

Physics-informed deep learning approach for modeling crustal deformation

Received: 21 April 2022

Accepted: 9 November 2022

Published online: 19 November 2022

 Check for updates

Tomohisa Okazaki¹✉, Takeo Ito², Kazuro Hirahara¹ & Naonori Ueda¹

The movement and deformation of the Earth's crust and upper mantle provide critical insights into the evolution of earthquake processes and future earthquake potentials. Crustal deformation can be modeled by dislocation models that represent earthquake faults in the crust as defects in a continuum medium. In this study, we propose a physics-informed deep learning approach to model crustal deformation due to earthquakes. Neural networks can represent continuous displacement fields in arbitrary geometrical structures and mechanical properties of rocks by incorporating governing equations and boundary conditions into a loss function. The polar coordinate system is introduced to accurately model the displacement discontinuity on a fault as a boundary condition. We illustrate the validity and usefulness of this approach through example problems with strike-slip faults. This approach has a potential advantage over conventional approaches in that it could be straightforwardly extended to high dimensional, anelastic, nonlinear, and inverse problems.

Geodetic observations made using different sensors and instruments, including global navigation satellite systems, have produced a significant amount of data on crustal deformation. Modeling such observed deformation is fundamental to understanding the mechanics of earthquake processes^{1,2}. A dislocation model of a slip on earthquake faults is commonly used for the forward and inverse modeling of coseismic and postseismic deformation and earthquake cycles^{3–5}. Over the past decades, analytical and semianalytical approaches have been developed for linear rheologies^{6–8}, and fully numerical methods have been constructed for complex structures, including nonlinear mechanical properties^{9–11}.

Recent advances in machine learning, especially deep learning techniques, have occurred due to the large amount of available data^{12–14}. Applications in geophysics have been implemented by utilizing accumulating seismic records^{15–17}. In contrast, deep learning has also promoted the application of machine-learning approaches to physical systems, specifically the solution of partial differential equations (PDEs). Automatic differentiation¹⁸ developed for the optimization of neural networks (NNs) plays a central role in the efficient computation of derivatives^{19,20}. Among them, physics-informed neural networks (PINNs) have been proposed for solving both the forward

and inverse problems of PDEs in a unified way²¹. PINNs represent continuous solutions without discretization and can be trained to conform to a physical law by incorporating the target PDEs and boundary/initial conditions into loss functions. Because of their simple implementation and applicability to different problem types, PINNs have received significant attention in physics and engineering²². In geophysics, a seismic inversion method was developed based on the similarity between automatic differentiation and the adjoint-state method²³. PINNs have been applied to synthetic models of seismic tomography²⁴ and full waveform inversions²⁵.

This study applies PINNs to dislocation models of crustal deformation. An essential characteristic of these models is that the displacement field is discontinuous across the fault surface and cannot be directly approximated by NNs that represent continuous functions. To resolve this difficulty, we set an appropriate coordinate system to separate the values on the two sides of the displacement discontinuity. This formulation enables precise modeling of crustal deformation, including near-fault locations. PINNs can be applied to complex structures and easily extended to high-dimensional, anelastic, and nonlinear problems, which serves as a potential advantage over conventional approaches as follows.

¹RIKEN Center for Advanced Intelligence Project, Seika, Japan. ²Graduate School of Environmental Studies, Nagoya University, Nagoya, Japan.

✉ e-mail: tomohisa.okazaki@riken.jp

Analytical approaches use Green's functions (GFs) to study crustal deformation from arbitrary slip distributions. They provide continuous solutions and explicit dependencies on model parameters. Exact expressions of elastostatic GFs have been obtained for several crustal structures, such as homogeneous half-space⁶, layered half-space²⁶, and layered spherical Earth²⁷. Viscoelastic rheology is addressed using the correspondence principle. Deformation due to a finite fault is expressed as a convolution of GFs and slip distribution on the fault. The surface topography is addressed by a series expansion assuming a small slope^{28–30}. However, the known analytical solutions have been limited to simple structures.

Semianalytical approaches have been developed to model composite rheologies and quasi-static earthquake cycles^{7,8,31}. A curved fault is typically expressed as a sum of planar sub-faults because its GFs require a deliberate derivation³². The general topography is addressed by sophisticated formulations with full-space GFs³³. However, the assumption of linear responses is a major limitation of these methods. Although a general representation of anelastic deformation using GFs was formulated for nonlinear quasi-static problems³⁴, GFs have only been derived in a homogeneous half-space³⁵.

Realistic problems are solved through fully numerical methods. In particular, the finite element method (FEM) is suitable for modeling complex tectonic settings such as subduction zones^{2,10,36}, mountain regions³⁷, and nonlinear rheologies¹¹. PINNs share common advantages with FEM: topography and heterogeneity can be modeled, and nonlinear PDEs are implemented in a similar manner to linear PDEs²¹. Therefore, PINNs are frequently compared with FEM³⁸. The FEM generates a discretized mesh, which restricts the model resolution. A large amount of memory required to store information at every grid point can make modeling complex structures difficult. In contrast, PINNs directly model a continuous field, which potentially achieves high accuracy by trained on arbitrarily large number of points. NNs can be trained with minibatch iteratively without storing all training points, which could be a computational advantage in modeling realistic large-scale problems. PINNs have another advantage in directly solving infinite domain problems without imposing boundary conditions on model boundaries; however, FEM can solve only finite domain problems and often requires modeling domains that are far larger than the target domains with careful treatment of boundary conditions to reduce boundary effects.

In a pioneering application of NNs to the forward modeling of crustal deformation³⁹, NNs were trained on simulation results of semianalytical methods^{6,26} to interpolate solutions at arbitrary locations and model parameters such as source depth and viscosity. This approach considerably accelerates the estimation of deformation in simple problems to which semianalytical approaches can be applied, whereas PINNs can address complex problems without any existing solver despite a longer computational time.

Results

Physics-informed neural network modeling

We consider slips on strike-slip faults, a model that could be used to describe repeated destructive earthquakes. For simplicity, we assume infinitely long strike-slip faults in linear elastic media. By taking the x - and z -axes in the horizontal direction and the y -axis in the vertical direction, we suppose a displacement $u(x, y)$ to be parallel to and invariant in the z -direction. We denote a medium as V , a fault (dislocation surface, DS) as Σ , and the Earth's surface (free surface, FS) as S (Fig. 1a). Because $u(x, y)$ is discontinuous across Σ , we use the polar coordinates (r, θ) whose branch cut is defined along Σ . In this way, $u(r, \theta)$ is continuous in the entire domain and the slip on Σ can be represented as a constraint between Σ_+ and Σ_- (Fig. 1b).

PINN modeling consists of three building blocks (Fig. 1c). First, an NN surrogates a continuous displacement field $u(r, \theta)$. Second, derivatives of the output u with respect to the input variables (r, θ) are

evaluated using automatic differentiation¹⁸. Finally, a loss function L is defined by the sum of the squared residual of the governing equation and boundary conditions: the equilibrium of linear elasticity L_{PDE} , the displacement discontinuity on the fault $L_{\text{DS-U}}$, the traction continuity on the fault $L_{\text{DS-T}}$, and the free-surface condition on the Earth's surface L_{FS} . The NN parameters are updated to decrease the loss function L using a stochastic gradient method. In this study, training is iterated until $L < 10^{-6}$ is satisfied for fixed grid points. See 'Methods' for the details of the NN architecture, mathematical expressions of L , and optimization procedure.

Modeling applications

We first consider vertical faults in a homogeneous half-space, for which analytical solutions are known. In homogeneous media, deformation is independent of a shear modulus μ . The following three models are considered. Models 1A and 1C represent a surface fault (Fig. 2a, left). Model 1B represents a buried fault extending from the locking to an infinite depth (Fig. 2a, right), which has been used for strain accumulation along a plate boundary with a relative motion during interseismic periods⁴⁰. By taking $\theta = 0$ upward, the surface and buried faults are expressed by branch cuts at $\theta = 0$ and $\theta = \pi$, respectively. Fault slips are uniform in Models 1A and 1B, whereas a distributed slip of a tapered stress is assumed in Model 1C (Fig. 2b). In the following modeling, all quantities are normalized by characteristic scales: spatial coordinates x and y by a fault depth d , displacement u by a maximum slip amount s_0 , and shear modulus μ by a reference value μ_0 . For example, strain is expressed in the unit of s_0 / d . The surface displacement and strain ε_{xz} are shown in Fig. 2c, d with analytical solutions³ (2-D displacement and strain are shown in Supplementary Fig. 1). The root-mean-square errors (RMSEs) are shown in Table 1. PINN solutions are generally accurate and the RMSE is sufficiently smaller than typical u and ε values. The RMSE is larger in u than in ε by approximately five times. This is because the governing PDE constrains not u itself but its spatial derivatives. Deformation is localized in the distributed slip (Model 1C) compared to that in the uniform slip (Model 1A), which results in a higher strain near the fault.

PINNs have advantages in the modeling of complex crustal structures. In particular, continuous geometry and changes in mechanical properties can be expressed without any approximation such as series expansion and discretization. Figure 3a shows the vertical section of an example dislocation model; a fault and the Earth's surface are curved, and the mechanical property varies continuously in space. Two slip distributions, a uniform slip in Model 2A and a distributed slip identical to Model 1C (Fig. 2b) in Model 2B, are considered on the fault.

The obtained strain fields ε_{yz} are shown in Fig. 3b (2-D displacement and strain are shown in Supplementary Fig. 2). The strain field in Model 2A is continuous and independent of the fault surface (dislocation surface) and was concentrated at the lower tip of the fault (dislocation line). Thus, crustal deformation for a uniform slip is completely determined by a dislocation line, which can be observed for different fault geometries (Supplementary Fig. 3). This property is well known for a plane fault in a homogeneous half-space³, whereas the present result applies to a curved fault in a heterogeneous medium. In contrast, the strain field for Model 2B is discontinuous on the fault surface. The surface displacement and strain are shown in Fig. 3c, d with FEM solutions (see Supplementary Text 1 for the FEM modeling). The discrepancy in these models is larger than that in Models 1A–1C but not more than twice that in Model 1C (Table 1). In Model 2B, a high strain is distributed on the fault surface (Fig. 3b), which results in localized displacement u (Fig. 3c) and high strain near the fault on the Earth's surface (Fig. 3d).

The Earth's crust is composed of various rock types with different mechanical properties. Strain accumulates at material boundaries where earthquakes tend to occur. Some earthquake faults are located

within damage zones extending to a considerable depth. Therefore, the modeling of discontinuous media is of practical importance. In this study, a displacement field is modeled with two NNs in the individual material regions (V_1 and V_2). Two terms L_{MB-u} and L_{MB-T} are added to the loss function L to impose appropriate conditions on the material boundary B . See ‘Methods’ for more details.

Here, we consider that the mechanical property changes discontinuously on either side of a buried vertical fault. Model 3A has a discontinuity across the fault and Model 3B has a compliant fault zone (Fig. 4a). The contrast in shear modulus is two in both models. The obtained strain fields ϵ_{xz} are shown in Fig. 4b (2-D displacement and strain are shown in Supplementary Fig. 4). Strain is discontinuous across the material boundary B . The surface displacement and strain are shown in Fig. 4c, d with analytical solutions³, and RMSEs are shown in Table 1. PINN solutions are accurate in Model 3A but exhibit a systematic overestimation at a long distance from the fault in Model 3B. This would be because the material boundary B isolates V_2 from the fault Σ on which a displacement discontinuity (i.e. Dirichlet boundary condition) is imposed, which leads to error accumulation at a long distance. This suggests that multiple material discontinuities can complicate PINN’s convergence property.

Computational costs

Table 1 summarizes model structures and computational costs on a single CPU (Intel Core i7, 3.60 GHz, 4 cores, 8 processors, and 16 GB memory) of example problems in this study. We note that NNs consist of 8 hidden layers with 40 nodes and the batch size of training is 256 in

V and 64 on Σ , S , and B in all problems (see ‘Methods’ for details). The number of iterations required to achieve the desired precision ($L < 10^{-6}$) has significant dependence on model structures. The geometry and mechanical properties had minor effects (Model 2A), whereas the distributed fault slip led to many iterations (Models 1C and 2B). This might be related to the property that strain only depends on the location of dislocation lines for uniform slips, as discussed previously. The material discontinuities (Models 3A and 3B) did not significantly increase the number of iterations but increased computational time per iteration by approximately 1.6 folds because the use of two NNs doubles the number of NN parameters. The column ‘Transfer’ indicates that the trained NN parameters on similar but simple problems are used as initial weights of the target problems (see ‘Methods’ for details). In experiments, training without transfer increased computational time by 1.15, 1.38, 4.68, and 2.13 folds in Models 1C, 2B, 3A, and 3B, respectively. This indicates that the transfer of NN parameters is particularly effective for discontinuous material problems.

The computational time in the FEM modeling is 256 and 715 s for Models 2A and 2B, respectively, which is significantly shorter than that in the PINN modeling (Table 1). Here, the number of cells and nodes in the FEM models is 790,392 and 138,110, respectively (Supplementary Text 1). Computational cost is currently a common challenge of PINN forward modeling^{41,42}. Fast and stable algorithms for PINN optimization should be investigated. In FEM, sophisticated mesh generation schemes including fault interfaces have been developed⁴³. Knowledge and experience on conventional solvers would play an essential role in the progress.

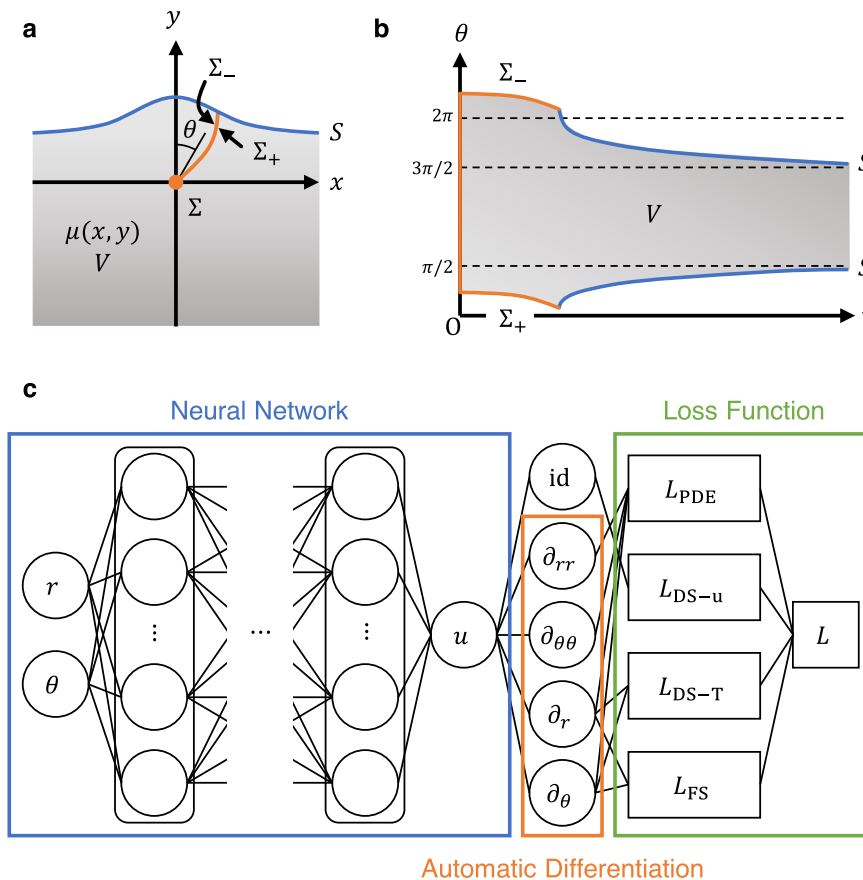


Fig. 1 | Physics-informed neural network (PINN) modeling of antiplane dislocations. **a** Cross-sectional view of an infinitely long strike-slip fault model. Shaded region, orange line, and blue line represent elastic medium V of shear modulus $\mu(x, y)$, fault Σ , and Earth’s surface S , respectively. Σ_+ and Σ_- represent the two sides

of Σ . The fault moves perpendicular to the page. **b** Polar coordinate representation of the model region. **c** PINN structure. A neural network surrogates a displacement $u(r, \theta)$ in the polar coordinates. Derivatives of u are calculated using automatic differentiation and constitute a loss function L .

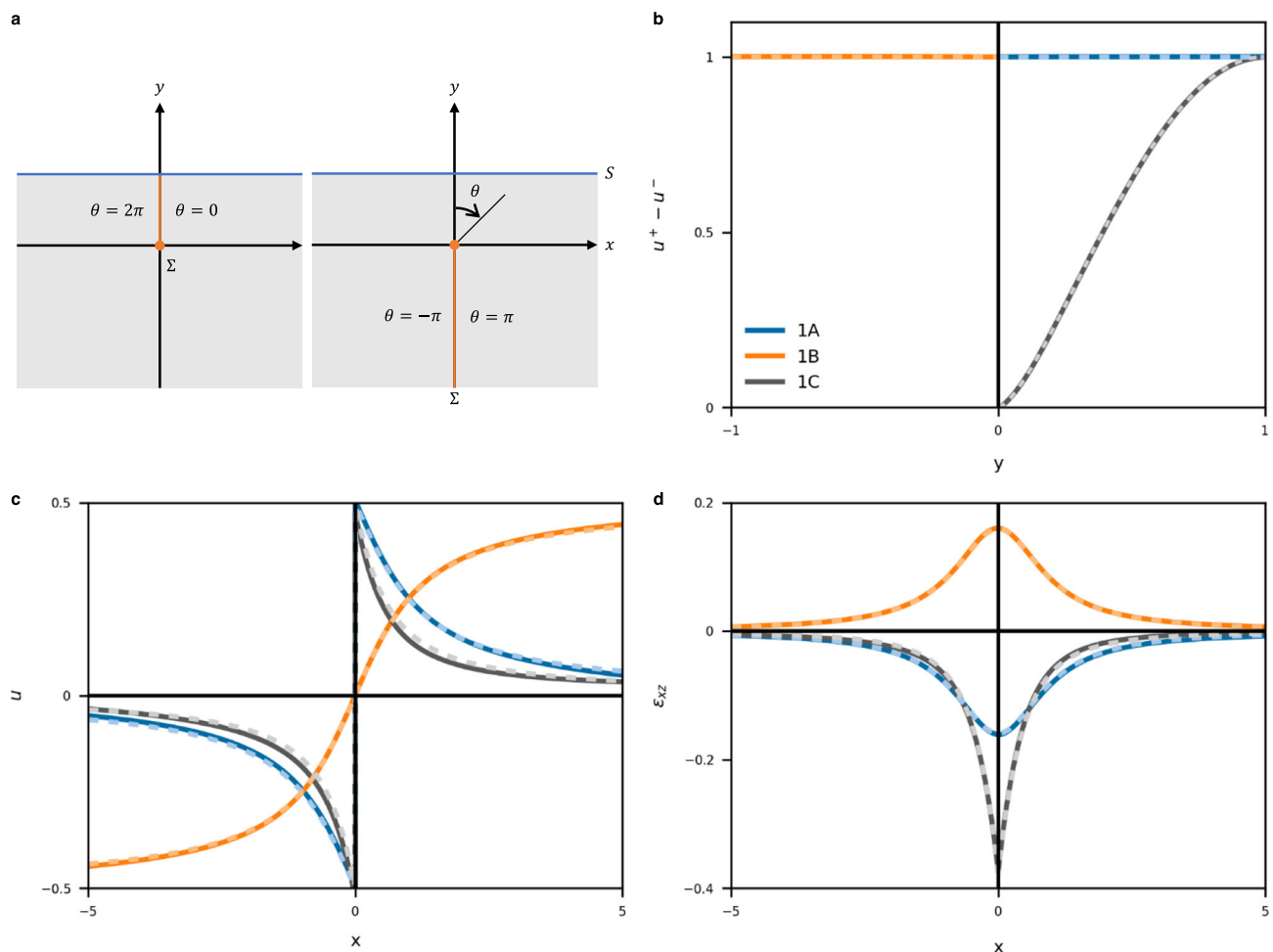


Fig. 2 | Model and estimated results in a homogeneous half-space. a Model structures. The orange and blue lines represent fault Σ and Earth's surface S , respectively. Surface (left) and buried (right) faults are distinguished by the branch cut. **b** Assumed displacement discontinuities on the fault. **c** Estimated surface

displacements. **d** Estimated surface strains. In (b)–(d), blue, orange and gray lines indicate Models 1A, 1B, and 1C, respectively. Solid and dashed lines indicate PINN and analytical solutions, respectively.

Table 1 | Model structures, root-mean-square errors (RMSEs), and computational costs

Model	Transfer	Fault	RMSE (u)	RMSE (ϵ)	Iterations	Time (s)	T/I (ms)
1A	–	Surface	6.733e-3	1.213e-3	22,875	825	36.07
1B	–	Buried	3.372e-3	6.257e-4	19,487	705	36.18
1C	1A	Surface	1.409e-2	2.732e-3	140,229	4892	34.89
2A	–	Surface	2.557e-2	3.727e-3	24,612	894	36.32
2B	2A	Surface	1.947e-2	3.957e-3	436,583	15,899	36.42
3A	1B	Buried	5.706e-3	1.013e-3	7803	466	59.72
3B	1B	Buried	1.523e-2	2.636e-3	37,583	2185	58.14

u displacement, ϵ strain, T/I time per iteration.

Discussion

This study focused on a dislocation model of strike-slip faults; however, PINNs can model dip-slip faults and quasi-static processes with a slight increase in the input and output variables of NNs. Inverse problems can be formulated by adding a data misfit term to the loss function; the simple implementation of inversion analyses is a major advantage of PINNs over conventional linear solvers. In fact, PINNs have currently succeeded more in inverse modeling than in forward modeling²². Because geophysical data are typically noisy, a Bayesian

approach⁴⁴ may be required for stable inversion. PINNs can also be applied to general rheologies, such as viscoelasticity, poroelasticity, and power law creep, by changing the loss term based on the governing equation. The difference in regional rheologies (e.g., elasticity in the crust and viscoelasticity in the upper mantle) can be treated by defining loss functions in each subregion, which is similar to discontinuous media.

This study presented successful applications of PINNs for modeling crustal deformation in antiplane problems with simple NN

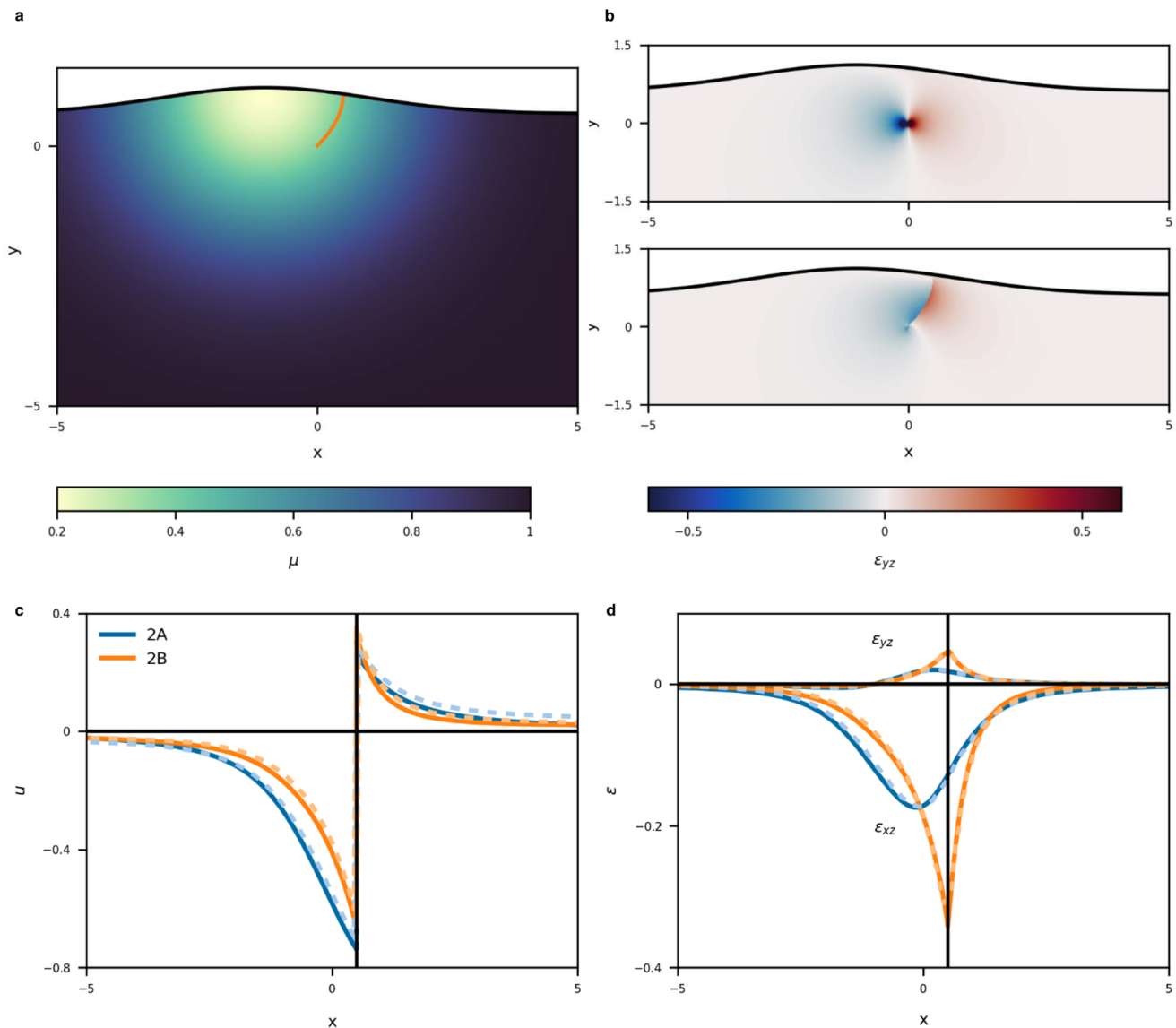


Fig. 3 | Model and estimated results of a curved, heterogeneous structure. **a** Model structure. Colors represent the shear modulus. The orange curve represents the fault surface. **b** Estimated strains for Models 2A (top) and 2B (bottom).

c Estimated surface displacements. **d** Estimated surface strains. In (c) and (d), blue and orange lines indicate Models 2A and 2B, respectively. Solid and dashed lines indicate PINN and FEM solutions, respectively.

architecture and optimization procedure. However, realistic problems require large and higher dimensional model space, and various material properties and rheologies require an increasing number of corresponding NNs. It has been recognized that PINNs can sometimes converge to erroneous solutions in time-dependent modeling^{45,46}. These factors would incur more computational costs for optimization of the loss function to achieve sufficient accuracy. Understanding the method for stable and fast optimization is key for the application of PINNs to large-scale geophysical problems. PINNs²¹ are newcomers to machine learning, and studies aimed at realizing faster and more efficient optimization have been accelerating^{45–48}. Therefore, our proposed approach based on PINNs may be a powerful tool for realizing a wide variety of modeling applications in crustal deformation.

Methods

Dislocation model

In this study, we consider antiplane dislocations, which model infinitely long strike-slip faults. By taking the x - and z -axes in the horizontal direction and the y -axis in the vertical direction, we suppose that a displacement $u(x, y)$ is parallel to and invariant in the z -direction. We

denote a medium as V , a fault (dislocation surface, DS) as Σ , and the Earth's surface (free surface, FS) as S (Fig. 1a). The normal vectors to Σ and S are denoted by $\mathbf{n}^{\text{DS}} = (n_x^{\text{DS}}, n_y^{\text{DS}}, 0)$ and $\mathbf{n}^{\text{FS}} = (n_x^{\text{FS}}, n_y^{\text{FS}}, 0)$, respectively. In an isotropic linear elastic medium, the system of governing equations is given by³

$$\mu \nabla^2 u + \nabla \mu \cdot \nabla u = 0 \text{ in } V, \tag{1}$$

$$u^+ - u^- = s \text{ on } \Sigma, \tag{2}$$

$$\boldsymbol{\sigma}^+ \cdot \mathbf{n}^{\text{DS}} = \boldsymbol{\sigma}^- \cdot \mathbf{n}^{\text{DS}} \text{ on } \Sigma, \tag{3}$$

$$\boldsymbol{\sigma} \cdot \mathbf{n}^{\text{FS}} = 0 \text{ on } S, \tag{4}$$

where μ is the shear modulus, s is the slip on Σ , $\boldsymbol{\sigma}$ is the stress tensor, and the superscripts $+$ and $-$ represent the opposite sides of Σ . The first equation represents the equilibrium equation of linear elasticity, the second represents the displacement discontinuity on the

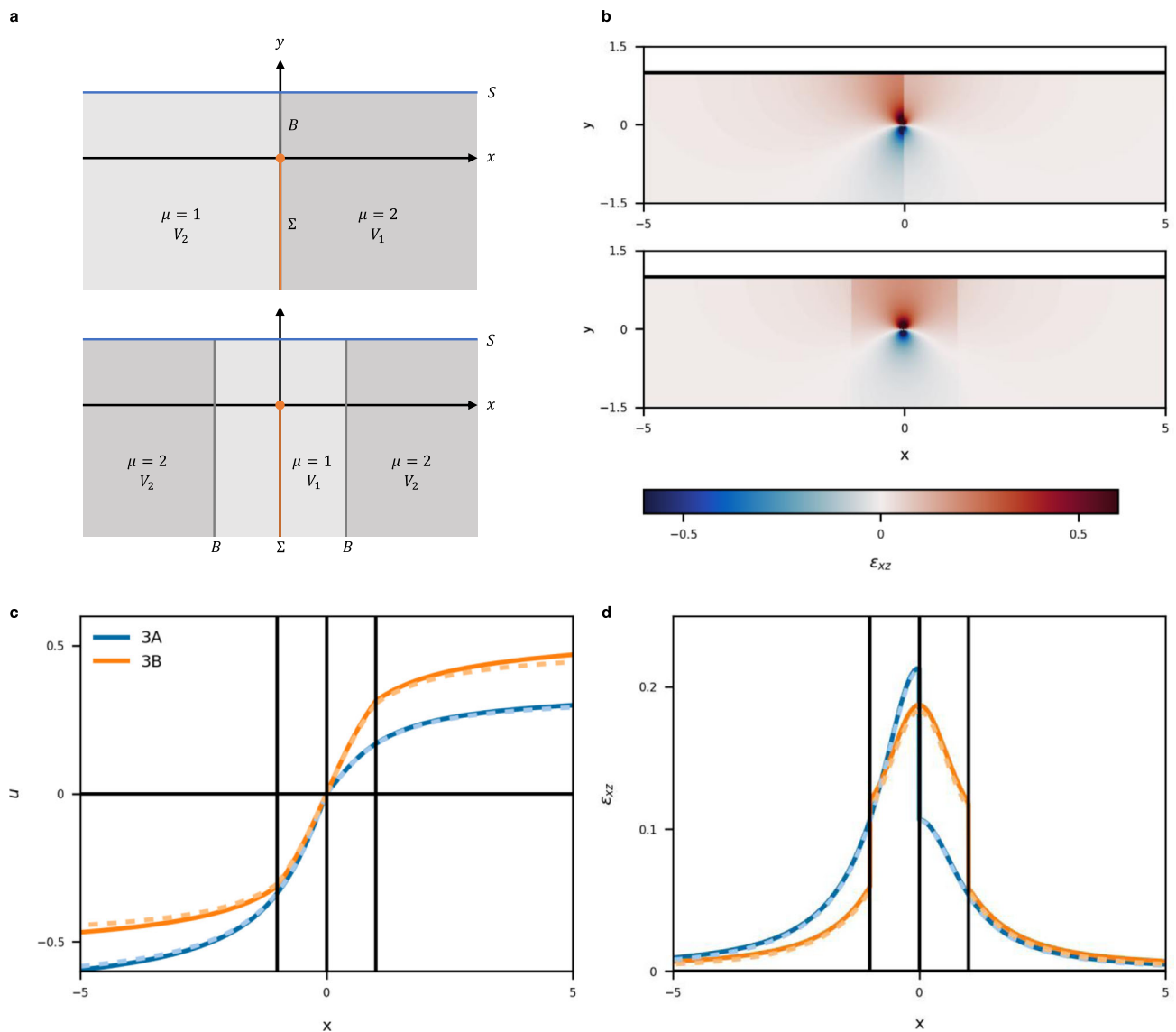


Fig. 4 | Model and estimated results of discontinuous structures. **a** Model structures of Models 3A (top) and 3B (bottom). Shaded regions represent elastic media V_1 and V_2 of shear modulus μ . The orange, blue, and gray lines represent fault Σ , Earth's surface S , and material boundary B , respectively. **b** Estimated strains for

Models 3A (top) and 3B (bottom). **c** Estimated surface displacements. **d** Estimated surface strains. In (c) and (d), blue and orange lines indicate Models 3A and 3B, respectively. Solid and dashed lines indicate PINN and analytical solutions, respectively.

fault, the third represents the traction continuity on the fault, and the fourth represents the free-surface condition on the Earth's surface.

Next, we consider that two regions V_1 and V_2 with shear moduli μ_1 and μ_2 , respectively, are in contact with a material boundary (MB) B (Fig. 4a). By denoting the displacements in V_1 and V_2 as u_1 and u_2 , respectively, the boundary conditions are expressed as

$$u_1 = u_2 \text{ on } B, \tag{5}$$

$$\sigma_1 \cdot \mathbf{n}^{\text{MB}} = \sigma_2 \cdot \mathbf{n}^{\text{MB}} \text{ on } B, \tag{6}$$

where $\mathbf{n}^{\text{MB}} = (n_x^{\text{MB}}, n_y^{\text{MB}}, 0)$ is the normal vector to B . They represent the displacement and traction continuity at the material boundary, respectively.

Neural network modeling

The displacement field u is modeled by NNs. $u(x, y)$ is discontinuous across a fault surface Σ (dislocation surface) and the stress can diverge at the fault tip (dislocation line), which prevents NNs from generating an accurate approximation. We therefore use the polar coordinate system (r, θ) whose pole is located at the dislocation line, and define a branch cut along the dislocation surface. A curved fault is expressed by a branch cut as a function of r . Modeling of $u(r, \theta)$ by NNs separates the coordinate values of the two sides of Σ (Fig. 1b), which results in the accurate modeling of displacements near fault surfaces.

A material boundary induces discontinuity not in displacements but in strains (derivatives of u). Therefore, it is difficult to approximate displacement using a single NN. We train the NNs in individual material regions and impose boundary conditions to ensure consistency between them. This is similar to the domain decomposition introduced to accelerate the convergence⁴⁹.

Loss function

A loss function is defined as the residuals of the governing equations and the boundary conditions. Using the stress–strain relation $\sigma_{xz} = \mu u_x / 2$ and $\sigma_{yz} = \mu u_y / 2$ for antiplane strains, the individual loss terms in the polar coordinates are written as

$$L_{\text{PDE}} = [r^2 u_{rr} + r u_r + u_{\theta\theta} + \mu^{-1} (r^2 \mu_r u_r + \mu_\theta u_\theta)]^2, \quad (7)$$

$$L_{\text{DS-u}} = (u^+ - u^- - s)^2, \quad (8)$$

$$L_{\text{DS-T}} = r^2 [\mu^+ (n_x^{\text{DS}} u_x^+ + n_y^{\text{DS}} u_y^+) - \mu^- (n_x^{\text{DS}} u_x^- + n_y^{\text{DS}} u_y^-)]^2, \quad (9)$$

$$L_{\text{FS}} = r^2 (n_x^{\text{FS}} u_x + n_y^{\text{FS}} u_y)^2, \quad (10)$$

$$L_{\text{MB-u}} = (u_1 - u_2)^2, \quad (11)$$

$$L_{\text{MB-T}} = r^2 [\mu_1 (n_x^{\text{MB}} u_{1x} + n_y^{\text{MB}} u_{1y}) - \mu_2 (n_x^{\text{MB}} u_{2x} + n_y^{\text{MB}} u_{2y})]^2, \quad (12)$$

where $u_x = \sin\theta u_r + r^{-1} \cos\theta u_\theta$ and $u_y = \cos\theta u_r - r^{-1} \sin\theta u_\theta$. The subscripts of u represent partial derivatives (e.g. $u_x = \partial u / \partial x$ and $u_{rr} = \partial^2 u / \partial r^2$). Automatic differentiation¹⁸ enables exact and efficient calculations of partial derivatives. The powers of r are multiplied to remove singular values at the origin. The total loss function for continuous media is given by

$$L = L_{\text{PDE}} + L_{\text{DS-u}} + L_{\text{DS-T}} + L_{\text{FS}}, \quad (13)$$

and that for discontinuous media is given by

$$L = L_{\text{PDE}} + L_{\text{DS-u}} + L_{\text{DS-T}} + L_{\text{FS}} + L_{\text{MB-u}} + L_{\text{MB-T}}. \quad (14)$$

Optimization

We use the same NN structure for all examples by mainly following the original work²¹. Fully connected feedforward NNs consisting of 8 hidden layers with 40 nodes are used. The activation functions are the hyperbolic tangent function in the hidden layers and the identity function in the output layer. Xavier's initial value⁵⁰ is used as initial NN parameters. Moreover, when two or more similar problems are considered, the trained NN parameters on a simple problem are transferred to the initial NN parameters of complex problems. This can be interpreted as a variant of the curriculum learning⁴⁵, which aims at accelerating and stabilizing PINN optimization. The correspondences are listed in Table 1. The NN parameters are updated using the gradient-based algorithm Adam⁵¹ with standard learning rates ($\eta = 10^{-3}$, $\beta_1 = 0.9$, and $\beta_2 = 0.999$). PINNs do not require training data, and loss functions are calculated at arbitrary points in the model domain, which are called collocation points. The range of collocation points is set as $-5 \leq x \leq 5$, $-5 \leq y$, and the upper bound defined by the Earth's surface. The batch size is set to 256 in V and to 64 on Σ , S , and B . Collocation points are independently sampled in each training step.

The distribution of collocation points during training can have a significant influence on the model performance³⁸. The influence is inspected by observing the spatial distribution of a residual $R = r^2 u_{rr} + r u_r + u_{\theta\theta} + \mu^{-1} (r^2 \mu_r u_r + \mu_\theta u_\theta)$, the mean square of which is a loss term, L_{PDE} (Supplementary Fig. 5a–d). When an NN is trained on collocation points sampled from a uniform distribution, the residuals exhibit higher values near the fault. If collocation points are sampled from a probability distribution that concentrates on the fault, the residual is uniformly distributed. Therefore, concentrated sampling is

used in this study. Examples of collocation points in the analyzed models (Figs. 3 and 4) are shown in Supplementary Fig. 5e, f.

Because the values of a loss function can vary considerably with the sampled collocation points, fixed fine grids are prepared for model evaluation. The grid intervals are set to 0.1 in V and to 0.01 on Σ , S , and B . Minibatch training is iterated until the following conditions are satisfied: if the training loss L_{tra} on collocation points is less than 10^{-6} , the validation loss on the fixed grids L_{val} is calculated; if L_{val} is also less than 10^{-6} , training is finished (Supplementary Fig. 6).

Data availability

No data were used in this study.

Code availability

Source programs of PINN modeling are available in Supplementary Software 1. FEM solutions were generated using the open-source Python package, PyLith⁴³.

References

- Pollitz, F. F., Wicks, C. & Thatcher, W. Mantle flow beneath a continental strike-slip fault: postseismic deformation after the 1999 Hector Mine Earthquake. *Science* **293**, 1814–1818 (2001).
- Sun, T. et al. Prevalence of viscoelastic relaxation after the 2011 Tohoku-oki earthquake. *Nature* **514**, 84–87 (2014).
- Segall, P. *Earthquake and Volcano Deformation* (Princeton University Press, 2010).
- Savage, J. C. A dislocation model of strain accumulation and release at a subduction zone. *J. Geophys. Res.* **88**, 4984–4996 (1983).
- Matsu'ura, M., Jackson, D. D. & Cheng, A. Dislocation model for aseismic crustal deformation at Hollister, California. *J. Geophys. Res.* **91**, 12661–12674 (1986).
- Okada, Y. Internal deformation due to shear and tensile faults in a half-space. *Bull. Seismol. Soc. Am.* **82**, 1018–1040 (1992).
- Pollitz, F. F. Gravitational viscoelastic postseismic relaxation on a layered spherical Earth. *J. Geophys. Res.* **102**, 17921–17941 (1997).
- Smith, B. & Sandwell, D. A three-dimensional semianalytic viscoelastic model for time-dependent analyses of the earthquake cycle. *J. Geophys. Res.* **109**, B12401 (2004).
- Reches, Z., Schubert, G. & Anderson, C. Modeling of periodic great earthquakes on the San Andreas Fault: Effects of nonlinear crustal rheology. *J. Geophys. Res.* **99**, 21983–22000 (1994).
- Masterlark, T. Finite element model predictions of static deformation from dislocation sources in a subduction zone: Sensitivities to homogeneous, isotropic, Poisson-solid, and half-space assumptions. *J. Geophys. Res.* **108**, 2540 (2003).
- Freed, A. & Bürgmann, R. Evidence of power-law flow in the Mojave Desert mantle. *Nature* **430**, 548–551 (2004).
- Krizhevsky, A., Sutskever, I. & Hinton, G. E. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems* (eds Pereira, F., Burges, C. J., Bottou, L. & Weinberger, K. Q.) 1097–1105 (Curran Associates, Inc., 2012).
- Lake, B. M., Salakhutdinov, R. & Tenenbaum, J. B. Human-level concept learning through probabilistic program induction. *Science* **350**, 1332–1338 (2015).
- LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
- Yoon, C. E., O'Reilly, O., Bergen, K. J. & Beroza, G. C. Earthquake detection through computationally efficient similarity search. *Sci. Adv.* **1**, e1501057 (2015).
- Ross, Z. E., Meier, M.-A. & Hauksson, E. P wave arrival picking and first-motion polarity determination with deep learning. *J. Geophys. Res.: Solid Earth* **123**, 5120–5129 (2018).

17. Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y. & Beroza, G. C. Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nat. Commun.* **11**, 3952 (2020).
18. Baydin, A. G., Pearlmutter, B. A., Radul, A. A. & Siskind, J. M. Automatic differentiation in machine learning: a survey. *J. Mach. Learn. Res.* **18**, 1–43 (2018).
19. Berg, J. & Nyström, K. A unified deep artificial neural network approach to partial differential equations in complex geometries. *Neurocomputing* **317**, 28–41 (2018).
20. Zhu, Y., Zabaras, N., Koutsourelakis, P.-S. & Perdikaris, P. Physics-constrained deep learning for high-dimensional surrogate modeling and uncertainty quantification without labeled data. *J. Comput. Phys.* **394**, 56–81 (2019).
21. Raissi, M., Perdikaris, P. & Karniadakis, G. E. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **378**, 686–707 (2019).
22. Karniadakis, G. E. et al. Physics-informed machine learning. *Nat. Rev. Phys.* **3**, 422–440 (2021).
23. Zhu, W., Xu, K., Darve, E. & Beroza, G. C. A general approach to seismic inversion with automatic differentiation. *Computers Geosci.* **151**, 104751 (2021).
24. Waheed, U. B., Alkhalifah, T., Haghghat, E., Song, C., & Virieux, J., PINNtomo: Seismic tomography using physics-informed neural networks. Preprint at <https://arxiv.org/abs/2104.01588> (2021).
25. Rasht-Behesht, M., Huber, C., Shukla, K. & Karniadakis, G. E. Physics-informed neural networks (PINNs) for wave propagation and full waveform inversions. *J. Geophys. Res.: Solid Earth* **127**, e2021JB023120 (2022).
26. Fukahata, Y. & Matsu'ura, M. General expressions for internal deformation fields due to a dislocation source in a multilayered elastic half-space. *Geophys. J. Int.* **161**, 507–521 (2005).
27. Pollitz, F. F. Coseismic deformation from earthquake faulting on a layered spherical Earth. *Geophys. J. Int.* **125**, 1–14 (1996).
28. Mahrer, K. D. Approximating surface deformation from a buried strike-slip fault or shear crack in a mildly uneven half-space. *Bull. Seismol. Soc. Am.* **74**, 797–803 (1984).
29. McTigue, D. F. & Segall, P. Displacements and tilts from dip-slip faults and magma chambers beneath irregular surface topography. *Geophys. Res. Lett.* **15**, 601–604 (1988).
30. Williams, C. A. & Wadge, G. An accurate and efficient method for including the effects of topography in three-dimensional elastic models of ground deformation with applications to radar interferometry. *J. Geophys. Res.: Solid Earth* **105**, 8103–8120 (2000).
31. Wang, R., Lorenzo-Martin, F. & Roth, F. PSGRN/PSCMP—a new code for calculating co- and post-seismic deformation, geoid and gravity changes based on the viscoelastic-gravitational dislocation theory. *Comput. Geosci.* **32**, 527–541 (2006).
32. Sato, D. S. K., Romanet, P. & Ando, R. Paradox of modelling curved faults revisited with general non-hypersingular stress Green's functions. *Geophys. J. Int.* **223**, 197–210 (2020).
33. Ohtani, M. & Hirahara, K. Effect of the Earth's surface topography on quasi-dynamic earthquake cycles. *Geophys. J. Int.* **203**, 384–398 (2015).
34. Barbot, S. & Fialko, Y. A unified continuum representation of post-seismic relaxation mechanisms: semi-analytic models of afterslip, poroelastic rebound and viscoelastic flow. *Geophys. J. Int.* **182**, 1124–1140 (2010).
35. Barbot, S., Moore, J. D. P. & Lambert, V. Displacement and stress associated with distributed anelastic deformation in a half-space. *Bull. Seismological Soc. Am.* **107**, 821–855 (2017).
36. Ichimura, T. et al. An elastic/viscoelastic finite element analysis method for crustal deformation using a 3-D island-scale high-fidelity model. *Geophys. J. Int.* **206**, 114–129 (2016).
37. Langer, L., Gharti, H. N. & Tromp, J. Impact of topography and three-dimensional heterogeneity on coseismic deformation. *Geophys. J. Int.* **217**, 866–878 (2019).
38. Lu, L., Meng, X., Mao, Z. & Karniadakis, G. M. DeepXDE: a deep learning library for solving differential equations. *SIAM Rev.* **63**, 208–228 (2021).
39. DeVries, P. M. R., Thompson, T. B. & Meade, B. J. Enabling large-scale viscoelastic calculations via neural network acceleration. *Geophys. Res. Lett.* **44**, 2662–2669 (2017).
40. Savage, J. C. & Burford, R. O. Geodetic determination of relative plate motion in central California. *J. Geophys. Res.* **78**, 832–845 (1973).
41. Markidis, S. The old and the new: Can physics-informed deep-learning replace traditional linear solvers? *Front. Big Data* **4**, 669097 (2021).
42. Cuomo, S. et al. Scientific machine learning through physics-informed neural networks: where we are and what's next. Preprint at <https://arxiv.org/abs/2201.05624> (2022).
43. Aagaard, B. T., Knepley, M. G. & Williams, C. A. A domain decomposition approach to implementing fault slip in finite-element models of quasi-static and dynamic crustal deformation. *J. Geophys. Res.: Solid Earth* **118**, 3059–3079 (2013).
44. Yang, L., Meng, X. & Karniadakis, G. E. B-PINNs: Bayesian physics-informed neural networks for forward and inverse PDE problems with noisy data. *J. Comput. Phys.* **425**, 109913 (2021).
45. Krishnapriyan, A., Gholami, A., Zhe, S., Kirby, R. & Mahoney, M. W. Characterizing possible failure modes in physics-informed neural networks. *Adv. Neural Inf. Process. Syst.* **34**, 26548–26560 (2021).
46. Wang, S., Sankaran, S. & Perdikaris, P. Respecting causality is all you need for training physics-informed neural networks. Preprint at <https://arxiv.org/abs/2203.07404> (2022).
47. Jagtap, A. D., Kawaguchi, K. & Karniadakis, G. M. Adaptive activation functions accelerate convergence in deep and physics-informed neural networks. *J. Comput. Phys.* **404**, 109136 (2020).
48. Wang, S., Teng, Y. & Perdikaris, P. Understanding and mitigating gradient flow pathologies in physics-informed neural networks. *SIAM J. Sci. Comput.* **43**, A3055–A3081 (2021).
49. Jagtap, A. D. & Karniadakis, G. E. Extended physics-informed neural networks (XIPINNs): A generalized space-time domain decomposition based deep learning framework for nonlinear partial differential equations. *Commun. Comput. Phys.* **28**, 2002–2041 (2020).
50. Xavier, G., & Bengio, Y., Understanding the difficulty of training deep feedforward neural networks. In: *Proc. 13th International Conference on Artificial Intelligence and Statistics 9* (eds Teh, Y. W. & Titterton, M.), 249–256 (JMLR Workshop and Conference Proceedings, 2010).
51. Kingma, D. P. & Ba, J., Adam: a method for stochastic optimization. Preprint at <https://arxiv.org/abs/1412.6980> (2014).

Author contributions

T.O. designed the study, carried out the numerical modeling with PINN, and prepared the manuscript. T.I. carried out the numerical modeling with FEM. K.H. and N.U. advised the project. All authors discussed the results in the article.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-34922-1>.

Correspondence and requests for materials should be addressed to Tomohisa Okazaki.

Peer review information *Nature Communications* thanks Fred Pollitz and the other anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

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

© The Author(s) 2022