



A Bayesian source model for the 2004 great Sumatra-Andaman earthquake

Q. Bletery, A. Sladen, J. Jiang, M. Simons

► To cite this version:

Q. Bletery, A. Sladen, J. Jiang, M. Simons. A Bayesian source model for the 2004 great Sumatra-Andaman earthquake. *Journal of Geophysical Research: Solid Earth*, 2016, 121 (7), pp.5116 - 5135. 10.1002/2016JB012911 . hal-01401349

HAL Id: hal-01401349

<https://hal.science/hal-01401349>

Submitted on 20 Sep 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

RESEARCH ARTICLE

10.1002/2016JB012911

A Bayesian source model for the 2004 great Sumatra-Andaman earthquake

Quentin Bletery^{1,2}, Anthony Sladen¹, Junle Jiang^{3,4}, and Mark Simons³

Key Points:

- Image the slip distribution of the Sumatra-Andaman earthquake
- Quantify uncertainties on the slip distribution
- Analyze the possible slip models

Supporting Information:

- Supporting Information S1

Correspondence to:

Q. Bletery,
qbletery@uoregon.edu

Citation:

Bletery, Q., A. Sladen, J. Jiang, and M. Simons (2016), A Bayesian source model for the 2004 great Sumatra-Andaman earthquake, *J. Geophys. Res. Solid Earth*, 121, 5116–5135, doi:10.1002/2016JB012911.

Received 11 FEB 2016

Accepted 22 JUN 2016

Accepted article online 5 JUL 2016

Published online 16 JUL 2016

¹University of Nice Sophia Antipolis, CNRS, IRD, Observatoire de la Côte d'Azur, Géoazur UMR, Valbonne, France, ²Now at the Department of Geological Sciences, University of Oregon, Eugene, Oregon, USA, ³Seismological Laboratory, Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California, USA, ⁴Now at the Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California, USA

Abstract The 2004 M_w 9.1–9.3 Sumatra-Andaman earthquake is one of the largest earthquakes of the modern instrumental era. Despite considerable efforts to analyze this event, the different available observations have proven difficult to reconcile in a single finite-fault slip model. In particular, the critical near-field geodetic records contain variable and significant postseismic signal (between 2 weeks' and 2 months' worth), while the satellite altimetry records of the associated tsunami are affected by various sources of uncertainties (e.g., source rupture velocity and mesoscale oceanic currents). In this study, we investigate the quasi-static slip distribution of the Sumatra-Andaman earthquake by carefully accounting for the different sources of uncertainties in the joint inversion of available geodetic and tsunami data. To this end, we use nondiagonal covariance matrices reflecting both observational and modeling uncertainties in a fully Bayesian inversion framework. Modeling errors can be particularly large for great earthquakes. Here we consider a layered spherical Earth for the static displacement field, nonhydrostatic equations for the tsunami, and a 3-D megathrust interface geometry to alleviate some of the potential epistemic uncertainties. The Bayesian framework then enables us to derive families of possible models compatible with the unevenly distributed and sometimes ambiguous measurements. We infer two regions of high fault slip at 3°N–4°N and 7°N–8°N with amplitudes that likely reach values as large as 40 m and possibly larger. These values are a factor of 2 larger than typically found in previous studies—potentially an outcome of commonly assumed forms of regularization. Finally, we find that fault rupture very likely involved shallow slip. Within the resolution provided by the existing data, we cannot rule out the possibility that fault rupture reached the trench.

1. Introduction

The 2004 M_w 9.1–9.3 Sumatra-Andaman earthquake ruptured along a fault segment 1600 km long [Meltzner et al., 2006], initiating offshore northern Sumatra and propagating unilaterally northward beyond the Andaman islands (see review paper of Shearer and Bürgmann [2010], and references therein). The earthquake induced a massive tsunami responsible for casualties and damages along most of the coasts surrounding the Indian Ocean. The total duration of this exceptionally long earthquake was estimated from high-frequency seismology to be approximately 500 s [Ni et al., 2005; Ishii et al., 2005; Krüger and Ohrnberger, 2005; Yao et al., 2011]. Such a long duration made conventional teleseismic body wave inversions for the distribution of fault slip difficult because of the interference between direct and reflected waves. Surface waves [Ammon et al., 2005] and normal modes [Stein and Okal, 2005; Park et al., 2005; Clévéde et al., 2012] were also used to analyze the source, but the information content of low-frequency records is inherently much less than high-frequency records and thus only provided low resolution source models.

Geodetic data provided additional constraints on the inferred spatial distributions of fault slip [Vigny et al., 2005; Banerjee et al., 2005; Catherine et al., 2005; Hashimoto et al., 2006], in particular near-field campaign Global Positioning System (GPS) data [Gahalaut et al., 2006; Subarya et al., 2006; Banerjee et al., 2007; Pietrzak et al., 2007; Rhie et al., 2007; Chlieh et al., 2007]. However, these near-field records contain, to varying extent, significant postseismic signal, which induced large—and often neglected—uncertainties in the inferred coseismic slip models.

Records of the tsunami from distant tide gauges also provide constraints on the slip distribution. *Merrifield et al.* [2005] and *Lay et al.* [2005] extracted information from the first wave arrival by backprojection methods, while *Tanioka et al.* [2006] and *Piatanesi and Lorito* [2007] tried to extract information from the full waveforms by performing finite-fault source inversions. However, the low sampling rate of the gauges combined with ambiguity on the stations clocks [*Lorito et al.*, 2010] and nonlinearity in the tsunami wave propagation (essentially due to the shallow depths of the gauges) [*Piatanesi and Lorito*, 2007] strongly limit their utility. The tsunami was also recorded by satellite altimeters which observed water height anomalies in the Indian Ocean about 2 h after the earthquake. These observations are in the deep ocean and thus minimally affected by nonlinear propagation effects. *Ablain et al.* [2006], *Hirata et al.* [2006], and *Sladen and Hébert* [2008] demonstrated the utility of these observations, which provide resolution on the shallow portion of the fault, especially in the southern part of the rupture where the altimeters recorded the tsunami front. Several studies included these observations in joint inversions together with tide gauge [*Fujii and Satake*, 2007] and geodetic data [*Hoechner et al.*, 2008; *Lorito et al.*, 2010].

The models obtained in all these studies exhibit a very extended coseismic slip distribution initiating offshore northern Sumatra and ending, depending on the studies, at significantly different locations in the Andaman archipelago (from latitude 10°N to 14°N). Most models show two regions of large slip offshore northern Sumatra (latitude 3°N–4°N) and Nicobar Islands (latitude 7°N–9°N) with maximum amplitudes varying from one study to another (grossly from 10 m [*Ammon et al.*, 2005] to 40 m [*Rhie et al.*, 2007]). The moment magnitude derived from these models varies from M_w 9.0 to 9.3.

An intriguing feature of this exceptional earthquake is the very slow rupture inferred at the northern end of the fault where low seismic radiation was detected [*Ammon et al.*, 2005; *Tsai et al.*, 2005; *Vallée*, 2007] but where large geodetic offsets were measured [*Gahalaut et al.*, 2006] and intense aftershocks occurred [e.g., *Lay et al.*, 2005]. The anomalously large low-frequency content in seismic records associated with a northward migration of the centroid compared to the centroid obtained from surface waves [*Stein and Okal*, 2005; *Park et al.*, 2005; *Clévéde et al.*, 2012] confirms that large coseismic moment was released in this Andaman section. The atypical behavior of one near-field tide gauge at Port Blair (PB in Figure 1) suggests that the rise time could be on the order of 30 min [*Singh et al.*, 2006], although this inference relies only on one observation. Such a long rise time would produce a tsunami of smaller amplitude and would be difficult to model since we have very little constraint on the exact slip history. One interpretation of these observations is that the slow rupture had an early fast slip component and then continued slipping slowly [*Lay et al.*, 2005; *Singh et al.*, 2006; *Clévéde et al.*, 2012].

Because of the inherent challenges in modeling the seismic and tide gauge records, we focus this study on geodetic and tsunami data. The geodetic data are a compilation of different types of measurements: the data set of *Chlieh et al.* [2007]—including offsets obtained by GPS and altimetry as well as coral and geological measurements—complemented with far-field GPS data [*Vigny et al.*, 2005; *Kreemer et al.*, 2006; *Banerjee et al.*, 2007] and offsets obtained by hyperspectral images [*Smet et al.*, 2008]. The tsunami data consist of satellite altimetry observations [*Ablain et al.*, 2006]. To reduce epistemic uncertainties, we compute our Green's functions using a 3-D fault geometry considering a spherical layered Earth and predict the tsunami data using the nonhydrostatic wave equations.

As mentioned earlier, a challenge arises in inverting the available geodetic data due to the different time windows of the offsets measurements, as most near-field measurements are contaminated by variable and significant amounts of postseismic deformation. Previously, this complication was neglected [*Subarya et al.*, 2006; *Gahalaut et al.*, 2006], treated by assuming large error bars [*Pietrzak et al.*, 2007], roughly estimated from the few far-field continuous stations available [*Banerjee et al.*, 2007; *Chlieh et al.*, 2007], or considered sufficiently problematic that affected data were excluded [*Lorito et al.*, 2010]. Here we propose to extract relevant information from these data containing variable amounts of postseismic displacement by taking into account their associated—sometimes spatially correlated—uncertainties (e.g., two nearby measurements affected by the same source of postseismic deformation share coherent errors). To do so, we use nondiagonal covariance matrices [*Yagi and Fukahata*, 2008; *Duputel et al.*, 2014] reflecting both observational and modeling uncertainties in a fully Bayesian inversion framework in order to evaluate the range of possible slip models compatible with the observations.

Using a Bayesian formalism, we aim to answer some fundamental questions such as the likelihood that the rupture reached the surface or, more generally, to evaluate the ranges of possible fault slip amplitudes: many

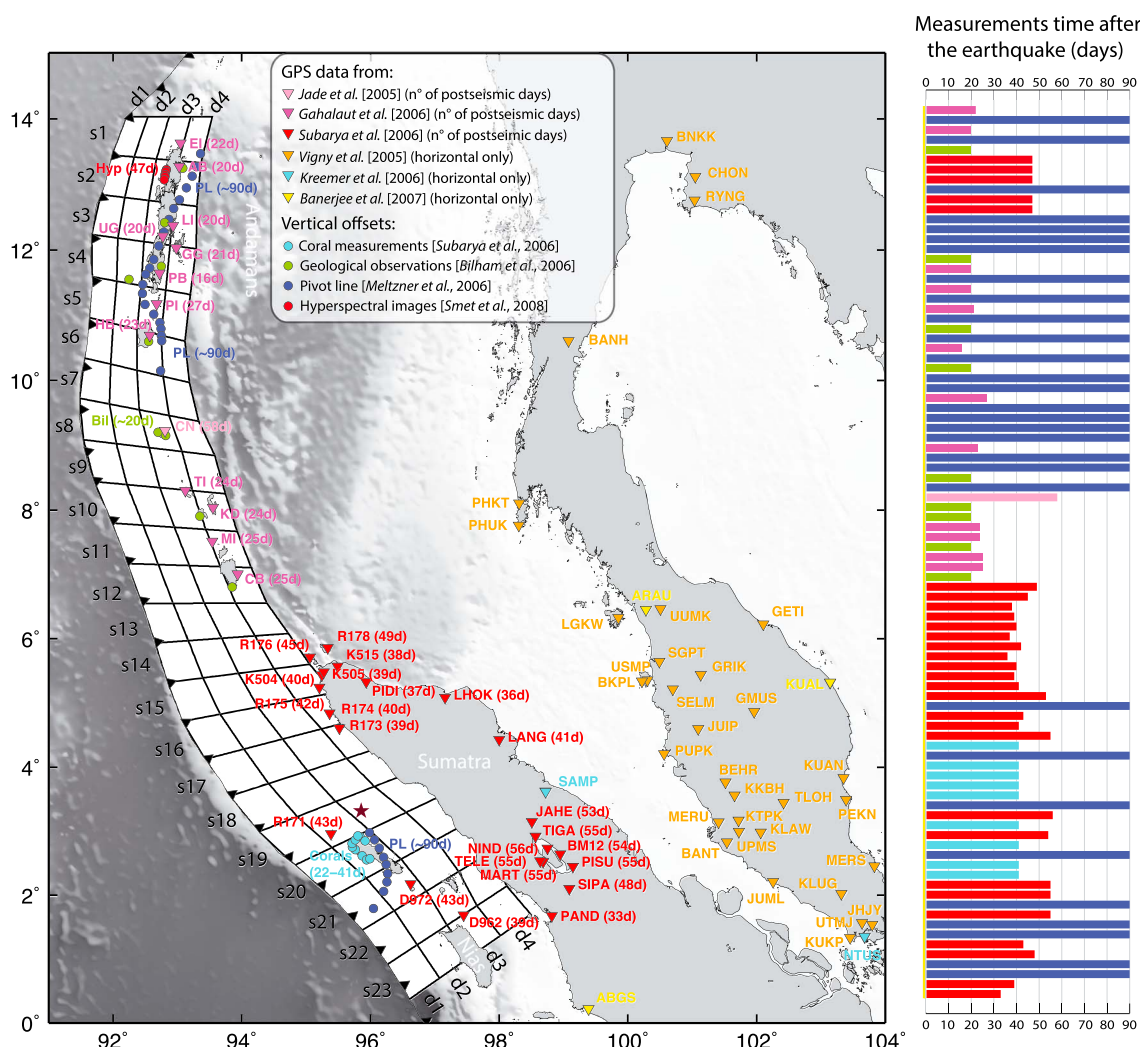


Figure 1. Near-field measurements of coseismic offsets. Inverted triangles are horizontal and vertical GPS measurements (except for stations R178, TELE, and data sets from Vigny *et al.* [2005], Banerjee *et al.* [2007], and Kreemer *et al.* [2006], which are horizontal only). Circles are vertical-only measurements. When data contains postseismic signal, the measurement time (in days) after the earthquake is given in brackets and represented in the bar plot at the right of the figure. s# and d# labels indicate patch numbers along strike and dip directions, respectively. The star indicates the epicenter (U.S. Geological Survey).

studies may have excluded solutions with very high slip amplitudes of 50–60 m, although such values were later documented or inferred in the case of the 2011 M_w 9.0 Tohoku-Oki earthquake [e.g., Fujiwara *et al.*, 2011; Sato *et al.*, 2011; Simons *et al.*, 2011; Bletery *et al.*, 2014]. Refining our knowledge of the slip history of this large event is also essential to explain the impact of the tsunami along the coasts of the Indian Ocean [Fritz and Borrero, 2006; Goff *et al.*, 2006; Okal *et al.*, 2006] and discriminate source and propagation effects from local coastal effects which might still be at play in the case of future events [e.g., Hébert *et al.*, 2007].

2. Data

2.1. Geodetic Data

We compile near-field campaign GPS data from Jade *et al.* [2005], Subarya *et al.* [2006], and Gahalaut *et al.* [2006] which recorded coseismic offsets very close to the fault. These measurements include between 16 and 58 days of postseismic signal (Figure 1), which may represent a significant part of the total signal (see, for instance, the compilation of postseismic estimates in Lin *et al.* [2013]). Near-field GPS data, their uncertainties and the extent of time following the earthquake spanned by the coseismic data are summarized in Table S1 in the supporting information. In addition to GPS, we include vertical near-field offsets measured on corals [Subarya *et al.*, 2006], other sea surface height landmarks [Bilham *et al.*, 2005], hyperspectral imagery [Smet *et al.*, 2008], and the pivot line defined by altimetry [Meltzner *et al.*, 2006]. Uncertainties and the temporal

extent of postseismic deformation associated with these data are summarized in Table S2. Uncertainties on the data from *Bilham et al.* [2005] are large. Altimetry images used to derive the pivot line were acquired as late as 90 days after the earthquake [Meltzner et al., 2006].

We also consider far-field horizontal GPS data from *Vigny et al.* [2005] and *Banerjee et al.* [2007]. Following the method of *Banerjee et al.* [2007], we retrieve the 5 days postseismic estimate to the coseismic solution of *Vigny et al.* [2005] in order to remove most of the postseismic signal contained in their 14 day average offsets. When offsets were also determined by *Kreemer et al.* [2006], we prefer these latter solutions because of the smaller associated uncertainties (see Table S3). The solutions of the different geodetic studies were all computed in the International Terrestrial Reference Frame 2000 and are thus consistent with each others. The location of all the far-field stations is shown in Figure S1.

2.2. Tsunami Observations

Sea surface height anomalies were recorded by several satellite altimeters in different parts of the Indian Ocean [Ablain et al., 2006]. Based on the signal-to-noise ratios of these time series, we restrict our analysis to records from Jason-1 and TOPEX-Poseidon satellites. These orbital tracks are relatively close to the source and provide snapshot profiles of the propagating tsunami about 2 h after the earthquake. These altimeter measurements record the direct wave front in the southern part of the profiles and reflected waves in the north (Figure 2). Thus, they provide better constraints on the rupture of the southern half than on the rupture of the northern half of the fault plane.

As noted by *Piatanesi and Lorito* [2007], inversion of tide gauge data is a highly nonlinear problem in the case of the Sumatra-Andaman earthquake as indicated by the large discrepancies between the prediction of the final model and the sum of the predictions of the individual subfaults. This problem was further highlighted by *Poisson et al.* [2011] who showed that inversions of tide gauge data [Fujii and Satake, 2007; Piatanesi and Lorito, 2007] provided a poorer fit on the tsunami height recorded by the Jason-1 satellite than geodetic and seismogeodetic studies [Banerjee et al., 2007; Rhie et al., 2007]. The inconsistency likely comes from the shallow depth of the water where the tide gauges are located, which invalidates common approximations, and makes simulations oversensitive to bathymetry inaccuracies. Also, tide gauges are often sheltered far inside the harbors, which strongly affects the simulations because of the complex structure of the ports. Moreover, *Lorito et al.* [2010] found systematic inconsistencies in the arrival times at the different gauges and suggest that station clocks might be inaccurate. Given all these issues, we choose not to include tide gauges in our inversion.

3. 3-D Fault Geometry and the Computation of the Green's Functions

We compute the theoretical deformation at the stations coordinates for 4459 rectangular subfaults (both for pure dip-slip and pure strike-slip dislocations) in a layered spherical Earth using the code STATIC1D—which is based on normal mode theory [Pollitz, 1996]—and the PREM Earth model [Dziewonski and Anderson, 1981]. The subfaults are square patches of 9.26 km on a side. To obtain sufficient numerical accuracy, we consider normal modes to harmonic degree 10,000. The coordinates of each slip patch (longitude, latitude, and depth) as well as their dip δ and strike ϕ angles follow the SLAB1.0 Sumatra subduction model [Hayes et al., 2012]. As this subduction model does not cover the northern Andaman area, we extended the model using tomography profiles of *Pesicek et al.* [2010]. To compute our Green's functions, we then add the contributions of the 4459 subfaults into 92 larger subfaults. These large subfaults are not exactly rectangles as δ and ϕ vary among the smaller component sources but are rather 3-D surfaces mimicking the geometrical complexities of the megathrust interface.

For the tsunami Green's functions, we first compute the ocean floor deformation—with a grid spacing of 40 arc seconds—for the same 4459 small subfaults as for geodetic predictions with the same code STATIC1D. As the contribution of horizontal motion (or bathymetry effect) is negligible in the Sumatra-Andaman subduction zone [Bletery et al., 2015], we then consider the vertical ocean floor deformation field as the initial state of sea surface height disturbance and compute the tsunami wave propagation with the code NEOWAVE [Yamazaki et al., 2011a, 2011b]. This code is based on the nonhydrostatic wave equations and accounts for dispersive effects of short-period waves. The resulting Green's functions are then time shifted depending on their respective distances to the hypocenter to account for the kinematics of the rupture propagation assuming a rupture velocity of 2.8 km/s determined from backprojection [Ishii et al., 2005], as the technique was shown

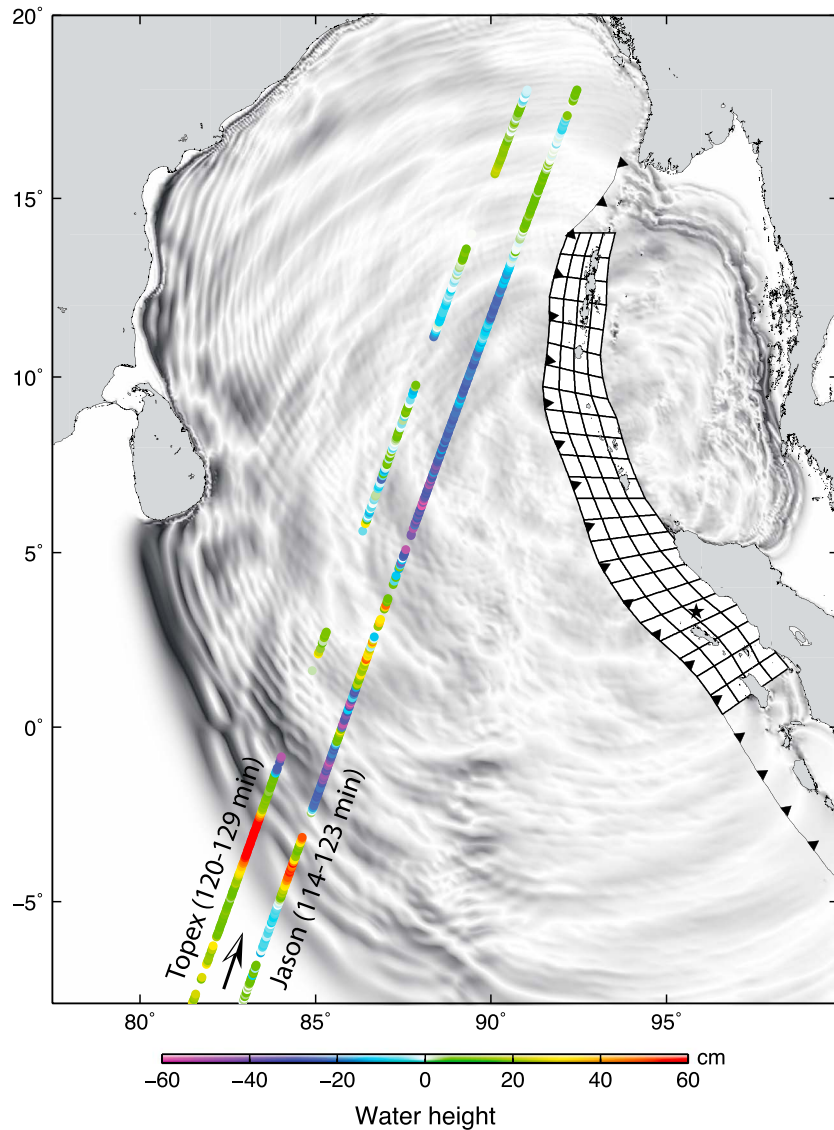


Figure 2. Water height anomaly recorded by satellite altimeters Jason-1 and TOPEX-Poseidon, approximately 2 h after the main shock is plotted with colored dots. Time of recording after the main shock is indicated for the southern and northern most points for each satellite (tracks were ascending). The approximate tsunami wave field 2 h after initiation is shown in black and white.

to reflect the rupture front propagation [Okuwaki *et al.*, 2014; Avouac *et al.*, 2015]. The hypocenter initiation time is taken to be 00:58:53 UTC (earthquake.usgs.gov).

4. Accounting for Data and Model Uncertainties

4.1. The Bayesian Framework

We use a Bayesian sampling approach to estimate the solution space of the slip models consistent with the data given the uncertainties we adopt on both the data and the Green's functions. This approach relies on the Cascading Adaptive Transitional Metropolis In Parallel (CATMIP) algorithm [Minson *et al.*, 2013] newly reimplemented in the code suite Altar to exploit the high efficiency of graphics processing units. The probability density function (PDF) $p(\mathbf{m}|\mathbf{d}_{\text{obs}})$ of a sampled model \mathbf{m} is evaluated based on its ability to fit the data \mathbf{d}_{obs} within given data and model uncertainties:

$$p(\mathbf{m}|\mathbf{d}_{\text{obs}}) \propto p(\mathbf{m}) \exp\left(-\frac{1}{2}(\mathbf{d}_{\text{obs}} - \mathbf{Gm})^T \mathbf{C}_x^{-1}(\mathbf{d}_{\text{obs}} - \mathbf{Gm})\right) \quad (1)$$

where \mathbf{G} is the matrix of the Green's functions, \mathbf{C}_χ is the misfit covariance that approximates both observational and modeling uncertainties, and $p(\mathbf{m})$ is the a priori information on the model. The CATMIP/AlTar algorithm is characterized by a very slow cooling process (the rejection criterion is refined by small steps) between the a priori state of information and the final PDFs designed to ensure good sampling of the whole solution space [Minson *et al.*, 2013]. The final PDFs obtained allow interpreting the results of a fault slip inversion with a robust estimation of the uncertainty on the inverted parameters [Simons *et al.*, 2011; Minson *et al.*, 2014; Jolivet *et al.*, 2014; Duputel *et al.*, 2015]. The sampling approach we use allows to consider non-Gaussian prior distributions $p(\mathbf{m})$. For instance, we use truncated uniform distributions—bounded between -0.5 m and 60 m (the slightly negative lower bound is taken to ensure good sampling near 0)—for the dip-slip component. We use Gaussian distributions with 0 mean and 1 m standard deviation for the strike-slip component.

In source inversions, data uncertainties are usually accounted for by weighting the different data based on their respective levels of confidence. We use a formalism that enables a more exhaustive exploration of the possible models accounting for both data and model prediction uncertainties. Under the assumption of Gaussian uncertainty distributions, data and model uncertainties can be conveniently represented by a simple sum of the associated covariance matrices [Tarantola, 2005]. The covariance matrix \mathbf{C}_χ in equation (1) may then be written

$$\mathbf{C}_\chi = \mathbf{C}_d + \mathbf{C}_p \quad (2)$$

where \mathbf{C}_d and \mathbf{C}_p are the covariance matrices associated with the data measurements and model prediction, respectively. \mathbf{C}_χ is nondiagonal because uncertainties between two data points can be strongly correlated. We further decompose \mathbf{C}_p into three terms reflecting the uncertainties associated with our imperfect knowledge of Earth structure ($\mathbf{C}_{\text{earth}}$), postseismic contamination (\mathbf{C}_{post}), and in the case of tsunami data, kinematics of the rupture (\mathbf{C}_{time}):

$$\mathbf{C}_\chi = \mathbf{C}_d + \mathbf{C}_{\text{earth}} + \mathbf{C}_{\text{post}} + \mathbf{C}_{\text{time}} \quad (3)$$

We discuss these different contributions below as well as how they can be estimated and included in the inversion.

4.2. \mathbf{C}_d

\mathbf{C}_d (Figure S2a) is the data covariance. The uncertainties associated with the geodetic measurements that compose our particular data set can be considered independent from each others, because the stations are far from each others. Therefore, error covariances can be considered as negligible compared to the other off-diagonal terms (in particular those of \mathbf{C}_{post} , see below). Thus, $\mathbf{C}_d^{\text{geod}}$ is diagonal and its diagonal terms are equal to the squares of the 1σ uncertainties $(\sigma_{\text{data}}^{\text{geod}})_i$, provided by the individual geodetic studies (see Tables S1–S3).

For tsunami observations recorded by altimetry, uncertainties are highly correlated because mesoscale oceanic variability is difficult to differentiate from the tsunami signal [Ablain *et al.*, 2006]. Such coherent noise is visible on the southern-most part of the altimetry profiles which are recorded before the tsunami arrives (Figure 2). We build $\mathbf{C}_d^{\text{alt}}$ (Figure S3a) to reflect the spatial correlations of the uncertainties:

$$(\mathbf{C}_d^{\text{alt}})_{ij} = (\sigma_{\text{data}}^{\text{alt}})_i (\sigma_{\text{data}}^{\text{alt}})_j \exp\left(-\frac{L_{ij}}{\lambda}\right) \quad (4)$$

where $(\sigma_{\text{data}}^{\text{alt}})_i$ is the absolute uncertainty on measurement i , L_{ij} is the distance between two measurements i, j , and λ is the correlation length. Based on the signal of the southern-most part of the profiles (which are not affected by the tsunami), we consider $(\sigma_{\text{data}}^{\text{alt}})_i = 10$ cm and $\lambda = 100$ km. As geodetic and tsunami observations are a priori independent from each other:

$$\mathbf{C}_d = \begin{pmatrix} \mathbf{C}_d^{\text{geod}} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_d^{\text{alt}} \end{pmatrix} \quad (5)$$

4.3. $\mathbf{C}_{\text{earth}}$

$\mathbf{C}_{\text{earth}}$ accounts for errors in the Green's functions due to imperfect knowledge of the Earth's interior and in particular errors in the rigidity assumed for the different Earth layers. We only consider $\mathbf{C}_{\text{earth}}$ for geodetic data since for tsunami data this source of error is considered negligible compared to \mathbf{C}_d and \mathbf{C}_{time} . We follow the

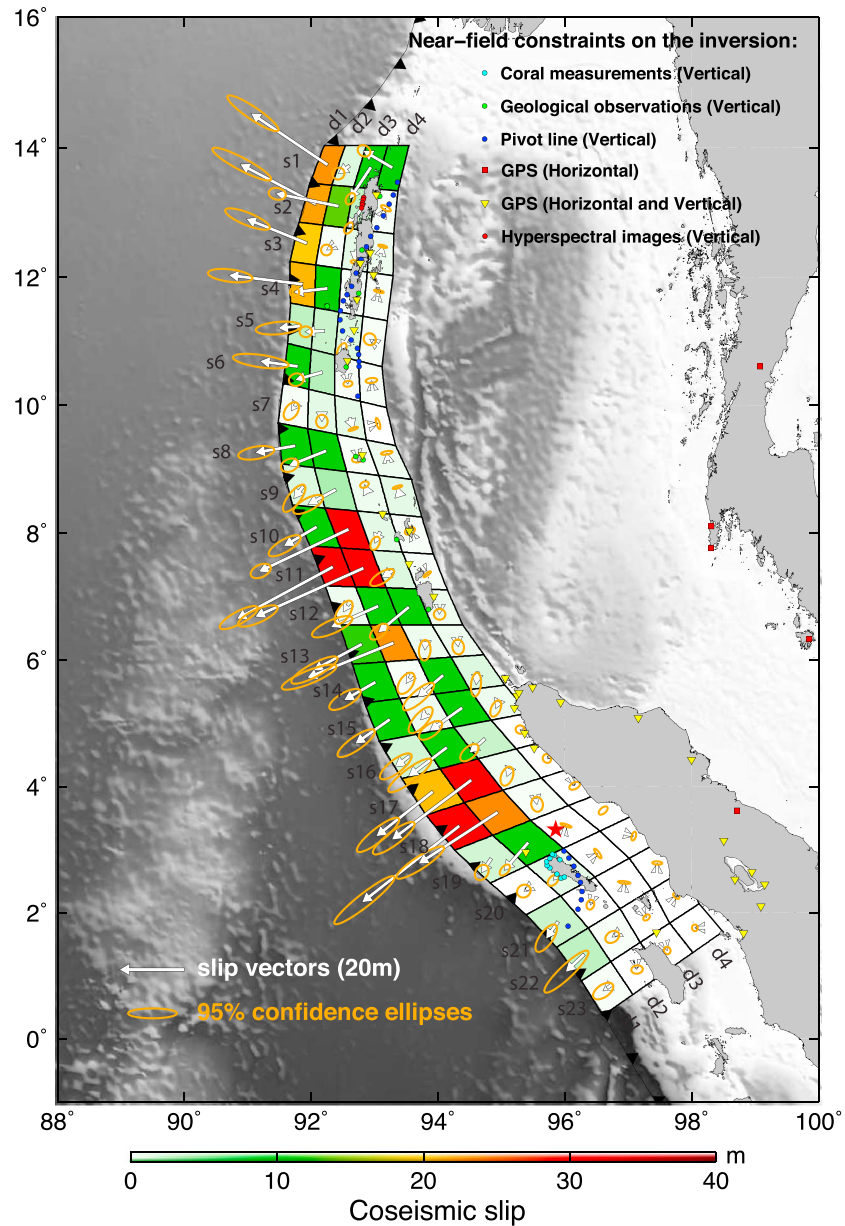


Figure 3. Average slip vectors (white arrows) obtained by joint inversion of geodetic and altimetric tsunami data and associated standard deviations (orange ellipses). The colors of the subfaults correspond to the amplitudes of the slip vectors in meters. *s#* and *d#* labels indicate patch numbers along strike and dip directions, respectively.

approach of Duputel *et al.* [2014] to iteratively compute this covariance matrix. We first precompute, by finite differences, the sensitivity kernels \mathbf{K}_μ^G of each Green's function in \mathbf{G} with respect to the shear moduli μ_j of the different layers in our 1-D spherical Earth model:

$$(\mathbf{K}_\mu^G)_{i,j,k} = \frac{\partial G_{i,k}}{\partial \ln \mu_j} \quad (6)$$

where *i* denotes the different data, *j* denotes the different Earth layers, and *k* denotes the different source parameters we aim to evaluate. Then, \mathbf{K}_μ is recalculated at each cooling step (when the rejection criterion is refined) based on the mean $\langle \mathbf{m} \rangle$ of the sampled models \mathbf{m} :

$$\mathbf{K}_\mu = \mathbf{K}_\mu^G \langle \mathbf{m} \rangle \quad (7)$$

Given a covariance matrix \mathbf{C}_μ on the shear moduli estimates μ , we derive our prediction covariance:

$$\mathbf{C}_{\text{earth}} = \mathbf{K}_\mu \mathbf{C}_\mu \mathbf{K}_\mu^T \quad (8)$$

Here we assume 8% independent uncertainties of the shear modulus estimates on the different layers such that \mathbf{C}_μ is diagonal. The resulting space of Earth models is shown in Figure S4.

4.4. \mathbf{C}_{post}

\mathbf{C}_{post} is a new term we introduce to account for potential contamination of our geodetic data by postseismic effects. As shown in Figure 1, our relatively dense network of near-field observations is sparsely sampled in time. Indeed, most near-field observations contain between 2 weeks and 2 months of postseismic signal. Thus, our data likely share spatially coherent noise.

To account for this coherent noise, we need to define what we call coseismic and postseismic deformation. We consider as coseismic slip, the slip that occurs in the time window between earthquake initiation time and 10 min later, which could potentially generate tsunamis. Any signal due to fault slip outside the first 10 min window is then considered as postseismic contamination and should be accounted for in \mathbf{C}_{post} (e.g., GPS daily solutions are considered to contain 1 day of postseismic contamination). Therefore, \mathbf{C}_{post} is null for tsunami data.

Assuming that postseismic deformation in the first few weeks after the earthquake is restricted to afterslip on our discretized fault (i.e., neglecting viscoelastic effects), we use our coseismic Green's functions to derive covariances between our data due to postseismic contamination. The assumption of afterslip is supported by studies made on the later postseismic deformation revealing motion dominated by afterslip on the fault during the first 2 years [Paul et al., 2007; Gahalaut et al., 2008] with the relative contribution of viscoelastic effects gradually increasing later [Paul et al., 2012; Gunawan et al., 2014]. As we do not want to introduce a priori information on the early postseismic slip distribution, we draw 1,000,000 samples \mathbf{m}_n from uniform distributions for the dip-slip component and from Gaussian distributions for the strike-slip component. To obtain a plausible correlation matrix, the sampling of the uniform distributions is bounded so that the associated coseismic moment magnitude is close to 9.2. The Gaussian distributions have 0 mean and 2 m of standard deviation. From these one million possible coseismic slip models, we compute the corresponding synthetic data and derive from them the co-seismic covariance matrix \mathbf{C}_{co} :

$$\mathbf{C}_{\text{co}} = \text{cov}(\mathbf{Gm}_1, \mathbf{Gm}_2, \dots, \mathbf{Gm}_{1,000,000}) \quad (9)$$

We then assume a logarithmic function of the afterslip over coseismic ratio:

$$r_i = \alpha \log\left(\frac{t_i}{\beta} + 1\right) \quad (10)$$

where t_i is the measurement time after the earthquake (in days); $\alpha = 0.21$ and $\beta = 14$ are determined empirically to predict a postseismic signal on the order of 15% of the coseismic signal after 15 days and 25% after 30 days, as observed on the few continuous stations available [Subarya et al., 2006]. Then, we can derive the afterslip covariance matrix \mathbf{C}_{post} (Figure S2b) from \mathbf{C}_{co} by normalizing its terms by the estimate of afterslip over coseismic ratio:

$$(\mathbf{C}_{\text{post}})_{ij} = \min(r_i, r_j)^2 (\mathbf{C}_{\text{co}})_{ij} \quad (11)$$

The covariance between the uncertainties of two data d_i and d_j depends on their common time of postseismic contamination, which is the minimum of t_i and t_j ; hence, we normalize $(\mathbf{C}_{\text{post}})_{ij}$ by the minimum of r_i and r_j . If d_i is a tsunami observation, its afterslip covariance with any other observation d_j will be null as anything that happens after the tsunamogenic phase is independent from the signal recorded in d_i . The diagonal of \mathbf{C}_{post} is simply \mathbf{C}_{co} normalized by the square of the afterslip over coseismic ratio.

The amplitudes of the variance in \mathbf{C}_{post} are actually much larger than those in \mathbf{C}_d for many near-field data (Figure S2) as GPS measurements uncertainties often do not exceed a few centimeters (Table S1). We note that some of the off-diagonal terms of \mathbf{C}_{post} —which are usually ignored—are on the order of the diagonal terms, suggesting strong spatial correlations of the errors due to postseismic signals. The addition of \mathbf{C}_{post} mitigates overfitting near-field data. For far-field GPS data, uncertainties on both the measurements and the postseismic contamination are very low (because they are daily solutions); but as the signal is also very low, the resulting resolving power is still very limited.

4.5. \mathbf{C}_{time}

\mathbf{C}_{time} is the covariance matrix associated with uncertainty on the rupture time of each subfault. This term is null for geodetic data as they are not affected by the kinematics of the rupture. In contrast, for tsunami data, each patch rupture time is calculated assuming the earthquake initiated at 00:58:53 UTC (source: <http://earthquake.usgs.gov/>) and that it propagated at a velocity of 2.8 km/s [Ishii et al., 2005]. Errors on these estimations would lead to nonnegligible errors in the tsunami Green's functions [Sladen and Hébert, 2008]. Because of nonlinearity, we estimate these errors by calculating Green's functions assuming a set of nine different rupture velocities in the range ($V_1 = 2.4$ km/s, $V_2 = 2.5$ km/s, $V_3 = 2.6$ km/s, ..., $V_9 = 3.2$ km/s) and two initiation times ($t_1 = 00:58:53$ UTC, $t_2 = 00:59:13$ UTC). Thus, we obtain 18 Green's functions matrices reflecting the values most commonly assumed in the literature [Ammon et al., 2005; Park et al., 2005; Ishii et al., 2005; Krüger and Ohrnberger, 2005; Ishii et al., 2007; Lorito et al., 2010]. We then draw 100,000 models \mathbf{m}_n from uniform (for the trench perpendicular component) and Gaussian (for the trench parallel one) distributions such as for the calculation of \mathbf{C}_{post} . We multiply each drawn model by each Green's function matrix to obtain 1,800,000 sets of predicted tsunami data and derive \mathbf{C}_{time} (Figure S3b) from their covariances:

$$\left\{ \begin{array}{l} \mathbf{d}_1^{V_1, t_1} = \mathbf{G}^{V_1, t_1} \mathbf{m}_1 \\ \mathbf{d}_1^{V_2, t_1} = \mathbf{G}^{V_2, t_1} \mathbf{m}_1 \\ \dots \\ \mathbf{d}_{100,000}^{V_9, t_2} = \mathbf{G}^{V_9, t_2} \mathbf{m}_{100,000} \end{array} \right\} \quad (12)$$

$$\mathbf{C}_{\text{time}} = \text{cov}(\mathbf{d}_1^{V_1, t_1}, \mathbf{d}_1^{V_2, t_1}, \dots, \mathbf{d}_{100,000}^{V_9, t_2}) \quad (13)$$

We find that the uncertainties associated with the rupture propagation are significantly larger than the uncertainties on the measurements (Figure S3, note the different color scales). Thus, the addition of \mathbf{C}_{time} should strongly affect the solution. We note as well that as for \mathbf{C}_{post} , the off-diagonal terms are on the order of the diagonal terms (Figure S3b). Large off-diagonal terms were to be expected since changing a subfault rupture initiation time result in shifting the synthetic tsunami wave which is going to affect neighboring data in a very coherent manner. While diagonal terms limit the information contained in the data, off-diagonal terms, on the contrary, bring information by stipulating that the error on one measurement is similar to the error on other measurements. Thus, while diagonal terms extend the solution space, off-diagonal terms actually shrink it and help constrain the solution.

5. Results

5.1. Bayesian Exploration of Probable Source Models

Accounting for the dominant sources of uncertainty both in the data and in the Green's functions, we apply the AlTar scheme to jointly invert tsunami and geodetic data for the rupture process of the Sumatra-Andaman earthquake. We obtain probability density functions (PDFs) for both the strike-slip (Figure S5) and dip-slip (Figure S6) components. These PDFs represent the solution space of the full inverse problem. An alternative—more visible but more restrictive—way to represent the obtained solution space is in the form of the average solution (Figure 3) and its associated 95% confidence ellipses derived from the standard deviations of the PDFs. The average model reveals a very extended rupture (roughly from latitude 2°N to latitude 14°N from south to north) with two areas of large slip (30–40 m) between latitudes 3°N and 4°N and latitudes 7°N and 8°N. Moreover, we infer significant slip (≈ 20 m) offshore Andaman islands. The mean moment is 8.37×10^{22} N m, corresponding to a moment magnitude $M_w = 9.25$. The standard deviation on the moment is very small -0.2×10^{22} N m ($M_w = 9.25 \pm 0.01$)—but it could vary more significantly depending on the assumed elastic structure. This moment magnitude is close to the magnitude (M_w 9.3) obtained by normal mode analyses [Stein and Okal, 2005; Park et al., 2005; Clévéde et al., 2012], which also depend on the assumed elastic structure.

The solution space (presented in map view in Figure 3 and as PDFs in Figures S5 and S6) is generally compatible with most published models, even though a straight comparison is difficult because of the different parameterizations of the fault geometry. We note that the location of the zones of largest slip zones (between latitudes 3°N and 4°N and latitudes 7°N and 8°N) are consistent with previous studies [e.g., Chlieh et al., 2007; Hoechner et al., 2008; Lorito et al., 2010]. Consistent with the models found by the studies including the pivot

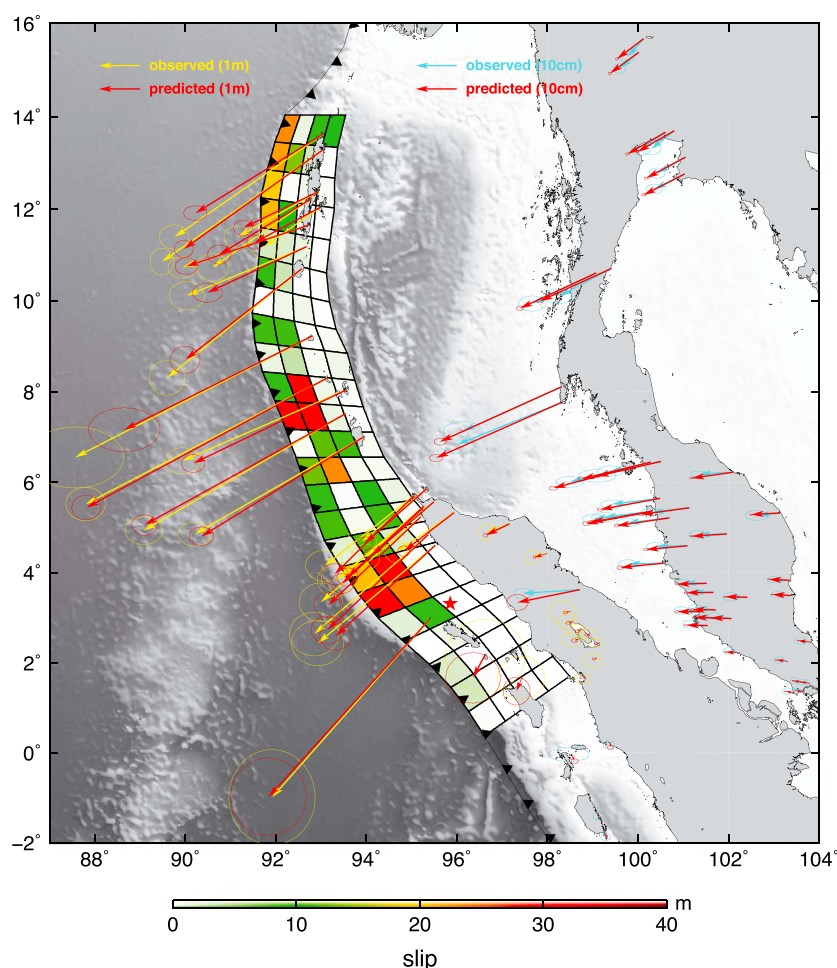


Figure 4. Near-field horizontal data fit associated with average slip model shown in Figure 3. Yellow arrows are near-field data, blue ones are farther-field data: they are plotted with different scales. Yellow ellipses are 2σ uncertainties derived from the diagonal terms of \mathbf{C}_x including data, prediction, and afterslip errors. Red vectors are offsets predicted by the average model (average slip at each subfault, independent to possible covariance with other subfaults) shown in Figure 3. Red ellipses are the 2σ uncertainties associated with the posterior covariances on the inverted parameters (prediction range for 95% of the models).

line determined by *Meltzner et al.* [2006], we find very low likelihood for any slip east of the line (Figures 3, S5, and S6). Our posterior models do not exclude the possibility for large slip reaching the trench. In fact, for some subfaults (s11d1, s17d1, and s18d1) the PDFs suggest over 20 m of shallow slip (Figure S6). Large shallow slip seems to be very likely as well in the northern part of the fault (subfaults s1-4d1). Moreover, we note, among the solutions, the possibility for slip over 50 m (see PDF of patch s18d1 on Figure S6). Actually, slip over 40 m is extremely likely (on patches s18d1, s11d1, or s11d2, Figure S6), even though our subfaults are relatively large. The value of maximum slip could even appear higher if data allowed us to solve for slip at smaller scales. This high likelihood for large slip is a remarkable feature because, before it started to be documented for the 2011 Tohoku-Oki earthquake [e.g., *Fujiwara et al.*, 2011; *Sato et al.*, 2011], such large slip amplitudes were thought to be very unlikely and slip inversions were made using strong smoothing regularization in order to exclude such solutions. It is possible that such high slip maxima are more common among giant earthquakes and that regularization strongly biases our knowledge of subduction earthquakes.

5.2. Data Fit

Near-field horizontal data are well explained by all models within the considered uncertainties (Figure 4). The red ellipses in Figure 4 show the prediction range for 95% (corresponding to 2σ standard deviations) of the slip models. We see that, in general, they intersect the data error ellipses (yellow and blue depending on the color scales) which account for all uncertainties contained in \mathbf{C}_x . Moreover, we note a systematic underprediction of

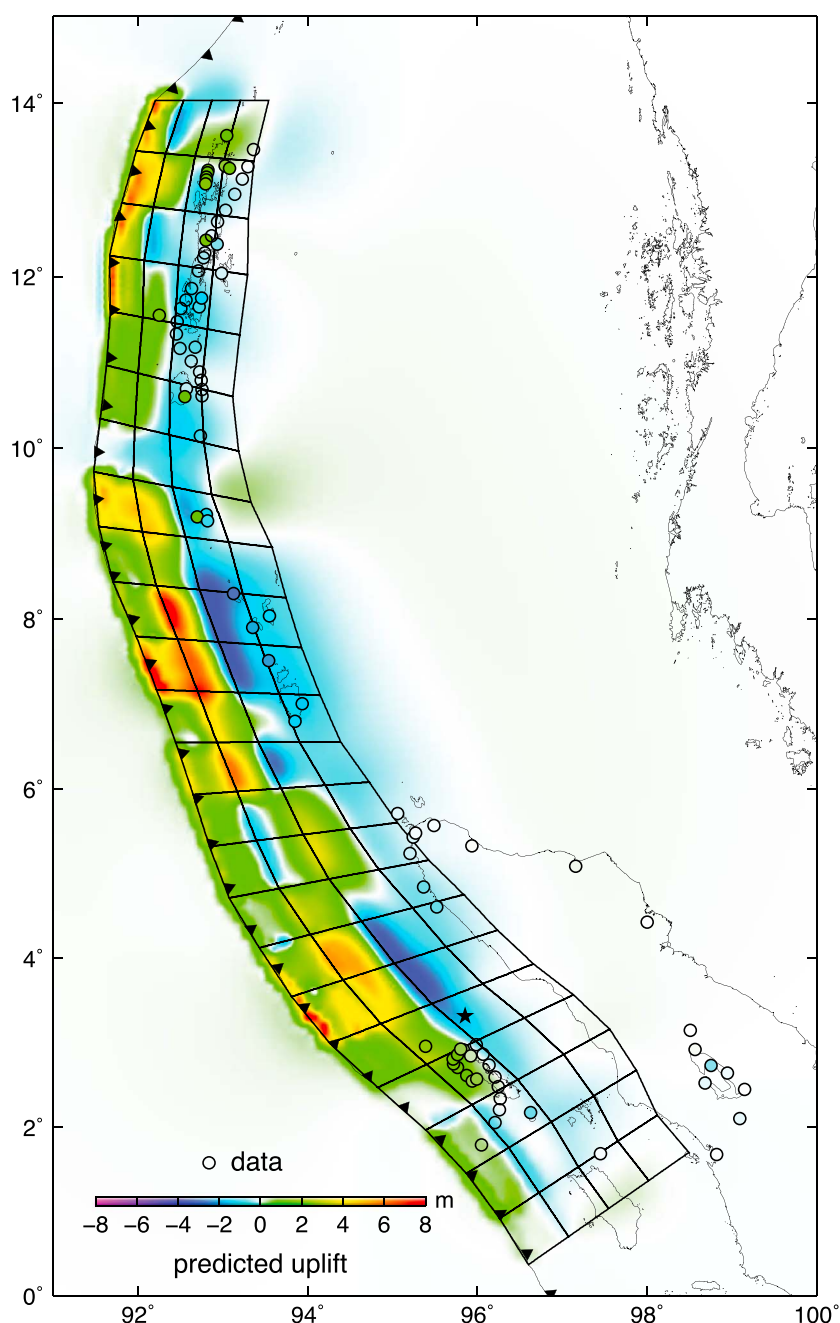


Figure 5. Fit of the near-field vertical data. Color map shows seafloor elevation change predicted by the average model. Colored dots show data.

near-field horizontal GPS data, which is consistent with probable contamination by postseismic processes. As we explicitly allow for spatially correlated misfit through our construction of $\mathbf{C}_{\text{post}^r}$, such coherent misfit shows the ability to avoid overfitting postseismic noise. We do not observe underprediction of the intermediate-field GPS data, which are derived from daily solutions and should therefore contain negligible postseismic signal.

Near-field vertical data are not as well explained (Figure 5). The likely reason for this misfit is the relatively large size of our patches compared to the data density, in particular beneath the Andaman islands (Figure 5). Also, the \mathbf{C}_{post} associated with the pivot line is particularly large because of the long postseismic contamination of the satellite images (see section 2.1). Far-field data are also well explained (Figure S7), although the most distant stations (Figure S8) seem to have an insufficient signal-to-noise ratio to bring any relevant constraint on the slip distribution.

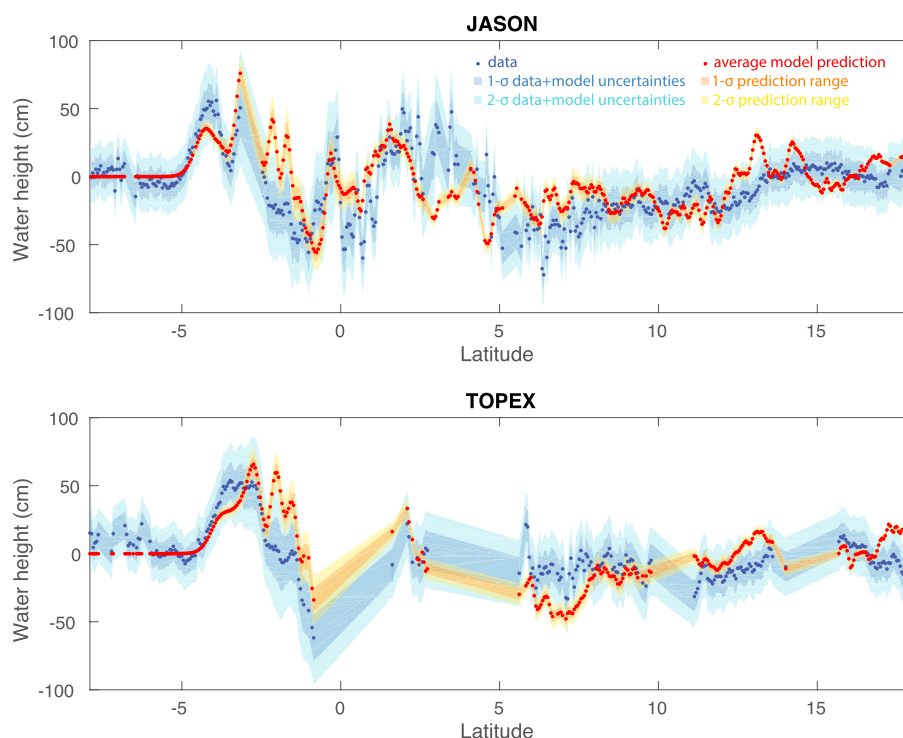


Figure 6. Altimetry records of the tsunami (blue dots) and average model prediction (red dots) assuming a rupture velocity of 2.8 km/s. Blue and light blue shaded areas show 1- σ and 2- σ , respectively, data and model uncertainties. Orange and yellow shades are 1- σ and 2- σ prediction ranges, respectively. Note that uncertainties associated with the off-diagonal terms of both \mathbf{C}_d and \mathbf{C}_{time} are not represented in this figure even though they probably affect the solution space as much as the diagonal terms (see Figure S3).

The prediction of the sea surface height anomaly by our average model along the tracks of the satellite altimeters (Figure 6) is consistent with the large uncertainties we have adopted both on the data and on the Green's functions. Vertical offsets reflect the variance (diagonal terms of \mathbf{C}_x) we imposed on the data and the modeling (represented by the 1 σ and 2 σ blue shaded areas in Figure 6). All models predict similar altimetric profiles: the 2 σ model uncertainty range (represented by the red shaded area) is narrow. This narrow prediction range likely comes from the constraints brought by the geodetic data and the off-diagonal terms of \mathbf{C}_x (Figure S3) which are on the same order than the diagonal terms and, therefore, significantly reduce the solution space. The prediction of the average model can be far from the data (as far as the variance allows), but all models must have large covariances, meaning that they are all in the same solution space region and predict similar fits. When neglecting \mathbf{C}_{post} and \mathbf{C}_{time} , the solution space is much reduced—the PDFs obtained (Figures S9 and S10) are much narrower than in Figures S6 and S7—while when neglecting the off-diagonal terms of \mathbf{C}_x , the solution space is extended (the PDFs in Figures S11 and S12 are broader than in Figures S5 and S6). This last observation illustrates that as discussed in paragraph 4.5, the off-diagonal terms of the covariance contain information that helps constraining the solution.

6. Discussion

6.1. Comparison With Independent Seismic Observations

To further interpret our set of posterior models, we consider several independent seismic observations from different phases of the seismic cycle: seismicity prior to the earthquake, high-frequency radiation during the earthquake, and the distribution of aftershocks. In Figure 7, we show seismicity located by the International Seismological Centre during a 44 year period preceding the earthquake (www.isc.ac.uk). The magnitude of the largest event recorded during this period is M_w 6.7, contrasting with the magnitude of the main shock (M_w 9.25) and its aftershocks (up to M_w 8.6). The spatial distribution of this seismicity clearly delineates the contours of the coseismic slip area (which is well defined in the posterior solutions due to the pivot line constraints). In 44 years, almost no earthquake (the catalog magnitude completeness is 4.0) was recorded in the 1600 km long 120 km large rupture area (Figure 7). This observation leads to the idea that the megathrust interface

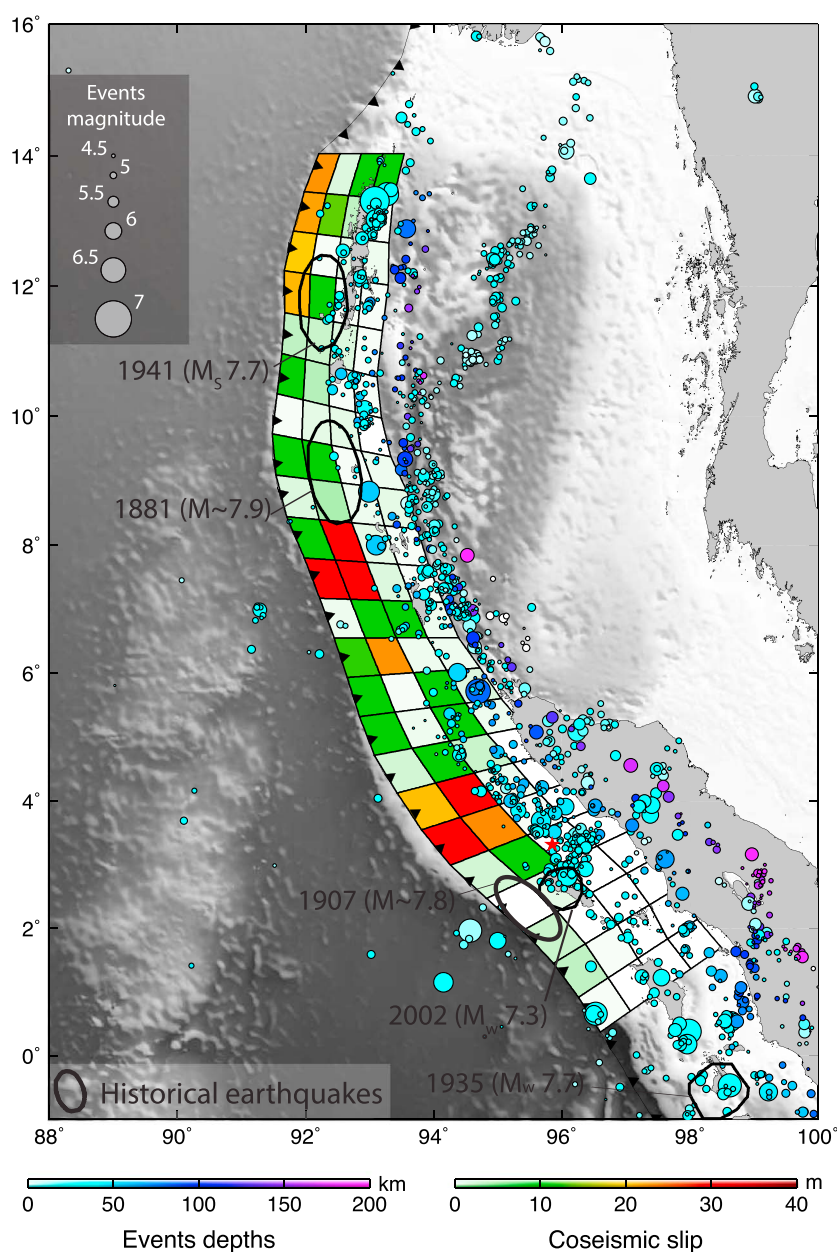


Figure 7. Comparison of our average slip model with regional seismicity before the Sumatra-Andaman earthquake starting from 1960 (International Seismological Center catalog). Contours of historical earthquakes are from *Briggs et al.* [2006] and *Kanamori et al.* [2010].

was completely locked during the interseismic period. However, at longer time scales, we see that historical earthquakes ($M \approx 7.9$ in 1881 and $M_s = 7.7$ in 1941 [Briggs et al., 2006, and references therein]) did occur in the coseismic slip zone beneath the Andamans (Figure 7) suggesting that stress accumulation was heterogeneous in this particular region.

Backprojection methods are useful to image high-frequency seismic radiation and complement finite-fault slip inversions which are sensitive to much lower frequency bands. Unfortunately, the resolution of backprojection models decreases with the distance to the epicenter—because imaging of late radiators relies on the late part of seismograms which are affected by many possible sources (e.g., earlier reflected waves)—but they provide valuable information on the southern half of the rupture. Looking at areas of high-frequency (>0.24 Hz) seismic radiation imaged by backprojection [Yao et al., 2011], we see that the large coseismic slip zones tend to be located just updip of the most intense short-period seismic radiators (Figure 8). This trend

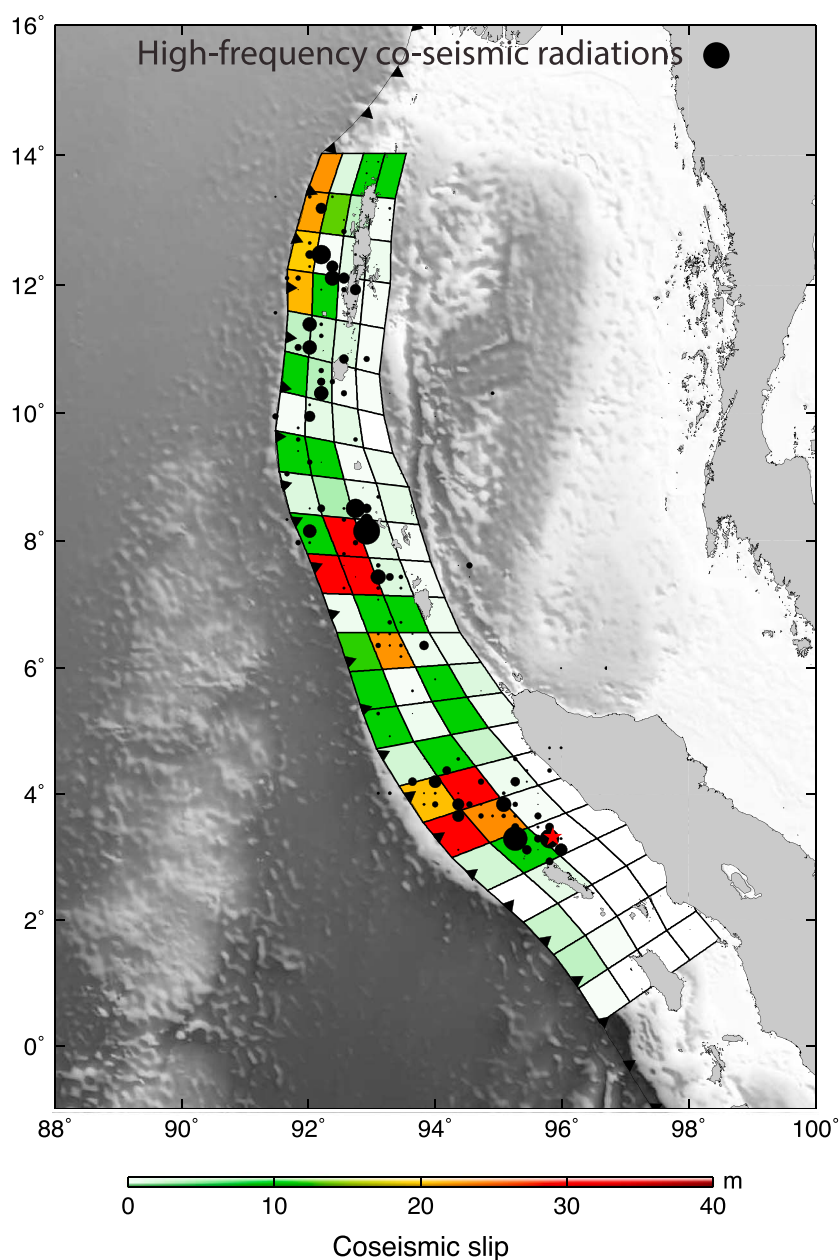


Figure 8. Comparison of our average slip model with coseismic high-frequency (>0.24 Hz) radiation (red circles) imaged by backprojection [Yao *et al.*, 2011].

was already highlighted by several studies at the large scale for several mega-earthquakes [Simons *et al.*, 2011; Meng *et al.*, 2011; Koper *et al.*, 2011; Lay *et al.*, 2012] and is here clear at a smaller scale for our two largest slip zones at latitudes 3°N – 4°N and 7°N – 8°N . This high-frequency radiation is thought to be due to sudden changes in rupture speed along sharp contrasts of fault rheology or geometry or along remnant stress concentrations from previous earthquakes [Simons *et al.*, 2011] and to delineate the downdip edge of the slip area [Simons *et al.*, 2011; Okuwaki *et al.*, 2014; Avouac *et al.*, 2015]. The spatial relationship we find in the case of the Sumatra-Andaman earthquake (Figure 8) strengthens the idea that the large slip areas we image correspond to locked asperities loaded in stress during the interseismic phase, further implying that the stress accumulation was not homogeneous even though pre-earthquake seismicity seems to indicate the fault was completely locked (Figure 7). This apparent contradiction could be solved by invoking a stress shadow effect in the fault areas surrounding strongly locked asperities [Hetland *et al.*, 2010; Hetland and Simons, 2010; Kanda *et al.*, 2013].

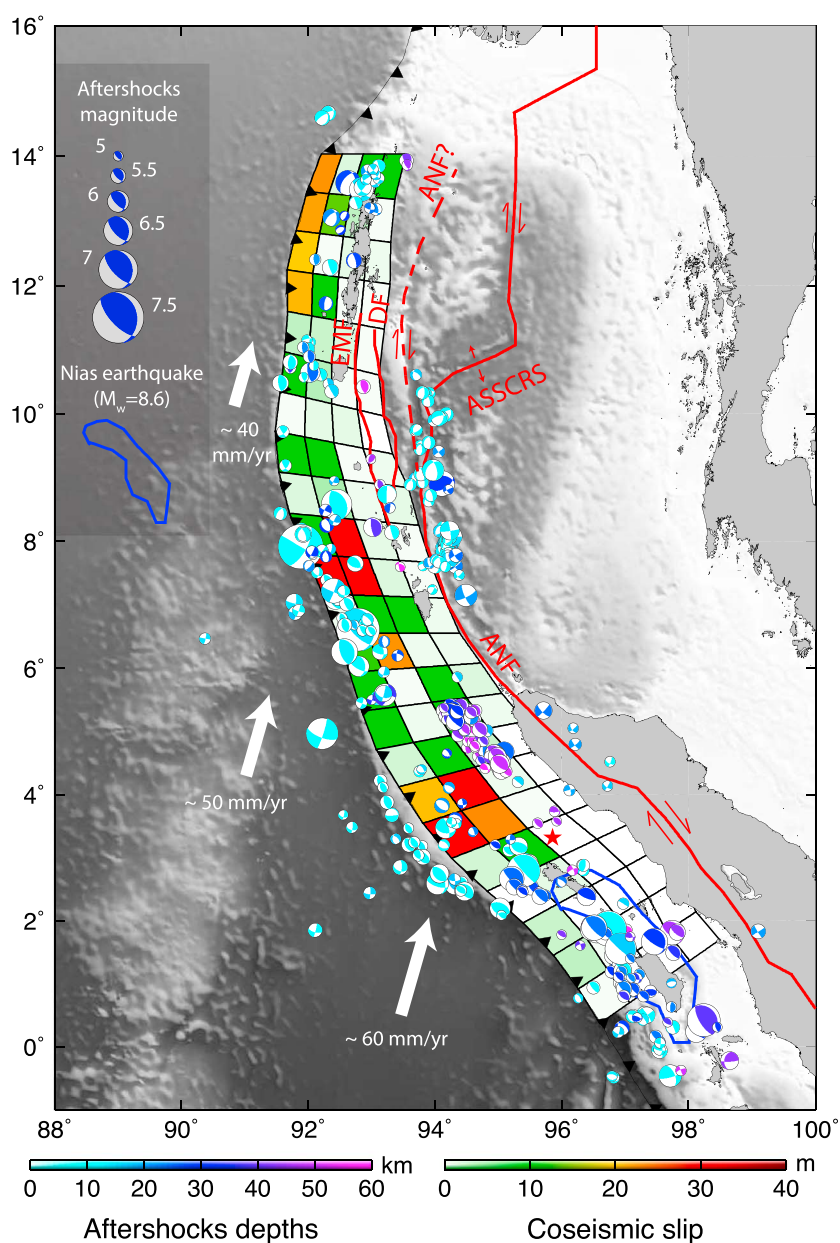


Figure 9. Comparison of our average slip model with aftershocks distribution during the 1 year period following the main shock (Global Centroid Moment Tensor catalog). Red lines show major faults in the hanging wall: the Andaman-Nicobar Fault (ANF), the Diligent Fault (DF), the Eastern Margin Fault (EMF) [Moeremans and Singh, 2015], and the Andaman Sea Spreading Center Rift System (ASSCRS) [Singh et al., 2013].

After large earthquakes, aftershocks are often observed in the vicinity of high slip patches [e.g., Hsu et al., 2006; Sladen et al., 2010]. The Sumatra-Andaman earthquake was followed by many aftershocks in the broad vicinity of the high slip area we image (Figure 9). If we ignore the 2005 M_w 8.6 Nias earthquake sequence to the South and the seismicity associated with the Andaman-Nicobar Fault (ANF) at the eastern end of the rupture, in the fore-arc basin (between latitudes 7°N and 11°N), aftershocks essentially correspond to the edges of the regions where we image significant slip. Aftershocks occurred intensively updip of the two high slip zones (3°N–4°N and 7°N–8°N), fewer at the bottom portion. Except for a swarm of dip-slip aftershocks northeast of the southern high slip zone (3°N–4°N), which seems to correspond to a backthrust fault [Cochran, 2010], we see very few aftershocks at the downdip edges of our large slip areas. We observe very few outerrise aftershocks in the subducting plate. According to the model proposed by Lay et al. [2009], this observation suggests that the rupture did not reach the surface. Our shallowest subfaults are large ($\approx 65 \text{ km} \times 65 \text{ km}$

on average) and the fact that some have average slip values of several tens of meters does not necessarily imply that rupture reached the surface. North of 10°N , we see moderate aftershocks ($M_w \leq 6.3$) activity mostly located downdip of the main slip area (Figure 9). Between latitudes 10°N and 11°N , focal mechanisms and depths are compatible with events occurring on the megathrust interface. This kind of aftershock distribution usually indicates the limits of the coseismic slip (e.g., Kato and Igarashi [2012] or Bletery et al. [2014], in the case of the M_w 9.0 Tohoku-Oki earthquake). These northern thrust events (patch line s6) are located just north of a section of the rupture where PDFs indicate very low likelihood for any slip (see subfaults line s7, Figure S6). North of this area is also where the GPS displacements start to change direction, approaching the highly oblique plate convergence direction, further suggesting that 10°N marks a radical change in style of rupture.

6.2. Oblique Slow Slip Beneath the Andaman Islands?

What happened in the northern part (defined above as north of line s7, latitude 10°N) of the rupture is still an open question. First, we note the reduced convergence rate between the Australian and the Sunda plates north of 8°N [Prawirodirdjo and Bock, 2004; Gahalaut et al., 2006] and the highly oblique convergence direction (11° from trench parallel) [Paul et al., 2001; Gahalaut et al., 2006; DeMets et al., 2010]. Between 8°N and 10°N , the obliquity is accommodated by the ANF, but north of 10°N , the ANF branches onto the Andaman Sea Spreading Centre rift system [Singh et al., 2013] and becomes distant from the subduction zone by several hundreds of kilometers (Figure 9). Therefore, it is unclear how the strike-slip component of the convergence is accommodated north of 10°N and to what extent there is partitioning across the fore arc. There are two large faults along the Andamans, the Diligent Fault (DF), and the Eastern Margin Fault (EMF). The DF seems to be in compression and too far east to accommodate partitioning [Cochran, 2010; Moeremans and Singh, 2015], while the EMF seems to be dominated by extension [Cochran, 2010] and not very active [Moeremans and Singh, 2015]. However, some studies extend the ANF further north [Cochran, 2010; Wang et al., 2014], making the ANF the best candidate to accommodate partitioning.

Our solution indicates that coseismic slip contains a strong strike-slip component just beneath the Andamans (see Figure 3 and posterior PDFs of subfaults s1–6d3 in Figure S5), which is well constrained by the systematic oblique signal in the Andaman GPS data (Figure 4). This coseismic strike-slip motion is followed by aftershocks presenting strike-slip mechanisms (Figure 9). Strike-slip aftershocks were, for instance, particularly intense near subfault s1d3 (the patch located beneath the northern tip of the Andaman archipelago) where we image the largest strike-slip motion (Figure 9). The depths of aftershocks indicate a probable superficial strike-slip fault, although the systematic oblique direction of all the Andaman GPS records (Figure 4) favors the idea of a deeper and large-scale coseismic oblique motion on the megathrust. To explain the obliquity of the GPS displacements, a strike-slip fault would need to move coherently over the 440 km of Andaman archipelago. The northern extension of the ANF [Wang et al., 2014] appears as the best candidate even though it does not appear in the relocated seismicity [Diehl et al., 2013]. Thus, either partitioning is accommodated by the ANF but it is not associated with any aftershock (Figure 9), or there is no partitioning on the coseismic slip, as documented in other subduction settings for coseismic or postseismic slip distributions [e.g., Pritchard et al., 2002; Sladen et al., 2010; Perfettini et al., 2010; Lin et al., 2013]. Yet to our knowledge, there is no documented case of large coseismic rupture occurring in a subduction zone with such an oblique convergence. Other aftershocks in the Andaman segment are characterized by inverse and even normal mechanisms, confirming the complexity of the regional seismic activity.

The anomalously large low-frequency (periods > 1000 s) seismic radiation observed in normal mode analyses led several authors to propose a scenario with a large component of slow slip in the Andaman section [Stein and Okal, 2005; Park et al., 2005; Clévéde et al., 2012]. These studies inferred significantly larger moments at low frequencies (periods of 600–3000 s) than in the standard frequency band used to analyze surface waves (periods of 300–500 s). They further found a northward migration of the centroid when lowering the frequency band, suggesting that the source of low-frequency radiation is in the northern part of the fault. A northern slow slip scenario was proposed by Lay et al. [2005] to explain the late signal in the only near-field tide gauge at Port Blair, in the Andamans (Port Blair location is indicated by the position of the GPS station PB in Figure 1). The delay in the signal was then shown to be primarily an artifact of the station clock drift, but once corrected, it still suggested slow rise time on the order of 30 min [Singh et al., 2006], consistent with normal mode analyses. The slow slip beneath the Andamans is also compatible with the amplitude of the geodetic records (Figure 4) which suggests that large moment was released in this portion.

The reliability of our inferred slip in the northern portion of the model (s1-6d1-4) might be questioned if we believe in the proposed slow slip scenario. The tsunami generated by a very slow rupture would be of lower amplitude than the one we simulated, leading to inconsistency in the tsunami Green's functions. Nevertheless, two arguments can be made to strengthen the consistency of the tsunami Green's functions. (1) The deep subfaults (s1-6d3-4) are actually constrained by geodetic data, which are not affected by the kinematics of the rupture (see sensitivity test in Figure S13). (2) The seismic radiators imaged by backprojections [Ni *et al.*, 2005; Ishii *et al.*, 2005; Krüger and Ohrnberger, 2005; Yao *et al.*, 2011] seem to indicate that at least a significant part of the slip was rapid (which is also what studies advocating for the slow slip scenario proposed [Lay *et al.*, 2005; Singh *et al.*, 2006]). Thus, while the standard deviations associated with the northern deep slip (s1-6d3-4) should be consistent with the true uncertainties, the uncertainties on the shallower subfaults (s1-6d1-2) are likely larger than our posterior PDFs estimates. For instance, based on the slip orientation of the subfaults beneath the Andamans (s1-6d3), one may reasonably think that the strike-slip component—and its associated uncertainty—is underestimated in subfaults s1-6d1-2 because if there is no partitioning in the Andaman section, the slip orientation should be relatively homogeneous and close to the convergence direction in the whole section.

7. Conclusion

The solutions of finite-fault slip inversions for the slip distribution of the 2004 Sumatra-Andaman earthquake show large variability depending on the inverted data sets. This variability likely comes from relative inconsistencies among the data, which are contaminated by different sources of errors. In an attempt to narrow down the range of possible source models for this earthquake, we considered an extended set of geodetic and tsunami observations in a joint inversion to image the distribution of fault slip. We accounted, as best as possible, for various sources of correlated uncertainties, including contamination due to afterslip, Earth structure inaccuracy, and errors in rupture velocity estimation. We explored the whole posterior solution space of the inverse problem which can be seen in PDF form (Figures S5 and S6) or in the form of its mean model with its associated uncertainties (Figure 3). This solution space reveals a well constrained moment magnitude of $M_w = 9.25$ as well as high likelihood for very large values of slip (40 m or more) on certain parts of the fault and does not exclude even larger slip maxima. Such high slip maxima might actually be common among giant earthquakes, but our past models may simply be biased by commonly applied regularization in the finite-fault modeling.

Our solution indicates two large slip areas (at latitudes 3°N–4°N and 7°N–8°N) that might have reached the surface. Identification of such large slip areas provides valuable information to understand the seismic cycle and the physics of the rupture, but more pragmatically, it is also what controls the size of the tsunami, particularly in the near field [e.g., Shimozono *et al.*, 2012; Yamazaki *et al.*, 2013]. We compare the average slip distribution we obtained with independent seismological records. The comparison shows clear correlations between likely large slip areas and high-frequency seismic radiators as well as with intense aftershock activity. These spatial relationships strengthen the idea that the event is dominated by the rupture of strongly locked asperities causing large slip both at the place they ruptured and in the surrounding fault portions.

Acknowledgments

This work was granted access to the HPC and visualization resources of Centre de Calcul Interactif hosted by Université Nice Sophia Antipolis and of the CITerra/FRAM cluster hosted by the California Institute of Technology. The study was partly supported by the French National Research Agency (ANR) EPOST project ANR-14-CE03-0002, the French Ministry of Research and Higher Education, the Université Nice Sophia Antipolis, the Centre National de la Recherche Scientifique (CNRS), and the California Institute of Technology. We are grateful to Y. Yamazaki and F. Pollitz for providing their respective codes NEOWAVE and STATIC1D, to H. Yao for providing material used in Figure 8, and to Z. Duputel, L. Rivera, J.-M. Nocquet, and B. Delouis as well as L. Feng, Y. Ben Zion, and an anonymous reviewer for their valuable comments. The ISC seismicity catalog shown in Figure 7 is available online (www.isc.ac.uk).

References

- Ablain, M., J. Dorandeu, P.-Y. Le Traon, and A. Sladen (2006), High resolution altimetry reveals new characteristics of the December 2004 Indian Ocean tsunami, *Geophys. Res. Lett.*, 33, L21602, doi:10.1029/2006GL027533.
- Ammon, C. J., et al. (2005), Rupture process of the 2004 Sumatra-Andaman earthquake, *Science*, 308(5725), 1133–1139.
- Avouac, J.-P., L. Meng, S. Wei, T. Wang, and J.-P. Ampuero (2015), Unzipping lower edge of locked Main Himalayan Thrust during the 2015, M_w 7.8 Gorkha earthquake, Nepal, *Nat. Geosci.*, 8, 708–711, doi:10.1038/ngeo2518.
- Banerjee, P., F. F. Pollitz, and R. Bürgmann (2005), The size and duration of the Sumatra-Andaman earthquake from far-field static offsets, *Science*, 308(5729), 1769–1772.
- Banerjee, P., F. Pollitz, B. Nagarajan, and R. Bürgmann (2007), Coseismic slip distributions of the 26 December 2004 Sumatra-Andaman and 28 March 2005 Nias earthquakes from GPS static offsets, *Bull. Seismol. Soc. Am.*, 97(1A), S86–S102.
- Bilham, R., R. Engdahl, N. Fiedl, and S. Satyabala (2005), Partial and complete rupture of the Indo-Andaman plate boundary 1847–2004, *Seismol. Res. Lett.*, 76(3), 299–311.
- Bletery, Q., A. Sladen, B. Delouis, M. Vallée, J.-M. Nocquet, L. Rolland, and J. Jiang (2014), A detailed source model for the M_w 9.0 Tohoku-Oki earthquake reconciling geodesy, seismology, and tsunami records, *J. Geophys. Res. Solid Earth*, 119, 7636–7653, doi:10.1002/2014JB011261.
- Bletery, Q., A. Sladen, B. Delouis, and L. Mattéo (2015), Quantification of tsunami bathymetry effect on finite fault slip inversion, *Pure Appl. Geophys.*, 172(12), 3655–3670.
- Briggs, R. W., et al. (2006), Deformation and slip along the Sunda megathrust in the great 2005 Nias-Simeulue earthquake, *Science*, 311(5769), 1897–1901.

- Catherine, J. K., V. K. Gahalaut, and V. K. Sahu (2005), Constraints on rupture of the December 26, 2004, Sumatra earthquake from far-field GPS observations, *Earth Planet. Sci. Lett.*, *237*(3), 673–679.
- Chlieh, M., et al. (2007), Coseismic slip and afterslip of the great M_w 9.15 Sumatra-Andaman earthquake of 2004, *Bull. Seismol. Soc. Am.*, *97*(1A), S152–S173.
- Clévéde, E., B. Bukchin, P. Favreau, A. Mostinskiy, A. Aoudia, and G. Panza (2012), Long-period spectral features of the Sumatra-Andaman 2004 earthquake rupture process, *Geophys. J. Int.*, *191*(3), 1215–1225.
- Cochran, J. R. (2010), Morphology and tectonics of the Andaman forearc, northeastern Indian ocean, *Geophys. J. Int.*, *182*(2), 631–651.
- DeMets, C., R. G. Gordon, and D. F. Argus (2010), Geologically current plate motions, *Geophys. J. Int.*, *181*(1), 1–80.
- Diehl, T., F. Waldhauser, J. R. Cochran, K. Kamesh Raju, L. Seeber, D. Schaff, and E. Engdahl (2013), Back-arc extension in the Andaman sea: Tectonic and magmatic processes imaged by high-precision teleseismic double-difference earthquake relocation, *J. Geophys. Res. Solid Earth*, *118*, 2206–2224, doi:10.1002/jgrb.50192.
- Duputel, Z., P. S. Agram, M. Simons, S. E. Minson, and J. L. Beck (2014), Accounting for prediction uncertainty when inferring subsurface fault slip, *Geophys. J. Int.*, *197*, 464–482.
- Duputel, Z., et al. (2015), The Iquique earthquake sequence of April 2014: Bayesian modeling accounting for prediction uncertainty, *Geophys. Res. Lett.*, *42*(19), 7949–7957.
- Dziewonski, A. M., and D. L. Anderson (1981), Preliminary reference earth model, *Phys. Earth Planet. Inter.*, *25*(4), 297–356.
- Fritz, H. M., and J. C. Borrero (2006), Somalia field survey after the December 2004 Indian Ocean tsunami, *Earthquake Spectra*, *22*(S3), 219–233.
- Fujii, Y., and K. Satake (2007), Tsunami source of the 2004 Sumatra-Andaman earthquake inferred from tide gauge and satellite data, *Bull. Seismol. Soc. Am.*, *97*(1A), S192–S207.
- Fujiwara, T., S. Kodaira, T. No, Y. Kaiho, N. Takahashi, and Y. Kaneda (2011), The 2011 Tohoku-Oki earthquake: Displacement reaching the trench axis, *Science*, *334*(6060), 1240–1240.
- Gahalaut, V., B. Nagarajan, J. Catherine, and S. Kumar (2006), Constraints on 2004 Sumatra-Andaman earthquake rupture from GPS measurements in Andaman-Nicobar Islands, *Earth Planet. Sci. Lett.*, *242*(3), 365–374.
- Gahalaut, V., et al. (2008), GPS measurements of postseismic deformation in the Andaman-Nicobar region following the giant 2004 Sumatra-Andaman earthquake, *J. Geophys. Res.*, *113*, B08401, doi:10.1029/2007JB005511.
- Goff, J., et al. (2006), Sri Lanka field survey after the December 2004 Indian Ocean tsunami, *Earthquake Spectra*, *22*(S3), 155–172.
- Gunawan, E., et al. (2014), A comprehensive model of postseismic deformation of the 2004 Sumatra-Andaman earthquake deduced from GPS observations in northern Sumatra, *J. Asian Earth Sci.*, *88*, 218–229.
- Hashimoto, M., N. Choosakul, M. Hashizume, S. Takemoto, H. Takiguchi, Y. Fukuda, and K. Fujimori (2006), Crustal deformations associated with the great Sumatra-Andaman earthquake deduced from continuous GPS observation, *Earth Planets Space*, *58*(2), 127–139.
- Hayes, G. P., D. J. Wald, and R. L. Johnson (2012), Slab1.0: A three-dimensional model of global subduction zone geometries, *J. Geophys. Res.*, *117*, B01302, doi:10.1029/2011JB008524.
- Hébert, H., A. Sladen, and F. Schindelé (2007), Numerical modeling of the great 2004 Indian Ocean tsunami: Focus on the Mascarene Islands, *Bull. Seismol. Soc. Am.*, *97*(1A), S208–S222.
- Hetland, E., and M. Simons (2010), Post-seismic and interseismic fault creep. II: Transient creep and interseismic stress shadows on megathrusts, *Geophys. J. Int.*, *181*(1), 99–112.
- Hetland, E., M. Simons, and E. Dunham (2010), Post-seismic and interseismic fault creep. I: Model description, *Geophys. J. Int.*, *181*(1), 81–98.
- Hirata, K., K. Satake, Y. Tanioka, T. Kuragano, Y. Hasegawa, Y. Hayashi, and N. Hamada (2006), The 2004 Indian Ocean tsunami: Tsunami source model from satellite altimetry, *Earth Planets Space*, *58*(2), 195–201.
- Hoechner, A., A. Y. Babeyko, and S. V. Sobolev (2008), Enhanced GPS inversion technique applied to the 2004 Sumatra earthquake and tsunami, *Geophys. Res. Lett.*, *35*, L08310, doi:10.1029/2007GL033133.
- Hsu, Y.-J., M. Simons, J.-P. Avouac, J. Galetzka, K. Sieh, M. Chlieh, D. Natawidjaja, L. Prawirodirdjo, and Y. Bock (2006), Frictional afterslip following the 2005 Nias-Simeulue earthquake, Sumatra, *Science*, *312*(5782), 1921–1926.
- Ishii, M., P. M. Shearer, H. Houston, and J. E. Vidale (2005), Extent, duration and speed of the 2004 Sumatra-Andaman earthquake imaged by the hi-net array, *Nature*, *435*(7044), 933–936.
- Ishii, M., P. M. Shearer, H. Houston, and J. E. Vidale (2007), Teleseismic P wave imaging of the 26 December 2004 Sumatra-Andaman and 28 March 2005 Sumatra earthquake ruptures using the Hi-net array, *J. Geophys. Res.*, *112*, B11307, doi:10.1029/2006JB004700.
- Jade, S., M. Ananda, P. D. Kumar, and S. Banerjee (2005), Co-seismic and post-seismic displacements in Andaman and Nicobar islands from GPS measurements, *Curr. Sci.*, *88*(12), 1980–1984.
- Jolivet, R., et al. (2014), The 2013 M_w 7.7 Balochistan earthquake: Seismic potential of an accretionary wedge, *Bull. Seismol. Soc. Am.*, *104*(2), 1020–1030.
- Kanamori, H., L. Rivera, and W. H. Lee (2010), Historical seismograms for unravelling a mysterious earthquake: The 1907 Sumatra earthquake, *Geophys. J. Int.*, *183*(1), 358–374.
- Kanda, R. V., E. A. Hetland, and M. Simons (2013), An asperity model for fault creep and interseismic deformation in northeastern Japan, *Geophys. J. Int.*, *192*(1), 38–57.
- Kato, A., and T. Igarashi (2012), Regional extent of the large coseismic slip zone of the 2011 M_w 9.0 Tohoku-Oki earthquake delineated by on-fault aftershocks, *Geophys. Res. Lett.*, *39*, L15301, doi:10.1029/2012GL052220.
- Koper, K., A. Hutko, and T. Lay (2011), Along-dip variation of teleseismic short-period radiation from the 11 March 2011 Tohoku earthquake (M_w 9.0), *Geophys. Res. Lett.*, *38*, L21309, doi:10.1029/2011GL049689.
- Kreemer, C., G. Blewitt, W. C. Hammond, and H.-P. Plag (2006), Global deformation from the great 2004 Sumatra-Andaman earthquake observed by GPS: Implications for rupture process and global reference frame, *Earth Planets Space*, *58*(2), 141–148.
- Krüger, F., and M. Ohrnberger (2005), Tracking the rupture of the M_w 9.3 Sumatra earthquake over 1,150 km at teleseismic distance, *Nature*, *435*(7044), 937–939.
- Lay, T., et al. (2005), The great Sumatra-Andaman earthquake of 26 December 2004, *Science*, *308*(5725), 1127–1133.
- Lay, T., H. Kanamori, C. J. Ammon, A. R. Hutko, K. Furlong, and L. Rivera (2009), The 2006–2007 Kuril Islands great earthquake sequence, *J. Geophys. Res.*, *114*, B11308, doi:10.1029/2008JB006280.
- Lay, T., H. Kanamori, C. J. Ammon, K. D. Koper, A. R. Hutko, L. Ye, H. Yue, and T. M. Rushing (2012), Depth-varying rupture properties of subduction zone megathrust faults, *J. Geophys. Res.*, *117*, B04311, doi:10.1029/2011JB009133.
- Lin, Y. N., et al. (2013), Coseismic and postseismic slip associated with the 2010 Maule earthquake, Chile: Characterizing the Arauco peninsula barrier effect, *J. Geophys. Res. Solid Earth*, *118*, 3142–3159, doi:10.1002/jgrb.50207.

- Lorito, S., A. Piatanesi, V. Cannelli, F. Romano, and D. Melini (2010), Kinematics and source zone properties of the 2004 Sumatra-Andaman earthquake and tsunami: Nonlinear joint inversion of tide gauge, satellite altimetry, and GPS data, *J. Geophys. Res.*, **115**, B02304, doi:10.1029/2008JB005974.
- Meltzner, A. J., K. Sieh, M. Abrams, D. C. Agnew, K. W. Hudnut, J.-P. Avouac, and D. H. Natawidjaja (2006), Uplift and subsidence associated with the great Aceh-Andaman earthquake of 2004, *J. Geophys. Res.*, **111**, B02407, doi:10.1029/2005JB003891.
- Meng, L., A. Inbal, and J.-P. Ampuero (2011), A window into the complexity of the dynamic rupture of the 2011 M_w 9 Tohoku-Oki earthquake, *Geophys. Res. Lett.*, **38**, L00G07, doi:10.1029/2011GL048118.
- Merrifield, M., et al. (2005), Tide gauge observations of the Indian Ocean tsunami, December 26, 2004, *Geophys. Res. Lett.*, **32**, L09603, doi:10.1029/2005GL022610.
- Minson, S., M. Simons, and J. Beck (2013), Bayesian inversion for finite fault earthquake source models. I—Theory and algorithm, *Geophys. J. Int.*, **194**(3), 1701–1726.
- Minson, S., M. Simons, J. Beck, F. Ortega, J. Jiang, S. Owen, A. Moore, A. Inbal, and A. Sladen (2014), Bayesian inversion for finite fault earthquake source models—II: The 2011 great Tohoku-Oki, Japan earthquake, *Geophys. J. Int.*, **198**(2), 922–940.
- Moeremans, R. E., and S. C. Singh (2015), Fore-arc basin deformation in the Andaman-Nicobar segment of the Sumatra-Andaman subduction zone: Insight from high-resolution seismic reflection data, *Tectonics*, **34**, 1736–1750, doi:10.1002/2015TC003901.
- Ni, S., H. Kanamori, and D. Helmberger (2005), Seismology: Energy radiation from the Sumatra earthquake, *Nature*, **434**(7033), 582–582.
- Okal, E. A., H. M. Fritz, P. E. Raad, C. Synolakis, Y. Al-Shijbi, and M. Al-Saifi (2006), Oman field survey after the December 2004 Indian Ocean tsunami, *Earthquake Spectra*, **22**(S3), 203–218.
- Okuwaki, R., Y. Yagi, and S. Hirano (2014), Relationship between high-frequency radiation and asperity ruptures, revealed by hybrid back-projection with a non-planar fault model, *Sci. Rep.*, **4**, 7120.
- Park, J., et al. (2005), Earth's free oscillations excited by the 26 december 2004 Sumatra-Andaman earthquake, *Science*, **308**(5725), 1139–1144.
- Paul, J., et al. (2001), The motion and active deformation of India, *Geophys. Res. Lett.*, **28**(4), 647–650.
- Paul, J., A. Lowry, R. Bilham, S. Sen, and R. Smalley (2007), Postseismic deformation of the Andaman islands following the 26 December, 2004 great Sumatra-Andaman earthquake, *Geophys. Res. Lett.*, **34**, L19309, doi:10.1029/2007GL031024.
- Paul, J., C. Rajendran, A. Lowry, V. Andrade, and K. Rajendran (2012), Andaman postseismic deformation observations: Still slipping after all these years?, *Bull. Seismol. Soc. Am.*, **102**(1), 343–351.
- Perfettini, H., et al. (2010), Seismic and aseismic slip on the Central Peru megathrust, *Nature*, **465**(7294), 78–81.
- Pesicek, J., C. Thurber, S. Widiyantoro, H. Zhang, H. DeShon, and E. Engdahl (2010), Sharpening the tomographic image of the subducting slab below Sumatra, the Andaman Islands and Burma, *Geophys. J. Int.*, **182**(1), 433–453.
- Piatanesi, A., and S. Lorito (2007), Rupture process of the 2004 Sumatra-Andaman earthquake from tsunami waveform inversion, *Bull. Seismol. Soc. Am.*, **97**(1A), S223–S231.
- Pietrzak, J., A. Socquet, D. Ham, W. Simons, C. Vigny, R. J. Labeur, E. Schrama, G. Stelling, and D. Vatvani (2007), Defining the source region of the Indian Ocean tsunami from GPS, altimeters, tide gauges and tsunami models, *Earth Planet. Sci. Lett.*, **261**(1), 49–64.
- Poisson, B., C. Oliveros, and R. Pedreros (2011), Is there a best source model of the Sumatra 2004 earthquake for simulating the consecutive tsunami?, *Geophys. J. Int.*, **185**(3), 1365–1378.
- Pollitz, F. F. (1996), Coseismic deformation from earthquake faulting on a layered spherical earth, *Geophys. J. Int.*, **125**(1), 1–14.
- Prawirodirdjo, L., and Y. Bock (2004), Instantaneous global plate motion model from 12 years of continuous GPS observations, *J. Geophys. Res.*, **109**, B08405, doi:10.1029/2003JB002944.
- Pritchard, M., M. Simons, P. Rosen, S. Hensley, and F. Webb (2002), Co-seismic slip from the 1995 July 30 M_w 8.1 Antofagasta, Chile, earthquake as constrained by InSAR and GPS observations, *Geophys. J. Int.*, **150**(2), 362–376.
- Rhie, J., D. Dreger, R. Bürgmann, and B. Romanowicz (2007), Slip of the 2004 Sumatra-Andaman earthquake from joint inversion of long-period global seismic waveforms and GPS static offsets, *Bull. Seismol. Soc. Am.*, **97**(1A), S115–S127.
- Sato, M., T. Ishikawa, N. Ujihara, S. Yoshida, M. Fujita, M. Mochizuki, and A. Asada (2011), Displacement above the hypocenter of the 2011 Tohoku-Oki earthquake, *Science*, **332**(6036), 1395–1395.
- Shearer, P., and R. Bürgmann (2010), Lessons learned from the 2004 Sumatra-Andaman megathrust rupture, *Annu. Rev. Earth Planet. Sci.*, **38**(1), 103–131.
- Shimozono, T., S. Sato, A. Okayasu, Y. Tajima, H. M. Fritz, H. Liu, and T. Takagawa (2012), Propagation and inundation characteristics of the 2011 Tohoku tsunami on the central Sanriku coast, *Coastal Eng. J.*, **54**(1), 1250004.
- Simons, M., et al. (2011), The 2011 magnitude 9.0 Tohoku-Oki earthquake: Mosaicking the megathrust from seconds to centuries, *Science*, **332**(6036), 1421–1425.
- Singh, S., M. Ortiz, H. Gupta, and D. Ramadass (2006), Slow slip below Port Blair, Andaman, during the great Sumatra-Andaman earthquake of 26 December 2004, *Geophys. Res. Lett.*, **33**, L03313, doi:10.1029/2005GL025025.
- Singh, S. C., R. Moeremans, J. McArdle, and K. Johansen (2013), Seismic images of the sliver strike-slip fault and back thrust in the Andaman-Nicobar region, *J. Geophys. Res. Solid Earth*, **118**, 5208–5224, doi:10.1002/jgrb.50378.
- Sladen, A., and H. Hébert (2008), On the use of satellite altimetry to infer the earthquake rupture characteristics: Application to the 2004 Sumatra event, *Geophys. J. Int.*, **172**(2), 707–714.
- Sladen, A., et al. (2010), Source model of the 2007 M_w 8.0 Pisco, Peru earthquake: Implications for seismogenic behavior of subduction megathrusts, *J. Geophys. Res.*, **115**, B02405, doi:10.1029/2009JB006429.
- Smet, S., R. Michel, and L. Bollinger (2008), Uplift of the 2004 Sumatra-Andaman earthquake measured from differential hyperspectral imagery of coastal waters, *J. Geophys. Res.*, **113**, B09403, doi:10.1029/2007JB005317.
- Stein, S., and E. A. Okal (2005), Seismology: Speed and size of the sumatra earthquake, *Nature*, **434**(7033), 581–582.
- Subarya, C., M. Chlieh, L. Prawirodirdjo, J.-P. Avouac, Y. Bock, K. Sieh, A. J. Meltzner, D. H. Natawidjaja, and R. McCaffrey (2006), Plate-boundary deformation associated with the great Sumatra-Andaman earthquake, *Nature*, **440**(7080), 46–51.
- Tanioka, Y., et al. (2006), Rupture process of the 2004 great Sumatra-Andaman earthquake estimated from tsunami waveforms, *Earth Planets Space*, **58**(2), 203–209.
- Tarantola, A. (2005), *Inverse Problem Theory and Methods for Model Parameter Estimation*, Siam.
- Tsai, V. C., M. Nettles, G. Ekström, and A. M. Dziewonski (2005), Multiple CMT source analysis of the 2004 Sumatra earthquake, *Geophys. Res. Lett.*, **32**, L17304, doi:10.1029/2005GL023813.
- Vallée, M. (2007), Rupture properties of the giant sumatra earthquake imaged by Empirical Green's function analysis, *Bull. Seismol. Soc. Am.*, **97**(1A), S103–S114.
- Vigny, C., et al. (2005), Insight into the 2004 Sumatra-Andaman earthquake from GPS measurements in Southeast Asia, *Nature*, **436**(7048), 201–206.

- Wang, Y., K. Sieh, S. T. Tun, K.-Y. Lai, and T. Myint (2014), Active tectonics and earthquake potential of the Myanmar region, *J. Geophys. Res.*, *119*(4), 3767–3822, doi:10.1002/2013JB010762.
- Yagi, Y., and Y. Fukahata (2008), Importance of covariance components in inversion analyses of densely sampled observed data: An application to waveform data inversion for seismic source processes, *Geophys. J. Int.*, *175*(1), 215–221.
- Yamazaki, Y., K. F. Cheung, and Z. Kowalik (2011a), Depth-integrated, non-hydrostatic model with grid nesting for tsunami generation, propagation, and run-up, *Int. J. Numer. Methods Fluids*, *67*(12), 2081–2107.
- Yamazaki, Y., K. F. Cheung, and Z. Kowalik (2011b), Depth-integrated, non-hydrostatic model with grid nesting for tsunami generation, propagation, and run-up, *Int. J. Numer. Methods Fluids*, *67*(12), 2081–2107.
- Yamazaki, Y., K. F. Cheung, and T. Lay (2013), Modeling of the 2011 Tohoku near-field tsunami from finite-fault inversion of seismic waves, *Bull. Seismol. Soc. Am.*, *103*(2B), 1444–1455.
- Yao, H., P. Gerstoft, P. M. Shearer, and C. Mecklenbräuker (2011), Compressive sensing of the Tohoku-Oki M_w 9.0 earthquake: Frequency-dependent rupture modes, *Geophys. Res. Lett.*, *38*, L20310, doi:10.1029/2011GL049223.