Cattania, C., Hainzl, S., Wang, L., Roth, F., Enescu, B. (2014): Propagation of Coulomb stress uncertainties in physics-based aftershock models. - *Journal of Geophysical Research*, *119*, 10, p. 7846-7864.

DOI: http://doi.org/10.1002/2014JB011183

AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Solid Earth

RESEARCH ARTICLE

10.1002/2014JB011183

Key Points:
- High impact of Coulomb stress uncertainties on aftershock modeling
- The variable orientation of receiver faults generates large stress heterogeneity
- Ensemble models based on a set of alternative slip models perform best

Supporting Information:
- Readme
- Text S1

Correspondence to:
C. Cattania,
camcat@gfz-potsdam.de

Citation:
Cattania, C., S. Hainzl, L. Wang, F. Roth, and B. Enescu (2014), Propagation of Coulomb stress uncertainties in physics-based aftershock models, *J. Geophys. Res. Solid Earth*, *119*, 7846–7864, doi:10.1002/2014JB011183.

Received 8 APR 2014
Accepted 19 SEP 2014
Accepted article online 23 SEP 2014
Published online 30 OCT 2014

Propagation of Coulomb stress uncertainties in physics-based aftershock models

Camilla Cattania[1], Sebastian Hainzl[1], Lifeng Wang[2], Frank Roth[1], and Bogdan Enescu[3,4]

[1]GFZ German Research Centre for Geosciences, Potsdam, Germany, [2]China Earthquake Networks Center, Beijing, China, [3]Earth Evolution Sciences Department, Faculty of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Japan, [4]Institute of Statistical Mathematics, Tokyo, Japan

Abstract Stress transfer between earthquakes is recognized as a fundamental mechanism governing aftershock sequences. A common approach to relate stress changes to seismicity rate changes is the rate-and-state constitutive law developed by Dieterich: these elements are the foundation of Coulomb-rate-and-state (CRS) models. Despite the successes of Coulomb hypothesis and of the rate-and-state formulation, such models perform worse than statistical models in an operational forecasting context: one reason is that Coulomb stress is subject to large uncertainties and intrinsic spatial heterogeneity. In this study, we characterize the uncertainties in Coulomb stress inherited from different physical quantities and assess their effect on CRS models. We use a Monte Carlo method and focus on the following aspects: the existence of multiple receiver faults; the stress heterogeneity within grid cells, due to their finite size; and errors inherited from the coseismic slip model. We study two well-recorded sequences from different tectonic settings: the M_w = 6.0 Parkfield and the M_w = 9.0 Tohoku earthquakes. We find that the existence of multiple receiver faults is the most important source of intrinsic stress heterogeneity, and CRS models perform significantly better when this variability is taken into account. The choice of slip model also generates large uncertainties. We construct an ensemble model based on published slip models and find that it outperforms individual models. Our findings highlight the importance of identifying sources of errors and quantifying confidence boundaries in the forecasts; moreover, we demonstrate that consideration of stress heterogeneity and epistemic uncertainty has the potential to improve the performance of operational forecasting models.

1. Introduction

Stress transfer between faults has been identified as a plausible mechanism for earthquake triggering, and a correlation between positive Coulomb stress changes and seismicity rate has repeatedly been observed [*King et al.*, 1994; *Harris*, 1998]. While controversies exist on the relative importance of various processes (such as coseismic slip, afterslip, and the redistribution of stresses by small events), all models are based on the same physical concept: the elastic response of the Earth's crust to one or more planar dislocations. *Okada* [1992] gave a set of solutions to model the deformation and stress field in an homogeneous elastic half-space, which has subsequently been extended to the more general case of a viscoelastic, layered half-space [*Wang et al.*, 2006]. In both cases, several simplifications are required to model the Coulomb stress field (ΔCFS), and various sources of errors are present. These include the following: uncertainties and finite resolution of the slip model, small-scale variability of elastic parameters in the crust, lack of knowledge of the orientation of the faults that can accommodate future seismicity, and discretization of the crust into cells of finite volume.

While a simple treatment may be sufficient to demonstrate a qualitative agreement between the distribution of seismicity and ΔCFS, careful consideration of uncertainties is required in order to construct models which are useful in the context of operational earthquake forecasting. In this work, we focus in particular on models which employ a rate-and-state seismicity evolution [*Dieterich*, 1994]: due to strong nonlinearity, such models are very sensitive to variability in input stress values, and several authors have shown that including stress variability leads to a significantly different spatiotemporal distribution of seismicity.

Marsan [2006] and *Helmstetter and Shaw* [2006] demonstrated that small-scale variability can lead to the suppression of stress shadows in the short term, a result that can be understood from the fact that according to the rate-and-state constitutive law, seismicity rate increases exponentially with stress change

immediately after a stress step, and hence, the largest values of a distribution dominate at short times. Therefore, small-scale variability can explain the observation of early aftershocks in areas which experience an average negative stress, including the vicinity of the rupture area on the main shock fault. Stress variability also affects the temporal evolution of seismicity. In the rate-and-state formulation, seismicity following a positive stress step is predicted to decay following an Omori law with $p = 1$ and c dependent on the value of stress change: due to the variability in c value, an effective value of $p < 1$ is observed when stresses of different magnitude are included [*Helmstetter and Shaw*, 2006]. These results indicate that a careful treatment of the variability in Coulomb stress may be a critical aspect for Coulomb-rate-and-state (CRS) models aimed at forecasting aftershock sequences. In fact, *Hainzl et al.* [2009] showed that when Coulomb stress uncertainties are included in CRS models of the Landers, 1992 $M_w = 7.3$ aftershock sequence, the correlation with observed seismicity improves and the inversion of rate and state parameters is more stable. They estimated uncertainties of ΔCFS by comparing the stress field obtained from existing published slip models, and they found the standard deviation of ΔCFS to be spatially similar to the absolute stress change; therefore, they modeled uncertainties by assuming a spatially constant coefficient of variation (CV, the ratio between standard deviation and absolute value). While this approach has the advantage of being simple to implement, it has few drawbacks: it introduces a new parameter (CV), and it may not provide an accurate spatial description of the variability of stress field, as pointed out by *Woessner et al.* [2012].

Our goal is to present methods by which uncertainties can be included into CRS models in a physically meaningful way and to perform an assessment of the variability of ΔCFS and its effect on the forecasts of aftershock sequences. We estimate uncertainties in input quantities by using prior information about fault orientation and slip characteristics, and we explicitly propagate them through the model using Monte Carlo simulations, thus without making any prior assumption on the distribution of ΔCFS. We focus on three sources of uncertainty: (1) the choice of receiver faults on which the stress tensor is resolved; (2) the existence of a range of stress values within each cell, due to its finite volume; and (3) uncertainties in the slip models, due both to its finite resolution and to modeling assumptions such as choice of geometry or inversion method. After outlining the modeling framework in section 2, we present methods to model the variability of ΔCFS: to model the presence of multiple receiver fault orientations, we sample the distribution of past focal planes; we then consider the effect of finite resolution of the calculation grid and suggest two techniques to reproduce this aspect of stress heterogeneity; finally, we outline methods to reproduce the effect of slip model uncertainties by adding a synthetic, small-scale slip and by comparing the output from different published slip models. In section 4, we study the effect of each source of uncertainty on the forecast, focusing on two case studies from different tectonic settings: the Parkfield, $M_w = 6.0$, strike slip earthquake and the Tohoku, $M_w = 9.0$, megathrust event, both of which have well-recorded aftershock sequences and for which several input slip models are available. After describing the sensitivity of the models to various sources of uncertainty, we discuss practical ways to combine this information into a single model, in particular drawing a distinction between aleatoric and epistemic uncertainties. In section 5 we test how these techniques affect model performance; a physical interpretation of these results and their significance in the context of medium term earthquake forecasting are discussed in the final part of the article.

2. Modeling Procedure

2.1. Coulomb Stress

The theory of static Coulomb stress triggering considers aftershocks as the consequence of the stresses induced by the static deformation of the crust following slip on a planar dislocation (main shock). According to the Coulomb-Mohr criterion, failure on a plane occurs when the sum of shear and frictional stress equals or exceeds the cohesive strength; therefore, seismicity is promoted when Coulomb stress (CFS = $\tau - \mu\sigma_{tot}$) increases and inhibited when it decreases. Here τ is the shear stress in the direction of slip, μ is the coefficient of friction, and σ_{tot} is the total normal pressure on a plane; the latter is given by $\sigma_{tot} = \sigma - p$, where σ is the direct effect of crust deformation and p the change in pore pressure due to the response of fluids. According to the apparent friction poroelastic model [*Cocco and Rice*, 2002, and references therein] $\Delta p = -B\Delta\sigma$, with B the Skempton coefficient which varies between 0 and 1. Hence, Coulomb stress changes can be expressed as

$$\Delta\mathrm{CFS} = \Delta\tau - \mu'\Delta\sigma \tag{1}$$

with $\mu' = \mu(1 - B)$. Typical values of μ are 0.6–0.8 [*Byerlee*, 1978].

2.2. Rate and State Evolution of Seismicity

Dieterich [1994] derived a constitutive law for the evolution of seismicity on a populations of faults governed by rate-and-state-dependent friction, subject to a stress perturbation. The seismicity rate R is given by

$$R = \frac{r_0}{\gamma \dot{\tau}_r} \tag{2}$$

where $\dot{\tau}_r$ is the constant shear stressing rate and γ is a constitutive parameter which evolves according to

$$\mathrm{d}\gamma = \frac{1}{A\sigma}\left[\mathrm{d}t - \gamma \mathrm{d}S\right] \tag{3}$$

with $S = \tau - (\mu - \alpha)\sigma_{\mathrm{tot}}$ is a positive nondimensional constitutive parameter [*Linker and Dieterich*, 1992]. It should be noted that earthquake nucleation in rate-and-state framework is based on physically different principles as the simple Coulomb failure criterion, in particular due to the dependence of friction on slip and slip rate. However, the quantity S, a linear combination of shear and normal stress, is analogous to CFS, and the concepts underling Coulomb stress transfer can be useful to interpret the spatial distribution of seismicity. Note that ΔS can be written as ΔCFS with modified coefficient of friction $\mu' = (\mu - \alpha)(1 - B)$ instead of $\mu' = \mu(1 - B)$. We use $\mu' = 0.3$ but also test the effect of choosing other values (Appendix A).

2.3. Model Implementation

When fitting the model to the catalog data, we determine the rate-and-state parameters $A\sigma$ and $t_a = A\sigma/\dot{\tau}_r$ with the highest likelihood value by using a simple grid search algorithm, described in *Hainzl et al.* [2009] (Appendix A). The background rate is fixed by the observed past seismicity (see section 2.4). In principle, a spatially variable background rate r_0 can be obtained from a smoothed, declustered catalog of past seismicity. However, CRS models with a nonhomogeneous background rate have been shown to be more sensitive to the details of the stress field [*Cocco et al.*, 2010]; thus, the uncertainties in the background model may be amplified. Moreover, using an homogeneous background rate facilitates the interpretation of the effect of other uncertainties, which is the main focus of this study. Therefore, the background rate r_0 is assumed to be spatially homogeneous.

In order to study the effect of stress variability on a subgrid scale and the use of different slip models, we use a Monte Carlo technique: we create 100 perturbed versions of the stress field and calculate the spatiotemporal distribution of seismicity for each iteration from the rate-and-state equations. The perturbed stress fields are obtained based on physical consideration of different sources of uncertainties, as described in section 3. In order to avoid the singularities of the Okada solutions, we follow the common approach of imposing a cutoff value for ΔCFS: we set a value of 10 MPa for Tohoku and 1 MPa for Parkfield.

2.4. Data

2.4.1. Parkfield

We use the Advanced National Seismic System (ANSS) catalog, combined with the more complete catalog of *Peng and Zhao* [2009] for the first 2 days following the main shock. The combined catalog has a completeness of $M_c = 2.0$ starting from 100 s after the main shock [*Peng and Zhao*, 2009]: we only fit data starting from the time of completeness up to 250 days from the main shock. We estimate the background rate based on the ANSS catalog from 1970 up to the main shock: rates are calculated for time windows of 30 days and 1 year, and the background rate is found visually ignoring increased values after large ($M_w \geq 6.0$) events, as shown in the supporting information (A). Due to the ongoing aftershock sequence of the Simeon earthquake, the rate at the time of the Parkfield main shock is higher than the background rate; we test the effect of this by starting the model with a nonsteady state rate of $r(0) = 1.144\ \mathrm{d}^{-1}$ and find no significant difference in model behavior: the results presented here are based on a background rate of $r_0 = 0.22\ \mathrm{d}^{-1}$ for the area indicated in Figure 1 and depth range between [0.5, 13.5] km (the calculation domain). We calculate Coulomb stresses from the slip model of *Wang et al.* [2012], unless otherwise stated.

2.4.2. Tohoku

We use the catalog from the Japan Meteorological Agency. We estimate a completeness magnitude of $M_c = 4.0$ starting from 5 days after the main shock: also in this case, we limit our study to the time period in which the catalog is complete, up to 250 days from the main shock. The background rate for the area in Figure 1 and depth between [3.0, 42.0] km is estimated from the Japan Meteorological Agency catalog from 2010 up to the main shock: we find a value of $r_0 = 0.2\ \mathrm{d}^{-1}$, as shown in the supporting information (A). Unless otherwise stated, stress calculations are based on the slip model of *Wang et al.* [2013], which uses the same geometry of *Pollitz et al.* [2011].

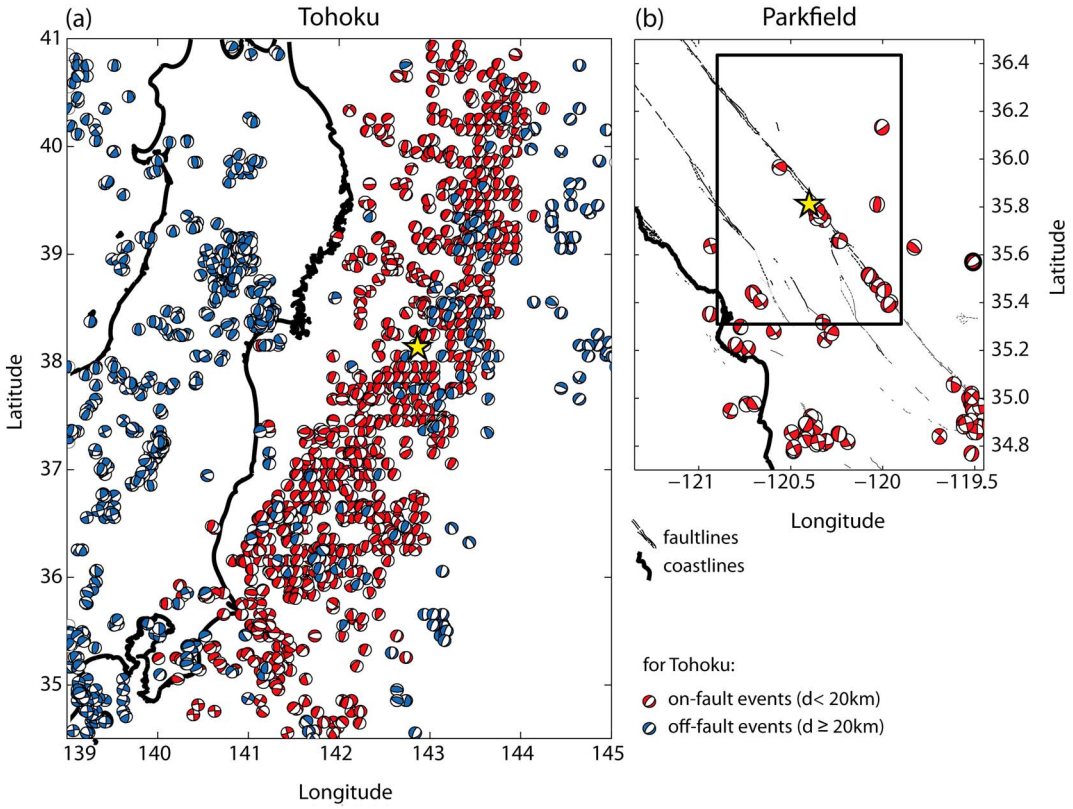

Figure 1. Maps of focal mechanisms in the area of the two main shocks, which we used as a set of receiver faults on which Coulomb stresses are resolved. (a) Focal mechanisms of $M_w \geq 3.1$ events in the area of the Tohoku earthquake, between February 1997 and the main shock. Events are color coded by distance from the fault: red events are considered on fault ($d < 20$ km) and those in blue are the intraplate events. This selection has been done from the observed clustering of focal mechanisms parameters (strike and dip) in Figure 2a. (b) Focal mechanisms of $M_w \geq 2.0$ events in the Parkfield area, between 1 January 1981 and the time of the main shock. The box in the Parkfield map indicates the model domain, and the depth range is [0.5, 13.5] km; for Tohoku, the entire area is the model domain, and the depth range is [3.0, 42.0] km. The choice of these areas was dictated by the distance at which corresponding Coulomb stresses are significant.

3. Treatment of Uncertainties

3.1. Receiver Fault Orientation

In principle, the total seismicity in a region of the crust should be obtained by adding the contribution from each of the fault planes in that volume, assuming that earthquakes occur on preexisting planes. However, it is in practice not possible to have a complete knowledge of all the faults. The two most commonly adopted assumptions are the following: (1) all events have the same focal plane orientation as the main shock and (2) earthquakes rupture along optimally oriented Coulomb failure planes, on which the total (background + coseismic) Coulomb stress is maximum. The first approach oversimplifies the geometry of most fault systems, in which faults exist with different orientation from the main shock. The second approach leads, in general, to more successful forecasts [*Cocco et al.*, 2010]; however, it relies on the assumption that faults exist everywhere with the optimal orientation and that only these faults will produce earthquakes. Considering only the contribution of the optimally oriented fault at each grid point is not consistent with rate-and-state modeling: while in this framework seismicity is predicted to be dominated by the largest ΔCFS values at short times ($t \ll t_a$), the contribution of faults with lower ΔCFS would have an impact at later times [*Helmstetter and Shaw*, 2006].

In order to account for the presence of multiple fault orientations, we use the catalog of past focal mechanisms in the area of interest, assuming that past focal planes provide a reliable sample of the existing fault structures. We use both possible planes from the catalog as possible rupture planes. Since the direction of slip on a preexisting plane depends on the external field, we only use the strike and dip values from the

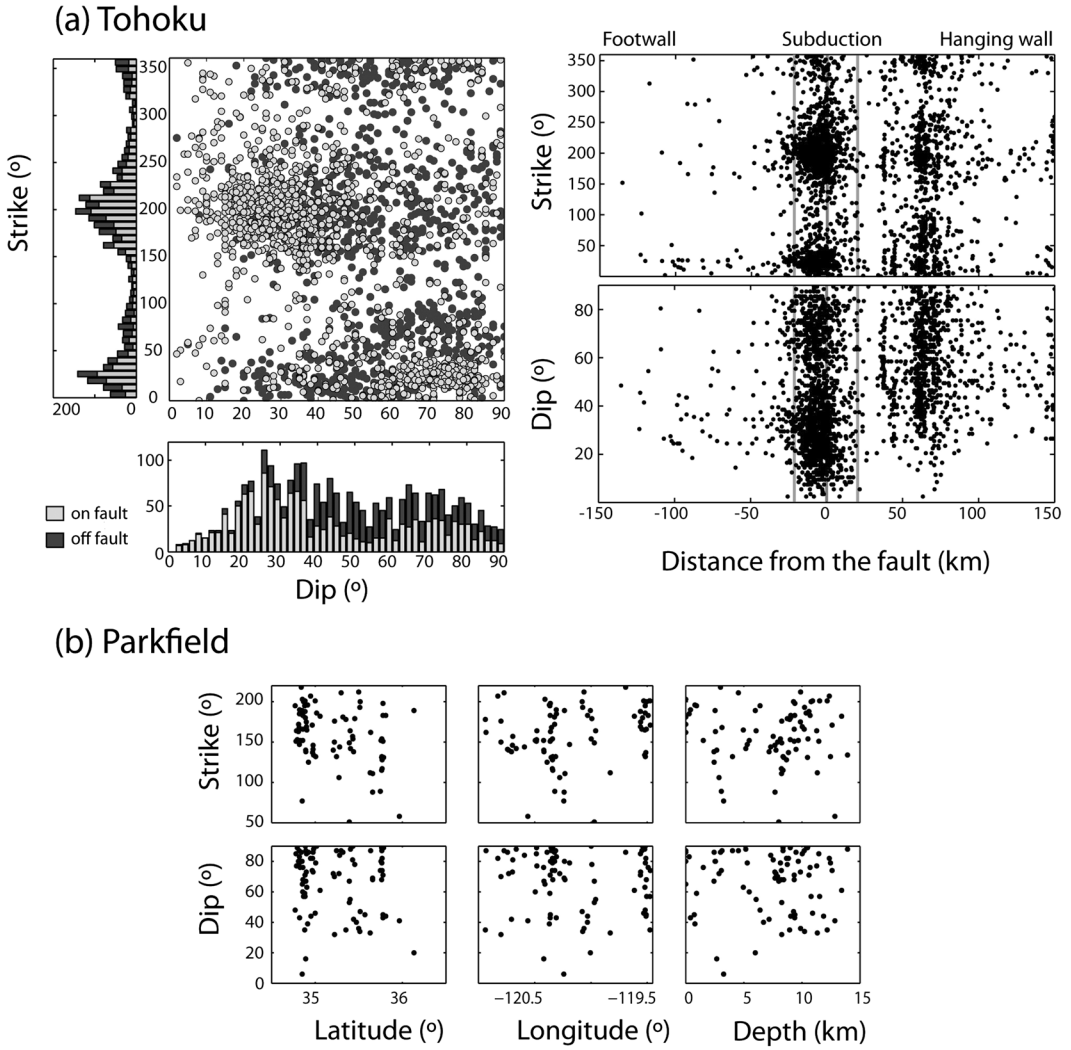

Figure 2. Statistical distribution of focal planes parameters for the planes used as receiver faults shown in Figure 1. (a) For Tohoku, we distinguish between three groups of events: within 20 km from the fault, more than 20 km above, and more than 20 km below, based on changes in focal plane distributions displayed in Figure 2a (right). Figure 2a (left) shows the distribution of strike and dip for on-fault and off-fault events (defined by the 20 km threshold); in particular, the distribution of dips shows a larger fraction of steep faults (dip = 40–90°) being activated further from the slab. (b) For Parkfield, strike and dip do not show clear spatial variations, and we do not distinguish between different zones.

catalog and choose the rake for which shear stress τ is highest. For Tohoku, we use a differential stress of 20 MPa [*Hasegawa et al.*, 2011], with orientation such that the main shock mechanism (strike = 195°, dip = 10°, and rake = 88°) is optimally oriented; for Parkfield, the differential stress is 10 MPa [*King et al.*, 1994], and the main shock plane (strike = 330°, dip = 89°, and rake = 180°) is optimally oriented.

3.1.1. Data

For Parkfield, we use the catalog of focal mechanisms from Yang, Haukkson, and Shearer [*Yang et al.*, 2012]. We select events occurring between the start of the catalog (1 January 1981) and the time of the main shock; we also perform spatial selection, using the spatial window corresponding to Figure 1 and use only events with $M_w \geq 2.0$, giving a total of 81 focal mechanisms. We investigate whether spatial trends in the orientation of these focal planes exists, and we find a rather homogeneous spatial distribution of focal mechanisms (as confirmed by Figures 1b and 2b): therefore, we sample the entire set of focal planes at each grid point.

For Tohoku, we use the *F* net moment tensor focal mechanisms solutions [*Okada et al.*, 2004], determined and provided by the National Research Institute for Earth Science and Disaster Prevention, Japan, and

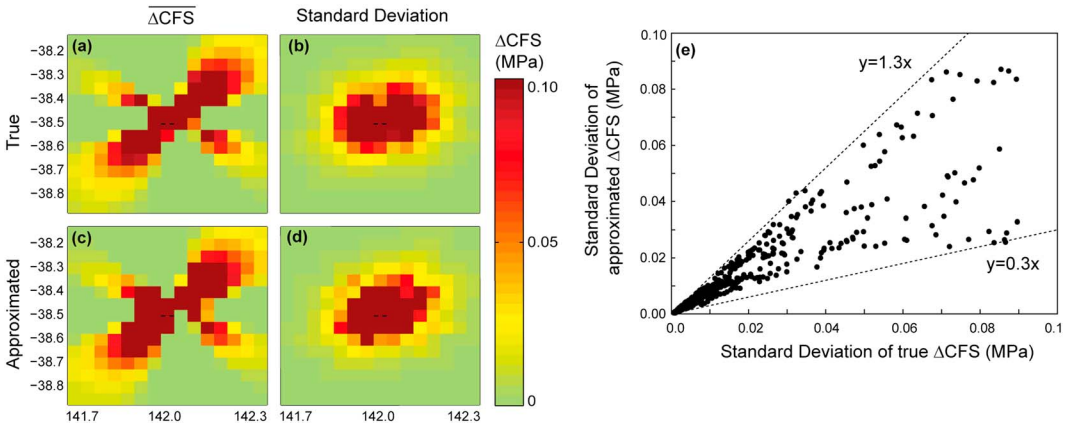

Figure 3. Synthetic test comparing the distribution of stress values due to finite cell size. (a, b) Mean and standard deviation of the distributions obtained from subdividing each cell into 125 subcells, by reducing the size by a factor of 5 in each dimension; (c, d) approximated values, obtained by first-order estimation of stress gradients described in equation (4). (e) Approximated values (obtained from estimating gradients) as a function of the real values (obtained from the subdividing the cells), with each point corresponding to a cell: for 95% of the points, the approximated values differ from the true values by less than a factor of 2. The distributions are generated using a synthetic slip model of a strike-slip event of $M_w = 6.0$ on a fault with strike = 90°, dip = 90°, and the depth interval for stress calculations of 4 km.

containing events of $M_w \geq 3.1$ from February 1997. The total number of focal mechanisms preceding the Tohoku earthquake is 2626. Compared to the Parkfield case, we find here a nonuniform spatial distribution of focal mechanisms, due to the larger area considered and to the presence of a wider range of fault structures, such as the subduction interface, splay faults, and intraplate faults in the hanging wall and footwall [*Asano et al.*, 2011]; in particular, we distinguish different distributions of strike and dip for subduction-related events and intra plate events on each side of the subduction interface (Figure 2a). In order to capture the local variability of receiver faults, we divide the catalog into three groups based on Figure 2a: events in the footwall, at a distance $d > 20$ km below the fault (208 events); events in the hanging wall, at $d > 20$ km from the fault (588 events); and subduction-related events, at $d < 20$ km (1830 events). The latest group includes a larger variety of focal planes, reflecting the coexistence of different fault structures within few tens of kilometers from the main shock plane. At each grid point, receiver faults are sampled from one of these groups, depending on the grid point location.

A caveat to this approach is that past seismicity did not occur uniformly on all existing faults but was higher on those which were best oriented in the past stress field so that the catalog of focal planes may not be a reliable proxy of existing faults. In the supporting information (B), we compare the distribution of focal planes from past events and aftershocks: we find that aftershocks tend to take place on the same planes which accommodated past seismicity, as opposed to the optimally oriented planes.

3.2. Finite Cell Volume

Uncertainties in CRS models also originate from discretization of the domain: the model expresses seismicity rates in a cell of finite volume, assuming that the value calculated at the center is uniform throughout the cell. This is not the case, especially in the vicinity of the faults where gradients in ΔCFS are highest. The most direct method to address this problem is to sample each cell at multiple locations, by dividing the cell into smaller subcells; while simple to implement, this method leads to a fast growth of problem size and computational requirements. We therefore suggest an alternative method to obtain a first-order estimate of the stress gradients within each cell and hence the range of possible stress values. For each cell, we estimate the range of values as follows:

$$\begin{aligned} \delta_j &= (\Delta CFS_j - \Delta CFS_0)/2 \\ \Delta CFS_{min} &= \Delta CFS_0 + \min\left(\delta_j\right) \\ \Delta CFS_{max} &= \Delta CFS_0 + \max\left(\delta_j\right) \end{aligned} \tag{4}$$

where ΔCFS_0 is the value at the center of the cell and $j = 0, 1, \ldots 6$ refers to the cell itself and the six neighbor cells. For each Monte Carlo iteration, we draw a value from the uniform distribution in the interval

Table 1. List of Input Slip Models[a]

Index	SRCMOD Identifier	Reference
	Tohoku	
1	-	*Wang et al.* [2013]
2	s2011HONSHU01SHAO	*Shao et al.* [2011]
3	s2011HONSHU02SHAO	*Shao et al.* [2011]
4	s2011HONSHU03SHAO	*Shao et al.* [2011]
5	s2011HONSHU04SHAO	*Shao et al.* [2011]
6	s2011TOHOKU01FUJI	*Fujii et al.* [2011]
7	s2011TOHOKU01GUSM	*Gusman et al.* [2012]
8	s2011TOHOKU01HAYE	*Hayes [2011]*
9	s2011TOHOKU01IDEx	*Ide et al.* [2011]
10	s2011TOHOKU01LAYx	*Lay et al.* [2011]
11	s2011TOHOKU01SATA	*Satake et al.* [2013]
12	s2011TOHOKU01WEIx	*Wei and Sladen* (Caltech, Tohoku 2011)
13	s2011TOHOKU01YAMA	*Yamazaki et al.* [2011]
14	s2011TOHOKU01YUEx	*Yue and Lay* [2013]
15	s2011TOHOKU02FUJI	*Fujii et al.* [2011]
16	s2011TOHOKU02GUSM	*Gusman et al.* [2012]
17	s2011TOHOKU02SATA	*Satake et al.* [2013]
18	s2011TOHOKU02WEIx	*Wei et al.* (Caltech: Tohoku 2011)
19	s2011TOHOKU03SATA	*Satake et al.* [2013]
20	s2011TOHOKU03WEIx	*Wei et al.* [2012]
	Parkfield	
1	-	*Wang et al.* [2012]
2	s2004PARKFI01DREG	*Dreger et al.* [2005]
3	s2004PARKFI01CUST	*Custodio et al.* [2005]
4	s2004PARKFI01JIxx	*Ji* (Caltech, Parkfield 2004)

[a]Figures of all slip models can be found in the supporting information (C).

$[\Delta CFS_{min}, \Delta CFS_{max}]$. A synthetic test (Figure 3) shows that the two methods yield similar distributions of stress values: in the remaining part of the paper, results from the approximated method (equation (4)) will be presented. The grid resolutions have been chosen based on the slip model patch size and the total domain size, so as to maintain computational costs reasonable. We used the following values: 0.1° (horizontal), 3 km (vertical) for Tohoku and 0.025° (horizontal), 1 km (vertical) for Parkfield.

3.3. Slip Model Uncertainties

The reliability of slip models is an important aspect to consider when modeling Coulomb stresses [*Hainzl et al.*, 2009; *Woessner et al.*, 2012]. Multiple sources of errors impact slip models, including the following: (1) the nonuniqueness of the inversion; (2) incomplete and erroneous data; (3) inversion algorithm; (4) choice of fixed parameters, including fault geometry, fault discretization, and smoothing; and (5) choice of other physical quantities such as rupture speed and velocity structure. The sensitivity of slip models to these aspects is the subject of several studies [*Yagi and Fukahata*, 2011; *Hartzell et al.*, 2007; *Beresnev*, 2003], and it is becoming more common to provide slip models with estimates of uncertainties [*Sudhaus and Jónsson*, 2009]; techniques such as Bayesian inversion are increasingly being used to incorporate data errors and obtain a family of possible models [*Minson et al.*, 2013], which should be included in a forecasting model to estimate confidence limits. We will distinguish between two types of errors: the finite resolution of an individual slip model and the differences between published slip models.

3.3.1. Subpatch Scale Slip

The resolution of a slip model is limited by the patch size and smoothing, and even when a slip model presents slip on isolated patches, it can be a modeling artifact and has large uncertainties [*Hartzell et al.*, 2007; *Beresnev*, 2003]. In order to consider small-scale slip, we create a set of possible slips by adding a random high-frequency slip, while leaving the low frequencies unchanged. Several authors have suggested that slip is fractal, with power spectrum characterized by a power law decay: $P(k) \propto k^{-2(4-D)}$, with D the fractal dimension and k the wave number. In this work, we used the procedure described in [*Kieling et al.*, 2014] to create a set of slip models with synthetic slip at small wavelengths while keeping the original slip

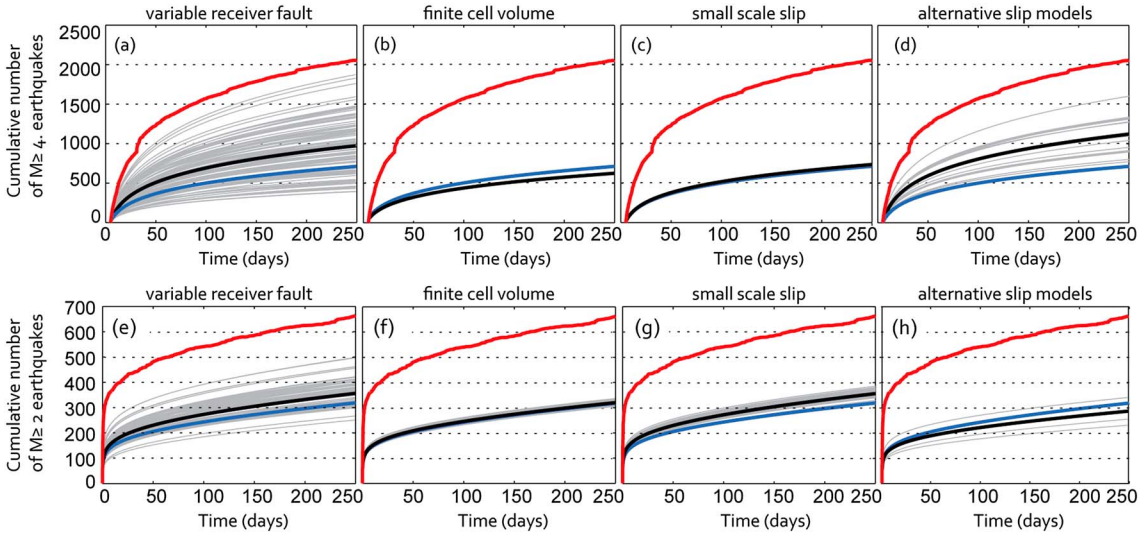

Figure 4. Sensitivity test showing the cumulative number of events versus time, with fixed rate-and-state parameters ($A\sigma = 28$ kPa, $t_a = 8000$ d for Tohoku and $A\sigma = 6$ kPa, $t_a = 10000$ d for Parkfield). Each thin line indicates a Monte Carlo iteration; the black line is their average; the red line is the observed seismicity; the blue line is the forecast obtained without any source of uncertainty (hidden by the black line in Figures 4c and 4f). Forecast for Tohoku, (a) including variability in receiver faults, (b) including stress field gradients within each cell, (c) including subpatch scale slip in the slip model, and (d) using different published slip models. (e–h) Same as above, for Parkfield. For this test, rate-and-state parameters are not optimized, leading to the tendency to underestimate the total number of events.

distribution at wavelengths $k > k_c$, where k_c is the corner wave number given by $k_c = 10^{1.82-0.5M_w}$ [*Causse et al.*, 2010]; we chose a value of $D = 2$.

3.3.2. Choice of Slip Model

When adding synthetic, small wavelength slip, it is assumed that the original slip model is realistic in terms of geometry and large-scale slip distribution. On the other hand, a number of different slip models exist for each main shock, which can differ quite significantly. In order to study the sensitivity of CRS models to the input slip, we use a set of slip models from the finite-source model database maintained by ETH Zürich: in total, we compare a set of four models for Parkfield and 20 models for Tohoku (Table 1). Figures comparing all the slip models can be found in the supporting information (C).

4. Impact of Uncertainties on Forecasted Seismicity

In order to assess the impact of different sources of uncertainty on the forecasted seismicity, we run models with each source of uncertainty included separately. We compare five models, with consideration of the following: (1) uncertainties from receiver fault orientation; (2) uncertainties from finite cell volume; (3) uncertainties from unresolved, small-scale slip; and (4) uncertainties from different published slip models and finally a model without consideration of uncertainty. The goal of this exercise is to perform a sensitivity analysis and not to fit the best possible model to the data; therefore, we set the rate-state parameters to some reasonable but not optimized values ($A\sigma = 6$ kPa, $t_a = 10,000$ d for Parkfield and $A\sigma = 28$ kPa, $t_a = 8000$ d for Tohoku). We run 100 iterations of the model, altering each of the sources of variability in turn. In models which do not consider receiver fault uncertainty, we use the average main shock focal mechanism; in models which do not consider slip model uncertainty, we use the original slip model without addition of short wavelength slip.

Figure 4 shows the effect of uncertainties on the forecasted temporal evolution of seismicity. In Figures 4a and 4e, each line corresponds to the choice of a different receiver focal plane; it is clear, for both case studies, that this choice has a more profound effect on the forecast than the other sources of uncertainty. We find that considering the finite volume of cells has very limited impact on the temporal evolution of seismicity for Tohoku and a moderate effect for Parkfield (Figures 4b and 4f); the effect of small-scale slip is also very small (Figures 4c and 4g) and negligible compared to the impact of using different published slip models (Figures 4d and 4h).

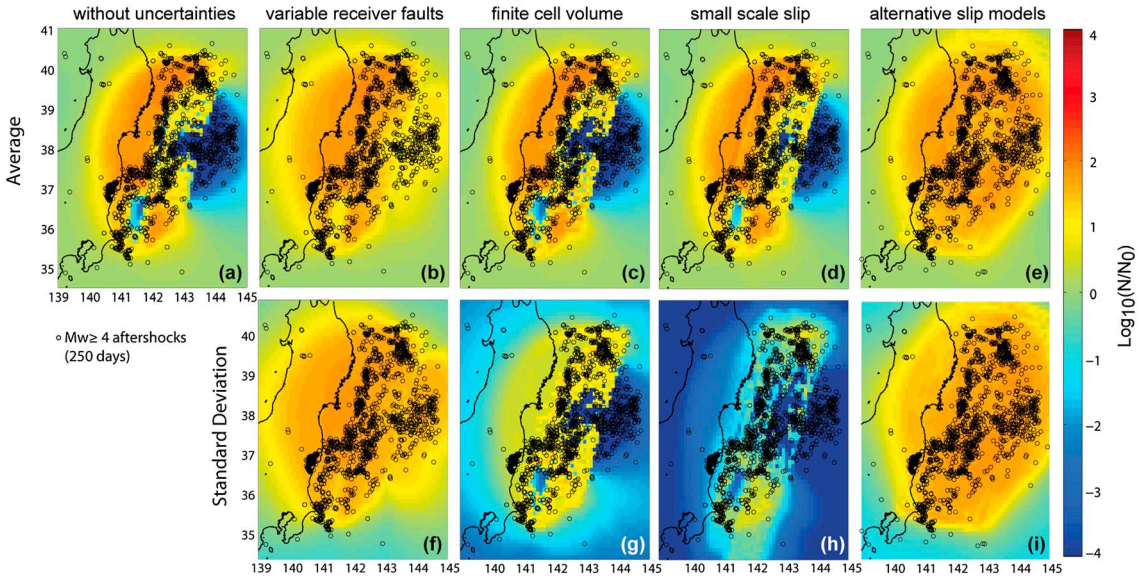

Figure 5. Maps of forecasted seismicity for Tohoku. The color scale indicates the number of events forecasted between 5 and 250 days (N), divided by the background number of events (N_0), on a logarithmic scale. (a) Uncertainties not included; mean forecast obtained including uncertainties from (b) receiver fault orientation, (c) finite cell volume, (d) small-scale slip in the slip model, and (e) different published slip models. (f–i) Standard deviation corresponding to the figures above, calculated at each grid point from the set of Monte Carlo iterations.

We find that the spatial distribution is also affected most significantly by the choice of receiver fault (Figures 5–7). In particular, in the average map, the seismicity lows in the area of coseismic slip disappear for both main shocks (Figures 5b and 7b). The variation of the forecast across iterations is large not only in the vicinity of the fault but also in the far field. For Tohoku, in the offshore region at lat $\sim$ 38°, lon $\sim$ 144° presents a large standard deviation (Figure 5f), in an area which is quiescent when uncertainties are not included (Figure 5a). This indicates that faults exist with orientations on which seismicity is promoted, as confirmed by the observed aftershocks in these areas.

When considering the finite cell size, a smoother forecast is obtained (Figures 5c and 7c); moreover, the on-fault area of reduced seismicity has a smaller size compared to the case in which uncertainties are

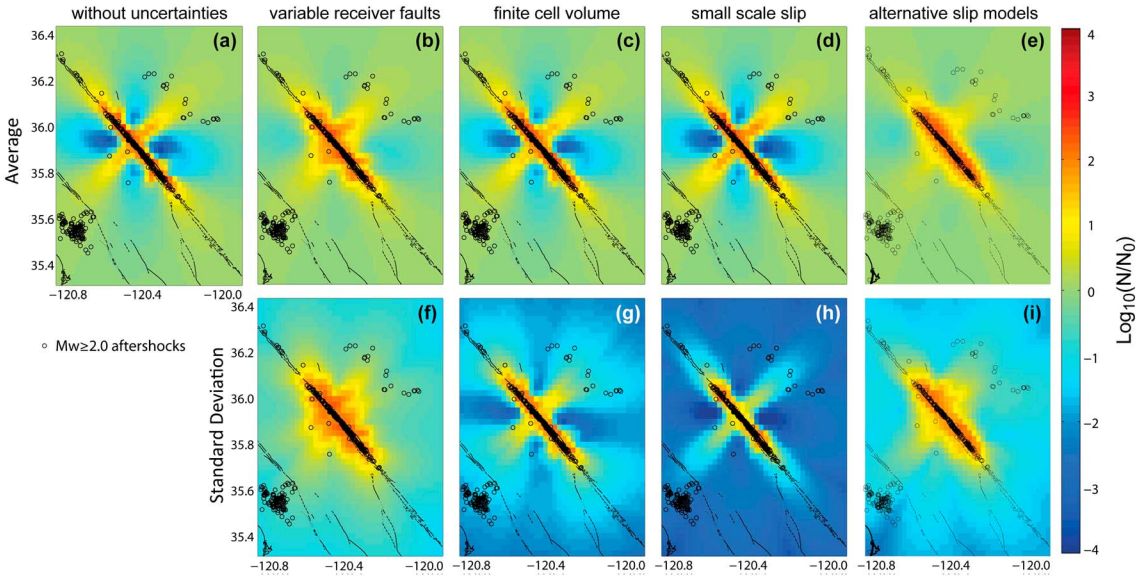

Figure 6. Maps of forecasted seismicity for Parkfield. Individual maps correspond to those in Figure 5.

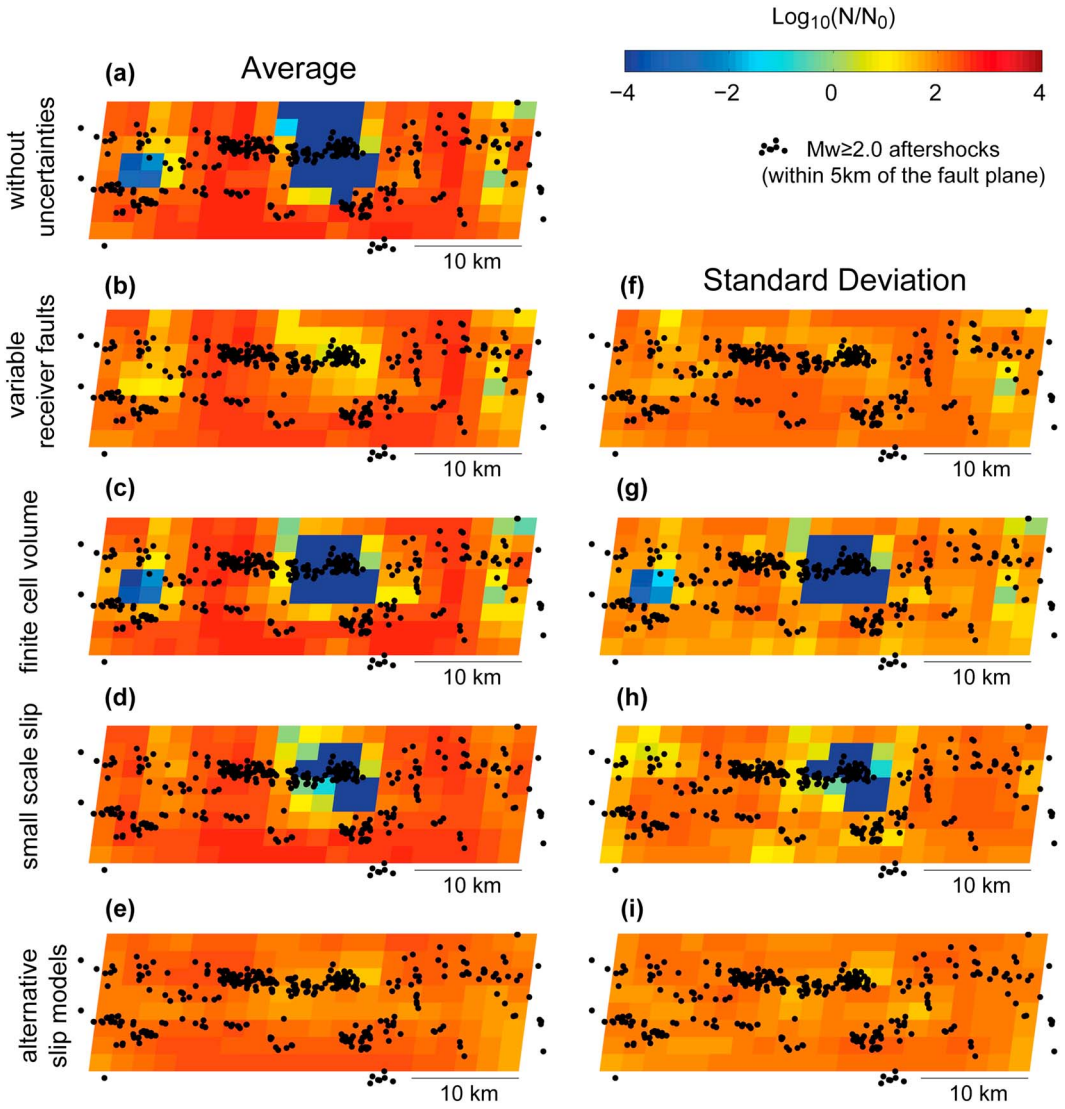

Figure 7. Forecasted seismicity for Parkfield, projected on the main shock fault plane; patches correspond to those in the slip model. Individual plots correspond to those in Figures 5 and 6.

negligible. This results indicate that sharp boundaries between areas of increased and decreased seismicity is an artifact due to coarse grid size with respect to the near-field gradients and that the area of stress shadows can be overestimated due to discrete sampling of the stress field.

Comparison of spatial forecasts obtained with and without slip model uncertainty (Figures 5d, 5a, 6d, and 6a) indicates that the impact of small-scale slip is very small. For both main shocks, a significant decrease in seismicity (stress shadow) is observed on the main shock fault; the standard deviation is also small in this area, indicating that this feature is consistent across most of the Monte Carlo iterations obtained from different perturbed slip models. This results indicates that slip heterogeneity, as simulated in our model, is not sufficient to induce seismicity close to the largest patch of coseismic slip on fixed receiver faults. On the other hand, the differences between published slip models are significant: we find that the standard deviation in the expected number of events across models obtained from different published slip distributions is similar to the absolute value of their average (Figures 5e, 5i, 7e, and 7i).

The two main shocks significantly differ in terms of size and complexity; however, we find that the same sources of uncertainties among those analyzed in this work are most significant in both cases. We see a particularly large variability due to receiver fault orientation for the Tohoku sequence, with the total number

of events varying by a factor of $\sim$ 4.5; this result reflects the higher structural complexity of a subduction zone, in which a wider range of receiver fault orientations is present. The differences due to the choice of slip model are also larger for Tohoku, but this result may simply be due to the larger number of slip models compared. The effect of finite cell volume depends on the grid resolution compared to the scale of the stress field: this aspect is slightly more important for Parkfield (Figures 4b and 4f), due to the fact that the grid used for Tohoku produced a smoother forecast than the one used for Parkfield (Figures 5 and 6).

4.1. Combining Models

Having analyzed the distribution of forecasts arising from different factors in the input data, the following question arises: how to combine the forecasts from each Monte Carlo iteration into a single forecast? To handle variability in the input data correctly, it is important to draw a distinction between the physical interpretation of different types of uncertainties. With subgrid variability, we refer to the spatial heterogeneity of physical properties in a crustal volume, for example, the presence of faults with different orientations or inhomogeneous elastic properties. These spatial variations, which can be interpreted as aleatoric uncertainties, cause a real variability in Coulomb stress values, which gives rise to physically observable characteristics of seismic sequences: an example of such phenomena is the observation of increased seismicity rates in"stress shadows," i.e., areas with negative average ΔCFS. In this context, each Monte Carlo iteration can be viewed as representative of a small cell subvolume, the exact location of which cannot be resolved. With these assumptions, the resulting forecast should not be considered as the average of the several possible and alternative forecasts but instead as a simulation of the behavior emerging from real heterogeneity within each grid cell. We consider the sampling of a set of existing receiver faults, as well as the consideration of heterogeneity due to finite cell volume, as ways to reproduce this physical heterogeneity.

The use of a set of possible slip models is an example of epistemic uncertainties, and it requires a different interpretation: the forecasts obtained in this case are alternative to each others, and their average does not have a clear physical meaning. The distribution of such forecasts, on the other hand, can be used to estimate the sensitivity of the model to the input slip model and therefore to provide an estimation of the final uncertainty of the forecast. Additionally, ensemble models can be constructed by combining models based on alternative assumptions or input data; and while such procedure may not have an obvious physical interpretation, recent studies [*Marzocchi et al.*, 2012] have proven it to be a valuable tool from the point of view of operational earthquake forecasting. In their study, the authors use Bayesian averaging to combine models from the 5 year Regional Earthquake Likelihood Models for California: these are statistical models submitted by different authors and based on different modeling strategies. Here we test whether the same approach can also be effective to combine forecasts from alternative input data within a single model class.

We follow the approach of *Marzocchi et al.* [2012] in creating an ensemble model based on a weighted average of forecasts from all available slip models. We weigh models using the Score Model Averaging method [*Marzocchi et al.*, 2012]:

$$R_{\text{ens}} = \sum_{j}^{J} w_j R_j \tag{5}$$

with weights given by

$$w_j = \frac{\delta_j^{\text{corr}} S_j}{\sum_k^J \left[\delta_k^{\text{corr}} S_k\right]} \tag{6}$$

where δ_j^{corr} is a factor accounting for the correlation between models and S_j is given by $S_j = 1/|\text{LL} - \text{LL}^*|$, where LL is the model LogLikelihood, and LL* is the LogLikelihood of a reference model, in this case a uniform Poisson model with seismicity rate given by the background rate. For more details about the method, see *Marzocchi et al.* [2012]. Due to the logarithmic scaling with Likelihood, this weighting scheme tends not to favor the best performing model as strongly as other schemes. In light of the fact that small-scale slip variability plays a secondary role compared to the differences between published slip models, we do not include this aspect, for computational efficiency: the individual models in our ensemble are all obtained from the original slip model, without addition of small-scale slip.

Table 2. Comparison of Rate and State Parameters and Model Performance[a]

Model	$A\sigma$ (kPa)	t_a (d)	ΔLL	N (N_{obs} = 2266)
		Tohoku		
Without uncertainties	100	19,500	0	547
With uncertainties	18	18,000	5.65	2,547
		Parkfield		
	$A\sigma$ (kPa)	t_a (d)	ΔLL	N (N_{obs} = 674)
Without uncertainties	8	10,000	0	254
With uncertainties	4	10,000	14.43	526

[a]The last column is the total number of forecasted aftershocks and N_{obs} is the observed number.

5. Model Performance

We quantify the performance of various models: first, we assess the effect of including aleatoric uncertainties (due to finite cell size and variable receiver fault orientations); second, we present a comparison of the performance of models obtained from the slip models listed in Table 1; finally, we study an ensemble model obtained combining the forecasts from alternative slip models.

Model performance can be evaluated by using the LogLikelihood, which is defined as the probability of the observed outcome (the observed earthquake catalog) given a specified model; given equal a priori

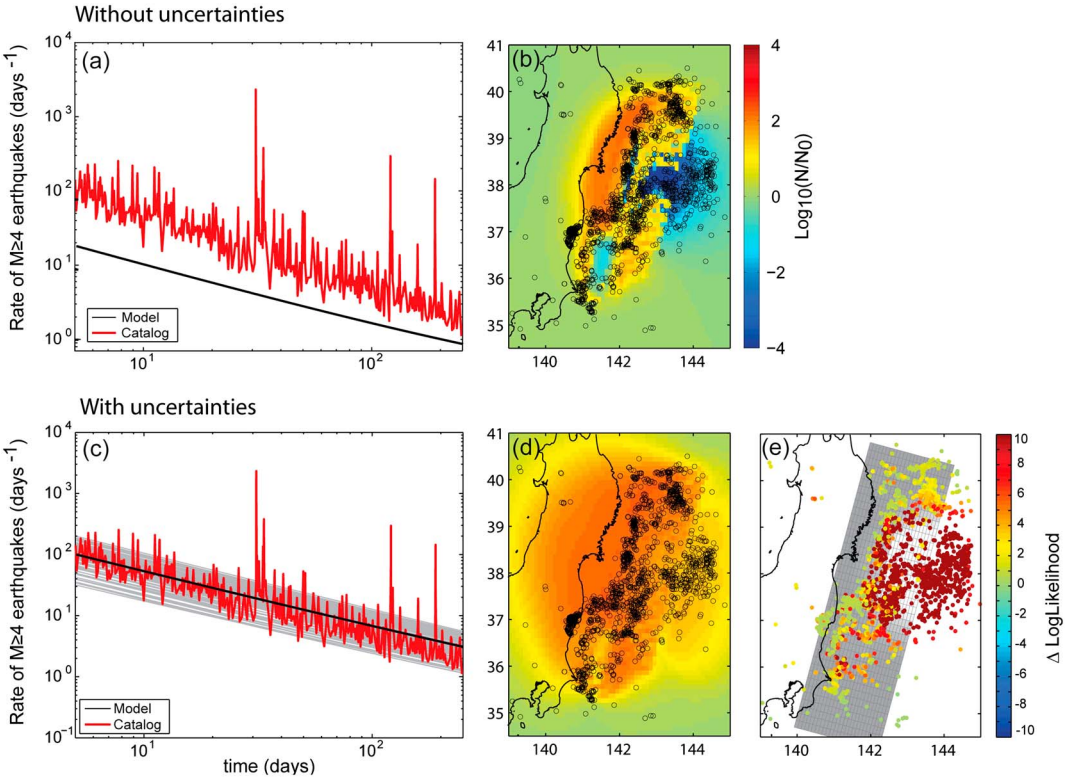

Figure 8. Comparison of the modeled seismicity (c, d) with and (a, b) without aleatoric uncertainty, for Tohoku. Figures 8a and 8c show temporal decay of seismicity; red line: observed decay, black line: modeled seismicity, and grey lines: individual Monte Carlo iterations. Figures 8b and 8d show spatial distribution of modeled number of events divided by background number of events in 250 days, on a logarithmic scale. (e) Map of individual aftershocks, color coded by change in LogLikelihood when aleatoric uncertainties are included: the large majority of events is better modeled by including uncertainties. In the background, in black and white: the slip model (saturated at slip = 20 m).

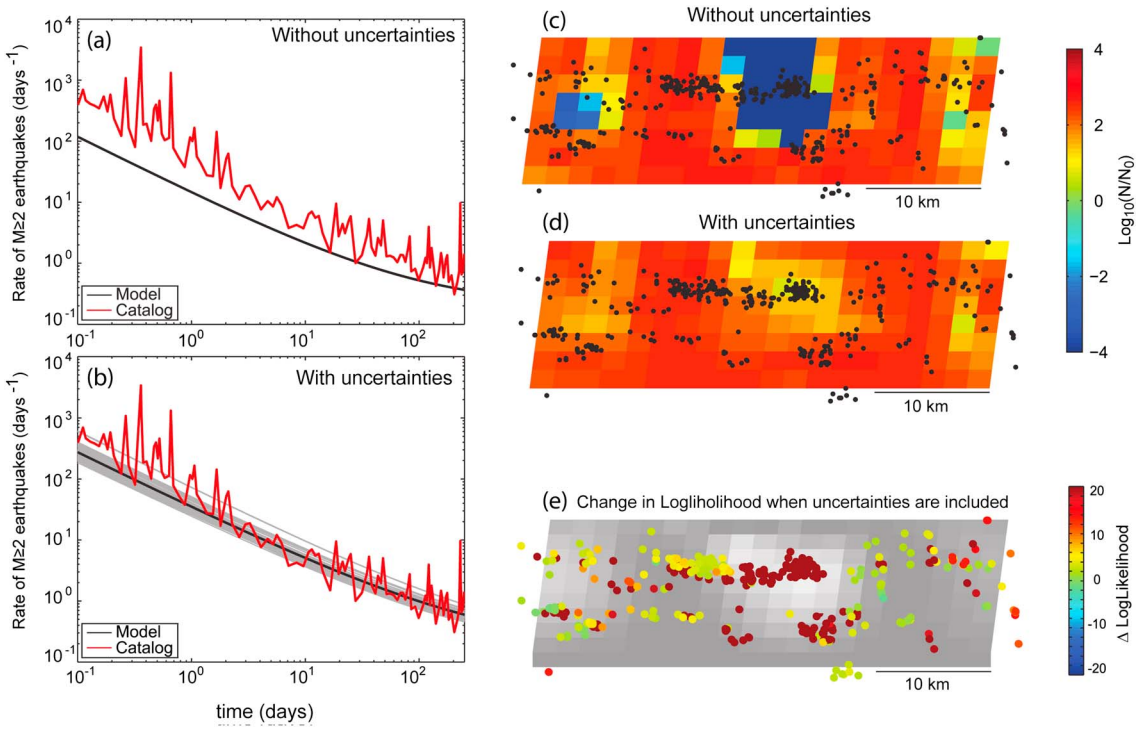

Figure 9. Comparison of the modeled seismicity (b, d) with and (a, c) without aleatoric uncertainty, for Parkfield. Lines and colors scales correspond to those in Figure 8. (e) Also in this case, most of the events are better modeled when aleatoric uncertainties are included; comparison with the underlying slip model (saturated at slip = 0.3 m) shows that in particular, the events occurring on the rupture area are better explained by including stress heterogeneity.

probability, Bayes theorem implies that models with higher LogLikelihood are more likely to be correct [*Ogata*, 1983]. For a stationary Poisson process, the LogLikelihood is defined as

$$\mathrm{LL} = \sum_{j}^{N} \log\left[R(\mathbf{x_j}, t_j)\right] - \int_{t_0}^{t_1} \int_{\mathrm{volume}} R(\mathbf{x}, t)\, \mathrm{d}\mathbf{x}\mathrm{d}t \tag{7}$$

where R is the rate at the time t_j and locations $\mathbf{x_j}$ of the observed earthquakes (with index j) and t_0 and t_1 are the start time and end time of the forecast. The values of R_j depend on the exact location of the observed events; in order to account for the uncertainty in earthquake location, we smooth the position over a set of grid points, assuming a Gaussian distribution with standard deviation in the horizontal and vertical position as given in the catalog. We compare models against each other using the change in Log-Likelihood per event, given by $\Delta\mathrm{LL} = (\mathrm{LL} - \mathrm{LL}_0)/N$, where LL_0 is the lowest LogLikelihood among the models tested.

5.1. Performance of Models Including Stress Heterogeneity

We compare the LogLikelihood of a model which includes stress variability to a simple model which assumes optimally oriented planes and does not account for finite volume of the cells; both models use the same slip distribution [*Wang et al.*, 2012, 2013], without the addition of synthetic slip. We invert for rate-and-state parameters using the entire period of the forecast (250 days). As previously observed by *Hainzl et al.* [2009], we find that overall model performance improved significantly when we account for the heterogeneity of the stress field (Table 2): the change in LogLikelihood per event is 5.65 for Tohoku and 14.43 for Parkfield. We find that models which consider aleatoric uncertainty provide a better description of both the spatial and temporal distribution of seismicity (Figures 8 and 9). Moreover, these models better estimate the total number of events for both sequences (Table 2 and Figures 8 and 9).

5.2. Performance of the Available Slip Models

We evaluate the ability of published slip models to reproduce the observed seismicity. We perform two types of tests: (1) nonprospective mode, we allow the model to find the best set of parameters to fit the data in the entire modeled period: the inversion of model parameters is done by maximizing the LogLikelihood

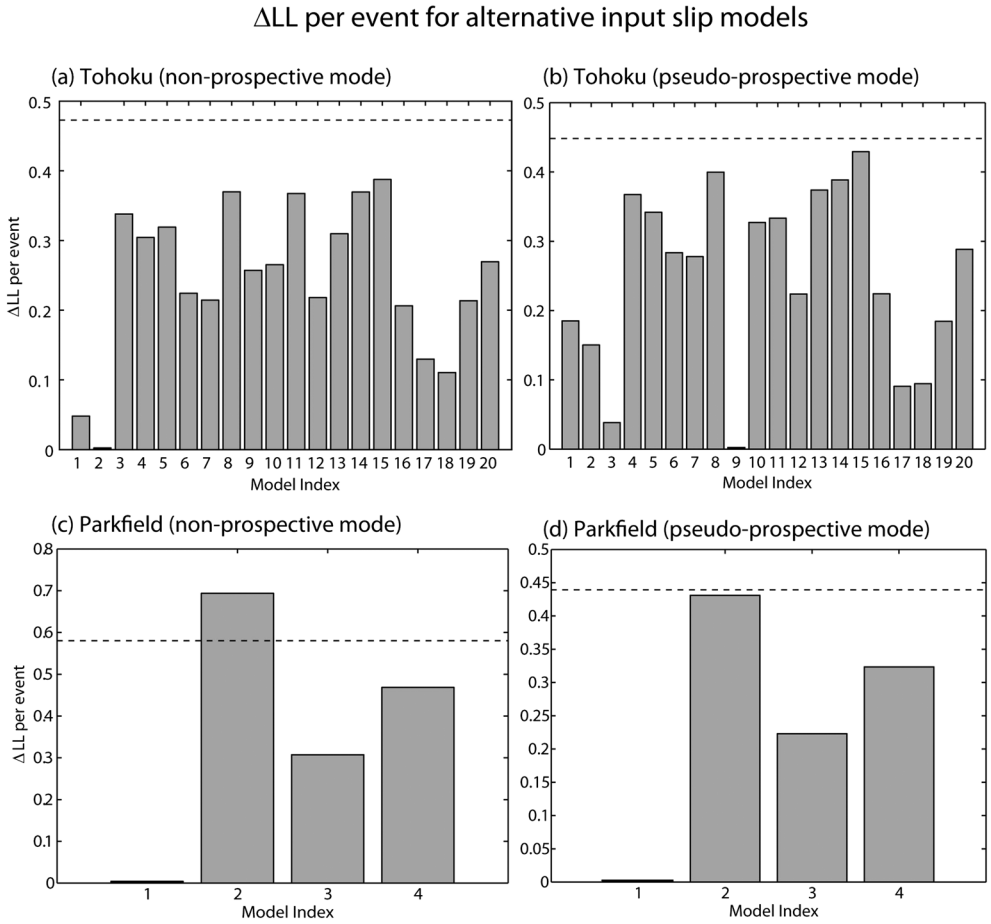

Figure 10. Comparison of ΔLL for Tohoku, in (a) nonprospective mode and (b) pseudo-prospective mode. Comparison of ΔLL for Parkfield: (c) nonprospective mode and (d) pseudo-prospective mode. Each histogram corresponds to a slip model, in the order in which they are listed in Table 1. The dotted line indicates the performance of the ensemble model.

for the test period of 250 days. (2) Pseudo-prospective mode, we use the first 5 days of complete catalog to fit rate-and-state parameters and model the seismicity for the remaining time. The goal of the first tests is to compare the performance across slip models in the ideal case in which the best parameters are known; the second test, on the other hand, reproduces the condition of a forecast scenario. For each model, we include aleatoric uncertainties from receiver fault and grid size and use the average value.

Model performance results are summarized in Figure 10. We do not find a clear correlation between slip model complexity (as indicated by number of fault segments or number of patches) and model performance. We find that for Tohoku, both in the nonprospective and the pseudo-prospective case, the ensemble model has a higher LogLikelihood than any of the models in the sample. For Parkfield, we find that the ensemble model outperforms all the individual models in the pseudo-prospective test; only one model outperforms the ensemble model in the nonprospective test.

6. Discussion

We performed a sensitivity analysis of CRS models to identify the most important sources of uncertainties in Coulomb stress calculations; furthermore, we suggested practical methods to include these uncertainties in the model, by considering their physical origin and drawing a distinction between aleatoric and epistemic sources of errors.

We find that various uncertainties in Coulomb stress have a profound impact on CRS models; in particular, the orientation of receiver faults and the choice of input slip model have a predominant effect.

The heterogeneity of stress field due to the presence of multiple receiver fault orientation plays a first-order role in both spatial distribution and temporal evolution of the forecast. This is not surprising considering that resolving the stress tensor on different planes can give rise to very different stress fields, both in the near field and in the far field (see supporting information (D)); on the other hand, the perturbations in the stress fields due to finite cell size and by small-scale slip act on a smaller scale and are overall less significant.

We find that some important features of the forecast, such as the existence of stress shadows, are altered if we consider seismicity as emerging from the contribution of faults at different orientations which respond independently to the stress field. As previously suggested by various authors [*Marsan*, 2006; *Helmstetter and Shaw*, 2006], variability of the stress field in the context of rate-and-state models can provide an explanation for the observation of seismicity in regions of average negative stress change. Our results indicate that the complexity of fault systems, with faults of different orientations coexisting in a crustal volume, can generate enough stress heterogeneity to explain these observations, and they are in agreement with the results from *Toda and Enescu* [2011] and *Enescu et al.* [2012], who noticed that receiver fault variability plays an important role for the Tohoku aftershock sequence. The suppression of stress shadows also accounts for the better performance of the models in which aleatoric uncertainty is taken into account: in such models seismicity does not shut down completely in the area of the coseismic rupture, or other areas of negative average stress change, and seismicity observed at these locations is better explained. Due to definition of LogLikelihood (which becomes singular if an event is observed in an area where the seismicity rate is predicted to be exactly 0), this reduction of stress shadows has a dramatic effect on the results of LogLikelihood-based tests.

We also observe that models which do not include aleatoric uncertainty underestimate the total number of events by a factor of 3–4, while better estimations are obtained when stress variability is taken into account. This difference can be understood from the values of $A\sigma$, the frictional resistance parameter which determines how strongly seismicity is enhanced by positive ΔCFS or inhibited by negative ΔCFS. The parameter search yields higher values of $A\sigma$ for models which do not include uncertainties (Table 2). This result is due to the presence of stress shadows: since the LogLikelihood heavily penalizes models with areas of almost complete seismicity shutdown, high values of $A\sigma$ will be favored so that stress shadows are less effective at inhibiting seismicity. At the same time, higher values of $A\sigma$ cause positive stress changes to produce fewer earthquakes, leading to the underestimation of total the number of aftershocks.

These findings have important implications for physics-based operational forecasting models. The most common choices of receiver fault orientation (the use of a fixed receiver or of optimally oriented planes) vastly underestimate the variability of Coulomb stress, and our sensitivity study shows that these assumptions have a dramatic effect in forecasted distribution of seismicity. Another approach consists of including information about the local geology by use of a gridded fault model [*Steacy et al.*, 2005; *Segou et al.*, 2013; *Toda and Enescu*, 2011]; however, this method also assumes a single fault orientation at each grid point and therefore neglects the aleatoric uncertainty associated with receiver faults. The performance comparison indicates that considering this aspect leads to significantly higher LogLikelihood scores.

Woessner et al. [2011] compared the performance of statistical and physical models for the 1992 Landers sequence, and they also found that physics-based models perform significantly better when they include a stochastic component; however, they observed that CRS models including uncertainties exhibited an excessive increase in seismicity in areas of large, negative stress changes. This behavior is due to the assumption that the standard deviation of ΔCFS at each grid point is directly proportional to its absolute value [*Hainzl et al.*, 2009]: since the distribution of ΔCFS values is wide in areas of large and negative stress, large and positive values are likely to be sampled in the Monte Carlo iterations, and they dominate early seismicity. On the other hand, stress variability is most likely underestimated in regions where the average stress field is close to 0. The models presented here do not make a priori assumptions on the distribution of ΔCFS: this approach is physically more consistent, and it does not lead to increased seismicity in stress shadows (Figure 6).

In CRS models, a distinction between the treatment of aleatoric and epistemic uncertainties has not previously been made. While the methods used in the two cases are similar (averaging in the first case, weighted averaging in the second), the difference should be kept in mind. When dealing with aleatoric uncertainties, the focus should be on realistically reproducing the physical stress variability; on the other hand, the combination of models with epistemic uncertainties is guided by an attempt to obtain the best model performance, as done in the ensemble models.

We find that the model ensemble generally outperforms all individual models. In particular, the results of the pseudo-prospective experiment (Figure 10) indicate that ensemble models are promising in the context of operational earthquake forecasting: without knowing a priori which model will perform best, the ensemble model allows to obtain a better performance than each model, including the best.

The LogLikelihood of a forecast can be very sensitive to the occurrence of events in areas with very low forecasted seismicity: this implies that the better performance of ensemble models may be due to fact that by overlaying a set of alternative forecast maps, both areas of high and low seismicity are reduced. However, by comparing the Likelihood of individual events, we verify that the improvement in LogLikelihood is not driven by a small number of events located in stress shadows.

In this work we attempt to address some of the limitations of CRS models and to better describe the spatial complexity of the stress field. However, a number of simplifications still exist in the model. We fitted rate-and-state parameters $A\sigma$, t_a to the data by maximizing the LogLikelihood and therefore did not consider them as a source or error; however, in a forecasting setting the values of these parameters are not constrained a priori. In Appendix A, we estimate the sensitivity of the model to these parameters and also test the effect of the coefficient of friction μ. We find that in the time domain, the uncertainty due to unknown parameters $A\sigma$ and t_a is comparable to the variability of receiver fault and to the input slip model; however, the spatial variability introduced by these parameters is minor compared to the effect of uncertainties in the stress calculations.

Our approach to estimating fault orientation is arguably rather crude. We performed some spatial selection for the Tohoku area, in order to account for different faulting style being predominant in each of the tectonic areas (wedge, fault interface, and outer rise). However, we did not make an attempt to describe smaller-scale spatial variations of fault structures, which could, for example, be achieved by weighting the focal planes from the catalog by distance to each grid point or by complementing the information from focal mechanisms with that from mapped faults [*Steacy et al.*, 2005; *Segou et al.*, 2013]. Focal planes in a volume of few kilometers (i.e., within a grid cell) are more homogeneous than on the entire scale of our domain [*Hardebeck*, 2006], and therefore, the variability of ΔCFS may have been overestimated in our study.

We find that considering small-scale variation of main shock slip does not affect the distribution of the forecast significantly; in particular, the effect is much smaller than the variability due to the differences between published slip models. The effect of small-scale slip variability may have been underestimated for two reasons: (1) the grid size we used limits our definition of near field, and it may be too large to be affected by slip below a similar length and (2) the heterogeneity in fault geometry, or fault roughness, has not been included: the fault was always treated as a uniform plane. However, this aspect should also be taken into account: our treatment of slip heterogeneity could be extended to include fault roughness by creating synthetic slip models. Our results indicate that the orientation of receiver planes plays a first-order role, also in the vicinity of the fault: by symmetry, this suggests that including variability in the orientation of the rupturing faults may also enhance stress heterogeneity.

Another simplification of the model is the assumption that faults are uniformly distributed, while in reality they tend to cluster around major fault structures: the model could be improved by including the major existing faults on which seismicity concentrates. This could be achieved, within the rate-and-state framework, by estimating a spatially variable background rate, based on smoothed seismicity. However, such background rate calculation may suffer from large uncertainties due to limited sample size and is very sensitive to the spatial pattern of Coulomb stress [*Cocco et al.*, 2010]. Another simplification in our model is the choice of spatially uniform rate-and-state parameters: while a simple 1-D depth dependence could be introduced, the consideration of the full 3-D spatially varying parameters would require a large number of parameters and lead to potentially unstable results. Finally, we point out that the model has intrinsic limitations due to the underlying physical assumptions. Various physical processes are neglected, including secondary triggering, afterslip, viscoelastic relaxation, and redistribution of stresses by fluids: all these processes can modify the stress field postseismically, while we assume a stationary field after each earthquake. We verified that the improvement in model performance obtained by including aleatoric uncertainties is a robust feature that persists when afterslip is also taken into account; a more detailed treatment of the role of time-dependent processes, which are particularly large following the Parkfield earthquake, will be addressed in a future investigation. The existence of afterslip should especially be kept in mind when evaluating the performance of different published slip models: since some of the earthquakes may be triggered

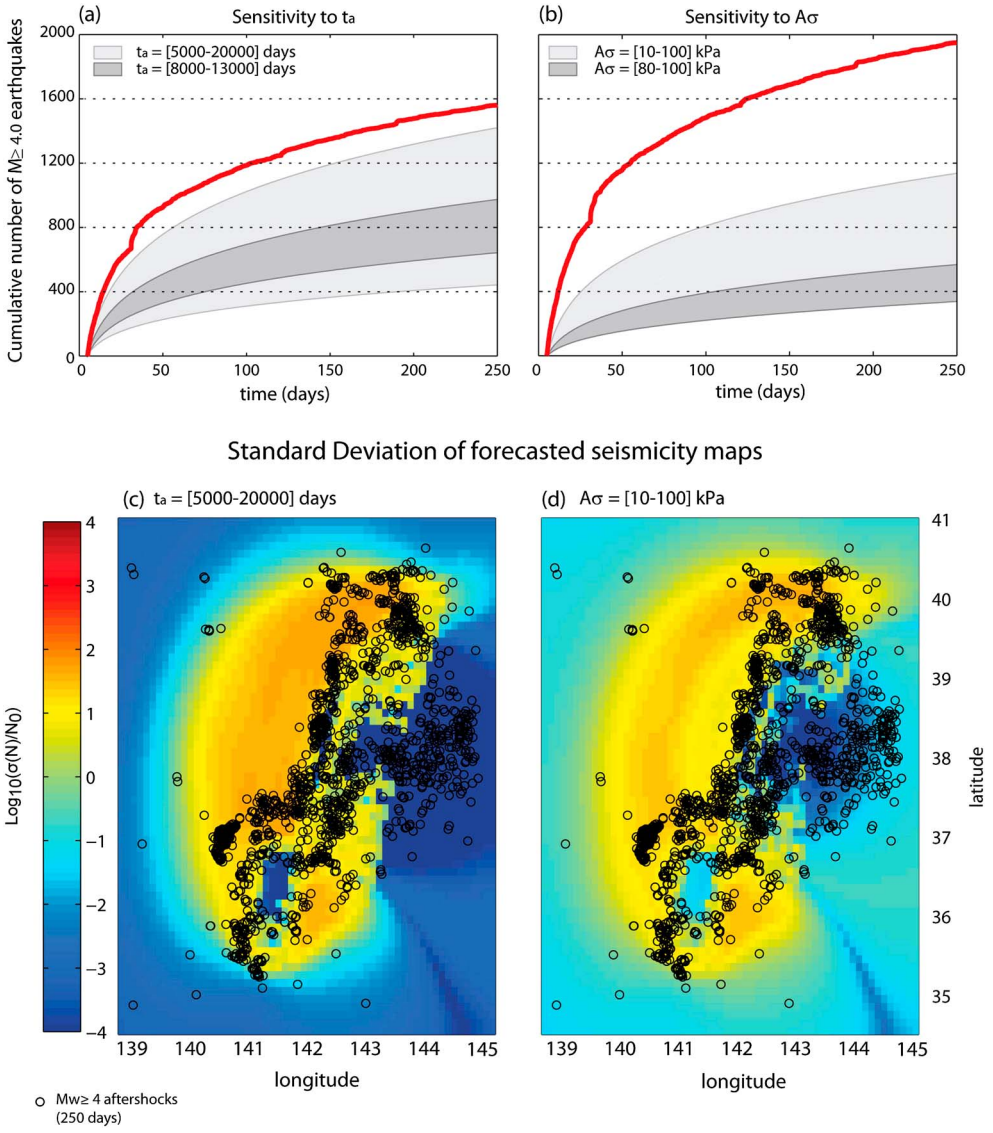

Figure A1. Sensitivity of the forecast to rate-and-state parameters $A\sigma$ and t_a, for the Tohoku sequence. (a) Temporal evolution of seismicity for typical values of t_a and $A\sigma = 28$ kPa. (b) Temporal evolution with variable $A\sigma$ and $t_a = 8000$ d. The red line is the observed cumulative number of events. (c and d) Standard deviation of forecasted number of events ($\sigma(N)$), divided by the background seismicity, on a log scale. t_a and $A\sigma$ are in the same range as in Figures A1a and A1b; t_a was varied in intervals of 500 d and $A\sigma$ in intervals of 10 kPa. In all cases, stresses were resolved on a fixed plane given by the average main shock focal plane; other sources of uncertainty were not included.

by postseismic stresses, the ranking of the slip models in their ability to explain observed seismicity may not reflect how accurately they describe coseismic slip.

7. Conclusions

In this work, we study the impact of different physical sources of uncertainties on CRS models of aftershock sequences, and we introduce a consistent framework to propagate epistemic and aleatoric uncertainties through such models.

We find that variability in ΔCFS gives rise to a large spread in the forecast, both in terms of spatial location and total number of events; in particular, receiver fault orientation plays a first-order role. By incorporating the stress heterogeneity due to variable receiver faults, we obtain a significant performance improvement. These results indicate that a careful consideration of receiver fault orientation, combined with a

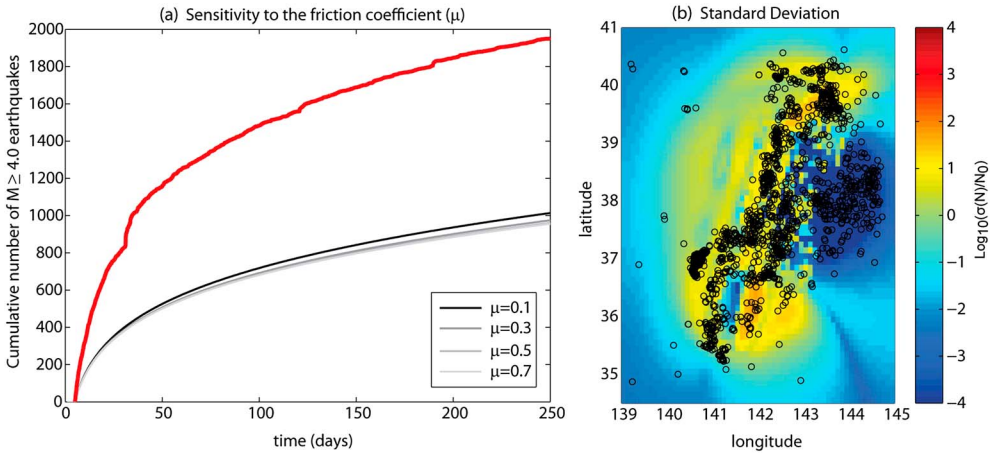

Figure A2. Sensitivity of the forecast to the coefficient of friction μ, for the Tohoku sequence. (a) Temporal evolution of seismicity with μ between 0.1 and 0.9. The red lines is the observed seismicity. (b) Standard deviation of forecasted number of events ($\sigma(N)$), divided by the background seismicity, on a log scale.

stochastic treatment of its variability, is an important aspect of CRS models and should be implemented in future applications.

We find that the choice of slip model is also an important source of uncertainty, and we explore the possibility of using ensemble averaging to combine models based on alternative slip distributions. This technique has previously been used to combine aftershock models based on entirely different principles [*Marzocchi et al.*, 2012]; here we suggest that the same methods can be used to address epistemic uncertainties within a single model class. In agreement with *Marzocchi et al.* [2012], we find that also in this case ensemble models systematically outperform the individual models, and they are a promising tool in the context of operational earthquake forecasting.

Appendix A: Model Sensitivity to $A\sigma$, t_a, and μ

The main focus of this study is the sensitivity of CRS models to uncertainties in the input data and model discretization, while the model parameters $A\sigma$ and t_a are optimized by maximum likelihood estimation. However, in the case of prospective forecast, the uncertainty of these parameters will matter. A detailed discussion of the sensitivity of CRS models to the background rate, and the rate-and-state parameters $A\sigma$ and t_a, can be found in *Cocco et al.* [2010]. In order to estimate the role of model parameters compared to the uncertainties discussed above, we estimate the spread of the forecast for the Tohoku sequence, due to three parameters: $A\sigma$, t_a, and the friction coefficient μ.

Typical values of the frictional resistance parameter $A\sigma$ are between 10 and 100 kPa [*Hainzl et al.*, 2010]. The parameter t_a, the aftershock duration time, is typically of the order of tens of years; here we used values between 5000 and 20,000 days (13.7–54.8 years). Figures A1a and A1b show that variations of $A\sigma$ and t_a within these ranges lead to significant differences in forecasted number of events, comparable to the variation due to the different published slip models or to the variability from receiver fault uncertainty (Figure 4). These values give a measure of the variability of the forecasts when $A\sigma$ and t_a are completely unknown; on the other hand, these values can be better constrained during the aftershock sequence. Considering a narrower range of $A\sigma$ between 80 and 100 kPa and t_a between 8000 and 13,000 days, the sensitivity of the model to the parameters has a smaller impact than the receiver fault and the differences between published slip models. In the spatial domain, we find that the variability due to these parameters is modest compared to the effect of the variable receiver fault and to the choice of slip model (comparison between Figures A1c, A1d, and 5); this can be expected since these parameters only affect the seismic productivity at each grid point, but they do not modify the spatial features of the stress field.

We also tested the effect of the coefficient of friction μ. Figure A2 shows that for μ between 0.1 and 0.9, the impact of this parameter on the model is comparable to the variation due to small-scale slip or to finite grid resolution but modest compared to the most significant sources of uncertainty (Figures 4 and 5).

Acknowledgments
We thank the Editor and two anonymous reviewers for their helpful comments. This research was carried out in the framework of REAKT Project (strategies and tools for Real-Time EArthquake RisK ReducTion) founded by the European Community via the Seventh Framework Program for Research (FP7), contract 282862. C.C. is a member of the Helmholtz graduate research school GeoSim. W.L. is partially supported by NSFC-41204064. We are grateful to the Japan Meterological Agency for providing the catalog data used in this study. The data used for this study are available from the following websites: ANSS catalog (determined by Northern California Seismic Network, USGS, Menlo Park and Berkeley Seismological Laboratory, University of California, Berkeley), http://quake.geo.berkeley.edu/anss/; focal mechanisms catalog for Southern California http://www.data.scec.org/research-tools/alt-2011-yang-hauksson-shearer.html; focal mechanisms catalog for Japan (*F* net moment tensor focal mechanisms solutions), http://www.fnet.bosai.go.jp/event/search.php; Slipmodelsdatabase (maintained by ETH Zürich), and http://equake-rc.info/SRCMOD.

References

Asano, Y., T. Saito, Y. Ito, K. Shiomi, and H. Hirose (2011), Spatial distribution and focal mechanisms of aftershocks of the 2011 off the Pacific coast of Tohoku Earthquake, *Earth Planets Space*, *63*(7), 669–673, doi:10.5047/eps.2011.06.016.

Beresnev, I. (2003), Uncertainties in finite-fault slip inversions: To what extent to believe? (a critical review), *Bull. Seismol. Soc. Am.*, *93*(6), 2445–2458, doi:10.1785/0120020225.

Byerlee, J. (1978), Friction of rocks, *Pure Appl. Geophys.*, *116*(4–5), 615–626, doi:10.1007/BF00876528.

Causse, M., F. Cotton, and P. M. Mai (2010), Constraining the roughness degree of slip heterogeneity, *J. Geophys Res.*, *115*, B05304, doi:10.1029/2009JB006747.

Cocco, M., and J. R. Rice (2002), Pore pressure and poroelasticity effects in Coulomb stress analysis of earthquake interactions, *J. Geophys Res.*, *107*(B2), ESE 2-1–ESE 2-17, doi:10.1029/2000JB000138.

Cocco, M., S. Hainzl, F. Catalli, B. Enescu, A. M. Lombardi, and J. Woessner (2010), Sensitivity study of forecasted aftershock seismicity based on Coulomb stress calculation and rate- and state-dependent frictional response, *J. Geophys Res.*, *115*, B05307, doi:10.1029/2009JB006838.

Dieterich, J. (1994), A constitutive law for rate of earthquake its application to earthquake clustering, *J. Geophys Res.*, *99*, 2601–2618.

Enescu, B., S. Aoi, S. Toda, W. Suzuki, K. Obara, K. Shiomi, and T. Takeda (2012), Stress perturbations and seismic response associated with the 2011 M9.0 Tohoku-oki earthquake in and around the Tokai seismic gap, central Japan, *Geophys. Res. Lett.*, *39*, L00G28, doi:10.1029/2012GL051839.

Hainzl, S., B. Enescu, M. Cocco, J. Woessner, F. Catalli, R. Wang, and F. Roth (2009), Aftershock modeling based on uncertain stress calculations, *J. Geophys Res.*, *114*, B05309, doi:10.1029/2008JB006011.

Hainzl, S., G. Zöller, and R. Wang (2010), Impact of the receiver fault distribution on aftershock activity, *J. Geophys Res.*, *115*, B05315, doi:10.1029/2008JB006224.

Hardebeck, J. L. (2006), Homogeneity of small-scale earthquake faulting, stress, and fault strength, *Bull. Seismol. Soc. Am.*, *96*(5), 1675–1688, doi:10.1785/0120050257.

Harris, A. (1998), Introduction to special section: Stress triggers, stress shadows, and implications for seismic hazard, *J. Geophys Res.*, *103*(B10), 24,347–24,358.

Hartzell, S., P. Liu, C. Mendoza, C. Ji, and K. M. Larson (2007), Stability and uncertainty of finite-fault slip inversions: Application to the 2004 Parkfield, California, earthquake, *Bull. Seismol. Soc. Am.*, *97*(6), 1911–1934, doi:10.1785/0120070080.

Hasegawa, A., K. Yoshida, and T. Okada (2011), Nearly complete stress drop in the 2011 Mw 9.0 off the Pacific coast of Tohoku Earthquake, *Earth Planets Space*, *63*(7), 703–707, doi:10.5047/eps.2011.06.007.

Helmstetter, A., and B. E. Shaw (2006), Relation between stress heterogeneity and aftershock rate in the rate-and-state model, *J. Geophys Res.*, *111*, B07304, doi:10.1029/2005JB004077.

Kieling, K., R. Wang, and S. Hainzl (2014), Broadband ground motion simulation using energy constrained rise time scaling, *Bull. Seismol. Soc. Am.*, in press.

King, G. C. P., R. S. Stein, and J. Lin (1994), Static stress changes and the triggering of earthquakes, *Bull. Seismol. Soc. Am.*, *84*(3), 935–953.

Linker, M. F., and J. H. Dieterich (1992), Effects of variable normal stress on rock friction: Observations and constitutive equations, *J. Geophys Res.*, *97*(92), 4923–4940.

Marsan, D. (2006), Can coseismic stress variability suppress seismicity shadows? Insights from a rate-and-state friction model, *J. Geophys Res.*, *111*, B06305, doi:10.1029/2005JB004060.

Marzocchi, W., J. D. Zechar, and T. H. Jordan (2012), Bayesian forecast evaluation and ensemble earthquake forecasting, *Bull. Seismol. Soc. Am.*, *102*(6), 2574–2584, doi:10.1785/0120110327.

Minson, S. E., M. Simons, and J. L. Beck (2013), Bayesian inversion for finite fault earthquake source models. I—Theory and algorithm, *Geophys. J. Int.*, *194*(3), 1701–1726, doi:10.1093/gji/ggt180.

Ogata, Y. (1983), Estimation of the parameters in the modified Omori formula for aftershock frequencies by the maximum likelihood procedure, *J. Phys. Earth*, *31*, 115–124.

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

Okada, Y., K. Kasahara, S. Hori, K. Obara, S. Sekiguchi, H. Fujiwara, and A. Yamamoto (2004), Recent progress of seismic observation networks in Japan—Hi-net, F-net, K-NET and KiK-ne, *Earth Planets Space*, *56*, 15–28.

Peng, Z., and P. Zhao (2009), Migration of early aftershocks following the 2004 Parkfield earthquake, *Nat. Geosci.*, *2*(12), 877–881, doi:10.1038/ngeo697.

Pollitz, F. F., R. Bürgmann, and P. Banerjee (2011), Geodetic slip model of the 2011 M9.0 Tohoku earthquake?, *Geophys. Res. Lett.*, *38*, L00G08, doi:10.1029/2011GL048632.

Segou, M., T. Parsons, and W. Ellsworth (2013), Comparative evaluation of physics-based and statistical forecasts in Northern California, *J. Geophys. Res. Solid Earth*, *118*, 6219–6240, doi:10.1002/2013JB010313.

Steacy, S., S. Nalbant, J. McCloskey, C. Nostro, O. Scotti, and D. Baumont (2005), Onto what planes should Coulomb stress perturbations be resolved?, *J. Geophys Res.*, *110*, B05S15, doi:10.1029/2004JB003356.

Sudhaus, H., and S. Jónsson (2009), Improved source modelling through combined use of InSAR and GPS under consideration of correlated data errors: Application to the June 2000 Kleifarvatn earthquake, Iceland, *Geophys. J. Int.*, *176*(2), 389–404, doi:10.1111/j.1365-246X.2008.03989.x.

Toda, S., and B. Enescu (2011), Rate/state Coulomb stress transfer model for the CSEP Japan seismicity forecast, *Earth Planets Space*, *63*(3), 171–185, doi:10.5047/eps.2011.01.004.

Wang, L., S. Hainzl, G. Zöller, and M. Holschneider (2012), Stress- and aftershock-constrained joint inversions for coseismic and postseismic slip applied to the 2004 M6.0 Parkfield earthquake, *J. Geophys Res.*, *117*, B07406, doi:10.1029/2011JB009017.

Wang, L., J. Liu, J. Zhao, and J. Zhao (2013), Co- and post-seismic modeling of the 2011 M9 Tohoku-Oki earthquake, and its impact on China mainland, *Earthquake*, *33*(4), 238–246.

Wang, R., F. Lorenzo-Martín, and F. Roth (2006), PSGRN/PSCMP—A new code for calculating co- and post-seismic deformation, geoid and gravity changes based on the viscoelastic-gravitational dislocation theory, *Comput. Geosci.*, *32*(4), 527–541, doi:10.1016/j.cageo.2005.08.006.

Woessner, J., et al. (2011), A retrospective comparative forecast test on the 1992 Landers sequence, *J. Geophys. Res.*, *116*, B05305, doi:10.1029/2010JB007846.

Woessner, J., S. Jónsson, H. Sudhaus, and C. Baumann (2012), Reliability of Coulomb stress changes inferred from correlated uncertainties of finite-fault source models, *J. Geophys Res.*, *117*, B07303, doi:10.1029/2011JB009121.

Yagi, Y., and Y. Fukahata (2011), Introduction of uncertainty of Green's function into waveform inversion for seismic source processes, *Geophys. J. Int.*, *186*(2), 711–720, doi:10.1111/j.1365-246X.2011.05043.x.

Yang, W., E. Hauksson, and P. M. Shearer (2012), Computing a large refined catalog of focal mechanisms for southern California (1981–2010): Temporal stability of the style of faulting, *Bull. Seismol. Soc. Am.*, *102*(3), 1179–1194, doi:10.1785/0120110311.