Nakata et al. Earth, Planets and Space (2016) 68:20
DOI 10.1186/s40623-016-0396-2

Earth, Planets and Space
a SpringerOpen Journal

LETTER Open Access

Geodetic inversion for spatial distribution of slip under smoothness, discontinuity, and sparsity constraints

Ryoko Nakata[1*], Tatsu Kuwatani[2], Masato Okada[3] and Takane Hori[1]

Abstract

In geodetic data inversion, insufficient observational data and smoothness constraints for model parameters make it difficult to clearly resolve small-scale heterogeneous structures with discontinuous boundaries. We therefore developed a novel regularization scheme for the inversion problem that uses discontinuity, sparsity, and smoothness constraints. In order to assess its usefulness and applicability, the proposed method was applied to synthetic displacements calculated by a ring-shaped and sharply varying afterslip distribution on a plate interface. The afterslip was obtained from reasonable numerical simulation of earthquake generation cycle with a rate- and state- dependent friction law and realistic three-dimensional plate geometry. The obtained afterslip distribution was heterogeneous, and the discontinuous boundary was sharper than that obtained by using smoothness constraint only. The same inversion test was conducted with a smoothly varying circular slip distribution with large slips inside the ring-shaped distribution. The method accurately reproduces the smooth distribution of the slip area as well as the ring-shaped distribution. Therefore, the method could be applied to any slip distribution, with both discontinuous and continuous boundaries. Adopting this method for measured data will make it possible to obtain detailed heterogeneous distributions of physical structures on fault planes. The proposed method is therefore applicable to various geophysical inversion problems that exhibit discontinuous heterogeneity.

Keywords: Smoothness constraint, Discontinuous constraint, Sparsity constraint, Regularized optimization, Geodetic inversion

Introduction

Earthquake generation is related to the heterogeneous distribution of fault slips and locking conditions on the fault planes (or plate interfaces); for example, a heterogeneous distribution is observed in coseismic slip distribution (e.g., Ide 2007), source areas of large interplate earthquakes with magnitudes greater than seven (Yamanaka and Kikuchi 2004), and small repeating earthquakes (Uchida and Matsuzawa 2011). In these cases, the slip distribution of each event varies in a piecewise smooth manner, but changes sharply across the boundaries between these heterogeneous areas and the surrounding area.

Conversely, images obtained through geodetic inversion analysis of Global Navigation Satellite System (GNSS) data in most previous studies imposed smooth slip distributions with continuous boundaries; for example, slip deficit and backslip distributions during interseismic periods (e.g., Suwa et al. 2006; Hashimoto et al. 2012; Loveless and Meade 2011), and afterslip distributions of large interplate earthquakes (e.g., Yagi et al. 2001; Miyazaki et al. 2004; Ozawa et al. 2012; Yamagiwa et al. 2015). Commonly, when conducting geodetic data inversion, available observational data are insufficient, and the smoothness constraint is treated as a priori information based on methods such as Bayes' theorem (e.g., Yabuki and Matsu'ura 1992) and Kalman filtering (e.g., Segall and Matthews 1997). In addition to convenience and traditional reasons, the smoothness constraint is derived by considering the process of earthquake rupture, where the spatial

* Correspondence: nakatar@jamstec.go.jp
[1]R & D Center for Earthquake and Tsunami, Japan Agency for Marine-Earth Science and Technology, 3173-25, Showa-machi, Kanazawa-ku, Yokohama, Kanagawa 236-0001, Japan
Full list of author information is available at the end of the article

variation of fault slip must be smooth to some degree (Yabuki and Matsu'ura 1992).

Recently, sparsity constraints have been introduced into geodetic inversions (e.g., Evans and Meade 2012). Sparse modeling is a form of statistical analysis for solving a problem by introducing the sparsity of the solution as a priori information. Evans and Meade (2012) applied the absolute value (L1) regularization approach (Tibshirani 1996) to the coseismic slip and afterslip of the 2011 Tohoku-Oki earthquake for geodetic data inversion. Using onshore GNSS observational data, compact and sharply varying afterslip distributions along the coast were obtained, and the slip areas were significantly different from other studies that used a smoothness constraint (e.g., Iinuma et al. 2012; Ozawa et al. 2012; Yamagiwa et al. 2015). The maximum slip estimated using a sparsity constraint was comparable to other studies. However, we have questions regarding the application of other shapes of slip distribution, because the estimated maximum slip was 1.7 times greater than the input value in a resolution test with a ring-shaped input slip distribution (Evans and Meade 2012).

Furthermore, another constraint focusing on long-term velocity discontinuity across faults has been introduced (Johnson and Fukuda 2010). Mixed linear–nonlinear Bayesian inversion was applied to the distributions of locked and creeping patches on faults, by incorporated a locking parameter for each patch that indicates whether a patch is completely locked (no slip) or creeping (Johnson and Fukuda 2010). If the amount and quality of the data are sufficient, introduction of such additional parameters is valid for resolving spatially complex slip distributions. However, our purpose in this study is to propose a method, which analyze rather small amounts of low-quality data from the slip on the plate interface beneath the ocean bottom. Thus, we consider a more efficient method, which requires fewer parameters to be estimated.

It is preferable to use the same method to resolve small-scale heterogeneous slip distributions composed of zero-slip (locked) areas and non-zero-slip areas with both continuous and discontinuous boundaries when we only have data of insufficient quality. In this study, we focus on regularized optimization with appropriate prior constraints for geodetic data (afterslip) inversion. We propose a novel evaluation function that uses smoothness, discontinuity, and sparsity constraints. The use of a smoothness constraint is required to reproduce a smooth distribution among non-zero-slip subfaults. The discontinuity constraint is expected to express sharp boundaries between slip and zero-slip areas. The sparsity constraint controls the number of non-zero-slip subfaults, and may be useful to obtain a minimum slip area. By combining our methods with appropriate hyperparameter selection, it is possible to obtain a smooth distribution at a continuous slip area, and a discontinuous distribution at boundaries between non-zero-slip and zero-slip areas.

Firstly, we describe evaluation functions using three constraints. Secondly, we describe the production of synthesized data for numerical tests. Then, the results are presented and discussed. Finally, we conclude with suggestions for future work.

Optimization using smoothness, discontinuity, and sparsity constraints

A numerical inversion test was conducted to assess how efficiently the proposed method can reproduce the original distribution of slip from noise-overlapped synthetic displacement data. For comparison, we present results obtained using three types of evaluation functions. First, we used an evaluation function that only included the L2 smoothness regularization term (Eq. 2) called *smoothness*. Second, we used an evaluation function that only included the L1 sparsity regularization term (Eq. 3), called *sparsity*. The third function is the proposed evaluation function that includes smoothness, discontinuity, and sparsity constraints (Eq. 4), called *SDS constraints*.

Smoothness constraints

The relationship between displacement on a free surface and slip on a plate interface is expressed as:

$$d_k = \sum_{l=1}^{N} G_{kl} s_l + \varepsilon_k \tag{1}$$

where d_k is the observed displacements at the k-th station on the Earth's surface, s_l is dip-slip of the l-th subfault on the plate boundary, N is the number of subfaults, G_{kl} is the Green's function representing displacement at station k due to unit slip on subfault l, and ε_k is the error (including observation noise) at the k-th station. We divided the plate interface into small rectangular subfaults, each of which was approximated by three triangles to calculate the angular dislocation (Comninou and Dundurs 1975). Green's functions are represented on a subfault, which is the combined effect of three angular dislocations within a subfault in an elastic, homogeneous, and isotropic half space.

If only a smoothness constraint is used, as in most previous studies involving geodetic inversion, the evaluation function can be expressed as:

$$E(\boldsymbol{s};\alpha,\beta) = \frac{\beta}{2}\sum_{k=1}^{K}\left(d_k - \sum_{l=1}^{N} G_{kl} s_l\right)^2 + \frac{\alpha}{2}\sum_{i\sim j}\left(s_i - s_j\right)^2, \tag{2}$$

where β is a precision hyperparameter, α is a smoothness hyperparameter, K is the number of observed GNSS displacements, and $\Sigma_{i\sim j}$ is the summation of all pairs of

neighboring cells. The first term on the right-hand side of Eq. 2 represents the reproducibility between model parameters and observations, and the second term represents the L2 regularization term, which indicates the smoothness of the model parameters ($\boldsymbol{s}$). Based on Kuwatani et al. (2014a), we used first-derivative regularization for the smoothness constraint. The hyperparameters, α and β, can be selected through the maximization of marginal likelihood technique proposed by Kuwatani et al. (2014a). For the numerical inversion test, we calculated the true values of these hyperparameters from the true slip $\boldsymbol{s}_t$. We refer to the true values of hyperparameters α and β as α_t $\left(\alpha_t = s_t^T C s_t / N\right)$ and $\beta_t (\beta_t = 1/\sigma^2)$, respectively. The matrix C is an $N \times N$ symmetric matrix that sums the differences of all pairs of nearest-neighbor subfaults (see Additional file 1). See Eq. (16) in Kuwatani et al. (2014a) for details. The square of the standard deviation (σ^2) used to represent the observational noise levels added to the observational data helps to determine β_t. The values α_t and β_t were compared to α and β to evaluate the validity of our method for determining hyperparameters.

Sparsity constraints

We incorporated two additional constraints as prior information into the analysis in order to accurately restore the heterogeneity of the slip distribution. One is the sparsity constraint that is derived from prior information that the slip area is considered to be smaller than the zero-slip area. In general, the subfault area is set to be wider than the target slip area, which increases the number of model parameters in the inversion analysis and enhances underdetermination of the inversion problem.

If only a sparsity constraint is used (e.g., Evans and Meade 2012; Honma et al. 2014), the evaluation function can be written as:

$$E(\boldsymbol{s};\lambda) = \sum_{i=1}^{N}\left(d_k - \sum_{l=1}^{N} G_{kl} s_l\right)^2 + \lambda \sum_{l=1}^{N} |s_l|, \tag{3}$$

where λ is a sparsity hyperparameter that controls the effect of the sparsity constraint. The second term on the right-hand side of Eq. 3 represents the sparsity constraint. A large value of λ decreases non-zero components, while λ = 0 introduces no sparsity. L1 regularization produces a compact representation of slip, and may be considered an alternative end-member to smoothed-L2 regularized solutions (Evans and Meade 2012).

To select the appropriate value of sparsity hyperparameter λ, a leave-one-out cross-validation technique can be used, and we adopted the λ value that minimized the mean squared residual (MSR) for the evaluation function with a sparsity constraint (Eq. 3). It is noted that we cannot rule out the possibility of inaccurate hyperparameter selection by the cross-validation technique, because GNSS observational data may not be sufficiently independent, causing overfitting.

Combination of smoothness, discontinuity, and sparsity constraints

As described in Introduction, if we assume only the smoothness constraint, then the estimated slip distribution must be smooth regardless of its true distribution. The evaluation function should be designed to avoid the smoothness constraint at boundaries between non-zero-slip and zero-slip areas. On the other hand, the sparsity constraint cannot reproduce a smooth distribution.

We thus propose a new evaluation function by incorporating three constraints: smoothness, discontinuity, and sparsity constraints as prior information for inversion (Fig. 1). The new evaluation function can be expressed as:

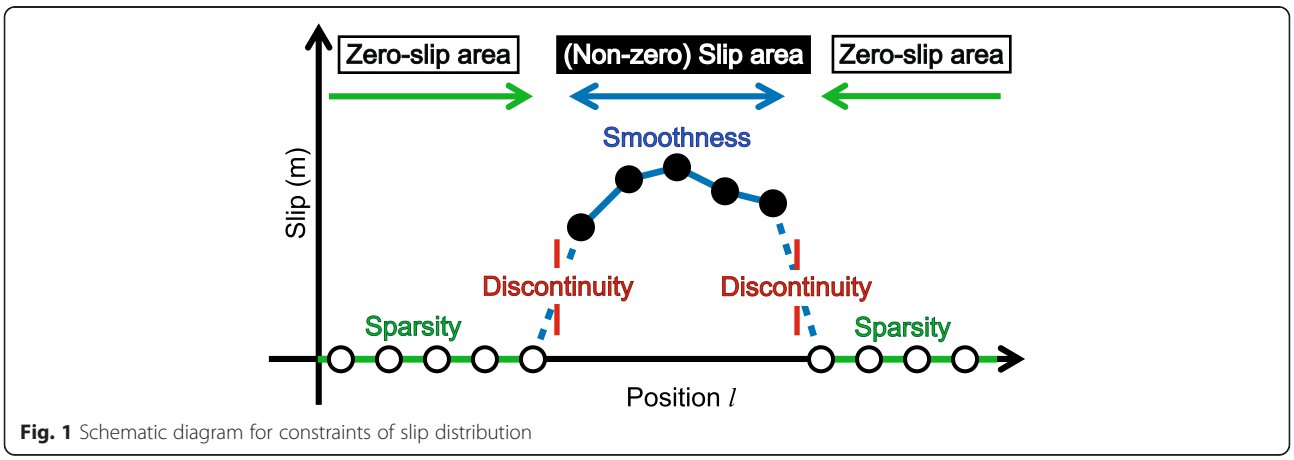

Fig. 1 Schematic diagram for constraints of slip distribution

$$E(\boldsymbol{s};\alpha',\beta',\nu)=\frac{\beta'}{2}\sum_{k=1}^{K}\left(d_k-\sum_{l=1}^{N}G_{kl}s_l\right)^2+\frac{\alpha'}{2}\sum_{i\sim j}\left(1-\delta'(s_i,s_j)\right)\left(s_i-s_j\right)^2+\nu\sum_{l=1}^{N}|s_l| \quad (4)$$

where β' is a precision hyperparameter, α' is a smoothness hyperparameter for the non-zero-slip area, ν is a sparsity hyperparameter, and $\delta'(s_i, s_j)$ behaves like a delta function, and is defined as follows:

$$\delta'(s_i,s_j)=\begin{cases}1, & \text{if } s_i=0 \text{ or } s_j=0\\ 0, & \text{if } s_i\neq 0 \text{ and } s_j\neq 0\end{cases}. \quad (5)$$

The proposed evaluation function in Eq. 4 consists of three terms: the first term on the right-hand side is the reproducibility between model parameters and observations, and the second term constrains the smoothness and discontinuity of model parameters ($\boldsymbol{s}$). For the neighboring pair, if there is zero slip in both cells, or if either of the neighboring cells is zero, then the function $\delta'(s_i, s_j)$ becomes equal to one; this results in the removal of the smoothness constraint term. The slip of both the neighboring cells must be non-zero for the smoothness constraint to be effective. Thus, the coefficient $(1\text{-}\delta'(s_i, s_j))$ serves as the cutoff tool for the smoothness constraint at the boundary between non-zero and zero-slip areas. The third term on the right-hand side of Eq. 4 represents the sparsity constraint.

The model parameter ($\boldsymbol{s}$), which minimizes the evaluation function in Eq. 4, is expected to be the best solution in terms of satisfying the reproducibility of the observation, the continuity of the slip with a discontinuous boundary, and the sparsity of a slip area. We take possible candidate sets of the model parameters in order to minimize the evaluation function in Eq. 4, i.e., the probability distribution function of the posterior probability $p(\boldsymbol{s}|\boldsymbol{d};\alpha',\beta',\nu)$ in terms of the Bayesian estimation. Following Kuwatani et al. (2014b), we used the Metropolis algorithm (Metropolis et al. 1953), which is a type of Markov chain Monte Carlo (MCMC) method. As a solution for the evaluation function in Eq. 4, we present a posterior mean (PM) solution defined by the mean value of a number of candidate sets, which are sampled using the posterior distribution. For all calculations, we set initial values of model parameters ($\boldsymbol{s}$) equal to zero. We also checked whether the estimated values were independent of the initial values of the model parameters.

The hyperparameters, α', β', and ν in the evaluation function in Eq. 4 control the behavior of the estimated model parameter ($\boldsymbol{s}$). The hyperparameters must be appropriately determined prior to minimization of the evaluation function in order to estimate accurate model parameters. Because of nonlinearity within the evaluation function (Eq. 4), the maximization of marginal likelihood technique for all three hyperparameters is computationally high-cost and non-stable. Therefore, the hyperparameters were appropriately selected as described below.

The hyperparameter selections were conducted in three steps. First, we selected λ by cross-validation, then we determined α' and β'. Finally, we calculated ν using the value of β' and λ. To determine the smoothness and precision hyperparameters, α' and β', we only used non-zero model parameters ($\boldsymbol{s}$) estimated by Eq. 3, so that the effect of the pre-smoothness coefficient $(1\text{-}\delta'(s_i, s_j))$ can be ignored. We selected the hyperparameters assuming that the two parameters, α' (smoothness for non-zero subfaults) and λ (number of zero subfaults), are almost uncorrelated. The assumption is considered to be satisfied in the present study, because α' is related to only non-zero-slip subfaults and λ is related to zero-slip subfaults. The sparsity hyperparameter ν was calculated from the sparsity hyperparameter λ of Eq. 3, by multiplying the coefficient of the reproducibility $\frac{\beta'}{2}$ of the proposed evaluation function (Eq. 4); this is because $(\nu = (\beta'/2)\cdot\lambda)$. In addition to α_t and β_t, the true values of hyperparameters α' and β' are referred to as α'_t and β'_t, respectively. In order to examine the validity of the selection of three hyperparameters (α', β', and ν), we check if α' and β' are comparable to α'_t and β'_t, respectively.

Synthetic displacements as observational data

Slip distribution resulting from realistic numerical simulation

We assessed the accuracy of our proposed method using example data obtained in previous studies (Nakata et al. 2012; 2014) (Table A1, see Additional file 1 for what was done in each study). The data were obtained by simulating an earthquake generation cycle based on the rate- and state-dependent friction law with a realistic three-dimensional plate geometry (Nakata et al. 2012). We conducted numerical simulations of large interplate earthquakes and afterslips beneath the Hyuganada offshore region, southwest Japan. Source areas of large earthquakes were approximated to represent a circle of 10 km radius that was assumed to be frictionally heterogeneous. Frictional parameters at the boundary

between the seismic source area and its surrounding area were discontinuous and sharp. This resulted in an earthquake with a moment magnitude (Mw) of 6.8 with no afterslip within a 10 km radius of the seismic source (Fig. 2a and 2b). The distribution of the frictional property on the plate interface was reflected in the afterslip distribution during the simulation. Thus, the area of the simulated afterslip enclosed the seismic source of the Mw 6.8 earthquake, and exhibited a ring-shaped afterslip distribution with a sharp boundary (Nakata et al. 2012). The afterslip depth-profile crossing the center of the seismic source area showed two peaks, and the local minimum corresponded to the large coseismic slip area (Fig. 2c, gray and black lines). In addition, the boundaries were sharp between the locked area (zero and nearly-zero-slip within the afterslip area) and the afterslip area, and between the afterslip area and the outlying areas.

A similar ring-shaped distribution of afterslip was estimated from the observed data; following the Tokachi-Oki earthquake of Mw 8.0 in 2003, a large afterslip surrounded the coseismic slip area, and a continuous boundary was inferred (Miyazaki et al. 2004). Thus, the ring-shaped distribution is not just a characteristic of numerical simulations, and is considered a typical spatial distribution of afterslip and an example of discontinuous heterogeneity on the plate interface. Hence, here we use the ring-shaped afterslip distribution of Nakata et al. (2012) as observation data. We refer to the afterslip as true slip (s_t).

Nakata et al. (2014) numerically reproduced the spatial and temporal distribution of a simulated afterslip on a plate interface following a Mw 6.8 earthquake using synthetic displacements as observed data, and a Network Inversion Filter based on Kalman filtering methods (Segall and Matthews 1997). The estimated afterslip distribution demonstrated a continuous and smooth, but weakly heterogeneous distribution that roughly corresponded to the coseismic slip area. The maximum slip was less than the true value, and the local minimum within the afterslip area was nearly half that of the surrounding peak values (Fig. 2c, blue line), although no afterslip was observed in the forward simulation results in Nakata et al. (2012) (Fig. 2c, black line). The

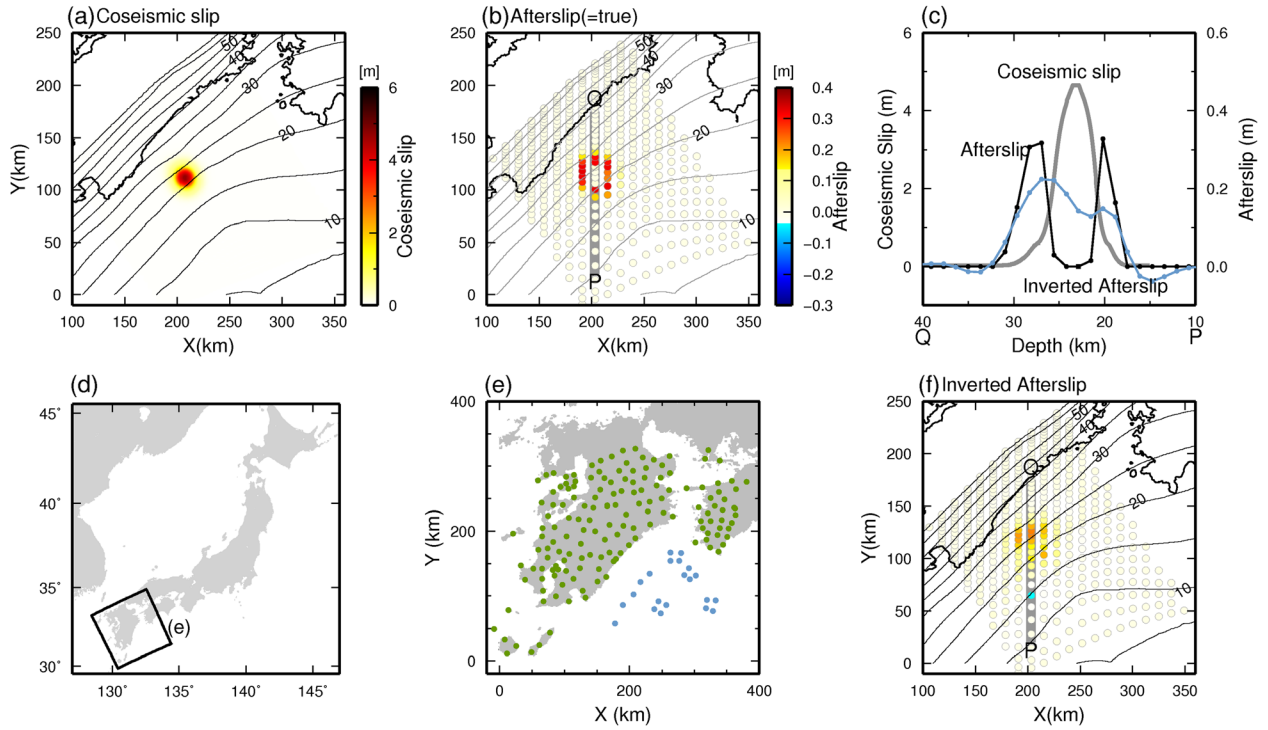

Fig. 2 **a** Coseismic slip distribution in the *X*–*Y* plane from Nakata et al. (2012). Contours indicate depth (in km) of the upper surface of the descending plate (Baba et al. 2002). **b** True image of afterslip distribution in the *X*–*Y* plane for a duration of 80 days. Subfaults were coarsely relocated from Nakata et al. (2012). **c** Depth profile along *X* = 205 km (*gray vertical* shadow PQ in Fig. 2b). *Gray line* and the left axis show coseismic slip (Fig. 2a). *Black line* with dots and the right axis show afterslip (true slip, Fig. 2b). *Blue line* shows afterslip estimated from previous inversion analysis (Nakata et al. 2014) (Fig. 2f). **d** Map of Japan: the inset represents the study area shown in Fig. 2e. **e** Station distribution used for calculating displacements. Existing GNSS stations (*green dots*, 156 stations) and planned stations on ocean floor (*blue dots*, 21 stations). **f** Afterslip distribution in the *X*–*Y* plane estimated from our previous inversion analysis using stations shown in Fig. 2e (Nakata et al. 2014). The maps were created using the software Generic Mapping Tools (Wessel and Smith 1998)

discrepancy between the results of the forward simulation and the inversion was likely caused by an excessive smoothness constraint in the inversion process.

Synthetic observations

To assess the accuracy of our proposed methods (Eq. 4), we used synthetic displacements calculated from two patterns of slip distribution. One is the ring-shaped distribution (Fig. 2b) as an example of discontinuous heterogeneity on the plate interface. The other (Fig. 3a and b) is a smoothly varying circular slip distribution as an example image obtained through previous inversion studies. It was prepared by filling in zero-slip areas of the ring-shaped slip distribution. To prepare the smoothly varying circular distribution, we set 0.33 m of slip at the central seven subfaults (X = 200–210 km, 14–23 km depth) in the true ring-shaped slip distribution. We then calculated the moving average at every nine subfaults for the entire fault area. Slip values were normalized to the maximum slip of the ring-shaped distribution. The universal applicability of the proposed method is demonstrated by the results obtained for these different slip distributions.

Similarly to Nakata et al. (2014), we calculated synthetic displacements on the free surface using results from these two slips. This was calculated using the position of existing GNSS stations operated by the Geospatial Information Authority of Japan, and the proposed positions of stations that are planned for future installation as a part of the ocean floor cable network. The latter were included to increase spatial resolution. In all, we used 156 existing continuous GNSS stations on land and 21 stations on the ocean floor (Fig. 2e). In this case, K is equal to 531 as there are three components for 177 stations (note that this is not an underdetermined problem, but there are some low quality data from stations far from the afterslip area), and N is equal to 436, as dip-slip only occurs for 436 subfaults (To shorter the calculation time, we did not consider strike-slip on the fault plane). The parameters N and the sizes of the subfaults were the same as in our previous study (Nakata et al. 2014; 12.15 km in the X direction in Fig. 2).

In order to represent observation noise in the synthetic displacement data, we added independent random numbers that followed a normal distribution (mean = 0, standard deviation = 0.3 mm). This was carried out for both the horizontal and vertical components. To avoid the difficulty resulting from randomness in the displacement data, the noise level was slightly less than that of the observed data. For comparison, the synthetic displacements of all components ranged from -3.80 to 2.28 cm. For simplicity, we did not distinguish between stations on land and on the ocean floor or between horizontal and vertical components, and used the same noise levels for all stations; note that the noise level differs from Nakata et al. (2014), and no noise was included in the results shown in Fig. 2c and 2f. We estimated a slip distribution on the plate interface using synthetic displacement data as observed data, which were obtained from the simulation results of Nakata et al. (2012).

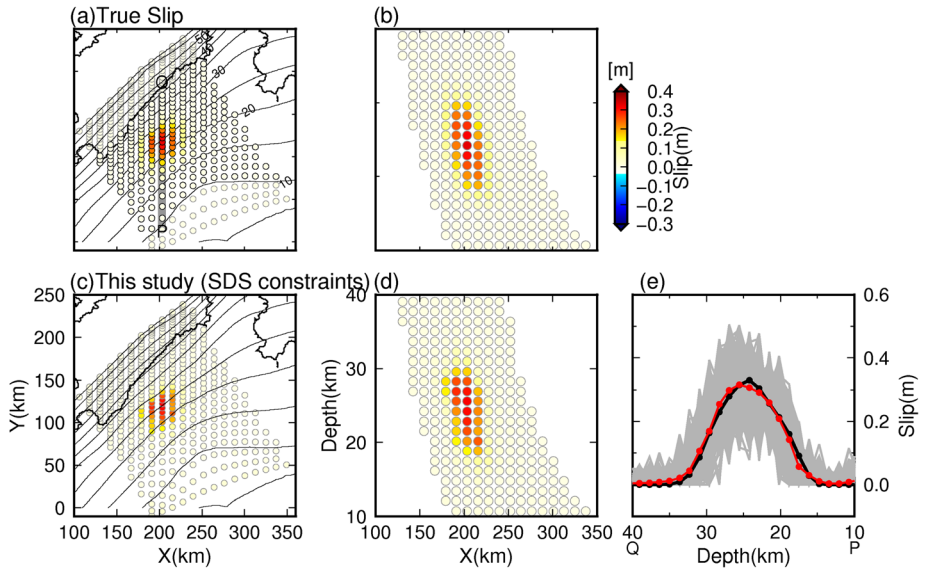

Fig. 3 a–b True slip distribution as an example of smoothly varying distribution. **c–e** As in Fig. 4g–i, but using smoothly varying circular distribution (Fig. 3a)

Results and discussion

For the ring-shaped slip distribution test using the smoothness regularization, the estimated spatial distribution showed a wider slip area with a smoother boundary compared to the true slip distribution (Fig. 4a and b). In addition, there were some falsely imaged subfaults with negative or positive slips at the edge grids of the modeled region. As observed in the previous inversion by Nakata et al. (2014), by using the smoothness constraints, it was not possible to resolve the two peaks seen in the depth profile of the true slip distribution (Fig. 4c). The estimated maximum slip was approximately 40 % less than the true value. Local minimum slip within the slip area was significantly larger than the true values that are zero, or are approximately at the zero level. In addition, the boundary between the slip area and the outside area is too smooth, and small negative slips are falsely identified around the boundary. These discrepancies between the estimations and the true value result from excessive homogenization of the original heterogeneity of the slip area, which in turn results from the smooth regularization.

For the sparsity constraint, we tested λ within the range 1×10^{-11} to 1×10^{-2}. The MSR was minimized when $\lambda = 1.6 \times 10^{-6}$ (Figure A1, see Additional file 1). As a result, the spatial distribution estimated by the sparsity

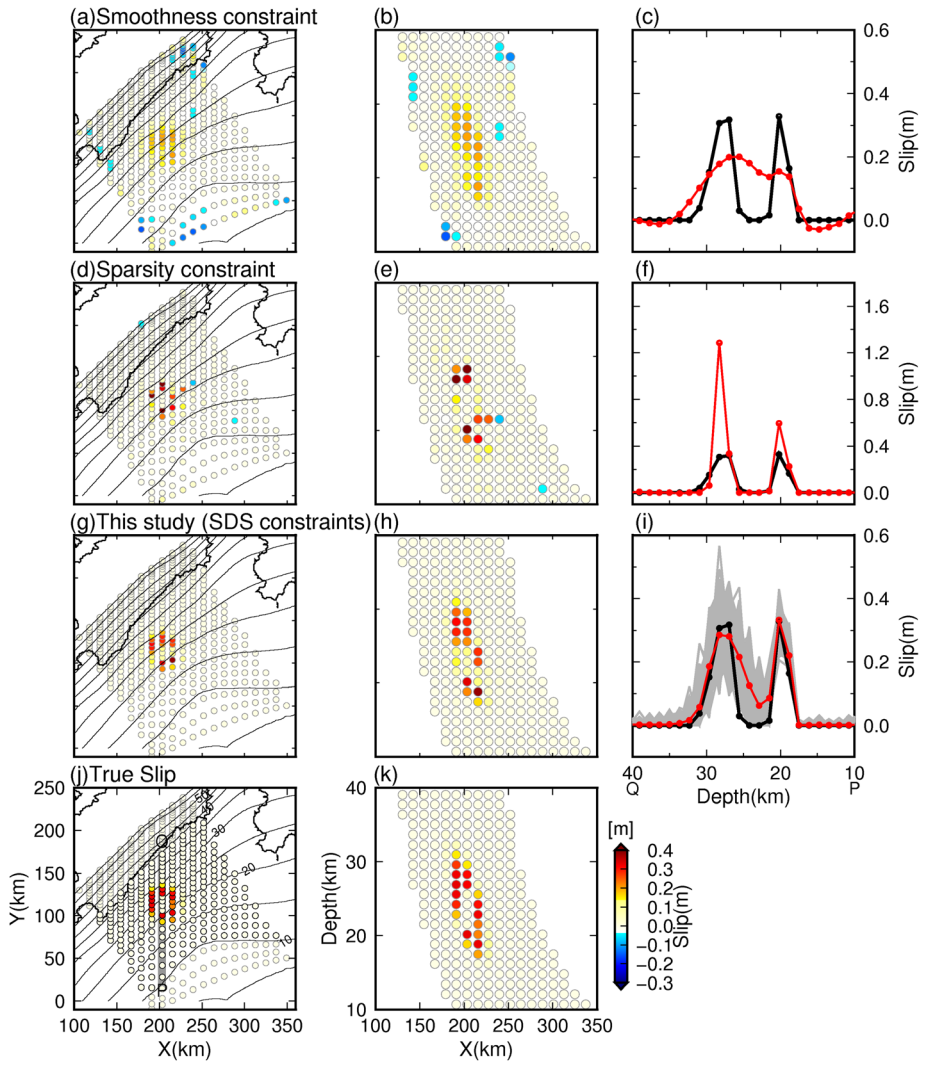

Fig. 4 Slip distributions on the plate interface estimated from synthetic displacement data of ring-shaped slip distribution. **a** Afterslip distribution in the *X*–*Y* plane obtained using the smoothness constraint. **b** Afterslip distribution in the *X*–depth plane obtained using the smoothness constraint. **c** Depth profile along $X = 205$ km obtained using the smoothness constraint: red and black lines with dots represent results using smoothness constraint and true slip, respectively. **d**–**f** As for (**a**)–(**c**), but using sparsity constraint. **g**–**i** As for (**a**)–(**c**), but using SDS constraints. Red line shows PM solution of the MCMC calculations; gray line shows the top 5% of solutions that minimize the evaluation function (Eq. 4). **j**–**k** As for (**a**)–(**b**), but showing true slip distribution. Figure 4j is the same as that of Fig. 2b

regularization showed a jagged slip area compared to the true image (Fig. 4d, e, and f). However, the main slip area was roughly similar to the true slip distribution, when compared with the smoothness estimation. Several subfaults with non-zero-slip true values were mistakenly estimated as zero-slip subfaults and vice versa. In the depth profile (Fig. 4f), the sparsity estimation successfully resolved the two peaks present in the true slip distribution, and the coseismic slip area was successfully estimated to be almost zero-slip. However, the estimated slip amplitudes were less accurate. The estimated maximum slip was one-and-a-half times greater than the value of true slip in the case of the shallower peak, while it was five times larger than the value of true slip for the deeper peak. These characteristics are attributed to the sparsity constraint greatly amplifying the original heterogeneity.

The spatial distribution estimated using the SDS constraints exhibited a slip area of similar shape and extent to that of the true image (Fig. 4g, h, and i). Additionally, there were a few artificial false slips, whereas the results from the smoothness and sparsity constraint showed several false slips outside each afterslip area (Fig. 4a and d). The two peaks present in the depth profile were much more clearly resolved compared to those obtained via smoothness regularization. Slip in the coseismic slip area was estimated to be a non-zero value, which was smaller than that estimated through smoothness regularization. The maximum slip values of both peaks were approximately equal to the true values.

Numerous model runs of the SDS inversion were conducted with different noise patterns. These were generated from different random number seeds and initial values of model parameters. As a result, the maximum a posteriori (MAP) solution by the MCMC calculations is dependent on the observation noise, initial values of model parameters, and the sampling process in the MCMC method. In contrast, PM solutions are independent of initial settings, randomness of noise, and sampling process. In the depth profile in Fig. 4i, several solutions of the top 5% of solutions that minimize the evaluation function (Eq. 4) showed a few zero- and nearly-zero-slip subfaults within the coseismic slip area (gray lines in Fig. 4i).

The root mean square errors (RMSE) of the model parameters, defined as $\sum_{l=1}^{N}\sqrt{(s_l - s_t)^2/N}$, for smoothness, sparsity, and SDS estimations were 0.0462, 0.0643, and 0.0195 m, respectively. Based on the characteristics described above and the lowest RMSE, the SDS estimation provides the best solution for model parameters among the three evaluation functions. This is attributed to the features of the SDS constraints that are inherited from smoothness regularization and sparsity regularization. SDS constraints can accurately reproduce the heterogeneous distributions of slip areas that have smoothness of slip, a discontinuous boundary of zero-slip area, and sparsity.

The Mw of estimated afterslip with smoothness, sparsity, and SDS constraints were 6.87, 6.87, and 6.90, respectively. The true value was calculated to be 6.86. Although, here we used a large mesh size, the moment magnitude in the SDS estimation was a reasonable value without underestimation.

For the SDS constraints, hyperparameters α', β', and ν were selected to be 39, 8.2×10^6, and 6.6, respectively. The true values of hyperparameters α_t' and β_t' were 22 and 1.0×10^7, respectively. The determined value of β' was similar to the true value, and the determined α' value, which is known to be difficult to determine precisely (e.g., Kuwatani et al. 2014a; Nakanishi-Ohno et al. 2014), was within the same order of magnitude as the true value. Compared to α_t' and β_t', the hyperparameter selection of α' and β' were apparently successful. The numbers of zero-slip and nearly-zero-slip subfaults were approximately similar to the true numbers of subfaults, which indicates that the selected values of sparsity hyperparameters λ and ν are valid. For the smoothness constraint, the determined hyperparameters were $\alpha = 575$ and $\beta = 4.7 \times 10^6$, whereas $\alpha_t = 196$ and $\beta_t = 1.0 \times 10^7$.

To test the validity of the selected hyperparameter values, we conducted SDS analysis using many sets of α' and ν values. The tested values were multiples of hyperparameters α' and ν resulting from the ring-shaped distribution (Fig. 4g). The tests were conducted within the range of $10^{-2} \times \alpha'$ to $10^2 \times \alpha'$, and $10^{-4} \times \nu$ to $10^2 \times \nu$. As a result, the RMSE was small, with values in and around $1 \times \alpha'$ and $1 \times \nu$ (Fig. 5). In particular, the RMSE was sensitive to the variation of α'. Although there is room for improvement in selecting hyperparameters, the results

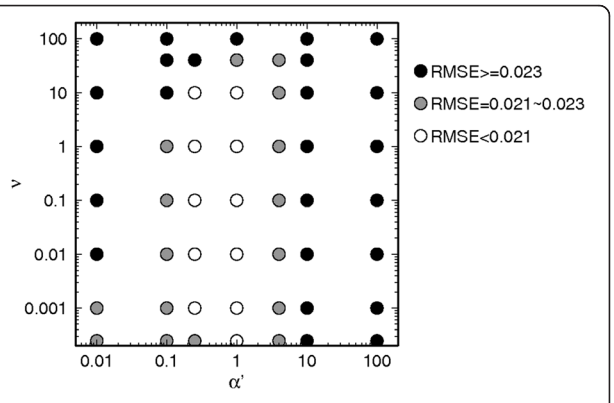

Fig. 5 RMSE (m) for various values of α' and ν using SDS constraints for the ring-shaped slip distribution

confirm that our technique is appropriate for hyperparameter selection of α' and ν (and β' and λ).

We also conducted the inversion test with a smoothly varying circular slip distribution (Fig. 3). The SDS constraints accurately reproduced the slip values and the smooth distribution of the slip area; the RMSE was 0.0129 m, and there were no false slips. On the other hand, the smoothness evaluation function (Eq. 2) reproduced the slip area and values of slip as accurately as the SDS analysis, but showed many false slips at the edge grids of the modeled region. As shown in Figs. 3d, e, and 4h, i, our SDS method could reproduce both discontinuously and continuously varying slip distributions without any additional parameters. SDS analysis is therefore able to resolve small-scale heterogeneous structures.

We assumed seafloor stations and analyzed a small Mw of afterslip with smaller measurement errors than that of realistic data. When we set the standard deviation to be comparable to realistic data sets, or used displacement data without seafloor stations, the SDS constraints, as well as the other constraints, could not resolve two peaks in the depth profile. However, when the Mw of an afterslip is larger (Mw > 7.4) and the measurement error is comparable to realistic data, the SDS method gave better estimations than other methods.

Conclusions and future work

We demonstrated that the spatial distribution of both ring-shaped and smoothly varying circular slip distributions can be reproduced by the same method. Our method was optimized using constraints for smoothness, discontinuity, and sparsity, when we only have data of insufficient quality. The slip distribution that minimized the evaluation function better reproduced images in terms of both area and amplitude compared to other methods. In contrast, experiments in which only the sparsity constraint was used overestimated slip amplitudes and underestimated the slip area. Using the SDS constraints proposed here, we were able to avoid the difficulty caused by underdetermined problems, and clearly resolve the small heterogeneity with a discontinuous boundary, even if the grid size was not sufficiently small.

By using SDS constraints, the discontinuous boundary between the heterogeneous area and the surrounding area could be clearly visualized without any false slips at the plate edge. The heterogeneous distribution of the slip controls the processes of stress accumulation and release on a plate interface. In particular, a discontinuous boundary indicates areas of stress concentration, which lead to initiation of rupture. Further, it is critical to understand the time of occurrence of earthquakes, the rupture patterns, and nucleation processes, and to forecast slip transition in the future. The distribution of frictional properties on a plate interface were reflected in the slip distribution. Future work should aim to obtain precise slip distribution on the plate interface from measured observational geodetic data. More precise estimates of slip distribution may be helpful in monitoring plate coupling, and for making improvements in our forward simulation model.

The proposed inversion method is applicable not only to geodetic inversions, but also to other geophysical and geochemical inversions such as seismic tomography, electromagnetic tomography, and pressure–temperature inversion of metamorphic rocks. Using our method, it is possible to obtain images that are clearer, sharper, and show appropriate smoothness. Moreover, this method may be particularly effective for underdetermined problems, in which observational data prove to be insufficient for estimating model parameters. We will develop more efficient and high-speed methodology in order to expand the applicability of the SDS method.

Additional file

Additional file 1: This file includes explanation of the matrix C, Table A1, and Figure A1. Table A1. What was done in Nakata et al. (2012), Nakata et al. (2014), and this study. Figure A1. MSR and λ obtained from cross-validation technique. We adopt the λ value that minimizes the mean squared residual for the evaluation function (Eq. 3). We used the value indicated by the red star. (PDF 184 kb)

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

RN carried out the calculation and drafted the manuscript. TK programed the main part of the code, participated in the design of the study, and helped draft the manuscript. MO participated in the design of the study. TH conceived and participated in the design of the study. All authors read and approved the final manuscript.

Acknowledgements

We thank Dr. K. Nagata for his insightful comments regarding the optimization algorithm. We also thank Dr. S. Ikeda and Dr. M. Honma for discussions regarding the super-resolution imaging of a black hole (Honma et al. 2014). This work was partly supported through the "Initiative for High-Dimensional Data-Driven Science through Deepening of Sparse Modeling" by MEXT KAKENHI (Grants No. 26120538, No. 25120005, and No. 25120009) and JSPS Grants No. 25280090, and No. 15K20864. We used the positions of Global Navigation Satellite System (GNSS) stations operated by the Geospatial Information Authority of Japan. Generic Mapping Tools (Wessel and Smith 1998) was used to draw all figures. Comments from J. Loveless, two anonymous reviewers, and the editor, Y. Aoki, were constructive and helped improve this manuscript.

Author details

[1]R & D Center for Earthquake and Tsunami, Japan Agency for Marine-Earth Science and Technology, 3173-25, Showa-machi, Kanazawa-ku, Yokohama, Kanagawa 236-0001, Japan. [2]Department of Solid Earth Geochemistry, Japan Agency for Marine-Earth Science and Technology, 2-15, Natsushima-cho, Yokosuka, Kanagawa 237-0061, Japan. [3]Graduate School of Frontier Sciences, University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa-shi, Chiba 277-8561, Japan.

Received: 31 July 2015 Accepted: 23 January 2016
Published online: 11 February 2016

References

Baba T, Tanioka Y, Cummins PR, Uhira K (2002) The slip distribution of the 1946 Nankai earthquake estimated from tsunami inversion using a new plate model. Phys Earth Planet Inter 132:59–73

Comninou M, Dundurs J (1975) Angular dislocation in a half space. J Elast 5:203–216. doi:10.1007/BF00126985

Evans EL, Meade BJ (2012) Geodetic imaging of coseismic slip and postseismic afterslip: Sparsity promoting methods applied to the great Tohoku earthquake. Geophys Res Lett 39:L11314. doi:10.1029/2012GL051990

Hashimoto C, Noda A, Matsu'ura M (2012) The Mw 9.0 northeast Japan earthquake: total rupture of a basement asperity. Geophys J Int 189:1–5. doi:10.1111/j.1365-246X.2011.05368.x

Honma M, Akiyama K, Uemura M, Ikeda S (2014) Super-resolution imaging with radio interferometer using sparse modeling. Publ Astron Soc Jpn 66:5. doi:10.1093/pasj/psu070

Ide S (2007) Slip inversion. In: Kanamori H (ed) Treatise on Geophysics, Vol. 4: Earthquake seismology. Elsevier, Amsterdam, pp 193–224. ISBN 978-0-444-51932-0

Iinuma T, Ohzono M, Ohta 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. J Geophys Res 117:B07409. doi:10.1029/2012JB009186

Johnson KM, Fukuda J (2010) New methods for estimating the spatial distribution of locked asperities and stress-driven interseismic creep on faults with application to the San Francisco Bay Area, California. J Geophys Res 115: B12408. doi:10.1029/2010JB007703

Kuwatani T, Nagata K, Okada M, Toriumi M (2014a) Markov-random field modeling for linear seismic tomography. Phys Rev E 90:042137

Kuwatani T, Nagata K, Okada M, Toriumi M (2014b) Markov-random field modeling for mapping geofluid distributions from seismic velocity structures. Earth Planets Space 66:5. doi:10.1186/1880-5981-66-5

Loveless JP, Meade BJ (2011) Spatial correlation of interseismic coupling and coseismic rupture extent of the 2011 MW = 9.0 Tohoku-oki earthquake. Geophys Res Lett 38:L17306. doi:10.1029/2011GL048561

Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equation of state calculations by fast computing machines. J Chem Phys 21(6):1087–1092

Miyazaki S, Segall P, Fukuda J, Kato T (2004) Space time distribution of afterslip following the 2003 Tokachi-oki earthquake: Implications for variations in fault zone frictional properties. Geophys Res Lett 31:L06623. doi:10.1029/2003GL019410

Nakanishi-Ohno Y, Nagata K, Shouno H, Okada M (2014) Distribution estimation of hyperparameters in Markov random field models. J Phys A 47:045001. doi:10.1088/1751-8113/47/4/045001

Nakata R, Hyodo M, Hori T (2012) Numerical simulation of afterslips and slow slip events that occurred in the same area in Hyuga-nada of southwest Japan. Geophys J Int 109:1213–1220. doi:10.1111/j.1365-246X.2012.05552.x

Nakata R, Miyazaki S, Hyodo M, Hori T (2014) Reproducibility of spatial and temporal distribution of aseismic slips in Hyuga-nada of southwest Japan. Mar Geophys Res 35:311–317. doi:10.1007/s11001-013-0100-z

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. J Geophys Res 117:B07404. doi:10.1029/2011JB009120

Segall P, Matthews M (1997) Time dependent inversion of geodetic data. J Geophys Res 102(B10):22391–22409

Suwa Y, Miura S, Hasegawa A, Sato T, Tachibana K (2006) Interplate coupling beneath NE Japan inferred from three-dimensional displacement field. J Geophys Res 111:B04402. doi:10.1029/2004JB003203

Tibshirani R (1996) Regression shrinkage and selection via the Lasso. J Roy Stat Soc B 58(1):267–288

Uchida N, Matsuzawa T (2011) Coupling coefficient, hierarchical structure, and earthquake cycle for the source area of the 2011 off the Pacific coast of Tohoku earthquake inferred from small repeating earthquake data. Earth Planets Space 63:675–679. doi:10.5047/eps.2011.07.006

Wessel P, Smith WHF (1998) New, improved version of Generic Mapping Tools released. EOS Trans Am Geophys Union 79:579–579. doi:10.1029/98EO00426

Yabuki T, Matsu'ura M (1992) Geodetic data inversion using a Bayesian information criterion for spatial distribution of fault slip. Geophys J Int 109:363–375

Yagi Y, Kikuchi M, Sagiya T (2001) Co-seismic slip, post-seismic slip, and aftershocks associated with two large earthquakes in 1996 in Hyuga-nada, Japan. Earth Planets Space 53:793–803

Yamagiwa S, Miyazaki S, 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. Geophys Res Lett 42:66–73. doi:10.1002/2014GL061735

Yamanaka Y, Kikuchi M (2004) Asperity map along the subduction zone in northeastern Japan inferred from regional seismic data. J Geophys Res 109:B07307. doi:10.1029/2003JB002683