A precise hypocenter determination method using network correlation coefficients and its application to deep low-frequency earthquakes

Kazuaki Ohta and Satoshi Ide

Department of Earth and Planetary Science, University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-0033, Japan

(Received March 13, 2008; Revised April 24, 2008; Accepted April 30, 2008; Online published September 8, 2008)

A knowledge of the precise locations of deep low-frequency earthquakes (LFEs) along subduction zones is essential to be able to constrain the spatial extent of various slow earthquakes and the underlying physical processes. We have developed a hypocenter determination method that utilizes the summed cross-correlation coefficient over many stations, denoted a network correlation coefficient (*NCC*). The method consists of two parts: (1) an estimation of relative hypocenter locations for every pair of events by a grid search, and (2) a linear least squares inversion for self-consistent relative hypocenter locations for the initial centroid. We have applied this method to ten LFEs in the Tokai region, Japan. Statistically significant values of *NCC* indicate the relative locations for many pairs, which in turn determine the self-consistent locations. While the catalog depths are widely distributed, the relocated hypocenters fall within a 2-km depth range, which implies that LFEs in the Tokai region occur on the plate interface, similar to LFEs in western Shikoku.

Key words: Low-frequency earthquakes, precise hypocenter determination, cross-correlation coefficient, Nankai subduction zone, Tokai region.

1. Introduction

Since the discovery of low-frequency tremors along the Nankai subduction zone in western Japan by Obara (2002), many studies focused on studying various unusual earthquakes in this area, such as low-frequency earthquakes (LFE) (Katsumata and Kamaya, 2003; Shelly *et al*., 2006), very low-frequency earthquakes (Ito *et al*., 2007), and slow slip events (Hirose and Obara, 2005). These unusual events may provide a clue to understanding subduction in general because similar phenomena are widely observed (Schwartz and Rokosky, 2007; Ide *et al*., 2007a).

Of the group classified as slow earthquakes, LFEs have been studied with relatively precise event locations due to their frequent occurrence and isolated signals. Shelly *et al*. (2006) determined the locations of LFE in western Shikoku using double-difference tomography and relative hypocenter location (Zhang and Thurber, 2003) and found that the hypocenters are distributed at a depth of 30–35 km, parallel to intraslab earthquakes, suggesting that these events occur on the plate boundary. Shelly *et al*. (2007a, b) also demonstrated that deep low-frequency tremors can be represented as a swarm of LFEs. Consequently, a knowledge of the properties of LFEs would lead to a better understanding of the whole tremor sequence.

Unfortunately, precise locations are not available in other regions of the Nankai trough, such as the eastern Shikoku, Kii peninsula, and Tokai regions, where LFE hypocenters determined by Japan Meteorological Agency (JMA) have a wide depth distribution. In the Tokai region, the depth distribution ranges from 20 to 50 km (Fig. 1) and resembles the tremor source distribution determined by Kao *et al*. (2005) in the Cascadia subduction zone. Since the wide depth distribution suggests a different physical interpretation, such as fluid movement in the overriding plate, the important question is whether or not the apparent wide depth distribution of LFEs along the Nankai subduction zone is real.

For signals in noisy records, the summation of cross-correlation coefficients for network stations, denoted the network correlation coefficient (*NCC*), is a useful tool, as demonstrated by Shelly *et al*. (2007a, b). These researchers detected many small LFEs in a continuous tremor sequence using *NCC* and known LFEs as template events. Their method, called matched filter analysis (Gibbons and Ringdal, 2006), can be modified to determine the relative location between event pairs. In this paper, we report our development of a new hypocenter determination method based on this concept. This new method is applied to a small set of real data, LFEs in the Tokai region (Fig. 1), to verify its effectiveness.

2. Method

2.1 Overview

The new method consists of two main steps, as illustrated in Fig. 2. First, we estimate the relative hypocenter location for each pair of events by maximizing the summation of waveform cross-correlation, or *NCC*, for all stations. We then determine a set of hypocenter locations that are consistent with the relative locations by solving a least square problem.

2.2 Determination of relative hypocenter location between a pair of events

We first determine the relative hypocenter location between a pair of earthquakes, events i and j. The event i

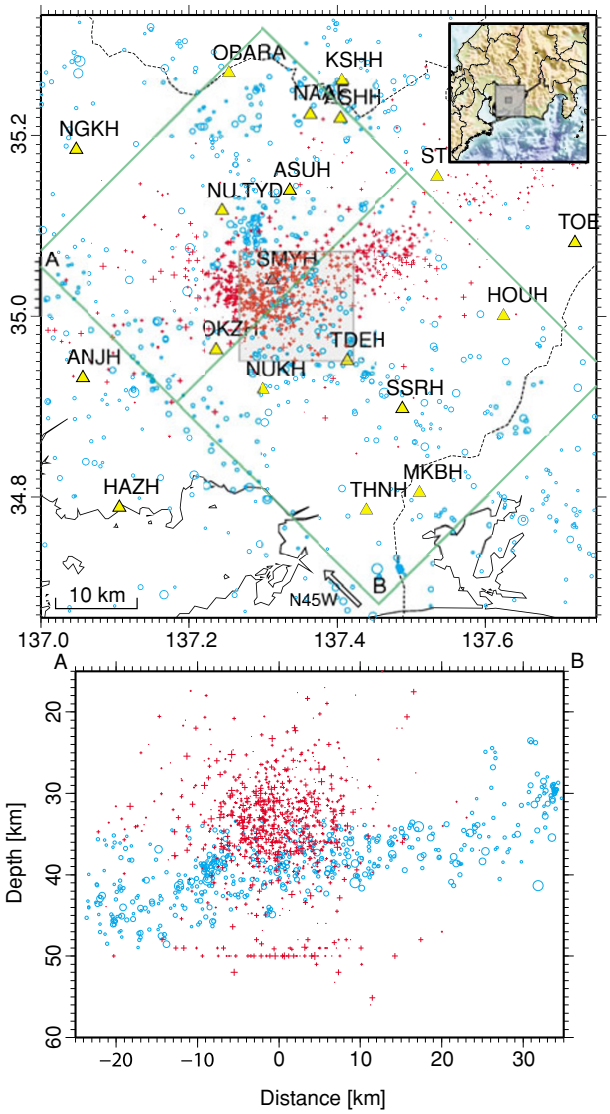

Fig. 1. (Top) Map showing the study area and the locations of LFEs (red crosses) and regular intraplate earthquakes (blue circles), determined by JMA. LFEs used for example are shown within the small gray square. The location of study area in Japan is shown in the inset. Triangles are stations. (Bottom) Cross-sectional view of the hypocenter distribution within the thick green rectangular oriented in the plate subduction direction.

is used as a reference event, with known location and origin time that are tentatively assumed to be the JMA catalog values. Ground velocity waveform in the l-th direction from the event i recorded at a seismic station n, $u^{iln}(t)$, is digitized as

$$u_m^{iln} = u^{iln}(\tau^i + t(\mathbf{x}_i^0, \mathbf{x}_n) - \Delta T_{\rm pre} + (m-1)\delta t), \quad m = 1, \ldots, M, \tag{1}$$

where δt is the sampling interval, and M is the sample number. The start time is calculated from the origin time τ^i, the theoretical travel time of a body wave from the source location $\mathbf{x}_i^0$ to the station $\mathbf{x}_n$, $t(\mathbf{x}_i^0, \mathbf{x}_n)$, and a presignal time $\Delta T_{\rm pre}$. The other event, j, is a target event. Its seismogram is similarly prepared with a time shift calculated from the relative hypocenter location between two events, $\Delta\mathbf{x}_{ij}$, and the origin time shifting, Δt_{ij},

$$u_m^{jln} = u^{jln}(\tau^j + t(\mathbf{x}_i^0 + \Delta\mathbf{x}_{ij}, \mathbf{x}_n) + \Delta t_{ij} - \Delta T_{\rm pre} + (m-1)\delta t), \quad m = 1, \ldots, M. \tag{2}$$

The cross-correlation coefficients between two waveforms are calculated for each component at each station. The total summation of these correlation coefficients is a function of $\Delta\mathbf{x}_{ij}$ and Δt_{ij} and expressed as

$$NCC(\Delta\mathbf{x}_{ij}, \Delta t_{ij}) = \sum_{ln} \frac{\sum_m u_m^{iln} u_m^{jln}}{\sqrt{\sum_m \left(u_m^{iln}\right)^2 \sum_m \left(u_m^{jln}\right)^2}} \tag{3}$$

This is the network correlation coefficient, NCC, defined by Gibbons and Ringdal (2006) and used by Shelly *et al.* (2007a, b) to detect LFEs in a tremor sequence.

The NCC is high only when waveforms from two events correlate for all stations and components. Therefore, the relative location, $\Delta\mathbf{x}_{ij}^{NCC}$, is determined by the maximum of the NCC. Since the NCC has many local maxima that can be a source of uncertainty, we determine the global maximum using a grid search. Applying this procedure to every pair of N events yields $N \times (N-1)$ relative locations, $\Delta\mathbf{x}_{12}^{NCC}, \ldots, \Delta\mathbf{x}_{ij}^{NCC}, \ldots, \Delta\mathbf{x}_{NN-1}^{NCC}$.

2.3 Least-squares inversion for self-consistent hypocenter locations

Once relative locations between events are measured precisely and the location of one event is given, the absolute locations of all events are in principle automatically determined. However, the uncertainty of the relative hypocenter locations is neither negligible nor homogeneous. We assume that the relative locations for all combinations form a data vector with Gaussian errors, each of which has a variance depending on the value of the NCC. Unknown parameters are hypocenter locations $\mathbf{x}_m (m = 1, \ldots, N)$, determined by minimizing,

$$E = \sum_{i,j,i\neq j} w_{ij} \left(\mathbf{x}_j - \mathbf{x}_i - \Delta\mathbf{x}_{ij}^{NCC}\right)^2, \tag{4}$$

where w_{ij} is a weighting factor. Since the absolute location is not constrained by the above equation, we also assume that the centroid of the hypocenters is unchanged from the catalog value so that

$$\sum_i \mathbf{x}_i = \sum_i \mathbf{x}_i^0. \tag{5}$$

The weighting factor is calculated based on the probability that a relative location is correct. Although a large maximum of NCC generally means high reliability, even waveforms of Gaussian noise may occasionally show a very large value in numerous iterations. If the NCC is measured from discrete time series of M samples of Gaussian white noise with a unit standard deviation at N stations, the NCC also has a Gaussian distribution with a standard deviation σ_{NCC} of $\sqrt{N/M}$, which is an expectation of the square of Eq. (3). Hence, the probability that the maximum NCC exceeds $r\sigma_{NCC}$ in a grid search for $N_{\rm g}$ points is given as

$$P^{N_{\rm g}}(r) = 1 - \left[\int_{-\infty}^{r} \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} ds\right]^{N_{\rm g}}, \tag{6}$$

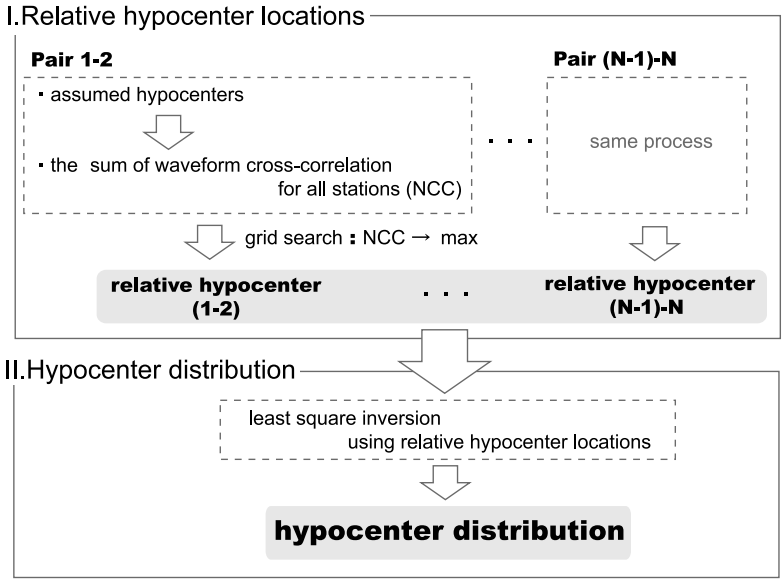

Fig. 2. A schematic diagram showing the main two steps of the new method in the case of determining hypocenter distribution of N events.

if calculations at different grid points are independent. For example, in a grid search for 100^4 points, 100 grid points in three-dimensional (3D) space and time, the maximum exceeds $5\sigma_{NCC}$ and $6\sigma_{NCC}$, with probabilities of about 100% and 10%, respectively. The standard deviation of NCC for real data is measured from NCC values at all grid points.

When the ratio between the maximum and the standard deviation of the NCC is r_{ij} for the event pair i and j, we assume the variance of the estimation error to be the sum of variances of noise and signals,

$$\sigma_{\rm d}^2(r_{ij}) = P^{N_{\rm g}}(r_{ij})\sigma_{\rm N}^2 + (1 - P^{N_{\rm g}}(r_{ij}))\sigma_{\rm S}^2, \qquad (7)$$

where $\sigma_{\rm S}^2$ and $\sigma_{\rm N}^2$ are the variances of signals and random noises, respectively. It should be noted that even in the case of identical signals, because we adopted a grid search scheme, there is a quantization error that is dependent on the grid point interval Δl. On the other hand, the variance of noise depends on the size of search space, L, in which any points are equally selected as the target location. Thus, $\sigma_{\rm N}^2$ and $\sigma_{\rm S}^2$ are respectively written as

$$\sigma_{\rm N}^2 = \frac{1}{L}\int_{-L/2}^{L/2} l^2 dl = \frac{L^2}{12},\ \sigma_{\rm S}^2 = \frac{1}{\Delta l}\int_{-\Delta l/2}^{\Delta l/2} l^2 dl = \frac{\Delta l^2}{12}. \qquad (8)$$

The reciprocal of this variance, $\sigma_{\rm d}^2(r_{ij})$, is the weighting factor w_{ij} in Eq. (4).

3. Application to LFEs in the Tokai Region

3.1 Study area and waveform data

We apply this method to LFEs in the Tokai region (Fig. 1). More than 1000 LFEs have been detected and located by JMA between 2002 and 2007, and the locations listed in the catalog are widely distributed. Before applying the new method, we attempted to determine hypocenter locations using the double-difference method (Waldhauser and Ellsworth, 2000) with cross correlation, similar to the analysis of LFEs carried out in western Shikoku by Shelly *et al.* (2006). However, few cross-correlation coefficients exceed 0.7, a threshold value used in Shelly *et al.* (2006), and we were unable to obtain reliable results.

As a first small data set, we selected ten events from the original data set of 1000 LFEs. These events have relatively large amplitude and impulsive waveforms and occurred within a small area (Fig. 1). The data are three-component velocity seismograms observed at the Hi-net stations. We then calculated the cross correlation using the vertical component of the P-wave and two horizontal components of the S-wave. Each seismogram is bandpass filtered between 2 and 8 Hz. The number of stations available for each event pair is different and averages about ten. The time window used for calculating a cross correlation is from 1.5 s before to 2.5 s after the theoretical P- or S-wave arrival times, which are calculated assuming a horizontally layered structure based on local seismic reflection and refraction surveys (Iidaka *et al.*, 2003).

3.2 Relative hypocenter location determined by *NCC*

We first determined the relative hypocenter locations for all combinations of these ten events. We assume that the differences of latitude, longitude, and depth between two events are within 0.1°, 0.1°, and 10 km, respectively, and search the location $\Delta\mathbf{x}_{ij}^{NCC}$, which gives the maximum NCC in this range. The range of the origin time shift Δt_{ij} is between -2 and $+2$ s. A grid search for $\Delta\mathbf{x}_{ij}$ and Δt_{ij} revealed the maximum, with grid point intervals of 0.001°, 0.001°, 0.1 km, and 0.04 s, in latitude, longitude, depth, and time directions, respectively.

Figure 3 shows the spatial distributions and histograms of NCC for two pairs of events. In the first case (Fig. 3(a)), the maximum is fivefold the standard deviation, σ_{NCC}, and the local maxima of similar values are widely distributed in the search space. As already explained, the maximum of such low values is well explained by Gaussian noise and many iterative calculations. Therefore, this estimation is almost meaningless. On the other hand, in the second example (Fig. 3(b)), the localized maximum is $9.3\sigma_{NCC}$. The cross-correlation coefficient for each trace is not that high

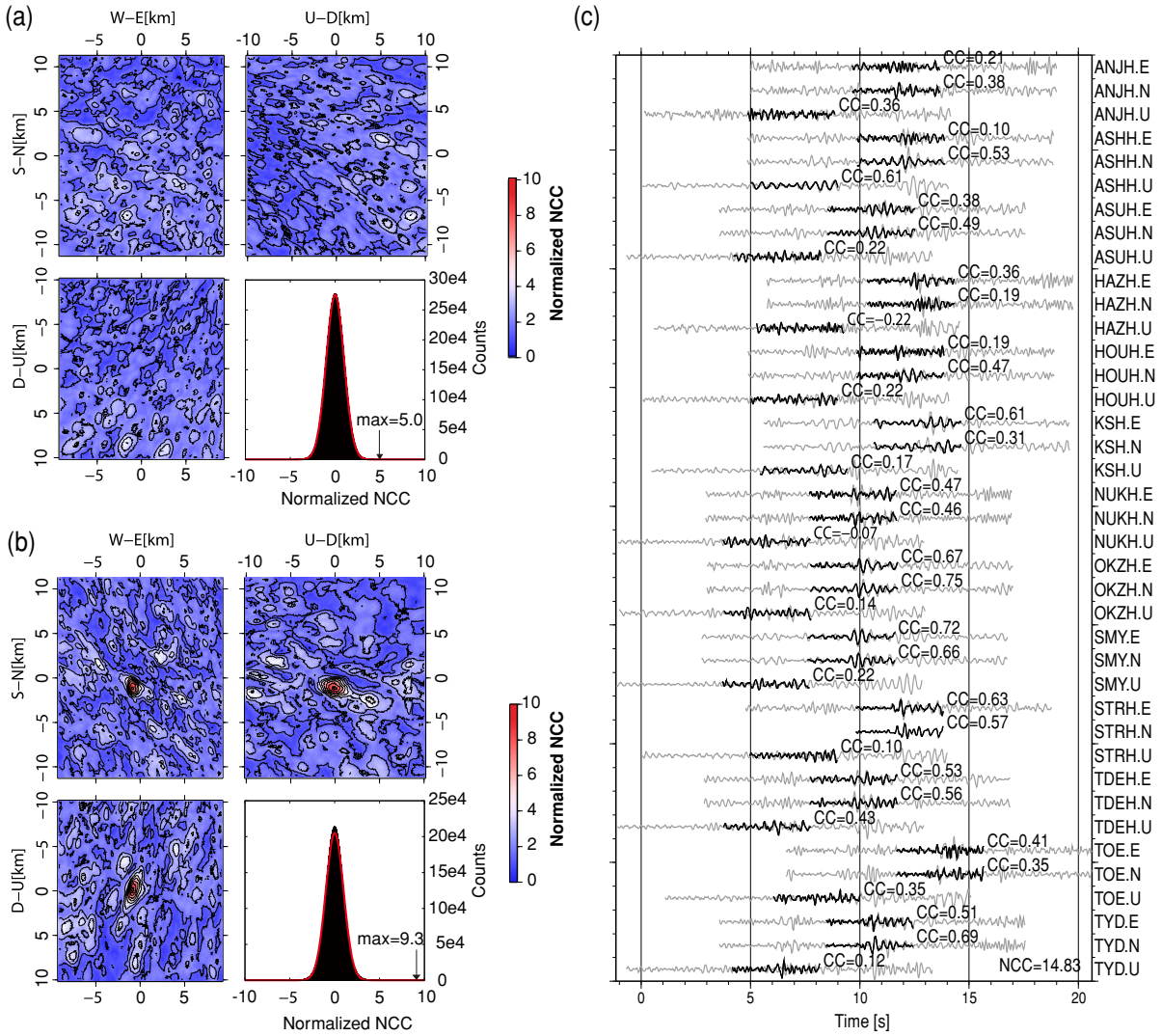

Fig. 3. The contour map and histogram of *NCC* for the event pair which have (a) a statistically insignificant maximum value of *NCC* (fivefold the standard deviation) and (b) a statistically high maximum (9.3-fold). In the contour maps, the maximum *NCC* for temporal grids at each spatial grid point is shown, normalized by the standard deviation. The red line of each histogram panel shows a Gaussian distribution. (c) For the high *NCC* maximum pair shown in (b), waveforms of the reference event (black) and the target event (gray) are compared for each component of 12 Hi-net stations. The correlation coefficient (*CC*) is shown for each trace, and the total sum, *NCC*, is shown in the lower right of the panel. Station names and components are written to the right of each trace.

(Fig. 3(c)), which even in the best case prevents the application of standard methods, such as the double-difference technique. Nevertheless, the maximum *NCC* is statistically significant; the probability that Gaussian noise leads to this value, $\sim 5.8 \times 10^{-10}$, is obtained using Eq. (6). Therefore, this location of the maximum *NCC* [(latitude, longitude, depth) = (-0.011°, -0.009°, -0.1 km)] certainly indicates the relative location between two events. These values are slightly different from the relative location in the catalog, (-0.0030°, 0.0077°, -0.67 km). When we switch the reference and the target, the maximum *NCC* of $8.9\sigma_{NCC}$ is obtained at (0.011°, 0.009°, 0 km), which is almost the opposite.

Among 90 pairs of events, 21 and 35 combinations have *NCC* larger than $7\sigma_{NCC}$ and $6.5\sigma_{NCC}$, with probabilities given by Eq. (6) of about 0.1% and 3%, respectively.

3.3 Relocation

A linear inversion of the relative locations with variances given by Eq. (7) determines the self-consistent hypocenter locations of the ten LFEs, as shown in Fig. 4. The hypocenter distribution is more concentrated than the catalog locations of JMA, particularly in terms of depth, which spans a 2-km range, suggesting that the depth variation in the catalog is due to errors in the picks and velocity model.

Figure 4 also shows that the LFEs are separated into two groups: a main group of seven events and a two-event group to the west of the main group. Within each group, the LFEs are strongly connected to each other, as shown by "bonds"—i.e., high *NCC* values larger than $7\sigma_{NCC}$, in the lower-right panel of Fig. 4. In the main group, several bonds of more than a 1-km distance separate events into two subgroups, with an east-west distribution. The location of one event that has no high *NCC* bond with other events is not well determined. The shallow depth of this event, 27.4 km, in the JMA catalog suggests that the true location may be far from those of the other events. Another possibility is that the mechanism of this event is different. We assume that the mechanisms of LFEs in this area are similar, as suggested

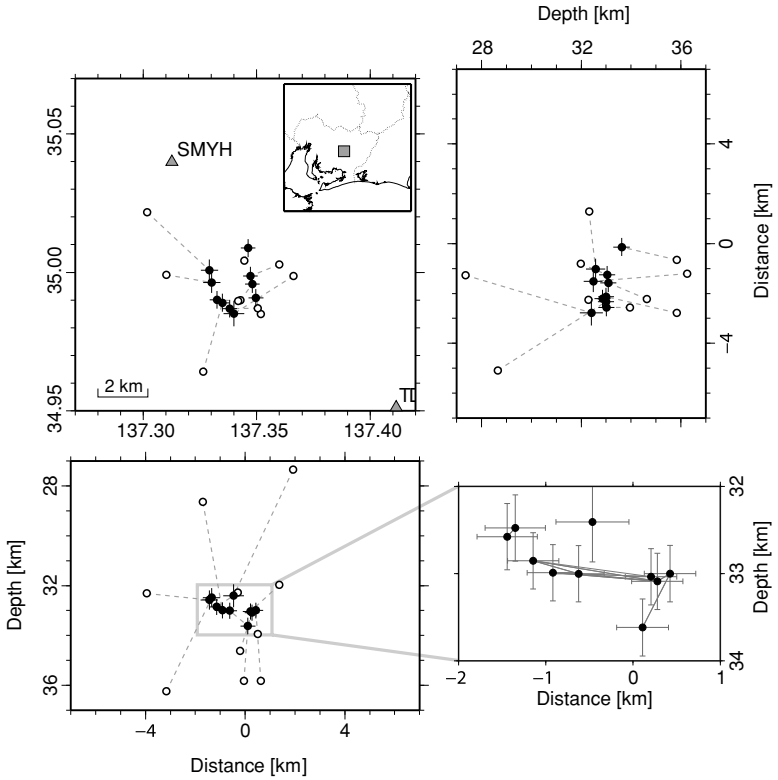

Fig. 4. The hypocenter locations of ten LFEs used for the analysis. A black and white circle connected by a dashed line are the relocated and the catalog locations for one event. Crosses show the standard deviation of the model parameters. The close-up panel shows the connection between pairs that have highly correlated waveforms. A "bond" connects two events that have *NCC* larger than $7\sigma_{NCC}$.

in Shikoku (Ide *et al*., 2007b), and while the many statistically significant *NCC* values support this assumption, we cannot exclude the possibility of mechanism variation for some events.

4. Discussion and Conclusion

When we calculated the cross correlation of seismic waves between two events simultaneously for many stations, we were able to precisely evaluate the relative location even in the presence of noise. Although the value of the correlation coefficient is small at each station, the sum, the network correlation coefficient (*NCC*), acquires a statistically significant value, which cannot be explained by Gaussian noise. The new method is robust and applicable for the events with a low signal-to-noise ratio if the source mechanisms are similar.

In this paper, we apply this method to ten LFEs in the Tokai region. The relocated hypocenters are more concentrated than those in the JMA catalog, suggesting that the apparent wide distribution in the catalog is an artifact. Since the subducting Philippine Sea plate changes strike beneath the study area, suggesting some complexity, and the assumed structure is just 1D, it is difficult to be sure precisely where these events occur; however, the localization in depth suggests that these LFEs occur along some irregularity; the plate interface is an obvious candidate. Despite apparent differences in the original catalog and the existence of short- and very long-term slow slip events (Hirose and Obara, 2006; Miyazaki *et al*., 2006), LFEs in the Tokai region may have the same characteristics as those in western Shikoku, where LFEs are considered to be shear slip on the plate interface (Shelly *et al*., 2006; Ide *et al*., 2007b). Application to large event sets will increase the number of bonds and may connect the separate groups in this study. Since the method is applicable to all events with a low signal-to-noise ratio, we expect to be able to apply this method not only to LFEs but also to low-frequency tremors.

Acknowledgments. We are grateful to Greg Beroza for his enlightening comments. We thank David Schaff and Yoshihiro Ito for their careful and constructive reviews. This work is supported by Grant-in-Aid for Scientific Research, Ministry of Education, Sports, Science and Technology, Japan and JSPS Bilateral Joint Project.

References

Gibbons, S. J. and F. Ringdal, The detection of low magnitude seismic events using array-based waveform correlation, *Geophys. J. Int*., **165**, 149–166, 2006.

Hirose, H. and K. Obara, Repeating short- and long-term slow slip events with deep tremor activity around the Bungo channel region, southwest Japan, *Earth Planets Space*, **57**, 961–972, 2005.

Hirose, H. and K. Obara, Short-term slow slip and correlated tremor episodes in the Tokai region, central Japan, *Geophys. Res. Lett*., **33**, L17311, doi:10.1029/2006GL026579, 2006.

Ide, S., G. C. Beroza, D. R. Shelly, and T. Uchide, A scaling law for slow earthquakes, *Nature*, **447**, 76–79, 2007a.

Ide, S., D. R. Shelly, and G. C. Beroza, The mechanism of deep low frequency earthquakes: Further evidence that deep non-volcanic tremor is generated by shear slip on the plate interface, *Geophys. Res. Lett*., **34**, L03308, doi:10.1029/2006GL028890, 2007b.

Iidaka, T., T. Iwasaki, T. Takeda, T. Moriya, I. Kumakawa, E. Kurashimo, T. Kawamura, F. Yamazaki, K. Koike, and G. Aoki, Configuration of subducting Philippine Sea plate and crustal structure in the central Japan region, *Geophys. Res. Lett*., **30**, doi:10.1029/2002GL016517, 2003.

Ito, Y., K. Obara, K. Shiomi, S. Sekine, and H. Hirose, Slow Earthquakes

Coincident with Episodic Tremors and Slow Slip Events, *Science*, **315**, 503–506, 2007.

Kao, H., S.-J. Shan, H. Dragert, G. Rogers, J. F. Cassidy, and K. Ramachandran, A wide depth distribution of seismic tremors along the northern Cascadia margin, *Nature*, **436**, 841–844, 2005.

Katsumata, A. and N. Kamaya, Low-frequency continuous tremor around the Moho discontinuity away from volcanoes in the southwest Japan, *Geophys. Res. Lett.*, **30**, doi:10.1029/2002GL015981, 2003.

Miyazaki, S., P. Segall, J. J. McGuire, T. Kato, and Y. Hatanaka, Spatial and temporal evolution of stress and slip rate during the 2000 Tokai slow earthquake, *J. Geophys. Res.*, **111**, B03409, doi:10.1029/2004JB003426, 2006.

Obara, K., Nonvolcanic deep tremor associated with subduction in southwest Japan, *Science*, **296**, 1679–1681, 2002.

Schwartz, S. Y. and J. M. Rokosky, Slow slip events and seismic tremor at circum-Pacific subduction zones, *Rev. Geophys.*, **45**, RG3004, doi:10.1029/2006RG000208, 2007.

Shelly, D. R., G. C. Beroza, S. Ide, and S. Nakamula, Low-frequency earthquakes in Shikoku, Japan and their relationship to episodic tremor and slip, *Nature*, **442**, 188–191, 2006.

Shelly, D. R., G. C. Beroza, and S. Ide, Non-volcanic tremor and low-frequency earthquake swarms, *Nature*, **446**, 305–307, 2007a.

Shelly, D. R., G. C. Beroza, and S. Ide, Complex evolution of transient slip derived from precise tremor locations in western Shikoku, Japan, *Geochem. Geophys. Geosyst.*, **8**, Q10014, doi:10.1029/2007GC001640, 2007b.

Waldhauser, F. and W. L. Ellsworth, A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, California, *Bull. Seismol. Soc. Am.*, **90**, 1353–1368, 2000.

Zhang, H. and C. H. Thurber, Double-difference tomography: The method and its application to the Hayward fault, California, *Bull. Seismol. Soc. Am.*, **93**, 1875–1889, 2003.

K. Ohta (e-mail: ohta@eps.s.u-tokyo.ac.jp) and S. Ide