Earth, Planets and Space

FULL PAPER

Open Access

Real-time automatic uncertainty estimation of coseismic single rectangular fault model using GNSS data

Keitaro Ohno[1,4], Yusaku Ohta[1,2,3*], Satoshi Kawamoto[4], Satoshi Abe[4], Ryota Hino[1,3], Shunichi Koshimura[3], Akihiro Musa[5,6] and Hiroaki Kobayashi[7]

Abstract

Rapid estimation of the coseismic fault model for medium-to-large-sized earthquakes is key for disaster response. To estimate the coseismic fault model for large earthquakes, the Geospatial Information Authority of Japan and Tohoku University have jointly developed a real-time GEONET analysis system for rapid deformation monitoring (REGARD). REGARD can estimate the single rectangular fault model and slip distribution along the assumed plate interface. The single rectangular fault model is useful as a first-order approximation of a medium-to-large earthquake. However, in its estimation, it is difficult to obtain accurate results for model parameters due to the strong effect of initial values. To solve this problem, this study proposes a new method to estimate the coseismic fault model and model uncertainties in real time based on the Bayesian inversion approach using the Markov Chain Monte Carlo (MCMC) method. The MCMC approach is computationally expensive and hyperparameters should be defined in advance via trial and error. The sampling efficiency was improved using a parallel tempering method, and an automatic definition method for hyperparameters was developed for real-time use. The calculation time was within 30 s for 1×10^6 samples using a typical single LINUX server, which can implement real-time analysis, similar to REGARD. The reliability of the developed method was evaluated using data from recent earthquakes (2016 Kumamoto and 2019 Yamagata-Oki earthquakes). Simulations of the earthquakes in the Sea of Japan were also conducted exhaustively. The results showed an advantage over the maximum likelihood approach with a priori information, which has initial value dependence in nonlinear problems. In terms of application to data with a small signal-to-noise ratio, the results suggest the possibility of using several conjugate fault models. There is a tradeoff between the fault area and slip amount, especially for offshore earthquakes, which means that quantification of the uncertainty enables us to evaluate the reliability of the fault model estimation results in real time.

Keywords: Global Navigation Satellite System (GNSS), Real-time GNSS, Bayesian inversion, Uncertainties estimation, MCMC, REGARD

Introduction

Rapid understanding of the magnitude of large earthquakes and their associated fault area is crucial for disaster response. Numerous studies have shown the advantages of the high-rate (typically 1 Hz sampling) Global Navigation Satellite System (GNSS) as a broadband sensor that can directly measure displacement without saturation (Larson et al. 2003; Ohta et al. 2006; Larson 2009). Using onshore high-rate GNSS data

*Correspondence: yusaku.ohta.d2@tohoku.ac.jp
[1] Research Center for Prediction of Earthquakes and Volcanic Eruptions, Graduate School of Science, Tohoku University, 6-6 Aza-Aoba, Aramaki, Aoba-ku, Sendai 980-8578, Japan
Full list of author information is available at the end of the article

Ohno *et al. Earth, Planets and Space*   (2021) 73:127

Page 2 of 18

enables the rapid estimation of earthquake magnitude and tsunamigenic potential within a short period (Blewitt et al. 2006, 2009). The 2011 Tohoku-Oki earthquake (M_w 9.0) and its associated tsunami reaffirm the importance of accurate tsunami early warning systems (Ozaki 2011). Numerous studies conducted to rapidly determine the size of the earthquake based on real-time GNSS data followed this event (Crowell et al. 2012; Ohta et al. 2012, 2015; Melgar et al. 2012; Ohta 2016; Tanaka et al. 2019). Melgar et al. (2012) proposed a new method called fastCMT (fast Centroid Moment Tensor) to determine the moment tensor and centroid location of earthquakes using real-time GNSS data. Crowell et al. (2012) investigated two different methods for inverting heterogeneous coseismic slip distribution using real-time GNSS time series. One method used pre-defined fault planes, such as the plate interface, while the other used the fastCMT solution proposed by Melgar et al. (2012), which estimates the CMT solution using GNSS data. Ohta et al. (2012, 2015) proposed an algorithm to detect and estimate permanent displacement due to the coseismic slip from real-time kinematic GNSS (RTK-GNSS) time series. They applied the developed algorithm to the 2011 Tohoku-Oki earthquake, and pointed out that the estimated moment release reached M_w 8.7 within five minutes after the origin time, which is close to the actual moment magnitude (M_w 9.0). Numerous other efforts have also shown the advantage of real-time GNSS data for the estimation of large events, including the studies of Hoechner et al. (2013), Colombelli et al. (2013), Melgar and Bock (2015), and Murray et al. (2019).

The Geospatial Information Authority of Japan (GSI) operates a continuous nationwide GNSS network, the GNSS Earth Observation Network System (GEONET), across Japan. Based on the lessons learned from the 2011 Tohoku-Oki earthquake, the GSI and Graduate School of Science, Tohoku University, have jointly developed a new crustal deformation analysis system using GEONET data for quasi real-time earthquake size estimation, known as the GEONET analysis system for rapid deformation monitoring (REGARD) (Kawamoto et al. 2016, 2017). This system consists of three major subsystems which perform the following functions: (1) real-time estimation of the displacement time series using real-time GNSS data, (2) automated extraction of displacement fields caused by a large event, and (3) automated coseismic fault model estimation via an approximated single rectangular fault model and coseismic slip distribution along the subducting plate interface.

The single rectangular fault model is one of the simplified models that can explain coseismic fault motion. It is difficult to explain the complex rupture processes, such as inhomogeneous slip distribution, for large earthquakes

(approximately > M_w 8.0). If only the GEONET sites (site spacing: 20–25 km) are used, most M_w 7 events can be explained by the single rectangular fault model as first-order approximation. This model has the disadvantage of using a single rectangular fault patch to represent complex large earthquake processes (Kawamoto et al. 2017) and modeling a non-interplate earthquake (~ M_w 7.5) whose geometry is difficult to assume in advance (e.g., 2016 Kumamoto and 2019 Yamagata-Oki earthquakes). Such events have frequently occurred in the past 20 years in Japan [e.g., 2004 Chuetsu earthquake (M_{jma} 6.6); 2007 Noto earthquake (M_{jma} 6.9); 2007 Chuetsu-Oki earthquake (M_{jma} 6.9); 2008 Iwate–Miyagi Nairiku earthquake (M_{jma} 7.2); 2016 Tottori-ken Chubu earthquake (M_{jma} 6.6)]. Thus, accurate and rapid estimation of the single rectangular fault model is highly important.

The estimation of a single rectangular fault model is a nonlinear problem, because all components should be estimated simultaneously. To solve this problem, the REGARD system adopts the maximum likelihood approach with a priori information (Matsu'ura and Hasegawa, 1987) based on tectonic background. For the 2016 Kumamoto earthquake (M_w 7.0), REGARD successfully estimated the single rectangular fault model automatically in real time (Kawamoto et al. 2016). However, a problem that required a solution appeared during the estimation process. To estimate the single rectangular fault model, the REGARD system should provide initial values of the unknown parameters. For the 2016 Kumamoto earthquake, REGARD assumed a strike-slip fault model as the initial model with conjugated fault parameters and required a long convergence time to obtain a reliable coseismic fault model (Kawamoto et al. 2016). Since the estimation strongly depends on the a priori information of the fault model, estimating the uncertainties of the coseismic fault model is difficult but necessary for the initial response of decision makers, such as governmental agencies.

Recently, numerous geodetic studies have adopted Bayesian approaches (Ito et al. 2012; Dettmer et al. 2014; Minson et al. 2014a, b; Jiang and Simons 2016; Ohno and Ohta 2018). These studies have also utilized those approaches for uncertainty estimation. The uncertainties of unknown parameters should contain both the observations and the modeling errors and be estimated without a priori constraints. To realize the Bayesian modeling, many previous studies adopted the Markov Chain Monte Carlo (MCMC) method, which produces a correlated sample for estimating expectations with respect to a target distribution. However, most previous studies could not use the MCMC in real time, because it is computationally expensive and requires hyperparameters to be set up in advance via trial and error. In

Ohno *et al. Earth, Planets and Space*   (2021) 73:127

Page 3 of 18

contrast, Minson et al. ([2014a, b](#)) investigated an inversion strategy capable of using real-time high-rate GNSS data to simultaneously solve for a distributed slip model and fault geometry in real time. By adopting an analytical Bayesian approach, they solved the complex inversion problem (including calculating the uncertainties) in real time. Their method first determines the fault geometry (strike and dip of fault plane) and then estimates the slip distribution. It efficiently estimates the fault geometry and fault slip simultaneously, although it mainly focuses on the slip distribution and its uncertainties on the "most likely" fault geometry. Since moderate-sized non-interplate earthquakes frequently occur in Japan and some of them occur along the blind fault (e.g., 2008 Iwate–Miyagi Nairiku earthquake; Ohta et al. [2008](#)), the geometry of the earthquake faults should not be assumed in advance. Furthermore, the uncertainties of the fault geometry and other unknown parameters, such as the rake angle and slip amount, will be traded-off between them. Thus, accurate estimation of a single rectangular fault model that includes uncertainties associated with fault model geometry is still required.

In this study, a new method was developed to rapidly predict the parameters of a coseismic single rectangular fault, along with the associated uncertainties, in real time using a full Bayesian approach. The algorithm was developed to replace the pre-existing fault model estimation function of REGARD. In the remainder of the paper, we first describe how to overcome the problems in the

Bayesian inversion

Let θ be a model parameter vector that contains fault parameters of a single rectangular fault model (Okada [1992](#)) and d be a permanent displacement data vector with three components (horizontal and vertical components) based on real-time GNSS observations and analysis. We define latitude and longitude as the center of the fault and depth as the top of the fault. Bayesian estimation of the unknown model parameters conditional to the observations is based on Bayes' theorem:

$$p(\theta|d) = \frac{p(d|\theta)p(\theta)}{p(d)}, \tag{1}$$

where $p(d)$ is the PDF of the observations, $p(\theta)$ is the prior PDF of the model parameters (i.e., conditional PDF), $p(d|\theta)$ is the likelihood function, and $p(\theta|d)$ is the posterior PDF of the model parameters. Note that $p(d)$ is constant, because observations are fixed values, and hence, the posterior PDF is proportional to the product of the prior PDF and the likelihood function. The likelihood function measures the degree of fit between observed d and calculated data $\widehat{d}(\theta)$. The residuals are given by $r(\theta) = \widehat{d}(\theta) - d$. Given that the number of stations is N, the dimension of d is $3N$. Assuming that the estimation error follows a normal distribution function and standard deviations $\sigma_i (i = East(E), North(N), UD(U))$ at each GNSS station are identical for each horizontal (EN) and vertical (U) component, the likelihood function is defined as follows:

$$p(d|\theta) = L_{EN}(\theta, \sigma_{EN})L_U(\theta, \sigma_U) = \frac{1}{\sqrt{(2\pi\sigma_{EN}^2)^{2N}}}\exp\left(-\frac{1}{2\sigma_{EN}^2}r_{EN}{}^{T}r_{EN}\right)\frac{1}{\sqrt{(2\pi\sigma_U^2)^{N}}}\exp\left(-\frac{1}{2\sigma_U^2}r_U{}^{T}r_U\right), \tag{2}$$

MCMC method for real-time estimations. Then, we present the application of the proposed method to past and simulated events. Finally, we discuss the utilization and significance of the obtained uncertainties.

Methods

We developed a new algorithm for the quantitative assessment of the estimated coseismic single rectangular fault model using a Bayesian approach in real time. Our algorithm used MCMC to obtain a probability density function (PDF) of fault parameters. To reduce the calculation time and accelerate the automatic adjustment of the hyperparameters for real-time prediction, we adopted a multi-stage approach with a parallel tempering technique. We refer to this new algorithm as the "Real-time automatic uncertainty estimation of the coseismic single rectangular fault model based on GNSS data" (hereafter, RUNE) and describe it in detail in this section.

where $L_i(\theta, \sigma_i)$ indicates the likelihood functions separated into horizontal and vertical components, and T indicates the transposition. The prior PDF $p(\theta)$ represents *priori* information about the model parameters, assuming a uniform distribution in each assumed search range. Bayes' theorem (Eq. ([1](#))) states that the prior PDF is updated to the posterior PDF by considering the observational data in the form of a likelihood function. By sampling the posterior PDF with the MCMC sampler, the fault parameter estimations can be obtained as PDFs.

The basic Metropolis–Hasting (M–H) method was adopted (Metropolis et al. [1953](#); Hastings [1970](#)) for sampling in this study. Specifically, sampling is performed by repeating the following procedures:

> Step 1: Generate a transition candidate, $\theta^{new}(=$ reference θ + perturbation $\Delta\theta$).

Ohno *et al. Earth, Planets and Space* (2021) 73:127

Page 4 of 18

Step 2: Calculate $\widehat{d}(\theta)$ and $\widehat{d}(\theta^{\text{new}})$ based on Okada's method ([1992](#)) and calculate an acceptance probability, α.

Step 3: Update the reference θ to θ^{new} with a probability α. If probability α is not satisfied, the old θ is used.

We assume that perturbation $\Delta\theta$ is a uniform random number with $\pm\Delta\theta^{\text{max}}/2$ as the maximum step size. In the M–H method, the acceptance probability α is defined as follows:

$$\alpha = min\left(1, \frac{p(d|\theta^{\text{new}})p(\theta^{\text{new}})}{p(d|\theta)p(\theta)}\right). \tag{3}$$

In addition to the fault parameters, the stress drop value and variance reduction (VR) defined by Eqs. ([4](#)) and ([5](#)) were also calculated as supporting information:

$$\Delta\sigma = 2c\mu D/\sqrt{A}, \tag{4}$$

$$\text{VR} = 100\left(1 - r^{\text{T}}r/d^{\text{T}}d\right), \tag{5}$$

where c, μ, D, and A denote geometrical factor, rigidity, slip amount, and fault area, respectively. Here, we assume that c and μ are 0.5 and 30 GPa, respectively. The VR corresponds to an index that shows the degree of coincidence between the observed and calculated values. Thus, a high VR value indicates a high likelihood.

Parallel tempering

Using just a single Markov chain entails a significantly long mixing time to obtain a reliable posterior PDF. To accelerate the sampling efficiency to achieve real-time predictions, we introduced a parallel tempering technique (Geyer [1991](#); Jasra et al. [2007](#)). Let $\left\{p_j(d|\theta)\right\}_{j=1,2,\ldots,n}$ be a family of PDFs defined as follows:

$$p_j(d|\theta) = p(d|\theta_j)^{1/T_j}, \tag{6}$$

where subscript j is the "chain number" and T_j $(j = 1, 2, \ldots, n)$ are the temperatures that satisfy $1 = T_1 < T_2 < \cdots < T_n$. In this study, we used eight parallel chains (n = 8). Note that T_1 corresponds to the target PDF such that the functional form of the PDF becomes smoother with an increasing chain number or temperature. The parallel tempering technique involves the following steps:

Step 1: Obtain the next sample θ_j in each chain using the M–H method described in "[Bayesian inversion](#)" Section,

Step 2: Randomly select four temperature values without overlap (T_j and $T_{j'}$; T_k and $T_{k'}$),

Step 3: Exchange the current samples between T_j and $T_{j'}$ with a probability, β. Repeat this step for T_k and $T_{k'}$. The exchange occurs with one pair per step. However, two pairs per step can be exchanged to promote exchange in real time. Exchange probability β between T_j and $T_{j'}$ is defined as follows:

$$\beta = min\left(1, \frac{p(d|\theta_{j'})^{1/T_j}p(d|\theta_j)^{1/T_{j'}}}{p(d|\theta_j)^{1/T_j}p(d|\theta_{j'})^{1/T_{j'}}}\right). \tag{7}$$

Each temperature was set using a geometric progression ($T_j = 100^{j-1/7}$) based on the idea that σ_i is multiplied by 10 at the maximum temperature.

Sampling flow for real-time use

To use the method described in "[Parallel tempering](#)" Section, we assume several hyperparameters in advance, i.e., the initial model sample (θ^{int}), standard deviations of the likelihood functions (σ_{EN}, σ_U; Eq. ([2](#))), maximum step size ($\Delta\theta^{\text{max}}$), and the number of Markov chains. The values of the hyperparameters strongly depend on the size of the earthquake, coverage of the GNSS stations, and the distance from the fault to a GNSS station. Thus, automatic adjustment of the hyperparameters is essential for real-time predictions. For this purpose, an "automatic hyperparameter setting phase" was introduced before the main sampling phase to automatically determine the values of hyperparameters. Figure [1](#) shows an overview of the sampling flow including the automatic hyperparameter setting phase. The following are the details of each phase and its preparations.

Initial parameter settings

In Japan, an earthquake early warning (EEW) can be issued by the Japan Meteorological Agency (JMA) based on the short-period seismometer data that cover the entire Japanese territory. Thus, the EEW was adopted as a priori information for the expected coseismic fault model as the initial value of the MCMC approach. Table [1](#) lists the settings of the initial values and search ranges for the fault parameters. For the fault area and slip amount assumption in the expected coseismic fault model, we used the scaling law proposed by Utsu ([2001](#)). As a priori information for the location of the fault, we set the prior PDFs for the latitude, longitude, and depth as the normal distribution based on the EEW information with pre-determined variances that depend on the fault length and width. There was no information on the fault mechanism for the timing of the EEW. Thus, possible fault
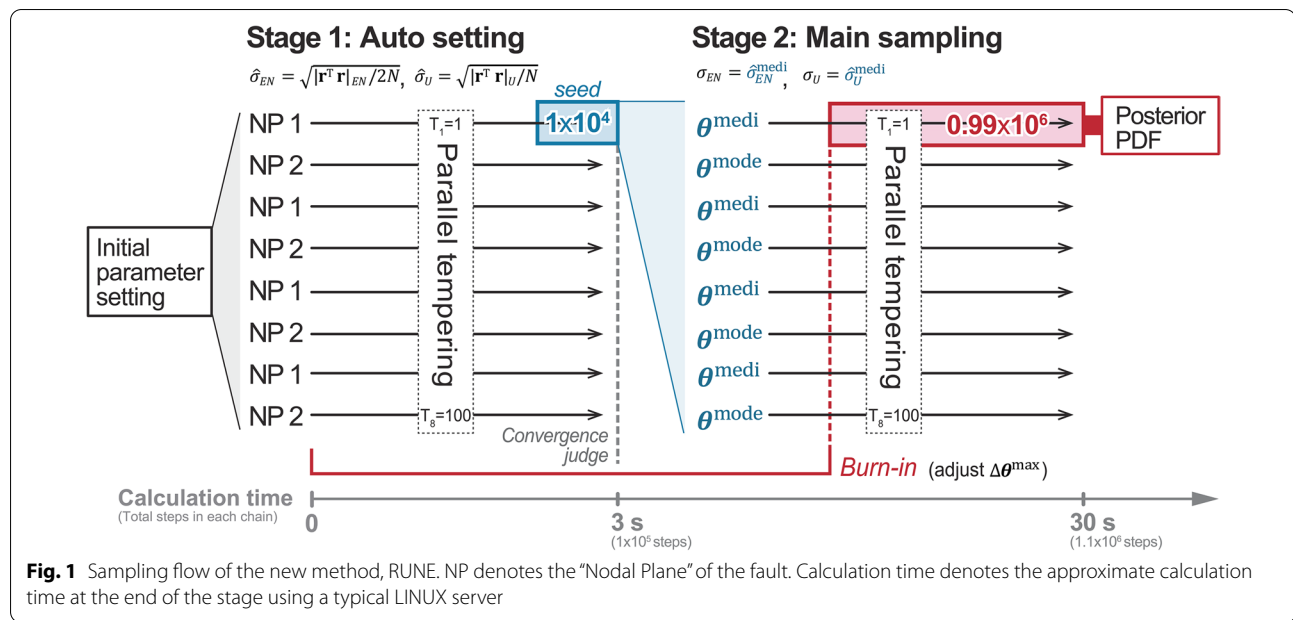
Ohno *et al. Earth, Planets and Space*   (2021) 73:127

Page 5 of 18

Fig. 1 Sampling flow of the new method, RUNE. NP denotes the "Nodal Plane" of the fault. Calculation time denotes the approximate calculation time at the end of the stage using a typical LINUX server

Table 1 Initial parameter settings for the analysis. Maximum step sizes are adjusted from these values in the Burn-in phase

	Initial value θ^{int}	Maximum step size $\Delta\theta^{max}$	Prior PDF
Latitude [°]	LAT	$\sqrt{LEN_M * WID_M}*0.1$	$N(LAT, \sqrt{LEN_{M-1} * WID_{M-1}}^2)$
Longitude [°]	LON	$\sqrt{LEN_M * WID_M} * 0.1$	$N(LON, \sqrt{LEN_{M-1} * WID_{M-1}}^2)$
Depth [km]	DEP	1	$N(DEP, 20^2)$
Strike [°]	NP1, NP2	10	$U(0, 360)$
Dip [°]	NP1, NP2	10	$U(0, 90)$
Rake [°]	NP1, NP2	10	$U(-180, 180)$
Length [km]	LEN_M	$LEN_M*0.1$	$U(0, \infty)$
Width [km]	WID_M	$WID_M*0.1$	$U(0, \infty)$
Slip amount [m]	$SLIP_M$	$SLIP_M*0.1$	$U(0, \infty)$

LAT, LON, DEP, M: values estimated by EEW

LEN_M, WID_M, $SLIP_M$: values based on scaling law by Utsu (2001). The units are converted for each line

NP1, NP2: Nodal planes based on Terakawa and Matsu'ura (2009)

geometries such as the fault strike, dip, and rake angles could be assumed. This information was derived and simplified from a tectonic stress pattern database at 10 and 40 km in depth, as reported in Terakawa and Matsu'ura (2009), whose data are identical to those in the current REGARD system (Kawamoto et al. 2017). Furthermore, for a more efficient unknown parameter search, only the results in the range of 0.2–21.2 MPa were accepted for the stress drop based on the empirical rule (Kanamori and Anderson 1975). These lower and upper limits were determined based on the stress drop value calculated with Eq. (4) using $[LEN_{M+1}, WID_{M+1}, SLIP_{M-1}]$ and $[LEN_{M-1}, WID_{M-1}, SLIP_{M+1}]$ for the lower and upper limits, respectively. Here, LEN_M, WID_M, and $SLIP_M$

denote values based on the scaling law proposed by Utsu (2001). The scaling law can translate from M to the fault length (LEN_M), width (WID_M), and slip amount ($SLIP_M$) of single rectangular faults. Here, LEN_{M+1} and LEN_{M-1} denote that the fault length expected from the scaling law is assumed to be the magnitude estimated by EEW plus one and minus one, respectively. Therefore, $[LEN_{M+1}, WID_{M+1}, SLIP_{M-1}]$ and $[LEN_{M-1}, WID_{M-1}, SLIP_{M+1}]$ correspond to the smallest and largest stress drop values, respectively. The lower and upper limits of the stress drop do not depend on M, because the scaling law by Utsu (2001) adopts a uniform stress drop value. For the same reason, with the constraint on the stress drop value, only the samples that

Ohno *et al. Earth, Planets and Space* (2021) 73:127

Page 6 of 18

satisfy "fault length > fault width" were accepted based on the relationship between fault length and width (e.g., Wells and Coppersmith 1994).

Stage 1: Auto setting phase

In the previous stage, we assumed the appropriate initial value of the coseismic fault model using the EEW issued by the JMA as well as empirical information. In the subsequent stage, we attempted to automatically determine the appropriate values of σ_{EN}, σ_U, and $\Delta\boldsymbol{\theta}^{\max}$. Based on the assumption in Eq. (2), the likelihood function can be simplified and used to derive the maximum likelihood estimates, i.e., $\widehat{\sigma}_i (i = EN, U)$, by setting $\partial L_i / \partial \sigma_i = 0$ and solving for σ_i. After substituting $\widehat{\sigma}_{EN} = \sqrt{\boldsymbol{r}_{EN}^{\mathrm{T}} \boldsymbol{r}_{EN} / 2N}$ and $\widehat{\sigma}_U = \sqrt{\boldsymbol{r}_U^{\mathrm{T}} \boldsymbol{r}_U / N}$ into Eq. (2), the likelihood function can be described as follows (Dosso and Wilmut 2006; Dettmer et al. 2014):

$$p(\boldsymbol{d}|\boldsymbol{\theta}) = L_{EN}(\boldsymbol{\theta})L_U(\boldsymbol{\theta}) \propto \exp\left(-\tfrac{2N}{2}\log_e\left(\boldsymbol{r}_{EN}^{\mathrm{T}}\boldsymbol{r}_{EN}\right)\right)\exp\left(-\tfrac{N}{2}\log_e\left(\boldsymbol{r}_U^{\mathrm{T}}\boldsymbol{r}_U\right)\right). \tag{8}$$

In the auto setting phase, Eq. (8) is used as a likelihood function for sampling, while $\widehat{\sigma}_i$ is calculated. This procedure is equivalent to assuming that the estimation error is the residual mean square, allowing us to determine an implicit value of σ_i in real time. To ensure that the setting does not only depend on the EEW and temperature of each chain, the $\Delta\boldsymbol{\theta}^{\max}$ values of all eight chains were automatically adjusted such that the acceptance rate is between 30 and 40% (Roberts and Rosenthal 2001). If the acceptance rate is smaller than 30%, $\Delta\boldsymbol{\theta}^{\max}$ performs constant multiplication at 10% decrease. Otherwise, if the acceptance rate exceeds 45%, $\Delta\boldsymbol{\theta}^{\max}$ performs constant multiplication at 5% increase. This auto-adjustment procedure occurs at every 1000 steps. Note that the ratio of $\Delta\boldsymbol{\theta}^{\max}$ between each fault parameter depends on the initial decision, because $\Delta\boldsymbol{\theta}^{\max}$ is adjusted by multiplying the whole unknown parameters. To evaluate the end of the auto setting phase, the following regulations were introduced: when the median value of the VR samples at every 1×10^4 steps (i.e., "1 batch" in this study) is larger than 90%, or if 10 batches are sampled, the auto setting phase is terminated to move on to the next stage. We refer to the last batch of stage 1 in T_1 as the "seed batch" for the next stage (Fig. 1).

Stage 2: Main sampling phase

In stage 2, stable sampling of the posterior PDF was performed based on the information from stage 1. To obtain a stable sampling result, $\sigma_i (i = EN, U)$ was fixed to the sample median value of the seed batch, $\widehat{\sigma}_i^{\mathrm{medi}}$. The median and mode values for $\boldsymbol{\theta}$ ($\boldsymbol{\theta}^{\mathrm{medi}}$ and $\boldsymbol{\theta}^{\mathrm{mode}}$) of the

"seed batch" were assigned to eight chains as the initial fault parameters in stage 2 (Fig. 1). For $\Delta\boldsymbol{\theta}^{\max}$, the optimization was continued based on the results of stage 1 until the end of the first batch in stage 2. After the first batch in stage 2, the $\Delta\boldsymbol{\theta}^{\max}$ value was fixed for the following batches. In general, because the result of the initial part of the chain strongly depends on the initial value of the unknown parameters, the chain eliminates these values (i.e., "Burn-in"). In our algorithm, we assumed that Burn-in occurs in stage 1 until the end of the initial batch in stage 2. Sampling was terminated when 100 batches (1×10^6 steps) were obtained. Thus, 99 batches obtained in this stage were used for rendering the posterior PDFs (Fig. 1).

Computation time

Typically, MCMCs consume a large amount of computational resources to increase the number of samples; thus, they are not suitable for real-time analysis. In this study, all the programs were coded in Fortran90, the memory access speed was increased using single precision for unnecessary components, and the code was parallelized using OpenMP (open multiprocessing) with multiple CPUs. For parallel computing, eight parallel chains were assigned to each thread and calculated with eight threads. The non-interfering loop calculations were also parallelized. The computation time depended on the computer performance, number of threads, amount of parallel tempering, number of MCMC samples, and the number of observation sites.

The computation time required for RUNE is 30 s when a Xeon Gold 6126 (2.6 GHz) processor is used with eight fixed parallel chains and 1.1×10^6 samples (i.e., the case for a maximum number of samples in stage 1 and 2). As the number of observation points used in this study was 100 or 200 when using 8 threads, the calculation was completed in less than 30 s: 15 s for 100 stations and 18 s for 200 stations (Additional file 1: Figure S1).

Results

To evaluate the performance of the proposed algorithm, we applied it to actual RTK-GNSS displacement data, which recorded several past non-interplate large earthquakes, e.g., the 2016 Kumamoto (M_w 7.0) and 2019 Yamagata-Oki earthquakes (M_w 6.4). For these events, we used the actual RTK-GNSS displacement data estimated by the REGARD system. We also conducted 60

numerical simulations of the earthquakes of M_w 6.7–7.8, which occurred in the Sea of Japan.

2016 Kumamoto earthquake (M_w 7.0)

The 2016 Kumamoto earthquake (M_w 7.0) occurred at 16:25:05 UTC on April 15, 2016. REGARD was operated in real time during the earthquake to automatically estimate the single rectangular fault model. The estimated fault model was consistent with the post-earthquake results based on the Interferometric synthetic-aperture radar (InSAR) and GNSS data. Kawamoto et al. (2016), however, identified a problem that required a solution for the inversion component based on the occurrence of this event. In the current REGARD system, the maximum likelihood estimation was adopted to estimate the single rectangular fault model. To estimate the model, including the fault geometry, the REGARD system assumes the initial values for unknown parameters for the inversion analysis, because it is a nonlinear problem. Thus, the obtained result depends on these initial values. For the 2016 Kumamoto earthquake, REGARD assumed a nodal plane opposite to the actual direction based on the focal mechanism database. Due to this erroneous assumption, a long convergence time was required for this event. By applying RUNE to this large inland event, we evaluated its magnitude and characteristics, but determining the fault location and geometry was difficult. We used 200 stations near the epicenter for the analysis. We estimated the displacement at each observation point at 5 s intervals from actual 1 Hz RTK-GNSS time-series data (difference between the average of 1 min before the event and the average of the 20 s travel time window), and applied RUNE to the data independently. This procedure of repeated displacement estimation is similar to that of REGARD, and it is important for dealing with large earthquakes whose displacement evolves over time. On the other hand, the application presented in this paper is for relatively small events whose displacements do not evolve over time; however, we also used the repeat estimation for comparison with REGARD.
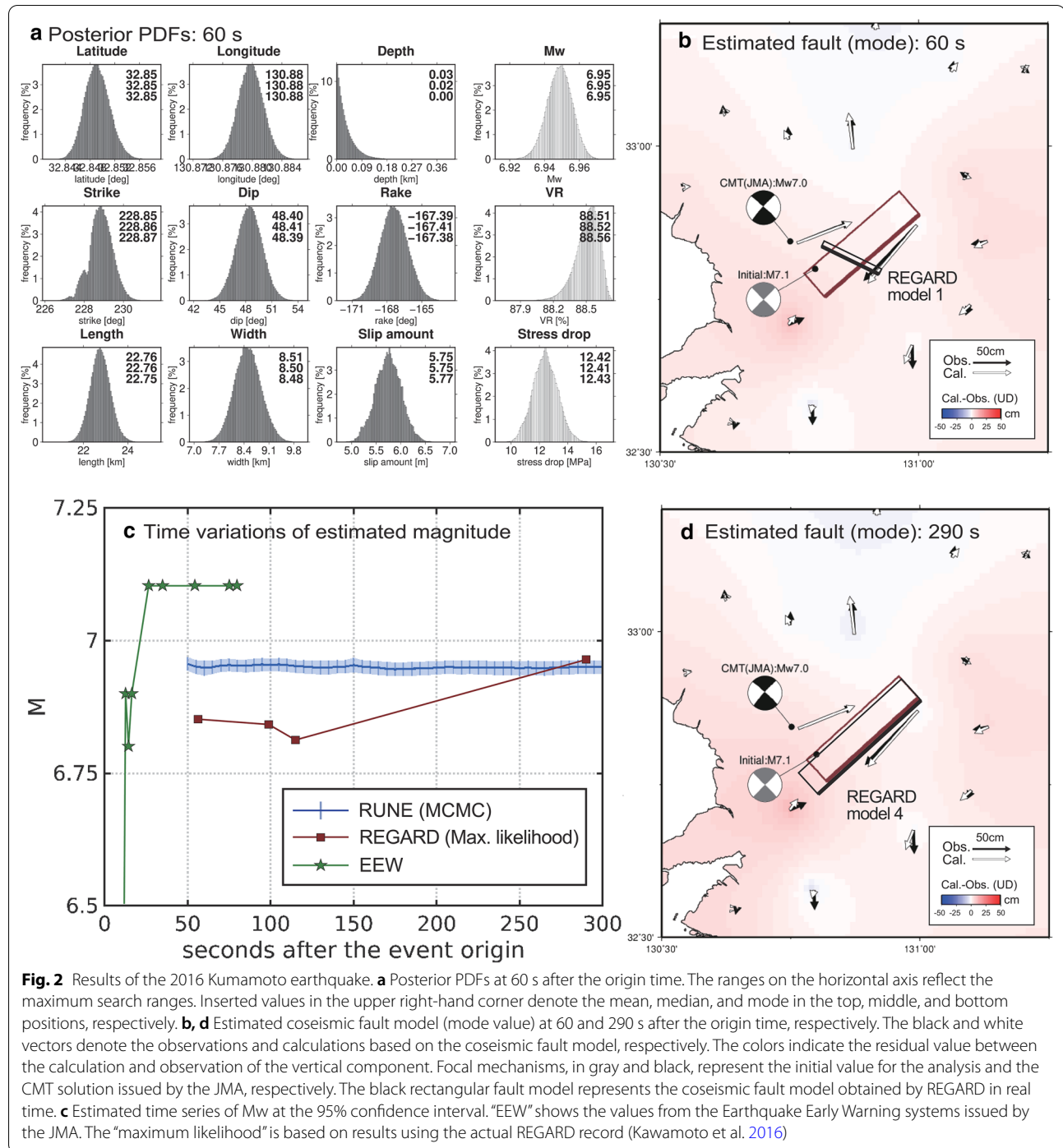
Figure 2a, b, d shows the results for 60 and 290 s after the origin time (here "time" indicates the time of the displacement data, not the time at which the result was obtained) (all marginal PDFs at 60 s: Additional file 1: Figure S2). Figure 2c shows a comparison of the time history of the estimated moment magnitude for three different sources. Our algorithm estimated a northwest dipping strike-slip fault along the Futagawa fault 60 s after the origin time, which is consistent with the estimation of the post-processed fault model (Yarai et al. 2016). In contrast, REGARD yielded a conjugate fault (Fig. 2b), which was northeast dipping because of the incorrect initial value assumption. The estimation condition did

not change until 290 s after the origin time (Fig. 2c, d). REGARD yielded a reasonable fault model after 290 s from the origin time, which was similar to the results of RUNE. In the 2016 Kumamoto earthquake, the source time function was less than 30 s (Kubo et al. 2016). The passage of the large seismic wave around the epicenter also terminated in less than 60 s. The comprehensive permanent displacement pattern did not change in 60 s after the origin time. Thus, the resulting conjugate fault estimated by REGARD may reflect a strong dependency on the initial value. These results indicate the advantages of the algorithm developed in this study, especially the weaker initial value dependency than the maximum likelihood estimation for nonlinear problems.

2019 Yamagata-Oki earthquake (Mw 6.4)

The 2019 Yamagata-Oki earthquake (M_w 6.4) occurred at 14:22:20 UTC on June 18, 2019. This event was an intraplate earthquake that occurred off Yamagata Prefecture along the eastern margin of the Sea of Japan. The focal mechanism proposed by the JMA suggests that reverse fault motion occurred with a northeast to southwest extension axis (Fig. 3b). This earthquake triggered a tsunami, which was observed by tide gauges on the coast of the Sea of Japan (~100 mm; Earthquake Research Committee 2019). Furthermore, the REGARD system automatically estimated the permanent displacement and a fault model with southeast dipping (Fig. 3b). However, the width of the estimated fault was small (1.8 km), and the ratio of the fault length to the width was quite large. The estimation result strongly depends on the initial values; thus, it is difficult to determine how accurately this model was estimated. By applying RUNE to this event, we assessed whether our approach could quantify the estimation uncertainties of the fault parameters. We used 50 stations near the epicenter from the same dataset used for REGARD estimation, considering the low signal-to-noise (S/N) ratio.

Figure 3 shows the estimated results (all marginal PDFs: Additional file 1: Figure S3). Posterior PDFs have multiple peaks in strike, dip, and rake angle, and large tails in length, width, and slip amount. The median values suggest a reverse fault with southeast dipping. Figure 3d shows the normalized fault location frequency, which is the distribution of each sample's fault plane normalized by maximum frequency. The distribution location is close to the location of the centroid moment tensor solution estimated by the JMA and has a complex shape with the nodes in the northwest–southeast and northeast–southwest directions. Figure 3c shows the marginal PDFs between the dip, rake, and strike angles. Multiple peaks in each posterior PDF correspond to four clusters. These clusters correspond to a reverse fault with northwest

Fig. 2 Results of the 2016 Kumamoto earthquake. **a** Posterior PDFs at 60 s after the origin time. The ranges on the horizontal axis reflect the maximum search ranges. Inserted values in the upper right-hand corner denote the mean, median, and mode in the top, middle, and bottom positions, respectively. **b, d** Estimated coseismic fault model (mode value) at 60 and 290 s after the origin time, respectively. The black and white vectors denote the observations and calculations based on the coseismic fault model, respectively. The colors indicate the residual value between the calculation and observation of the vertical component. Focal mechanisms, in gray and black, represent the initial value for the analysis and the CMT solution issued by the JMA, respectively. The black rectangular fault model represents the coseismic fault model obtained by REGARD in real time. **c** Estimated time series of Mw at the 95% confidence interval. "EEW" shows the values from the Earthquake Early Warning systems issued by the JMA. The "maximum likelihood" is based on results using the actual REGARD record (Kawamoto et al. 2016)

dipping, a reverse fault with southeast dipping, and two strike-slip faults perpendicular to them. The same behaviors were observed even when the number of samples was increased.

These results suggest that some fault parameters such as the depth, dip, and fault length and width, were not well-constrained for this event. Using RUNE, we showed that the observed data did not have sufficient resolution
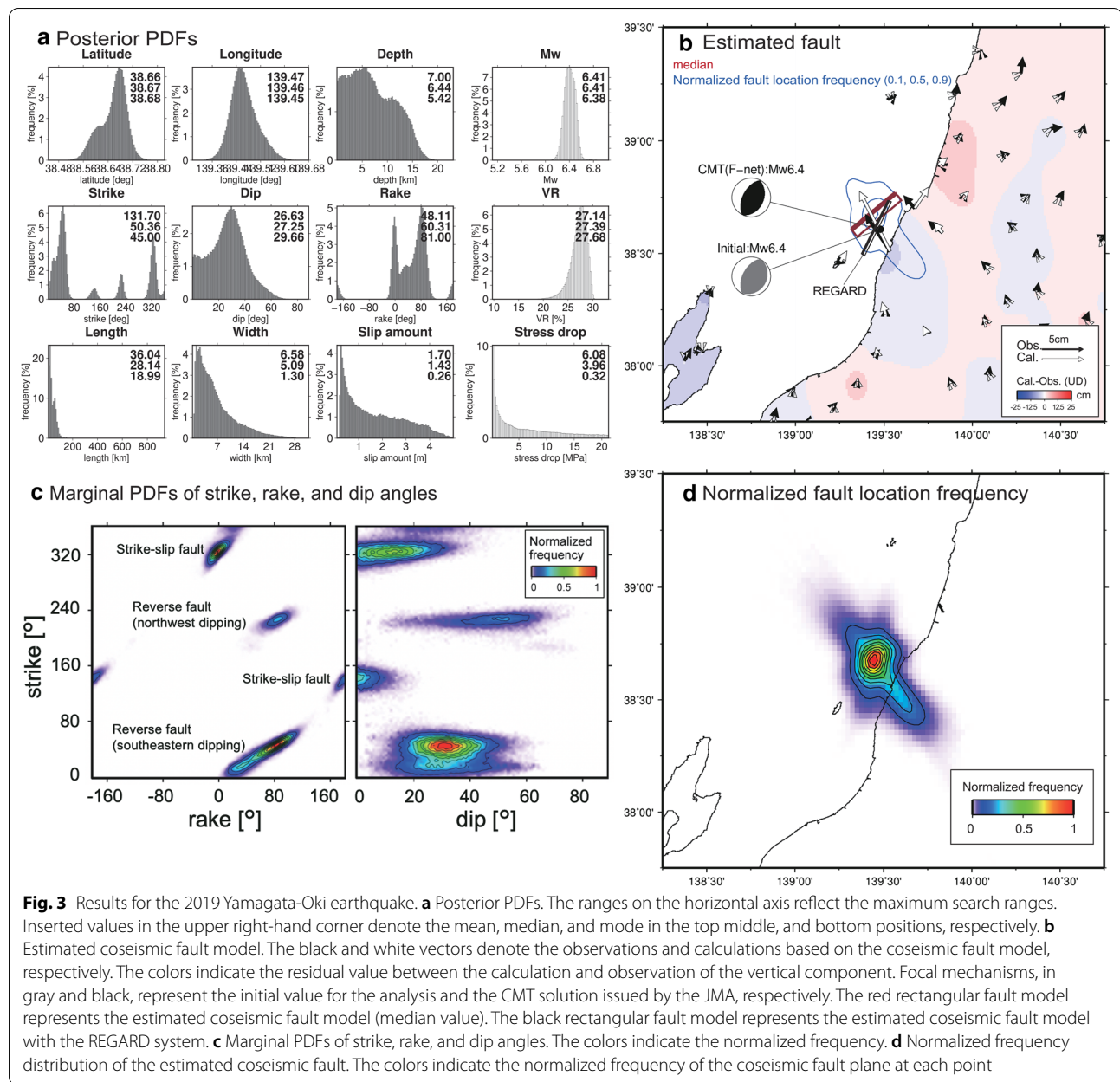
Ohno *et al. Earth, Planets and Space*   (2021) 73:127

Page 9 of 18

Fig. 3 Results for the 2019 Yamagata-Oki earthquake. **a** Posterior PDFs. The ranges on the horizontal axis reflect the maximum search ranges. Inserted values in the upper right-hand corner denote the mean, median, and mode in the top middle, and bottom positions, respectively. **b** Estimated coseismic fault model. The black and white vectors denote the observations and calculations based on the coseismic fault model, respectively. The colors indicate the residual value between the calculation and observation of the vertical component. Focal mechanisms, in gray and black, represent the initial value for the analysis and the CMT solution issued by the JMA, respectively. The red rectangular fault model represents the estimated coseismic fault model (median value). The black rectangular fault model represents the estimated coseismic fault model with the REGARD system. **c** Marginal PDFs of strike, rake, and dip angles. The colors indicate the normalized frequency. **d** Normalized frequency distribution of the estimated coseismic fault. The colors indicate the normalized frequency of the coseismic fault plane at each point

to reject the possibility of the four models, and the reverse fault with southeast dipping was superior to the other models. Thus, it is possible to discuss the reliability of the estimated results from a more diverse perspective by acquiring not only variance reduction but also uncertainties in real time.

Simulated intraplate earthquakes in the Sea of Japan (M_w 6.7–7.8)

In this section, the detectability of intraplate earthquakes was evaluated by numerical simulations using RUNE. The target area is the eastern margin of the Sea of Japan, where numerous non-interplate tsunamigenic earthquakes have occurred, such as the 1940 Shakotan-Oki (M_w 7.5), 1964 Niigata (M_w 7.5), 1983 Middle Sea of Japan (M_w 7.9), and 1993 Hokkaido Nansei-Oki (M_w 7.8) earthquakes (Tanioka et al. 1995). A single rectangular fault model can successfully model a non-interplate earthquake whose geometry is difficult to assume in advance.

To obtain the input coseismic fault models, we adopted 60 fault models proposed by the Japanese government (MLIT, CAO, and MEXT 2014: Additional file 1: Table S1, Figure S4a). Some of the 60 assumed scenarios

are composed of multiple rectangular faults (estimation was performed as a single rectangular fault). The range of the magnitude is M_w 6.7–7.8, which is larger than the 2019 Yamagata-Oki earthquake that occurred in the same area ("2019 Yamagata-Oki earthquake (Mw 6.4)" Section). The assumed fault models have the same mechanism of reverse fault slip as the 2019 Yamagata-Oki earthquake. The input data of RUNE include the displacement data and EEW. We calculated the expected coseismic permanent displacement for 50 GEONET stations surrounding the focal area based on Okada (1992) by adding the expected observation error via a Gaussian distribution with a standard deviation of 2 and 5 cm for the horizontal and vertical components, respectively. For setting the EEW information, we used the average value of the constituting fault centers as latitude and longitude, uniform 10 km as depth, and the closest value to the input in 0.5 intervals as magnitude. Then, we applied RUNE to the 60 cases independently.

Figure 4 shows the estimated results for the 60 cases. For better visualization, Fig. 4a, b shows odd and even case IDs, respectively. The red rectangular faults and blue contours indicate the median and the normalized frequency distribution, respectively. In almost all cases, the estimations were located around the assumed fault models. In the case of multiple rectangular faults' inputs, the coseismic faults were estimated offshore as a single rectangular fault rather than assumed faults (F24 in Fig. 4b). However, the normalized frequency distribution varied depending on the arrangement of the observation stations. The uncertainty of the spread of the fault plane was larger for faults located offshore.

Figure 5 shows the estimated M_w as a box plot, whose horizontal axis is the distance to the nearest observation station. The larger the distance to the nearest observation station, the larger the estimated range of M_w. The green line in Fig. 5 shows the envelopes of M_w required to cause a horizontal displacement of 5 cm (~ maximum displacement observed in the 2019 Yamagata-Oki earthquake) at the nearest observation station for six types of reverse fault mechanisms. In other words, these envelopes indicate that the larger the distance from the epicenter, the larger the M_w required to observe a horizontal displacement of 5 cm. Since the given simulated noise is uniform, the case closer to the envelope line suggests a smaller S/N
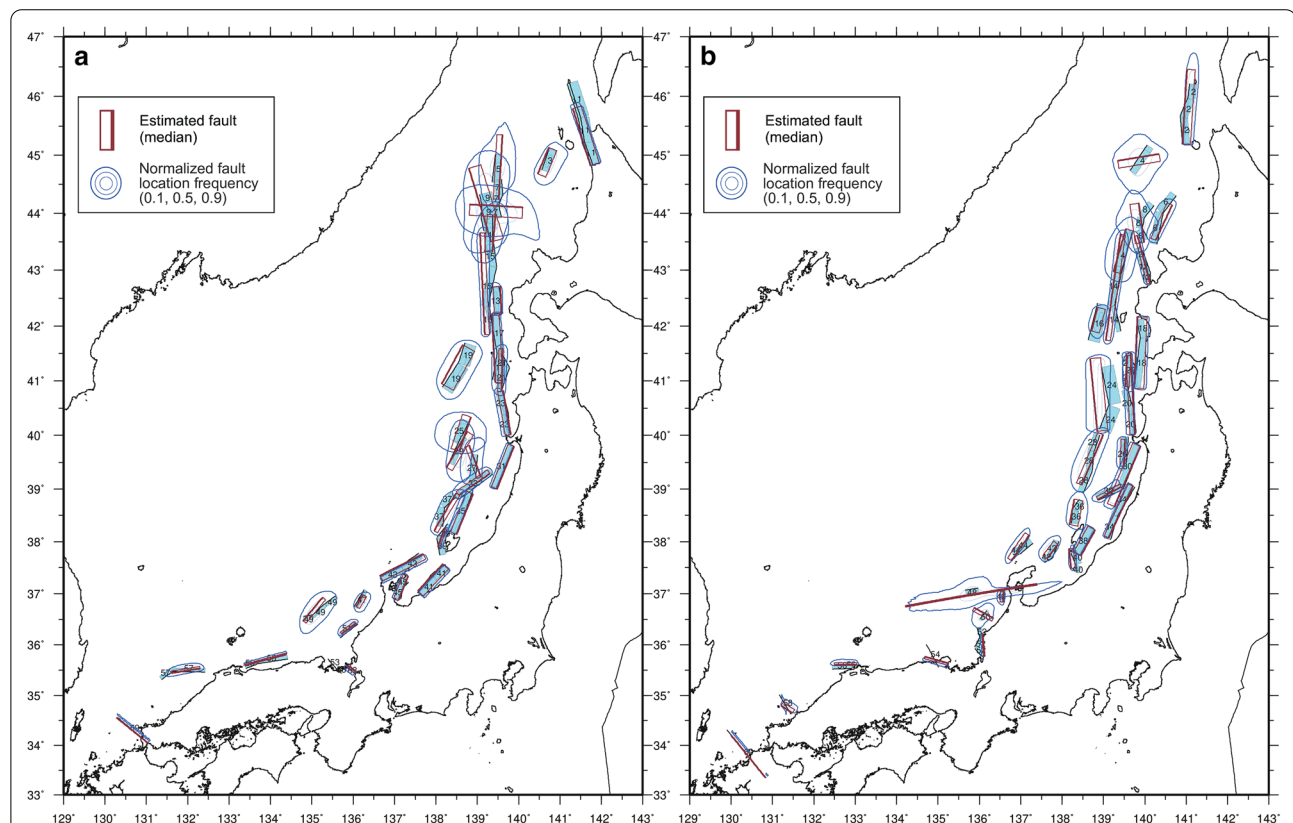

Fig. 4 Results for the simulated earthquakes. **a** Results of odd IDs. **b** Results of even IDs. The red rectangular fault model represents the estimated coseismic fault model (median). The blue contours indicate the normalized frequency distribution of the estimated coseismic fault. The contour of 0.9 is the outermost one and is indicated by a bold line
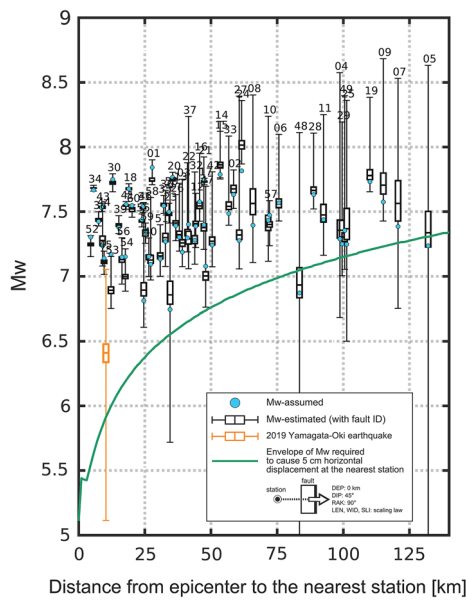
Fig. 5 Relationship between the distance from the epicenter to the nearest observation station and estimated M_w. Blue dots indicate the assumed M_w. Each box plot indicates the result of each fault ID. The yellow box indicates the result of the 2019 Yamagata-Oki earthquake. The green line indicates the theoretical envelope of M_w required to cause 5 cm horizontal displacement at the nearest station. The insert indicates the positional relationship between the station and the fault. We used the scaling law of Utsu (2001) to assume the size and calculated theoretical displacement according to Okada (1992). The small step in the green line near zero on the horizontal axis corresponds to a change in the displacement pattern at the ground surface, because the assumed fault has a finite size according to the scaling law

Table 2 VR and standard deviation of likelihood function in 60 simulations. VR values indicate the mode values

ID	VR [%]			S.D. of likelihood function [cm]	
	All	Horizontal	Vertical	Horizontal	Vertical
1	97.72	98.75	97.04	1.81	4.78
2	92.14	95.33	90.51	2.00	4.97
3	17.57	45.91	− 1.21	1.92	4.59
4	4.78	17.45	− 0.42	1.99	4.78
5	1.28	6.83	− 0.06	1.85	5.35
6	49.13	81.42	− 0.26	2.01	5.34
7	5.22	19.08	0.62	1.93	5.25
8	15.02	37.93	1.53	2.15	5.00
9	17.85	45.62	1.57	2.13	5.29
10	22.23	52.40	1.32	2.23	5.44
11	16.02	48.63	− 0.21	1.73	4.84
12	70.84	89.13	16.41	2.03	4.62
13	68.55	88.56	22.77	2.08	5.05
14	90.85	97.47	72.06	2.04	6.77
15	93.17	97.20	80.81	2.16	4.97
16	78.54	93.60	13.08	1.94	4.88
17	97.12	97.26	97.05	2.36	4.75
18	87.43	95.54	49.96	3.18	7.35
19	64.23	85.18	3.81	1.99	4.21
20	94.16	97.95	49.30	2.22	4.54
21	34.06	71.51	2.25	1.96	5.59
22	45.82	74.34	− 0.06	2.14	4.72
23	85.83	94.94	33.64	1.98	4.31
24	89.30	96.51	10.28	2.19	4.99
25	6.14	24.41	0.37	1.94	5.54
26	63.49	90.03	6.87	1.81	5.34
27	19.96	49.71	0.94	1.91	4.77
28	56.12	83.70	2.58	1.83	4.57
29	8.98	29.69	0.55	1.95	5.14
30	92.76	94.37	87.66	4.24	5.07
31	89.59	97.38	70.11	1.92	5.78
32	37.88	76.17	− 1.53	1.80	5.20
33	52.50	80.24	6.98	2.06	4.93
34	97.37	98.68	95.62	2.32	4.90
35	94.29	92.27	94.77	1.98	4.78
36	46.89	76.59	16.27	1.95	5.14
37	25.96	60.96	2.31	1.96	5.40
38	87.78	95.68	76.56	1.90	5.26
39	79.49	95.35	10.65	1.90	5.71
40	78.79	88.85	61.45	2.41	4.83
41	94.75	97.99	71.49	2.23	4.45
42	25.89	50.46	4.81	2.04	4.32
43	97.33	98.43	96.63	2.12	5.47
44	84.38	87.33	83.39	1.98	5.53
45	77.21	87.16	65.20	2.33	6.70
46	22.15	49.04	2.94	1.99	4.60
47	14.95	43.59	3.35	1.89	5.49

ratio. Therefore, the smaller the S/N ratio of the observation data, the larger the estimated range of M_w.

Table 2 shows the $\sigma_i (i = EN, U)$ values of the likelihood function fixed in RUNE stage 2 in each case. The σ_i values are close to the noise of the simulated observation data: the standard deviations of Gaussian noises are 2.0 and 5.0 cm for the horizontal and vertical components, respectively. Since the same Green's function (Okada 1992) was used to simulate the data and to estimate the fault models, there were no modeling errors. Their similarities indicate that the dynamic adjustment of hyperparameters in RUNE worked well. Table 2 also shows the median value of VR in each case. Although some cases have high VR values exceeding 95%, the estimated faults may not always reproduce the assumed fault model. For example, the F01 result whose VR is 97.72% cannot estimate the fault plane extending to the north direction, where there is no station (Fig. 4a). This result suggests that a model with higher VR may not always be a model with higher accuracy.

Ohno *et al. Earth, Planets and Space*   (2021) 73:127

Page 12 of 18

Table 2 (continued)

ID	VR [%]			S.D. of likelihood function [cm]	
	All	Horizontal	Vertical	Horizontal	Vertical
48	1.91	8.71	− 0.36	1.80	4.79
49	11.11	39.30	− 0.78	1.92	5.35
50	3.57	16.42	− 0.42	1.97	5.37
51	49.18	78.17	8.26	2.08	5.07
52	74.24	82.40	72.39	3.48	13.11
53	50.50	57.61	40.69	3.96	5.86
54	71.97	85.30	55.34	2.15	10.27
55	92.73	97.58	36.78	2.06	4.38
56	53.05	78.60	6.66	2.13	4.69
57	49.69	78.71	12.14	1.94	4.90
58	39.30	73.12	− 5.68	2.17	5.40
59	92.79	97.66	62.18	2.00	4.48
60	97.90	99.25	93.27	3.13	5.74
Ave. (1σ)	56.36 (33.01)	72.82 (26.98)	31.37 (35.96)	2.15 (0.48)	5.35 (1.34)

Discussion

Practical use of the estimated uncertainties for real-time purposes

In "Results" Section, the proposed method was applied to both past and synthetic intraplate events. The method provides a reliable coseismic fault model with the uncertainty of each parameter in real time. In this section, we discuss using our method with the obtained uncertainties of the fault model in a real-time purpose.

As mentioned in "Introduction" Section, the GSI and Tohoku University have developed the REGARD system for coseismic fault estimation in real time using GNSS data. One of the purposes of the REGARD system is to estimate an accurate moment magnitude, because GNSS provides unsaturated magnitude values for large events. This information is used by the JMA as supporting information when issuing tsunami warnings/advisories. The moment magnitude is the basic starting information to decide whether to issue a tsunami warning/advisory. Real-time GNSS data can provide a relatively accurate moment magnitude value despite the instability of the other parameters. The proposed method can also provide the PDF of the moment magnitude (Figs. 2, 3, Additional file 1: Figure S5). To avoid under/overestimation of the tsunami warning/advisory, information on the uncertainty of the moment magnitude is required.

Currently, the JMA only uses information on the moment magnitude from the REGARD system, but the GNSS can provide not only the moment magnitude value but also the finite fault model. In contrast, the onshore GNSS data cannot strictly distinguish between the fault slip amount and fault area. Figure 6 shows the correlation between the slip amount and fault area for all four events with lines indicating the stress drop values. All events show a clear trade-off between the slip amount and fault area, whereas the magnitude is relatively well-constrained. For example, the moment magnitude was stably estimated for the 2019 Yamagata-Oki event despite a relatively low VR value compared to the other events. The 2016 Kumamoto earthquake was also characterized

Fig. 6 Marginal PDFs between the estimated fault area and slip amount, **a** without 60 simulations except for simulated F16 earthquake in the Sea of Japan and **b** with 60 simulations. The contours denote the normalized frequency of the marginal PDF between 0.1 and 0.9. The thick black line denotes the calculated moment magnitude. The thin dashed lines denote the stress drop values calculated using Eq. (4). Gray hatched areas denote the non-searchable areas constrained by the stress drop value

by a trade-off between the VR value and moment magnitude. The range, however, was smaller than that for the other events. For the 2016 Kumamoto earthquake, station coverage was sufficient to estimate the fault model compared to the other offshore events. Fault area information is necessary to estimate an accurate initial sea surface distribution (Ohta et al. 2012; Tsushima and Ohta 2014; Tsushima et al., 2014). For practical real-time monitoring, a comparison with past large earthquakes in such figures to understand the trade-offs would be useful to validate the potential uncertainties in the fault area.

As mentioned in "2019 Yamagata-Oki earthquake (Mw 6.4)" Section, the posterior PDFs of the fault parameters also enable us to discuss the reliability of the estimated results. The VR value is required for the validation of the estimation reliability. However, the results of the numerical simulations in "Simulated intraplate earthquakes in the Sea of Japan (Mw 6.7–7.8)" Section suggest that the VR value may not always support accuracy. Therefore, discussing the results together with their uncertainties provides new evidence, unlike discussing the results only with the estimated values, depending on the initial value and its VR.

The results of REGARD, including the single rectangular fault model, were also adopted as sources to estimate tsunami inundation, and resulting damage. This system for tsunami inundation and resulting damage has been developed based on the results of the "The research group of real-time estimation for tsunami inundation" (Ohta et al. 2018). The system has been adopted as part of the Disaster Information Systems (DIS) operated by Japan's Cabinet Office (Musa et al. 2018; Ohta et al. 2018) and will use the generated information for initial governmental response after tsunamigenic earthquakes. The quantitative uncertainties of the fault models are directly related to the uncertainties of the vertical seafloor deformation and the tsunami hazard (Additional file 1: Figure S6). The method developed in this study for the coseismic fault model has the potential to enhance tsunami prediction (i.e., by the addition of the uncertainty information).

In this study, we only focused on the estimation of the single rectangular fault model for moderate-sized earthquakes ($\sim M_w$ 7.5). It is still necessary to understand the intermediate large non-interplate earthquakes, such as inland and intraplate earthquakes, to obtain accurate

real-time predictions. However, this simple model cannot explain more complex rupture processes, including inhomogeneous slip distribution. Future work should develop and implement a method to estimate the fault geometry and slip distribution simultaneously, as proposed by Minson et al. (2014a).

Origin of quantified uncertainties

In "Practical use of the estimated uncertainties for real-time purposes" Section, we discussed using the obtained uncertainties of the fault model with our method. The uncertainty of each parameter includes both the observation and model errors. The adjusted hyperparameters ($\sigma_i (i = EN, U)$ and $\Delta\boldsymbol{\theta}^{max}$), a priori information, and the number of stations may affect uncertainties.

First, we discuss the effect of the standard deviation of the likelihood function σ_i. To obtain the standard deviation in real time, we adopted Eq. (8), where the σ_i values strongly depend on the stability and size of the displacements. Thus, σ_i cannot be identical to the typical RTK-GNSS observation error; the difference between the analytical solution of Okada (1992) and the actual observation is used as σ_i (Eq. (8)). Figure 5 shows the comparison between the estimated M_w of the 2019 Yamagata-Oki earthquake ("2019 Yamagata-Oki earthquake (Mw 6.4)" Section) and the results of the numerical simulations ("Simulated intraplate earthquakes in the Sea of Japan (Mw 6.7–7.8)" Section). The range of M_w of the 2019 Yamagata-Oki earthquake is larger, although the fixed σ_i of the likelihood function is smaller ([EN, U] = [0.58 cm, 1.12 cm], Table 2), because in the 2019 Yamagata-Oki Earthquake, the S/N ratio was smaller, and/or there was non-Gaussian local noise at each station. Furthermore, kinematic time series after large earthquakes are affected by the dynamic rupture process and seismic wave passage. Adjustment of σ_i with Eq. (8) can avoid overfitting of such data.

Second, we discuss the effect of the utilization of a priori information. RUNE has prior PDFs on the latitude, longitude, and depth, based on the EEW value ("Sampling flow for real-time use" Section). A normal distribution of the magnitude of the earthquake with a standard deviation is given for the epicenter (Table 1). This is the constraint that increases the reliability of the epicenter estimated from the seismic waveform data


(See figure on next page.)

Fig. 7 Results for the simulated earthquake F16 in different EEW positions. **a**, **c**, and **e** Posterior PDFs. The ranges on the horizontal axis reflect the maximum search ranges. Inserted values in the upper right hand corner denote the mean, median, and mode in the top, middle, and bottom positions, respectively. Solid red and black dashed lines denote the given and initial fault parameter, respectively. **b**, **d**, and **f** Estimated coseismic fault model. The black and white vectors denote the observations and calculations based on the coseismic fault model, respectively. The colors indicate the residual value between the calculation and observation of the vertical component. Focal mechanisms, in gray, represent the initial values for the analysis. The red and black rectangular fault models represent the estimated (median value) and given coseismic fault model, respectively. The blue contours indicate the normalized frequency distribution of the estimated coseismic fault
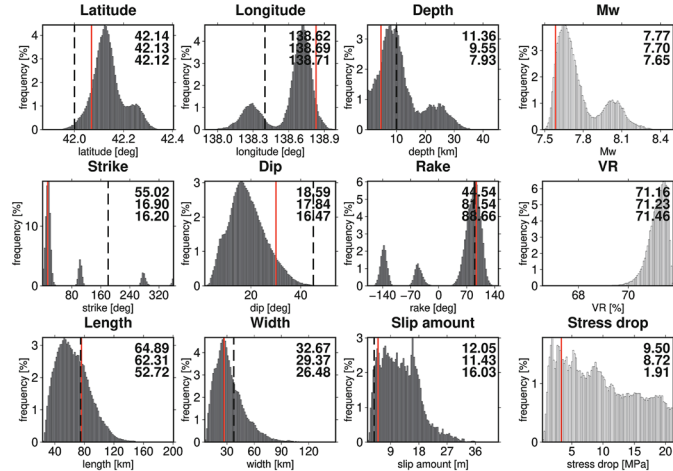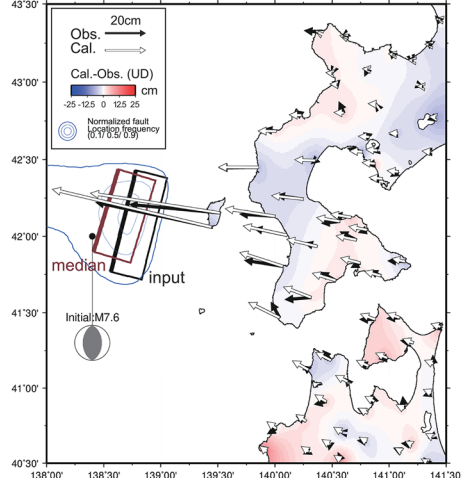
Ohno *et al. Earth, Planets and Space* (2021) 73:127

Page 14 of 18

EEW (Lat., Lon.)=(138.9, 42.0)

a Posterior PDFs
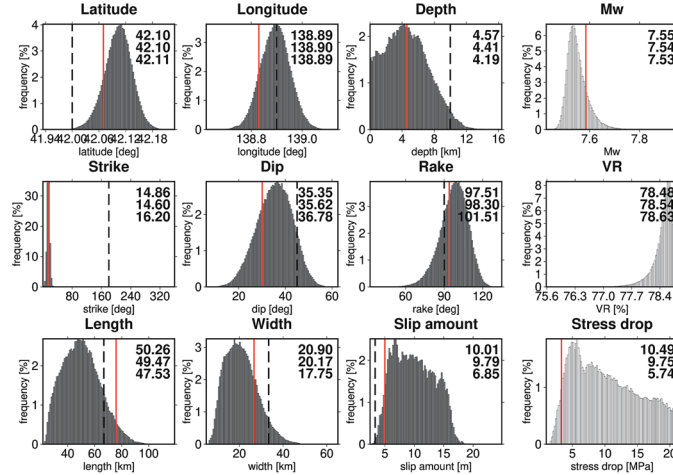
b Estimated fault
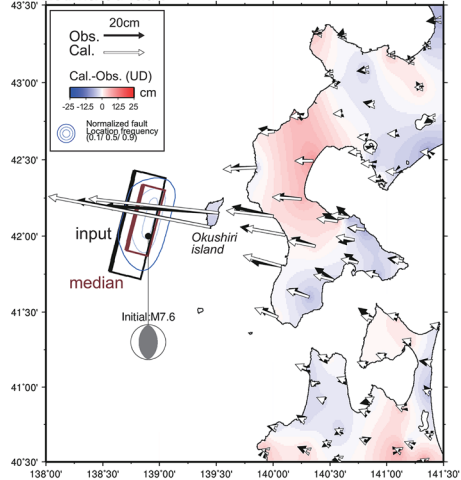
EEW (Lat., Lon.)=(139.5, 42.0)

c Posterior PDFs
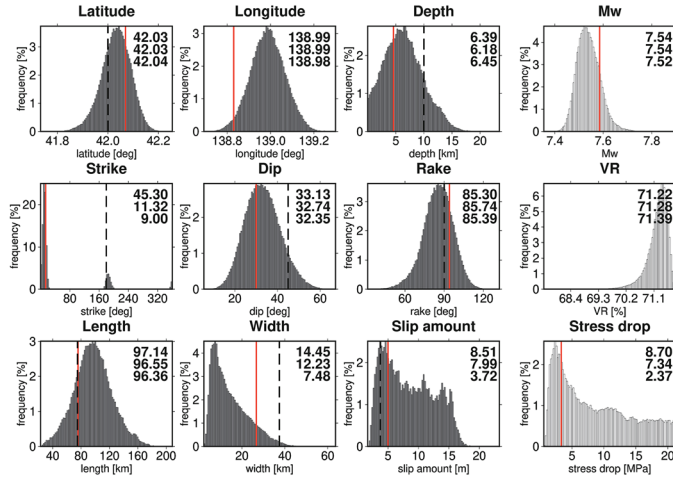
d Estimated fault
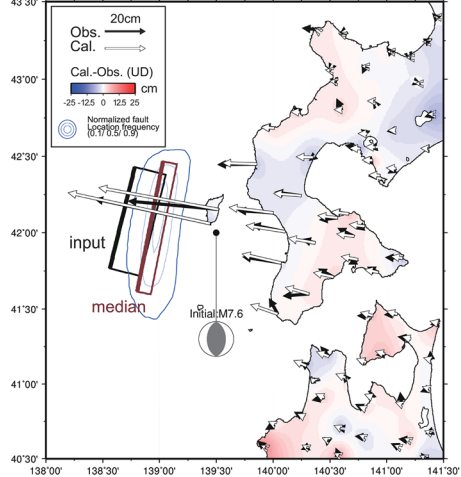
EEW (Lat., Lon.)=(138.4, 42.0)

e Posterior PDFs

f Estimated fault

Ohno *et al. Earth, Planets and Space*   (2021) 73:127

Page 15 of 18

when the magnitude of the earthquake is small (the S/N ratio of GNSS is small or spatial coverage of the GNSS site is not good). In such an estimation, GNSS data do not have the resolution to converge the estimation to the assumed fault position beyond the a priori position constraint. Figure 7 shows the estimate results for the simulated earthquake F16 at Sea of Japan in different EEW positions. If the hypocenter location by the EEW is significantly different from actual one, the estimated fault location is likewise estimated in a location that slightly deviates from actual one, although the value of VR does not change significantly. Although MCMC has an advantage similar to that of grid search, the MCMC search depends on the likelihood function and the constraint. Therefore, it is difficult to quantify the variations in the parameters in one estimation and in those with a limited number of samples.

In RUNE, the prior PDF was designed by applying it to the actual data of the 2019 Yamagata-Oki earthquake, which was a small earthquake. One of the methods to solve the difficulties in search efficiency is the parallel tempering approach, which is also implemented in RUNE. To further increase the search efficiency, the parallel chains can be started from different initial values (latitude, longitude), and the number of parallel chains can be increased. Furthermore, although EEW data were used as a priori information in this study, the CMT solution based on the seismic data can be used without any change in the code. Moreover, although past estimation results with high VR were not applied to the time series data (the 2016 Kumamoto earthquake), they can be used as a priori information for stability.

Next, we discuss the effect of the number of stations. REGARD consists of three subsystems, but RUNE predicts the replacement of the fault model estimation subsystem, and the selection of observation points to be used is not included in the current algorithm. However, owing to the tendency of MCMC to have a small initial value dependency and the current policy of minimizing the a priori constraint of RUNE, the spatial arrangement of used observation points affects the search behavior. In the application to cases, where the S/N is small, if the observation point arrangement is adopted unnecessarily far from the fault position, the fault length may unrealistically increase to explain them (Fig. 4b and Table 2, fault ID: 48). Besides, when using Okada's model (1992), the homogeneous and isotropic elastic half-space assumption becomes problematic when using GNSS sites very far from the actual focal area. The relationship between the observation point selection and search behavior requires a future systematic study.

We can also optimize RUNE for objectives other than those mentioned above, because the settings affect the estimation results. For example, we can fix the standard deviations of the likelihood function in stage 2 with an additional parameter, i.e., $\sigma_i = \max\left(\widehat{\sigma}_i^{\mathrm{medi}}, \mathrm{Obs.err_i}\right)$. This parameter is dependent on the observation error and can prevent overfitting. Depending on the problem, it may be possible to use this parameter to effectively and properly define the observation error. If the time scale associated with estimating the observation error is much larger (1 day to 1 year), we can apply our approach to post-seismic long-time scale events. Furthermore, we can utilize other a priori information as prior PDF easily (CMT solution, distribution of aftershocks). The algorithm itself can be applied to post hoc analysis without any changes in the code.

Finally, we discuss the contribution of the Green's function. In this study, we adopted the conventional elastic half-space analytic solution (Okada 1992) as the Green's function. The uncertainties of unknown parameters should contain both the observation and modeling errors. However, in this study, we chose this simple analytic solution as the Green's function because of its relatively low computational costs compared with more realistic solutions, such as the layered spherical earth (Pollitz, 1996), three-dimensional finite element (FE) method (e.g., Kyriakopoulos et al. 2013), and elastic/viscoelastic FE method (e.g., Ichimura et al. 2016). Considering future improvements in computing ability and other factors, these approaches are promising for obtaining more reliable results in real time.

Conclusions

Uncertainties in earthquake models are significant indicators of model reliability. One way to estimate uncertainties in nonlinear problems is to use a full Bayesian inversion approach such as the MCMC. However, the MCMC is computationally expensive and the hyperparameters should be set up in advance via trial and error (initial model sample, standard deviations of the likelihood function, and maximum step size). To overcome these problems, a new method was developed to estimate the coseismic fault model in real time, combined with the uncertainties of the obtained fault model, using a full Bayesian inversion approach. Efficient sampling techniques were adopted using a parallel tempering method and a multistage approach. The first stage was only used to automatically estimate the optimal hyperparameter values for the second stage. In the second stage, the algorithm focused on sampling for the posterior PDFs. Using a standard 1U server, the calculation time was

Ohno *et al. Earth, Planets and Space*   (2021) 73:127

Page 16 of 18

determined to be less than 30 s for 1×10^6 samples, which is ideal for use with a real-time system, such as REGARD.

The reliability of the method was evaluated using data from past earthquakes (the 2016 Kumamoto and 2019 Yamagata-Oki earthquakes) and 60 simulated earthquakes in the Sea of Japan. The results showed the advantage of the new method over the existing one for initial value dependence in nonlinear problems. In applications to data with a small S/N ratio, the estimated results suggested the possibility of using several conjugate fault models. There was a trade-off between the fault area and slip amount, especially for earthquakes that occur offshore. This suggests a lack of sensitivity for earthquakes generated offshore using only onshore GNSS data. These results imply that quantification of the uncertainty enables us to discuss the reliability of the fault model estimation in real time.

To apply the developed method to real earthquakes, discussion with decision makers is necessary, because the required information varies depending on current conditions. Awareness of the limitations of the estimated results is also required. Finally, the new method can be used as a tool to determine which fault parameters are not constrained based on the current network and how to optimize the new site for a more robust model estimation.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40623-021-01425-0.

Additional file 1: Table S1. Fault parameters of 60 simulated earthquakes. **Figure S1.** Computation times required for the "RUNE" algorithm (1.1×10^6 steps per chain using eight parallel chains). **Figure S2**. All marginal PDFs for the 2016 Kumamoto earthquake (60 s after the origin time). **Figure S3.** All marginal PDFs for the 2019 Yamagata-Oki earthquake. **Figure S4.** Map of 60 simulated earthquakes. **Figure S5.** Marginal PDFs between VR and M_W. **Figure S6.** Expected seafloor crustal deformation based on the estimated coseismic fault model.

Authors' information

KO is a past graduate student at the Graduate School of Science, Tohoku University, and is now part of the staff at the Geospatial Information Authority of Japan. YO and RH are an associate professor and a professor, respectively, at the Graduate School of Science, Tohoku University. S. KA and SA are a member of staff at the Geospatial Information Authority of Japan. S. KO is a professor at the International Research Institute of Disaster Science, Tohoku University.

AM is a visiting professor at the Cyberscience Center, Tohoku University. HK is a professor at the Graduate School of Information Sciences, Tohoku University.

Authors' contributions

KO and YO designed the study, developed the programs, and analyzed the data. S. KA, SA, and YO developed the REGARD system. AM and HK contributed to speeding-up the developed programs. YO, RH, HK, and S.KO assisted in the interpretation of the results. All authors read and approved the final manuscript.

Availability of data and materials

All 30-s GEONET raw data for post-processing can be obtained via the GSI website. Other data used and/or analyzed in this study are available from the corresponding author upon reasonable request.

Declarations

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Research Center for Prediction of Earthquakes and Volcanic Eruptions, Graduate School of Science, Tohoku University, 6-6 Aza-Aoba, Aramaki, Aoba-ku, Sendai 980-8578, Japan. [2]Division for the Establishment of Frontier Sciences of Organization for Advanced Studies, Tohoku University, Sendai, Japan. [3]International Research Institute of Disaster Science, Tohoku University, Sendai, Japan. [4]Present Address: Geospatial Information Authority of Japan, Tsukuba, Japan. [5]Cyberscience Center, Tohoku University, Sendai, Japan. [6]NEC Corporation, Tokyo, Japan. [7]Computer Architecture Laboratory, Graduate School of Information Sciences, Tohoku University, Sendai, Japan.

References

Blewitt G, Kreemer C, Hammond WC, Plag H-P, Stein S, Okal E (2006) Rapid determination of earthquake magnitude using GPS for tsunami warning systems. Geophys Res Lett 33:L11309. https://doi.org/10.1029/2006GL026145

Blewitt G, Hammond WC, Kreemer C, Plag H-P, Stein S, Okal E (2009) GPS for real-time earthquake source determination and tsunami warning systems. J Geod 83:335–343. https://doi.org/10.1007/s00190-008-0262-5

Crowell BW, Bock Y, Melgar D (2012) Real-time inversion of GPS data for finite fault modeling and rapid hazard assessment. Geophys Res Lett 39(9):1–6. https://doi.org/10.1029/2012GL051318

Colombelli S, Allen R M, Zollo A (2013) Application of real-time GPS to earthquake early warning in subduction and strike-slip environments, J Geophys Res Solid Earth, 118, 3448–3461. https://doi.org/10.1002/jgrb.50242

Dettmer J, Benavente R, Cummins PR, Sambridge M (2014) Trans-Dimensional Finite-Fault Inversion Geophys J Int 199:735–751

Dosso SE, Wilmut MJ (2006) Data uncertainty estimation in matched-field geoacoustic inversion. IEEE J Oceanic Eng 31:470–479. https://doi.org/10.1109/JOE.2006.875099

Earthquake Research Committee (2019) Evaluation of Earthquake Offshore Yamagata Prefecture on June 18, 2019. https://www.jishin.go.jp/main/chousa/19jun_yamagata/index-e.htm.

Ohno *et al. Earth, Planets and Space* (2021) 73:127

Page 17 of 18

Geyer CJ (1991) Markov Chain Monte Carlo Maximum Likelihood, Interface Foundation of North America. Retrieved from the University of Minnesota Digital Conservancy, https://hdl.handle.net/11299/58440.

Hastings WK (1970) Monte carlo sampling methods using markov chains and their applications. Biometrika 57(1):97–109. https://doi.org/10.2307/2334940

Hoechner A, Ge M, Babeyko AY, Sobolev SV (2013) Instant tsunami early warning based on real-time GPS-Tohoku 2011 case study. Nat Hazards Earth Syst Sci 13(5):1285–1292. https://doi.org/10.5194/nhess-13-1285-2013

Ichimura T, Agata R, Hori T, Hirahara K, Hashimoto C, Hori M, Fukahata Y (2016) An elastic/viscoelastic finite element analysis method for crustal deformation using a 3-D island-scale high-fidelity model. Geophys J Int 206(1):114–129. https://doi.org/10.1093/gji/ggw123

Ito T, Gunawan E, Kimata F, Tabei T, Simons M, Meilano I, Agustan A, Ohta Y, Nurdin I, Sugiyanto D (2012) Isolating along-strike variations in the depth extent of shallow creep and fault locking on the northern Great Sumatran Fault. J Geophys Res 117:B06409. https://doi.org/10.1029/2011JB008940

Jasra A, Stephens DA, Holmes CC (2007) On population-based simulation for static inference. Stat Comput 17(3):263–279. https://doi.org/10.1007/s11222-007-9028-9

Jiang J, Simons M (2016) Probabilistic imaging of tsunamigenic seafloor deformation during the 2011 Tohoku-Oki Earthquake. J Geophys Res Sol Earth 12:9050–9076. https://doi.org/10.1002/2016JB013760

Kanamori H, Anderson DL (1975) Theoretical basis of some empirical relations in seismology. Bull Seismol Soc Am 65(5):1073–1095

Kawamoto S, Hiyama Y, Ohta Y, Nishimura T (2016) First result from the GEONET real-time analysis system (REGARD): the case of the 2016 Kumamoto earthquakes. Earth Planets Space 68(190):1–12. https://doi.org/10.1186/s40623-016-0564-4

Kawamoto S, Ohta Y, Hiyama Y, Todoriki M, Nishimura T, Furuya T, Sato Y, Yahagi T, Miyagawa K (2017) REGARD: A new GNSS-based real-time finite fault modeling system for GEONET. J Geophys Res Solid Earth 122:1324–1349. https://doi.org/10.1002/2016JB013485

Kubo H, Suzuki W, Aoi S, Sekiguchi H (2016) Source rupture processes of the 2016 Kumamoto, Japan, earthquakes estimated from strong-motion waveforms. Earth Planets Space 68(161):1–13. https://doi.org/10.1186/s40623-016-0536-8

Kyriakopoulos C, Masterlark T, Stramondo S, Chini M, Bignami C (2013) Coseismic slip distribution for the M w 9 2011 Tohoku-Oki earthquake derived from 3-D FE modeling. J Geophys Res. 118:3837–3847

Larson KM (2009) GPS seismology. J Geod 83(3–4):227–233. https://doi.org/10.1007/s00190-008-0233-x

Larson KM, Bodin P, Gomberg J (2003) Using1-HzGPSdatato measure deformations caused by the Denali fault earthquake. Science 300:1421–1424. https://doi.org/10.1126/science.1084531

Matsu'ura M, Hasegawa Y (1987) A maximum likelihood approach to nonlinear inversion under constraints. Phys Earth Planets Inter 47:179–187

Melgar D, Bock Y (2015) Kinematic earthquake source inversion and tsunami runup prediction with regional geophysical data. J Geophys Res Solid Earth 120:3324–3349. https://doi.org/10.1002/2014JB011832

Melgar D, Bock Y, Crowell BW (2012) Real-time centroid moment tensor determination for large earthquakes from local and regional displacement records. Geophys J Int 188(2):703–718

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

Ministry of Land, Infrastructure, Transport and Tourism (MLIT), Cabinet Office (CAO), Ministry of Education, Culture, Sports, Science and Technology (MEXT) (2014) Study meeting on large scale earthquakes in the Japan Sea. http://www.mlit.go.jp/river/shinngikai_blog/daikibojishinchousa/. (in Japanese).

Minson SE, Simons M, Beck JL, Ortega F, Jiang J, Owen SE, Moore AW, Inbal A, Sladen A (2014a) Bayesian inversion for finite fault earthquake source models—II: The 2011 great Tohoku-Oki. Japan Earthquake Geophys J Int 198(2):922–940. https://doi.org/10.1093/gji/ggu170

Minson SE, Murray JR, Langbein JO, Gomberg JS (2014b) Real-time inversions for finite fault slip models and rupture geometry based on high-rate GPS data. J Geophys Res Solid Earth 119:3201–3231. https://doi.org/10.1002/2013JB010622

Murray JR, Bartlow N, Bock Y, Brooks B, Foster J, Freymueller J, Hammond W, Hodgkinson K, Johanson I, López-Venegas A, Mann D, Mattioli G,

Melbourne T, Mencin D, Montgomery-Brown E, Murray M, Smalley R, Thomas V (2019) Regional global navigation satellite system networks for crustal deformation monitoring. Seismol Res Lett 91(2A):552–572. https://doi.org/10.1785/0220190113

Musa A, Watanabe O, Matsuoka H, Hokari H, Inoue T, Murashima Y, Ohta Y, Hino R, Koshimura S, Kobayashi H (2018) Real-time tsunami inundation forecast system for tsunami disaster prevention and mitigation. J Supercomput 74(7):3093–3113. https://doi.org/10.1007/s11227-018-2363-0

Ohno K, Ohta Y (2018) Trial of quantitative uncertainty estimation for the single rectangular coseismic fault model deduced from real-time GNSS data using the MCMC approach: a case study of the 2011 Tohoku-Oki earthquake. J Geod Soc Japan 64:39–50 (**(in Japanese)**)

Ohta Y (2016) Study on the improvement of real-time kinematic GNSS data analysis and development of rapid coseismic fault determination method based on those data. J Geod Soc Japan 62(1):1–19. https://doi.org/10.11366/sokuchi.62.1. (**(in Japanese)**)

Ohta Y, Meilano I, Sagiya T, Kimata F, Hirahara K (2006) Large surface wave of the 2004 Sumatra-Andaman earthquake captured by the very long baseline kinematic analysis of 1-Hz GPS data. Earth Planets Space 58(2):153–157. https://doi.org/10.1186/BF03353372

Ohta Y, Ohzono M, Miura S, Iinuma T, Tachibana K, Takatsuka K, Miyao K, Sato T, Umino N (2008) Coseismic fault model of the 2008 Iwate-Miyagi Nairiku earthquake deduced by a dense GPS network. Earth Planets Space 60(12):1197–1201

Ohta Y, Kobayashi T, Tsushima H, Miura S, Hino R, Takasu T, Fujimoto H, Iinuma T, Tachibana K, Demachi T, Sato T, Ohzono M, Umino N (2012) Quasi real-time fault model estimation for near-field tsunami forecasting based on RTK-GPS analysis: application to the 2011 Tohoku-Oki earthquake (Mw 9.0). J Geophys Res 117:B02311. https://doi.org/10.1029/2011JB008750

Ohta Y, Kobayashi T, Hino R, Demachi T, Miura S (2015) Rapid coseismic fault determination of consecutive large interplate earthquakes: The 2011 Tohoku-Oki sequence (2015). Int Assoc Geodesy Symposia. https://doi.org/10.1007/1345_2015_109

Ohta Y, Inoue T, Koshimura S, Kawamoto S, Hino R (2018) Role of real-time GNSS in near-field tsunami forecasting. J Disaster Res 13(3):453–459. https://doi.org/10.20965/jdr.2018.p0453

Okada Y (1992) Internal deformation due to shear and tensile faults in a half-space. Bull Seismol Soc Am 82(2):1018–1040

Ozaki T (2011) Outline of the 2011 off the pacific coast of Tohoku earthquake (M_w 9.0) -Tsunami warnings/advisories and observations. Earth Planets Space 63:827–830

Pollitz FF (1996) Coseismic deformation from earthquake faulting on a layered spherical earth. Geophys J Int 125(1):1–14. https://doi.org/10.1111/j.1365-246X.1996.tb06530.x

Roberts GO, Rosenthal JS (2001) Optimal scaling for various Metropolis-Hastings algorithms. Stat Sci 16(4):351–367

Tanaka Y, Ohta Y, Miyazaki S (2019) Real-time coseismic slip estimation via the GNSS carrier phase to fault slip approach: a case study of the 2016 Kumamoto earthquake. Geophys Res Lett 46(3):1367–1374

Tanioka Y, Satake K, Ruff L (1995) Total analysis of the 1993 Hokkaido Nansei-oki earthquake using seismic wave, tsunami, and geodetic data. Geophys Res Lett 22(1):9–12

Terakawa T, Matsu'ura M (2009) A methodological breakthrough in stress inversion: Estimation of 3-D tectonic stress patterns from CMT data inversion. Zisin J Seismol Soc Japan 2:339–346

Tsushima H, Ohta Y (2014) Review on near-field tsunami forecasting from offshore tsunami data and onshore GNSS data for tsunami early warning. J Dis Res 9(3):339–357. https://doi.org/10.20965/jdr.2014.p0339

Tsushima H, Hino R, Ohta Y, Iinuma T, Miura S (2014) tFISH/RAPiD: Rapid improvement of near-field tsunami forecasting based on offshore tsunami data by incorporating onshore GNSS data. Geophys Res Lett 41:3390–3397. https://doi.org/10.1002/2014GL059863

Utsu T (2001) Seismology (Jishingaku 3rd edn). Kyoritsu Shuppan, ISBN-13: 978-4320046375

Wells DL, Coppersmith KJ (1994) New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. Bull Seismol Soc Am 84(4):974–1002

Yarai H, Kobayashi T, Kawamoto S, Morishita Y, Fujiwara S, Munekane H, Hiyama Y, Yamaguchi K, Suzuki A, Abe S, Miyahara B, Miura Y, Kakiage Y, Ueshiba

Ohno *et al. Earth, Planets and Space*   *(2021) 73:127*

Page 18 of 18

H (2016) Overview of crustal deformation associated with the 201–6 Kumamoto earthquakes. Earth Planets Space. https://doi.org/10.1186/s40623-016-0563-5

Publisher's Note

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