i+1}) - \boldsymbol{d} - \sqrt{\alpha_{i+1}}\,\boldsymbol{e}_{j,i}\right)^{\mathsf{T}}(\alpha_{i+1}\boldsymbol{C}_{dd})^{-1}\left(\boldsymbol{g}(\boldsymbol{\Theta}_{j,i+1}) - \boldsymbol{d} - \sqrt{\alpha_{i+1}}\,\boldsymbol{e}_{j,i}\right)$. [a]Data assimilation results can be found in Supporting Information S1. [b]Computed with both horizontal and vertical data.

of the mantle and thus remove it from the state vector (Experiments 6 and 7, Section 4.6). Note that we generally consider the activation energy as an unknown in our remaining experiments as the activation energy cannot be directly observed and laboratory estimates are considerably uncertain (e.g., up to 75 kJ/mol uncertainty, Hirth and Kohlstedt (2003)). Moreover, we investigate the magnitude of the stress power for the shear zone (Experiments 8 to 12, Section 4.4), and the influence of a temperature-dependent viscosity of the shear zone (Experiments 4, 13 and 14, Section 4.5). We run additional data assimilation experiments to investigate the influence of the number of shear zone segments (Experiments 4, 15, and 16, Text S1 in Supporting Information S1), a power-law rheology of the lower crust (Experiments 17 to 22, Text S2 in Supporting Information S1), and the potential bias of only assimilating either horizontal and vertical data (Experiments 23 and 24, Text S3 in Supporting Information S1).

In our ESMDA experiments, we use an ensemble size, N_e, of 100 and set the number of data assimilation steps, N_{mda}, to 4. For various ensemble sizes (50–200) we find that the cost function converges after 5 data assimilation steps (Figure S12 in Supporting Information S1). However, our posterior models show little difference (i.e., approximately 1% difference in posterior displacements) between experiments with a total of 4 or 5 data assimilation steps. Therefore, we select 4 data assimilation steps.

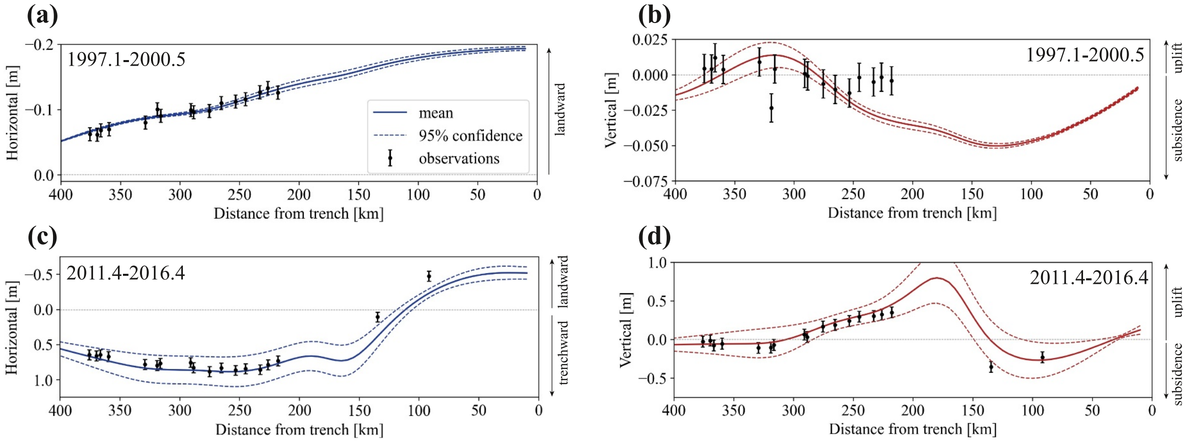


Figure 3. (a), (b) Interseismic and (c), (d) 5-yr postseismic displacements as a function of distance from the trench from Experiment 4 (i.e., the best performing experiment). Solid lines and uncertainty bands represent the mean and two standard deviations (i.e., 95% confidence level) of the ensemble of modeled displacements.

We assume equal weighting of horizontal and vertical data at all stations. We purposely do not use the uncertainty following the parameterization of the time series in our assimilation procedure as this underrepresents the total observational error (Marsman et al., 2023). Instead, we assume a constant observation error that encompasses both the representation and instrumentation errors as shown in Figures 1c and 1d. We assume an observation error of 7 cm for postseismic data, which encompasses both the data and model errors (based on postseismic inversion results by Fukuda and Johnson (2021)). For the interseismic period, we assign an observation error of 1 cm, such that the ratio between the data and observational error is similar to that of postseismic data (Figures 1c and 1d). Note that we include a higher quantity of postseismic data (1,094 observations over 10 years) than interseismic data (324 observations over 3 years). Moreover, there is significantly more on-land GNSS data available (1,260 observations) than offshore GNSS/A data (158 observations). Therefore postseismic- and interseismic data are weighted approximately 3 to 1, and onshore and offshore data are weighted approximately 8 to 1.

4. Results and Analysis

4.1. Best Model Results

We begin by providing detailed results of our best performing data assimilation experiment (Experiment 4, Table 1). We define this experiment as the best performing experiment as it results in a low cost function (Equation 10). Note that the number of parameters included in the state vector can greatly influence the cost function, hence we also carefully analyzed the match with the interseismic and postseismic observations. We define the "best-fit model" as the forward model assigned with mean ensemble parameter values from the best performing experiment, which is Experiment 4. The best-fit model is characterized by the following elements: (a) temperature-dependent power-law creep in the mantle, (b) temperature-independent power-law creep in the shear zone, (c) a division in the mantle wedge, (d) elastic crust, and (e) four shear zone segments. These modeling choices follow from our data assimilation experiments detailed in Sections 4.2 to 4.5 and Texts S1–S3 in Supporting Information S1.

During the interseismic phase, our best-fit model accurately predicts landward horizontal motion (Figure 3a). The ensemble captures the general deformation pattern shown by the interseismic verticals, that is, subsidence in the forearc and uplift in the backarc (Figure 3b) due to sufficient flexure of the overriding plate. However, the ensemble of models exhibits too much subsidence on the eastern coast compared to the observations, considering the observation error. Interseismic vertical displacements are sensitive to details in the modeling approach, particularly in the viscoelastic structure (Figure S22 in Supporting Information S1), elasticity parameters and the locking pattern (e.g., Trubienko et al., 2013). Additionally, interseismic deformation is sensitive to other non-tectonic processes present in the data, such as groundwater extraction. The actual representation error, that partly accounts for missing physics, is likely larger and hence the observation error should be larger than our assumed 1 cm at some locations, such as the eastern coast and where groundwater extraction takes place. Overall, interseismic results match well with the observational data.

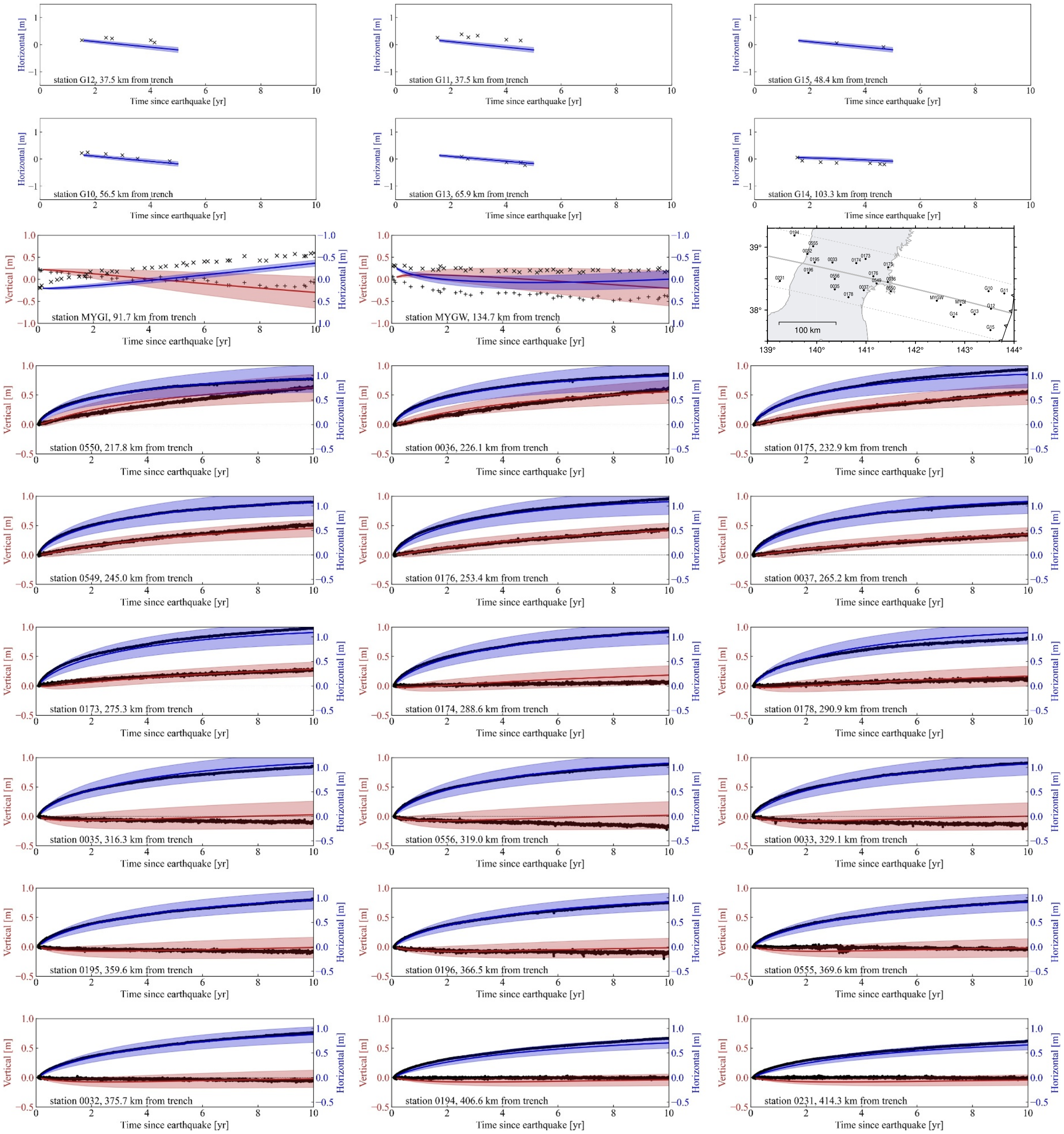


Figure 4. Postseismic time series of all stations resulting from Experiment 4 (best performing ensemble). Solid lines and uncertainty bands represent the mean and two standard deviations (i.e., 95% confidence level) of the ensemble of modeled displacements. Sign convention horizontals: trenchward is positive and landward is negative. Sign convection verticals: uplift is positive and subsidence is negative.

Postseismically, the ensemble shows onland trenchward motion and near-trench landward motion (Figure 3c). The models exhibit near-trench uplift, offshore subsidence, coastal uplift and backarc subsidence, in agreement with the data (Figure 3d). This is a typical pattern produced by the combined effect of afterslip and viscoelastic relaxation (Luo & Wang, 2022). The models capture the viscous response in the onshore data in both the horizontal (blue lines Figure 4) and vertical directions (red lines Figure 4) very well within 95% confidence levels. Near-trench horizontal motion between 2 and 5 years after the earthquake is quite linear and is generally captured well (Figure 4). However, the evolution of the offshore stations is not always captured within the 95% confidence level at stations MYGW and MYGI. This is because the model struggles to accurately represent the actual physics and thus the assumed representation error is too low. In summary, we find a good match with both postseismic onshore and offshore data and can capture the viscous response in the data.

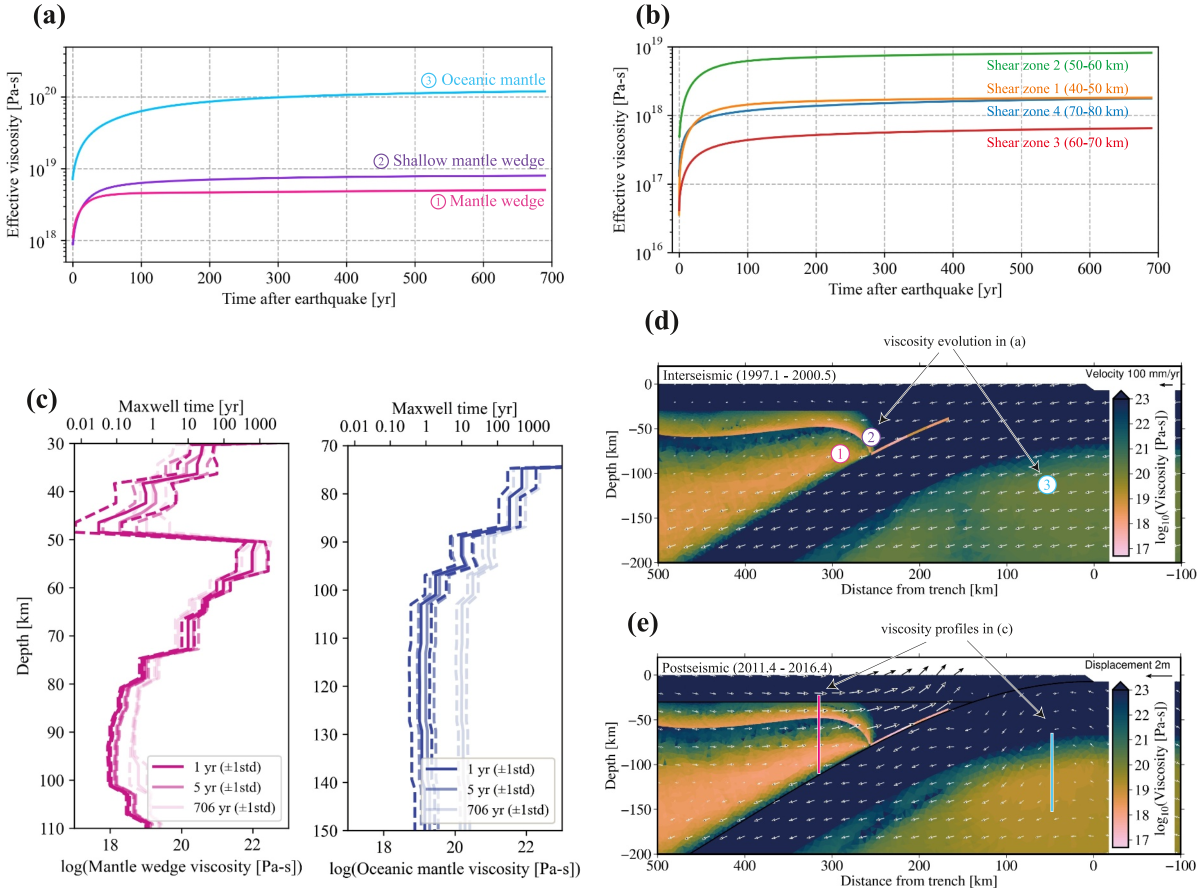


Figure 5. Simulation results on the effective viscosity of the best performing data assimilation test (Experiment 4), with a power-law rheology in both the mantle ($n_m = 3.0$) and shear zone ($n_{SZ} = 2.5$). (a) Effective viscosity evolution at selected locations in the mantle wedge, shallow mantle wedge and oceanic mantle shown in panel (c) (b) Effective viscosity evolution of the shear zone. (c) Effective viscosity as a function of depth at 1, 5 and 706 years after the earthquake along select profiles in the mantle wedge and oceanic mantle. The locations of the viscosity profiles can be found in panel (e) (d) Effective viscosity during the interseismic phase. (e) Effective viscosity and displacements during the postseismic phase, 5 years after the earthquake. Effective viscosities in the mantle and shear zone are calculated using Equations 16 and 17, respectively. The results in a, b, d and e are obtained from a single model with input mean parameter values from Experiment 4.

Next, we analyze the spatiotemporal viscosity structure of the mantle and shear zone. We find low effective viscosities in the shallow mantle wedge (<50 km depth) of 7.0×10^{18} and 8.6×10^{18} Pa·s directly after the earthquake and 1 year after, respectively (Figures 5a and 5e). Our models particularly contain low viscosities in the shallow part of the mantle (<50 km depth, Figure 5c), whereas previous studies typically infer lower viscosities at larger depths ranging between 80 and 300 km depth excluding the volcanic arc (Agata et al., 2019; Dhar et al., 2022; Freed et al., 2017). The effective viscosities of the oceanic mantle are slightly smaller than 10^{19} Pa·s directly after the earthquake, and increase toward 10^{20} Pa·s interseismically (Figures 5a, 5d and 5e). Previous studies employing a Burgers rheology have inferred a short-term viscosity of the order 10^{18}–10^{19} Pa·s and a steady-state viscosity of 10^{20} Pa·s for the oceanic mantle (e.g., Hu et al., 2014; e.g., Iinuma et al., 2016; Sun et al., 2014), which agrees with our estimates. However, we do not require a separate low-viscosity sub-slab asthenosphere, indicating that there is less viscoelastic relaxation in the sub-slab asthenosphere required by our models (see discussion in Section 5.4). Our best-fit model shows that afterslip is concentrated downdip of the coseismic rupture between 40 and 50 km depth and decreases toward zero at 80 km depth (Figure 12b), which is compatible with the previously published models with stress-driven afterslip. Our results suggest a small increase in afterslip with depth between 60 and 70 km depth, particularly present within the first 3 years after the earthquake. The viscosity of the shear zone segments are ordered from low to high from η_{SZ3}, η_{SZ1}, η_{SZ4}, to η_{SZ2} (Figure 5b) which directly control the afterslip pattern together with the effective shear stress. The increase in afterslip between 60 and 70 km is attributed to the low viscosity for this shear zone segment (red line in Figure 5b). The peak in afterslip between 40 and 50 km is 2.8 , 3.7, 4.1 and 4.5 m after 1, 3, 6 and 10 years,

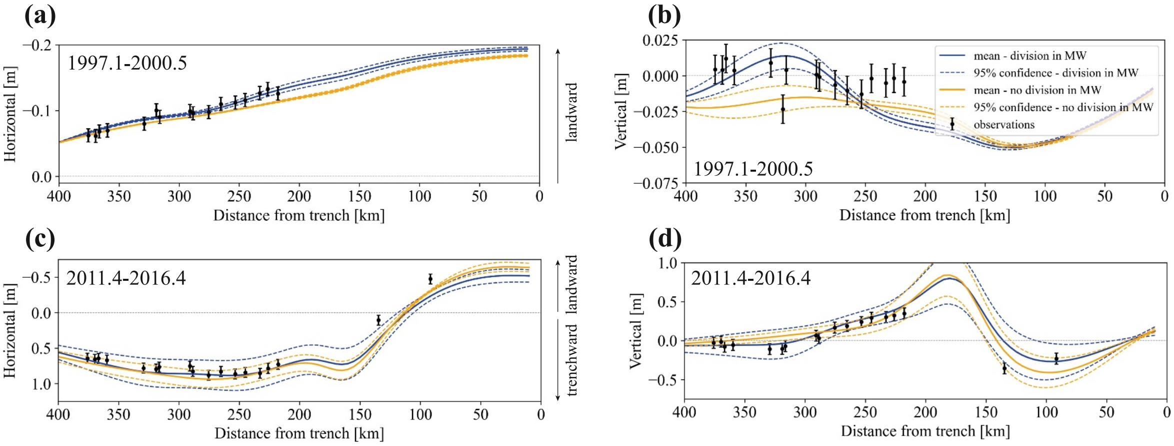


Figure 6. Cumulative (a) horizontal and (b) vertical interseismic displacements after approximately 3 years where we either do (blue, Experiment 4) or do not (orange, Experiment 5) consider a division in the mantle wedge. (c) Cumulative horizontal and (d) vertical postseismic displacements 5 years after the earthquake where we either do (blue) or do not (orange) consider a division in the mantle wedge.

respectively (Figure 12b). We further discuss the implications of the shear zone rheology and afterslip in our models in Section 5.1. The spatiotemporal viscoelastic structure of the mantle and the inferred afterslip pattern result in a good match with both interseismic and postseismic observations. Hence our best-fit model can successfully explain different scales across the earthquake cycle with a single rheology.

4.2. A Division in the Mantle Wedge

We also investigate the need for a division in the mantle wedge (Experiments 4 and 5, Table 1). We first focus on modeling results without the division in the mantle wedge (Experiment 5). During the interseismic phase, the ensemble of models without the division in the mantle wedge produces landward motion, but they are slightly too small near the eastern coast (orange lines in Figure 6a). Moreover, the ensemble shows onshore subsidence (orange lines in Figure 6b) as there is not enough flexure of the overriding plate. Postseismically, these models can accurately reproduce landward near-trench motion and onshore trenchward motion (orange lines in Figure 6c). The postseismic vertical pattern is captured well within 95% confidence at almost all sites, except for a few sites between 300 and 350 km from the trench (orange lines in Figure 6d). In short, the models without the division fail to reproduce the correct interseismic pattern, particularly the vertical pattern, and fall short in explaining postseismic backarc subsidence.

The viscosity beneath the overriding plate between 50 and 100 km depth is well above 10^{21} Pa·s 5 years after the earthquake (Figure S13 in Supporting Information S1) without the mantle wedge division. Hence, the effective elastic thickness of the lithosphere (both on short and longer timescales) is significantly higher than the elastic upper plate thickness inferred in previous models that is typically assumed between 20 and 50 km (Agata et al., 2019; Dhar et al., 2022; Freed et al., 2017; Fukuda & Johnson, 2021; Hu et al., 2016; Iinuma et al., 2016; Muto et al., 2019; Suito, 2017; Sun et al., 2014; Yamagiwa et al., 2015). As a result, the continental lithosphere does not bend enough during the interseismic phase. The flexural rigidity of the lithosphere can be reduced by decreasing the viscosity within the shallow part of the mantle wedge, thereby reducing the effective elastic thickness. This can be achieved by including a division in the mantle wedge, allowing us to assign a different pre-exponent factor in the shallow part of the mantle wedge (A_s) than in the deeper part (A_w). Consequently, a (large) jump in viscosities with depth is allowed, and lower viscosities (and thus faster viscoelastic relaxation) can be achieved within the shallow part of the mantle wedge. We define the division according to the 800°C isoline (Figure 2) so that there is a smooth transition in the viscosity structure as the viscosities are scaled by the temperature.

Models with a division in the mantle wedge reproduce a postseismic pattern similar to models without the division (blue lines in Figures 6c and 6d). Additionally, the mean of the ensemble with the division better captures backarc subsidence than without the division. The ensemble with the division captures interseismic landward motion at all

10.1029/2024JB029665

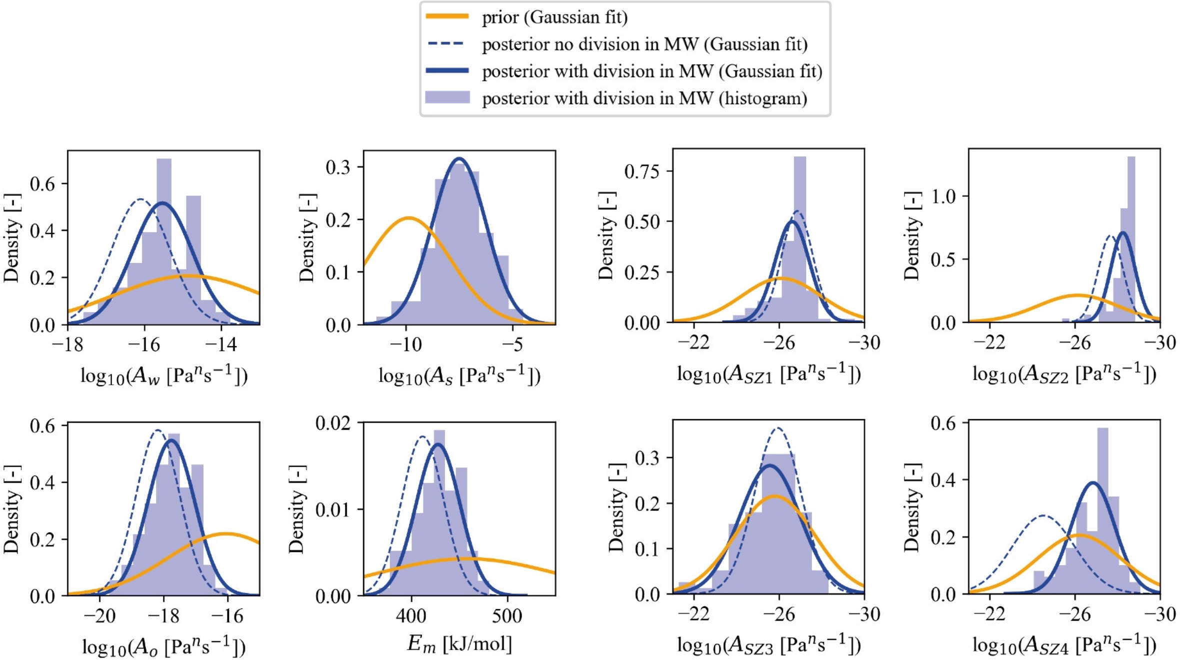


Figure 7. Prior and posterior probability distributions of the creep parameters and shear zone viscosity parameters for the case where we either do (blue, solid lines) or do not (blue, dashed lines) consider a division in the mantle wedge with power-law creep in the mantle ($n_m = 3.0$) and power-law creep in the shear zone ($n_{SZ} = 2.5$) (Experiments 4 and 5).

sites (blue lines Figure 6a) and produces enough flexure to match interseismic uplift in the backarc (blue lines Figure 6b).

The 1-D and 2-D distribution of parameters used to calculate the ensemble of model displacements are shown in Figures 7 and 8. For all parameters, we find a different posterior distribution (blue lines in Figure 7) than the prior (orange lines in Figure 7), indicating that the data has enough information to constrain these parameters. We find a good match for the creep parameters A_w, A_o and E_m with laboratory derived values (Figure 8) (Hirth & Kohlstedt, 2003). In our experiment without the division in the mantle wedge (dashed lines in Figure 7), we find slightly smaller estimates for A_w, A_o and E_m than with the division (solid blue lines in Figure 7) in the mantle wedge, but there is still an overlap with laboratory values. The spread and shape of the 2-D scatter plots (Figure 8) give an indication of a correlation between parameters and how well ESMDA can estimate individual parameters. If there is no trend between the parameters, or if a vertical or horizontal pattern is observed, then the correlation is low. For example, a high correlation between the pre-exponent factor and activation energy is found (ρ between 0.94 and 1.00), evident from the trend in the 2-D scatter plots (Figure 8). A smaller pre-exponent factor results in larger viscosities, whereas a smaller activation energy results in smaller viscosities. In both cases, we find similar viscosity estimates in the oceanic and mantle wedge (excluding the shallow mantle wedge part) (Figure 5e, Figure S13a in Supporting Information S1), explaining the high correlation. High viscosities in the shallow part of the mantle wedge result in less viscoelastic relaxation in models without the division. Consequently, more afterslip is required to match the observations (Figure S13b in Supporting Information S1). Moreover, without the division, the effective elastic plate thickness in the model becomes too large, and thus the plate is not able to bend enough during both the interseismic and postseismic phases. The division in the mantle enables more precise control over viscoelastic relaxation and lithospheric bending. Therefore, models with the division in the mantle wedge result in a closer match to the observations, particularly with the interseismic vertical observations and postseismic vertical observations in the backarc (more than 300 km from the trench) than models without the division.

4.3. Power-Law Versus Linear Viscous Rheology

We systematically explore different rheologies within the (shallow) mantle wedge, the oceanic mantle and the shear zone with our data assimilation experiments (Experiments 1–4, Table 1). In these models, we include a division in the shallow mantle wedge, which improves the match with the observations as explained in

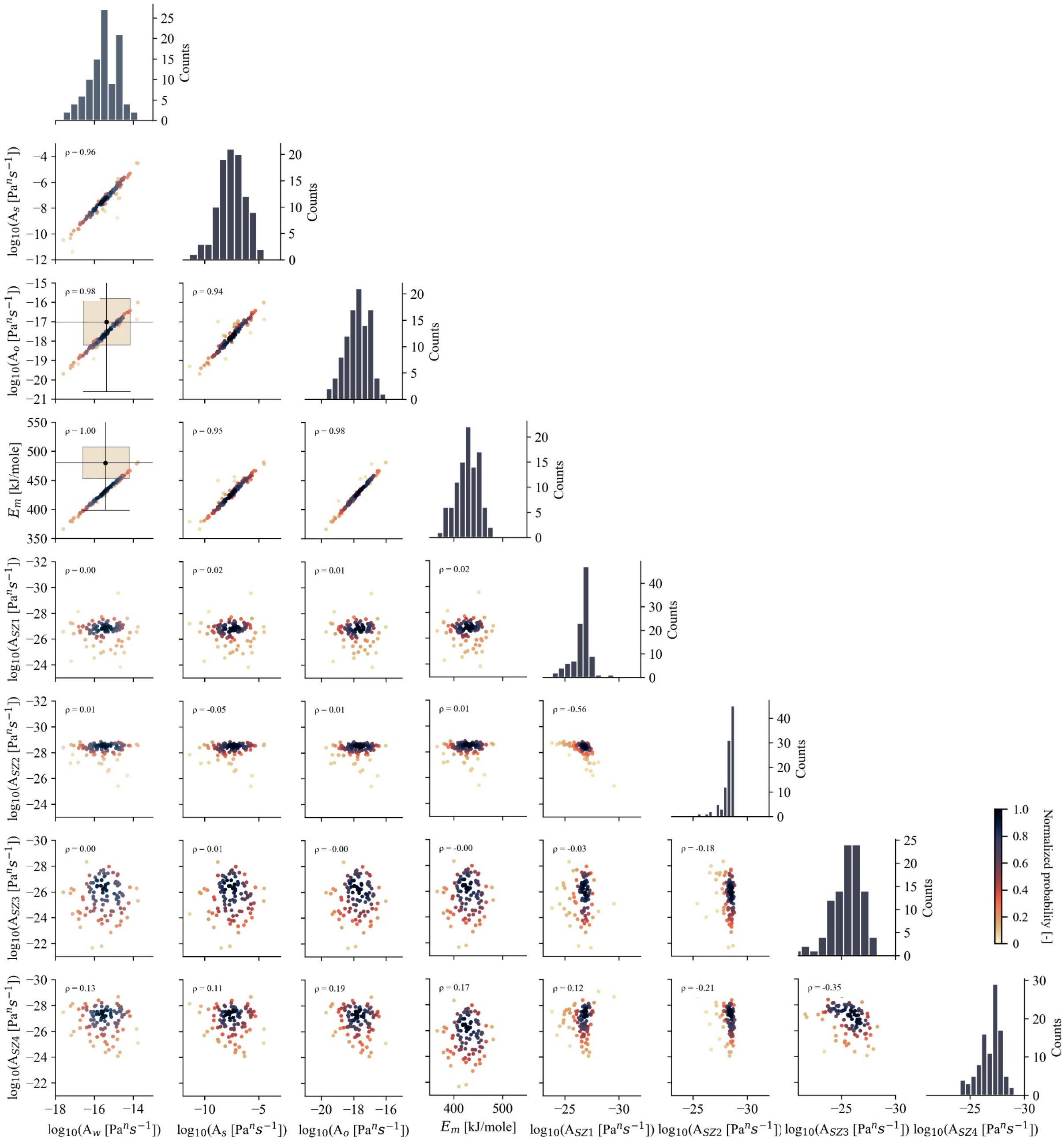


Figure 8. Two-dimensional posterior probability distributions where we consider a division in the mantle wedge with power-law creep in the mantle ($n_m = 3.0$) and power-law creep in the shear zone ($n_{SZ} = 2.5$) (Experiment 4). The whisker boxplots represent the laboratory values of the activation energy of the mantle, the pre-exponent factor of the mantle wedge (assuming a water content of 0.1%) and the pre-exponent factor of the oceanic mantle (assuming a water content of 0.005%) (Hirth & Kohlstedt, 2003).

Section 4.2. We test the following rheology combinations: (a) linear Maxwell mantle and shear zone, (b) power-law mantle and a linear shear zone, (c) linear mantle and a power-law shear zone, and (d) power-law mantle and shear zone. In each experiment, we estimate eight parameters using ESMDA: the pre-exponent factor of the shallow mantle wedge, the mantle wedge, the oceanic mantle, the four shear zone segments, and the activation energy of the mantle. For a linear Maxwell rheology, we set the stress power to 1.0. For the power-law rheology of the mantle we set the stress power to 3.0 (Karato et al., 1986) and for the power-law rheology of the shear zone we set the stress power to 2.5 (see Section 4.4). Generally, we do not include the stress power in the state vector of the data assimilation, as this would result in larger uncertainties in our posterior pdfs. We analyze

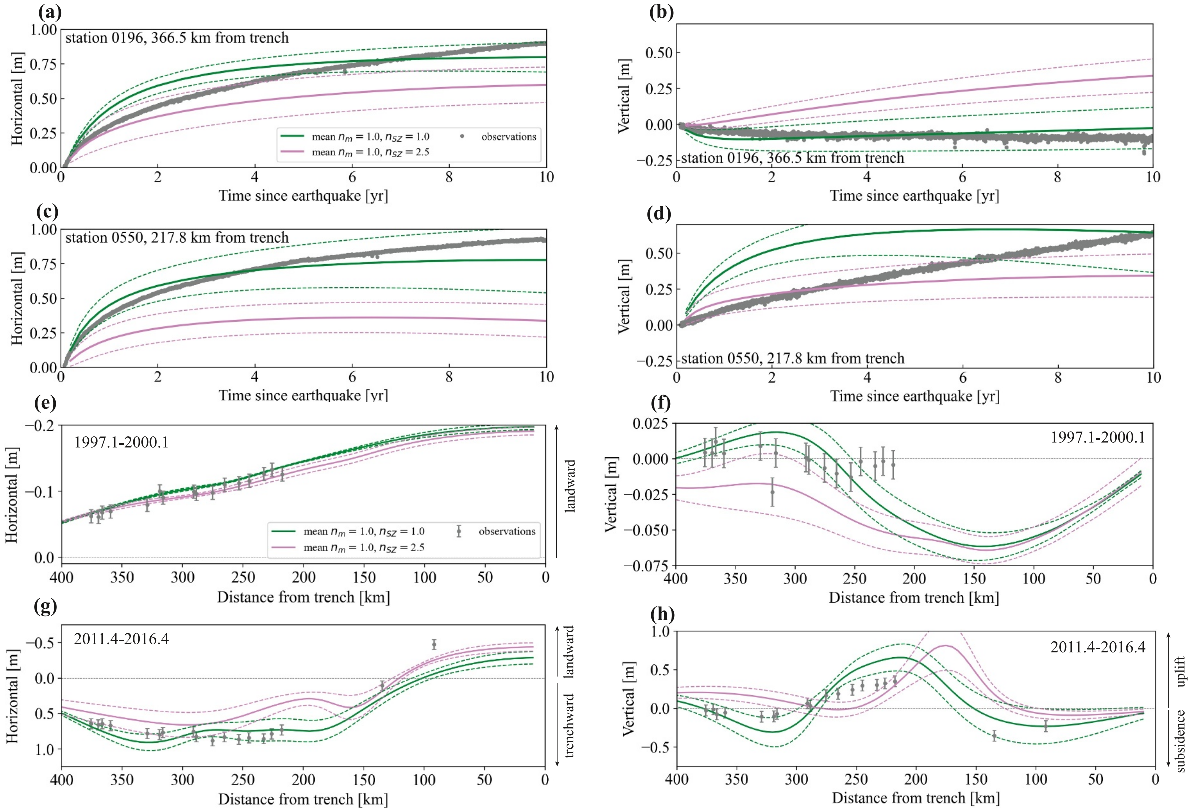


Figure 9. Impact of modeling power-law (pink lines) or linear (green lines) creep in the shear zone with linear creep in the mantle on the interseismic and postseismic displacements (Experiments 1–2). The solid lines represent the mean of the posterior ensemble and the dashed lines represent the 95% confidence interval (2σ uncertainty). (a), (c) Postseismic horizontal and (b), (d) vertical times series of stations 0196 and 0550 (see Figures 1 and 4 for locations). (e), (f) Cumulative 3-year interseismic and (g), (h) 5-year postseismic displacements. Error bars indicate 1σ data uncertainties.

how well each rheology combination can explain the data using ESMDA. Note that for each combination, ESMDA seeks the best combination of parameters to minimize the cost function. For ease of our evaluation, we focus on the mean of the posterior ensemble of modeled displacements and the associated 95% confidence level (i.e., two standard deviations) for each combination.

For each rheology combination, we observe shortening of the overriding plate with landward motion (Figures 9 and 10e), along with subsidence in the forearc and uplift in the backarc due to flexure before the earthquake (Figures 9 and 10f, except for the pink line). All our models exhibit excessive interseismic subsidence on the eastern coast. As explained in Section 4.1, this may be attributed to both model errors and non-tectonic processes not captured by the model, resulting in a larger representation error than assumed. During the postseismic phase, our models show onland trenchward motion and near-trench landward motion (Figures 9 and 10a, 10c and 10g). Generally, our models show offshore subsidence, coastal uplift and backarc subsidence (Figures 9 and 10h); however, the location and magnitude may differ between the different choices of rheology. Even though all the models with different rheology combinations generally produce the right pattern, the details in the postseismic viscous response and steady-state viscosity structure allow us to identify the best model and potential pitfalls of certain rheology choices.

First, we focus on the model results with linear creep in the mantle and either linear or power-law creep in the shear zone. The ensemble results with power-law creep in the shear zone deteriorate in comparison to the case with linear creep, and the cost function increases by more than 30% (Table 1). This shows that the transient response in afterslip does not lead to better results when modeling a linear viscous mantle. In particular, we observe that the ensemble results do not align well with interseismic verticals with power-law creep in the shear zone (Figure 9f). This discrepancy is due to relatively low shallow viscosities ($>10^{17}$ Pa·s above 50 km depth) in the mantle wedge (not shown), resulting in a weak lithosphere. Consequently, the upper plate experiences little

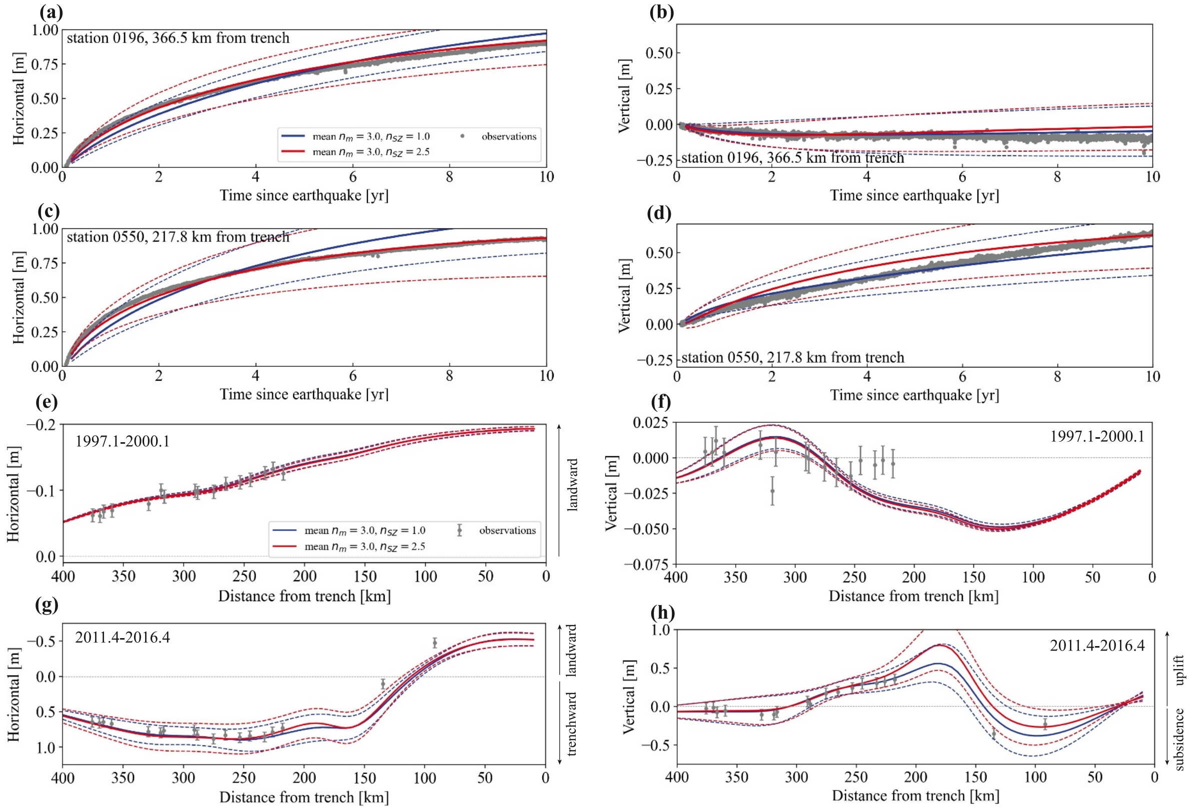


Figure 10. Impact of modeling power-law (blue lines) or linear creep (red lines) in the shear zone with power-law creep in the mantle on the interseismic and postseismic displacements (Experiments 3–4). The solid lines represent the mean of the posterior ensemble and the dashed lines represent the 95% confidence interval (2σ uncertainty). (a), (c) Postseismic horizontal and (b), (d) vertical times series of stations 0196 and 0550 (see Figures 1 and 4 for locations). (e), (f) Cumulative 3-year interseismic and (g), (h) 5-year postseismic displacements. Error bars indicate 1σ data uncertainties.

flexure and can easily be dragged down by the slab to which it is mechanically coupled. Moreover, the fully linear model is unable to explain the beginning and end of the postseismic time series (green lines in Figures 9a–9d) as the model relaxes too quickly. When considering linear creep in the mantle and power-law creep in the shear zone, we observe a too slow response during the postseismic phase (pink lines in Figures 9a–9c) indicating insufficient relaxation. Overall, these results show that models with linear creep in the mantle cannot accurately explain the temporal evolution of postseismic data. Additionally, with power-law creep in the shear zone, the models also poorly match the interseismic verticals.

Models with power-law creep in the mantle result in a closer match to both interseismic and postseismic data (Figure 10) than with linear creep in the mantle. Particularly, the viscous response improves as both the beginning and end of the postseismic time series are well matched with the observations (Figures 10a–10d), and the resultant steady-state viscosity structure can explain the interseismic data as well (Figures 10e and 10f). Differences in modeled displacements resulting from power-law creep (red lines in Figure 10) and linear creep (blue lines in Figure 10) in the shear zone are minimal, with a large overlap between the 95% confidence intervals. Moreover, the cost functions of these two experiments are very similar, making it difficult to assess whether linear or power-law creep in the shear zone is necessitated by the data. However, early postseismic horizontals (<3 years) are somewhat underpredicted and later postseismic horizontals (>5 years) are somewhat overpredicted with linear creep in the shear zone, indicating that relaxation is too slow, which is a general pattern shown at most of the station locations (Figure S15 in Supporting Information S1). When we assume power-law creep in both the mantle and shear zone, we observe an improvement of the postseismic horizontal displacements at all station locations (Figure 4), as more relaxation takes place early on. Note that the mean of the data cost function ensemble is slightly higher when we assume power-law creep in the shear zone than with linear creep (Experiments 3 and 4, Table 1). This is due to the fact the postseismic vertical displacements slightly deteriorate when we assume a power-law in the shear zone. However, we find it essential to accurately match the early postseismic horizontals.

Overall, we find that a power-law creep in both the mantle and shear zone provides the best explanation of the data.

Nonlinear crustal flow within the lower crust has been shown to agree with observations following the 2010 Maule earthquake (Peña et al., 2020) and to contribute to a secondary zone of coseismic and early postseismic uplift (van Dinther et al., 2019). Moreover, low-temperature dislocation creep has also been inferred following the 2013 Baluchistan earthquake (Cheng et al., 2022). We test several configurations in our model with a power-law rheology within the lower crust. Following our previous analysis, we focus on models that include power-law creep in both the mantle and shear zone. When we include viscoelastic flow within the lower crust, we find that both interseismic and postseismic results deteriorate - particularly, no interseismic uplift is predicted and interseismic landward motion is overestimated due to a weaker plate (see Text S2 in Supporting Information S1). Hence, we only consider an elastic lower crust in our models.

4.4. The Shear Zone Stress Power

We run several data assimilation tests with a different stress power of the shear zone (Experiments 3, 4 and 9–12, Table 1). In a separate test (Experiment 8, Table 1), we include n_{SZ} as an additional unknown parameter in the state vector. Here, we obtain a posterior mean and standard deviation of 2.61 ± 0.14. However, the uncertainty of our ensemble results increases significantly when we include n_{SZ} as an additional parameter in the state vector. Therefore, we prefer to use a pre-defined stress power in our experiments. The closest match with the data is obtained when we include power-law creep in the shear zone with a shear zone stress power n_{SZ} of 2.5 (Figure 11). For a higher stress power than 2.5, relaxation occurs faster in the shear zone, and thus we generally observe larger trenchward motion and larger uplift and subsidence at the GNSS station locations. For the linear Maxwell case (i.e., stress power 1.0), the modeled displacements evolve more gradually and are unable to match the fast changes in the early horizontal observation times series. However, for $n_{\mathrm{SZ}} = 2.0$, we observe faster simulated displacements than with $n_{\mathrm{SZ}} = 2.5$. This is probably because ESMDA minimizes the cost function, leading to adjustments in other parameters besides the shear zone stress power, such that faster motion is simulated.

When we set the shear zone stress power to 1.0, afterslip is more evenly distributed with depth and a small peak is found around 55 km depth than when we set the shear zone stress power to 2.5 or 4.0 (Figure 12). More equivalent moment release is predicted with increasing stress power, particularly updip. Not only does the afterslip distribution change with a different shear zone stress power but the viscoelastic structure is optimized to minimize the cost function. For a higher shear zone stress power (≥ 4.0), less viscoelastic relaxation is observed in the shallow part of the mantle wedge (Figure S16 in Supporting Information S1), and the match with interseismic vertical observations deteriorates (Figure 11h). The former would suggest a trade-off between afterslip and bulk viscoelastic relaxation, which will be explored in Section 4.6.

4.5. Temperature-Dependent Shear Zone Viscosity

Our results from Section 4.4 show that a stress power of 2.5 is preferred, indicating that dislocation creep is driving deformation within the shear zone. One of the key parameters controlling dislocation creep, and thus the viscosity of the shear zone and afterslip, may be temperature. To investigate this, we run additional data assimilation experiments with a temperature dependence in the viscosity of the shear zone, in a similar fashion to the mantle wedge (Experiments 13 and 14, Table 1). The shear zone viscosity can be described using Equation 16 multiplied by the Arrhenius term $\exp\left(\frac{E_{\mathrm{SZ}}}{\mathrm{RT}}\right)$, where E_{SZ} is the activation energy in the shear zone. The state vector for data assimilation contains the mantle creep parameters (A_w, A_S, A_o, E_m) and the shear zone creep parameters ($A_{\mathrm{SZ}}, E_{\mathrm{SZ}}$ in Experiment 13 or $A_{\mathrm{SZ1}}, A_{\mathrm{SZ2}}, A_{\mathrm{SZ3}}, A_{\mathrm{SZ4}}, E_{\mathrm{SZ}}$ in Experiment 14).

The temperature varies significantly along the shear zone—from ~120°C in the shallowest part (40 km), ~150°C at 50 km depth, ~200°C at 50 km depth to ~1000°C in the deepest part (80 km) (Figure 2b). When we only model a single shear zone segment, we find a smooth transition from high viscosities in the shallow part of the shear zone to lower values in the deeper part (e.g., $5 \cdot 10^{18}$ to 10^{17} Pa·s postseismically, Figure 13b). The effective shear stress is much higher downdip of the coseismic rupture than at the deepest part of the shear zone. This shows that the temperature gradient within the shear zone primarily controls the effective viscosity gradient. Consequently, a uniform-like distribution of afterslip is estimated and thus more afterslip is estimated in the deeper part of the shear zone relative to the temperature-independent case (Figures 13e and 13f).

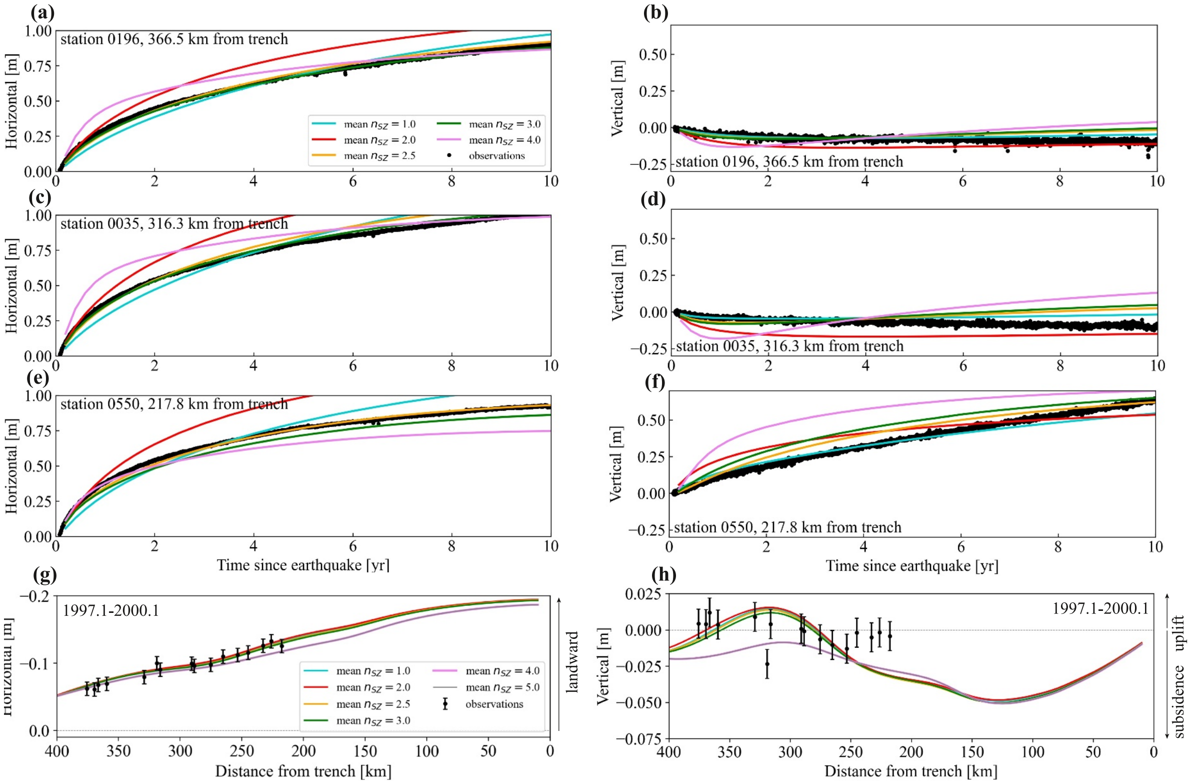


Figure 11. Impact of the shear zone stress power n_{SZ} on postseismic time series and interseismic displacements (Experiments 3, 4 and 9–12). The left column (a, c, e) shows the horizontal displacement times series at stations 0196, 0035 and 0550 (see Figures 1 and 4 for locations) and (g) shows the 3-year cumulative horizontal interseismic displacement. The right column (b, d, f) shows the vertical displacement time series at stations 0196, 0035 and 0550 and (f) show the 5-year cumulative vertical interseismic displacement. The ensemble uncertainties can be viewed in Figures S17 and S18 in Supporting Information S1.

If we divide the shear zone into four segments and thus estimate four different pre-exponent factors of the shear zone, we obtain very similar afterslip and shear zone viscosity estimates as the temperature-independent case (Figures 13c, 13d and 13g). When we model a temperature-dependent viscosity within the shear zone, we obtain very similar displacements as when we model a temperature-independent viscosity within the shear zone. However, by modeling a temperature-dependent viscosity, we slightly underpredict very early postseismic motion and cannot simulate enough postseismic backarc subsidence (Figure S19 in Supporting Information S1). This implies that the modeled viscosities in the shallowest part of the shear zone are likely overestimated. The transient deformation is not well represented by the model with only temperature dependence and could have been better captured by incorporating a Kelvin element in the rheology for example, Introducing more shear zone

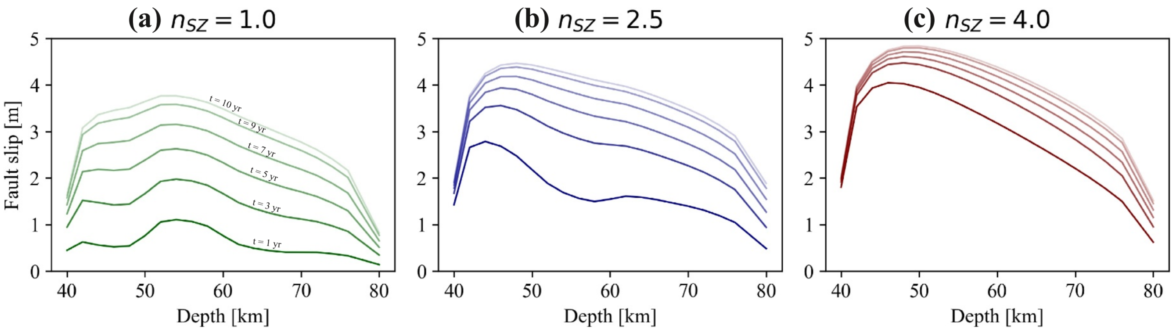


Figure 12. Impact of the shear zone stress power n_{SZ} on the postseismic slip at the subduction interface at different times after the earthquake resulting from data assimilation experiments 3, 4 and 11 (subpanels (a)–(c), respectively). Note that for each experiment, the results are computed with a single model with input mean parameters resulting from the ensemble.

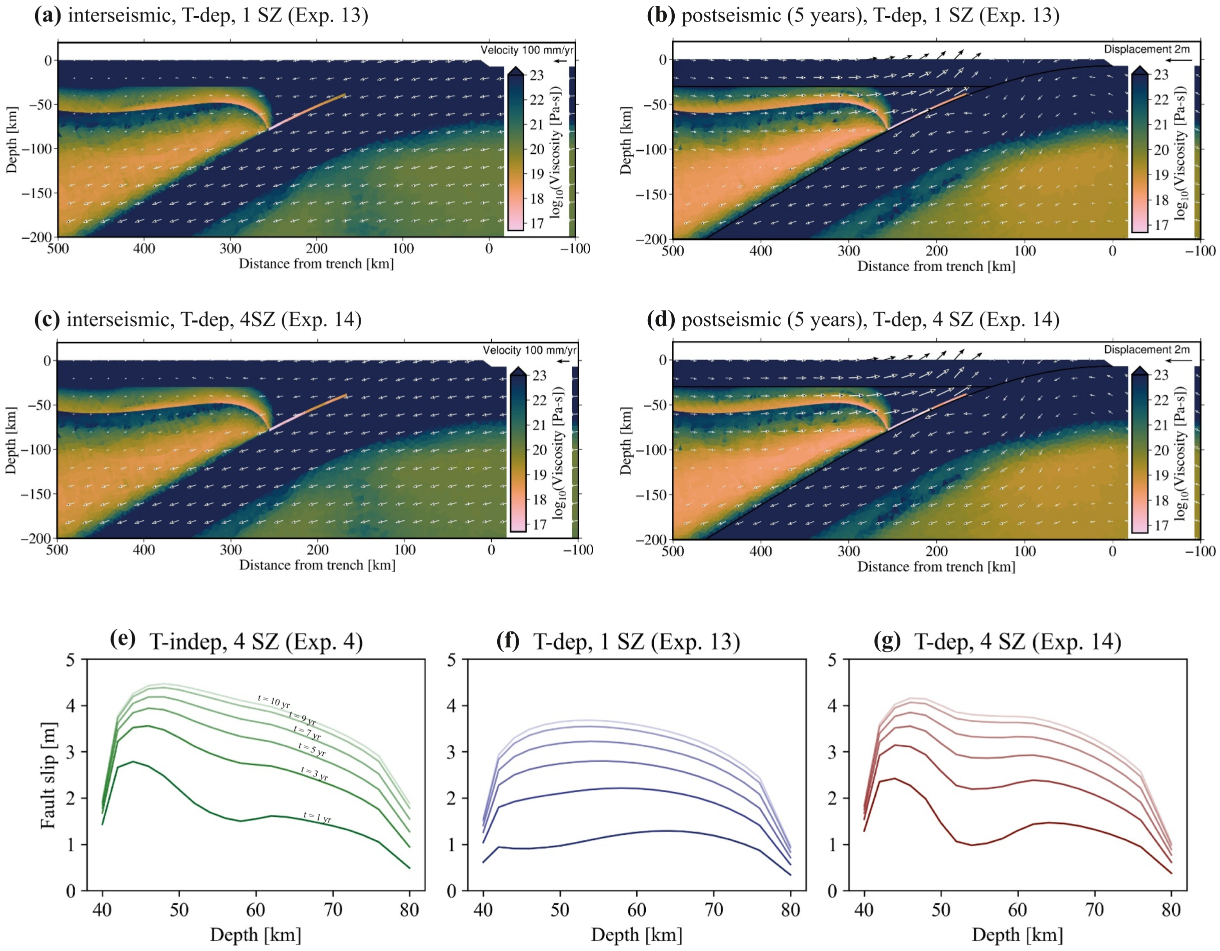
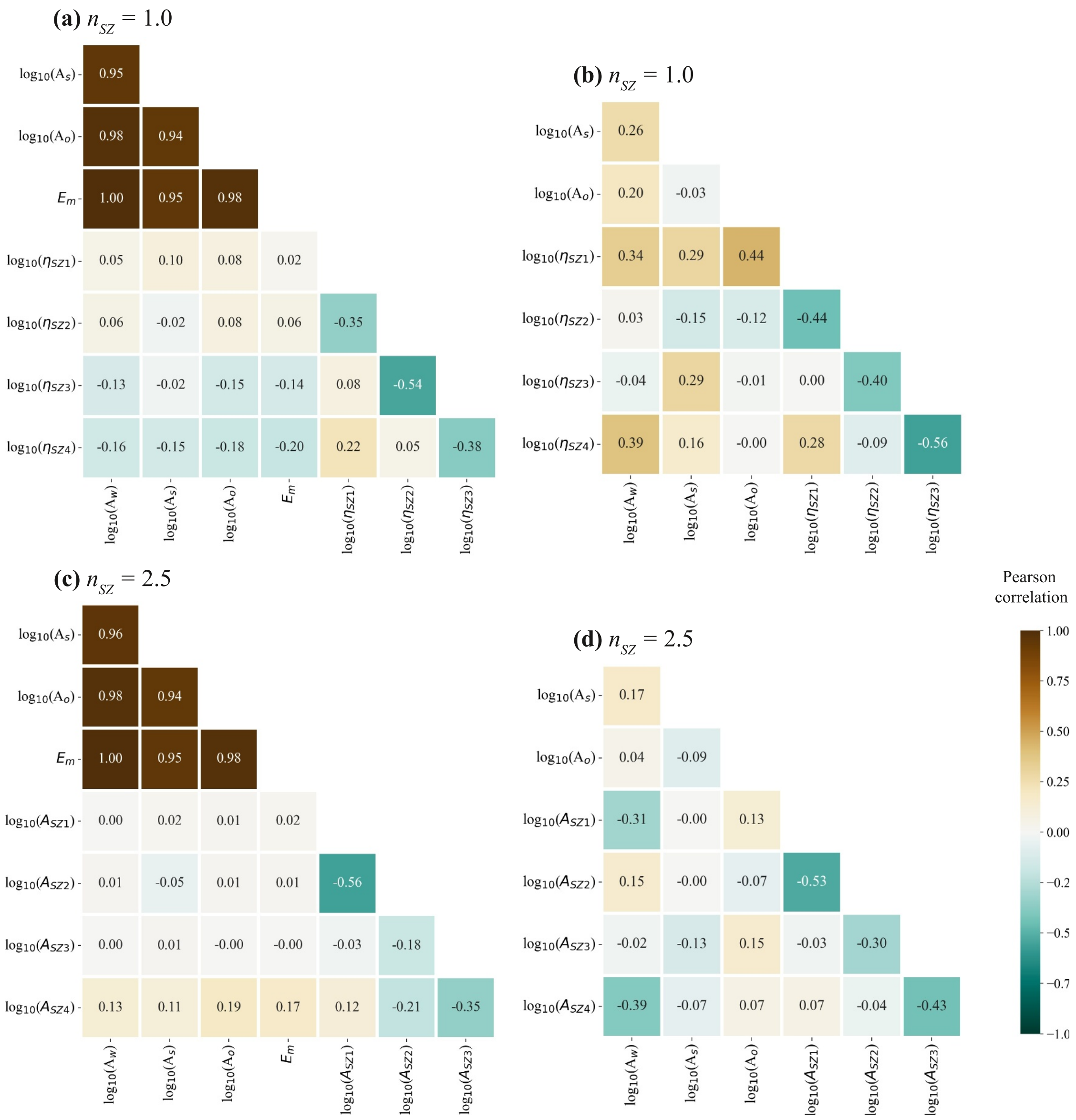


Figure 13. Impact of modeling a temperature-dependent viscosity for the shear zone. (a) Effective viscosity during the interseismic phase with a temperature-dependent viscosity in a single shear zone segment (Experiment 13). (b) Effective viscosity during the postseismic phase (5 years after the earthquake) with a temperature-dependent viscosity in a single shear zone segment (Experiment 13). (c) Effective viscosity during the interseismic phase with a temperature-dependent viscosity in four shear zone segments (Experiment 14). (d) Effective viscosity during the postseismic phase (5 years after the earthquake) with a temperature-dependent viscosity in four shear zone segments (Experiment 14). Afterslip distributions at select times when modeling (e) temperature-independent viscosity in four shear zone segments, (f) temperature-dependent viscosity in a single shear zone segment, and (g) temperature-dependent viscosity in four shear zone segments. Note that for each experiment, the results are computed with a single model with input mean parameters resulting from the ensemble.

segments naturally introduces more degrees of freedom into the estimation problem and a potential to overfit the data. We find that introducing these extra degrees of freedom is necessary as a temperature dependence of the viscosity alone cannot explain the data as well and it is physically more likely that a large portion of afterslip is found directly downdip of the coseismic rupture due to high stresses (e.g., Fukuda & Johnson, 2021; Muto et al., 2019). This implies that the segmentation in the pre-exponent factor represents some missing physics. We further discuss the implications of the segmentation and the shear zone width on our results in Section 5.1.

4.6. Trade-Off Between Afterslip and Viscoelastic Relaxation

As our models are nonlinear, we cannot simply disentangle contributions from afterslip and viscoelastic relaxation from the modeled surface deformation. We fix the stress powers of both the shear zone and mantle (n_{SZ} and n_m, respectively). Thus, afterslip is controlled by the pre-exponent factors in the shear zone A_{SZ1}, A_{SZ2}, A_{SZ3}, and A_{SZ4}, whereas viscoelastic relaxation is controlled by the creep parameters A_w, A_s, A_o, and E_m. These creep parameters determine the viscosity structure of the mantle and thus also influence interseismic deformation. To investigate if there is a trade-off between afterslip and viscoelastic relaxation, we focus on the correlation (ρ) between parameters representing afterslip and viscoelastic relaxation.



(a) $n_{SZ} = 1.0$

	$\log_{10}(A_w)$	$\log_{10}(A_s)$	$\log_{10}(A_o)$	E_m	$\log_{10}(\eta_{SZ1})$	$\log_{10}(\eta_{SZ2})$	$\log_{10}(\eta_{SZ3})$
$\log_{10}(A_s)$	0.95						
$\log_{10}(A_o)$	0.98	0.94					
E_m	1.00	0.95	0.98				
$\log_{10}(\eta_{SZ1})$	0.05	0.10	0.08	0.02			
$\log_{10}(\eta_{SZ2})$	0.06	-0.02	0.08	0.06	-0.35		
$\log_{10}(\eta_{SZ3})$	-0.13	-0.02	-0.15	-0.14	0.08	-0.54	
$\log_{10}(\eta_{SZ4})$	-0.16	-0.15	-0.18	-0.20	0.22	0.05	-0.38

(b) $n_{SZ} = 1.0$

	$\log_{10}(A_w)$	$\log_{10}(A_s)$	$\log_{10}(A_o)$	$\log_{10}(\eta_{SZ1})$	$\log_{10}(\eta_{SZ2})$	$\log_{10}(\eta_{SZ3})$
$\log_{10}(A_s)$	0.26					
$\log_{10}(A_o)$	0.20	-0.03				
$\log_{10}(\eta_{SZ1})$	0.34	0.29	0.44			
$\log_{10}(\eta_{SZ2})$	0.03	-0.15	-0.12	-0.44		
$\log_{10}(\eta_{SZ3})$	-0.04	0.29	-0.01	0.00	-0.40	
$\log_{10}(\eta_{SZ4})$	0.39	0.16	-0.00	0.28	-0.09	-0.56

(c) $n_{SZ} = 2.5$

	$\log_{10}(A_w)$	$\log_{10}(A_s)$	$\log_{10}(A_o)$	E_m	$\log_{10}(A_{SZ1})$	$\log_{10}(A_{SZ2})$	$\log_{10}(A_{SZ3})$
$\log_{10}(A_s)$	0.96						
$\log_{10}(A_o)$	0.98	0.94					
E_m	1.00	0.95	0.98				
$\log_{10}(A_{SZ1})$	0.00	0.02	0.01	0.02			
$\log_{10}(A_{SZ2})$	0.01	-0.05	0.01	0.01	-0.56		
$\log_{10}(A_{SZ3})$	0.00	0.01	-0.00	-0.00	-0.03	-0.18	
$\log_{10}(A_{SZ4})$	0.13	0.11	0.19	0.17	0.12	-0.21	-0.35

(d) $n_{SZ} = 2.5$

	$\log_{10}(A_w)$	$\log_{10}(A_s)$	$\log_{10}(A_o)$	$\log_{10}(A_{SZ1})$	$\log_{10}(A_{SZ2})$	$\log_{10}(A_{SZ3})$
$\log_{10}(A_s)$	0.17					
$\log_{10}(A_o)$	0.04	-0.09				
$\log_{10}(A_{SZ1})$	-0.31	-0.00	0.13			
$\log_{10}(A_{SZ2})$	0.15	-0.00	-0.07	-0.53		
$\log_{10}(A_{SZ3})$	-0.02	-0.13	0.15	-0.03	-0.30	
$\log_{10}(A_{SZ4})$	-0.39	-0.07	0.07	0.07	-0.04	-0.43

Figure 14. Correlations between posterior parameter pairs. In (a) and (b) we assume linear creep in the shear zone and set $n_{SZ} = 1.0$ (Experiments 3 and 7). In subplots (c) and (d) we assume power-law creep in the shear zone and set $n_{SZ} = 2.5$ (Experiments 4 and 6). In (b) and (d) we remove E_m from the state vector and fix it to 420 kJ/mol.

We find a very strong correlation (ρ between 0.95 and 1.00) between the pre-exponent factor and activation energy of the mantle when we model either linear or power-law creep in the shear zone (Figures 8, Figures 14a and 14c). Since both parameters scale the viscosity, a wide range of these parameters is acceptable as their combinations produce similar viscosity structures. This strong correlation hinders the individual estimation of the activation energy and water fugacity, which is partly represented by the pre-exponent factor. The wide range of possible values of the pre-exponent factor and activation energy of the mantle may overshadow potential trade-offs between other parameter pairs. Because there is a strong correlation between the pre-exponent factor and activation energy, the estimated parameters are more sparsely distributed. Meaningful correlations between other parameter pairs are not easily detectable. To alleviate this, we run additional ESMDA tests where we fix the mantle activation energy at 420 kJ/mol (Karato et al., 1986) and thereby reduce the number of estimated parameters. As we fix a parameter, correlation differences occur due to a change in the marginal relation between the remaining parameters. Consequently, some relations may become more evident with the reduction of the number of estimated parameters. Note that fixing the activation energy may lead to biased estimates for the remaining parameters as they need to adjust to compensate for the fixed value of the activation energy. By not fixing the activation energy, we can account for its inherent uncertainties. In the case where we assume power-law creep in the shear zone (Experiment 6), we do not find any significant correlations between the mantle creep parameters and the shear zone pre-exponent factors (Figure 14d). This suggests that afterslip and bulk viscoelastic relaxation can be independently resolved from the observations if the activation energy is narrowly constrained. In the case where we assume linear creep in the shear zone (Experiment 7), we find moderate correlations between the creep parameters and shear zone viscosity of the shallowest shear zone segment (ρ between 0.29 and 0.44, Figure S14 in Supporting Information S1, Figure 14b). This would suggest a relatively minor trade-off between afterslip and viscoelastic relaxation when assuming linear creep in the shear zone.

We run additional forward models to further investigate the resolvability between afterslip and viscoelastic relaxation. As Wang et al. (2012) pointed out, afterslip and bulk viscoelastic relaxation have a similar spatial footprint on postseismic horizontal deformation. Our forward model simulations show that decreasing the shear zone viscosity (by increasing the pre-exponent factor) results in more trenchward motion at distances greater than 100 km from the trench at 5 years after the earthquake (Figure S20 in Supporting Information S1), due to faster relaxation. Similarly, decreasing the viscosity in the continental and oceanic mantle (by increasing the pre-exponent factor) leads to more trenchward motion at distances ≥100 km from the trench at 5 years after the earthquake (Figure S21 in Supporting Information S1), also due to faster relaxation. These effects contribute strongly to a positive correlation between the shear zone and mantle creep parameters. Decreasing the shear zone and shallow mantle viscosity (by increasing the pre-exponent factor) and increasing the sub-slab mantle viscosity (by decreasing the pre-exponent factor) results in more uplift between approximately 150 and 300 km from the trench and more subsidence in the backarc at 5 years after the earthquake (Figures S20 and S21 in Supporting Information S1). This contributes to a negative correlation between the shear zone parameters and the oceanic mantle pre-exponent factor. Decreasing the mantle wedge viscosity (by increasing the pre-exponent factor) results in more onshore uplift at 5 years after the earthquake (Figure S21 in Supporting Information S1), whereas decreasing the shear zone viscosity leads to more subsidence beyond 250 km from the trench (Figure S20 in Supporting Information S1), contributing to some negative correlation between the pre-exponent factors of the mantle wedge and shear zone. In the power-law case, we observe that the correlation between the viscosity of the shear zone (particularly the shallowest part) and bulk pre-exponent factors drops with respect to the linear case (Figures 14b and 14d). With linear creep in the shear zone, similar spatial patterns can be achieved, but the temporal response in both horizontal and vertical deformation would particularly differ from the power-law shear zone models (Figures S23 to S26 in Supporting Information S1) as the power-law dictates the temporal response. Particularly for horizontal deformation, the power-law decay becomes important during the first 2 years after the earthquake, illustrated by Figure 11, because coseismic stresses cause lower viscosities.

Overall, the forward model tests show that the horizontal spatial signature of afterslip and viscoelastic relaxation are similar, but the vertical spatial signature—particularly in onshore displacements - can somewhat contribute to the resolvability. More importantly, the temporal signature, particularly of the first 2 years in horizontal deformation, provides the means to resolve afterslip from viscoelastic relaxation. We can successfully disentangle afterslip and viscoelastic relaxation from each other by modeling a power-law decay in the shear zone, given a power-law in the mantle.

5. Discussion

5.1. Implications of the Shear Zone Rheology

Our findings support a power-law behavior for the rheology within the shear zone, characterized by a stress power of 2.5 (Section 4.4). A linear Maxwell rheology in the shear zone cannot capture the fast changes in the data. This suggests that dislocation creep, opposed to diffusion creep, controls the governing deformation within the shear zone. Moreover, our results reveal that high stress powers, equal to or greater than 4.0, are unfavorable. This implies that a very rapid response due to fast relaxation is unlikely below the rupture zone (i.e., at 40 km depth).

The underlying perspective in our approach to simulate (re)locking and afterslip shows some similarities with that of more realistic friction formulations, such as rate-and-state friction (Dieterich, 1979; Ruina, 1983), as suggested by D'Acquisto and Govers (2023). In the rate-and-state formulation, the asperity is treated as rate-weakening portion of the megathrust, exhibiting rapid unstable sliding during the earthquake, whereas we simulate unstable behavior in a kinematic fashion. Downdip of the coseismic rupture, we simulate aseismic slip facilitated by a viscous shear zone, which shows some similarities with rate-strengthening behavior in the sense that both approaches promote stable sliding. Both approaches describe the evolution of afterslip and it has been shown that the postseismic decay of frictional approaches is very similar to that of a power-law approach (Montési, 2004). Both approaches simulate brittle and ductile behavior along the megathrust. High power-law exponents are compatible with brittle behavior, whereas lower power-law exponents are compatible with ductile behavior. Low power-law exponents found in the ductile shear zone are thus not unexpected and are compatible with laboratory test results (Barbot & Zhang, 2023). Previous studies that use rate-and-state friction laws to model afterslip following the 2011 Tohoku-Oki earthquake, typically find a transition from velocity weakening to velocity strengthening at depths between 30 and 50 km (Agata et al., 2019; Dhar et al., 2022; Fukuda & Johnson, 2021; Muto et al., 2019). The differences in estimated afterslip between this study and previous studies is further discussed in Section 5.2.

As mentioned earlier, the temperature within the shear zone varies drastically, ranging from ~120°C in the shallowest part (40 km), through ~150°C at 50 km depth, ~200°C at 50 km depth, to ~1,000°C in the deepest part (80 km) (Figure 2b). As our results suggest that dislocation creep governs deformation over diffusion creep, this could imply that dislocation creep occurs in both low-temperature and high-temperature regimes. Low-temperature dislocation creep could still occur under considerable differential stress and may be further enhanced by the presence of water—conditions that are typically found downdip of the coseismic rupture. At first glance, our results favor a temperature-independent viscosity of the shear zone over a temperature-dependent viscosity. To reproduce similar results to the temperature-independent case, it is necessary to introduce multiple segments within the shear zone for the temperature-dependent case. The resulting effective viscosity profiles within the shear zone do not display a simple correlation with increasing depth or temperature given the geometry of the shear zone, as illustrated in Figure 13. Another factor to consider is that there is a scaling relation between the shear zone width and viscosity (e.g., Govers et al., 2018), that is, the same amount of afterslip can be reproduced with a larger shear zone width if the viscosity increases accordingly. For example, the pre-exponent factor A_{SZ} (Equation 16) would be 10 times larger for a shear zone width of 300 m (i.e., 10 times smaller than in our models), and the viscosities would be 10 times lower. In our models we assume a constant thickness of the shear zone, although it is actually expected to broaden with depth/temperature (Scholz, 1988; Sibson, 1983). A correlation between shear zone viscosity and temperature could thus hold, taking into account thickness variations, and confirm that the rheology of the shear zone is thermally controlled. The segmentation needed in the shear zone is thus required to represent changes in the shear zone width when we consider temperature-dependent viscosities. On the other hand, in our simulations with a temperature-independent viscosity, the segmentation of the shear zone represents changes in the shear zone width as well as changes in temperature captured by the pre-exponent factor.

Our model setup of the megathrust provides a first-order representation of the seismogenic zone and a downdip shear zone with variable properties. Our results demonstrate that we can individually resolve the properties of four the four shear zone segments of uniform length, as well as the bulk mantle properties. The varying properties of the shear zone capture factors such as temperature, lithology, frictional parameters, and shear zone geometry. Future work will focus on exploring additional scenarios, such as segments of variable length or increased compartmentalization, to investigate potential connections between rheological properties to the metamorphic and lithological stratification of the megathrust.

5.2. Finding a Unique Solution for Afterslip and Viscoelastic Relaxation

As our model is 2-D, afterslip inherently only takes place downdip of the coseismic rupture. Our best-fit results ($n_{SZ} = 2.5$) suggest that the afterslip is concentrated near the downdip end of the seismogenic zone (40–45 km depth) and gradually decays downdip, which is compatible with 2-D (Muto et al., 2019) and 3-D studies (Agata et al., 2019; Dhar et al., 2022; Fukuda & Johnson, 2021; Hu et al., 2016) that model stress-dependent afterslip.

However, differences in modeling approaches (e.g., coseismic slip pattern, afterslip modeling, rheology) and constraints from geodetic data lead to somewhat different afterslip magnitudes and locations, as is shown in Figure 15. Particularly, we find less afterslip is necessary than suggested by Fukuda and Johnson (2021) to match the observations. We require more afterslip than inferred by Muto et al. (2019), as we also include the first year as

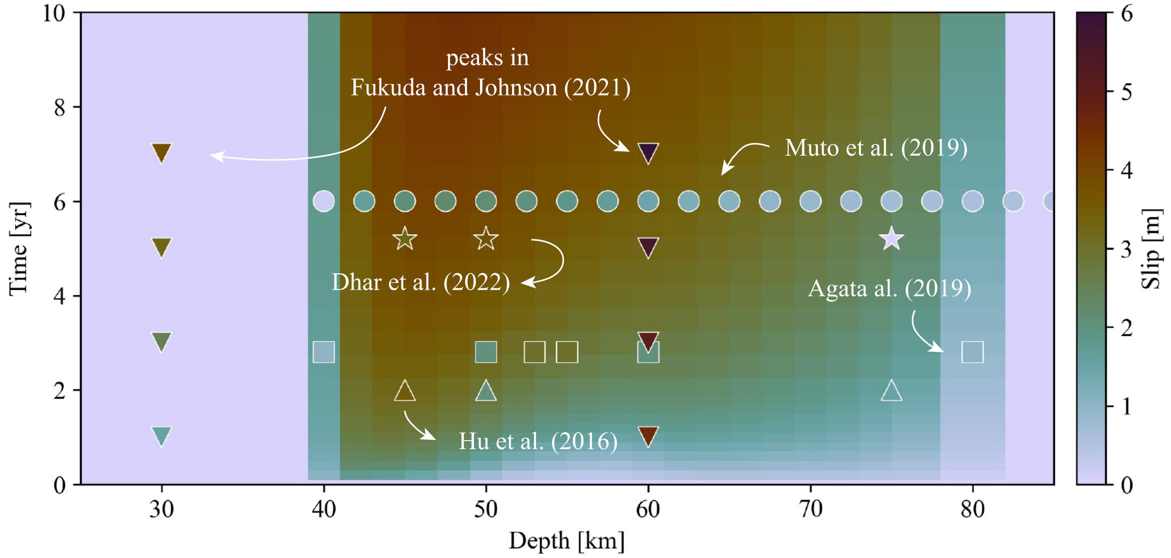


Figure 15. Spatiotemporal distribution of afterslip for our best-fit model (calculated with input mean parameter values from Experiment 4). Estimates of afterslip from previous studies with stress-driven afterslip are included.

a geodetic constraint. Our estimates align with those of Dhar et al. (2022) and Hu et al. (2016), yet we find more shallow afterslip compared to Agata et al. (2019).

Another factor to consider is that our models are 2-D and use a smoothed coseismic slip pattern based on the 3-D slip pattern by Hooper et al. (2013). Afterslip patterns are particularly sensitive to coseismic stress changes, which can explain some differences between our modeling results in terms of afterslip and viscoelastic relaxation. Additionally, in our 2-D models we do not allow an overlap between regions of afterslip and the coseismic rupture. Some overlap has been suggested as a considerable moment release in the form of aftershocks was recorded within the rupture zone near the downdip end (Dielforder et al., 2023). However, our simplified approach does capture most of the moment release that is concentrated at the downdip end of the coseismic rupture. Our modeling approach is compatible with that of the classical asperity model, where afterslip and coseismic slip occur in spatially distinct zones: coseismic slip in velocity-weakening regions and afterslip in velocity-strengthening regions and diminishes in velocity-strengthening regions (Avouac, 2015). Some studies following the 2011 Tohoku-Oki earthquake infer that regions of afterslip and coseismic slip are spatially distinct (Evans & Meade, 2012; Iinuma et al., 2016), but others report a small overlap (Freed et al., 2017; Johnson et al., 2012; Ozawa et al., 2012; Perfettini & Avouac, 2014).

To investigate whether an overlap between afterslip and coseismic slip would affect our conclusions, we ran an additional data assimilation experiment (Experiment 27, Table 1). In this experiment, we extend the shear zone to 30 km depth so that it overlaps with the coseismic segment of the megathrust. We separately solve for the pre-exponent of this new segment between 30 and 40 km depth. We find surface displacements that are (visually) indistinguishable from Experiment 4 (Figure 3, Figures S27a–S27d in Supporting Information S1). The resulting viscosity of the shallowest shear zone is high (10^{21} Pa·s) in comparison to the deeper shear zones ($\sim 10^{17}$ Pa·s) (Figure S27e in Supporting Information S1). As a result, afterslip is minimal in the 30–40 km depth region and the peak in afterslip is still found directly downdip of the coseismic rupture (Figure S27f in Supporting Information S1). We thus conclude that (deep) afterslip occurs downdip of the coseismic rupture.

Rheological choices may also influence our results. Unlike previous postseismic studies of the 2011 Tohoku-Oki earthquake that employ a Burgers rheology incorporating power-law component in both Maxwell and Kelvin elements (Agata et al., 2019; Barbot, 2020; Dhar et al., 2022, 2023; Muto et al., 2019), our model uses only a power-law Maxwell rheology. The stress dependence in our viscosity already captures most of the transient deformation without the need for additional complexity, such as a Kelvin element. Note that we exclude the deformation within the first month following the earthquake, a period that likely experienced significant transient deformation due to viscoelastic relaxation and other processes like poroelastic rebound and brittle creep. Our results do not indicate the need for a transient power-law component. It is worth noting that trade-offs between model parameters may lead to alternative viable solutions.

Three-dimensional studies applying kinematic afterslip (Freed et al., 2017; Iinuma et al., 2016; Sun et al., 2014; Wang et al., 2018; Yamagiwa et al., 2015) have identified patches of afterslip near the northern and southern edges of the coseismic rupture zone and estimate minimal afterslip downdip of the coseismic rupture. The non-uniqueness of published solutions for afterslip and viscoelastic relaxation may originate from trade-offs between contributions from afterslip and viscoelastic relaxation to surface displacements (Fukuda & Johnson, 2021). It has been suggested that the spatial footprint of afterslip and viscoelastic relaxation (Freed et al., 2017; Fukuda & Johnson, 2021), as well as a modeling approach with mechanical coupling between afterslip and viscoelastic relaxation, provide the means to find a unique solution (Fukuda & Johnson, 2021). In this study, we delve deeper into the interplay between afterslip and viscoelastic relaxation by examining the specific parameters that govern them.

As described in Section 4.6, we quantify trade-offs between the shear zone pre-exponent parameters, A_{SZ1}, A_{SZ2}, A_{SZ3}, and A_{SZ4}, and the mantle creep parameters, A_w, A_s, A_o, and E_m, which control afterslip and viscoelastic relaxation of the mantle, respectively. The spatial distribution of horizontal postseismic deformation contributes to the non-uniqueness of the solution, whereas vertical deformation and interseismic deformation offer constraints (Figures S20 and S21 in Supporting Information S1). Our results show that when we assume power-law creep in the shear zone, as opposed to linear creep, we obtain weak correlations between these parameters and thus can distinguish viscoelastic relaxation of the mantle and afterslip as opposed to a linear rheology (Figure 14). With power-law creep in the shear zone, relaxation in the shear zone occurs faster than with linear creep. Because of this difference in the relaxation times, we obtain a closer match to the data, particularly for early postseismic horizontal deformation, and we can separately determine the parameters controlling afterslip and viscoelastic relaxation. Not only does the temporal signature change with power-law creep due to faster viscoelastic relaxation, a different solution in the spatial distribution of afterslip is preferred. Specifically, more afterslip is obtained directly downdip of the coseismic rupture (Figure 12). Our results suggest that the spatial footprint of vertical observations can help to distinguish bulk viscoelastic relaxation from afterslip. In addition to the spatial footprint of the vertical data, the temporal response of horizontal data, especially in the first 2 years after the earthquake, provide the means to disentangle the two processes from each other, which is successfully captured by modeling a power-law in the shear zone (rather than linear creep).

5.3. Low Viscosities in the Shallow Mantle Wedge

Our results indicate that low viscosities ($\sim 10^{18}$ Pa·s within 10 years after the earthquake) in the continental lithospheric mantle (30–60 km depth) are required, particularly to achieve a good match with vertical postseismic backarc (>300 km from the trench) and vertical interseismic observations. This reduces the effective thickness of the overriding plate and thus the flexural rigidity, allowing the plate to bend more easily during both interseismic and postseismic phases, and enabling shallow viscoelastic relaxation to occur postseismically. Without the division in the mantle wedge, more afterslip is required to match the postseismic observations, but the match with observations deteriorates. Hence there is a risk of overestimating afterslip when shallow viscoelastic relaxation is not possible.

An artifact of the input temperature model (Morishige, 2022) is the "viscous belly" seen at distances greater than 400 km from the trench (Figure 2b). This overthickening of the overriding lithosphere is caused by cooling of the lithosphere that takes place over millions of years, which is longer than the age of the subduction zone (Hall, 2012; Wilson & van Keken, 2023). As relatively low temperatures are present in the shallow mantle wedge (approximately 400–800°C between 30 and 60 km depth, Figure 2b), high viscosities at a relatively shallow depth would be present in our models without the division in the mantle wedge. Our results underline the necessity of a division in the mantle wedge to simulate a different rheology in the shallow part of the mantle wedge compared to the deeper part. This way, low shallow viscosities are simulated to match interseismic and postseismic observations. Without the artificial overthickening of the continental lithospheric mantle, shallow temperatures would likely be higher, leading to lower viscosities, and possibly making the division in the mantle wedge unnecessary. However, we observe large strain rates in the shallow mantle within the forearc (within 300 km from the trench) in our model, indicating that most relaxation takes place within the forearc. Due to the viscous belly, viscosities past 400 km from the trench are approximately 1 order higher than within 300 km from the trench at ~50 km depth during the postseismic period (Figure 5e). The contribution of viscoelastic relaxation below the backarc to surface displacements is significantly less given the strain rates, which may also be an indirect effect from the viscous belly.

Another reason for the need of low viscosities in the mantle wedge in our models could be due to the presence of structural heterogeneities, such as the volcanic arc and/or compliance in the upper plate. Inferring lower viscosities around the volcanic arc (at approximately 300 km) (Hu et al., 2014; Muto et al., 2016) or introducing a compliance contrast in the overriding plate in the vicinity of the volcanic arc (Itoh et al., 2019, 2021) have been shown to lead to a better fit with the data. Note that we also include a compliance contrast in the overriding plate at 425 km from the trench, but likely still require a division in the mantle wedge as this structural heterogeneity is introduced far away from the volcanic arc. Contrarily, Freed et al. (2017) argue that vertical stratification of the viscosity structure offers a satisfactory explanation of the data and that lateral heterogeneities are not needed. Additionally, a horizontal distribution of lower viscosities in the forearc can be explained by water transport due to chemical reactions at the volcanic arc and corner flow (Iwamori, 2007). Our model results are in line with Freed et al. (2017) considering that horizontal layering of mantle wedge rheology is needed, but the specific details vary due to different modeling strategies. However, vertical stratification at the volcanic arc or horizontal stratification in the forearc are also plausible solutions that may lead to similar surface displacements.

5.4. Two Causes of Rapid Postseismic Relaxation in the Sub-Slab Asthenosphere

Our models accurately replicate the observed landward motion of the overriding plate within 100 km of the trench during the early postseismic period (Figure 6c), with rates faster than the observed interseismic velocity (Watanabe et al., 2014). These motions are primarily driven by viscoelastic relaxation in the sub-slab asthenosphere (Suito (2017), Figure S21 in Supporting Information S1), while relocking of the megathrust is required to transfer the imprint of this deep relaxation to the overriding plate (D'Acquisto & Govers, 2023). Sun et al. (2014) were the first to reproduce these observations by introducing a separate layer in their model beneath the oceanic plate with a very low viscosity of 2.5×10^{17} Pa·s, although they do not explicitly mention that this layer enhances near-trench landward motion. Subsequent postseismic modeling studies also required low sub-slab viscosities ranging between 2.5×10^{17} Pa·s and 3.0×10^{18} Pa·s (Freed et al., 2017; Iinuma et al., 2016; Suito, 2017; Sun & Wang, 2015). Additionally, Sun et al. (2024) inferred sub-slab viscosities as low as 5.0×10^{16} Pa·s to explain enhanced landward motion adjacent the coseismic rupture zone. Studies deploying a power-law rheology found viscosities in the sub-slab asthenosphere from as low as 10^{16-17} Pa·s within a few days after the earthquake to 10^{17-18} Pa·s 1 year after (Agata et al., 2019; Dhar et al., 2023; Muto et al., 2019). The viscosities in these power-law models evolve because of relaxation of coseismic stresses.

We also find a temporary zone with low sub-slab asthenospheric viscosities after the earthquake near the trench below the oceanic plate (Figure 5e). However, our estimated viscosity is significantly larger than previously found; 7.0×10^{18} Pa·s right after the earthquake, and 8.6×10^{18} Pa·s 1 year later (Figure 5a). The reason for this difference is that our models have an additional driver of near-trench landward motions resulting from our earthquake cycle approach. In our model, subduction is driven by the far-field velocity of the oceanic plate and by the velocity of the deep end of the slab. Following an earthquake, the megathrust is immediately and fully locked. During the interseismic period, the oceanic plate moves landward at full plate tectonic velocity in the far-field, and much slower near the trench where it is locked to the overriding plate. The oceanic plate is thus shortened interseismically. Unlocking resolves the slip deficit on the part of the megathrust that was previously locked. All other parts of the model respond elastically to this instantaneous slip because viscoelastic domains cannot relax accumulated stresses on the coseismic timescale. Therefore, the oceanic plate can hardly move relative to the sub-slab asthenosphere so that it cannot immediately recover from interseismic shortening during the megathrust earthquake. This occurs during the early postseismic phase when the elastic oceanic plate catches up by accelerated landward motion relative to the sub-slab asthenosphere. This additional driver of postseismic deformation is absent in models driven by coseismic stresses. Motions throughout the oceanic plate and slab are partly driven by the velocity boundary conditions imposed at the edges of the incoming oceanic plate and the outgoing slab, reflecting a constant ridge push and slab pull. We observe that interseismic strain accumulation is limited to approximately 200 km seaward of the trench. This occurs at about 200 km from the edges of the incoming oceanic plate, which suggests that our boundary conditions do not impact our findings regarding motion near the trench.

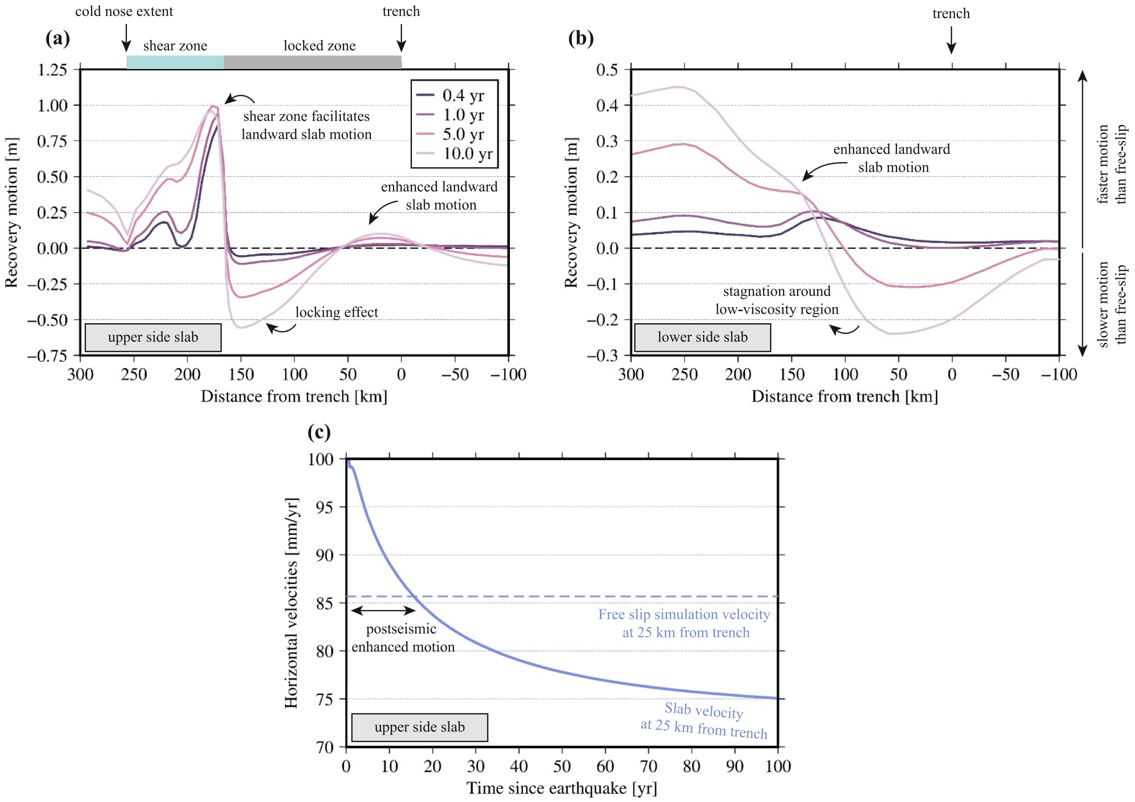


Figure 16. Postseismic recovery motion along the upper (a) and lower (b) sides of the slab. Recovery motion is defined as the horizontal motion since the earthquake subtracted by the horizontal motion of another model with only free-slip (no locking, no earthquakes). (c) Horizontal velocity along the upper side of the slab as a function of time.

Figure 16 illustrates what we refer to as "recovery motion" in our model. This term describes the horizontal motion of the oceanic plate relative to the horizontal motion from another simulation where the megathrust interface slips freely at all times. Positive recovery motion indicates that the oceanic plate is moving faster than in the free-slip model, while negative recovery motion means that it is moving slower. We observe positive recovery motion of the shallow oceanic plate between −20 and 60 km from the trench (Figure 16a). The horizontal velocity of the top of the oceanic plate at horizontal coordinate −25 km is substantially higher than the free-slip velocity during the first ~17 years after the earthquake (Figure 16c). This is a consequence of the horizontal un-shortening of the oceanic plate. The base of the oceanic plate also shows enhanced landward motion during the first year after the earthquake (Figure 16b). The base of the oceanic plate between −20 and 60 km from the trench recovers much more quickly from interseismic shortening than the top.

6. Conclusions

Our two-dimensional earthquake cycle model of the Japan Trench subduction zone megathrust rupture highlights the importance of using a self-consistent model, in terms of slip deficit accumulation and release, to simulate interseismic and postseismic deformation. We investigate the rheological makeup of the shear zone and mantle to consistently explain the surface motion observed before and after the 2011 Tohoku-Oki earthquake. We use data assimilation, in the form of an ensemble smoother with multiple data assimilation, to find a single set of parameters to explain both sets of observations. We test several combinations involving a linear Maxwell rheology and a nonlinear power-law rheology in both the shear zone and mantle. Nonlinear dislocation creep with a stress power of 3.0 within the mantle is favored over linear creep to match the postseismic transient in the observations. Both interseismic and postseismic observations suggest that shallow viscoelastic relaxation (30–60 km depth) of the continental mantle is required to allow sufficient plate bending. Additionally, our estimated rheological parameters of the mantle wedge and sub-slab asthenosphere, such as the activation energy and pre-exponent factor, align with those derived from laboratory tests. Our results show that (deep) afterslip occurs downdip of the coseismic rupture. Nonlinear creep with a stress power of 2.5 within the shear zone is preferred over linear creep to accommodate fast changes within the very early postseismic phase (within 3 years after the earthquake). A segmentation of the shear zone is required to account for viscosity changes along the shear zone. Our inferred shear zone rheology implies that shear zone creep is thermally controlled—assuming that the shear zone broadens with increasing depth and temperature.

We demonstrate that the type of rheological law used affects the uniqueness of the solution between bulk relaxation of the mantle and afterslip on the megathrust. A linear Maxwell rheology of the shear zone coupled with a nonlinear power-law rheology of the mantle reveals trade-offs between parameters controlling afterslip and bulk relaxation. If we impose a nonlinear rheology for both the shear zone and mantle, we can find a unique geodetic footprint due to the spatiotemporal signature of both afterslip and mantle relaxation.

Postseismically, viscosities in the sub-slab asthenosphere drop about 1 order of magnitude with respect to the steady-state viscosity before the earthquake. We put forward a combination of two processes to explain rapid postseismic relaxation near the trench. First, enhanced landward motion is partly driven by viscous flow of the sub-slab asthenosphere, which was previously suggested as the sole driver. The second process is captured by our simulation of repeated earthquake cycles. The incoming oceanic plate and outgoing slab are continuously pushed and pulled by velocity boundary conditions we kinematically apply at the two edges. Elastic shortening stresses accumulated during the interseismic phase are only partly released coseismically. The remaining elastic stress is released postseismically by enhanced landward motion of the oceanic plate, which is expressed in the form of afterslip along the deeper upper part of the slab as well as the base of the slab. Due to the latter effect, the sub-slab landward viscous flow is less pronounced to explain the observations and the postseismic sub-slab viscosities are larger than previously suggested (over 7.0×10^{18} Pa·s) in our cyclic models. These results demonstrate the importance of incorporating earthquake cycles using a consistent formulation to better constrain the processes driving surface deformation during the interseismic and postseismic phases.

Data Availability Statement

Input and output files that were used for the models of this paper are digitally stored in the Yoda repository of Utrecht University and are accessible in compliance with FAIR (Findable, Accessible, Interoperable, Reusable) principles available at Marsman et al. (2024). The finite element modeling package, GTECTON, is available at Govers and van de Wiel (2024). The GNSS displacements used in this study are available at https://www.gsi.go.jp/ENGLISH/geonet_english.html, subject to the policies of the Geospatial Information Authority of Japan. Figures are made using Generic Mapping Tools (Wessel et al., 2019), Adobe Illustrator, and Python 3.8.

Acknowledgments
The authors thank the editor Satoshi Ide and the associate editor for handling the manuscript. The authors thank two anonymous reviewers for their insightful comments and constructive feedback, which have improved this paper. This work was funded by the Dutch Research Council NWO grant ALWGO.2019.001.

References

Agata, R., Barbot, S. D., Fujita, K., Hyodo, M., Iinuma, T., Nakata, R., et al. (2019). Rapid mantle flow with power-law creep explains deformation after the 2011 Tohoku mega-quake. *Nature Communications*, *10*(1), 1385. https://doi.org/10.1038/s41467-019-08984-7

Altamimi, Z., Collilieux, X., & Métivier, L. (2011). ITRF2008: An improved solution of the international terrestrial reference frame. *Journal of Geodesy*, *85*(8), 457–473. https://doi.org/10.1007/s00190-011-0444-4

Avouac, J.-P. (2015). From geodetic imaging of seismic and aseismic fault slip to dynamic modeling of the seismic cycle. *Annual Review of Earth and Planetary Sciences*, *43*(1), 233–271. https://doi.org/10.1146/annurev-earth-060614-105302

Barbot, S. (2020). Frictional and structural controls of seismic super-cycles at the Japan trench. *Earth Planets and Space*, *72*(1), 63. https://doi.org/10.1186/s40623-020-01185-3

Barbot, S. (2023). Constitutive behavior of rocks during the seismic cycle. *AGU Advances*, *4*(5), e2023AV000972. https://doi.org/10.1029/2023AV000972

Barbot, S., & Zhang, L. (2023). Constitutive behavior of olivine gouge across the brittle-ductile transition. *Geophysical Research Letters*, *50*(24). https://doi.org/10.1029/2023gl105916

Bürgmann, R., & Dresen, G. (2008). Rheology of the lower crust and upper mantle: Evidence from rock mechanics, geodesy, and field observations. *Annual Review of Earth and Planetary Sciences*, *36*(1), 531–567. https://doi.org/10.1146/annurev.earth.36.031207.124326

Cheng, G., Barnhart, W. D., & Li, S. (2022). Power-law viscoelastic flow of the lower accretionary prism in the makran subduction zone following the 2013 baluchistan earthquake. *Journal of Geophysical Research: Solid Earth*, *127*(11), e2022JB024493. https://doi.org/10.1029/2022JB024493

D'Acquisto, M., Broerse, T., Marsman, C. P., & Govers, R. (2023). Reconciling the conflicting extent of overriding plate deformation before and during megathrust earthquakes in South America, Sunda and northeast Japan. *Geophysical Journal International*, *235*(1), 879–908. https://doi.org/10.1093/gji/ggad262

D'Acquisto, M., & Govers, R. (2023). Offshore landward motion shortly after a subduction earthquake implies rapid relocking of the shallow megathrust. *Geophysical Research Letters*, *50*(1), e2022GL101638. https://doi.org/10.1029/2022GL101638

DeMets, C., Gordon, R. G., & Argus, D. F. (2010). Geologically current plate motions. *Geophysical Journal International*, *181*(1), 1–80. https://doi.org/10.1111/j.1365-246X.2009.04491.x

Dhar, S., Muto, J., Ito, Y., Miura, S., Moore, J. D. P., Ohta, Y., & Iinuma, T. (2022). Along-arc heterogeneous rheology inferred from post-seismic deformation of the 2011 Tohoku-Oki earthquake. *Geophysical Journal International*, *230*(1), 202–215. https://doi.org/10.1093/gji/ggac063

Dhar, S., Muto, J., Ohta, Y., & Iinuma, T. (2023). Heterogeneous rheology of Japan subduction zone revealed by postseismic deformation of the 2011 Tohoku-Oki earthquake. *Progress in Earth and Planetary Science*, *10*(1), 9. https://doi.org/10.1186/s40645-023-00539-1

Dielforder, A., Bocchini, G. M., Kemna, K., Hampel, A., Harrington, R. M., & Oncken, O. (2023). Megathrust stress drop as trigger of aftershock seismicity: Insights from the 2011 Tohoku earthquake, Japan. *Geophysical Research Letters*, *50*(3), e2022GL101320. https://doi.org/10.1029/2022GL101320

Dieterich, J. H. (1979). Modeling of rock friction: 1. Experimental results and constitutive equations. *Journal of Geophysical Research*, *84*(B5), 2161–2168. https://doi.org/10.1029/JB084iB05p02161

Emerick, A. A., & Reynolds, A. C. (2013). Ensemble smoother with multiple data assimilation. *Computers and Geosciences*, *55*, 3–15. https://doi.org/10.1016/j.cageo.2012.03.011

Evans, E. L., & Meade, B. J. (2012). Geodetic imaging of coseismic slip and postseismic afterslip: Sparsity promoting methods applied to the great Tohoku earthquake. *Geophysical Research Letters*, *39*(11), L11314. https://doi.org/10.1029/2012GL051990

Evensen, G. (2018). Analysis of iterative ensemble smoothers for solving inverse problems. *Computational Geosciences*, *22*(3), 885–908. https://doi.org/10.1007/s10596-018-9731-y

Evensen, G. (2019). Accounting for model errors in iterative ensemble smoothers. *Computational Geosciences*, *23*(4), 761–775. https://doi.org/10.1007/s10596-019-9819-z

Evensen, G., Vossepoel, F. C., & van Leeuwen, P. J. (2022). *Data Assimilation Fundamentals* (1 ed.). Springer. https://doi.org/10.1007/978-3-030-96709-3

Freed, A. M., & Bürgmann, R. (2004). Evidence of power-law flow in the Mojave desert mantle. *Nature*, *430*(6999), 548–551. https://doi.org/10.1038/nature02784

Freed, A. M., Hashima, A., Becker, T. W., Okaya, D. A., Sato, H., & Hatanaka, Y. (2017). Resolving depth-dependent subduction zone viscosity and afterslip from postseismic displacements following the 2011 Tohoku-Oki, Japan earthquake. *Earth and Planetary Science Letters*, *459*, 279–290. https://doi.org/10.1016/j.epsl.2016.11.040

Fujiwara, S., Tobita, M., & Ozawa, S. (2022). Spatiotemporal functional modeling of postseismic deformations after the 2011 Tohoku-Oki earthquake. *Earth Planets and Space*, *74*(1), 13. https://doi.org/10.1186/s40623-021-01568-0

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

Govers, R., Furlong, K. P., van de Wiel, L., Herman, M. W., & Broerse, T. (2018). The geodetic signature of the earthquake cycle at subduction zones: Model constraints on the deep processes. *Reviews of Geophysics*, *56*(1), 6–49. https://doi.org/10.1002/2017RG000586

Govers, R., & van de Wiel, L. (2024). Initial public release of GTecton platform (Version 2024.05) [Software]. *Initial public release of GTecton platform (Version 2024.05)*. https://doi.org/10.5281/zenodo.11471903

Govers, R., & Wortel, M. J. R. (1993). Initiation of asymmetric extension in continental lithosphere. *Tectonophysics*, *223*(1), 75–96. https://doi.org/10.1016/0040-1951(93)90159-H

Govers, R., & Wortel, M. J. R. (1999). Some remarks on the relation between vertical motions of the lithosphere during extension and the necking depth parameter inferred from kinematic modeling studies. *Journal of Geophysical Research*, *104*(B10), 23245–23253. https://doi.org/10.1029/1999JB900201

Hall, P. S. (2012). On the thermal evolution of the mantle wedge at subduction zones. *Physics of the Earth and Planetary Interiors*, *198–199*, 9–27. https://doi.org/10.1016/j.pepi.2012.03.004

Hashima, A., Becker, T. W., Freed, A. M., Sato, H., & Okaya, D. A. (2016). Coseismic deformation due to the 2011 Tohoku-Oki earthquake: Influence of 3-D elastic structure around Japan. *Earth Planets and Space*, *68*(1), 159. https://doi.org/10.1186/s40623-016-0535-9

Hashima, A., & Sato, T. (2017). A megathrust earthquake cycle model for northeast Japan: Bridging the mismatch between geological uplift and geodetic subsidence. *Earth Planets and Space*, *69*(1), 23. https://doi.org/10.1186/s40623-017-0606-6

Hayes, G. P., Moore, G. L., Portner, D. E., Hearne, M., Flamme, H., Furtney, M., & Smoczyk, G. M. (2018). Slab2, a comprehensive subduction zone geometry model. *Science*, *362*(6410), 58–61. https://doi.org/10.1126/science.aat4723

Hearn, E. H., Bürgmann, R., & Reilinger, R. E. (2002). Dynamics of İzmit earthquake postseismic deformation and loading of the Düzce earthquake Hypocenter. *Bulletin of the Seismological Society of America*, *92*(1), 172–193. https://doi.org/10.1785/0120000832

Hirahara, K. (2002). Interplate earthquake fault slip during periodic earthquake cycles in a viscoelastic medium at a subduction zone. *Pure and Applied Geophysics*, *159*(10), 2201–2220. https://doi.org/10.1007/s00024-002-8731-x

Hirth, G., & Kohlstedt, D. (2003). Rheology of the upper mantle and the mantle wedge: A view from the experimentalists. *Inside the Subduction Factory*, *138*, 83–105. https://doi.org/10.1029/138gm06

Hoechner, A., Sobolev, S. V., Einarsson, I., & Wang, R. (2011). Investigation on afterslip and steady state and transient rheology based on postseismic deformation and geoid change caused by the Sumatra 2004 earthquake. *Geochemistry, Geophysics, Geosystems*, *12*(7), Q07010. https://doi.org/10.1029/2010GC003450

Honsho, C., Kido, M., Tomita, F., & Uchida, N. (2019). Offshore postseismic deformation of the 2011 Tohoku earthquake revisited: Application of an improved GPS-Acoustic positioning method considering horizontal gradient of sound speed structure. *Journal of Geophysical Research: Solid Earth*, *124*(6), 5990–6009. https://doi.org/10.1029/2018JB017135

Hooper, A., Pietrzak, J., Simons, W., Cui, H., Riva, R., Naeije, M., et al. (2013). Importance of horizontal seafloor motion on tsunami height for the 2011 Mw=9.0 Tohoku-Oki earthquake. *Earth and Planetary Science Letters*, *361*, 469–479. https://doi.org/10.1016/j.epsl.2012.11.013

Hu, Y., Bürgmann, R., Freymueller, J. T., Banerjee, P., & Wang, K. (2014). Contributions of poroelastic rebound and a weak volcanic arc to the postseismic deformation of the 2011 Tohoku earthquake. *Earth Planets and Space*, *66*(1), 106. https://doi.org/10.1186/1880-5981-66-106

Hu, Y., Bürgmann, R., Uchida, N., Banerjee, P., & Freymueller, J. T. (2016). Stress-driven relaxation of heterogeneous upper mantle and time-dependent afterslip following the 2011 Tohoku earthquake. *Journal of Geophysical Research: Solid Earth*, *121*(1), 385–411. https://doi.org/10.1002/2015JB012508

Hu, Y., & Wang, K. (2012). Spherical-Earth finite element model of short-term postseismic deformation following the 2004 Sumatra earthquake. *Journal of Geophysical Research*, *117*(B5), B05404. https://doi.org/10.1029/2012JB009153

Iinuma, T., Hino, R., Uchida, N., Nakamura, W., Kido, M., Osada, Y., & Miura, S. (2016). Seafloor observations indicate spatial separation of coseismic and postseismic slips in the 2011 Tohoku earthquake. *Nature Communications*, *7*(1), 13506. https://doi.org/10.1038/ncomms13506

Itoh, Y., Nishimura, T., Wang, K., & He, J. (2021). New megathrust locking model for the southern Kurile subduction zone incorporating viscoelastic relaxation and non-uniform compliance of upper plate. *Journal of Geophysical Research: Solid Earth*, *126*(5), e2020JB019981. https://doi.org/10.1029/2020JB019981

Itoh, Y., Wang, K., Nishimura, T., & He, J. (2019). Compliant volcanic arc and backarc crust in southern Kurile suggested by interseismic geodetic deformation. *Geophysical Research Letters*, *46*(21), 11790–11798. https://doi.org/10.1029/2019GL084656

Iwamori, H. (2007). Transportation of H2O beneath the Japan arcs and its implications for global water circulation. *Chemical Geology*, *239*(3), 182–198. https://doi.org/10.1016/j.chemgeo.2006.08.011

Japan Meteorological Agency. (2000). Recent seismic activity in the Miyakejima and Niijima-Kozushima region, Japan —The largest earthquake swarm ever recorded—. *Earth Planets and Space*, *52*(8), i–viii. https://doi.org/10.1186/BF03351657

Japan Meteorological Agency. (2013). Japan Meteorological Agency (English version). In *National catalogue of the active volcanoes in Japan* (The fourth edition).

Johnson, K. M., Fukuda, J. i., & Segall, P. (2012). Challenging the rate-state asperity model: Afterslip following the 2011 M9 Tohoku-Oki, Japan, earthquake. *Geophysical Research Letters*, *39*(20), L20302. https://doi.org/10.1029/2012GL052901

Julve, J., Barbot, S., Moreno, M., Tassara, A., Araya, R., Catalán, N., et al. (2024). Recurrence time and size of Chilean earthquakes influenced by geological structure. *Nature Geoscience*, *17*(1), 79–87. https://doi.org/10.1038/s41561-023-01327-8

Karato, S.-i. (2008). *Deformation of Earth materials: An introduction to the Rheology of Solid Earth*. Cambridge University Press. https://doi.org/10.1017/S0016756809006323

Karato, S.-I., & Jung, H. (2003). Effects of pressure on high-temperature dislocation creep in olivine. *Philosophical Magazine*, *83*(3), 401–414. https://doi.org/10.1080/0141861021000025829

Karato, S.-I., Paterson, M. S., & FitzGerald, J. D. (1986). Rheology of synthetic olivine aggregates: Influence of grain size and water. *Journal of Geophysical Research*, *91*(B8), 8151–8176. https://doi.org/10.1029/JB091iB08p08151

Karato, S.-i., & Wu, P. (1993). Rheology of the upper mantle: A synthesis. *Science*, *260*(5109), 771–778. https://doi.org/10.1126/science.260.5109.771

Kawakatsu, H., Kumar, P., Takei, Y., Shinohara, M., Kanazawa, T., Araki, E., & Suyehiro, K. (2009). Seismic evidence for sharp lithosphere-asthenosphere boundaries of oceanic plates. *Science*, *324*(5926), 499–502. https://doi.org/10.1126/science.1169499

Kirby, S. H., & Kronenberg, A. K. (1987). Rheology of the lithosphere: Selected topics. *Reviews of Geophysics*, *25*(6), 1219–1244. https://doi.org/10.1029/rg025i006p01219

Koohkan, M., & Bocquet, M. (2012). Accounting for representativeness errors in the inversion of atmospheric constituent emissions: Application to the retrieval of regional carbon monoxide fluxes. *Tellus B: Chemical and Physical Meteorology*, *64*(1), 19047. https://doi.org/10.3402/tellusb.v64i0.19047

Li, S., Fukuda, J. i., & Oncken, O. (2020). Geodetic evidence of time-dependent viscoelastic interseismic deformation driven by megathrust locking in the southwest Japan subduction zone. *Geophysical Research Letters*, *47*(4), e2019GL085551. https://doi.org/10.1029/2019GL085551

Loveless, J. P., & Meade, B. J. (2011). Spatial correlation of interseismic coupling and coseismic rupture extent of the 2011 MW = 9.0 Tohoku-Oki earthquake. *Geophysical Research Letters*, *38*(17), L17306. https://doi.org/10.1029/2011GL048561

Luo, H., & Wang, K. (2021). Postseismic geodetic signature of cold forearc mantle in subduction zones. *Nature Geoscience*, *14*(2), 104–109. https://doi.org/10.1038/s41561-020-00679-9

Luo, H., & Wang, K. (2022). Finding simplicity in the complexity of postseismic coastal uplift and subsidence following great subduction earthquakes. *Journal of Geophysical Research: Solid Earth*, *127*(10), e2022JB024471. https://doi.org/10.1029/2022JB024471

Mallick, R., Lambert, V., & Meade, B. (2022). On the choice and implications of rheologies that maintain kinematic and dynamic consistency over the entire earthquake cycle. *Journal of Geophysical Research: Solid Earth*, *127*(9), e2022JB024683. https://doi.org/10.1029/2022JB024683

Marone, C. J. (1998). Laboratory-derived friction laws and their application to seismic faulting. *Annual Review of Earth and Planetary Sciences*, *26*(1), 643–696. https://doi.org/10.1146/annurev.earth.26.1.643

Marone, C. J., Scholtz, C. H., & Bilham, R. (1991). On the mechanics of earthquake afterslip. *Journal of Geophysical Research*, *96*(B5), 8441–8452. https://doi.org/10.1029/91JB00275

Marsman, C., Vossepoel, F., van Dinther, Y., & Govers, R. (2023). Estimating geodynamic model parameters from geodetic observations using a particle method. *Geophysical Journal International*, *236*(3), 1183–1205. https://doi.org/10.1093/gji/ggad450

Marsman, C., Vossepoel, F. C., D'Acquisto, M., van Dinther, Y., van de Wiel, L., & Govers, R. (2024). Model output for the paper "Unravelling processes and rheology of the Tohoku earthquake cycle using Bayesian inference"(version 1) [Dataset]. *Model output for the paper "Unravelling Processes and Rheology of the Tohoku Earthquake Cycle Using Bayesian Inference"*. https://doi.org/10.24416/UU01-E4PRQX

Masuti, S., Barbot, S. D., Karato, S.-i., Feng, L., & Banerjee, P. (2016). Upper-mantle water stratification inferred from observations of the 2012 Indian Ocean earthquake. *Nature*, *538*(7625), 373–377. https://doi.org/10.1038/nature19783

Melosh, H. J., & Raefsky, A. (1980). The dynamical origin of subduction zone topography. *Geophysical Journal International*, *60*(3), 333–354. https://doi.org/10.1111/j.1365-246X.1980.tb04812.x

Melosh, H. J., & Raefsky, A. (1983). Anelastic response of the Earth to a dip slip earthquake. *Journal of Geophysical Research*, *88*(B1), 515–526. https://doi.org/10.1029/JB088iB01p00515

Mitsui, Y., & Heki, K. (2013). Scaling of early afterslip velocity and possible detection of tsunami-induced subsidence by GPS measurements immediately after the 2011 Tohoku-Oki earthquake. *Geophysical Journal International*, *195*(1), 238–248. https://doi.org/10.1093/gji/ggt256

Montési, L. G. J. (2004). Controls of shear zone rheology and tectonic loading on postseismic creep. *Journal of Geophysical Research*, *109*(B10), B10404. https://doi.org/10.1029/2003JB002925

Montési, L. G. J., & Hirth, G. (2003). Grain size evolution and the rheology of ductile shear zones: From laboratory experiments to postseismic creep. *Earth and Planetary Science Letters*, *211*(1), 97–110. https://doi.org/10.1016/S0012-821X(03)00196-1

Moreno, M. S., Bolte, J., Klotz, J., & Melnick, D. (2009). Impact of megathrust geometry on inversion of coseismic slip from geodetic data: Application to the 1960 Chile earthquake. *Geophysical Research Letters*, *36*(16), L16310. https://doi.org/10.1029/2009GL039276

Morishige, M. (2022). The thermal structure of subduction zones predicted by plate cooling models with variable thermal properties. *Geophysical Journal International*, *229*(3), 1490–1502. https://doi.org/10.1093/gji/ggac008

Morishige, M., & Kuwatani, T. (2020). Bayesian inversion of surface heat flow in subduction zones: A framework to refine geodynamic models based on observational constraints. *Geophysical Journal International*, *222*(1), 103–109. https://doi.org/10.1093/gji/ggaa149

Muto, J., Moore, J. D. P., Barbot, S., Iinuma, T., Ohta, Y., & Iwamori, H. (2019). Coupled afterslip and transient mantle flow after the 2011 Tohoku earthquake. *Science Advances*, *5*(9), eaaw1164. https://doi.org/10.1126/sciadv.aaw1164

Muto, J., Shibazaki, B., Iinuma, T., Ito, Y., Ohta, Y., Miura, S., & Nakai, Y. (2016). Heterogeneous rheology controlled postseismic deformation of the 2011 Tohoku-Oki earthquake. *Geophysical Research Letters*, *43*(10), 4971–4978. https://doi.org/10.1002/2016GL068113

Nakagawa, H. (2009). Development and validation of GEONET new analysis strategy (version 4). In *Journal of the Geospatial information Authority of Japan* (Vol. 118). [in Japanese].1.

Nishimura, T. (2014). Pre-Co-and post-seismic deformation of the 2011 Tohoku-oki earthquake and its implication to a paradox in short-term and long-term deformation. *Journal of Disaster Research*, *9*(3), 294–302. https://doi.org/10.20965/jdr.2014.p0294

Ozawa, S., Nishimura, T., Munekane, H., Suito, H., Kobayashi, T., Tobita, M., & Imakiire, T. (2012). Preceding, coseismic, and postseismic slips of the 2011 Tohoku earthquake, Japan. *Journal of Geophysical Research: Solid Earth*, *117*(B7). https://doi.org/10.1029/2011JB009120

Peña, C., Heidbach, O., Moreno, M., Bedford, J., Ziegler, M., Tassara, A., & Oncken, O. (2020). Impact of power-law rheology on the viscoelastic relaxation pattern and afterslip distribution following the 2010 Mw 8.8 Maule earthquake. *Earth and Planetary Science Letters*, *542*, 116292. https://doi.org/10.1016/j.epsl.2020.116292

Perfettini, H., & Avouac, J. P. (2014). The seismic cycle in the area of the 2011 Mw9.0 Tohoku-Oki earthquake. *Journal of Geophysical Research: Solid Earth*, *119*(5), 4469–4515. https://doi.org/10.1002/2013JB010697

Pollitz, F., Banerjee, P., Grijalva, K., Nagarajan, B., & Bürgmann, R. (2008). Effect of 3-D viscoelastic structure on post-seismic relaxation from the 2004 M= 9.2 Sumatra earthquake. *Geophysical Journal International*, *173*(1), 189–204. https://doi.org/10.1111/j.1365-246X.2007.03666.x

Ranalli, G. (1987). Rheology of the Earth: Deformation and flow processes in geophysics and geodynamics.

Ruina, A. (1983). Slip instability and state variable friction laws. *Journal of Geophysical Research*, *88*(B12), 10359–10370. https://doi.org/10.1029/JB088iB12p10359

Sasajima, R., Shibazaki, B., Iwamori, H., Nishimura, T., & Nakai, Y. (2019). Mechanism of subsidence of the Northeast Japan forearc during the late period of a gigantic earthquake cycle. *Scientific Reports*, *9*(1), 5726. https://doi.org/10.1038/s41598-019-42169-y

Schmerr, N. (2012). The Gutenberg Discontinuity: Melt at the lithosphere-asthenosphere boundary. *Science*, *335*(6075), 1480–1483. https://doi.org/10.1126/science.1215433

Scholz, C. H. (1988). The brittle-plastic transition and the depth of seismic faulting. *Geologische Rundschau*, *77*(1), 319–328. https://doi.org/10.1007/BF01848693

Sibson, R. H. (1983). Continental fault structure and the shallow earthquake source. *Journal of the Geological Society*, *140*(5), 741–767. https://doi.org/10.1144/gsjgs.140.5.0741

Silverii, F., Cheloni, D., D'Agostino, N., Selvaggi, G., & Boschi, E. (2014). Post-seismic slip of the 2011 Tohoku-oki earthquake from GPS observations: Implications for depth-dependent properties of subduction megathrusts. *Geophysical Journal International*, *198*(1), 580–596. https://doi.org/10.1093/gji/ggu149

Sobolev, S. V., & Muldashev, I. A. (2017). Modeling seismic cycles of great megathrust earthquakes across the scales with focus at postseismic phase. *Geochemistry, Geophysics, Geosystems*, *18*(12), 4387–4408. https://doi.org/10.1002/2017GC007230

Suito, H. (2017). Importance of rheological heterogeneity for interpreting viscoelastic relaxation caused by the 2011 Tohoku-Oki earthquake. *Earth Planets and Space*, *69*(1), 21. https://doi.org/10.1186/s40623-017-0611-9

Suito, H., & Freymueller, J. T. (2009). A viscoelastic and afterslip postseismic deformation model for the 1964 Alaska earthquake. *Journal of Geophysical Research*, *114*(B11), B11404. https://doi.org/10.1029/2008JB005954

Sun, T., & Wang, K. (2015). Viscoelastic relaxation following subduction earthquakes and its effects on afterslip determination. *Journal of Geophysical Research: Solid Earth*, *120*(2), 1329–1344. https://doi.org/10.1002/2014JB011707

Sun, T., Wang, K., & He, J. (2024). Geodetic signature of a weak Lithosphere-Asthenosphere Boundary in postseismic deformation of large subduction earthquakes. *Earth and Planetary Science Letters*, *630*, 118619. https://doi.org/10.1016/j.epsl.2024.118619

Sun, T., Wang, K., Iinuma, T., Hino, R., He, J., Fujimoto, H., et al. (2014). Prevalence of viscoelastic relaxation after the 2011 Tohoku-Oki earthquake. *Nature*, *514*(7520), 84–87. https://doi.org/10.1038/nature13778

Tanaka, Y., Klemann, V., Fleming, K., & Martinec, Z. (2009). Spectral finite element approach to postseismic deformation in a viscoelastic self-gravitating spherical Earth. *Geophysical Journal International*, *176*(3), 715–739. https://doi.org/10.1111/j.1365-246X.2008.04015.x

Tobita, M. (2016). Combined logarithmic and exponential function model for fitting postseismic GNSS time series after 2011 Tohoku-Oki earthquake. *Earth Planets and Space*, *68*(1), 41. https://doi.org/10.1186/s40623-016-0422-4

Trubienko, O., Fleitout, L., Garaud, J.-D., & Vigny, C. (2013). Interpretation of interseismic deformations and the seismic cycle associated with large subduction earthquakes. *Tectonophysics*, *589*, 126–141. https://doi.org/10.1016/j.tecto.2012.12.027

Uchida, N., & Matsuzawa, T. (2013). Pre- and postseismic slow slip surrounding the 2011 Tohoku-Oki earthquake rupture. *Earth and Planetary Science Letters*, *374*, 81–91. https://doi.org/10.1016/j.epsl.2013.05.021

Usami, K., Ikehara, K., Kanamatsu, T., & McHugh, C. M. (2018). Supercycle in great earthquake recurrence along the Japan Trench over the last 4000 years. *Geoscience Letters*, *5*(1), 11. https://doi.org/10.1186/s40562-018-0110-2

van Dinther, Y., Preiswerk, L. E., & Gerya, T. V. (2019). A secondary zone of uplift due to megathrust earthquakes. *Pure and Applied Geophysics*, *176*(9), 4043–4068. https://doi.org/10.1007/s00024-019-02250-z

van Keken, P. E., & Wilson, C. R. (2023). An introductory review of the thermal structure of subduction zones: I—Motivation and selected examples. *Progress in Earth and Planetary Science*, *10*(1), 42. https://doi.org/10.1186/s40645-023-00573-z

Wada, I., & Wang, K. (2009). Common depth of slab-mantle decoupling: Reconciling diversity and uniformity of subduction zones. *Geochemistry, Geophysics, Geosystems*, *10*(10), Q10009. https://doi.org/10.1029/2009GC002570

Wang, K. (1995). Coupling of tectonic loading and earthquake fault slips at subduction zones. *Pure and Applied Geophysics*, *145*(3), 537–559. https://doi.org/10.1007/BF00879588

Wang, K., He, J., Dragert, H., & James, T. S. (2001). Three-dimensional viscoelastic interseismic deformation model for the Cascadia subduction zone. *Earth Planets and Space*, *53*(4), 295–306. https://doi.org/10.1186/BF03352386

Wang, K., Hu, Y., & He, J. (2012). Deformation cycles of subduction earthquakes in a viscoelastic Earth. *Nature*, *484*(7394), 327–332. https://doi.org/10.1038/nature11032

Wang, K., Sun, T., Brown, L., Hino, R., Tomita, F., Kido, M., et al. (2018). Learning from crustal deformation associated with the M9 2011 Tohoku-Oki earthquake. *Geosphere*, *14*(2), 552–571. https://doi.org/10.1130/ges01531.1

Watanabe, S.-i., Sato, M., Fujita, M., Ishikawa, T., Yokota, Y., Ujihara, N., & Asada, A. (2014). Evidence of viscoelastic deformation following the 2011 Tohoku-Oki earthquake revealed from seafloor geodetic observation. *Geophysical Research Letters*, *41*(16), 5789–5796. https://doi.org/10.1002/2014GL061134

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

Wilson, C. R., & van Keken, P. E. (2023). An introductory review of the thermal structure of subduction zones: II—Numerical approach and validation. *Progress in Earth and Planetary Science*, *10*(1), 68. https://doi.org/10.1186/s40645-023-00588-6

Wu, P. (2004). Using commercial finite element packages for the study of earth deformations, sea levels and the state of stress. *Geophysical Journal International*, *158*(2), 401–408. https://doi.org/10.1111/j.1365-246X.2004.02338.x

Yamagiwa, S., Miyazaki, S. i., Hirahara, K., & Fukahata, Y. (2015). Afterslip and viscoelastic relaxation following the 2011 Tohoku-Oki earthquake (Mw9.0) inferred from inland GPS and seafloor GPS/Acoustic data. *Geophysical Research Letters*, *42*(1), 66–73. https://doi.org/10.1002/2014GL061735

Yokota, Y., Ishikawa, T., & Watanabe, S.-i. (2018). Seafloor crustal deformation data along the subduction zones around Japan obtained by GNSS-A observations. *Scientific Data*, *5*(1), 180182. https://doi.org/10.1038/sdata.2018.182

Yuzariyadi, M., & Heki, K. (2021). Enhancement of interplate coupling in adjacent segments after recent megathrust earthquakes. *Tectonophysics*, *801*, 228719. https://doi.org/10.1016/j.tecto.2021.228719

Zhao, D., Matsuzawa, T., & Hasegawa, A. (1997). Morphology of the subducting slab boundary in the northeastern Japan arc. *Physics of the Earth and Planetary Interiors*, *102*(1), 89–104. https://doi.org/10.1016/S0031-9201(96)03258-X

References From the Supporting Information

Iinuma, T., Hino, R., Kido, M., Inazu, D., Osada, Y., Ito, Y., et al. (2012). Coseismic slip distribution of the 2011 off the Pacific Coast of Tohoku Earthquake (M9.0) refined by means of seafloor geodetic data. *Journal of Geophysical Research*, *117*(B7), B07409. https://doi.org/10.1029/2012JB009186