Geophysical Journal International

Geophys. J. Int. (2022) **228,** 711–727
Advance Access publication 2021 September 03
GJI Seismology

https://doi.org/10.1093/gji/ggab359

Assessing the role of selected constraints in Bayesian dynamic source inversion: application to the 2017 M_w 6.3 Lesvos earthquake

Filip Kostka,[1] Jiří Zahradník,[1] Efthimios Sokos[2] and František Gallovič[1]

[1] *Department of Geophysics, Faculty of Mathematics and Physics, Charles University, Prague,* 18000*, Czech Republic. E-mail: filip.kostka2@gmail.com*
[2] *Department of Geology, Seismological Laboratory, University of Patras, Patras,* 26504*, Greece*

Accepted 2021 September 1. Received 2021 August 20; in original form 2021 February 8

SUMMARY

A dynamic finite-fault source inversion for stress and frictional parameters of the M_w 6.3 2017 Lesvos earthquake is carried out. The main shock occurred on June 12, offshore the southeastern coast of the Greek island of Lesvos in the north Aegean Sea. It caused 1 fatality, 15 injuries, and extensive damage to the southern part of the island. Dynamic rupture evolution is modelled on an elliptic patch, using the linear slip-weakening friction law. The inversion is posed as a Bayesian problem and the Parallel Tempering Markov Chain Monte Carlo algorithm is used to obtain posterior probability distributions by updating the prior distribution with progressively more constraints. To calculate the first posterior distribution, only the constraint that the model should expand beyond the nucleation patch is used. Then, we add the constraint that the model should reach a moment magnitude similar to that obtained from our centroid moment tensor inversion. For the final posterior distribution, 15 acceleration records from Greek and Turkish strong motion networks at near regional distances ($\approx$ 30–150 km) in the frequency range of 0.05–0.15 Hz are used. The three posterior distributions are compared to understand how much each constraint contributes to resolving different quantities. The most probable values and uncertainties of individual parameters are also calculated, along with their mutual trade-offs. The features best determined by seismograms in the final posterior distribution include the position of the nucleation region, the mean direction of rupture (towards WNW), the mean rupture speed (with 68 per cent of the distribution lying between 1.4 and 2.6 km s^{-1}), radiated energy (12–65 TJ), radiation efficiency (0.09–0.38) and the mean stress drop (2.2–6.5 MPa).

Key words: Earthquake dynamics; Waveform inversion; Probability distributions.

1 INTRODUCTION

Dynamic inversions of earthquake rupture aim at finding parameters governing frictional and stress conditions on a fault. This can be done in two ways. In the first approach, the stress on the fault is calculated from the history of slip obtained via kinematic inversion, and the two fields are then used to estimate parameters of the constitutive law relating slip and friction (e.g. Fukuyama & Mikuno 1993; Ide & Takeo 1997; Pulido & Irikura 2000; Peyrat *et al.* 2001; Tinti *et al.* 2005; Burjánek & Zahradník 2007). A more recent approach, used here, is the fully dynamic inversion (e.g. Peyrat & Olsen 2004; Di Carli *et al.* 2010; Ruiz & Madariaga 2011, 2013; Díaz-Mojica *et al.* 2014; Twardzik *et al.* 2014; Herrera *et al.* 2017; Gallovič *et al.* 2019a, b; Mirwald *et al.* 2019; Gallovič *et al.* 2020). In this approach, simulations in which the elastodynamic equation is coupled with the constitutive law are used and the parameters describing the law and the initial stress on the fault are searched directly. The history of slip during the rupture is obtained as a by-product and it is guaranteed to be consistent with physical laws. Solving a fully dynamic inversion problem thus solves an associated kinematic inversion problem. However, dynamic inversions also permit interpretation of the earthquake properties in terms of physics. This is crucial for understanding processes of rupture nucleation, propagation and arrest, eventually enabling realistic simulations of near-source ground motions (e.g. Aochi & Ulrich 2015).

There are several issues with dynamic inversions that complicate interpretations of their results, and prevent their widespread use. First, the appropriate form of the constitutive law that describes friction on geological faults is still a topic of intense research. Widely applied empirical friction laws have been derived from small-scale laboratory experiments (Dieterich 1979; Ruina 1983; Ohnaka & Yamashita 1989; Chen & Spiers 2016, etc.), but little is known about their applicability to the Earth's crust (see Perfettini *et al.* 2003; Marone *et al.* 2009; Viesca & Garagash 2015, for an extensive

discussion on the topic). Since the differences between the relevant laws are negligible at low frequencies and hard to distinguish within the precision and accuracy of current seismological data, we use the linear slip-weakening law (Ida 1972), which was introduced to regularize problems in fracture mechanics. Owing to its simplicity, it has been used in almost all dynamic inversions published to date. For an overview of other important friction laws, we refer the reader to Bizzarri (2011).

Second, just as kinematic inversions, dynamic inversions are non-unique and it is desirable to describe their uncertainty (Ruiz & Madariaga 2013; Gallovič *et al.* 2019a, b). We achieve this by casting the problem in a probabilistic, Bayesian framework and expressing the information about model parameters in the form of a *posterior probability density function*. This function provides a formal basis for analysing uncertainties of model parameters and their trade-offs. Another advantage of the Bayesian framework is that it enables researchers to clearly formulate their prior assumptions (in the form of the *prior probability density function*) and to identify how they affect the resulting inference.

Third, running a fully dynamic rupture simulation (representing the forward part of the inverse problem) is computationally demanding. Due to the non-linear relationship between model parameters and data, many such simulations must be run to solve the inverse problem, even in non-Bayesian methods. That is why the dynamic rupture solver must be as fast as possible. To achieve this, we use the highly efficient finite difference code FD3D_TSN, which utilizes GPU acceleration and requires approximately 1 s of single GPU computational time per 1 s of rupture propagation (Premus *et al.* 2020). An important strategy for making the inversion feasible is to keep the dimension of the parameter space low. There have been only a few inversions directly seeking a discretized distribution of stress and friction on the fault (Fukuyama & Mikumo 1993; Peyrat *et al.* 2001; Peyrat & Olsen 2004; Corish *et al.* 2007; Gallovič *et al.* 2019a,b, 2020). Instead, simple parametrizations are typically considered, such as models consisting of one or two elliptic subfaults (2011Di Carli *et al.* 2010; Ruiz & Madariaga 2011, 2013; Díaz-Mojica *et al.* 2014; Twardzik *et al.* 2014; Herrera *et al.* 2017).

Finally, and most importantly for this study, interpreting the results of the inversion is difficult because it is not clear which constraints are responsible for the appearance of particular features. For example, anticorrelation between the average slip and the ruptured area is likely to be observed in every earthquake with a well-constrained seismic moment. In contrast, fine geometric and temporal features of the rupture propagation may only be constrained by using detailed seismic waveforms specific to the earthquake. We investigate this issue by using progressively more information to calculate three posterior distributions. For the first posterior distribution, we use only the constraint that the rupture breaks at least 10 per cent of the available area and lasts more than 1 second. This condition removes uninteresting models that produce negligible wave radiation. For the second distribution, we add information about the moment magnitude. The resulting distribution will characterize models with similar mechanism and magnitude compared to the one from moment tensor inversion, regardless of the observed waveforms. For the final posterior distribution, we use both moment magnitude and waveforms observed at near-regional seismic stations. Comparing the three posterior distributions allows us to separate features that are determined by the rupture condition, those determined by magnitude, and those determined by waveforms.

We apply our method to the M_w 6.3 Lesvos earthquake that occurred on 12 June 2017, 12:28 GMT, offshore the southeastern coast of the Greek island of Lesvos in the Lesvos Basin, Aegean Sea. According to the Geophysical Institute of the National Observatory of Athens (GI-NOA), it was a shallow crustal event with a hypocentral depth of 12.0 ± 1.7 km. The stress state in the area is characterized as transtensional, with minimum principal stress axis $\boldsymbol{\sigma}_3$ oriented in the NNE-SSW direction (Konstantinou *et al.* 2017). The earthquake likely ruptured the eastern segment of the Lesvos Basin fault, oriented perpendicular to $\boldsymbol{\sigma}_3$, dipping SSW with a normal faulting mechanism (Kiratzi 2018). We show a map of the epicentral area in Fig. 1.

Most of the damage caused by the earthquake occurred on the southern coast of Lesvos. In what has been called the 'Vrisa paradox' (Papadimitriou *et al.* 2018), the heaviest structural damage was observed in the small village of Vrisa about 20 km towards NW from the main shock epicentre, despite the presence of closer towns and villages. (e.g. Plomari, Akrasi, Vatera, etc.). One woman in Vrisa died, and at least 15 people were injured. This has been attributed to site effects, vulnerable infrastructure (Lekkas *et al.* 2017), and the large spatial extent of slip and source directivity. The last two have been examined by a kinematic inversion of seismic data (Kiratzi 2018), a kinematic inversion of GPS data (Chousianitis & Konca 2018), and an analysis of the aftershock sequence (Papadimitriou *et al.* 2018). Both kinematic inversions conclude that the slip was concentrated in a large patch with unilateral propagation of rupture from the hypocentre towards the northwest, as also indicated by centroid position reported shortly after the event at EMSC by Sokos & Zahradník (2017). Here we use data from local Greek and Turkish stations to reanalyse the earthquake in a fully dynamic, Bayesian framework.

2 METHOD

2.1 Forward problem

The forward problem consists of a dynamic rupture simulation and a calculation of synthetic waveforms in a layered isotropic medium. For the former, we use the Fortran code FD3D_TSN (Premus *et al.* 2020), which uses finite differences on a staggered grid of 4th order in space and of 2nd order in time. The code solves the elastodynamic equation in a 3-D box, allowing for discontinuous displacement (slip) on a pre-defined fault. The mechanical conditions on the fault, which is placed at one of the vertical faces of the box and implemented using the traction-at-split-node approach, are governed by the linear slip-weakening friction law (Ida 1972). This law relates shear traction $\boldsymbol{T}(\boldsymbol{x}, t)$ and slip $\mathbf{s}(\boldsymbol{x}, t)$ at each point $\boldsymbol{x}$ on the fault and time t and consists of two parts:

1. The rupture criterion: The slip rate at $\boldsymbol{x}$ is zero until the magnitude of shear traction at that point reaches the strength T_u.
2. Constitutive law: The on-fault traction during slip is a function of the accumulated slip $d(\boldsymbol{x}, t) = \int_0^t \|\dot{\boldsymbol{s}}(\boldsymbol{x}, \tau)\| d\tau$, and the slip-rate direction:

$$\boldsymbol{T}(\boldsymbol{x}, t) = -f(d(\boldsymbol{x}, t)) \frac{\dot{\boldsymbol{s}}(\boldsymbol{x}, t)}{\|\dot{\boldsymbol{s}}(\boldsymbol{x}, t)\|}, \tag{1}$$

where (see Fig. 2a):

$$f(d) = \begin{cases} T_u \left(1 - \frac{d}{D_c}\right) + T_d & \text{for } 0 \le d \le D_c \\ T_d & \text{for } d > D_c \end{cases}. \tag{2}$$

Here D_c is the so-called characteristic slip distance and T_d is dynamic friction.

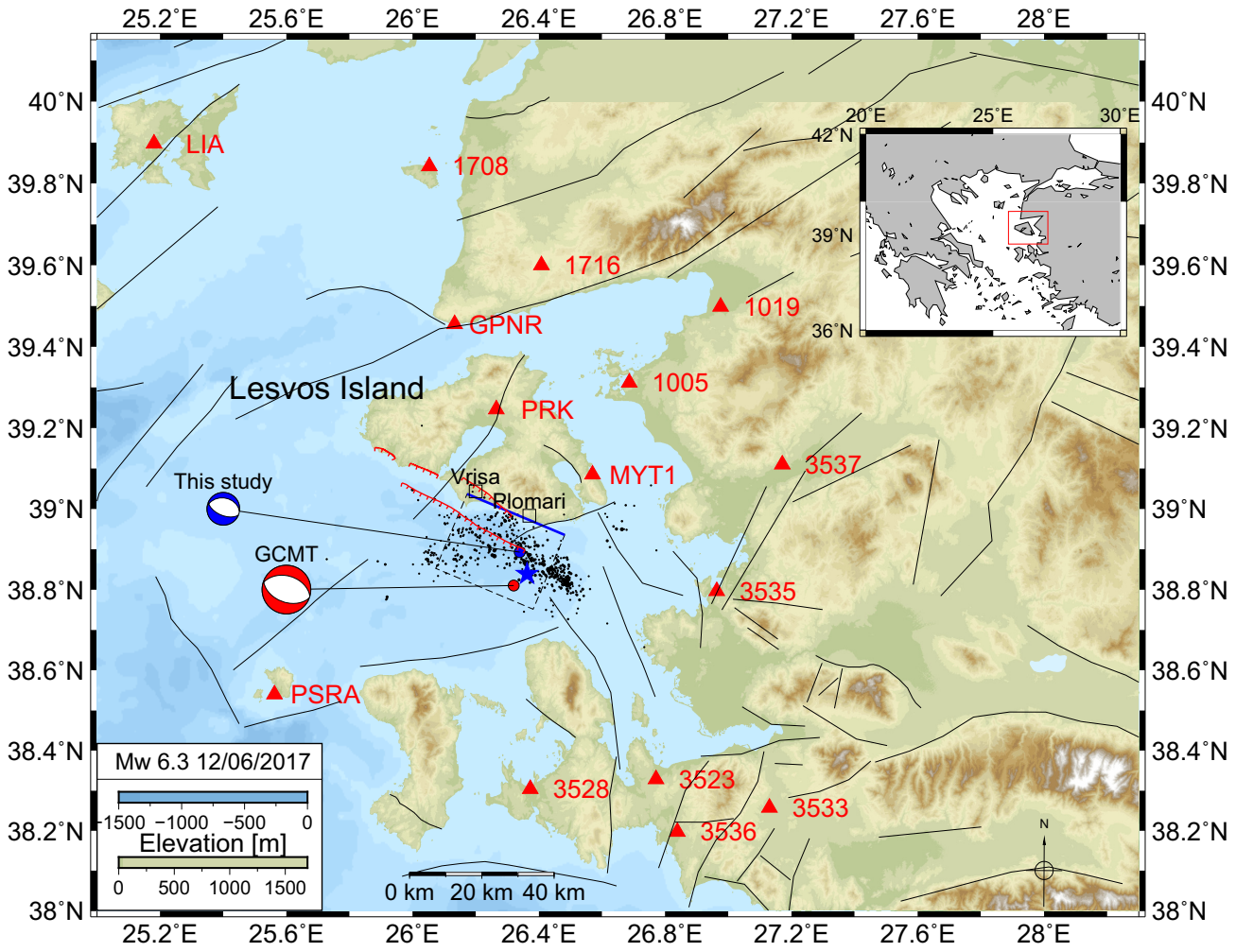

Figure 1. Map of the epicentral area of the 2017 M_w 6.3 Lesvos earthquake (for wider geographic view, see the inset). Strong-motion stations used for the dynamic inversion are shown as red triangles. The black rectangle shows the projection of the assumed fault plane used in the inversion. The blue solid line is the top fault edge at the surface. The blue beachball shows the centroid moment tensor of the M_w 6.3 main shock inferred in this study, while the red beachball shows the GCMT centroid for the same event. Black dots are the aftershocks within two months after the event, as determined by GI-NOA. The hypocentre located by GI-NOA is denoted as a blue star. Fault traces from The European Database of Seismogenic Faults (EDSF, Basili *et al.* 2013) are shown as black lines. The red lines show inferred and mapped faults from Chatzipetros *et al.* (2013) and Ganas *et al.* (2013).

The initial shear traction $\boldsymbol{T}_i$ points in the up-dip direction (to represent normal faulting), approximating the centroid rake direction of the event (–83°, see Section 2.5). The direction of traction is not fixed, but changes only negligibly during the dynamic simulation. We set T_d to zero, as is commonly done in dynamic inversions (e.g. Ruiz & Madariaga 2011, 2013; Twardzik *et al.* 2014; Gallovič *et al.* 2019a, b). Indeed, we have verified that shifting $\|\boldsymbol{T}_i\|$, T_u, T_d and T_f by a fixed value has a negligible effect on the simulation results.

The boundary conditions at the remaining faces of the computational box consist of a free surface enforced by the stress-imaging technique (Levander 1988; Graves 1996; Kristek *et al.* 2002) at the top face, and perfectly matched layers (Berenger 1994) as absorbing conditions at the remaining faces. We note that a significant speed-up in FD3D_TSN is achieved by assuming that the fault is vertical. To partially compensate for neglecting the actual dip of the fault, we stretch the along-dip positions of the velocity model interfaces, so that they conform to the actual depths along the fault. Since the synthetic ruptures do not reach the surface, the error caused by ignoring the actual dip in the first stage of the calculation is negligible (Gallovič *et al.* 2019b).

As a result of the dynamic rupture simulation, we obtain the evolution of traction and slip rate at each gridpoint on the fault. The slip rates are then convolved with Green's functions pre-calculated using the Axitra software (Cotton & Coutant 1997). But unlike the FD3D simulation, we supply Axitra with the centroid dip and rake to calculate the Green's functions (see application and validation of this approach by Gallovič *et al.* 2019b). The resulting elementary seismograms are summed over every gridpoint and as a result, synthetic displacements on specified stations are obtained. As a final step, we apply the fourth-order causal Butterworth filter to each of the seismograms, the same as used for the data (see Section 2.5).

2.2 Dynamic model parametrization

The distribution of friction and initial stress on the fault is defined by a single elliptic patch model (Ruiz & Madariaga 2011, 2013; Díaz-Mojica *et al.* 2014; Herrera *et al.* 2017, see Fig. 2b). It is relatively simple, which is appropriate due to the small complexity of the event indicated by (i) previous studies of the earthquake

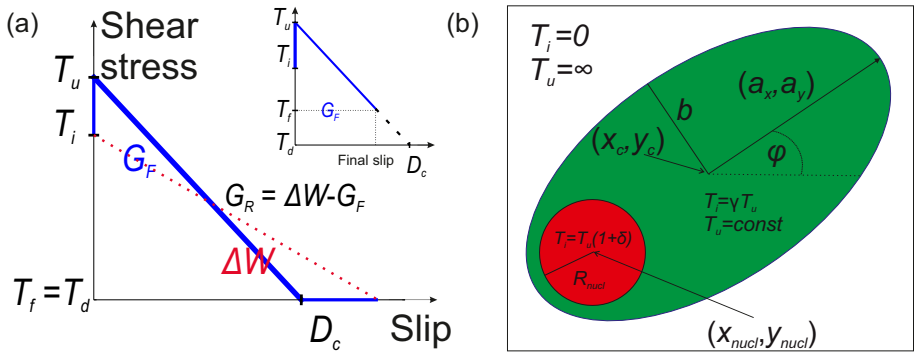

Figure 2. (a) The linear slip weakening friction law. The magnitude of shear stress at each point $\boldsymbol{x}$ on the fault has the initial value $T_i(\boldsymbol{x})$ (see Fig. b) and obeys linear elasticity until it is larger or equal than T_u (shear strength). It then decreases linearly with accumulated slip d until it reaches the final stress $T_f = T_d$ at $d = D_c$. The inset shows the case when slip stops before reaching D_c, in which case $T_f > T_d$. Stress drop is defined as $T_i - T_f$. The area ΔW under the red dotted line is the available strain energy surface density. The area G_F under the blue curve is the fracture energy surface density (energy dissipated from ΔW per area of rupture). The quantity $G_R = \Delta W - G_F$ is the radiated energy density. Integrals of these quantities over the fault plane are the available strain energy change ΔE, fracture energy E_F, and radiated energy E_R, respectively. For the extension to 3-D, see Ripperger *et al.* (2007). (b) The elliptic parametrization used for the dynamic inversion. The shear strength T_u is finite and constant inside an elliptic patch and infinite elsewhere. The geometry of the ellipse is defined by the along-strike and along-dip coordinates of its centre, x_c and y_c, the along-strike and along-dip components of one of the semi-major axes, a_x, a_y, and the length of the semi-minor axis, b. The angle between the semi-major axis and the horizontal line is denoted by ϕ. The geometry of the nucleation zone (red) is parametrized by the coordinates of its centre x_{nucl}, y_{nucl} and the radius R_{nucl}. The magnitude of the shear stress T_i inside the nucleation zone is defined relatively to T_u by the positive parameter δ: $T_i = T_u\,(1 + \delta)$. Elsewhere on the patch, T_i has a constant value of $\gamma\, T_u$. Outside of the patch, it is zero.

(Chousianitis & Konca 2018; Kiratzi 2018), (ii) our moment tensor inversion (see Text S1) and (iii) the limited frequency range we use. In addition, the limited number of parameters makes the elliptic patch model suitable for a Bayesian inversion.

The geometry of the patch is parametrized by five parameters (Fig. 2b): The position of its centre along strike, x_c, and along dip (measured from bottom to top), y_c, the components a_x, a_y of the vector connecting the centre and the tip of the semi-major axis, and the length of the semi-minor axis b. The strength is set to a constant value of T_u inside the patch and to a very large value outside of it, so the patch is the only region in which rupture may propagate. The characteristic slip distance is set to a constant value of D_c everywhere on the fault.

The magnitude of the initial shear traction T_i inside the patch is a piecewise constant function with the value of $T_u(1 + \delta)$, with positive δ, on a small circular nucleation zone from which rupture begins, and the value of $T_u\gamma$, with γ between 0 and 1, on the rest of the patch. Finally, the geometry of the nucleation zone is determined by three parameters: the along-strike and along-dip locations x_{nucl}, y_{nucl} of its centre and its radius R_{nucl}.

To sum up, we parametrize the model by 12 model parameters (Fig. 2b): 8 that describe the geometry of the elliptic patch and the nucleation zone (x_c, y_c, a_x, a_y, b, R_{nucl}, x_{nucl}, y_{nucl}), 2 that determine the magnitude of the initial shear stress (γ, δ) and 2 that define frictional properties (T_u, D_c).

2.3 The inverse problem

2.3.1 Bayesian framework

We describe our knowledge about the parameters in terms of a *probability density function* on the space of model parameters M. This is a function $f\colon M \to \mathbb{R}_0^+$ such that the true value of the parameter combination $\boldsymbol{m} = (m_1, m_2 \ldots, m_{12})$ is contained in a set $A \subset M$ with probability

$$P(A) = \int_A f(m_1, m_2 \ldots, m_{12})\, \mathrm{d}\boldsymbol{m}\,. \tag{3}$$

For the sake of brevity, we also refer to probability density functions as *probability distributions*, or simply *distributions*.

Even before analysing the constraints $\boldsymbol{y}$ on the earthquake in detail, we have some prior knowledge about the parameters. For example, we know that the hypocentre of the earthquake is located near Lesvos island, we have some estimates about the extent of the rupture, maximum stress drop, etc. We formalize this knowledge in terms of the *prior probability density function* $\rho_{pr}(\boldsymbol{m})$, which we fully describe in Section 2.4. We then update our prior knowledge by taking $\boldsymbol{y}$ into account. Mathematically, this is represented by passing to the *posterior probability density function* $\rho(\boldsymbol{m}|\boldsymbol{y})$ using *Bayes' formula*:

$$\rho(\boldsymbol{m}|\boldsymbol{y}) = cL(\boldsymbol{m}|\boldsymbol{y})\,\rho_{pr}(\boldsymbol{m})\,. \tag{4}$$

Here, $L(\boldsymbol{m}|\boldsymbol{y})$ is the *likelihood function*, which represents the probability density function of $\boldsymbol{y}$ given that the true model parameters are equal to $\boldsymbol{m}$, and c is a normalizing constant such that $\rho(\boldsymbol{m}|\boldsymbol{y})$ integrates to 1.

In this study, we divide our updating process into three steps. In the first step, we assume that we have no information about the earthquake other than that it lasted for at least one second and ruptured more than twice the area of the nucleation zone. This rupture condition, which we label r, leaves only models that successfully rupture beyond their nucleation zone and excludes uninteresting models that cannot produce any ground displacements at seismic stations. Formally, the likelihood function for this condition can be written as

$$L_0(\boldsymbol{m}|r) \propto I_R(\boldsymbol{m})\,, \tag{5}$$

where R is the set of all models that meet the rupture condition and I_R is its indicator function. Plugging eq. (5) into eq. (4) with $\boldsymbol{y} = r$, we obtain the posterior distribution

$$\rho_0(\boldsymbol{m}|r) = c_0\, I_R(\boldsymbol{m})\,\rho_{pr}(\boldsymbol{m})\,, \tag{6}$$

where c_0 is a normalizing constant. Since the exact shape of R is unknown, neither the value of c_0 nor the precise form of I_R can be determined beforehand. However, this is not important when estimating the distribution with a suitable Monte Carlo method, such as the Parallel Tempering algorithm (see Section 2.3.2).

In the second step, we add to $\boldsymbol{y}$ an estimate of the earthquake's moment magnitude M_w^0, so that $\boldsymbol{y} = (r, M_w^0)$. When we solve the forward problem with model $\boldsymbol{m}$, we obtain the corresponding moment magnitude $M_w(\boldsymbol{m})$. Due to observational and modelling errors, neither M_w^0 nor $M_w(\boldsymbol{m})$ will be exact. However, if we suppose that the errors are normally distributed, it turns out (Tarantola 2005) that the combined error is also normally distributed with variance $\sigma_{M_w}^2$ which is equal to the sum of the original variances. The likelihood function is then:

$$L_1\left(\boldsymbol{m}|M_w^0\right) \propto \exp\left(-\frac{\left(M_w\left(\boldsymbol{m}\right)-M_w^0\right)^2}{2\sigma_{M_w}^2}\right). \tag{7}$$

Using ρ_0 as the prior distribution and updating it using eq. (4), we obtain the posterior distribution

$$\begin{aligned}\rho_1\left(\boldsymbol{m}|r, M_w^0\right) &= c_1\, L_1\left(\boldsymbol{m}|M_w^0\right)\rho_0\left(\boldsymbol{m}|r\right)\\ &= c_1 I_R\left(\boldsymbol{m}\right)\, \exp\left(-\frac{\left(M_w\left(\boldsymbol{m}\right)-M_w^0\right)^2}{2\sigma_{M_w}^2}\right)\\ &\quad\times \rho_{pr}\left(\boldsymbol{m}\right),\end{aligned} \tag{8}$$

where c_1 is again a normalizing constant. This distribution is independent of ground displacement observations, except for the information about the moment magnitude. As such, it allows us to identify model parameters and features that are required for the earthquake to rupture with approximately the given magnitude, but are not necessarily specific to the actual earthquake.

In the final step, we supplement the preceding constraints with the observed seismograms $\boldsymbol{d_{obs}}$, that is $\boldsymbol{y} = (r, M_w^0, \boldsymbol{d}_{obs})$. Assuming normally distributed errors as before, the likelihood function for this update is

$$L_2\left(\boldsymbol{m}|\boldsymbol{d_{obs}}\right) \propto \exp\left(-\frac{\|\boldsymbol{d}\left(\boldsymbol{m}\right)-\boldsymbol{d}_{obs}\|^2}{2\sigma_d^2}\right), \tag{9}$$

where $\boldsymbol{d}(\boldsymbol{m})$ are the synthetic seismograms calculated from the model $\boldsymbol{m}$, and σ_d^2 is the total observational and modelling variance. Using this likelihood function and using ρ_1 as the prior distribution, eq. (4) now yields

$$\begin{aligned}\rho_2\left(\boldsymbol{m}|r, M_w^0, \boldsymbol{d_{obs}}\right) = c_2\, I_R\left(\boldsymbol{m}\right)\exp\Bigg(&-\frac{\|\boldsymbol{d}\left(\boldsymbol{m}\right)-\boldsymbol{d}_{obs}\|^2}{2\sigma_d^2}\\ &-\frac{\left(M_w\left(\boldsymbol{m}\right)-M_w^0\right)^2}{2\sigma_{M_w}^2}\Bigg)\,\rho_{pr}\left(\boldsymbol{m}\right),\end{aligned} \tag{10}$$

where c_2 is another normalizing constant. In Section 3.2, we compare the prior distribution and the three posteriors to extract the information contained in each of the constraints.

Since the model space M has 12 dimensions, the probability distributions are hard to visualize and interpret. Nevertheless, if we decompose M as a Cartesian product of spaces A and B such that $M = A \times B$, we can calculate, for any distribution ρ on M, its *marginal distribution* $\rho_A(\boldsymbol{m}_A|\boldsymbol{y})$ for $\boldsymbol{m}_A \in A$ by integrating $\rho(\boldsymbol{m}_A, \boldsymbol{m}_B|\boldsymbol{y})$ over B:

$$\rho_A\left(\boldsymbol{m}_A|\boldsymbol{y}\right) = \int_B \rho\left(\boldsymbol{m}_A, \boldsymbol{m}_B\,|\boldsymbol{y}\right) d\boldsymbol{m}_B. \tag{11}$$

Plots of the marginal distributions give us a picture about the original (joint) distribution, but they must be approached with caution because some information is lost by the integration. In particular, that the marginal distribution has a maximum at some $\boldsymbol{m}_A \in A$ is neither a sufficient nor necessary condition for the joint distribution to have a maximum at $(\boldsymbol{m}_A, \boldsymbol{m}_B)$ for some $\boldsymbol{m_B} \in B$. To simplify notation in the following text, we use the same symbols ρ_{pr}, ρ_0, ρ_1 and ρ_2 for both the marginal and the joint distributions.

2.3.2 Sampling the posterior distributions with the Parallel Tempering algorithm

Even though eqs (6), (8) and (10) show the functional form of the sought posterior distributions, a major complication immediately arises when one wants to calculate them in practice. The values of the functions $I_R(\boldsymbol{m})$, $M_w(\boldsymbol{m})$ and $\boldsymbol{d}(\boldsymbol{m})$ which appear in these formulas are only available through costly numerical simulations; we do not have analytical expressions for them. If we try to numerically evaluate the distributions on a regular grid spanning the model space, the calculation quickly becomes unfeasible as the dimension of the model space increases and most time is spent on models with negligible probability density. Moreover, the values of the normalizing constants are unknown.

Monte Carlo methods solve this problem by efficiently drawing *samples* of a posterior distribution and using the obtained ensemble to approximately characterize the distribution. A popular technique used for Monte Carlo sampling is the *Metropolis–Hastings* (MH) algorithm (Metropolis *et al.* 1953; Hastings 1970; see also Sambridge & Mosegaard 2002), which uses a random walker that moves through the model space according to a prescribed proposal distribution, and accepts or rejects models according to a mathematically derived rule. The rule guarantees that given enough steps, the accepted models will sample the target distribution. However, the original MH algorithm works best for distributions with a single local maximum. For distributions with multiple local maxima, the walker may get trapped in a close neighborhood of a particular one without ever exploring the others.

One of the extensions of the MH algorithm that solve this problem, already utilized for dynamic source inversion by Gallovič *et al.* (2019a, b), is the Parallel Tempering algorithm (also known as Replica exchange Monte Carlo; Swendsen & Wang 1986). This algorithm samples the target distribution by performing the MH algorithm on multiple chains, each of which, if working independently, would sample a tempered distribution $\rho_T(\boldsymbol{m}|\boldsymbol{y}) := k_T\,[\rho(\boldsymbol{m}|\boldsymbol{y})]^{\frac{1}{T}}$. Here T is a parameter called *temperature*, generally different for each chain, and k_T are normalizing constants, the knowledge of which is not required by the algorithm. The target distribution is sampled by chains at $T = 1$. The chains at higher temperatures are auxiliary and their purpose is to help the $T = 1$ chains to jump over areas of low probability. This is achieved by allowing the chains at different temperatures to randomly swap their models. Indeed, since the tempered distributions, sampled by chains at $T > 1$, are flatter than the target distribution (converging to the uniform distribution as $T \to \infty$), these chains can traverse areas that would be unavailable for chains at $T = 1$. For more details, see Sambridge (2013).

Once a large enough ensemble of samples is obtained, it can be used to approximate the important integrals characterizing the distribution, including the normalizing constant. For example, the mean and the variance of the distribution are estimated by calculating the mean and the variance, respectively, of the sampled ensemble. Samples of the marginal probability density functions (see eq. 11)

are obtained by projecting samples of the joint distribution to the target space of interest.

To approximately reconstruct the original distribution from its samples, we use the *kernel density estimation* technique (KDE, see, e.g. Zambom & Dias 2013). In the case of 1-D spaces (generalization to higher dimensions is straightforward), the distribution is estimated by the formula:

$$\rho_{est}(x) = \frac{1}{|S|\,h} \sum_{s \in S} K\left(\frac{x-s}{h}\right), \tag{12}$$

where S is the collection of the obtained samples, $|S|$ is its size, h is a positive real parameter called the *bandwidth* and $K(x)$ is a smooth even function with unit integral, called the *kernel*. Here, we use the Gaussian kernel:

$$K(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right). \tag{13}$$

Compared to the more standard method of plotting histograms of the samples, KDE estimates are smooth functions and on average, converge faster to the true distribution as the number of samples increases (Wasserman 2004). Nevertheless, the dependence on the bin width is replaced by the dependence on the bandwidth h. We choose h equal to 1/40 of the respective parameter range. This value is similar to the bandwidths calculated according to the rule of thumb suggested by Scott (1992).

2.4 Prior distribution on the model space

The joint prior distribution ρ_{pr} of the 12 model parameters is assumed to be uniform on the set $P \subset M$ defined by the following four constraints:

1. All parameters lie within the intervals specified in Table 1.
2. The parameter b specifying the length of the semi-minor axis is smaller or equal to the length of the semi-major axis a:

$$b \leq a = \sqrt{a_x^2 + a_y^2}\,. \tag{14}$$

3. The centre of the nucleation zone lies within the patch.
4. The patch is completely contained within the fault.

The purpose of constraint #2 is to avoid ambiguity in specifying the semi-major and semi-minor axes. Constraint #3 ensures that rupture always initiates within the patch, which represents a weakened area on the fault. Constraint #4 ensures that the patch always has an elliptic shape. We note that we allow both a_x and a_y to take positive and negative values, permitting occurrences of pairs $(a_x,\ a_y)$ and $(-a_x,\ -a_y)$ which represent the same model. This introduces redundancy into the inversion, but it is nevertheless convenient as it eliminates unnecessary barriers in the Monte Carlo sampling.

We sample the prior distribution ρ_{pr} with a random number generator, uniformly generating models within the bounds given by constraint #1 and then only accepting those that satisfy constraints #2–4. KDE estimates of 1-D marginal priors of ρ_{pr}, reconstructed from 100 000 samples, are shown as black curves in Fig. 3. Since the 1-D distributions of parameters a_x and a_y are not very informative, we instead show derived quantities a (length of the semi-major axis) and ϕ, which is the angle between the vector (a_x,a_y) and the horizontal line, $\phi = \mathrm{atan2}(a_y, a_x)$. After obtaining the samples, we identify values of ϕ differing by 180°, as they describe the same model. Similarly, since x_c and y_c necessarily have very similar distributions to x_{nucl} and y_{nucl}, respectively (due to constraint #3), we instead show two derived quantities that characterize the connecting vector $\Delta\boldsymbol{r} = (x_c - x_{nucl}, y_c - y_{nucl})$: C_{dist}, which is the length of $\Delta\boldsymbol{r}$ relative to the elliptic patch (equal to 0 when the nucleation is at the centre of the patch and equal to 1 when it is at the boundary), and C_{ang}, which is the angle between $\Delta\boldsymbol{r}$ and the horizontal line, $C_{ang} = \mathrm{atan2}(\Delta r_y,\ \Delta r_x)$.

We note that while the *joint* prior distribution function of the model parameters is uniform on P, *marginal* distributions of the geometric model parameters (except for R_{nucl}) shown in Fig. 3 are non-uniform. There are two reasons for this. First, due to constraints #2–4, P has a non-rectangular shape, and the effective bounds over which the joint distribution is integrated (eq. 11) may depend on the value of the parameter for which the marginalization is carried out. Secondly, the derived parameters a, ϕ, C_{dist} and C_{ang} were obtained by a coordinate transformation, and the joint distribution is therefore modified by a non-uniform Jacobian factor.

2.5 Data and model setup

Based on the full-waveform centroid moment tensor (CMT) inversion performed before the dynamic inversion by the software ISOLA (Zahradník & Sokos 2018) and summarized in Text S1, we adopt the fault plane geometry and mechanism with strike/dip/rake values 113°/40°/–83° and $M_w^0 = 6.24$ (eqs 7, 8 and 10). The map projection of the fault plane is shown in Fig. 1.

Within 150 km from the centroid, the earthquake was recorded by 55 strong-motion stations of the Geodynamic Institute of the National Observatory of Athens (GI-NOA) and Bogazici University Kandilli Observatory and Earthquake Research Institute (KOERI). We excluded stations very close to each other (with almost the same waveforms) to reduce station redundancy.

For the computation of the Green's functions, we adopt a five-layer model after Karagianni *et al.* (2002), Fig. S6. However, many of the recordings are affected by significant basin and site effects, which are not included in our Green's functions. To select stations suitable for the inversion, we used Axitra to calculate synthetic seismograms for a point source located at the GI-NOA hypocentre in the low-frequency range of 0.05–0.08 Hz and excluded stations with visibly poor fits with the observed seismograms. For example, stations in the city of Izmir were removed because of a significant path effect likely caused by the presence of the Izmir basin, which cannot be reproduced in our 1-D model. In the end, we selected the 15 stations shown in Fig. 1. The event has good angular coverage of stations to the east (azimuths from –10° to 170°), but poor coverage to the west as that direction corresponds to the open sea.

Acceleration records were tapered with a rectangular time window starting at the origin time of 2017/06/12 12:28:37 GMT and with a duration of 80 s. The records were then bandpass filtered between 0.05 and 0.15 Hz by the 4th order Butterworth filter, and integrated into displacements. The lower frequency bound is necessary to remove low-frequency instrumental noise from the data. The upper bound is chosen to lower the influence of the imperfect velocity model, as well as uncertainties in fault geometry.

We set the standard deviation σ_{M_w} (eqs 7, 8, 10) to 0.1, representing roughly the variability of magnitudes from different studies/agencies (Table S1). The total seismogram standard deviation σ_d (eqs 9 and 10) is set to 2.5 cm. A similar relative estimate of seismogram uncertainty was also used in the dynamic inversions by Gallovič *et al.* (2019a, b). We note that it is close to the theoretical estimate by Hallo & Gallovič (2016) based on synthetic simulations with randomly varied velocity models. To account for the uncertainty in the origin time and to balance for relatively weak or

Table 1. Prior ranges of model parameters, and modes and 68 per cent highest density regions (HDR) of posterior distributions ρ_0, ρ_1 and ρ_2. The symbols denoted by asterisks are the model parameters directly searched in the inversion, the rest is derived from them using explicit formulas or obtained as a result of dynamic rupture simulations (emergent quantities). In the last three columns, we use the notation $[_l M_u]$, where l and u are the lower and upper bounds of the 68 per cent HDR and M is the mode of the distribution. When the HDR is not an interval, we instead write M $[S]$, where S is the total size of the HDR.

Parameter name	Meaning	Prior min.	Prior max.	Mode and 68 per cent HDR		
				ρ_0	ρ_1	ρ_2
Model parameters						
a_x (km)*	Along-strike component of the semi-major axis	−17.5	17.5	−11.1 [18.8]	−6.0 [17.0]	−8.6 [15.5]
a_y (km)*	Along-dip component of the semi-major axis	−17.5	17.5	−11.3 [18.4]	5.9 [16.5]	$[_{-7.5} - 0.3_{7.5}]$
a (km)	Length of the semi-major axis	–	–	$[_{10.1}14.3_{17.4}]$	$[_{5.9}8.0_{12.6}]$	$[_{7.11}9.0_{13.5}]$
b (km)*	Length of the semi-minor axis	0.0	24.75	$[_{4.7}8.9_{12.6}]$	$[_{2.4}3.8_{5.7}]$	$[_{2.7}4.7_{6.5}]$
ϕ (°)	Angle of the semi-major axis	–	–	135 [101]	0.541 [118]	174 [86]
x_c (km)*	Along-strike position of the elliptic centre	1	34	$[_{12.9}17.6_{22.1}]$	$[_{10.2}17.4_{23.3}]$	$[_{15.2}17.6_{20.0}]$
y_c (km)*	Along-dip position of the elliptic centre	1	34	$[_{13.2}17.7_{22.3}]$	$[_{11.7}17.6_{23.6}]$	$[_{18.0}20.9_{22.9}]$
C_{dist}	Elliptic distance of the nucleation zone and the elliptic centre	–	–	$[_{0.46}0.78_{0.96}]$	$[_{0.35}0.60_{0.85}]$	$[_{0.33}0.62_{0.84}]$
C_{ang} (°)	Angle of the vector connecting the nucleation zone with the elliptic centre	–	–	268 [232]	270 [227]	$[_{115}180_{226}]$
$\gamma*$	Initial background shear traction in the elliptic patch, relative to T_u	0.2	1.0	$[_{0.50}0.96_{0.99}]$	0.68 [0.51]	$[_{0.24}0.41_{0.61}]$
T_u (MPa)*	Shear strength	1.0	18.0	$[_{10.4}16.9_{17.8}]$	$[_{8.8}14.0_{17.1}]$	$[_{3.9}6.4_{12.0}]$
$\delta*$	Initial shear stress excess at the nucleation zone, relative to T_u	0.0	0.2	$[_{0.09}0.18_{0.20}]$	$[_{0.10}0.18_{0.19}]$	$[_{0.11}0.17_{0.19}]$
R_{nucl} (km)*	Radius of the nucleation zone	0.1	3.0	$[_{1.5}2.7_{2.9}]$	$[_{1.6}2.8_{3.0}]$	$[_{1.7}2.5_{2.9}]$
D_c (m)*	Characteristic slip-weakening distance	0.05	0.5	0.08 [0.26]	$[_{0.05}0.08_{0.28}]$	$[_{0.05}0.08_{0.25}]$
x_{nucl} (km)*	Along-strike position of the nucleation zone	1	34	$[_{10.0}16.5_{25.1}]$	$[_{9.1}14.2_{24.2}]$	$[_{16.9}19.3_{22.3}]$
y_{nucl} (km)*	Along-dip position of the nucleation zone	1	34	$[_{11.0}16.7_{25.7}]$	$[_{11.0}17.9_{24.7}]$	$[_{16.6}20.1_{23.0}]$
S_{el} (km^2)	Area of the elliptic patch	–	–	$[_{97}285_{550}]$	$[_{53}109_{182}]$	$[_{77}137_{215}]$
$R_{nucl}/\tilde{R}_m$	Ratio of R_{nucl} to the estimate of the critical nucleation length.	–	–	$[_{0.56}1.18_{2.60}]$	$[_{0.78}1.23_{2.60}]$	$[_{0.90}1.24_{2.10}]$
κ	Similarity parameter	–	–	$[_{0.29}1.32_{6.10}]$	$[_{0.33}1.23_{3.17}]$	$[_{0.51}1.11_{1.97}]$
Emergent quantities						
M_w	Moment magnitude	–	–	$[_{6.43}6.93_{7.34}]$	$[_{6.14}6.27_{6.39}]$	$[_{6.08}6.18_{6.28}]$
S_r (km^2)	Ruptured area	–	–	$[_{0}36_{463}]$	$[_{57}105_{161}]$	$[_{76}126_{187}]$
$\Delta\sigma_E$ (MPa)	Slip-weighted mean stress drop	–	–	$[_{4.8}8.1_{12.1}]$	$[_{4.1}7.4_{10.9}]$	$[_{2.2}3.6_{6.5}]$
$\overline{V_r}$ (km/s)	Slip-weighted mean rupture speed	–	–	$[_{2.3}3.0_{4.7}]$	$[_{1.4}2.7_{3.4}]$	$[_{1.4}1.7_{2.6}]$
E_R (TJ)	Radiated energy	–	–	8 [272]	$[_{15}68_{210}]$	$[_{12}35_{65}]$
E_F (TJ)	Fracture energy	–	–	$[_{0}18_{217}]$	$[_{0}26_{149}]$	$[_{38}80_{174}]$
$\overline{G_F}$ (MJ/m^2)	Average fracture energy surface density	–	–	$[_{0.0}0.4_{1.4}]$	$[_{0}0.3_{1.3}]$	$[_{0.1}0.4_{1.3}]$
η	Radiation efficiency	–	–	$[_{0.75}0.92_{1.00}]$	0.90 [0.44]	$[_{0.09}0.20_{0.38}]$
$\tilde{\psi}$ (°)	Slip-weighted mean rupture direction	–	–	89 [222]	90 [222]	180 [103]
T_r (s)	Equivalent duration of rupture	–	–	5.7 [8.2]	$[_{2.4}3.6_{6.0}]$	$[_{4.4}6.2_{8.3}]$

strong nucleation of dynamic rupture, we uniformly shift synthetic seismograms at all stations by the time Δt within $(-3.2\,s, 3.2\,s)$ that results in the minimum misfit.

3 RESULTS

Let us describe the sampling of the individual posterior distributions ρ_0, ρ_1 and ρ_2. For the rupture-constrained posterior distribution ρ_0, we took 100 000 samples of the prior distribution and accepted only models that satisfied the predefined rupture condition r, that is with rupture lasting at least one second and breaking more than twice the area of the nucleation zone. About 70 per cent of prior samples did not meet the condition, leaving approximately 30 000 samples of ρ_0.

We sampled the posterior distributions ρ_1 and ρ_2 by two independent runs of the Parallel Tempering algorithm, using eqs (8) and (10), respectively. Initial models (different for each chain) were randomly picked from the prior distribution ρ_{pr}. MCMC proposals for each parameter had normal probability densities centred around its present value with a standard deviation equal to 2 per cent of its allowed range. We used parallel computing on 6 GPUs of our in-house cluster, with 2 MPI threads per GPU and 8 MCMC chains per MPI thread. The temperature of two chains in each thread was set to 1, so they sampled the target distribution, while the remaining temperatures were randomly sampled from a log-normal distribution between 1 and 100. An additional MPI thread controlled the swapping of models among the chains.

For each distribution, the number of models visited by all chains was approximately 2 500 000 within 35 d. To ensure more accurate sampling, only every 10th accepted sample was recorded and the first sixth of the recorded samples were discarded to account for the so-called burn-in phase. Finally, we discarded the samples that did not satisfy the rupture condition and kept only models sampled at $T = 1$. This way, about 50 000 samples were obtained for ρ_1 and ρ_2 each.

3.1 Best-fitting models

To provide the reader with an intuitive sense of the parameters and robust features of the sampled models, we first examine three

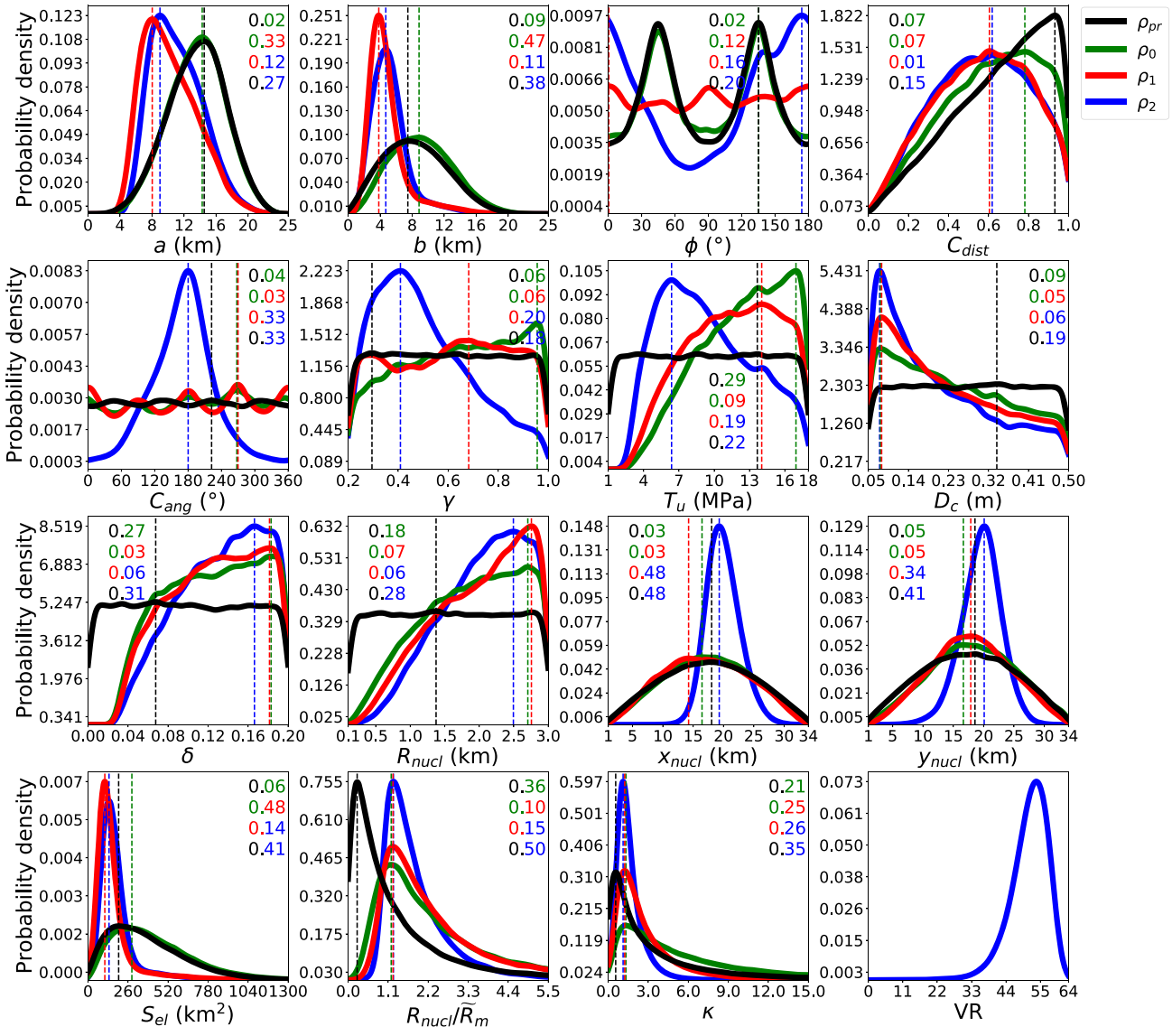

Figure 3. KDE estimates of 1-D marginal distributions for the model parameters and quantities derived from them, obtained with the Parallel Tempering algorithm. Different colours correspond to the prior distribution (ρ_{pr}) and posterior distributions incrementally constrained by the rupture condition (ρ_0), moment magnitude (ρ_1) and seismograms (ρ_2) (see legend). Vertical bars show the modes of each distribution. The numbers show the Hellinger distances between distributions, with colour-coding representing the respective pair of distributions under comparison (for example, the distance between ρ_{pr} and ρ_0 is shown as a black-and-green number). Only the 12 parameters in the first three rows are independent, the quantities $S_{el} = \pi ab$, $R_{nucl}/\tilde{R}_m$ (eq. 17) and κ (eq. 18) were calculated from them. The KDE of seismogram variance reduction (VR) for samples of ρ_2 is shown at the bottom right.

models sampled from ρ_2 that best fit the observed seismograms. Spatial distributions of their slip, stress drop, and rupture time are shown in Fig. 4. The fits between observed and synthetic waveforms for these models have variance reductions of 63–64 per cent. In Fig. 5 we plot the waveform fit of the best-fitting model along with KDE estimates of synthetic seismograms generated from the whole ρ_2 ensemble. The KDE of variance reduction for ρ_2 is shown in Fig. 3. We note that 65, 30 and 3 per cent of models have variance reductions larger than 50, 55 and 60 per cent, respectively.

The feature in which the models are the most similar is the centre coordinates of their nucleation zones, within 2 km. The dimensions of the elliptic patch are similar for the best-fitting models, though model #1 is somewhat more elongated than the other two. However, the rupture in model #1 only breaks 85 per cent of the whole patch, stopping before reaching its lower left tip. The inclination angles ϕ of each elliptic patch differ quite strongly and this parameter is rather broadly distributed in the ρ_2 ensemble. Nevertheless, for all three models, the nucleation zone is to the right of the centre of the ellipse, as is the case for 90 per cent of models in the ensemble. The rupture propagates predominantly to the WNW, as it is soon arrested at the ESE edges of the patch. However, in contrast to the relatively robust directional preference with respect to the strike direction, the data do not seem to distinguish between upward or downward spreading. We note that kinematic inversions of Kiratzi (2018) and Chousianitis & Konca (2018) suggest even more pronounced unilateral rupture propagation.

The average slip-weighted stress drops $\Delta\sigma_E$ of the three models are 5.7, 3.4 and 3.1 MPa, respectively. Model #1 has the lowest slip-weighted average rupture speed (1.25 km s^{-1} versus 1.38 km s^{-1} and 1.91 km s^{-1}). It also has larger values of both T_u and D_c than the

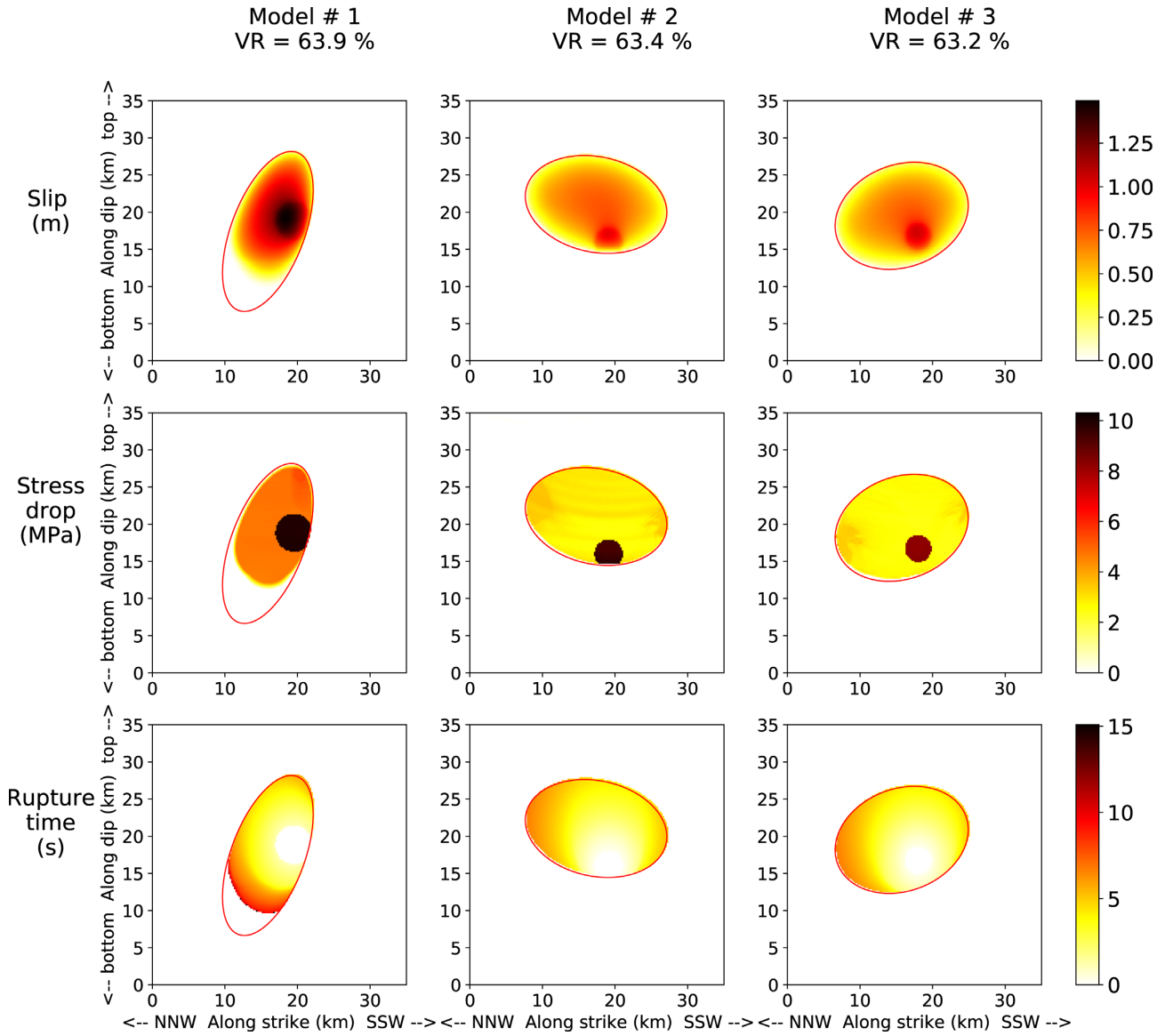

Figure 4. Distributions of the final slip, stress drop and rupture time (rows) for the three best-fitting models (columns). Only the positive part of the stress drop is shown. The red ellipses show the boundaries of the elliptic patches within which rupture may propagate.

other two models (9.6 MPa versus 8.5 MPa versus 7.8 MPa, and 0.43 m versus 0.12 m versus 0.14 m). Consequently, model #1 also has the largest average fracture energy surface density, obtained by dividing the fracture energy by the total ruptured area (1.6 MJ m^{-2} versus 0.7 MJ m^{-2} versus 0.6 MJ m^{-2}). The total radiated energies (see Fig. 2 here or Fig. 1 and eq. C2 of Ripperger *et al.* 2007) of all three models range from 20 TJ (model #1) to 40 TJ (model #2). As we discuss below, this is the quantity best resolved by the seismograms. The seismic moments of the models are similar (3.4–3.7 $\times 10^{18}$ Nm). All three models have pronounced nucleation, which is a common feature in the ρ_2 ensemble.

3.2 Characteristics of the posterior distributions

In this section, we analyse various parameters and how well they are constrained when incrementally updating the prior distribution ρ_{pr} by the rupture condition (posterior distribution ρ_0), the constraint on the moment magnitude (ρ_1), and the observed seismograms (ρ_2). A comparison of KDE estimates of marginal distributions of ρ_{pr}, ρ_0, ρ_1 and ρ_2 for model parameters (or parameters derived from them, such as a, ϕ, C_{ang} and C_{dist}) is shown in Fig. 3. Fig. 6 shows marginal posterior distributions for *emergent* quantities, that is quantities that emerge as outputs of rupture simulation, such as the moment magnitude M_w or the radiated energy E_R.

To compare the information contained in ρ_{pr}, ρ_0, ρ_1 and ρ_2, we measure their mutual dissimilarity in terms of the *Hellinger distance* H (Shemyakin 2014):

$$H(\rho_i, \rho_j) = \left(1 - \int_{\mathbb{R}} \sqrt{\rho_i(x)\,\rho_j(x)}dx\right)^{1/2}. \quad (15)$$

It is a metric on the space of probability distribution densities, equal to 0 for identical distributions and 1 for disjoint distributions. For the sake of conciseness, we use the notation $H_{i,j} = H(\rho_i, \rho_j)$. Large values of $H_{1,2}$, for example, suggest that the information contained in seismograms is quite different from that contained in the moment magnitude. Evaluating data informativeness by comparing

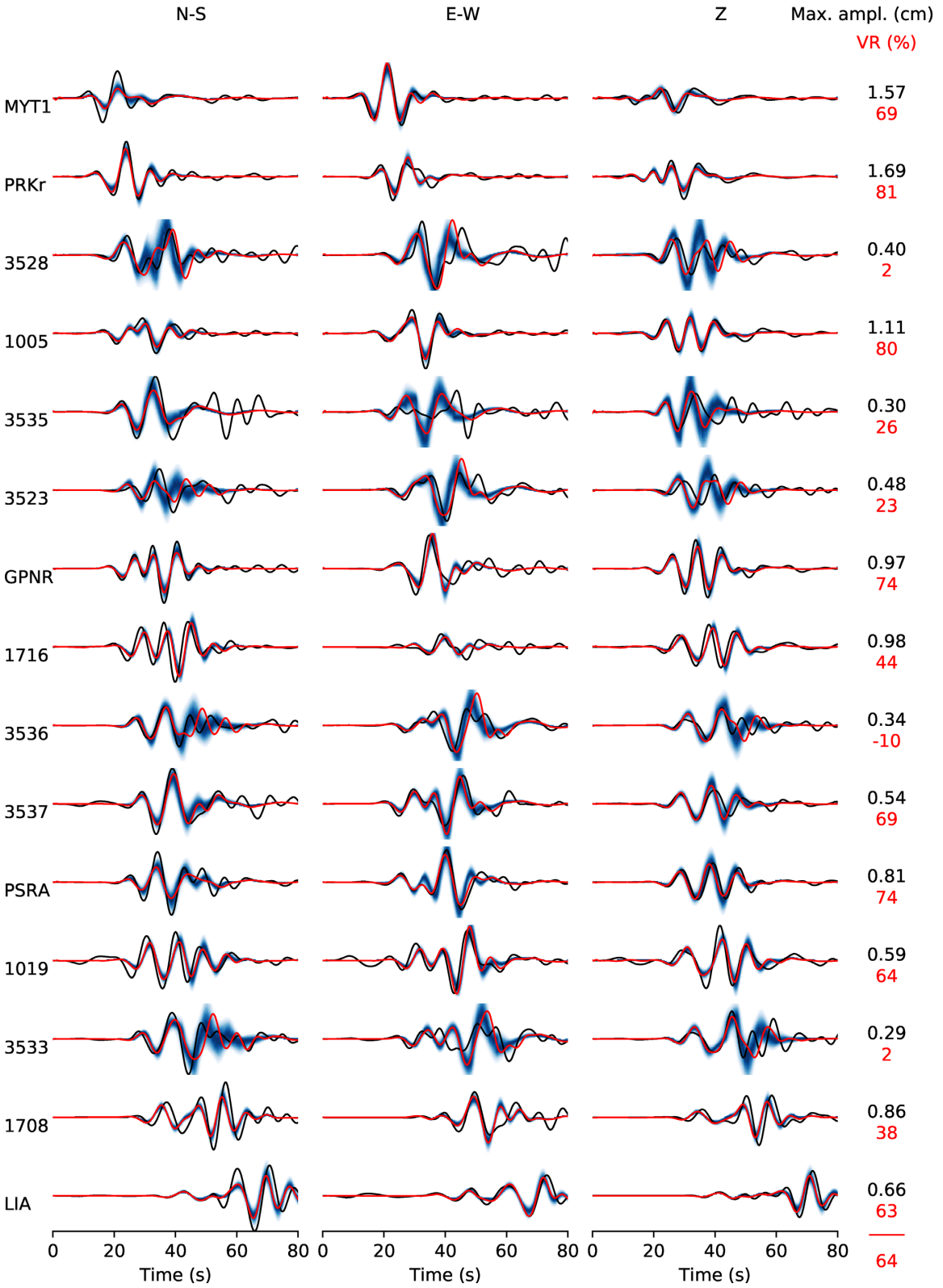

Figure 5. Displacement seismograms at the 15 selected stations in the frequency range of 0.05–0.15 Hz. Observed seismograms are black, synthetic seismograms of the best-fitting model are red. KDE estimates of synthetic seismogram posterior distribution, calculated at each time step, are blue. Numbers on the right show the maximum amplitude of the observed seismograms (black) and variance reduction at individual stations for the best-fitting model (red). The red number at the bottom is the overall variance reduction of the best-fitting model. The time axis begins on 12 June 2017, 12:28:38.26 GMT.

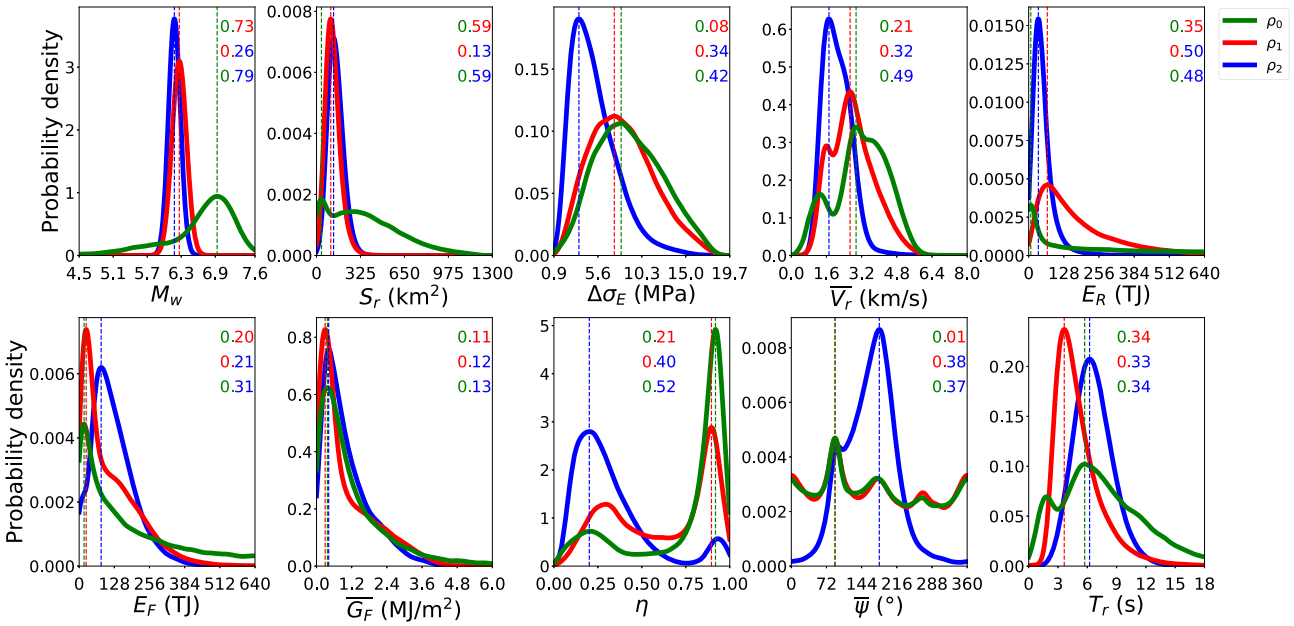

Figure 6. KDE estimates of 1-D marginal distributions for emergent quantities, obtained with the Parallel Tempering algorithm. Different colours correspond to the posterior distributions incrementally constrained by the rupture condition (ρ_0), moment magnitude (ρ_1), and seismograms (ρ_2) (see legend). Vertical bars show the modes of each distribution. The numbers show the Hellinger distances between distributions, with colour-coding representing the respective pair of distributions under comparison.

Table 2. Model parameters and derived quantities sorted according to the value of the Hellinger distance $H_{0,1}$ between posteriors ρ_0 and ρ_1, and the Hellinger distance $H_{1,2}$ between posteriors ρ_1 and ρ_2. The ratios of the sizes of the respective HDRs, $\mathrm{rHDR}_{0,1}$ and $\mathrm{rHDR}_{1,2}$, are also shown.

Rank	Parameter	$H_{0,1}$	$\mathrm{rHDR}_{0,1}$	Parameter	$H_{1,2}$	$\mathrm{rHDR}_{1,2}$
1	S_{el}	0.48	0.28	x_{nucl}	0.48	0.36
2	b	0.47	0.42	y_{nucl}	0.34	0.47
3	a	0.33	0.93	C_{ang}	0.33	0.48
4	κ	0.25	0.49	κ	0.26	0.51
5	ϕ	0.12	1.17	γ	0.20	0.72
6	$R_{nucl}/\tilde{R}_m$	0.10	0.89	T_u	0.19	0.98
7	T_u	0.09	1.13	ϕ	0.16	0.72
8	C_{dist}	0.07	1.02	$R_{nucl}/\tilde{R}_m$	0.15	0.66
9	R_{nucl}	0.07	0.90	S_{el}	0.14	1.07
10	γ	0.06	1.04	a	0.12	0.95
11	D_c	0.05	0.91	b	0.11	1.15
12	y_{nucl}	0.05	0.93	R_{nucl}	0.06	0.91
13	x_{nucl}	0.03	1.00	D_c	0.06	0.87
14	δ	0.03	0.94	δ	0.05	0.91
15	C_{ang}	0.03	0.98	C_{dist}	0.01	1.00

Table 3. Emergent quantities sorted according to the value of the Hellinger distance $H_{0,1}$ between posteriors ρ_0 and ρ_1, and the Hellinger distance $H_{1,2}$ between posteriors ρ_1 and ρ_2. The ratios of the sizes of the respective HDRs, $\mathrm{rHDR}_{0,1}$ and $\mathrm{rHDR}_{1,2}$, are also shown.

Rank	Parameter	$H_{0,1}$	$\mathrm{rHDR}_{0,1}$	Parameter	$H_{1,2}$	$\mathrm{rHDR}_{1,2}$
1	M_w	0.73	0.28	E_R	0.50	0.27
2	S_r	0.59	0.22	η	0.40	0.66
3	E_R	0.35	0.72	$\bar{\psi}$	0.38	0.46
4	T_r	0.34	0.44	$\Delta\sigma_E$	0.34	0.63
5	η	0.21	1.79	T_r	0.33	1.07
6	$\overline{V_r}$	0.21	0.83	$\overline{V_r}$	0.32	0.61
7	E_F	0.20	0.69	M_w	0.29	0.82
8	$\overline{G_F}$	0.11	0.94	E_F	0.21	0.91
9	$\Delta\sigma_E$	0.08	0.93	S_r	0.13	1.07
10	$\bar{\psi}$	0.01	1.00	$\overline{G_F}$	0.12	0.94

the prior and posterior distributions was also emphasized by Minson *et al.* (2014) in their Bayesian kinematic inversion of the great Tōhoku earthquake. In Tables 2 and 3 we list each model parameter and emergent quantity, respectively, sorted by $H_{0,1}$ and $H_{1,2}$. We note that since H is a metric, it satisfies the triangle inequality. For example, $|H_{0,1} - H_{1,2}| \leq H_{0,2} \leq H_{0,1} + H_{1,2}$.

To quantify the uncertainty of the parameters, we evaluate the size of their 68 per cent *highest density regions* (HDR). These are the regions that contain 68 per cent of the posterior density, such that the density within them is always larger than the density outside (Hyndman 1996). In particular, they always contain the largest mode of the distribution. For Gaussian distributions, the size of the HDR corresponds to two times their standard deviation. We list the modes and the HDRs of ρ_0, ρ_1 and ρ_2 in Table 1. We use the notation $p\ [_l M_u]$ to denote that the distribution of parameter p has mode M and l and u are the lower and upper bounds of its HDR. To compare the relative resolution of a parameter in distribution ρ_i and in its updated distribution ρ_j, we calculate the inverse ratio of their HDR sizes:

$$\mathrm{rHDR}_{i,j} = \frac{|\mathrm{HDR}\,(\rho_j)|}{|\mathrm{HDR}\,(\rho_i)|}\,. \tag{16}$$

Low values of $\mathrm{rHDR}_{i,j}$ signify good relative resolution. We list the values of $\mathrm{rHDR}_{0,1}$ and $\mathrm{rHDR}_{1,2}$ in Table 2 (for model parameters) and in Table 3 (for emergent quantities).

2-D marginal distributions for selected pairs of quantities are shown in Fig. 7. To quantify their correlation, we use the Spearman correlation coefficient (SC). 2-D distributions for a more extensive list of quantities are shown in Figs S7–S9.

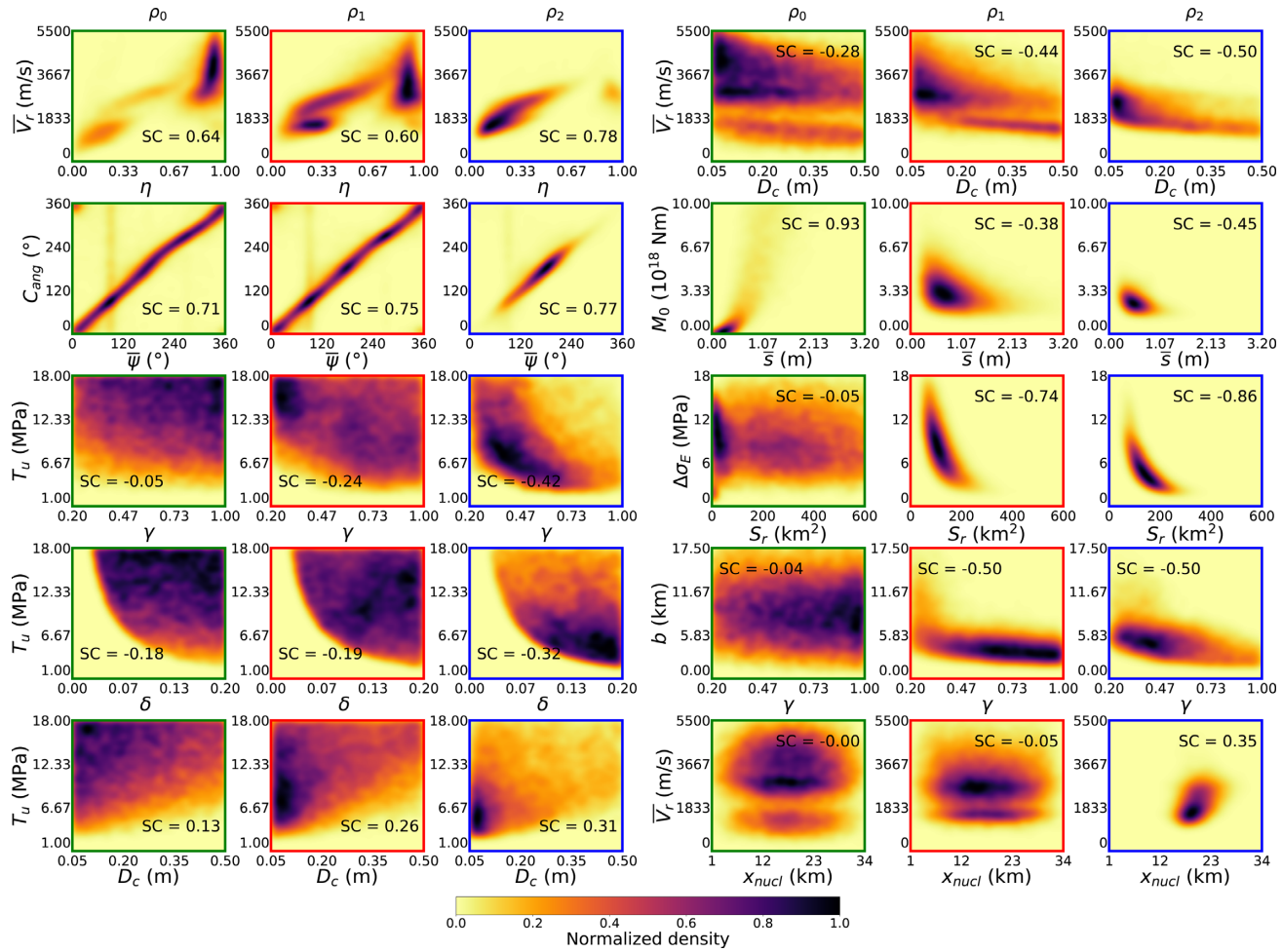

Figure 7. KDE estimates of 2-D marginal distributions of the posteriors ρ_0, ρ_1 and ρ_2 for selected pairs of quantities. The distributions are normalized by their respective maxima. SC denotes the value of the Spearman correlation coefficient.

3.2.1. *Model parameters*

The model parameters whose distributions appreciably change upon transition from ρ_{pr} to ρ_0 ($H_{pr,0} > 0.1$) are T_u, δ and R_{nucl} (Fig. 3). In ρ_{pr}, the distributions are flat, while in ρ_0 they are increasing (the sharp decay to zero at the boundaries is an artefact of the KDE method). This transition from a uniform to a monotonous distribution is also seen in parameters γ (increasing) and D_c (decreasing), although there the change is only small. Distributions of these five parameters change only negligibly after the update from ρ_0 to ρ_1 (Fig. 3), which shows that they are not further affected by the magnitude. Nevertheless, distributions of parameters T_u and γ do visibly change after the seismogram update ($H_{1,2} \approx 0.2$), developing peaks in the lower halves of their allowed ranges.

The only model parameters that strongly change in the transition from ρ_0 to ρ_1 are the lengths of the semi-major axis a and the semi-minor axis b, with $H_{0,1} = 0.33$ and 0.47, respectively. The ρ_1 distributions of both parameters peak at lower values than the ρ_0 distributions. In the case of a, the shift in the mode position is the main contributor to the change, and the uncertainty does not decrease very much ($\mathrm{rHDR}_{0,1} = 0.93$). In the case of b, there is a significant reduction of uncertainty ($\mathrm{rHDR}_{0,1} = 0.42$), the largest among model parameters in the $\rho_0 \rightarrow \rho_1$ transition. Nevertheless, the total area of the elliptic patch, $S_{el} = \pi ab$ (whose KDE is also shown in Fig. 3), gets resolved even more sharply than a or b individually ($\mathrm{rHDR}_{0,1} = 0.28$). The distributions of these three quantities are adjusted only a little when passing to the posterior ρ_2. Hence, they are mostly determined by moment magnitude.

The only other parameter for which ρ_0 and ρ_1 appreciably differ ($H_{0,1} > 0.1$) is the inclination angle of the patch, ϕ. In ρ_0, it peaks near 45° and 135°, as more elliptic patches fit within the fault (prior constraint #4 in Section 2.4) at these angles. Contrarily, the ρ_1 distribution of ϕ is much flatter, probably because the elliptic patches sampled from this posterior tend to be smaller, so their orientation only weakly affects the probability of them fitting within the fault. In ρ_2, the parameter has a single peak around 0° (≡ 180°), suggesting that elongation of the patch along strike is preferred when the seismograms are included. However, this peak is rather broad (the size of the HDR is 86°, which is 48 per cent of its possible range and $\mathrm{rHDR}_{1,2} = 0.72$). As we have shown in Section 3.1, ϕ considerably varies even among the best fitting models.

Finally, the ρ_2 distributions of the parameters x_{nucl}, y_{nucl} and C_{ang} are strongly determined by seismograms only, with $H_{1,2}$ of 0.48, 0.34 and 0.33, respectively. The along-strike location of the nucleation zone, x_{nucl}, has the greatest reduction in uncertainty among model parameters ($\mathrm{rHDR}_{1,2} = 0.36$). Its HDR size is 5.4 km, compared to the 6.4 km we observe in the along-dip location, y_{nucl} ($\mathrm{rHDR}_{1,2} = 0.47$). The distance of the most probable nucleation zone location and the GI-NOA hypocentre (see the 2-D distribution

of (x_{nucl}, y_{nucl}) in Fig. S7c) is 6.3 km. Finally, parameter C_{ang}, which is better resolved than ϕ (with the absolute HDR size equal to 31 per cent of its allowed range and $\mathrm{rHDR}_{1,2} = 0.48$), peaks around 180°. This corresponds to the nucleation zone being located to the ESE from the centre of the ellipse. We consider the fact that we can resolve the position of the nucleation zone and the rupture direction as evidence that the seismograms are sensitive to finite-fault features of the source.

3.2.2 *Emergent quantities*

For emergent quantities, the differences between the ρ_0 posterior and the other two posteriors are often more pronounced than for most model parameters and their Hellinger distances are larger on average (see Fig. 6 and Table 3). The ρ_0 posterior for M_w is wide and asymmetric, spanning values from 4.5 to 7.6, with a peak at 6.93. The ρ_1 posterior for M_w has a symmetric peak at 6.27, only slightly larger than the target value of $M_w^0 = 6.24$. The size of its HDR is 0.25. The ρ_2 posterior has a slightly lower mode (6.18) and smaller HDR (0.2).

Ruptures sampled from ρ_0 are fast, with mean slip-weighted rupture speed $\overline{V_r} \sim [_{2.3}3.0_{4.7}]$ km s^{-1} and tend to have large radiation efficiency η ($[_{0.75}0.92_{1.00}]$). Both quantities (which are strongly correlated in all distributions, see Fig. 7) are shifted towards lower values in the remaining two posteriors, especially in ρ_2. Since the shear wave speed β is between 3.36 and 3.95 km s^{-1} on most of the fault, an appreciable fraction of models sampled from ρ_0 and even ρ_1 have supershear average rupture speed. For ρ_2, this is true only for a negligible number of models. We note that for some models in ρ_2, the rupture speed surpasses β on a small part of the fault, but the supershear propagation is unsustained.

The total ruptured area S_r has similar distributions to S_{el} and like that parameter, it is strongly resolved by magnitude ($\mathrm{rHDR}_{0,1} = 0.22$, the lowest among all emergent quantities). However, the distributions of S_r are slightly denser in lower values, since not all models rupture the whole elliptic patch. Note that some distributions of the quantities η, $\overline{V_r}$, S_r and T_r have two peaks; the minor peaks correspond to the partially ruptured models.

The distributions of the slip-weighted mean stress drop $\Delta\sigma_E$ are similar for ρ_0 and ρ_1. In ρ_2, the peak of the distribution gets shifted to lower values by about 50 per cent and narrowed ($\mathrm{rHDR}_{1,2} = 0.63$), so that $\Delta\sigma_E$ $[_{2.2}3.6_{6.5}]$ MPa.

The quantity whose uncertainty is reduced the most when seismograms are used, having $\mathrm{rHDR}_{1,2} = 0.27$, is the radiated energy E_R ($[_{12}35_{65}]$ TJ in ρ_2). Its Hellinger distance $H_{1,2}$ is 0.50, the largest among both model and emergent quantities. In contrast, the fracture energy E_F ($[_{38}80_{174}]$ TJ in ρ_2) and its average area density $\overline{G_F}$ ($[_{0.1}0.4_{1.3}]$ MJ m^{-2}) have $\mathrm{rHDR}_{1,2} = 0.91$, $H_{1,2} = 0.21$ and $\mathrm{rHDR}_{1,2} = 0.94$, $H_{1,2} = 0.12$, respectively. This makes E_F and $\overline{G_F}$, together with S_r, the emergent quantities whose distributions are the least affected by the seismogram information.

To explore rupture directivity, we examine the slip-weighted mean direction of rupture $\bar{\psi}$, defined as the angle between the horizontal line and the slip-weighted mean rupture velocity vector $\frac{\int_{patch} \mathbf{p}/|\mathbf{p}|^2 s dS}{\int_{patch} s dS}$, where $\boldsymbol{p}(\boldsymbol{x})$ is the gradient of rupture time. It is strongly correlated to the model parameter C_{ang}, with SC > 0.7 for all posterior distributions (Fig. 7). Marginal distributions of ρ_0 and ρ_1 for $\bar{\psi}$ are nearly identical. Local peaks are distributed at multiples of 90°, but only the peaks at 0° and 180° have the same height. This is not surprising, because the dynamic problem is symmetric only with respect to the strike. The distribution of $\bar{\psi}$ changes substantially in the $\rho_1 \rightarrow \rho_2$ transition ($H_{1,2} = 0.38$) . Its $\mathrm{rHDR}_{1,2}$ is 0.46, the second biggest reduction in uncertainty among emergent quantities, after E_r. The HDR of the ρ_2 distribution lies approximately between 112° and 215°, which shows a strong seismogram preference for rupture spreading in the negative strike direction (towards WNW).

The final important quantity which is clearly affected by the waveform data is the equivalent duration of rupture T_r, calculated by dividing the seismic moment of each model by the maximum of its source time function. Even though the sizes of its ρ_1 and ρ_2 HDRs are similar ($\mathrm{rHDR}_{1,2} = 1.07$), the ρ_2 posterior peaks at ≈ 6.2 s, which is about 70 per cent larger than for ρ_1. We plot the KDEs of source time functions sampled from ρ_1 and ρ_2 in Fig. 8, in which the different duration is clearly visible. While the ρ_1 STFs (Fig. 8a) form a diffuse large peak at the beginning, the ρ_2 STFs (Fig. 8b) are sharper, consisting of two lower peaks connected by a 3–4 s long plateau.

4 DISCUSSION

We have performed a detailed Bayesian analysis of the dynamic inversion of the 2017 M_w 6.3 Lesvos earthquake, assuming an elliptic patch model. Thanks to careful sampling with the Parallel Tempering algorithm, we have identified how different quantities become resolved as progressively more information is added into the inversion. Specifically, we have analysed the chain of distributions $\rho_{pr} \rightarrow \rho_0 \rightarrow \rho_1 \rightarrow \rho_2$, corresponding to updating the prior (ρ_{pr}) by constraints on the rupture condition (ρ_0), magnitude (ρ_1) and seismograms (ρ_2), according to the Bayes' formula (4). We emphasize that if the inversion were performed by considering only the final posterior distribution, ρ_2, one could mistakenly attribute the resolution of a particular parameter only to the seismograms, even though the parameter may have gotten resolved mainly by the rupture condition or the magnitude information. Therefore, careful interpretation of results is needed.

We found that the final marginal posteriors ρ_2 of every model parameter are appreciably different from the prior distributions ρ_{pr}, with the minimum value of Hellinger distance $H_{pr,2}$ equal to 0.15 for parameter C_{dist} (see the black-and-blue numbers in Fig. 3). However, the update causing the biggest change varies among parameters. The quantities whose distributions are mainly affected by the rupture criterion (i.e. the greatest change happens in the $\rho_{pr} \rightarrow \rho_0$ transition, while the ρ_1 and ρ_2 distributions are very close to ρ_0) are δ and R_{nucl}. Parameters a and b are mostly constrained by the moment magnitude ($\rho_0 \rightarrow \rho_1$), as is the total ruptured area S_r. The distributions of the model parameters x_{nucl}, y_{nucl}, C_{ang} and γ, and emergent quantities E_r, η, $\bar{\psi}$ $\Delta\sigma_E$ and $\overline{V_r}$ are mainly affected by seismograms ($\rho_1 \rightarrow \rho_2$). Finally, the distributions of parameters D_c and C_{dist} do not change very much in either transition, but the small changes accumulate and their ρ_2 distributions are appreciably different from ρ_{pr}.

The seismogram information seems to push both T_u and γ towards lower values. This may be because seismograms require lower values of the background initial stress $T_u\gamma$, in accordance with the preference of lower values of the slip-weighted stress drop $\Delta\sigma_E$ in ρ_2 (Fig. 6). Indeed, T_u and γ are weakly negatively correlated in ρ_2 (see Fig. 7), suggesting a tendency to keep $T_u\gamma$ constrained. Low values of γ mean large contrast between the stress at the nucleation zone and the rest of the patch. They are also more likely to lead to subshear rupture speeds (Andrews 1976; Dunham 2007;

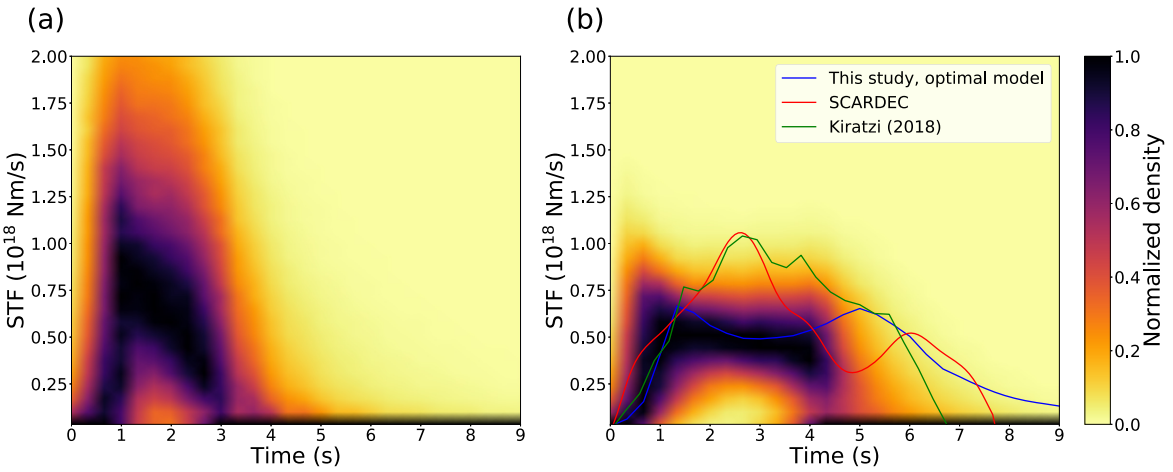

Figure 8. KDE estimates of source–time functions for (a) ρ_1 and (b) ρ_2, normalized at each time instance by the respective maximum. The source time function for the best-fitting model of this study is shown as the blue curve. The source–time functions from the SCARDEC database (Vallée & Douet 2016) and Kiratzi (2018) are shown as the red and green curves, respectively.

Xu *et al.* 2015), which could be significant for fitting the observed seismograms.

The change in distributions of parameters T_u, δ and R_{nucl} due to the rupture condition (the $\rho_{pr} \rightarrow \rho_0$ transition, see subsection 3.2.1) calls for an explanation. A natural way to interpret it is that there exists a parameter-dependent minimum nucleation radius, R_m, such that a model does not meet the rupture condition unless $R_{nucl} > R_m$. Since there is no analytic formula for R_m, we adopt an approximate closed-form expression for the critical radius from Galis *et al.* (2015):

$$\tilde{R}_m = \sqrt{\frac{3^3}{2^{11}}}\,\pi\,\frac{\mu D_c}{T_u \gamma^{3/2}\sqrt{(1+\delta-\gamma)}}. \tag{17}$$

It was derived for runaway ruptures, which break the whole patch, and should be considered as an upper bound for R_m. We plot the KDE of the non-dimensional ratio $R_{nucl}/\tilde{R}_m$ in Fig. 3. The distribution of the parameter changes significantly between ρ_{pr} (in which it is essentially decreasing from zero since the small peak is an artefact of the KDE method) and ρ_0, in which it peaks at ≈ 1.2. The remaining posteriors also peak near that value. Importantly, less than 1 per cent of all models sampled from ρ_0 have $R_{nucl}/\tilde{R}_m$ smaller than ≈ 0.3. We interpret this as the nucleation condition choosing only models with $R_{nucl}/\tilde{R}_m$ greater than this minimum value. It also explains why the ρ_0 distribution of D_c is decreasing, while the ρ_0 distributions of parameters T_u, δ and R_{nucl} are increasing. Further, since the parameters in eq. (17) can mutually compensate their effects on $\tilde{R}_m$, we can expect that they are not independent in the posteriors (as they are in ρ_{pr}). This is most notable in the 2-D marginal distributions of the pair (δ, T_u), which have a roughly semi-circular shape with non-zero values only when both T_u and δ are large enough (Fig. 7).

Another non-dimensional parameter, which was shown to control the transition between failed and successful rupture in elongated homogeneous barrier patches with sufficient width (Madariaga & Olsen 2000), is the *similarity* parameter κ, defined in our parametrization as:

$$\kappa = \frac{b T_u \gamma^2}{\mu D_c}. \tag{18}$$

Ruiz & Madariaga (2011, 2013), Twardzik *et al.* (2014), Herrera *et al.* (2017), and Mirwald *et al.* (2019) inverted for its value and obtained relatively similar values for well-fitting models, between 1.0 and 2.0. We plot its marginal distributions in Fig. 3. We find that its HDR in the ρ_2 posterior is 0.5–2.0, compared to its ρ_{pr} HDR of 0.14–3.5, which is a good reduction in uncertainty. The peak in all three posterior distributions is centred around similar values ($\approx 1.1 - 1.3$), but it gets narrower in both the $\rho_0 \rightarrow \rho_1$ ($\mathrm{rHDR}_{0,1} = 0.49$) and the $\rho_1 \rightarrow \rho_2$ ($\mathrm{rHDR}_{1,2} = 0.51$) transitions. The first sharpening is likely related to the resolution of parameter b by the magnitude information ($\mathrm{rHDR}_{0,1} = 0.42$), while the second is either caused by the resolution of γ ($\mathrm{rHDR}_{1,2} = 0.72$), or it cannot be attributed to a single parameter.

The quantity whose uncertainty is reduced the most when seismograms are used is the radiated energy E_R. That E_R is better resolved by seismograms than the fracture energy E_F seems reasonable, as it is the energy carried away from the fault by elastic waves that reach seismic stations. In addition, since for a fully fractured patch, the fracture energy is proportional to $T_u D_c S_r$, and S_r is well determined, there are two ways in which E_F could be resolved. Either both T_u and D_c need to be well resolved or they may be individually unresolved but mutually anticorrelated so that their product is approximately constant. Instead, neither parameter is strongly resolved and their correlation coefficients are small but positive in all posterior distributions (Fig. 7). The chosen frequency band (0.05–0.15 Hz) may be the reason for this weak (relative to other parameters) resolution. As higher frequencies become available, finer details of the earthquake might become visible, possibly permitting better resolution of D_c (Guatteri & Spudich 2000) and consequently E_F.

As seen above, statistical dependencies between pairs of parameters can be explored by examining 2-D marginal distributions. However, like the 1-D marginals, these need to be interpreted carefully. Some dependencies are observed in all three posteriors (Fig. 7), such as the positive correlation between $\overline{V_r}$ and η, expected from simple rupture-mechanical models (e.g. Kanamori & Brodsky 2004) or the weak negative correlation between $\overline{V_r}$ and D_c (noted also by Mirwald *et al.* 2019, in their inversion of the 2017 M_w 7.1 Puebla-Morelos earthquake). However, conservation of correlations across updates of distributions should not be taken for granted, even when we might expect the correlations from physical considerations. For example, since $M_0 = \mu S_r \bar{s}$, where $\bar{s}$ is the average slip, we would

expect M_0 to be positively correlated to $\bar{s}$. Indeed, the two parameters are strongly positively correlated in ρ_0 (SC $>$ 0.9). But when passing to ρ_1 and ρ_2, M_0 and $\bar{s}$ become weakly negatively correlated instead. We note that for M_0 and S_r, the parameters remain positively correlated, but the correlation weakens in the transition from ρ_0 to ρ_1 and ρ_2 (see Fig. S8).

On the other hand, quantities that are almost completely uncorrelated in ρ_0 may become correlated in ρ_1 and ρ_2 (Fig. 7). For example, we observe a weak positive correlation between x_{nucl} and $\overline{V_r}$ in ρ_2, that is models sampled from ρ_2 tend to rupture faster when nucleating farther along strike. The quantities $\Delta\sigma_E$ and S_r are another example; a strong negative correlation (also observed by Mirwald *et al.* 2019) appears upon the passage from ρ_0 to the magnitude constrained ρ_1. This scaling is what we would expect from the circular crack model, in which $M_0 \propto \Delta\sigma(S_r)^{3/2}$ (Madariaga 1979). The anticorrelation between emergent parameters $\Delta\sigma_E$ and S_r translates to an anticorrelation between model parameters γ and b (Fig. 7). We note that one of these parameters is geometric, while the other controls initial shear stress. Therefore, fixing geometric model parameters (e.g. to those of the best fitting model, Twardzik *et al.* 2014), and exploring only stress-frictional parameters underestimates the uncertainty of the latter, though faster exploration is obtained as a trade-off.

5 CONCLUSION

We have carried out a Bayesian dynamical inversion of the 2017 M_w 6.3 Lesvos earthquake using a single elliptic model with stress and friction defined by 12 parameters. We have calculated three different posterior probability distributions for those parameters by requiring 0) only that the model they describe continues to rupture after nucleation, (1) that it produces moment magnitude close to a value determined from a CMT inversion and (2) that it produces displacement waveforms that fit recordings from near regional stations.

We compared the distributions to assess the degree to which each piece of information constrains various parameters. We found that the parameters δ (nucleation stress overshoot relative to shear strength) and R_{nucl} (radius of nucleation zone) are sensitive mainly to the rupture condition, so that neither magnitude nor seismograms strongly constrain them further. Lengths of the semi-major and semi-minor axes of the elliptic patch, a and b, are mainly determined by moment magnitude. Finally, the along-strike and along-dip coordinates of the nucleation zone x_{nucl}, y_{nucl}, the relative angle between nucleation and centre of the patch C_{ang} and the ratio γ of the initial traction to the shear strength are determined chiefly by seismograms.

We also examined emergent quantities, obtained as results of dynamic rupture simulations. The surface area of rupture is the emergent quantity best resolved by the magnitude information. The quantity best resolved by the seismogram information is the radiated energy, with its final posterior distribution peaking at 35 TJ and with 68 per cent of values between 12 and 65 TJ. Most models sampled from the final posterior are characterized by slow mean rupture velocity (1.4–2.6 km s^{-1}), low radiation efficiency (10–40 per cent), and low slip-weighted mean stress drop (2.2–6.5 MPa). The rupture is further characterized by pronounced nucleation and subshear propagation directed predominantly towards WNW.

The presented analysis of resolvability of source parameters is limited to the considered elliptic models and seismogram frequency range and station distribution similar to this study. Nevertheless, our approach to assessing the roles of various constraints in Bayesian inversion by comparing 1-D distributions can be generalized to other dynamic earthquake source studies. Even for more complex source models with a large number of possibly correlated model parameters, the approach could prove useful in analysing gross rupture properties, such as radiated energy, seismic moment, average stress drop, or average rupture velocity.

ACKNOWLEDGEMENTS

We are grateful to two anonymous reviewers for suggestions that helped improve the manuscript. We thank the personnel of the Hellenic Seismic Network, National Seismic Network of Turkey (DDA) and Bogazici University Kandilli Observatory and Earthquake Research Institute (KOERI) for their work on the installation, operation and maintenance of stations whose waveforms were used in this paper. We thank Jan Premus for codeveloping the dynamic rupture simulation code FD3D_TSN and assisting with its use. We thank Malcolm Sambridge for making available his Fortran subroutines for the Parallel Tempering algorithm. We acknowledge financial support from the Charles university (projects GA UK No. 1698218 and SVV 115–09/260581). This work was supported by The Ministry of Education, Youth and Sports (cze) of the Czech Republic through the e-INFRA CZ (ID:90140).

DATA AVAILABILITY

All waveforms and seismic event locations used in this paper are publicly available from the Hellenic Seismic Network (National Observatory of Athens doi:10.7914/SN/HL, ITSAK doi:10.7914/SN/HI), National Seismic Network of Turkey (DDA), and Bogazici University Kandilli Observatory and Earthquake Research Institute (KOERI) (doi:10.7914/SN/KO). The dynamic rupture simulation code FD3D_TSN is publicly available at https://github.com/JanPremus/fd3d_TSN. The Parallel Tempering subroutines are available, upon registration, at http://www.ieearth.org.au/codes/ParallelTempering/. The files containing all the probability distribution samples, as well as the Python scripts used for their visualization and statistical analysis, will be shared on request to the corresponding author.

REFERENCES

Andrews, D.J., 1976. Rupture velocity of plane strain shear cracks, *J. geophys. Res.,* **81**(32), 5679–5687.

Aochi, H. & Ulrich, T., 2015. A probable earthquake scenario near Istanbul determined from dynamic simulations, *Bull. seism. Soc. Am.,* **105**(3), 1468–1475.

Basili, R. *et al.*, 2013. The European Database of Seismogenic Faults (EDSF) compiled in the framework of the Project SHARE. http://diss.rm.ingv.it/share-edsf/, doi: 10.6092/INGV.IT-SHARE-EDSF.

Berenger, J.-P., 1994. A perfectly matched layer for the absorption of electromagnetic waves, *J. Comput. Phys.,* **114**(2), 185–200.

Bizzarri, A., 2011. On the deterministic description of earthquakes, *Rev. Geophys.,* **49,** RG3002.

Burjánek, J. & Zahradník, J., 2007. Dynamic stress field of a kinematic earthquake source model with k-squared slip distribution, *Geophys. J. Int.,* **171,** 1082–1097.

Chatzipetros, A., Kiratzi, A., Sboras, S., Zouros, N. & Pavlides, S., 2013. Active faulting in the north-eastern Aegean Sea Islands, *Tectonophysics,* **597-598,** 106–122.

Chen, J. & Spiers, C.J., 2016. Rate and state frictional and healing behavior of carbonate fault gouge explained using microphysical model, *J. geophys. Res.,* **121,** 8642–8665.

Chousianitis, K. & Konca, A.O., 2018. Coseismic slip distribution of the 12 June 2017 Mw = 6.3 Lesvos earthquake and imparted static stress changes to the neighboring crust, *J. geophys. Res.,* **123,** 8926–8936.

Corish, S., Bradley, C. & Olsen, K., 2007. Assessment of a nonlinear dynamic rupture inversion technique applied to a synthetic earthquake. *Bull. seism. Soc. Am.,* **97,** 901–914.

Cotton, F. & Coutant, O., 1997. Dynamic stress variations due to shear faults in a plane-layered medium, *Geophys. J. Int.,* **128**(3), 676–688.

Di Carli, S., François-Holden, C., Peyrat, S. & Madariaga, R., 2010. Dynamic inversion of the 2000 Tottori earthquake based on elliptic subfault approximations, *J. Geophys. Res.,* **115**(B12).

Díaz-Mojica, J., Cruz-Atienza, V.M., Madariaga, R., Singh, S.K., Tago, J. & Iglesias, A., 2014. Dynamic source inversion of the M6.5 intermediate depth Zumpango earthquake in central Mexico: a parallel genetic algorithm, *J. geophys. Res.,* **119**(10), 7768–7785.

Dieterich, J.H., 1979. Modeling of rock friction: 1. Experimental results and constitutive equations, *J. geophys. Res.,* **84**(B5), 2161–2168.

Dunham, E., 2007. Conditions governing the occurrence of supershear rupture under slip weakening friction, *J. geophys. Res.,* **112**(B7), doi:10.1029/2006JB004717.

Fukuyama, E. & Mikumo, T., 1993. Dynamic rupture analysis: inversion for the process of the 1990 Izu-Oshima, Japan, earthquake (M = 6.5), *J. geophys. Res.,* **98,** 6529–6542.

Galis, M., Pelties, C., Kristek, J., Moczo, P., Ampuero, J.-P. & Mai, P.M., 2015. On the initiation of sustained slip-weakening ruptures by localized stresses, *Geophys. J. Int.,* **200,** 890–909.

Gallovič, F., Valentová, Ľ., Ampuero, J.-P. & Gabriel, A.-A., 2019a. Bayesian dynamic finite-fault inversion: 1. Method and synthetic test, *J. geophys. Res.,* **124,** 6949–6969.

Gallovič, F., Valentová, Ľ., Ampuero, J.-P. & Gabriel, A.-A., 2019b. Bayesian dynamic finite-fault inversion: 2. Application to the 2016 Mw6.2 Amatrice, Italy, Earthquake, *J. geophys. Res.,* **124,** 6970–6988.

Gallovič, F., Zahradník, J., Plicka, V., Sokos, E., Evangelidis, Ch., Fountoulakis, I. & Turhan, F., 2020. Complex rupture dynamics on an immature fault during the 2020 Mw 6.8 Elazığ earthquake, Turkey, *Commun. Earth Environ,* **1**(40), doi:10.1038/s43247-020-00038-x.

Ganas, A., Oikonomou, I. & Tsimi, C., 2013. NOAfaults: a digital database for active faults in Greece, *Bull. Geol. Soc. Greece*. **47,** 518–530.

Graves, R.W., 1996. Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences, *Bull. seism. Soc. Am.,* **86,** 1091–1106.

Guatteri, M. & Spudich, P., 2000. What can strong-motion tell us about slip-weakening fault-friction law?, *Bull. seism. Soc. Am.,* **90**(1), 98–116.

Hallo, M. & Gallovič, F., 2016. Fast and cheap approximation of Green functions uncertainty for waveform-based earthquake source inversions, *Geophys. J. Int.,* **207,** 1012–1029.

Hastings, W.K., 1970. Monte Carlo sampling methods using Markov Chains and their applications. *Biometrika*., **57**(1), 97–109.

Herrera, C., Ruiz, S., Madariaga, R. & Poli, P., 2017. Dynamic inversion of the 2015 Jujuy earthquake and similarity with other intraslab events, *Geophys. J. Int,* **209**(2), 1 866–875.

Hyndman, R.J., 1996, Computing and graphing highest density regions, *Am. Stat.,* **50**(2), 120–126.

Ida, Y., 1972. Cohesive force across tip of a longitudinal-shear crack and Griffith specific surface-energy, *J. geophys. Res.,* **77,** 3796–3805.

Ide, S. & Takeo, M., 1997, Determination of constitutive relations of fault slip based on seismic wave analysis, *J. geophys. Res.,* **102 (** B12**)**, 27 379–27 391.

Kanamori, H. & Brodsky, E.E., 2004. The physics of earthquakes, *Rep. Prog. Phys.,* **67**(8), 1429–1496.

Karagianni, E.E. *et al.*, 2002. Rayleigh wave group velocity tomography in the Aegean area, *Tectonophysics,* **358,** 187–209.

Kiratzi, A., 2018. The 12 June 2017 Mw 6.3 Lesvos Island (Aegean Sea) earthquake: slip model and directivity estimated with finite-fault inversion, *Tectonophysics,* **724-725,** 1–10.

Konstantinou, K.I., Mouslopoulou, V., Liang, W.-T., Heidbach, O., Oncken, O. & Suppe, J., 2017. Present-day crustal stress field in Greece inferred from regional-scale damped inversion of earthquake focal mechanisms, *J. geophys. Res.,* **122,** 506–523

Kristek, J., Moczo, P. & Archuleta, R.J., 2002. Efficient methods to simulate planar free surface in the 3D 4th-order staggered-grid finite-difference schemes, *Stud. Geophys. Geod.,* **46,** 355–381.

Lekkas, E., Carydis, P., Skourtsos, E., Mavroulis, S., Andreadakis, E., Antoniou, V. & Spyrou, N.I., 2017. Factors controlling the distribution of building damage in the traditional Vrissa settlement induced by the 2017 June 12, Mw 6.3 Lesvos earthquake (Northeastern Aegean Sea, Greece), in *Proceedings of the PATA DAYS 2017: 8th International Workshop on Paleoseismology, Active Tectonics and Archeoseismology,* 13–16 November 2017, New Zealand.

Levander, A., 1988. Fourth-order finite-difference P-SV seismograms. *Geophysics,* **53,** 1425- 1436.

Madariaga, R., 1979. On the relation between seismic moment and stress drop in the presence of stress and strength heterogeneity, *J. geophys. Res.,* **84,** 2243–2250.

Madariaga, R. & Olsen, K.B., 2000. Criticality of rupture dynamics in 3D, *Pure appl. Geophys.,* **157**(11–12), 1981–2001.

Marone, C., Cocco, M., Richardson, E. & Tinti, E., 2009. The critical slip distance for seismic and aseismic fault zones of finite width, *Int. Geophys.,* **94,** 135–162.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A. & Teller, E., 1953. Equations of state calculations by fast computing machines, *J. Chem. Phys.,* **21,** 1087–1092.

Minson, S.E. *et al.*, 2014. Bayesian inversion for finite fault earthquake source models–II: the 2011 Great Tohoku-oki, Japan earthquake, *J. geophys. Int.,* **198**(2), 922–940.

Mirwald, A., Cruz-Atienza, V.M., Díaz-Mojica, J., Iglesias, A., Singh, S.K., Villafuerte, C. & Tago, J., 2019. The 19 September 2017 (Mw 7.1) intermediate-depth Mexican earthquake: a slow and energetically inefficient deadly shock. *Geophys. Res. Lett.,* **46,** 2054–2064.

Ohnaka, M. & Yamashita, T., 1989. A cohesive zone model for dynamic shear faulting based on experimentally inferred constitutive relation and strong motion source parameters, *J. geophys. Res.,* **94**(B4), 4089–4104.

Papadimitriou, P. *et al.*, 2018. The 12th June 2017 Mw = 6.3 Lesvos earthquake from detailed seismological observations, *J. Geodyn.,* **115,** 23–42.

Perfettini, H., Campillo, M. & Ionescu, I. 2003. On the scaling of the slip weakening rate of heterogeneous faults, *J. geophys. Res.,* **108**(B9), 2410.

Peyrat, S. & Olsen, K.B., 2004. Nonlinear dynamic rupture inversion of the 2000 Western Tottori, Japan, earthquake, *Geophys. Res. Lett.,* **31,** n/a.

Peyrat, S., Olsen, K.B. & Madariaga, R., 2001. Dynamic modeling of the 1992 Landers earthquake, *J. geophys. Res.,* **106**(B11), 26467–26482.

Premus, J., Gallovič, F., Hanyk, L. & Gabriel, A.-A., 2020. FD3D_TSN: fast and simple code for dynamic rupture simulations with GPU acceleration, *Seismol. Res. Lett.,* **91,** 2881–2889

Pulido, N. & Irikura, K., 2000. Estimation of dynamic rupture parameters from the radiated seismic energy and apparent stress, *Geophys. Res. Lett,* **27,** 3945–3948.

Ripperger, J., Ampuero, J.-P., Mai, M.P. & Giardini, D., 2007. Earthquake source characteristics from dynamic rupture with constrained stochastic fault stress, *J. geophys. Res.,* **112,** 1–17.

Ruina, A., 1983. Slip instability and state variable friction laws, *J. geophys. Res.,* **88,** 10 359–10 370.

Ruiz, S. & Madariaga, R., 2011. Determination of the friction law parameters of the Mw6.7 Michilla earthquake in northern Chile by dynamic inversion, *Phys. Res. Lett.,* **38,** L09317.

Ruiz, S. & Madariaga, R., 2013. Kinematic and dynamic inversion of the 2008 northern Iwate earthquake, *Bull. seism. Soc. Am.,* **103**(2A), 694–708.

Sambridge, M., 2013. A Parallel Tempering algorithm for probabilistic sampling and multimodal optimization, *Geophys. J. Int.,* **196,** 357–374.

Sambridge, M. & Mosegaard, K., 2002. Monte Carlo methods in geophysical inverse problems, *Rev. Geophys.,* **40**(3), 3–1.

Scott, D.W., 1992. *Multivariate Density Estimation: Theory, Practice, and Visualization,* John Wiley & Sons.

Shemyakin, 2014. Hellinger distance and non-informative priors, *Bayesian Anal.,* **9**(4), 923–938.

Sokos, E. & Zahradník, J., 2017. Lesvos June 12, 2017, Mw 6.3 event, a quick study of the source, EMSC report (https://www.emsc-csem.org/Files/news/Earthquakes_reports/Lesvos_Source_Study_Sokos_and_Zahradnik.pdf, accessed 15th Sep. 2021).

Swendsen, R.H. & Wang, J.S., 1986. Replica Monte Carlo simulation of spin glasses, *Phys. Rev. Lett.,* **57,** 2607–2609.

Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM.

Tinti, E., Spudich, P. & Cocco, M., 2005. Earthquake fracture energy inferred from kinematic rupture models on extended faults, *J. geophys. Res.,* **110,** B12303.

Twardzik, C., Das, S. & Madariaga, R., 2014. Inversion for the physical parameters that control the source dynamics of the 2004 Parkfield earthquake, *J. geophys. Res.,* **119**(9), 7010–7027.

Vallée, M. & Douet, V., 2016. A new database of source time functions (STFs) extracted from the SCARDEC method. *Phys. Earth planet. Inter.,* **257,** 149–157.

Viesca, R.C. & Garagash, D.I., 2015. Ubiquitous weakening of faults due to thermal pressurization, *Nat. Geosci.,* **8,** 875–879.

Wasserman, L., 2004. *All of Statistics: A Concise Course in Statistical Inference,* Springer.

Xu, J., Zhang, H. & Chen, X., 2015. Rupture phase diagrams for a planar fault in 3-D full-space and half-space, *Geophys. J. Int.,* **202**(3), 2194–2206.

Zahradník, J. & Sokos, E., 2018. ISOLA code for multiple-point source modeling—review, in *Moment Tensor Solutions—A Useful Tool for Seismotectonics,* pp. 1–28, ed. D'Amico, S., Springer.

Zambom, A.Z. & Dias, R., 2013. A review of kernel density estimation with applications to econometrics, *Int. Econ. Rev.,* **5**(1), 20–42.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

Figure S1: Map view of the stations used in the centroid moment tensor inversion (Text S1 and Figs S2–S5). The blue star is the GI-NOA hypocentre.

Figure S2: Horizontal grid search (36 trial positions, increments of 2 km) for the centroid at a depth of 8 km. The star shows the NOA epicentre for reference. The centroid position, characterized by the largest correlation between observed and synthetic displacements in the frequency range of 0.03 and 0.07 Hz, located 2 km west and 6 km north from the epicentre, is shown by the largest beachball, trial position #28. The balls are colour-coded according to the double-couple percentage (all >90 per cent). Note the stability of the focal mechanism across the grid.

Figure S3: Waveform fit for the CMT solution (position #28 in Fig. S2, strike/dip/rake = 113°/40°/–83°, frequency range from 0.03 0.07 Hz, the total variance reduction VR = 79 per cent). Real and synthetic displacements are shown as black and red lines, respectively. The blue numbers denote the variance reductions of the individual components.

Figure S4: (a) A three-point source model. Three point-source models with deviatoric moment tensors are searched by iterative deconvolution of ISOLA (Zahradník & Sokos 2018) software in the assumed fault plane, frequency range 0.05–0.10 Hz. The sources are shown by circles, sized according to their scalar moment, and colour-coded according to their rupture time ($t = 0$ is the origin time). The large black circle at the top is a referential moment scale, not a solution. The 3-point solution was calculated repeatedly, each time removing one station (jackknifing), thus producing multiple triplets of circles (some of them coincide) to estimate uncertainty. The plot demonstrates a stable evolution of the rupture from the southeast to the northwest. The process starts with an early subevent (turquoise-coloured) near the epicentre (marked by an isolated diamond near the trial position #52), continues with the major subevent in the middle of the fault, and ends with a third moment-release episode at about 6–8 seconds after origin time. Variance reduction from the stations shown in Fig. S5 varied during jackknifing from VR = 0.69–0.76. The focal mechanism of the major subevent was almost constant, the mechanisms of the remaining two subevents varied within 30° Kagan angle from the CMT solution during the jackknifing. (b) The mechanisms of the best solution from all stations. The largest subevent is at position #38, followed by positions #32 and #12. Strong stability of the mechanism is seen.

Figure S5: Waveform fit for the three-point source model of Fig. S4(b). The frequency range is 0.05–0.10 Hz, the total variance reduction VR = 71 per cent (when considering only the two largest subevents, or just a single largest subevent, variance reduction drops to VR = 66 and 60 per cent, respectively). The real and synthetic displacements are shown as black and red lines, respectively. The blue numbers denote the variance reductions of the individual components.

Figure S6: The seismic velocity and mass density model assumed in the inversion adapted from Karagianni *et al.* (2002) by removing its low-velocity channel. The quality factors are Q_p =300 and Q_s =150 for depths smaller than 32 km. At larger depths, $Q_P = Q_S$ = 1000.

Figure S7: KDE estimates of 2-D marginal posteriors of (a) ρ_0, (b) ρ_1 and (c) ρ_2 for pairs of model parameters. The letters SC stand for Spearman's correlation coefficient. The colour scale is relative to the maximum value of the distribution in each plot. The blue star and the red circle in the last panel of each figure denote the position of the GI-NOA hypocentre and the centroid inferred by our CMT inversion, respectively.

Figure S8: KDE estimates of 2-D marginal posteriors of (a) ρ_0, (b) ρ_1 and (c) ρ_2 for pairs of emergent quantities. The letters SC stand for Spearman's correlation coefficient. The colour scale is relative to the maximum value of the distribution in each plot.

Figure S9: KDE estimates of 2-D marginal posteriors of (a) ρ_0, (b) ρ_1 and (c) ρ_2 for pairs of model (vertical axes) and emergent (horizontal axes) quantities. The letters SC stand for Spearman's correlation coefficient. The colour scale is relative to the maximum value of the distribution in each plot.

Table S1: Comparison of centroid solutions found by different studies. Only the nodal plane used in this study is shown.

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