Geodetic inversion for spatial distribution of slow earthquakes under sparsity constraints

Takane Hori¹, Ryoko Nakata¹, Hideitsu Hino², Tatsu Kuwatani³,⁴, Shoichi Yoshioka², Masato Okada⁶

1. Research and Development Center for Earthquake and Tsunami, Japan Agency for Marine-Earth Science and Technology, Japan
2. Department of Computer Science, University of Tsukuba, Japan
3. Department of Solid Earth Geochemistry, Japan Agency for Marine-Earth Science and Technology, Japan
4. PRESTO, Japan Science and Technology Agency, Japan
5. Research Center for Urban Safety and Security, Kobe University, Japan
6. Graduate School of Frontier Sciences, The University of Tokyo, Japan

E-mail: horit@jamstec.go.jp

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 applied sparse modelling to geodetic data inversion. In this paper, we reported two examples: one is 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 the plate interface beneath the Hyuga-nada region in southwest Japan. The discontinuous boundary was sharper than that obtained by using smoothness constraint only. The other is used the fused regularization, a type of sparse modelling suitable for detecting discontinuous changes in the model parameters. We estimated spatial distribution of the long-term slow slip events beneath the Bungo channel in southwest Japan. We found that the largest slip abruptly becomes zero at the down-dip limit of the seismogenic zone, and is immediately reduced to half at the up-dip limit of the deep low-frequency tremors, and becomes zero near its down-dip limit. Such correspondences imply that some thresholds exist in the generation processes for both tremors and SSEs. These results suggest that geodetic data inversion with sparse modelling can detect such abrupt changes and discontinuous boundaries in the slip distribution of slow earthquakes.

1. Introduction

Earthquake generation is related to the heterogeneous distribution of fault slips and locking conditions on the fault planes, including plate interfaces. The slip distribution of each earthquake 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. 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 [1] and Kalman filtering [2]. The smoothness constraint is introduced into the geodetic analysis because drastic and abrupt changes in the total slip distributions are the result of excessive concentration of stress, which does not seem to be reasonable considering the process of earthquake rupture. However, slip distributions obtained from geodetic data are significantly smoother than those from seismic data. It is suggested that heterogeneous distribution of slip and corresponding stress concentration may be removed from its original distribution by the prior constraint in geodetic data analyses. This is a critical problem to accomplish our purpose that we estimate preparation process of earthquake generation and forecast the following slip process using crustal deformation data during inter-seismic periods. Therefore, we would like to extract information about slip distribution on a plate interface with adequate extent of variation between smoothness and heterogeneity from observation data. This corresponds to the “Computational theory” level, in Figure 1 [3]. For this purpose, we should create numerical model and represent algorithm for analysis. In data-driven science, we clearly distinguish between these two levels, "modeling" and "representation and algorithm" [3]. In this paper, we introduce two examples with different approaches to the different purpose of “computational theory” for seemingly two similar problems. The first example is directly represented above, and our goal is to extract adequately both smoothness and abrupt change in slip distribution to the extent of observation data quality. We analyze afterslip on a plate interface following a moment magnitude (Mw) 6.8 earthquake beneath the Hyuga-nada offshore region [4]. On the other hand, the goal of the second example is to determine the presence of abrupt changes and detect their locations without detail slip distribution. Here, we analyze long-term slow slip events (L-SSEs) beneath the Bungo Channel [5].

\[
\begin{align*}
\text{Computational Theory} & \quad \text{(specific fields of natural science, measurement science)} \\
\text{The goal and its appropriateness of data analysis is} & \quad \text{discussed to establish a logic of applicable strategy} \\
\text{Modeling} & \quad \text{(theoretical physics, mathematical science, statistical)} \\
\text{Based on the computational theory, the system is} & \quad \text{formulated mathematically} \\
\text{Representation and algorithm} & \quad \text{(statistics, machine learning, computer science)} \\
\text{Algorithms are developed to carry out computational} & \quad \text{problems offered by modeling} \\
\end{align*}
\]

**Figure 1.** Three levels of data-driven science [3].

**2. Geodetic data inversion**

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

\[
d_k = \sum_{l=1}^{N} G_{kl}s_l + \varepsilon_k \quad (k = 1, 2, \ldots, K),
\]

where \(K\) is the number of observed GNSS displacements, \(d_k\) is the observed displacements at the \(k\)-th station on the Earth’s surface, \(s_l\) is the dip-slip of the \(l\)-th subfault on the plate boundary, \(N\) is the number of subfaults, \(G_{kl}\) is Green’s function, which represents the displacement at station \(k\) due to a unit slip on subfault \(l\), \(\varepsilon_k\) is the error (including observation noise) at the \(k\)-th station.

We divided the plate interface into small rectangular subfaults. Green’s functions, which are the combined effect of three angular dislocations within a subfault in an elastic, homogeneous, and
isotropic half space, were represented on a subfault. In southwest Japan, the Philippine Sea plate is subducting beneath the Amur plate. We used a realistic three-dimensional geometry of the Philippine Sea plate [6]. For simplicity, only dip-slips on the fault plane were considered. To shorten the calculation time, we did not consider strike-slip on the fault plane.

3. Afterslip analyses under smoothness, discontinuity, and sparsity constraints

We assessed the accuracy of our proposed method using example data obtained by simulating an earthquake generation cycle based on the rate- and state-dependent friction law [7]. In the numerical simulation, 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 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 [7]. 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 (Figure 2). 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.

Figure 2. Slip distribution of the simulated coseismic slip (a) and afterslip (b) of a Mw6.8 earthquake. (c) Depth profile along X = 205 km (gray vertical shadow PQ in Figure 2b). Gray line and the left axis show coseismic slip (Figure 2a). Black line with dots and the right axis show afterslip (true slip, Figure 2b). Blue line shows afterslip estimated from previous inversion analysis [8].

3.1. Methods

We proposed a new evaluation function by incorporating three constraints: smoothness, discontinuity, and sparsity (SDS) constraints as prior information for inversion (Figure 3) [4]. The new evaluation function can be expressed as:

$$E(s; \alpha', \beta', \nu) = \beta' \sum_{k=1}^{R} \left( \frac{1}{2} \sum_{i=1}^{N} |s_{i} - \frac{1}{2} \sum_{k=1}^{R} c_{k} s_{i} s_{k} |^{2} + \frac{1}{2} \sum_{i,j \in \mathcal{N}_{k,j}} |s_{i} - s_{j}|^{2} + \nu \sum_{i=1}^{N} |s_{i}| \right)$$

where $\beta'$ is a precision hyperparameter, $\alpha'$ is a smoothness hyperparameter for the non-zero-slip area, $\nu$ is a sparsity hyperparameter. $\sum_{i,j \in \mathcal{N}_{k,j}}$ is the summation of all pairs of neighboring cells. $\delta'(s_{i}, s_{j})$ behaves like a delta function, and is defined as follows:

$$\delta'(s_{i}, s_{j}) = \begin{cases} 1, & \text{if } s_{i} = 0 \text{ or } s_{j} = 0 \\ 0, & \text{if } s_{i} \neq 0 \text{ and } s_{j} \neq 0 \end{cases}$$
The proposed evaluation function in equation (2) 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 ($g$). 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. The third term on the right-hand side of equation (2) represents the sparsity constraint.

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.

![Figure 3. Schematic diagram for constraints of slip distribution [4].](image)

### 3.2. Results and Discussions

Using the smoothness regularization, the estimated spatial distribution showed a wider slip area with a smoother boundary compared to the true slip distribution. By using the smoothness constraints, it was not possible to resolve the two peaks seen in the depth profile of the true slip distribution (Figure 4a). The estimated maximum slip was approximately 45% 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 only, the spatial distribution estimated by the sparsity regularization showed a jagged slip area compared to the true image (Figure 4b). 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, 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. 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 (Figure 4c). 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.

We also conducted the inversion test with a smoothly varying circular slip distribution. The SDS constraints accurately reproduced the slip values and the smooth distribution of the slip area.
4. Slow slip events using the Fused regularization

We estimated spatial distribution of the L-SSEs beneath the Bungo channel in southwest Japan. L-SSEs with durations of 0.5–5 years and recurrence interval of >5 years occur within this gap of the Nankai subduction zone. L-SSEs in this region occurred repeatedly around 1997, 2003, and 2010. The total slip distributions of these three L-SSEs estimated by geodetic data have been smoothed to some extent due to prior constraints on inversion analyses [9]. The slip areas of these three events are estimated to nearly coincide, filling a spatial gap between deep ETS and seismogenic zones.

The upper (up-dip) boundary of the L-SSEs area was located close to the sources of two M > 7 interplate earthquakes (the 1946 M 8.0 Nankai earthquake and the 1968 M 7.5 Hyuga-nada earthquake). This implies partitioning between seismogenic and aseismic slip areas as highlighted from slow slip events [10]. The lower (down-dip) boundary of the L-SSEs area partly overlaps with the source areas of deep tremors. Temporal correspondence has been reported between the occurrence of L-SSEs and the acceleration of deep tremors beneath the Bungo Channel [11].

4.1. Method

Using generalised fused regularization [12], which is a kind of sparse modelling suitable for detecting discontinuous changes in the model parameters, the evaluation function can be expressed as:

$$
\hat{s}^* = \arg \min \left\{ \|d - Gs\|_2^2 + \lambda \sum_{i \neq j} |s_i - s_j| + \gamma \cdot \lambda \|s\|_1 \right\}
$$

where $\lambda \geq 0$ is a regularisation parameter for fused smoothness and $\gamma \geq 0$ is a hyperparameter to control the ratio between the fusion and the sparsity penalty terms. A large value of $\gamma \cdot \lambda$ decreases the non-zero components, while $\gamma \cdot \lambda$=0 introduces no sparsity. The second term on the right-hand side of equation (4) represents the fused regularisation term, which indicates the smoothness and flatness of model parameters ($s$). The third term represents the sparsity constraint. Optimal solutions were obtained using a set of hyperparameters with minimum residuals via leave-one-out (LOO) cross validation [13].

We used digital data for the observed vertical and horizontal displacements of the 1997, 2003, and 2010 L-SSEs, which were in the Supporting Information of Yoshioka et al. [9]. Crustal displacements were observed at the positions of GNSS stations. In all, we used 33, 65, and 106 continuous GNSS
stations as the number of stations increased over time. We used only the total displacement of each component during periods from 1996.7 to 1998.5, 2001.9 to 2004.5, and 2009.5 to 2011.2. In this study, we focused only on the spatial distribution to estimate the detailed distribution of the total slips during each L-SSE event. $N$ is equal to 781 and the dip-slip only occurs for 781 subfaults.

4.2. Result and Discussions

The slip distribution on the plate interface was estimated using fused regularisation. For the 2010 L-SSEs, slips $> 0.2$ m were concentrated to a relatively narrow area, and slips $< 0.2$ m were mainly distributed on the deeper side (Figure 5a). There were no slips toward the shallower area (Figure 5a). The cross section of the estimated spatial distribution of the slips along line AB showed a clear step at the down-dip side and a steep boundary at the up-dip side (Figure 5b, red line). The 1997 and 2003 L-SSEs showed similar distributions to that of the 2010 L-SSE. However, due to fewer observation stations, the spatial resolutions were lower than that for 2010. Large ($> 0.2$ m) slips were estimated to have occurred at almost the same locations identified by Yoshioka et al. [9]. However, the slip distributions at the up-dip side significantly differed from the previous study [9].

The unique elements of our results are the up-dip boundary of the L-SSE area and the step-like distribution of the down-dip side. The L-SSE areas are clearly separated into two parts [the major (minor) slip area at the up-dip (down-dip) side] with two sharp boundaries (Figures 6 and 7). (i) The shallow boundary between the seismogenic zone and the major L-SSE area corresponds to the isotherm at 350 °C [14]. (ii) The deep boundary between the major L-SSE area and the minor L-SSE area corresponds to the up-dip limit of the tremor area. The sharp variation in the slip amount of L-SSEs and the correspondence of the boundaries with up-dip limit of the tremor area suggest that frictional properties, which control slip behaviour, may change sharply.

Figure 5. (a) Estimated slip distribution in the dip direction for the 2010 event. Black dots indicate GEONET stations. (b) Slip distribution along line AB. Red represent the 2010 event. Gray line represents the 2010 event estimated by smoothed inversion [9].
Figure 6. Comparison of L-SSE slip distribution in 2010 (red polygon with the solid and dashed lines) estimated in this study and spatial distribution based on findings from other seismological studies. Black dashed contours show the temperature (250–550 °C) on the Philippine Sea plate [14].

Figure 7. Blue, green, and red lines represent slip distributions of the 1997, 2003, and 2010 L-SSEs along line AB (as in Figure 5b). Yellow line shows the total tremor numbers in 10-km increments along line AB. Gray line shows the total number of aftershocks for the 1968 Hyuga-nada earthquake in 10-km increments along line AB.

5. 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 one evaluation function as shown in (2). That function can constrain simultaneously and properly smoothness, discontinuity, and sparsity depending on given data. The slip distribution that minimized the evaluation function better reproduced images in terms of
both area and amplitude. 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.

On the other hand, using the fused regularization, we found that the largest slip abruptly becomes zero at the down-dip limit of the seismogenic zone, and is immediately reduced to half at the up-dip limit of the deep low-frequency tremors, and becomes zero near its down-dip limit. Such correspondences imply that some thresholds exist in the generation processes for both tremors and SSEs. It suggests that geodetic data inversion with sparse modelling can detect such abrupt changes in the slip distribution. We can use this method to monitor the position of the abrupt changes, which represents the stress concentration level at the seismogenic zone.

Furthermore, it should be noted that sparse modelling can examine the extent that sharp edges, steps and small slip zones can fit the geodetic data, which only showed a smooth distribution in previous studies. Our results indicate that the resolution of the position of sharp edges estimated from geodetic data analysed with sparse modelling is comparable to the spatial resolution of the hypocentre distribution determined from seismic data. Such independent datasets of the spatial distribution are useful to hypothesise about the physical processes of both megathrust and slow earthquakes. Furthermore, a method based on the generalised fused regularisation should be available for any slip and slip-deficit distribution analyses on various types of faults, promoting the consideration of physical processes.

Both proposed inversion methods are 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 methods, it is possible to obtain images that are clearer, sharper, and show appropriate smoothness. We will develop more efficient and high-speed methodology in order to expand the applicability of the proposed methods.

Acknowledgments

This work was partly supported by the “Initiative for High-Dimensional Data-Driven Science through Deepening of Sparse Modeling” by JSPS KAKENHI (Grant Numbers JP16H01564, 25120005, 25120009, and 25120011) and JSPS KAKENHI (Grant Numbers JP16K17798, 16K16108, 25280090, and 15K20864) and JST PRESTO (Grant Number JPMJPR1676). The Figures and graphs are created by Generic Mapping Tools software (GMT v4.5.12; http://gmt.soest.hawaii.edu/) [15].

References

[1] Yabuki, T. and M. Matsu’ura (1992), Geodetic data inversion using a Bayesian information criterion for spatial distribution of fault slip. Geophys J. Int. 109, 363-375.
[2] Segall, P. and M. Matthews (1997), Time dependent inversion of geodetic data. J. Geophys. Res. 102(B10), 22391-22409.
[3] Igarashi, Y. et al. (2016), Three levels of data-driven science. Journal of Physics Conference Series 699, 12001-1-012001-13.
[4] Nakata R., T. Kuwatani, M. Okada, and T. Hori (2016), Geodetic inversion for spatial
distribution of slip under smoothness, discontinuity, and sparsity constraints, Earth Planets Space, 68:20, doi: 10.1186/s40623-016-0396-2.

[5] Nakata, R., H. Hino, T. Kuwatani, S. Yoshioka, M. Okada, and T. Hori (2017), Discontinuous boundaries of slow slip events beneath the Bungo Channel, southwest Japan, Sci. Rep. 7, doi:10.1038/s41598-017-06185-0.

[6] Baba, T, Y. Tanioka, P. R. Cummins, and K. Uhira (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.

[7] Nakata R., M. Hyodo, and T. Hori (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.

[8] Nakata R., S. Miyazaki, M. Hyodo, and T. Hori (2014), Reproducibility of spatial and temporal distribution of aseismic slips in Hyuga-nada of southwest Japan, Mar. Geophys. Res., 35, 311-317.

[9] Yoshioka, S., Y. Matsuoka, and S. Ide (2015), Spatiotemporal slip distributions of three long-term slow slip events beneath the Bungo Channel, southwest Japan, inferred from inversion analyses of GPS data, Geophys. J. Int., 201, 1437-1455.

[10] Yagi, Y. and M. Kikuchi (2003), Partitioning between seismogenic and aseismic slip as highlighted from slow slip events in Hyuga-nada, Japan. Geophys. Res. Lett. 30(2), 1087, doi:10.1029/2002GL015664.

[11] Hirose, H. et al. (2010), Slow Earthquakes Linked Along Dip in the Nankai Subduction Zone. Science 330, doi:10.1126/science.1197102.

[12] Tibshirani, R., M. Saunders, S. Rosset, J. Zhu and K. Knight (2005), Sparsity and smoothness via the fused lasso, J. Royal. Statist. Soc. B., 67, 91-108.

[13] Bishop, C. M. (2006), Pattern recognition and machine learning. Springer, New York.

[14] Ji, Y., S. Yoshioka, and T. Matsumoto (2016), Three-dimensional numerical modeling of temperature and mantle flow fields associated with subduction of the Philippine Sea plate, southwest Japan. J. Geophys. Res. Solid Earth 121, doi:10.1002/2016JB012912.

[15] Wessel, P. and W. H. F. Smith (1998), New, improved version of Generic Mapping Tools released. Eos, Trans Am Geophys Union 79, 579-579, doi:10.1029/98EO00426.