Geophysical Journal International

Geophys. J. Int. (2019) **217,** 1123–1140
Advance Access publication 2019 February 05
GJI Seismology

doi: 10.1093/gji/ggz073

Bayesian approach for detecting dynamically triggered very low-frequency earthquakes in the Nankai subduction zone and application to the 2016 M_w 5.9 off-Kii Peninsula earthquake, Japan

Masatoshi Miyazawa

Disaster Prevention Research Institute, Kyoto University, Kyoto 611-0011, *Japan. E-mail: miyazawa@rcep.dpri.kyoto-u.ac.jp*

Accepted 2019 February 2. Received 2019 January 31; in original form 2018 October 22

SUMMARY

Remote triggering of very low-frequency (VLF) earthquakes in the Nankai subduction zone by surface waves from a moderate to large, distant earthquake was examined using a Bayesian approach. The triggering of another type of tectonic slow earthquake/tremor and slow slip was previously discovered using simple signal processing techniques that exploit the different characteristic periods of the triggered and triggering event waveforms. Meanwhile, detecting low-amplitude VLF earthquakes embedded in high-amplitude seismic waves is challenging when they have similar characteristic periods, as established detection algorithms are inapplicable in such cases. Here we use a particle filter/smoother for time-series analysis combined with Markov chain Monte Carlo methods, to detect seismic signals and estimate the maximum likelihood source parameters of VLF events that might have been remotely triggered by surface waves from a moderate to large, distant earthquake, in data where the surface waves and target signals both have predominant periods between 10 and 100 s. This approach was applied to seismograms of the 2016 M_w 5.9 off-Kii Peninsula earthquake (Japan), recorded by the KiK-net borehole array in western Shikoku, Japan, where VLF events were previously reported. The borehole array is used because comparing surface and borehole records can identify vertically incident body waves based on depth-dependent phase differences, which do not appear in records of horizontally propagating surface waves. In particle filtering/smoothing, waveforms of a probable VLF event at each seismic station are predicted by a full-wavefield simulation in a 3-D structure model, and seismic interferometry that estimates traveltimes between surface and borehole sensors. A maximum-likelihood approach is used to estimate source parameters that best explain the surface observations predicted by the borehole observations and probable VLF signals. We found that six VLF events with magnitudes $3.4 \leq M_w \leq 4.3$ likely occurred in the seismogenic regions of ambient low-frequency earthquakes/tremors (LFEs), and were likely triggered by the arrivals of high-amplitude surface waves. The triggering stress changes are estimated from 0.4 to 1.5 kPa, similar to the values for triggered LFEs reported previously. Since there were no triggered LFEs in western Shikoku following the 2016 M_w 5.9 earthquake, these VLF earthquakes are more sensitive to stress changes caused by surface waves from large earthquakes. This is the first reported example of dynamic triggering of VLF earthquakes.

Key words: Computational seismology; Earthquake dynamics; Seismic interferometry; Subduction zone processes; Time-series analysis; Wave propagation.

1 INTRODUCTION

Slow tectonic events, such as slow earthquakes and slow slip events (SSEs), are phenomena that slowly release strain energy following a scaling law and originate in a conditionally stable slip area in a transition zone between the seismic and aseismic regions of subduction zones (e.g. Ide *et al.* 2007; Obara & Kato 2016). Slow earthquakes include low-frequency earthquakes/tremors (LFEs) and deep and shallow very low-frequency (VLF) earthquakes. LFEs and VLF earthquakes having characteristic periods of <1–100 s can be detected by seismographs, while SSEs with characteristic periods of days to months are detected by geodetic observations, such as the Global Navigation Satellite System (GNSS), strain meters and tilt meters.

All slow tectonic events are known to be sensitive to external stress perturbations due to passing seismic waves and earth tides

(e.g. Peng & Gomberg 2010). Observations indicate that high-amplitude surface waves from large earthquakes have transiently triggered LFEs in the Nankai subduction zone of southwest Japan (e.g. Miyazawa & Mori 2005, 2006), Cascadia (e.g. Rubinstein *et al.* 2007), the Hikurangi subduction zone (Fry *et al.* 2011) and Mexican subduction zones (Zigone *et al.* 2012). The amplitudes of triggered LFEs are logarithmically proportional to dynamic stress changes (Miyazawa & Brodsky 2008; Chao & Obara 2016). Slow-slip events are also remotely triggered by passing surface waves in southwest Japan (Itaba & Ando 2011) and Mexico (Zigone *et al.* 2012). Investigation of these phenomena provides us with a better understanding of the mechanisms of LFEs; for example the amplitudes of triggered LFEs are in phase with triggering waves associated with the stress changes that promote slip on the plate interface (e.g. Miyazawa & Mori 2006; Rubinstein *et al.* 2007; Miyazawa & Brodsky 2008).

Despite many observations of dynamic triggering of LFEs and SSEs in subduction zones, dynamic triggering of VLF events during the arrival of seismic waves from a large, distant earthquake has not been observed; a possible exception is the high VLF seismicity rate in southwest Japan within about an hour after the passage of seismic waves from the 2015 M_w 7.9 Ogasawara (Bonin) deep earthquake (National Research Institute for Earth Science and Disaster Prevention 2016). Previous studies have examined pairs of events with different characteristic periods, which are easily separated: LFEs with dominant periods $\leq$1 s triggered during the passage of surface waves with dominant periods of 10–20 s can be detected by high-pass filtering, even if the LFE amplitudes are four orders of magnitude lower than the surface wave amplitudes (e.g. Miyazawa & Brodsky 2008). SSEs can also be detected because an SSE releases the moment at periods of days and can therefore be observed geodetically (e.g. by GNSS), while surface waves are seismic phenomena, even though there is some difficulty in reading an event onset from geodetic data. In the case of a VLF earthquake, since both the VLF earthquake and the surface wave have similar predominant periods between 10 and 100 s, remote triggering is not easily detected by using existing methods.

In this work, a probabilistic framework is developed for the detection of VLF earthquakes triggered by distant surface waves, in which both phenomena have similar characteristic periods. Signal processing approaches such as cross-correlation to extract VLF earthquake signals are inapplicable due to the similar periods, even though background VLF earthquakes have been detected successfully by cross-correlation methods (e.g. Baba *et al.* 2018). Frequency–wavenumber analyses that utilize the different source locations of VLF earthquakes and surface waves cannot generally be used due to the poor coverage of the seismic array for earthquake observation. In this study, a state-space model for time-series analysis is used to extract VLF earthquakes from observed waveforms. Kitagawa *et al.* (2002) applied a similar spatial–temporal model based on prior knowledge of the propagation of the seismic signal, and successfully separated direct waves from small-amplitude reflected and/or refracted waves recorded by ocean-bottom seismometers. A particle filter/smoother (e.g. Kitagawa 1996; Wikle and Berliner 2007) is employed that can estimate non-Gaussian state-space models, as the emergent arrivals of the VLF signals may cause a non-Gaussian probabilistic distribution of amplitudes. In practice, parameter estimation in the state-space model corresponds to the detection of signals. A self-organizing state-space model to obtain parameters was originally proposed by Kitagawa (1998), followed by other derivative methods (e.g. Liu and west 2001; Yano 2008; Carvalho *et al.* 2010). In this study, since we cannot begin with the assumption that a triggered VLF earthquake exists, a Markov chain Monte Carlo method (MCMC) is developed for parameter estimation. A similar method, particle Markov chain Monte Carlo, which is a nonstandard combination of MCMC and sequential Monte Carlo methods, has also been developed to estimate parameters (e.g. Andrieu *et al.* 2010).

Creation of a state-space model requires predicting the waveforms of surface waves and VLF earthquakes that will be observed at stations and treating this as a priori information for waveform separation. The prediction of incoming surface waves therefore plays an essential role in the particle filter/smoother because of their high amplitudes relative to any possible VLF earthquakes. The analysis employs seismograms from KiK-net stations in Japan, each of which has two sensors: one at the surface and another in a vertical borehole. The borehole seismogram is used to predict surface waves at the surface and the surface seismogram is then used to explain the observations recorded on the ground. Waveforms from probable VLF earthquakes are synthetically obtained using an open-source Seismic Wave Propagation Code (OpenSWPC, Maeda *et al.* 2017).

This study proposes a Bayesian approach, which combines a particle filter/smoother and MCMC, to detect low-amplitude VLF earthquakes embedded in the observed seismograms. After exploring the limits of detectability through synthetic tests, the method is applied to detect probable triggers of VLF earthquakes by passing surface waves from the 2016 M_w 5.9 earthquake, off-Kii Peninsula, Japan, on the plate interface, which occur in the seismogenic region of historical and future megathrust earthquakes in the Nankai trough (e.g. Wallace *et al.* 2016; Tsuji *et al.* 2018). This earthquake is already known to have triggered adjacent SSEs and VLF earthquakes at the shallow extent of the plate interface (Araki *et al.* 2017; Kaneko *et al.* 2018, Nakano *et al.* 2018); however, the present study focuses on deep VLF earthquakes in western Shikoku, 200–400 km to the west (Fig. 1). Investigations of how VLF earthquakes are remotely and dynamically triggered can advance our understanding of the source mechanisms of VLF earthquakes, as has been done for LFEs, using methods that are independent from the conventional seismological approach.

2 METHODOLOGY AND ALGORITHM

We use a particle filter/smoother for waveform analysis and MCMC for parameter estimation. To detect a low-amplitude signal from a VLF earthquake characterized by its parameters, which might be embedded in the observed waveform data, the likelihood of an embedded waveform is obtained using a particle filter/smoother. MCMC is used to compute the probabilistic distributions of parameters and determine the parameter set with the maximum likelihood value. In this section, the method and algorithm used in the analysis are introduced.

2.1 Particle filter and smoother

In general, a state-space model can be applied to time-series prediction, interpolation, and parameter estimation. The particle filter/smoother is a sequential Monte Carlo method for time-series analyses, applicable to nonlinear and/or non-Gaussian state-space models. In this study, a particle filter/smoother is used to estimate the parameters in the model of seismic waveforms. A particle filter/smoother is introduced, based on Kitagawa & Sato (2001).

The generalized state-space model is defined by a system of equations:

$$x_t = f(x_{t-1}, \xi_x, v_t) \quad (1)$$

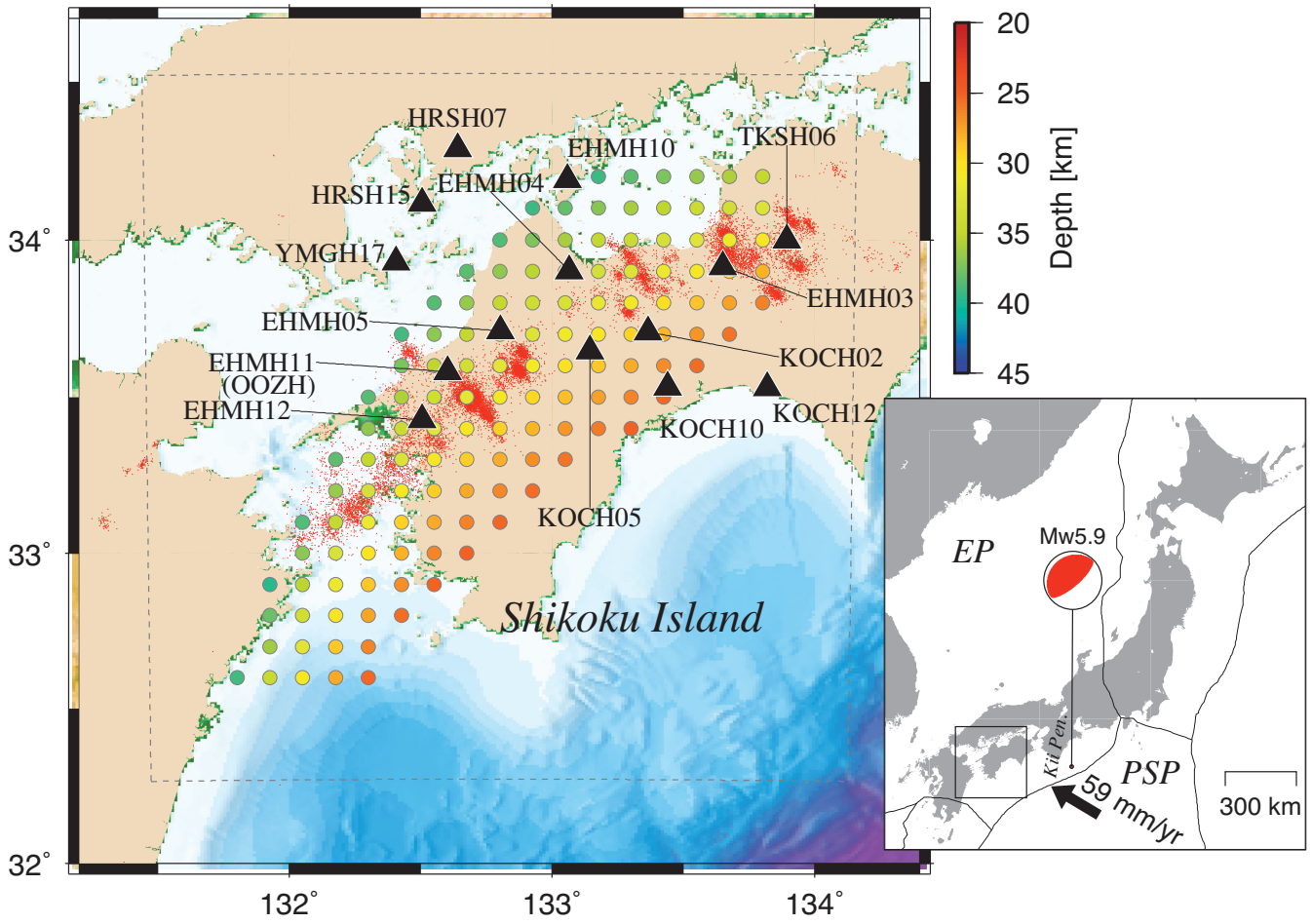

Figure 1. Research area in Western Japan. Fourteen KiK-net stations are indicated by triangles with station names. Coloured circles show the locations of virtual VLF earthquake sources ($n = 132$) on the plate interface. The locations of LFEs from the JMA catalogue (2000–2016) are indicated by red dots. Synthetic wavefields of VLF earthquakes are calculated in the region bounded by the dashed square. The inset map includes the focal mechanism of the 2016 M_w 5.9 off-Kii earthquake from the Global Centroid Moment Tensor (CMT) catalogue. Plate boundaries are from Bird (2003). The arrow in the inset map indicates the motion of the Philippine Sea Plate (PSP) relative to the Amur Plate (Eurasian Plate, EP), based on MORVEL plate motion estimates by DeMets *et al.* (2010).

and an observation equation:

$$y_t = h(x_{t-1}, \xi_y, w_t), \tag{2}$$

where x_t is a state variable at time $t = 1, 2, , T$, y_t is an observation at time t, v_t is system noise, w_t is observation noise, $\theta = (\xi_x, \xi_y)$ is a multidimensional parameter, and f and h can be linear or non-linear functions. eqs (1) and (2) indicate that state x and observation y are represented by the conditional probability distributions

$$x_t \sim p(x_t|x_{t-1}) \tag{3}$$

and

$$y_t \sim p(y_t|x_t), \tag{4}$$

respectively, where '$\sim$' is read as 'is distributed as' and $p(Y|X)$ is a conditional probability distribution for Y given X. The predictive distribution, filtering distribution, and smoothing distribution are represented by

$$p(x_t|y_{1:t-1}) = \int p(x_t|x_{t-1})p(x_{t-1}|y_{1:t-1})dx_{t-1}, \tag{5}$$

$$p(x_t|y_{1:t}) = \frac{p(y_t|x_t)p(x_t|y_{1:t-1})}{\int p(y_t|x_t)p(x_t|y_{1:t-1})dx_t}, \tag{6}$$

and

$$p(x_t|y_{1:T}) = p(x_t|y_{1:t}) \int \frac{p(x_{t+1}|x_t)p(x_{t+1}|y_{1:T})}{p(x_{t+1}|y_{1:t})} dx_{t+1}, \tag{7}$$

respectively.

A particle filter/smoother that proceeds through discrete (i.e. Monte Carlo) approximations is used to determine the exact posterior distributions. This method has two key advantages over a Kalman filter: the observation model (system of equations) can be non-linear, and the noise distribution can also be non-Gaussian. The density distribution is expressed by using N particles as follows:

$$x^i_{t|t-1} \sim p(x_t|y_{t-1}) \tag{8}$$

for $i = 1, 2, , N$. eq. (5) can be approximated by

$$p(x_t|y_{1:t-1}) \simeq \frac{1}{N} \sum_{i=1}^{N} \delta(x_t - x^i_{t|t-1}), \tag{9}$$

where δ is a unit delta function. Given an observation y_t and the ensemble of particles, $\{x^i_{t|t-1}\}^N_{i=1}$, the posterior probability of the particle is obtained from

$$P(x_t = x^i_{t|t-1}|y_{1:t}) = \frac{p(y_t|x^i_{t|t-1})p(x^i_{t|t-1}|y_{1:t-1})}{\sum_{i=1}^N p(y_t|x^i_{t|t-1})p(x^i_{t|t-1}|y_{1:t-1})} = \frac{w^i_t}{\sum_{i=1}^N w^i_t}, \quad (10)$$

where the weight w^i_t is defined by

$$w^i_t = p(y_t|x^i_{t|t-1}). \quad (11)$$

This is equivalent to approximating the true filter distribution as an empirical distribution function:

$$p(x_t|y_{1:t}) \simeq \sum_{i=1}^N \frac{w^i_t}{\sum_{i=1}^N w^i_t}\delta(x_t - x^i_{t|t-1}). \quad (12)$$

The particle smoother can be obtained by slightly modifying the particle filter, and is introduced in Section 2.2.

The conditional probability distribution of y_t given $y_{1:\,t-1}$ is obtained from

$$p(y_t|y_{1:t-1}) = \int p(y_t|x_t)p(x_t|y_{1:t-1})dx_t, \quad (13)$$

which is rewritten by using N particles as

$$p(y_t|y_{1:t-1}) = \frac{1}{N}\sum_{i=1}^N w^i_t. \quad (14)$$

Finally, the log-likelihood (LLH) function for θ is given by

$$l(\theta) = \log L(\theta) = \sum_{t=1}^T \log p(y_t|y_{1:t-1}, \theta) = \sum_{t=1}^T \log\left\{\sum_{i=1}^N w^i_t\right\} - T\log N. \quad (15)$$

The total log-likelihood function for M observations is given by summing the log-likelihood functions $l_m(\theta)$ over the station index m, as follows:

$$l(\theta) = \sum_{m=1}^M l_m(\theta). \quad (16)$$

2.2 Algorithm combining the particle smoother and MCMC

We now summarize the algorithm used to obtain the distribution of parameters using MCMC and particle smoother techniques. MCMC is used to find the parameters that maximize the log-likelihood function for the parameter θ, where the distribution is obtained by the particle smoother method. This study uses the fixed-lag particle smoother that is an extension of the particle filter (Kitagawa 1996; Clapp & Godsill 1999). The smoother is generally more useful for knowledge discovery in data processing and suitable for our research goals. In particular, the particle smoother algorithm only preserves the particles of past states. To avoid degeneracy due to the particle smoother, the Random-Walk Metropolis–Hastings sampling algorithm of Gilks & Berzuini (2001) is employed. We parallelized this algorithm for rapid computation, which is easily done. The log-likelihood functions are obtained by the steps of workflow as follows (Fig. 2).

(1) Generate an initial parameter θ that includes the source parameters of the VLF earthquake. Iterate from (2) to (4) until the distribution of $l(\theta)$ is obtained.
(2) Loop for t to obtain $l(\theta)$ using the particle smoother.

(2.1) Initialize time index $t = 0$ given observation $y_{1:\,T}$, parameter θ, and initial distribution $p(x_0|0)$.
(2.2) Generate state $x^i_0 \sim p(x_0|0)$.
Iterate from (2.3) to (2.6) until $t = T$.
(2.3) Compute particles at time t.

(1) Generate a random number necessary for particles.
(2) Compute distribution $x^i_{t|t-1}$ in eq. (8).
(3) Compute the weight in eq. (11) based on the observation eq. (2).
(4) Generate $S^i = \{\hat{x}^i_{t-L|t-1}, \cdots, \hat{x}^i_{t-1|t-1}, x^i_{t|t-1}\}$ for a given L, where $\hat{}$ denotes a particle resampled in the previous time step.

(2.4) Random-Walk Metropolis–Hastings sampling

(1) Resample S^i with weight w^i_t corresponding to $x^i_{t|t-1}$.
(2) Generate a new particle governed by a random walk,

$$\tilde{x}^i_{t|t-1} = x^i_{t|t-1} + e_t, \quad (17)$$

where $e_t \sim N(0, \Sigma)$ and Σ is the variation.
(3) Accept/reject the candidate following the decision rule:

$$\hat{x}^i_{t|t-1} = \begin{cases} \tilde{x}^i_{t|t-1} & \text{with probability} \\ & 1 \wedge \dfrac{p(y_t|\tilde{x}^i_{t|t-1})p(\tilde{x}^i_{t|t-1} \leftarrow x^i_{t|t-1})}{p(y_t|x^i_{t|t-1})p(x^i_{t|t-1} \leftarrow \tilde{x}^i_{t|t-1})} \\ x^i_{t|t-1} & \text{otherwise} \end{cases}, \quad (18)$$

where $a\bar{\wedge}b = \min\{a, b\}$.

(2.5) Obtain $\hat{x}^i_{t-L|t-1}$ from S^i, and the filtering distribution from eq. (12).
(2.6) Update $t \leftarrow t + 1$.
(2.7) Compute $l(\theta)$ using eq.(15).

(3) Metropolis–Hastings sampling for the MCMC
For $\hat{\theta}'$, the parameter obtained in the previous step, accept θ with a probability of $1 \wedge \dfrac{l(\theta)}{l(\hat{\theta}')}$, for the other $\hat{\theta}'$.
(4) Update $\theta \leftarrow \theta + \Delta\theta$, where $\Delta\theta \sim N_n(0, \Sigma_n)$ and Σ_n is the variation.

When we use multiple observations, we obtain $l_m(\theta)$ in process (2) for each $m = 1, , M$ and then sum to obtain $l(\theta)$ using eq. (16).

2.3 Practical description

In this study, the time-series are seismic waveforms recorded at the surface and in the borehole. To satisfy the Markov process, x_t is approximately given by a first-order Taylor expansion around $t - 1$:

$$x_t = x_{t-1} + \frac{\partial x_{t-1}}{\partial t}\Delta t + v_t, \quad (19)$$

where the second term with the partial derivative can be obtained numerically and Δt is the time step. Assuming a linear relationship between the two waveforms, since the borehole is vertical and the depth of the borehole sensor is mostly within 100–200 m of the surface, which is significantly shorter than the wavelength of surface waves (>50 km), we can approximate the surface waveform y^{suf}_t at the surface based on the waveform recorded in the borehole, x^{sur}_t. Any difference between sensor pairs is attributed to the observation

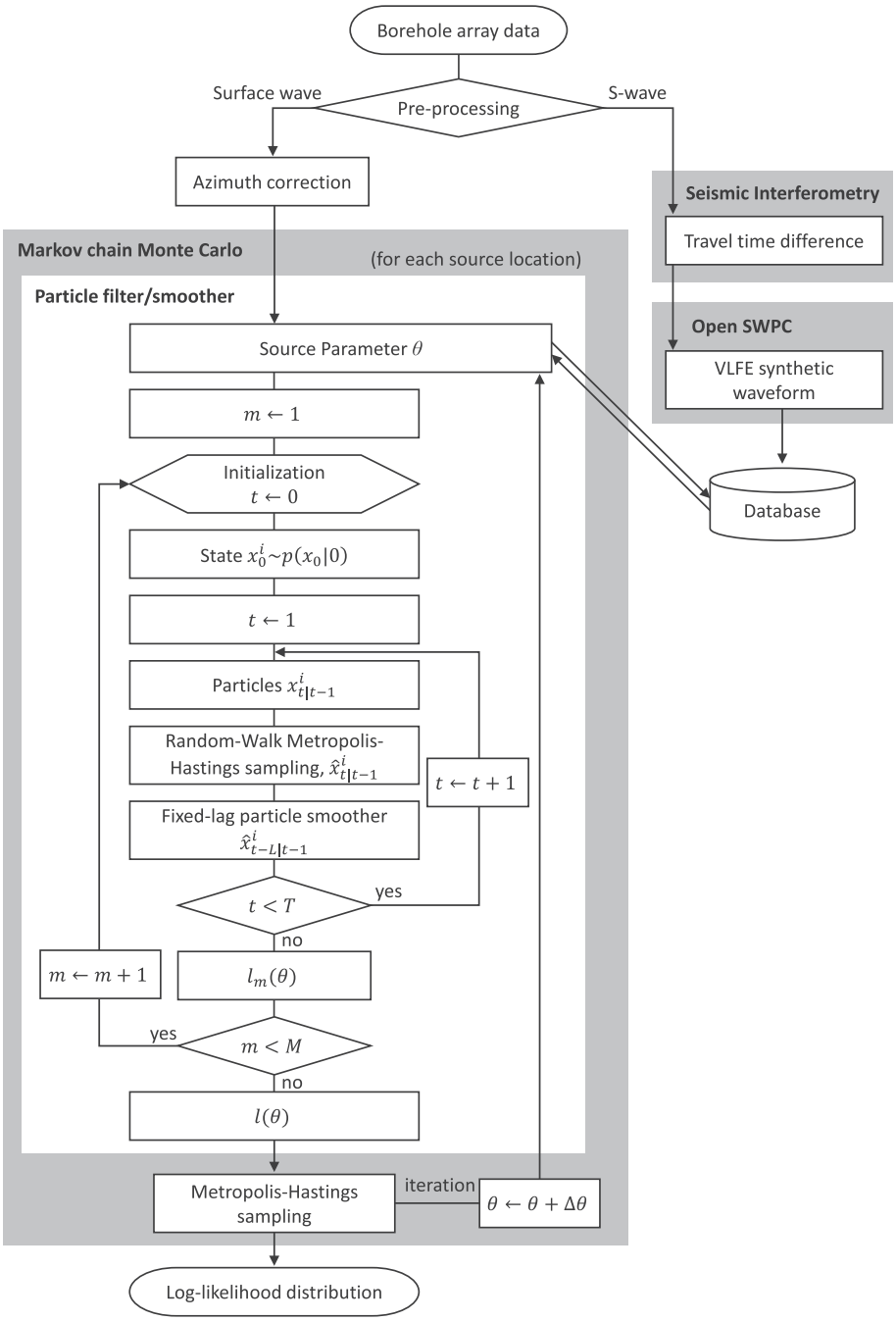

Figure 2. Workflow diagram for detecting VLF earthquakes.

of non-horizontal incident signals, y_t^{vlf} at the surface and x_t^{vlf} in the borehole, from other local events, or to the noise. Therefore, the observation equation (2) can be rewritten as

$$\begin{aligned} y_t &= y_t^{\mathrm{sur}} + y_t^{\mathrm{vlf}} \\ &= c x_t^{\mathrm{sur}} + w_t^{\mathrm{sur}} + h'(x_t^{\mathrm{vlf}}, \theta, w_t^{\mathrm{vlf}}), \quad \theta = (s, \alpha, t_0), \end{aligned} \tag{20}$$

where c is a constant, s, α and t_0 are the source location, relative magnitude and origin time of the VLF earthquake, respectively, and h' corresponds to changing a sensor location that observed a VLF earthquake (i.e. from borehole to the surface). For simplicity, it is assumed that the source parameters of the VLF earthquake, including the strike, dip, rake and source time function, are constants, and the magnitude, location and origin time are unknown. In general, hyperparameters can be estimated by maximum likelihood methods; however, in this study these values are determined empirically through trial and error. The noise terms w_t^{sur} and w_t^{vlf}, and random walk term e_t are assumed based on the average amplitude change at each time step. The weight in eq. (11) is obtained from the standard deviation characterized by the difference between state x and observation y. Most of the given distributions are Gaussian, while Cauchy-distributed system noise is assumed to predict the amplitude of a VLF event that is considerably smaller than the surface wave. This is because the waveform is simulated assuming source parameters that are not initially well-constrained, and we need to compensate for the discrepancy between the simulated and observed

VLF event waveforms in the generalized state-space model, which may be critically large.

3 SYNTHETIC TESTS

We assume that borehole array data are of sufficient quality to discriminate between waves arriving from different incidence angles. In the case of a VLF earthquake triggered by a distant earthquake, the observed waves from a VLF event at a source depth of around 30 km are mostly *S* waves in the region, and the incidence angle is almost vertical, which causes a traveltime difference between surface and borehole observations. Surface waves that propagate horizontally from a distant earthquake do not cause such a time delay. These characteristics can be utilized by applying this technique to seismic data acquired by KiK-net, a borehole seismic network in Japan. Taking the simple difference between the surface and borehole waveforms cannot highlight triggered events because this would eliminate both the surface waves and any triggered signal (if present).

We generate synthetic surface and borehole seismograms by superimposing surface waves made by tapering a sinusoidal wave on Ricker waves mimicking *S* waves from a VLF earthquake (Fig. 3). Both synthetic phase types have a predominant period of about 20 s. The surface waves recorded at the surface and in the borehole are assumed to be the same. The amplitude of the VLF event is about three orders of magnitude lower than that of the surface waves, which is represented by the scale factor $r = 1 \times 10^{-3}$. The surface waves observed at these sensors have no time shift, while the observed *S*-waves from the VLF earthquake have phase differences due to the traveltime difference between two sensors and reflection at the surface. The traveltime difference is assumed to be 0.5 s, but this value does not affect our results provided it is non-zero and distinguishable in the discretized data; that is >0.01 s when the sampling frequency is 100 Hz. White noise with amplitude two orders of magnitude lower than the VFL signal is embedded in the surface records.

In the particle smoother, y_t is the seismogram observed at the surface. We need to calculate $x^i_{t|t-1}$, the predicted wave in the borehole at time t. Given observations y_t in Fig. 3(a), reference VLF signals x^{vlf} and y^{vlf} in Fig. 3(b), and parameters t_0 and α, we solve for the VLF event signals at the surface and in the borehole. We do not solve for the source location because it can be included in t_0 in the present test. The predicted surface wave is obtained by removing the VLF event signals from the observation in the borehole and using eq. (19) to satisfy the Markov property. In practice, since a linear observation model is used and c is assumed to be 1.0 in eq. (20), we obtain the distribution by temporarily subtracting x^{srf}_t from y_t to avoid cumulative rounding errors and inherent drift apparent in the waveforms. The distribution of y_t is obtained in terms of likelihood for $t = 1, 2, , T$. If the parameters are incorrectly specified, then the predicted surface wave might also include VLF event signals. However, such an incorrect assumption yields a low likelihood. The median y^i_t for each t is chosen to generate the waveform y_t after applying particle smoothing. The number of particles is 1000 (i.e. $N =$ 1000), which is sufficient to obtain the distribution and constrain its 95 per cent confidence level. We assume that $L = 20$ for smoothing, following Kitagawa (1998). We search a parameter space in which t_0 ranges from –20.0 to 20.0 s and α ranges from 0.1 to 10.0, in which an observation corresponds to $t_0 = 0.0$ and $\alpha = 1.0$. For $\Delta\theta$, a Gaussian distribution is assumed whose components in time and amplitude follow linear and logarithmic scales and whose standard deviations are 1.0 s and 0.2, respectively. One standard deviation in the random-walk Metropolis–Hastings sampling in Section 2.2, is comparable to the smaller of the system noise or the observation noise.

Fig. 4(a) shows the sample distribution for parameters t_0 and α obtained by MCMC with 10 000 iterations, which is consistent with the log-likelihood distribution when we use a grid search for parameter estimation instead of MCMC (Fig. 4b). The high peak value at $t_0 = 0.0$ and $\alpha = 1.0$ means that we can successfully detect the VLF event signal embedded in the observed waveform. Local maxima around ±15 s correspond to one cyclic shift of the VLF event signal. The peak in time is resolvable to within several seconds. In general, the detection resolution at smaller α seems relatively poor compared with larger values, indicating the difficulty in detecting low-amplitude signals and the possibility of estimating the magnitude to within a confidence interval ranging from 0.5× to 2.0× the measured value or more. We can reliably detect the peak regardless of the phase relationship between the surface wave and the VLF event signal, because when we use other surface waves whose phases are random, similar results are obtained; this will be explored in the tests below. This robustness arises mainly because in the synthetic tests, we could perfectly estimate the incoming surface waves, removing the VLF event signals.

We performed additional tests to investigate the detectability of VLF event signals when the ratio of incoming surface-wave amplitudes to VLF signal amplitudes is less than 1×10^{-3}. Fig. 5 shows the log-likelihood distributions at ratios of $r = 1 \times 10^{-3}$, 1×10^{-4}, and 1×10^{-5}, where the parameter to be determined is either t_0 or α in other words, the parameter is t_0 while α is fixed at 1.0 in Fig. 5(a) and the parameter is α while t_0 is fixed at 0.0 in Fig. 5(b). In this test, 10 observations (i.e. $M = 10$) are employed, where surface waves are generated by randomly changing the phase. The peak used for time detection appears at $r \geq 1 \times 10^{-4}$. When we increase the amplitude of the white noise while keeping r constant (e.g. $r = 1 \times 10^{-3}$ and s/n $= 1 \times 10^{1}$), the detectability becomes very poor and is the same as when $r = 1 \times 10^{-4}$, which means that if the signal-to-noise ratio is sufficiently large (e.g. $\geq 10^{1}$) then we can detect the VLF event. Since no signal filtering is applied to remove noise during the synthetic tests, the results show the importance of removing noise before applying a particle smoother to detect a signal. In addition, stacking the likelihood values of multiple channels can improve the detection, regardless of the phase differences between the incoming and signal waves. As shown in the previous test, α seems poorly resolved relative to t_0.

4 APPLICATION TO THE 2016 M_{w} 5.9 EARTHQUAKE OFF KII PENINSULA, JAPAN

We apply this method to seismic data from the M_{w} 5.9 earthquake, which occurred on 1 April 2016 (UTC), off Kii Peninsula, Japan (Global Centroid Moment Tensor, CMT) (Fig. 1). This event satisfies all the following requirements for application of our method:

(1) The earthquake is located far from the array, so that body waves and phases from aftershocks do not contaminate the surface waves.

(2) The earthquake is large enough that we can expect triggering of VLF earthquakes on the basis of previous studies of LFEs, but the surface wave amplitude does not exceed the probable VLF event signal amplitude by four orders of magnitude.

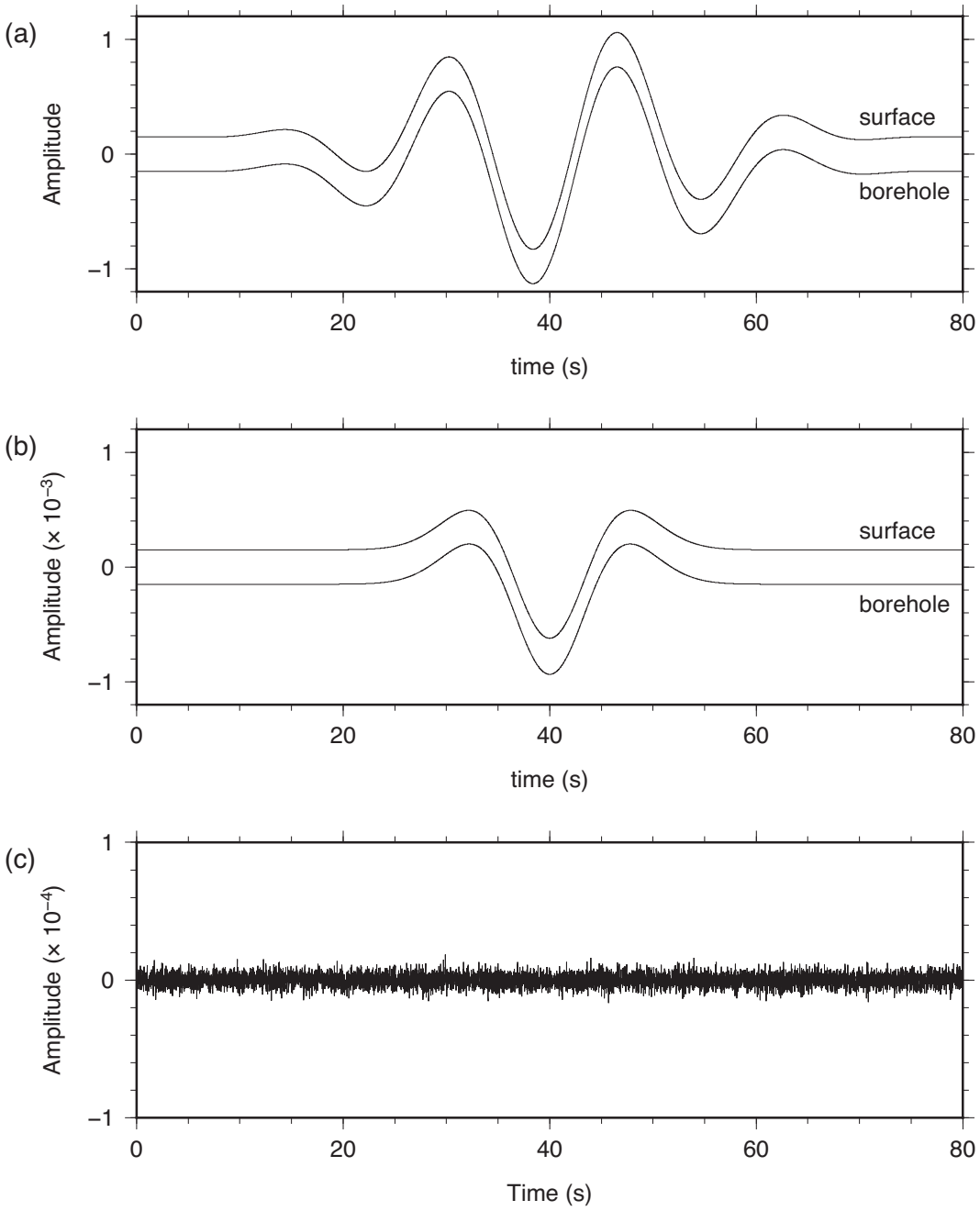

Figure 3. Waveforms used for synthetic tests. (a) Synthetic waveforms at the surface and in a borehole. (b) *S*-waves from a VLF earthquake. (c) White noise embedded in the surface record.

(3) The surface wave arrivals are observed at KiK-net borehole array stations.

The last condition is stringent because KiK-net is designed for observations of strong ground motion near the source, and recording relies on an event trigger system that records for 120 s from the trigger time; thus, even if an earthquake satisfies the first two conditions, the waveforms of surface waves will not be recorded if an event occurs far from the stations or no strong motion is observed in the region of interest. To date, the 2016 M_w 5.9 earthquake is the only event that meets all three requirements for a VLF earthquake search in western Japan.

4.1 Data selection and preprocessing

We select waveforms from KiK-net stations to search for triggered VLF earthquakes. Some KiK-net stations in the region, where VLF earthquakes have been reported previously, recorded the 2016 off-Kii earthquake. It is important to accurately predict the incoming surface waves at surface stations from borehole observations, so the observed borehole and surface waveforms should be quite similar. First, the recorded acceleration data are integrated to obtain velocity waveforms, which are bandpass filtered from 0.025 to 0.100 Hz. This frequency band includes the VLF event signal band, and these corner frequencies can remove the predominant high-frequency noise around 1 Hz. The horizontal components of

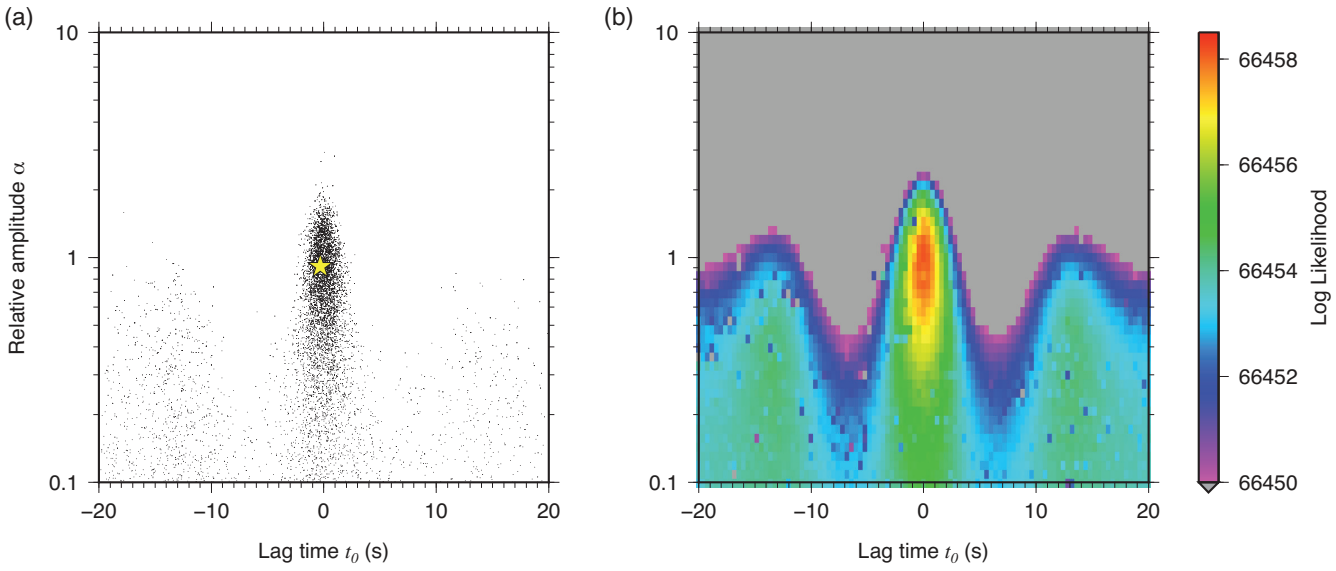

Figure 4. Parameter estimation. (a) Markov chain Monte Carlo samples for α and t_0. The star indicates the highest populated point in a grid area of 0.4 s times 0.04 in log-scale. (b) Log-likelihood distributions by grid search. In both cases, the peak of the log-likelihood plot appears at $\alpha = 1$ and $t_0 = 0$, which are presumed to be estimated.

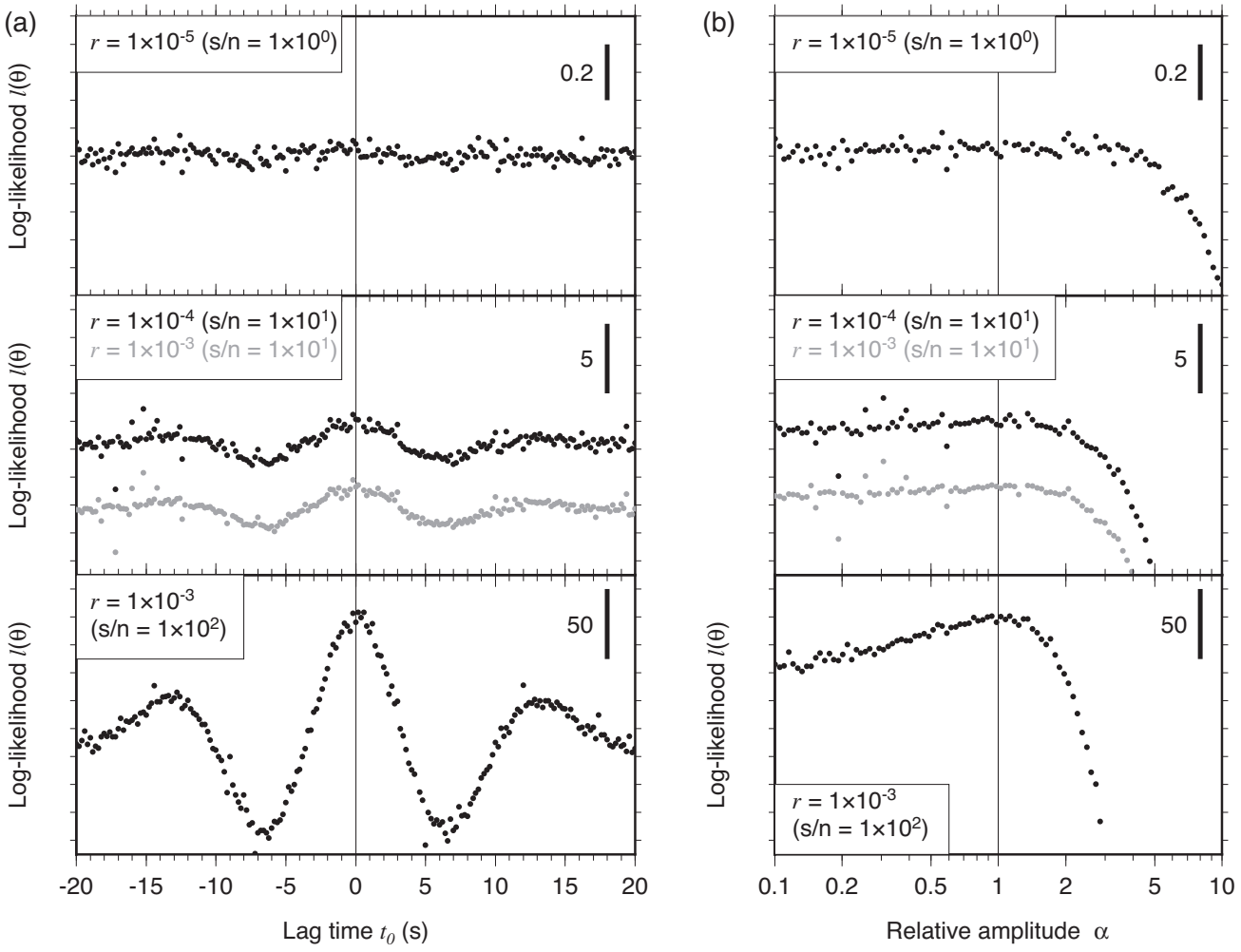

Figure 5. Log-likelihood distribution of the amplitude ratio r of VLF event signal to incoming wave. (a) and (b) show the lag time t_0 and the relative amplitude α, respectively, where the lag time $t_0 = 0$ and $\alpha = 1.0$ are treated as true values. The vertical bar in each window shows the amplitude of log-likelihood.

the borehole sensor are corrected for azimuthal deviations. Though the National Research Institute for Earth Science and Disaster Resilience (NIED) provides azimuths of borehole sensors determined from teleseismic phase arrivals (e.g. Shiomi *et al.* 2003), more precise values are obtained from the present test data by rotating the components in 0.01° increments to maximize the cross-correlation between the north–south (NS) and east–west (EW) components of the surface and borehole instruments. In this procedure, filtered surface wave arrivals are extracted from 100 to 200 s after the origin time, and then a 5 s cosine taper is applied to both ends of each seismogram to exclude early body wave arrivals. Hereafter, $t = 0$ is referred to as the origin time unless another meaning is specified. We also correct the amplitude of the borehole sensor so that $c = 1.0$ in eq. (20). Data are chosen for further analysis if the variance reduction of the velocity waveforms at the surface and borehole sensors is larger than 0.99, which roughly corresponds to a root-mean-square-error smaller than 1×10^{-2} mm s^{-1}. Station EHMH12 is required for reliable detection of VLF earthquakes in the western region because no other stations are located above the western LFEs and VLF earthquake region, so we extracted its waveform data from 100 to 180 s and omitted the last 20 s of data because of an unexplained increase in the variance. Hence, selected waveform data from 21 channels at 14 stations (M) are shown in Fig. 1. Table 1 gives the rotation angles (positive values indicate clockwise) used in this study, which are consistent with the values of the NIED. Raw and processed waveform data are shown in Fig. 6 and Fig. 7, respectively. Waveforms at the surface and in the borehole were plotted by applying a slight offset to the illustration because they are otherwise indistinguishable at this scale. Rayleigh waves propagating at 3.4 km s^{-1} are seen in this time window on the EW and up–down (UD) components of the seismograms, while Love wave amplitudes appear small, which may cause a poor correlation between surface and borehole waveforms and could explain why fewer waveforms were chosen from NS component channels than from either of the other two components.

4.2 Prediction of VLF event *S*-waves

The S waves from a VLF earthquake, x_t^{vlf}, are calculated numerically based on the parameter θ. We assume that the mechanism of a VLF earthquake has strike = 225°, dip = 15° and rake = 90°, for a fault plane consistent with the subducting slab geometry, and we assume a rise time of 20 s based on Ito *et al.* (2007). The magnitude is characterized by the relative amplitude coefficient α: if $\alpha = 1.0$, the moment magnitude $M_{\mathrm{w}} = 4.0$, as this is thought to be the upper limit of the expected magnitude of a large VLF event (Takeo, pers. comm., 2017). A grid search is used to determine the source location on the plate interface with a grid spacing of 0.125° longitude and 0.1° latitude, and a depth range of 25–40 km (Fig. 1), where VLF events have been found by previous studies (Ito *et al.* 2007, 2009; Baba *et al.* 2018). The Japan Integrated Velocity Structure Model (JIVSM, Koketsu *et al.* 2012) is used to determine the depth of the plate interface. A similar grid-search approach for detecting VLF earthquakes has been used in previous studies (e.g. Suda & Matsuzawa 2017; Baba *et al.* 2018).

We use the OpenSWPC software of Maeda *et al.* (2017) for elastic wave simulation with the JIVSM (Koketsu *et al.* 2012). This model is considered to appropriately estimate source waveforms of VLF events in the subduction zone because the synthetic amplitudes computed with JIVSM significantly differ from those computed with a 1-D velocity structure. Whereas amplitude or magnitude may be poorly estimated compared with the origin time according to the synthetic tests, the more realistic waveform prediction is required to minimize the uncertainties. This approach solves the equations of motion in 3-D Cartesian coordinates with viscoelastic constitutive equations using a finite-difference method to the fourth order in space and the second order in time. The spatial grid size is 0.5 km in the horizontal and vertical directions, and the time step is 0.025 s. Quadratic interpolation is employed such that the observed waveforms are sampled at 100 Hz. The computational region has an area of 250 km × 250 km and extends to 100 km depth (see Fig. 1). At the boundaries, we apply an absorbing boundary condition from the auxiliary differential equation. A complex frequency-shifted perfectly matched layer (Zhang & Shen 2010) is used to avoid artificial reflections. It is assumed that the calculated waveforms consist mainly of S waves.

The vertical grid spacing of 0.5 km is too crude to differentiate between synthetic waveforms at the surface and in the borehole, particularly because the borehole sensors are typically located within a few hundred meters of the surface. Arrival time errors may appear in synthetic records due to epistemic uncertainties in structure and the source–time function, but these can be corrected using the distribution of t_0. On the other hand, accurate estimation of traveltime differences between surface and borehole sensors at a single site is critically important in our analysis. We use seismic interferometry to measure this traveltime difference (e.g. Nakata & Snieder 2011, 2012) using S waves from the 2016 off-Kii event. The advantages of this method include that the traveltime can be measured at any KiK-net station, regardless of whether logging data are available, and it can also be measured for the long-period S waves with which we are concerned. We deconvolve the surface record u_s out of the borehole record u_b, which in the frequency domain is given by

$$G(x_b, x_s, \omega) = \frac{u_b(\omega)}{u_s(\omega)} \approx \frac{u_b(\omega) u_s^*(\omega)}{|u_s(\omega)|^2 + \varepsilon}, \tag{21}$$

where $G(x_b, x_s, \omega)$ is the Green's function that accounts for wave propagation from the surface sensor x_s to the borehole sensor x_b, and ε is a water-level parameter to stabilize the deconvolution. Fig. 8 shows an example of traveltime measurement between sensors at KiK-net station EHMH04 (Fig. 1). Upward- and downward-propagating S waves arrive at negative and positive lag times, respectively. The traveltime between two sensors is estimated by averaging the absolute traveltimes at negative and positive lag times across the two horizontal components. The measurement includes an inherent error of at least 0.01 s due to the sampling interval; thus, measured traveltimes at each station are consistent with values obtained from logging data (Table 1). The traveltimes measured by the deconvolution method are generally equal to or less than those estimated from logging data because of differences in the frequency content of the seismic waves used for the measurements.

Using these traveltime differences and assuming total reflection of upward-propagating S waves at the surface, we compute synthetic waves from a source at each gridpoint on the plate interface. Fig. 9 shows an example of synthetic waveforms at station EHMH04 from an M_{w} 4.0 VLF earthquake beneath the station. The ratio of the VLF event amplitude to the surface wave amplitude is about 1.0×10^{-3}, for which our synthetic tests suggest that the VLF event can be detected.

Table 1. Horizontal rotations and traveltime difference at KiK-net stations.

		Azimuth correction (°)		Traveltime difference (s)	
Station name	Component	NIED	This study	Log-data	This study
EHMH03	EW, NS	0	1.08	0.08	0.06
EHMH04	EW, NS, UD	−1	−0.79	0.36	0.31
EHMH05	UD			0.13	0.08
EHMH10	EW, UD	−1	1.92	0.12	0.11
EHMH11	UD			0.07	0.05
EHMH12	UD			N/A	0.15
HRSH07	EW, UD	2	2.22	0.15	0.14
HRSH15	UD			0.09	0.06
KOCH02	EW, UD	4	3.60	0.12	0.05
KOCH05	EW, UD	2	2.96	0.06	0.05
KOCH10	UD			0.07	0.07
KOCH12	UD			0.17	0.15
TKSH06	UD			0.14	0.13
YMGH17	UD			0.18	0.14

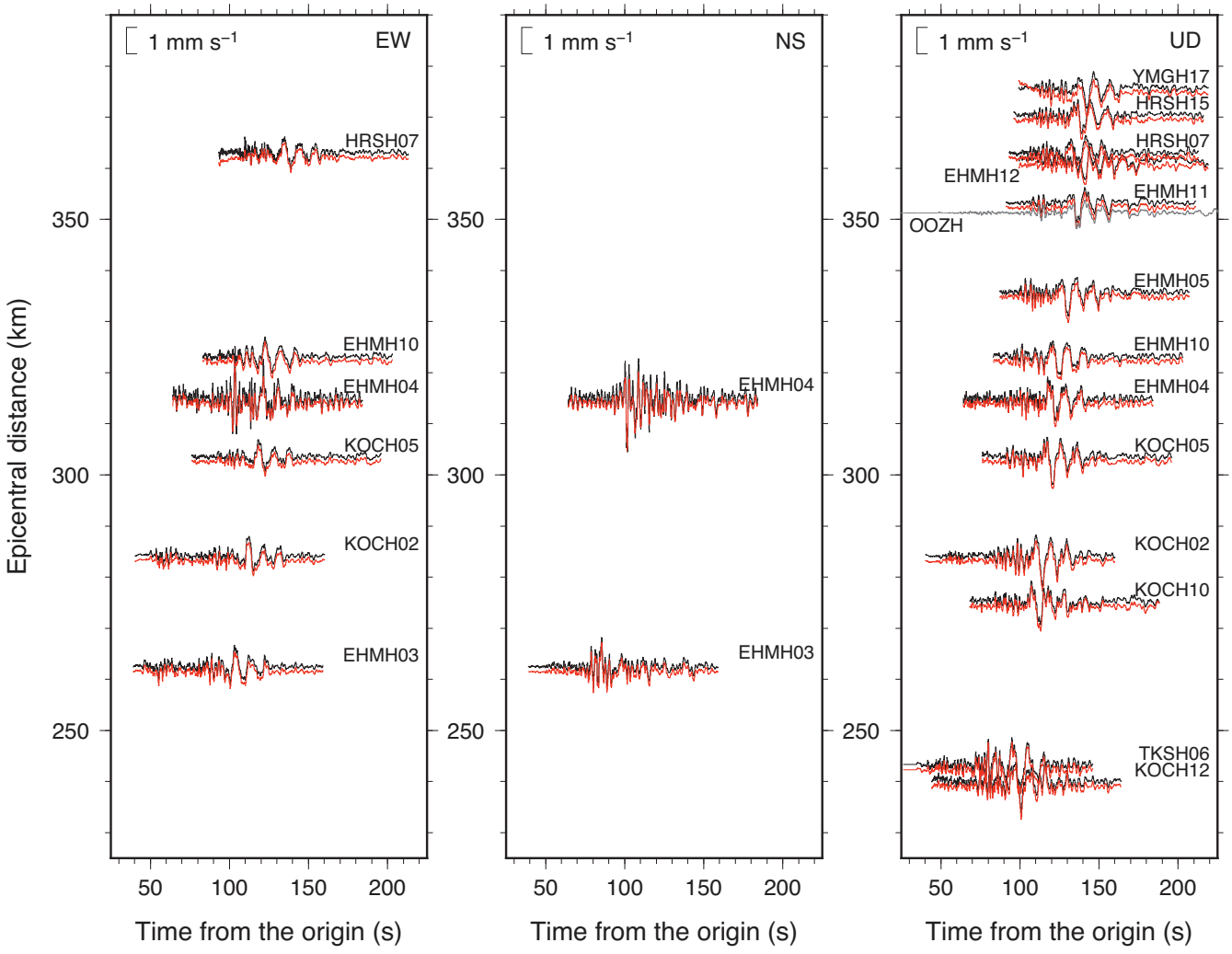

Figure 6. Velocity waveforms of the 2016 off-Kii earthquake, Japan (1 April 2016, 02:39:07.85 UTC) for the east–west (EW), north–south (NS) and up-down (UD) components. Surface (black) and borehole (red) data are shifted slightly based on epicentral distance for improved visual clarity. Waveform data are high-pass filtered at 0.01 Hz. Station locations are shown in Fig. 1. At EHMH11, continuous waveform data recorded by Hi-net station OOZH in the same borehole are shown in grey to illustrate the *P* and *S* arrivals.

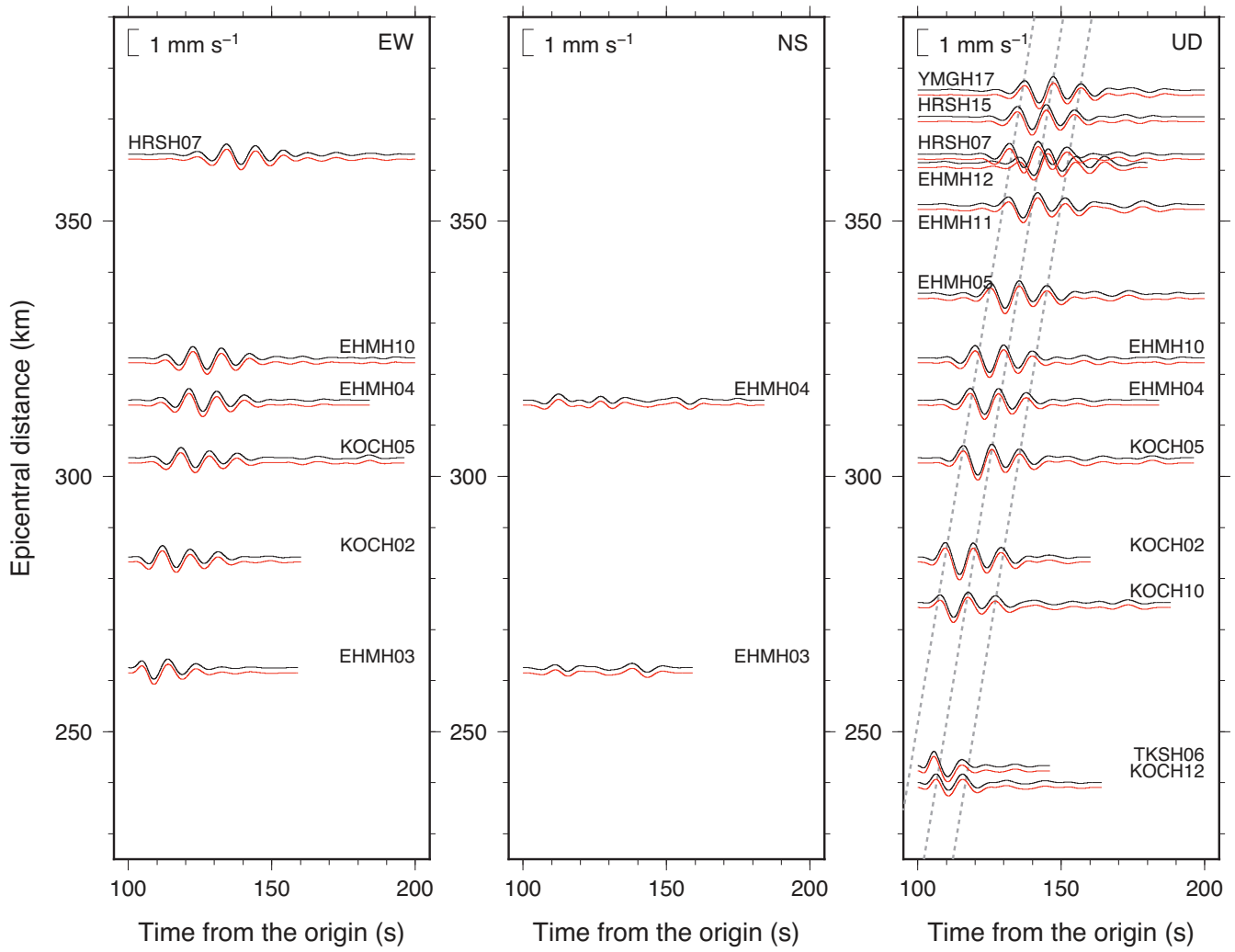

Figure 7. Preprocessed velocity waveforms of the 2016 off-Kii earthquake. Data shown are bandpass filtered from 0.025 to 0.100 Hz. Surface records are in black and borehole records in red. The three vertical velocity peaks of Rayleigh wave propagation at 3.4 km s^{-1} are indicated by dotted lines.

4.3 Estimated distribution of parameters and detected VLF earthquakes

We apply our particle smoother to data from the 2016 M_w 5.9 off-Kii earthquake and combine the grid search and MCMC methods to search the parameter space θ related to the VLF earthquake. Fig. 2 shows the workflow outlining the strategy for detection. We set $N = 100$, which seems small compared with the number of synthetic tests, but is sufficient to constrain the median of y_t^i since we are not interested in the distribution of y_t. In the MCMC process we compute 10 000 iterations to search the parameter space of t_0 and α. A grid-search approach is adopted to estimate the hypocentre. These calculations are performed using a parallel computing array, which considerably decreases the computation time. For $\Delta\theta$ the standard deviations of time and log-amplitude are 2.5 s and 0.2, respectively.

Fig. 10 shows the distributions of maximum log-likelihood values in 3 s time windows centred on the times indicated in each text label, where each value is a sum of the maximum log-likelihoods over all channels in the time window (i.e. eq. 16). This is equivalent to correcting for traveltime errors caused by assuming source locations only on discrete gridpoints and errors introduced by structural heterogeneities that are not described by the velocity model. Spatial distributions are smoothed by averaging values over 0.25°. Maps are generated in 1 s intervals (see Supporting Information for animation). Values in each time window are compared with Rayleigh wave arrivals estimated from three major phases in vertical velocity waveforms, while Love wave amplitudes are relatively small (Fig. 7). The estimated Rayleigh wave phase velocity is 3.4 km s^{-1} at these frequencies, while surface waves appear to propagate more slowly across the southern part of the array due to structural heterogeneities. The maximum log-likelihood values are generally small in the western region, indicating low detectability due to poor coverage by the seismic array. If vertically incident body waves originate from a source other than VLF events, the large log-likelihood values could appear at gridpoints closest to and near the epicentre at its origin time. Therefore, the brightly shaded regions closest to the off-Kii earthquake could correspond to observations of *S* waves or body waves from early aftershocks, but these cannot be found in either the JMA catalogue or in seismograms recorded closer to the hypocentre of the main shock.

We identify likely VLF event detections using the distribution of all spatiotemporal values in log-likelihood maps. When the maximum log-likelihood value exceeds a nominal threshold level of 549 648.50, which corresponds to 9.0 × MAD (median absolute

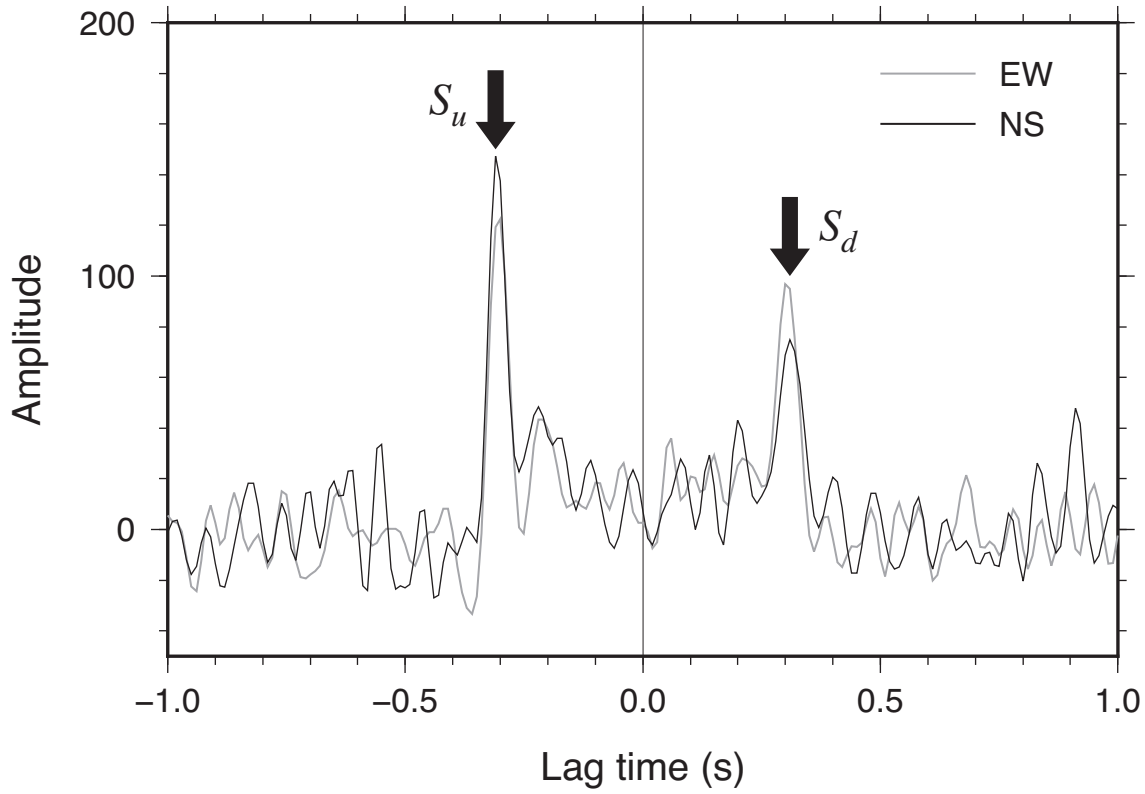

Figure 8. Example of *S* waves detected using deconvolution interferometry at KiK-net station EHMH04. Waveforms deconvolved from EW- and NS-component data are shown in grey and black, respectively. Upward- and downward-propagating *S* waves (S_u and S_d, respectively) appear at negative and positive lag times, respectively (arrow).

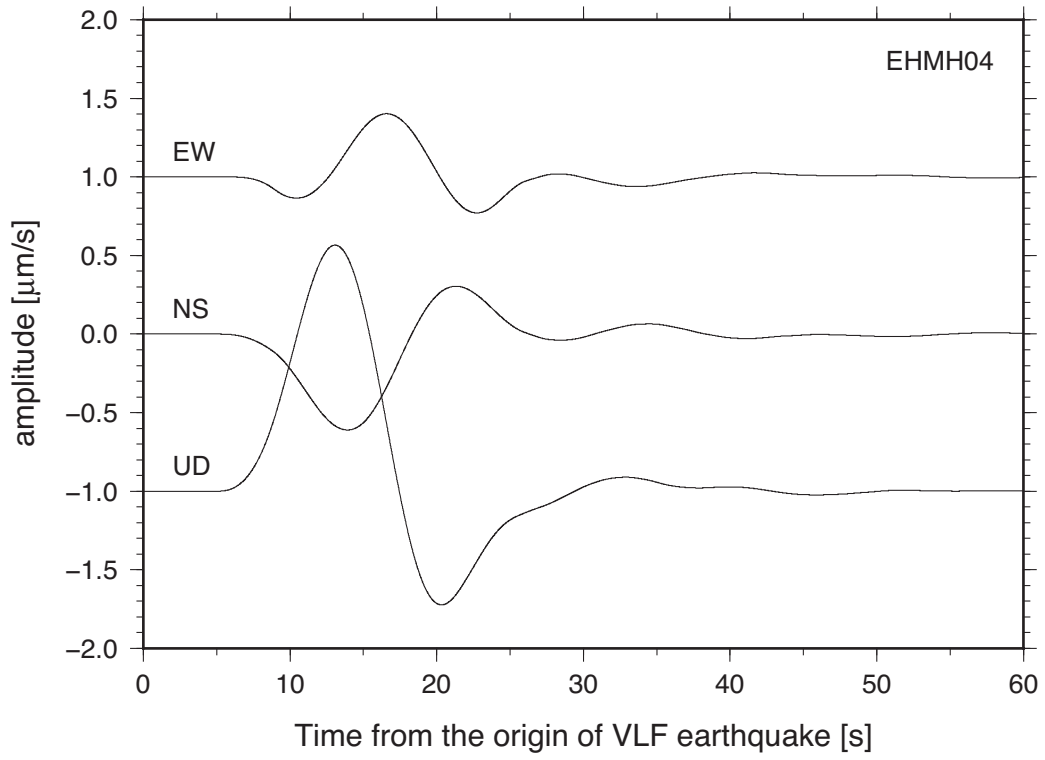

Figure 9. Synthetic waveforms at station EHMH04 for a M_w 4.0 VLF earthquake at a gridpoint beneath the station. Time $t = 0$ corresponds to the origin of the VLF event.

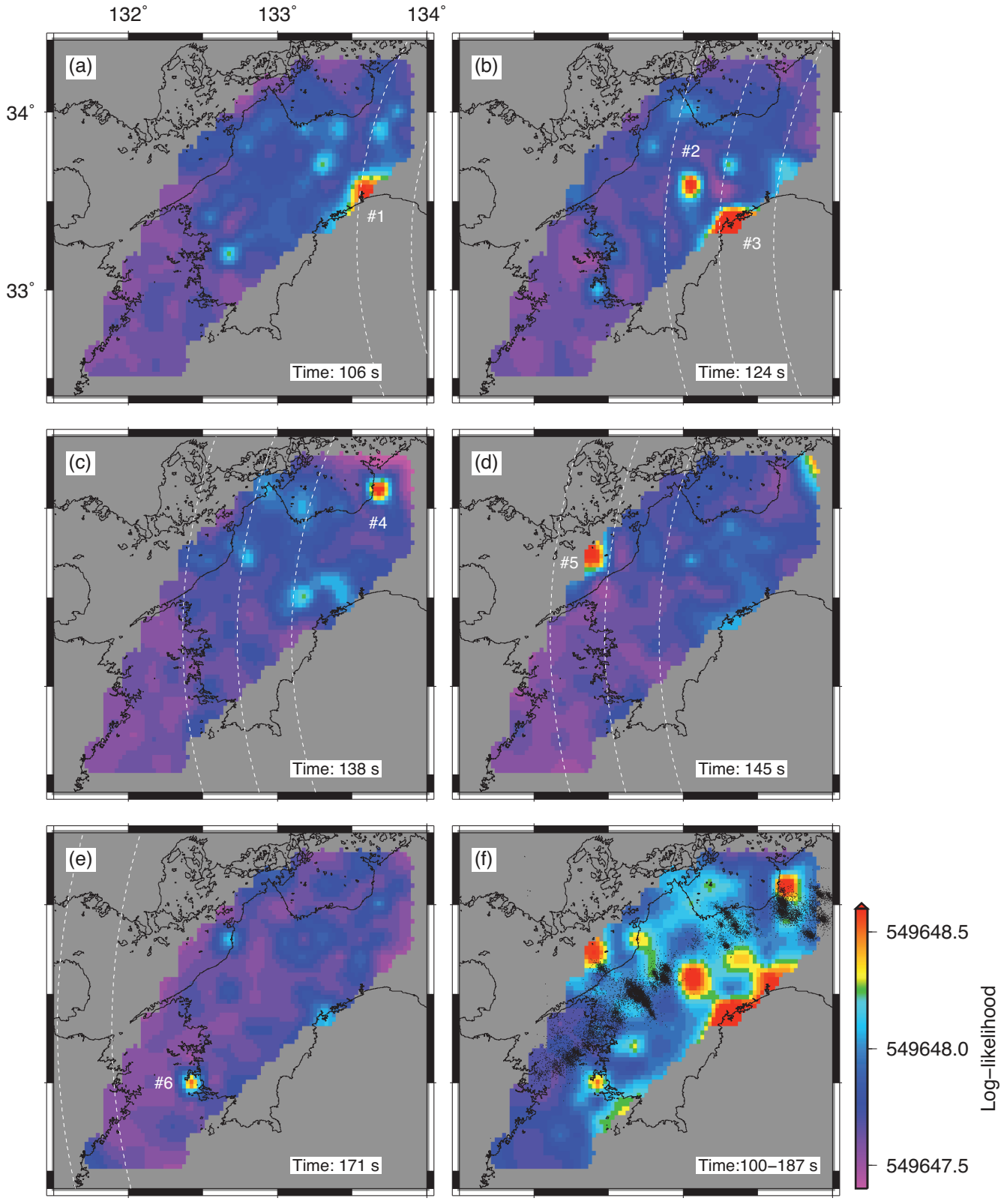

Figure 10. Log-likelihood distributions of surface waveforms at different times after the origin of the 2016 off-Kii earthquake. (a) Origin time of event #1, (b) events #2 and 3, (c) event #4, (d) event #5, (e) event #6 and (f) the maximum log-likelihood in the time range analysed. Maxima computed in 3s time windows centred on each point are spatially smoothed by the minimum curvature for (a)–(e). White dashed curves correspond to peak arrivals of Rayleigh waves in the vertical component that propagate at a phase velocity of 3.4 km s^{-1} (see Fig. 7). The locations of LFEs are indicated by black dots in (f).

deviation from the average), a VLF earthquake detection is declared. At least six events are detected from 100 to 180 s after the main shock origin time (Fig. 10; Table 2). Table 2 lists the catalogue of detected VLF earthquakes, where the moment magnitude is estimated from the relative amplitude α. The horizontal resolution of the hypocentre is about 10 km, and event depths are fixed on the plate interface of the structural model, JIVSM. The first event occurs beneath central Shikoku island at relatively shallow depths (Fig. 10a). The second event is found along the plate interface downward from the first event location (Fig. 10b). The third event is located near the first (Fig. 10b). The fourth to sixth events are scattered (Figs 10c–e). Events 1 and 3 are relatively shallow. The other four events seem to occur at the shallow edge (2 and 6) or deep edge (4 and 5) of the seismogenic region of LFEs (Figs. 1 and 10f). Four events (1, 2, 3 and 5) occurred during the passage of high-amplitude surface waves, while the others occurred afterward. Fig. 11 shows the extracted waveforms of VLF events after particle smoothing; their amplitudes are about three orders of magnitude smaller than those of the surface waves (Fig. 7). Note that the magnitudes are approximate and have large uncertainties.

5 DISCUSSION

5.1 Triggered VLF earthquakes and triggering mechanism

The background rate of VLF seismicity in western Shikoku (Fig. 1) is about 250 events per year when the 2016 M_w 5.9 earthquake occurred (Baba *et al.* 2018). The detection of only six events in a time window of 80 s during and after the passage of surface waves indicates either extremely high background activity or dynamic triggering.

We assumed that triggered VLF earthquakes were located in the same region as previous VLF events (Fig. 1). Events 1 and 3 have shallower locations, which suggests their true locations could lie outside of the assumed region, which reduces the reliability of the detections because the sources are outside of the seismic array. The two shallow events indicate that VLF events can occur in this region; otherwise, they might be due to rare coherent noise in observations. They cannot be false positive detections due to scattering of seismic waves in the structure, as the high log-likelihood values are not produced by incoming waves; similar reasoning rules out early aftershocks of the 2016 off-Kii earthquake. Events 4 and 5 are very close to VLF earthquakes detected by Ito *et al.* (2009). The complementary distribution of VLF earthquakes (2, 4, 5 and 6) and LFEs might be intuitive, but more precise locations are required to understand the spatial relationship (Fig. 10f).

The estimated magnitudes of the triggered VLF earthquakes range from 3.4 to 4.3, which again include large uncertainties, but these values might be larger than those of the background VLF events. If so, this relationship is similar to that of triggered vs. background LFEs (e.g. Miyazawa 2012), and triggered LFEs can be as much as one order of magnitude larger than ambient events when the triggering stresses are extremely large.

We model the initial physical rupture process of the triggered VLF earthquakes using a linear kernel approach that computes continuous waveforms spanning the full spectrum at depth (Miyazawa & Brodsky 2008), adapted for various phase velocities of surface waves (Miyazawa 2015); see the Appendix for details. For each triggered VLF event, we choose the Hi-net station closest to the event because Hi-net continuously records the ground motion. We calculate the dynamic stress change for VLF event slip at the depth of the plate boundary, as discussed in Section 4.2, and correct for the traveltimes of seismic waves from the 2016 off-Kii earthquake. Fig. 12 shows the relationship between the Coulomb failure stress changes (ΔCFF) after time corrections and the origin times of triggered VLF events, where we assume an apparent friction coefficient $\mu' = 0.4$. Since the origin time includes large errors due to unresolved fine-scale features in the velocity structure and uncertainties in the source time function, the value is tentative, with an error of ± 2 s. A VLF event seems to be triggered when stress changes are relatively high (0.4–1.5 kPa), comparable to the stress changes associated with earth tides. For the two shallow events 1 and 3, the origin time is consistent with the peak of ΔCFF. Events 4 and 6 are triggered when ΔCFF from the coda of the passing surface waves is large. Events 2 and 5 seem to occur before the corresponding peaks in the shear stress changes. The poorly resolved origin time of each event means that we could compare changes in amplitude during VLF events with stress changes resolved for the slip. In this case, which corresponds to events 2 and 5, the moment rate function of the source should be used to examine the relationship. The relationship between stress changes and event triggering is not as clear here as for LFEs (e.g. Miyazawa & Brodsky 2008; Miyazawa *et al.* 2008). Eventually, the origin times of triggered VLF earthquakes likely correspond to large stress changes from passing surface waves; however, since source–time functions are long compared with the durations of stress changes from incoming waves, the details of the initial rupture process remain unclear.

We found no evidence for triggering of LFEs during the passage of surface waves from the 2016 off-Kii earthquake, while the triggering stress changes were similar to those reported previously for LFEs triggering. The early surface waves are contaminated by the high-frequency S-wave coda, which inhibits detection using traditional high-pass filtering (e.g. Miyazawa & Mori 2006). This does not necessarily mean that no LFEs are dynamically triggered, but it is intriguing that only triggered VLF events have been detected so far. It may be that VLF earthquakes are more sensitive than LFEs to stress changes from surface waves, possibly because the dominant frequency of the triggering surface waves is similar to the rise time of the VLF source–time function. It is possible that more VLF earthquakes might be triggered by other large, distant earthquakes.

5.2 Triggering by the 2011 Tohoku-Oki earthquake

KiK-net recorded the 2011 M_w 9.0 Tohoku-Oki earthquake even in western Japan, at epicentral distances of >800 km. Some KiK-net stations and Hi-net stations in western Shikoku recorded LFEs with amplitudes of 1 μm s^{-1} remotely triggered by surface waves from the 2011 Tohoku-Oki earthquake because the event waveforms included significant high-frequency (>1 Hz) energy (Miyazawa 2011, 2012). The Tohoku-Oki earthquake might also have triggered VLF events, because they might be accompanied by LFEs. In this case, the surface wave amplitudes in western Japan reached about 1 m s^{-1}, six orders of magnitude higher than the expected amplitudes of VLF event signals. The root-mean-square error between the surface and borehole sensors is at least 1 mm s^{-1}, which our synthetic tests suggest is too large for our method to apply.

5.3 Resolution for detectability

Our tests suggest that a signal with amplitude three orders of magnitude lower than that of passing surface waves can be detected,

Table 2. Detected VLF events.

Event No.	Origin time from M_w 5.9 event (s)	M_w	Lon (°E)	Lat (°N)	Depth (km)
1	106	4.2	133.550	33.600	25.5
2	123	4.3	133.050	33.600	29.9
3	124	3.8	133.300	33.400	25.1
4	139	4.0	133.675	34.100	33.5
5	145	4.1	132.425	33.700	39.0
6	171	3.4	132.425	33.000	28.1

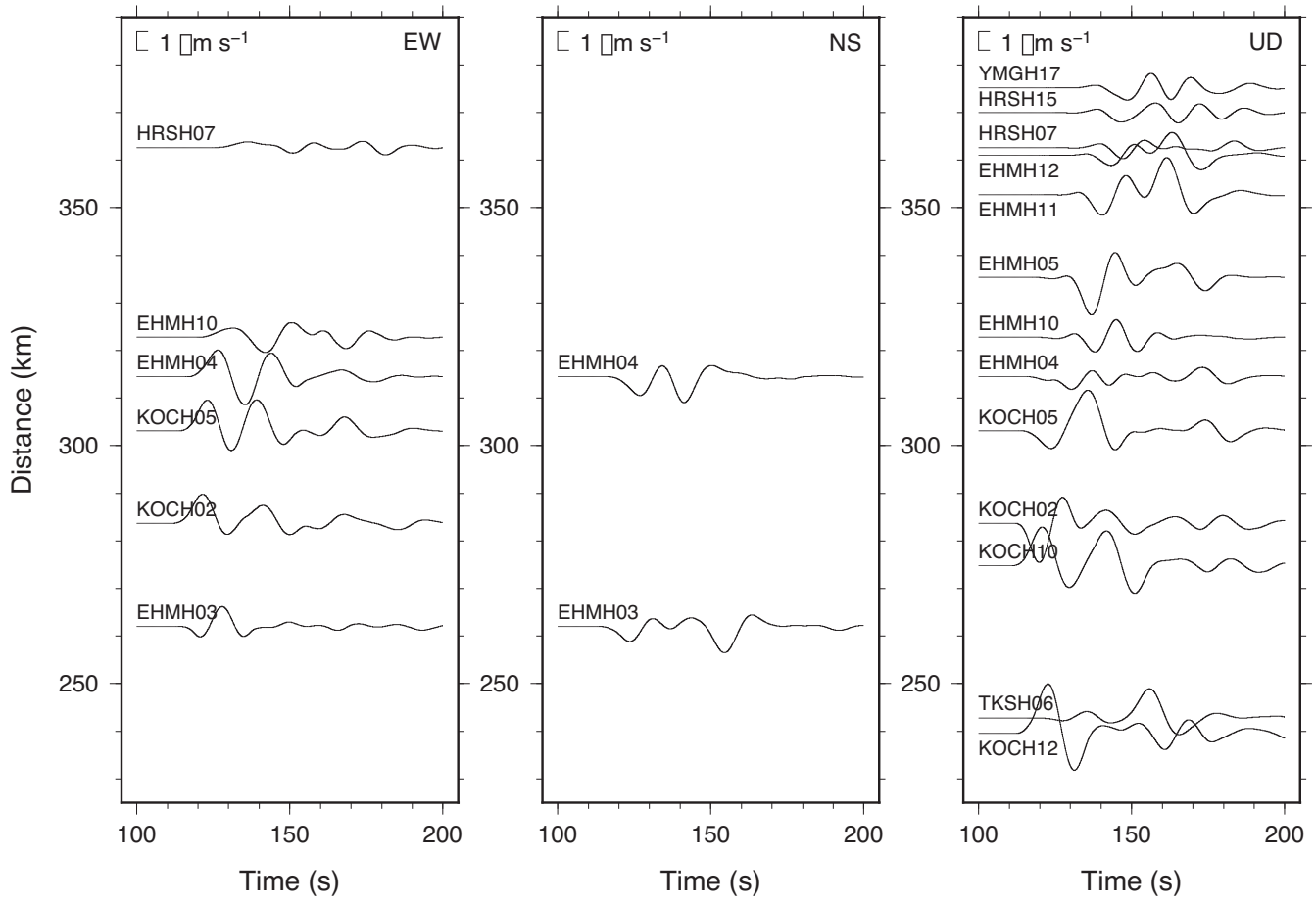

Figure 11. Waveforms of VLF event signals extracted from Fig. 7 by the particle smoother.

though the noise levels must be low. However, KiK-net accelerometers have poor amplitude resolution, about 1 μm s^{-2}, because these sensors are designed to record the phases of large earthquakes without clipping. When the acceleration time-series are integrated to produce velocity records, the resolution of VLF event signals observed by KiK-net might be 10 μm s^{-1}, which is nearly one order of magnitude higher than the peak amplitudes of detected VLF events; as a result, detection of VLF signals might be impossible using data from only one station. We overcome this limitation by stacking 21 channels from 14 stations. The high apparent log-likelihood values could also be caused by inaccurate predictions of waveforms of surface waves at surface stations based on borehole data and waveforms of VLF events, due to errors in the azimuthal corrections of the borehole sensors, errors in the structure model used to calculate synthetic waveforms of VLF earthquakes, and differences between assumed and actual mechanisms of VLF earthquakes. Although we have carefully attempted to remove these effects throughout our processing and analysis, unforeseen factors could introduce artificially high log-likelihood values. The magnitudes of detected events are likely less constrained than their triggering times.

6 CONCLUSIONS

Our novel combination of particle filtering and MCMC has resulted in the first known detections of dynamically triggered VLF events. Although the 2016 M_w 5.9 earthquake is the only event suitable for use with the present method, KiK-net records of teleseismic and/or continuous waves could allow us to detect additional triggered VLF seismicity, which would improve our understanding of VLF source mechanisms. Even without a vertical borehole array, this method could be used to detect triggered events with a dense seismic array whose average station spacing is smaller than the wavelength of incident phases, provided that the backazimuth of the triggering

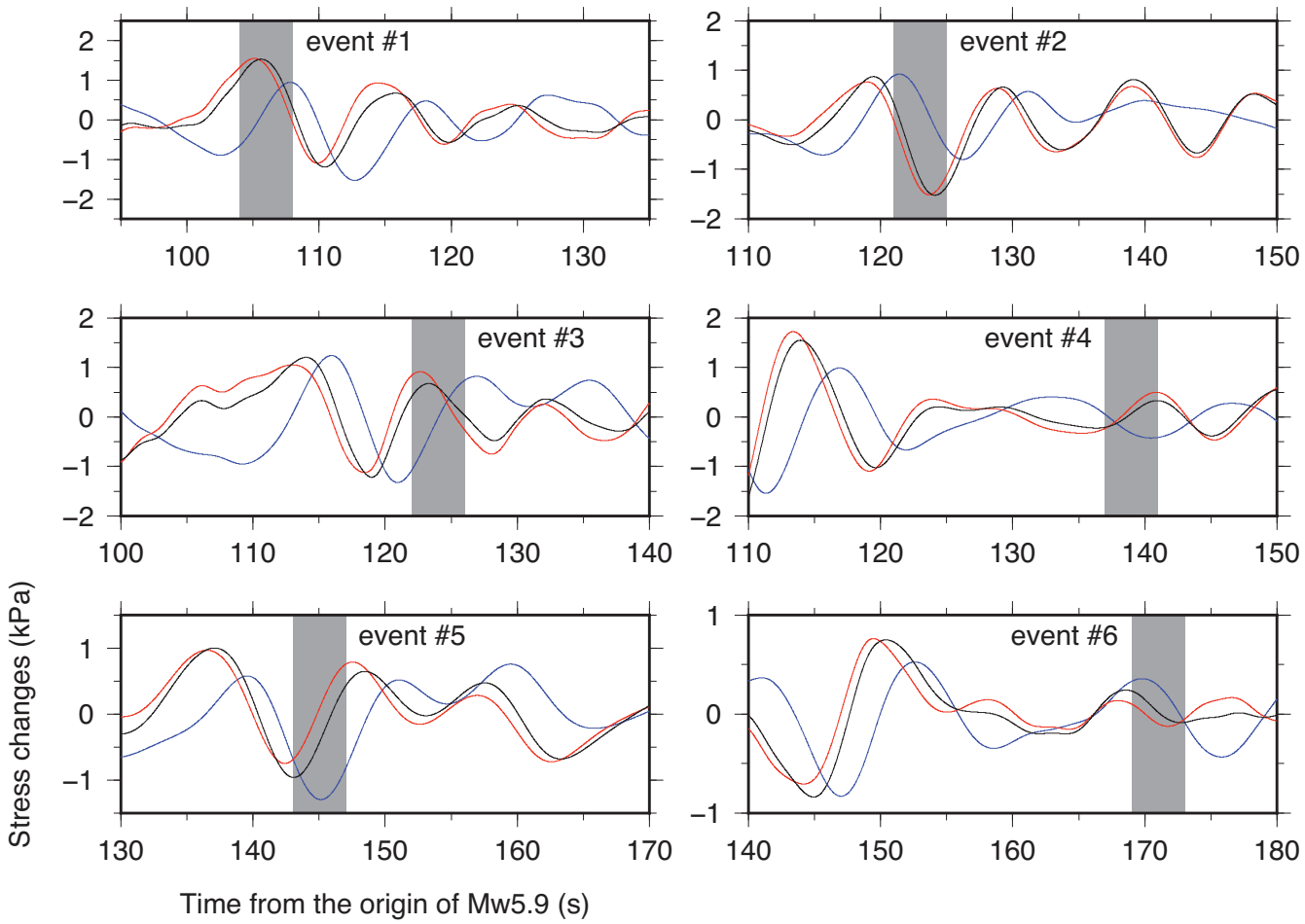

Figure 12. Stress changes during each triggered event. Changes in the Coulomb failure stress (ΔCFF), shear stress ($\Delta\tau$), and normal stress ($\Delta\sigma_n$) (negative values indicate compression) are shown by black, red and blue lines, respectively. The wide grey bar indicates the origin time of each triggered VLF earthquake.

teleseismic event is sufficiently different from that of the triggered event.

Although we used fixed source parameters in this work, source parameters could also be input to θ in our scheme and thereby used to elucidate the source mechanism of VLF earthquakes in detail. The optimization of numerical calculations to reduce the algorithm's computation time will be the subject of future work.

ACKNOWLEDGEMENTS

This work was supported by JSPS KAKENHI Grant Numbers JP16K05537 and JP16H06473 in Scientific Research on Innovative Areas "Science of Slow Earthquakes". Seismic waveform data are from KiK-net and Hi-net stations operated by the National Research Institute for Earth Science and Disaster Resilience (NIED), Japan. The seismicity catalogue of the Japan Meteorological Agency (JMA) was used in this work, along with the Global CMT catalogue (Dziewonski *et al.* 1981; Ekström *et al.* 2012). Figures were generated using Generic Mapping Tools (Wessel and Smith 1998). For data analysis, we used the computer systems of the Earthquake and Volcano Information Center of the Earthquake Research Institute (ERI), University of Tokyo, Japan. MM thanks two anonymous reviewers for careful and thoughtful reviews.

REFERENCES

Andrieu, C., Doucet, A. & Holenstein, R., 2010. Particle markov chain monte carlo methods, *J. R. Stat. Soc. B,* **72,** 269–342.

Araki, E., Saffer, D.M., Kopf, A.J., Wallace, L.M., Kimura, T., Machida, Y.,Ide & S. Davis, E., IODP Expedition 365 shipboard scientists, 2017. Recurring and triggered slow-slip events near the trench at the Nankai Trough subduction megathrust, *Science,* **356,** 1157–1160, doi:10.1126/siene.aan3120.

Baba, S., Takeo, A., Obara, K., Kato, A., Maeda, T. & Matsuzawa, T., 2018. Temporal activity modulation of deep very low frequency earthquakes in Shikoku, southwest Japan, *Geophys. Res. Lett.,* **45,** 733–738.

Carvalho, C.M., Johannes, M.S., Lopes, H.F. & Polson, N.G., 2010. Particle learning and smoothing, *Stat. Sci.,* **25,** 88–106.

Chao, K. & Obara, K., 2016. Triggered tectonic tremor in various types of fault systems of Japan following the 2012 Mw8.6 Sumatra earthquake, *J. geophys. Res.,* **121,** 170–187.

Clapp, T.C. & Godsill, S.J., 1999. Fixed-lag blind equalization and sequence estimation in digital communications systems using sequential importance sampling, *Proc. ICASSP99,* **5,** 2495–2498, doi:10.1109/ICASSP.1999.760637.

DeMets, C., Gordon, R.G. & Argus, D.F., 2010. Geologically current plate motions, *Geophys. J Int.,* **181,** 1–80, doi: 10.1111/j.1365-246X.2009.04491.x.

Dziewonski, A.M., Chou, T.-A. & Woodhouse, J.H., 1981. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. geophys. Res.,* **86,** 2825–2852.

Ekström, G., Nettles, M. & Dziewonski, A.M., 2012. The global CMT project 2004-2010: Centroid-moment tensors for 13,017 earthquakes, *Phys. Earth Planet. Inter.,* **200-201,** 1–9, doi:10.1016/j.pepi.2012.04.002.

Fry, B., Chao, K., Bannister, S., Peng, Z. & Wallace, L., 2011. Deep tremor in New Zealand triggered by the 2010 Mw8.8 Chile earthquake, *Geophys. Res. Lett.,* **38,** L15306, doi:10.1029/2011GL048319.

Gilks, W.R. & Berzuini, C., 2001. Following a moving target–Monte Carlo inference for dynamic Bayesian models, *J. R. Stat. Soc. B,* **63,** 127–146.

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

Itaba, S. & Ando, R., 2011. A slow slip event triggered by teleseismic surface waves, *Geophys. Res. Lett.,* **38,** L21306, doi:10.1029/2011GL049593.

Ito, Y., Obara, K., Matsuzawa, T. & Maeda, T., 2009. Very low frequency earthquakes related to small asperities on the plate boundary interface at the locked to aseismic transition, *J. geophys. Res.,* **114,** B00A13, doi:10.1029/2008JB006036.

Ito, Y., Obara, K., Shiomi, K., Sekine, S. & Hirose, H., 2007. Slow earthquakes coincident with episodic tremors and slow slip events, *Science,* **315,** 503–506, doi:10.1126/science.1134454.

Kaneko, L., , Ide, S. & Nakano, M., 2018. Slow earthquakes in the microseism frequency band (0.1-1.0 Hz) off Kii Peninsula, Japan, *Geophys. Res. Lett.,* **45,** 2618–2624.

Kitagawa, G., 1996. Monte Carlo filter and smoother for non-Gaussian nonlinear state space model, *J. Comput. Graph. Stat.,* **5,** 1–25, doi:10.2307/1390750.

Kitagawa, G., 1998. A self-organizing state-space model, *J. Am. Stat. Assoc.,* **93,** 1203–1215.

Kitagawa, G. & Sato, S., 2001. Monte Carlo smoothing and self-organising state-space model, in *Sequential Monte Carlo Methods in Practice,* pp. 177–195, eds Doucet, A., de Freitas, N. & Gordon, N., Springer.

Kitagawa, G., Takanami, T., Kuwano, A., Murai, Y. & Shimamura, H., 2002. Extraction of signal from high dimensional time series: analysis of ocean bottom seismograph data, in *Progress in Discovery Science, Lecture Notes in Artificial Intelligence,* Vol. **2281,** pp. 449–459, eds Arikawa, S. & Shinohara, A., Springer.

Koketsu, K., Miyake, H. & Suzuki, H., 2012. Japan integrated velocity structure model version 1, *Proc. 15WCEE,* 1773.

Liu, J. & West, M., 2001. Combined parameter and state estimation in simulation-based filtering, in *Sequential Monte Carlo Methods in Practice,* pp. 197–223, eds Doucet, A., de Freitas, N. & Gordon, N., Springer.

Maeda, T., Takemura, S. & Furumura, T., 2017. OpenSWPC: an open-source integrated parallel simulation code for modeling seismic wave propagation in 3D heterogeneous viscoelastic media, *Earth Planets Space,* **69,** 102, .

Miyazawa, M., 2011. Propagation of an earthquake triggering front from the 2011 Tohoku-Oki earthquake, *Geophys. Res. Lett.,* **38,** L23307, doi:10.1029/2011GL049795.

Miyazawa, M., 2012. Detection of seismic events triggered by P-waves from the 2011 Tohoku-Oki earthquake, *Earth Planets Space,* **64,** 1223–1229.

Miyazawa, M., 2015. Seismic fatigue failure may have triggered the 2014 Mw7.9 Rat Islands earthquake, *Geophys. Res. Lett.,* **42,** 2196–2203.

Miyazawa, M. & Brodsky, E.E., 2008. Deep low-frequency tremor that correlates with passing surface waves, *J. geophys. Res.,* **113,** B01307, doi:10.1029/2006JB004890.

Miyazawa, M., Brodsky, E.E. & Mori, J., 2008. Learning from dynamic triggering of low-frequency tremor in subduction zones, *Earth Planets Space,* **60,** e17–e20, doi:10.1186/BF03352858.

Miyazawa, M. & Mori, J., 2005. Detection of triggered deep low-frequency events from the 2003 Tokachi-oki earthquake, *Geophys. Res. Lett.,* **32,** L10307, doi:10.1029/2005GL022539.

Miyazawa, M. & Mori, J., 2006. Evidence suggesting fluid flow beneath Japan due to periodic seismic triggering from the 2004 Sumatra-Andaman earthquake, *Geophys. Res. Lett.,* **33,** L05303, doi:10.1029/2005GL025087.

Nakano, M. *et al.*, 2018. The 2016 Mw 5.9 earthquake off the southeastern coast of Mie Prefecture as an indicator of preparatory processes of the next Nankai Trough megathrust earthquake, *Prog. Earth Planet. Sci.,* **5,** 30.

Nakata, N. & Snieder, R., 2011. Near-surface weakening in Japan after the 2011 Tohoku-Oki earthquake, *Geophys. Res. Lett.,* **38,** L17302, doi:10.1029/2011GL048800.

Nakata, N. & Snieder, R., 2012. Estimating near-surface shear wave velocities in Japan by applying seismic interferometry to KiK-net data, *J. geophys. Res.,* **117,** B01308, doi:10.1029/2011JB008595.

National Research Institute for Earth Science and Disaster Prevention, 2016. Activity of Shallow Very-low-frequency Earthquakes in and around Japan (May - October, 2015), *Report of the Coordinating Committee for Earthquake Prediction, Japan,* **95,** 5–8.

Obara, K. & Kato, A., 2016. Connecting slow earthquakes to huge earthquakes, *Science,* **353,** 253–257, doi:10.1126/science.aaf1512.

Peng, Z. & Gomberg, J., 2010. An integrated perspective of the continuum between earthquakes and slow-slip phenomena, *Nature Geosci.,* **3,** 599–607, doi:10.1038/ngeo940.

Rubinstein, J.L., Vidale, J.E., Gomberg, J., Bodin, P., Creager, K.C. & Malone, S.D., 2007. Non-volcanic tremor driven by large transient shear stresses, *Nature,* **448,** 579–582.

Shiomi, K., Obara, K., Aoi, S. & Kasahara, K., 2003. Estimation of the azimuth of the Hi-net and KiK-net borehole seismometers, *Zisin,* **56,** 99–110 (in Japanese with English abstract), doi:10.4294/zisin1948.56.1_99.

Suda, N. & Matsuzawa, T., 2017. Long-term activity of deep very low-frequency earthquake in the Bungo Channel and Shikoku region, Joint WS on Slow Eqs. 2017, A2–10.

Tsuji, T., Minato, S., Kamei, R., Tsuru, T. & Kimura, G., 2017. 3D geometry of a plate boundary fault related to the 2016 Off-Mie earthquake in the Nankai subduction zone, Japan, *Earth Planet. Sci. Lett.,* **478,** 234–244.

Wallace, L.M. *et al.*, 2016. Near-field observations of an offshore Mw 6.0 earthquake from an integrated seafloor and subseafloor monitoring network at the Nankai Trough, southwest Japan, *J. geophys. Res.,* **121,** 8338–8351.

Wessel, P & Smith, W.H.F., 1998. New, improved version of generic mapping tools released, *EOS, Trans. Am. Geophys. Un.,* **579**.

Wikle, C.K. & Berliner, L.M., 2007. A Bayesian tutorial for data assimilation, *Phys. D: Nonlin. Phenom.,* **230,** 1–16, doi:10.1016/j.physd.2006.09.017.

Yano, K., 2008. A self-organizing state space model and simplex initial distribution search, *Comp. Stat.,* **23,** 197–216.

Zhang, W. & Shen, Y., 2010. Unsplit complex frequency-shifted PML implementation using auxiliary differential equations for seismic wave modeling, *Geophysis,* **75,** T141–T154, doi:10.1190/1.3463431.

Zigone, D. *et al.*, 2012. Triggering of tremors and slow slip event in Guerrero, Mexico, by the 2010 Mw 8.8 Maule, Chile, earthquake, *J. geophys. Res.,* **117,** B09304, doi:10.1029/2012JB009160.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

llh_movie.mp4

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

APPENDIX: DYNAMIC STRESS AND STRAIN CHANGES

The stress changes associated with passing surface waves can be obtained from surface observations. For simplicity, we assume fundamental mode Rayleigh and Love waves propagating in an isotropic elastic medium. The elements of the transfer kernels are not functions of time and the strain changes as a function of time are obtained by the linear product of the transfer kernel and the surface observations. Since the kernel is a function of the angular frequency of

the surface waves, the stress and strain changes can be easily computed in the frequency domain; see Miyazawa & Brodsky (2008) for derivations in detail.

For Rayleigh wave propagation, we assume a half-space. Below, k is the wavenumber, ω is the angular frequency, r is the radial distance, $d(=-z)$ is the depth, α is P-wave velocity, β is S-wave velocity and $c_R(=\omega/k)$ is the phase velocity of the Rayleigh wave. Dynamic stress changes due to passing Rayleigh waves are given by

$$e_{rr} = K_{rr} u_z^{\rm obs} \tag{A1}$$

$$e_{zz} = K_{zz} u_z^{\rm obs} \tag{A2}$$

$$e_{rz} = e_{zr} = K_{rz} u_r^{\rm obs}, \tag{A3}$$

where $u_z^{\rm obs}$ and $u_r^{\rm obs}$ are the observed vertical and radial particle motions at the surface, respectively. The transfer kernels are

$$K_{rr} = -k\left\{\exp(-\omega\hat{\eta_\alpha}d) + \left(\frac{c_R^2}{2\beta^2}-1\right)\exp(-\omega\hat{\eta_\beta}d)\right\} \Big/ \left\{c_R\hat{\eta_\alpha} + \frac{1}{c_R\hat{\eta_\beta}}\left(\frac{c_R^2}{2\beta^2}-1\right)\right\} \tag{A4}$$

$$K_{zz} = \left\{c_R\omega\hat{\eta_\alpha}^2\exp(-\omega\hat{\eta_\alpha}d) + \frac{\omega}{c_R}\left(\frac{c_R^2}{2\beta^2}-1\right)\exp(-\omega\hat{\eta_\beta}d)\right\} \Big/ \left\{c_R\hat{\eta_\alpha} + \frac{1}{c_R\hat{\eta_\beta}}\left(\frac{c_R^2}{2\beta^2}-1\right)\right\} \tag{A5}$$

$$K_{rz} = \frac{2\beta^2}{c_R^2}\omega\hat{\eta_\alpha}\exp(-\omega\hat{\eta_\alpha}d) + \frac{\omega}{2}\left(1-\frac{2\beta^2}{c_R^2}\right)\left(\hat{\eta_\beta}+\frac{1}{c_R^2\hat{\eta_\beta}}\right)\exp(-\omega\hat{\eta_\beta}d), \tag{A6}$$

where

$$\hat{\eta_\alpha} = \sqrt{\frac{1}{c_R^2}-\frac{1}{\alpha^2}} \tag{A7}$$

$$\hat{\eta_\beta} = \sqrt{\frac{1}{c_R^2}-\frac{1}{\beta^2}}. \tag{A8}$$

For a Love wave, we assume a low-velocity layer that overlies a half-space whose top layer has thickness H. Here, β_1 and μ_1 are the S-wave velocity and shear modulus of the top layer, and β_2 and μ_2 are those of the half-space; $c_L(=\omega/k)$ is the phase velocity of the Love wave and $\beta_1 < c_L < \beta_2$. The dynamic strain changes due to the passing Love wave are given by

$$e_{tt} = 0 \tag{A9}$$

$$e_{tz} = e_{zt} = K_{tz} u_t^{\rm obs} \tag{A10}$$

$$e_{tr} = e_{rt} = K_{tr} v_t^{\rm obs}, \tag{A11}$$

where $u_t^{\rm obs}$ and $v_t^{\rm obs}$ are the observed transverse particle motion and its velocity motion at the surface. The transfer kernels are

$$K_{tz} = \frac{\omega\eta_{\beta_1}}{2}\sin(\omega\eta_{\beta_1}d) \tag{A12}$$

$$K_{tr} = -\frac{1}{2c_L}\cos(\omega\eta_{\beta_1}d) \tag{A13}$$

for $0 \le d < H$ and

$$K_{tz} = \frac{\omega\hat{\eta}_{\beta_2}}{2}\cos(\omega\eta_{\beta_1}H)\exp(-\omega\hat{\eta}_{\beta_2}(d-H)) \tag{A14}$$

$$K_{tr} = -\frac{1}{2c_L}\cos(\omega\eta_{\beta_1}H)\exp(-\omega\hat{\eta}_{\beta_2}(d-H)) \tag{A15}$$

for $d \ge H$, where

$$\eta_{\beta_1} = \sqrt{\frac{1}{\beta_1^2}-\frac{1}{c_L^2}} \tag{A16}$$

$$\hat{\eta}_{\beta_2} = \sqrt{\frac{1}{c_L^2}-\frac{1}{\beta_2^2}}. \tag{A17}$$

Since the fundamental mode is assumed,

$$\omega\eta_{\beta_1} = \frac{1}{H}\tan^{-1}\left(\frac{\mu_2\hat{\eta}_{\beta_2}}{\mu_1\eta_{\beta_1}}\right), \quad 0 \le \omega\eta_{\beta_1} < \frac{\pi}{2H} \tag{A18}$$

which leads c_L, where

$$\beta_1 \le c_L < \left(\frac{1}{\beta_1^2}-\left(\frac{\pi}{2H\omega}\right)^2\right)^{-1/2}. \tag{A19}$$

From eqs (A1)–(A3),(A9)–(A11), and the stress–strain relationship, the temporal changes in the stress tensor at depth can be obtained. Finally, we obtain the change in the Coulomb failure stress as a function of time from

$$\Delta{\rm CFF}(t) = \Delta\tau + \mu'\Delta\sigma_n, \tag{A20}$$

where μ' is the apparent friction coefficient and $\Delta\tau$ and $\Delta\sigma_n$ are the shear stress change resolved in the slip direction and the normal stress change on the fault plane, respectively.

In this study, we assume $\mu = \lambda = 35$ GPa, $\beta = 3.9\,{\rm km\,s^{-1}}$, and $\alpha = \sqrt{3}\beta$ for a half-space, and $\mu_1 = 30$ GPa, $\mu_2 = \lambda_2 = 66$ GPa, $H = 32$ km, $\beta_1 = 3.2\,{\rm km\,s^{-1}}$ and $\beta_2 = 4.5\,{\rm km\,s^{-1}}$ for one layer over a half-space.