Patterns of earthquakes and aseismic slip on a heterogeneous strike-slip fault with static/kinetic friction and temperature-dependent creep

Xiaoyu Zhou¹ and Yehuda Ben-Zion¹

¹University of Southern California

May 28, 2024

Abstract

An earthquake simulator is developed to study the dynamics of seismicity and seismic/aseismic slip partitioning on a heterogeneous strike-slip fault using a generalized model of a discrete fault governed by static/dynamic friction and creep in an elastic half-space. Previous versions of the simulator were shown to produce various realistic seismicity patterns (e.g., frequency-magnitude event statistics, hypocenter and slip distributions, temporal occurrence) using friction levels and creep properties that vary in space but are fixed in time. The new simulator incorporates frictional heat generation by earthquake slip leading to temperature rises, subsequent diffusion cooling into the half space, and time-dependent creep on the fault. The model assumes a power law dependence of creep velocity on the local shear stress, with temperature-dependent coefficients based on the Arrhenius equation. Temperature rises due to seismic slip produce increased aseismic slip, which can lead to further stress concentrations, aftershocks, and heat generation in a feedback loop. The partitioning of seismic/aseismic slip and space-time evolution of seismicity are strongly affected by the temperature changes on the fault. The results are also affected significantly by the difference between the static and kinetic friction levels. The model produces realistic spatio-temporal distribution of seismicity, transient aseismic slip patterns, foreshock-mainshock-aftershock sequences, and a bimodal distribution of earthquakes with background and clustered events similar to observations. The simulator (EQsim) may be used to clarify relations between fault properties and different features of seismicity and aseismic slip, and to improve the understanding of failure patterns preceding large earthquakes.
Patterns of earthquakes and aseismic slip on a heterogeneous strike-slip fault with static/kinetic friction and temperature-dependent creep

Xiaoyu Zhou¹, Yehuda Ben-Zion¹,²
1. University of Southern California, Department of Earth Science
2. Statewide California Earthquake Center

Key Points

- We develop an earthquake simulator for failure events on a large heterogeneous strike-slip fault in an elastic half-space.

- The simulator includes a feedback mechanism between frictional heat generated by earthquake slip and subsequent creep slip on the fault.

- The feedback facilitates the generation of spatio-temporal clustering of earthquakes on the fault including realistic aftershock sequences.
Abstract
An earthquake simulator is developed to study the dynamics of seismicity and seismic/aseismic slip partitioning on a heterogeneous strike-slip fault using a generalized model of a discrete fault governed by static/dynamic friction and creep in an elastic half-space. Previous versions of the simulator were shown to produce various realistic seismicity patterns (e.g., frequency-magnitude event statistics, hypocenter and slip distributions, temporal occurrence) using friction levels and creep properties that vary in space but are fixed in time. The new simulator incorporates frictional heat generation by earthquake slip leading to temperature rises, subsequent diffusion cooling into the half space, and time-dependent creep on the fault. The model assumes a power law dependence of creep velocity on the local shear stress, with temperature-dependent coefficients based on the Arrhenius equation. Temperature rises due to seismic slip produce increased aseismic slip, which can lead to further stress concentrations, aftershocks, and heat generation in a feedback loop. The partitioning of seismic/aseismic slip and space-time evolution of seismicity are strongly affected by the temperature changes on the fault. The results are also affected significantly by the difference between the static and kinetic friction levels. The model produces realistic spatio-temporal distribution of seismicity, transient aseismic slip patterns, foreshock-mainshock-aftershock sequences, and a bimodal distribution of earthquakes with background and clustered events similar to observations. The simulator (EQsim) may be used to clarify relations between fault properties and different features of seismicity and aseismic slip, and to improve the understanding of failure patterns preceding large earthquakes.

Plain language summary
The physics governing the dynamics of seismicity is currently not fully understood because of the inherent complexity of the problem and the lack of data covering many large earthquake cycles. Large catalogs of seismic and aseismic events generated for different sets of model parameters can aid the understanding of seismicity patterns. Here we develop a simulator of earthquakes and aseismic slip events on a large heterogeneous strike-slip fault in a 3D elastic solid that includes temperature changes on the fault from frictional heat generated by earthquake slip. The simulator generates a wide variety of slip and seismicity patterns including transient aseismic slip following earthquakes, different forms of frequency-size event statistics, and spatio-temporal clustering of earthquakes on the fault including realistic aftershock sequences. The results produced by the simulator may be used to study correlations between different signals and clarify processes leading to large earthquakes.

1. Introduction
Earthquakes are a prime example of a complex natural process with far-from-equilibrium strongly nonlinear dynamics that is not well understood, with substantial societal and economic relevance for large populations worldwide (e.g., Ben-Zion et al., 2022). The lack of quantitative data on timescales capturing multiple large earthquake cycles is a fundamental impediment for progress in the field. Numerical simulations provide the only path for overcoming the lack of data and elucidating spatio-temporal patterns and long-term dynamics that extend the knowledge beyond sporadic case
studies and regional statistical laws. Simulation results that are constrained by a wide variety of observations can allow us to develop possible causal relations between physical variables and responses of faults, clarify processes leading to large events, and enhance the general understanding of earthquake and fault dynamics on various space-time scales. Laboratory results from friction, fracture, and other experiments provide fundamental insights on small-scale failure processes. However, it is unknown how to scale-up microscopic processes on small-scale laboratory samples under the experimental conditions to the complex microscopic processes in the crust with strong geometrical and material heterogeneities, non-steady-state loadings over wide ranges of periods and amplitudes, etc. It is therefore essential to develop and use earthquake simulators that are constrained with the available observations of seismicity. The simulators should be constrained with diverse observations to reduce the non-uniqueness inherent to any one type of results (e.g., frequency-size or temporal statistics).

Various types of earthquake simulators have been developed in the last few decades, focusing among other topics on understanding spatio-temporal seismicity patterns, clarifying relations between seismic and aseismic slip on faults, developing scaling relations, and simulating earthquake catalogs to aid probabilistic forecasting of large earthquakes. Example simulators include block-spring models and cellular automata with nearest-neighbor stress interactions (e.g., Burridge and Knopoff, 1967; Carlson and Langer, 1989; Rundle and Klein, 1993), discrete models of heterogeneous large faults with long-range stress interactions and slip patches that can fail independently of their neighbors (e.g., Ben-Zion and Rice, 1993; Robinson and Benites, 1995; Zöller and Ben-Zion, 2014; Zielke et al., 2015), continuum models of faults where patches can only fail as part of a coordinated failure with their neighbors (e.g., Rice and Ben-Zion 1996; Lapusta et al., 2000; Hillers et al., 2006; Erickson et al., 2020; Abdelmeguid and Elbanna, 2022), hybrid discrete-continuum models of fault networks (e.g., Ward, 1996; Pollitz, 2009; Richards-Dinger and Dieterich, 2012; Shaw et al., 2018; Im and Avouac, 2023), and hybrid models accounting for coupled evolution of earthquakes and faults (e.g., Lyakhovsky et al., 1997; 2001; Ben-Zion et al., 1999). For a more detailed overview of earthquake simulators, see Rundle et al. (2003), Ben-Zion (2008), Tullis et al. (2012) and Zielke and Mai (2023).

Ben-Zion and Rice (1993, 1995) and Ben-Zion (1996) showed that discrete fault models, described further in section 2, with a fundamental segmentation length represented by the fault discreteness and long-range stress interactions, provide powerful tools for studying multiple aspects of earthquake patterns including frequency-size and temporal statistics, spatial distribution of events along strike and with depth, and more. However, these models did not account for aftershock sequences and other aspects of clustering that are observed features of seismicity (e.g. Utsu, 2002; Zaliapin and Ben-Zion, 2022). Statistical analyses of regional and global earthquake catalogs demonstrate that earthquakes can be generally described as a superposition of background and triggered events (e.g., Kagan and Jackson, 1991; Ogata, 1999; Gu et al., 2013). This concept is at the core of the stochastic epidemic type aftershock sequence model (ETAS) that provides the standard statistical description of seismicity (Ogata, 1998, 1999).

Following Zaliapin et al. (2008) and Zaliapin and Ben-Zion (2013a,b), a robust data-driven way to separate between background and triggered events is to use a proximity measure between pairs of earthquakes defined (Baiesi and Paczuski, 2004) as
\[ \eta_{ij} = (t_j - t_i)10^{-qbm_i} \cdot (r_{ij})^{d_f} 10^{-(1-q)bm_i}, t_j > t_i \]

where \( t_j - t_i \) is the time between events \( i \) and \( j \), \( r_{ij} \) is the spatial distance between the events, \( m_i \) is the magnitude of the earlier event \( i \), and \( d_f \) is the fractal dimension of the hypocenter (or epicenter) distribution. In data analysis, it is convenient to use rescaled proximities \( R \) and \( T \) defined as

\[ R_{ij} = (r_{ij})^{d_f} 10^{-(1-q)bm_i}, T_{ij} = t_{ij} 10^{-qbm_i}. \]

Identifying for each earthquake \( j \) a parent event \( i \) that has the closest proximity \( \eta \) among all earlier events in the catalog based on Eq. (1), and plotting the rescaled nearest-neighbor proximities in 2D diagrams of \( \log R \) vs. \( \log T \) leads generally to a bi-modal distribution that facilitates the classification of seismicity to background and clustered (triggered) events (Zaliapin and Ben-Zion 2013a,b; 2016; Ruhl et al., 2016; Martínez-Garzón et al., 2019; Goebel et al., 2019). For observed examples of such bi-modal distributions, see Figure 4 of Zaliapin and Ben-Zion (2013a) and additional observational results in the papers cited above. The main goal of this work is to develop a deterministic earthquake simulator incorporating a few basic mechanical processes that can be used to study observed aspects of seismicity and aseismic slip, including clustering in relation to the assumed fault properties.

The developed simulator (EQsim) generalizes the discrete fault model of Ben-Zion (1996) with quenched (space-dependent) friction and creep properties by adding time-dependency to the creep coefficients. The time-dependency of creep coefficients and aseismic slip on the fault is implemented by calculating frictional heat from each seismic slip, converting the heat to a temperature increase in the slip zone, and coupling the evolving temperature to changes of creep coefficients using the Arrhenius equation. Between seismic slip episodes the temperature at each slip patch is reduced through fault-normal heat conduction to the bulk. The evolution of stress, displacement, and temperature field is calculated in a quasi-static manner. The dynamics of the system is governed by the distributions of the frictional and creep properties, parameters determining the temperature changes due to seismic slip, and the long-range stress transfer due to seismic and aseismic fault motion. The thermal feedback loop between seismic displacements and fault creep properties produces post-seismic slip that leads to clustering of earthquakes on the fault. A parameter-space study is used to develop correlations between fault properties and different aspects of seismicity. For a range of parameters, the simulations produce a bimodal distribution of earthquakes between background and clustered events similar to observations.

In the next section, we describe the EQsim and the relevant physical parameters that affect the results. Section 3 illustrates different types of simulation output, with an emphasis on earthquake clustering in space-time and the bimodal distribution of earthquakes between background and clustered events, which to our knowledge have not been simulated so far by any other physical model. The ability of the model to simulate a wide range of observations provides a path to understand processes and parameters that produce different types of seismicity. The results and possible continuations of the research are discussed in the final section 4.
2. Model

The model consists of a vertical cellular fault zone embedded in a 3-D elastic half-space (Fig. 1). The half-space has a quasi-brittle upper layer that can fail by a combination of frictional and creep processes, over a lower crust and upper mantle sections where constant sliding is imposed at a plate rate $v_{plate} = 35 \text{ mm/yr}$. The model assumes that the first order slip process in a large fault that is considerably narrower than the along-strike and depth dimensions zone can be mapped onto an ideal planar fault represented by a collection of cellular slip patches on a 2-D computational grid in a half-space (Ben-Zion and Rice 1993, 1995; Ben-Zion, 1996). At each grid position $(x, z; y=0)$, time-dependent frictional and creep deformations are calculated to represent the corresponding deformation of a volumetric fault zone section centered around line $(x + \Delta x/2, z + \Delta z/2)$ with a cross section $(\Delta x, \Delta z)$, where $\Delta x$ and $\Delta z$ denote length and depth of numerical cells along the fault, respectively.

![Figure 1](image1.png)

Figure 1. (a) A schematic representation of a heterogeneous strike-slip fault embedded in a 3D elastic half-space. Constant plate motion of 35 mm/yr is applied on both sides and below the computational grid, which has dimensions of 70 km by 17.5 km. Each cell in the grid has creep velocity based on local shear stress, activation energy, and temperature. (b) (Left) Initial Strength profile of the fault. Cells within
7.5 km on both sides are considered boundary cells where the creep velocity gradually increases to plate motion. Thermal effect is also turned off within the boundary region by setting activation energy to zero. (Right) Failure envelope as the minimum between static frictional strength and creep strength. The brittle-ductile transition depth can vary based on the chosen activation energy value.

The stress level at fault location \((i, j)\) and time \(t\), generated by boundary conditions and deforming fault patches, is computed from the slip distribution along the fault using a discretized boundary integral

\[
\tau(i, j, t) = \sum_{k, l} K(i, j; k, l)[V_p t - u(k, l, t)],
\]

(2)

where \(K(i, j; k, l)\) is stress transfer for a unit slip over a rectangular patch (Chinnery, 1961). The source dislocation term in the square bracket from Eq. (2) is the slip deficit of cell \((k, l)\) at time \(t\) with respect to the plate motion. The total slip \(u(k, l, t)\) at each cell location \((k, l)\) is a combination of aseismic slip from creep process operating between earthquakes and seismic slip generated by earthquake brittle failures.

Following Ben-Zion and Rice (1993, 1995), the brittle process is governed by a spatial distribution of static frictional strength \(\tau_s\), dynamic strength \(\tau_d\), and arrest stress \(\tau_a\). The model assumes that the \(\tau_s\) increases linearly with depth \(z\) and is given by

\[
\tau_s(z) = \tau_0 + f_s (\rho_s - \rho_w) g z,
\]

(3)

where \(\tau_0\) is cohesion, \(f_s\) is the static coefficient of friction assumed to have a uniform value of 0.75 along the fault, \(\rho_s = 2600 \text{ kg/m}^3\) and \(\rho_w = 1000 \text{ kg/m}^3\) are densities of rock and water, respectively, and \(g\) is the gravitational acceleration. Eq. (3) produces an effective normal stress gradient of 18 MPa/km. The stress at any fault location (cell) changes both gradually due to far-field tectonic loading and creep deformations at all fault locations, and abruptly due to brittle failures along the fault. When the stress level \(\tau(x, z)\) at the center of a cell reaches the static strength \(\tau_s\), brittle failure occurs. The strength drops to a dynamic level \(\tau_d < \tau_s\) and the cell slips as to reduce the local stress to an lower arrest level \(\tau_a < \tau_d\) to account for a dynamic overshoot (e.g., Madariaga, 1976).

The static strength, dynamic strength, and arrest stress are related as

\[
\varepsilon = (\tau_s - \tau_a)/(\tau_s - \tau_a),
\]

(4)

where \(\varepsilon\) is a strength-change parameter. Positive values of \(\varepsilon\) correspond to dynamic weakening during failure appropriate for brittle deformation in the upper crust, while negative values correspond to dynamic strengthening appropriate for regions sustaining ductile deformation (Mehta et al. 2006; Ben-Zion et al. 2011). Analytical and numerical results show that the limit value \(\varepsilon = 0\) represents a critical dynamic regime with scale-invariant response functions during failure (Fisher et al., 1997; Dahmen et al., 2009; Ben-Zion, 2012). Stress drops during earthquakes at failing cells increase the
stresses elsewhere along the fault, according to the long range stress transfer function $K(i, j; k, l)$ in Eq.(2), which may trigger more brittle slip instabilities. Failed cells in a composite event have a lower brittle failure strength $\tau_d$ until the end of the event, and are more likely to fail again. A composite event is finished when all cells have local stress below their strength level. Then, a variable evolutionary time step is used to advance the fields forward so that only one cell (hypocenter) reaches the static strength at the beginning of the next brittle event. The evolutionary time step has a maximum value of 3 days to ensure accurate calculations of all time-dependent physical values including stress, slip, temperature, and creep velocity.

The creep process is governed by a power law dependence of creep-velocity $v_c$ on the local stress and coefficients that depend exponentially on the local temperature and activation energy following the Arrhenius function,

$$v_c(x, z, t) = A \exp(-\frac{E(x,z)}{RT(x,z)}) \tau(x, z, t)^3.$$  \hspace{1cm} (5)

Here $A$ is a constant, $E$ is activation energy that has a baseline uniform initial value with random fluctuations along the fault. Grid points are selected randomly with a probability of 0.4 to have activation energy $E \pm r n$, where $r n = 10$ kJ/mol. $T$ is the local temperature, and $R$ is the universal gas constant. The creep parameters are chosen to generate a pine-tree shape strength profile as shown in Fig. 1, with a brittle-ductile transition depth near 10 km. The strength profile is calculated as the minimum between static frictional stress $\tau_s$, and creep stress $\tau_c$ by setting $V_c = V_p$ in Eq. (5). To avoid ongoing events near the boundaries of the model, values of the Arrhenius constant $A$ at fault locations within $x_{BD} = 7.5$ km on both sides of the computational grid are tuned so that the side boundaries mimic creeping behavior of the deep part of the fault.

A key aspect of the model is the mechanical coupling between seismic slip and aseismic slip, which is embedded in the thermal feedback loop connecting stress, slip, and creep velocity. When cells fail during an earthquake, frictional heat generated by the seismic slip increases the local temperature, causing local creep velocity to increase rapidly based on Eq. (5). Cells with accelerated creep velocities then increase the stress at the neighboring cells, which can trigger subsequent brittle failures. Combining the heat generation with a simple 1D cooling in the fault-normal direction, the model captures the basic interaction between seismic slip and aseismic creep by updating at each time step the temperature term in the Arrhenius equation. The fault-normal cooling is done using

$$T(x, z, t) = T_0(z) + (\frac{\tau_d}{\rho w_h} + \Delta T) \cdot \exp[-\frac{a(t_{\text{now}} - t_{\text{eq}})}{w_h l}].$$  \hspace{1cm} (6)

The first term $T_0$ is a background temperature profile with a surface value set to be 278.15 degrees $K$ and a 20 $K/km$ temperature depth gradient. The pre-exponential factor accounts for temperature increase in a slip zone of width $w_h$ with mass density $\rho$ and specific heat $c$, generated by seismic slip $d$ under stress $\tau$, and the residual temperature change $\Delta T$ from previous earthquakes. The exponential
function accounts for cooling by conduction normal to the fault, ignoring heat exchange along the fault (between cells), where $\alpha$ is the thermal diffusivity, $L$ is the distance into the half-space where the temperature field remains constant, $t_{EQ}$ is the time of the seismic slip, and $t_{Now}$ is a subsequent time under consideration. Eq. (6) is calculated using a finite difference method with numerical parameters chosen to meet the Courant-Friedrichs-Lewy condition to ensure numerical stability. The results were benchmarked against the analytical integral solution of Cardwell et al. (1978) for step increase of heat rate in a fault zone of a given width, with temperature decay from the fault zone to a uniform surrounding in a half space.

The model also implements a local co-seismic drop in the activation energy consistent with laboratory results for damaged and intact rocks (e.g., Atkinson, 1984; Chester, 1994; Kirby and Kronenberg, 1987; Tenthorey and Cox, 2006), and a subsequent logarithmic recovery of the activation energy after an earthquake accounting for rock healing. This is implemented using (Zöller et al., 2005),

$$E(x, z, t) = E_d + C \cdot (E_{max} - E_d)\log(1 + \frac{t_{hat} - t_{EQ}}{t_{heal}(x)}),$$

where $E_d$ is the activation energy of damaged rock, $E_{max}$ is the activation energy of intact rock, $C$ is a constant, and $t_{heal}$ is a time interval for healing of the activation energy. After a model earthquake leading to lower activation energy, failed cells creep faster for a period of time based on the prescribed healing time $t_{heal}$, which decreases linearly with depth. Cells participating in model earthquakes may also have repeating brittle failures if the local stress reaches the dynamic friction. By changing the slip zone width $w_h$, cooling distance $L$, Eq. (6), and co-seismic activation energy drop $dE = E_{max} - E_d$ the model can generate different seismic and aseismic slip patterns including aftershock sequences and bimodal distribution of clustered and background seismicity. Table 1 summarizes the geometrical and material parameters used in most simulations.

Table 1: Baseline parameter values used in model simulations.

<table>
<thead>
<tr>
<th>Geometrical and Material Parameters</th>
<th>Symbols</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance along strike</td>
<td>$X$</td>
<td>70 km</td>
</tr>
<tr>
<td>Number of cells along strike</td>
<td>$nL$</td>
<td>256</td>
</tr>
<tr>
<td>Depth</td>
<td>$Z$</td>
<td>17.5 km</td>
</tr>
<tr>
<td>Number of cells along depth</td>
<td>$nD$</td>
<td>64</td>
</tr>
<tr>
<td>Length of numerical cell</td>
<td>$\Delta x$</td>
<td>70 km/nL</td>
</tr>
<tr>
<td>Depth of numerical cell</td>
<td>$\Delta z$</td>
<td>17.5 km/nD</td>
</tr>
<tr>
<td>Plate velocity</td>
<td>$v_p$</td>
<td>35 mm/yr</td>
</tr>
<tr>
<td>Property</td>
<td>Symbol</td>
<td>Value</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>--------</td>
<td>----------------</td>
</tr>
<tr>
<td>Slip velocity</td>
<td>$v_s$</td>
<td>6 km/s</td>
</tr>
<tr>
<td>Rigidity of the half-space</td>
<td>$\mu$</td>
<td>30 GPa</td>
</tr>
<tr>
<td>Strength change parameter</td>
<td>$\varepsilon$</td>
<td>0.8</td>
</tr>
<tr>
<td>Cohesion</td>
<td>$\tau_0$</td>
<td>10 MPa</td>
</tr>
<tr>
<td>Static friction coefficient</td>
<td>$f_s$</td>
<td>[0.7,0.8], uniform distribution</td>
</tr>
<tr>
<td>Static stress gradient</td>
<td>$d\tau$</td>
<td>18 MPa/km</td>
</tr>
<tr>
<td>Earthquake stress drop</td>
<td>$\tau_s - \tau_a$</td>
<td>[2,6] MPa, uniform distribution</td>
</tr>
<tr>
<td>Arrhenius amplitude</td>
<td>$A$</td>
<td>6.31e-17</td>
</tr>
<tr>
<td>Activation energy of intact rock</td>
<td>$E_{max}$</td>
<td>130 kJ/mol</td>
</tr>
<tr>
<td>Activation energy of damaged rock</td>
<td>$E_d$</td>
<td>0.8$E$</td>
</tr>
<tr>
<td>Activation energy fluctuation</td>
<td>$rn(E)$</td>
<td>±12 kJ/mol</td>
</tr>
<tr>
<td>Maximum healing time of activation energy</td>
<td>$t_{Heal max}$</td>
<td>1 year</td>
</tr>
<tr>
<td>Minimum healing time of activation energy</td>
<td>$t_{Heal min}$</td>
<td>1 day</td>
</tr>
<tr>
<td>Specific heat</td>
<td>$c_p$</td>
<td>790 J/kgK</td>
</tr>
<tr>
<td>Diffusivity</td>
<td>$\kappa$</td>
<td>0.01 cm^2/s</td>
</tr>
<tr>
<td>Slip Zone Width</td>
<td>$w_h$</td>
<td>5 cm</td>
</tr>
<tr>
<td>Cooling distance</td>
<td>$L$</td>
<td>100 m</td>
</tr>
<tr>
<td>Surface temperature</td>
<td>$T_0$</td>
<td>278.15 K</td>
</tr>
<tr>
<td>Temperature gradient</td>
<td>$dT$</td>
<td>20 K/km</td>
</tr>
<tr>
<td>Creep boundary, horizontal</td>
<td>$x_{BD}$</td>
<td>7.5 km</td>
</tr>
</tbody>
</table>

### 3. Results

The simulations use a model with a cellular computational grid in an elastic half-space (Fig. 1a), where evolving stress, slip, and temperature fields are generated in response to ongoing loading imposed as velocity boundary conditions on the surrounding fault regions. A representative failure envelope is shown in Fig. 1b as the minimum between $\tau_s(z)$ and $\tau_{\text{creep}}(x, z) = v_{\text{plate}}$, where $\tau_s(z)$ is the frictional failure stress given by Eq. (3) and $\tau_{\text{creep}}$ is creep failure stress given by Eq. (5). Near the edges of the modeled fault section ($x \leq x_{BD} = 7.5 \text{ km}$ & $x \geq 70 \text{ km} - x_{BD}$), the temperature
dependence of creep is turned off by setting the activation energy in Eq. (4) to zero. The Arrhenius amplitude is adjusted so that the creep rate mimics the values at the deeper part of the fault as the edges of the model are approached.

The depth of the brittle-ductile transition controlling the depth distribution of the seismicity depends on the adopted value for activation energy. The presented results focus on the main portion of the computational grid ($x > 7.5 \, km$ & $x < 62.5 \, km$) not including the boundary edges.

The simulations cover several cases where we vary only a subset of parameters affecting the degree of heterogeneity of arrest stress $\tau_a$ during earthquake failures, the strength-change parameter $\varepsilon$, and evolving temperature on the faults (slip zone width $w_r$, cooling distance $L$). All other properties are kept fixed at the values listed in Table 1. The failure envelope used in the simulations corresponds to the intersection between the frictional (blue) and creep (red with random fluctuations between green and black) strength lines in Fig. 1b. We consider primarily two distributions of $\tau_a$ representing different types of heterogeneous fault structures. The first case (Fig. 2a) is associated with a uniform distribution with random fluctuations following $\tau_s - \tau_a = 4MPa \pm 2MPa$. The second case (Fig. 2b) represents a fault with relatively low stress drop segments separated by high stress drop barriers. Here, cells at free surface ($z = 0$) are selected randomly with a probability of 0.2 to initiate one-cell-thick segment barriers, which are extended to depth by 2-D random walks. Stress drop values within the segment barriers are taken from a uniform distribution between [8, 9] MPa. The rest of the fault has stress values between [2, 4] MPa. We consider several cases of strength-change parameter $\varepsilon$ and several values for slip zone width and cooling distances as described below.

![Figure 2. Model realizations with different distributions of arrest stresses leading to different earthquake stress drops. (a) Case 1 with a uniform random distribution of stress drop values between [2, 6] MPa. (b)](image-url)
Assumed stress drops representing a fault with quasi-vertical barriers having larger stress drops than the rest of the fault.

Figures 3a and 3b illustrate several model processes for case 1 with $\varepsilon = 0.8$, slip zone width $w_h = 5 \, \text{cm}$ and cooling distance $L = 100 \, \text{m}$ near middle along-strike positions that participate in relatively large events ($x = 35.4 \, \text{km}$ and $x = 32.7 \, \text{km}$) and at two different depths ($z = 4.5 \, \text{km}$ and $z = 10.3 \, \text{km}$). The dotted vertical lines in all panels indicate times of earthquakes with magnitudes larger than 6. The magnitudes are calculated from the seismic potency values (slip times the rupture area) of the events using the quadratic potency-magnitude scaling relation of Trugman and Ben-Zion (2024), which updates the earlier empirical quadratic relation of Ben-Zion and Zhu (2002) based on considerable more data. The obtained earthquake magnitudes are referred to as potency magnitudes and are denoted by $M_p$. The top panels show shear stress vs. time, with overall linear increasing values due to the tectonic loading, supposed fluctuations generated by earthquakes on the fault, and abrupt drops at times of earthquake ruptures that include the examined cells. The abrupt drops and other changes during model earthquakes are shown better in the zoomed-in plots on the right.

Figure 3(a). Time evolution of physical parameters at a cell located at 35.4 km along strike and 4.5 km deep. Blue vertical lines indicate the time of all the system size ($M_p>6$) earthquakes. The panels from top
to bottom show (i) stress, (ii) activation energy (black) and temperature change (red), (iii) creep velocity plotted with a log scale, and (iv) cumulative seismic slip (black) and cumulative aseismic slip (red).

The second panels from top show the activation energy (left) and temperature change (right) vs. model. Earthquakes spanning a given cell trigger abrupt decrease in activation energy and corresponding increase in temperature. The activation energy drop is followed by a logarithmic healing (Eq. 7). The time interval of healing in Fig. 3a is longer than that in Fig. 3b, as the healing period is prescribed to decrease with normal stress and hence depth. The amount of co-seismic temperature change is determined by shear stress and seismic slip (Eq. 6), so larger temperature changes are generated at deeper parts of the fault and by larger earthquakes. The third set of panels show the creep velocity calculated using (Eq. 5) and the evolving quantities shown in the first two panels. The creep velocity is dominated by the local temperature and activation energy as illustrated clearly in the zoomed-in plots. The bottom panels in Figures 3a and 3b show the cumulative seismic (left) and aseismic (right) slip at the examined cells. Every step increase in the seismic slip is followed by increasing creep slip velocity and aseismic slip. The ratio between cumulative seismic to aseismic slip decreases with depth as a result of the increased ambient temperature.

Figure 3(b). Corresponding results to (a) for a cell located at 32.7 km along strike and 10.3 km deep.
Figures 4a-d present results from a 100-year simulation for case 1 with $\varepsilon = 0.8$, slip zone width $w_s = 5 \text{ cm}$ and cooling distance $L = 100 \text{ m}$. The model hypocenters form clusters and gaps from the surface to roughly 14 km depth (Fig. 4a). Here and in the other results, the hypocenter locations are placed randomly within the cell where they occur. In this model realization, events with potency magnitudes $M_p \geq 6.0$ (red stars) have hypocenters in the depth range 4.5-7 km. During the 100-year simulation, 80 percent of the hypocenters, including those of the events with $M_p > 6$, fall in the top 10 km that define the main seismogenic zone. The seismicity below the nominal brittle-ductile transition boundary, at about $12 \pm 2 \text{ km}$ based on the failure envelope (Fig. 1b), is generated by relatively-deep brittle patches based on the random fluctuations of the activation energy, and to some extent also transient deepening of the brittle-ductile transition due to stress concentration near the bottom of the large events. The magnitude vs. time plot (Fig. 4c) exhibits temporal clustering after most but not all moderate to large events ($M_p \geq 5.5$), which is consistent with observational results (e.g., Ben-Zion, 2008, Fig. 4). Events with hypocenters near model boundaries ($x \leq x_{BD} = 7.5 \text{ km}$ & $x \geq 70 \text{ km} - x_{BD}$) are excluded from the catalog, which accounts for aftershock sequences without a mainshock (e.g. around $t = 3050 \text{ yr}$ in Fig. 4c). The frequency-magnitude distribution of the 100-year catalog (Fig. 4d) follows overall the characteristic earthquake behavior. The results consist of Gutenberg-Richter statistics $\log_{10} n(M) = a - bM$ for the low magnitude events ($M_p < 4$) and enhanced rate of larger events. The deviations from Gutenberg-Richter statistics are also clear in the limited density of earthquakes in the magnitude range 4-6 (Fig. 4c).

Figure 4. Simulation results for case 1 with uniform random distribution of arrest stress and $\varepsilon = 0.8$. (a) Locations of earthquakes (circles) during 100-year simulation with $M_p \geq 6.0$ events denoted by red stars. (b) Depth distribution of seismicity from the 100-year simulation. (c) Potency magnitude of events vs. time. The boxes and stars indicate $M_p \geq 6.0$ events with and without clear aftershock sequences,
The brittle-ductile additional aftershocks then along depth. The background of mainshock 6.32 km of discussed for 7.5 and simulation the (Fig. 1 case hypocenter seismogenic criticality, 2008) close creep (Fig. 1 and frequency-magnitude reduces heating more before, if from (We would the results we Ben-Zion about the events b-value). The (Fig. 1 and frequency-size distribution of low magnitude events in the top 5 km of the fault with minimal creep is characterized by b~1 (Fig. 5). As shown in previous studies (e.g. Fisher et al., 1997; Ben-Zion, 2008) and below, simulations with ε close to zero, representing fault failure processes operating near criticality, produce frequency-size event statistics over the entire seismogenic zone with b value of about 1.

![Figure 5. Frequency-size statistics of earthquakes in (a) the top 5 km of the fault and (b) the deeper seismogenic zone producing the case 1 results over the entire fault shown in Fig. 4d.](image)

Figures 6a-c illustrate results associated with an aftershock sequence of a mainshock with a M_p = 6.32 and hypocenter depth of 7.5 km in the simulation for case 1 discussed above (Fig. 4). The mainshock ruptured the entire seismogenic zone and was followed by an aftershock sequence consisting of roughly 10,000 events over 120 days. The aftershock locations and times are plotted in Fig. 6a on the background of the mainshock slip that exceeded 0.5 m in some places. The early aftershocks started around the hypocenter and expanded quickly to cover first the shallow part of the seismogenic zone and then greater depth. Temperature changes along the rupture zone led to aseismic slip that triggered additional aftershocks and produced an extended sequence. Some of the aftershocks occurred below the nominal brittle-ductile transition because of the rapid stress/strain transfer at the bottom of the
seismogenic zone. The aftershock decay rate follows generally the Omori-Utsu law (Utsu et al., 1995), with fluctuations that reflect in part secondary aftershock sequences as around day 15 (red arrow) when an $Mp = 5.93$ event occurs (Fig. 6b). The initial aftershocks and aseismic slip trigger early on an aftershock with $Mp = 5.0$ followed closely by a larger $Mp = 5.93$ event (Fig. 6c).

Figure 6. An aftershock sequence generated by case 1 with $\varepsilon = 0.8$ that lasts around 130 days. (a) Red to yellow colors show slip values of the mainshock with a hypocenter indicated by the red star. Locations of aftershocks are plotted with green and blue circles with color indicating the time of the aftershocks. (b) Number of aftershocks plotted using a 3-day time window. The fitting curve (red) uses the Omori-Utsu parameters $k = 10,000$, $c = 1$ hr, and $p = 1$. Fitting curves with $c$ values of 1 min and 1 day are given in the supplementary Fig. S1. (c) Potency magnitude vs. time of the entire aftershock sequence with an additional 20 events preceding the mainshock plotted for illustration. The magnitude 6.32 mainshock is highlighted as the red star.

Figures 7a-d present corresponding results to those of Fig. 4 for case 2 with quasi-random vertical barriers, $\varepsilon = 0.8$, slip zone width $w_h = 5$ cm and cooling distance $L = 100$ m. The quasi-random vertical barriers increase somewhat the range of Gutenberg-Richter event statistics (Fig. 7d) but the $b$-value of the simulated low magnitude event statistics is again about 2. Similar results for case 3 of $\tau_a$ distribution on the fault (Fig. S2) including several circular asperities on the fault are presented in the supplementary information (Fig. S3-4). Figures 8a-c illustrate various aspects of a mainshock-aftershock sequence in the simulation for case 2. In this case, the slip variations along the fault reflect the locations of the barriers as
illustrated in Fig. 8a for a mainshock with $M_p = 6.30$. The vertical barriers at roughly $x = 15 \text{ km}$ and $x = 35 \text{ km}$ significantly reduce the slip beyond them and lead to a wider range of moderate to large events, which accounts for the expanded range of Gutenberg-Richter event statistics shown in Fig. 8d. The aftershock decay rates in these and other cases follow overall the Omori-Utsu law (Fig. 8b). The stress drop variations in the vertical barriers produce larger gaps in the magnitudes of aftershocks in Fig. 8c compared to the results for the more uniform case 1 (Fig. 5c). The magnitude difference $\Delta M_p$ between the mainshocks and largest aftershocks in all simulated aftershock sequences for cases 1-3 (Figs. 4, 7, S2) range between 0.4 and 2 with an average value of 1.15, consistent with the empirical Båth law (Båth, 1965).

Figure 7. Results for case 2 with quasi-vertical stress drop barriers and $\varepsilon = 0.8$. (a) Locations of earthquakes (circles) during 100-year simulation with $M_p \geq 6.0$ events denoted by red stars. (b) Depth distribution of seismicity from the 100-year simulation. (c) Potency magnitude ($M_p$) of events vs. time. The boxes and stars indicate $M_p \geq 5.5$ events with and without clear aftershock sequences, respectively. (d) Frequency-magnitude distribution of simulated events in the magnitude range from 2.0 to 6.4. The reference red line has a slope of $b = 1.6$. 
Figure 8. Results for an aftershock sequence generated by case 2. (a) An aftershock sequence of about 130 days on the background of the slip value of the mainshock with a hypocenter marked by the red star. Locations of aftershocks are plotted with a color scale indicating their time. (b) Number of aftershocks in 3-day time windows and the Omori-Utsu fitting curve (red) with $k = 4,000$, $c = 1 \text{ hr}$, and $p = 1$. (c) Potency magnitude vs. time of the entire aftershock sequence with an additional 20 events preceding the mainshock. The magnitude 6.30 mainshock is highlighted as the red star.

In the simulations presented so far with $\varepsilon = 0.8$ (Eq. 4), the rupture areas are overall connected and various earthquake quantities follow standard scaling relations for crack-like ruptures. However, $\varepsilon$ values near 0 on a finite-fault leads to a critical behavior associated with scale-invariant slip patterns, power-law event statistics, and different scaling relations of various earthquake quantities (e.g., Fisher et al., 1997; Ben-Zion, 2012). The critical behavior is illustrated in Fig. 9 with results for case 1 but with $\varepsilon = 0.1$. This model realization produces more events than with $\varepsilon = 0.8$ (Fig. 4), and the frequency-size statistics follow the Gutenberg-Richter distribution with a $b$-value of 1 (other than the finite size bending at the upper magnitude range). In addition, the slip distributions of large events are considerably more heterogeneous (Figs. 10 and S5) and are associated with pulse-like ruptures and different scaling relations than those characterizing cases with relatively large $\varepsilon$ values (Fisher et al., 1997; Dahmen and Ben-Zion, 2009). Simulations for case 2 with quasi-vertical barriers and $\varepsilon = 0.1$ produce results similar to those shown in Fig. 9 with a $b$-value of 1 (Fig. S6).
Figure 9. Simulation results for case 1 with a strength change parameter $\varepsilon = 0.1$. (a) Locations of earthquakes (circles) during 100-year simulation with $M_p \geq 5.7$ events denoted by red stars. (b) Depth distribution of seismicity from the 100-year simulation. (c) Potency magnitude of events vs. time. The boxes and stars indicate $M_p \geq 5.7$ events with and without clear aftershock sequences, respectively. (d) Frequency-size distribution of the simulated events in the magnitude range from 3.0 to 6.4. The reference red line has a slope of $b = 1$. 
Figure 10. Slip distributions of events generated by case 1 uniform model with different strength change parameters. (a) $\varepsilon = 0.8$ and $M_p = 6.03$, (b) $\varepsilon = 0.1$ and $M_p = 5.82$. The stars denote the hypocenter locations.

![Figure 10](image)

Figure 11. Nearest-Neighbor (NN) diagrams based on synthetic catalogs for case 1 (uniform arrest stress distribution) with $\varepsilon = 0.8$ and three sets of parameters controlling the fault temperature and creep velocity. Each of the three realizations has only one changing parameter. The rows from top to bottom represent (i) the baseline case 1 with a cooling distance $L = 100$ m, slip zone width $w_h = 5$ cm, and activation energy of damaged rock $E_d = 0.8E_{\text{max}}$; (ii) a case with slip zone width $w_h = 100$ cm; (iii) a case with activation energy $E_d = 0.9E_{\text{max}}$. The columns from left to right show NN diagrams for (i) 10-year catalog including at least one $M_p \geq 6.0$ mainshock-aftershock sequence; (ii) 10-year catalog with no $M_p \geq 6.0$ mainshock-aftershock sequence; (iii) 50-year catalog containing multiple sequences. The 50-year window is chosen to include multiple earthquake cycles to produce representative NN results for the examined case.

Figure 11 presents nearest-neighbor diagrams of earthquake proximities (Eqs. 1a and 1b) using simulated results of three variations of case 1. Each row has results for one changing parameter related
to the evolving temperature and creep velocity on the fault. At each column from left to right, nearest-neighbor diagrams are generated using synthetic results of 10-year catalog containing one $M_p \geq 6$ mainshock-aftershock sequence, 10-year catalog without a large aftershock sequence, and 50-year catalog with multiple earthquake cycles. The results in the middle column without a large aftershock sequence are shifted to larger rescaled times and distances (weaker clustering) relative to the results in the left column with a large aftershock sequence, and resemble more the background mode of seismicity (e.g., Zaliapin and Ben-Zion, 2013a). Increasing the shear zone width $w_h$ (second row) reduces the temperature changes generated on the fault by earthquake slip, and hence the rate and amount of post-seismic creep. A lower co-seismic temperature change reduces the clustering of seismicity and changes the results in the NN diagram toward a background mode. Decreasing the co-seismic change of activation energy $dE = E_{max} - E_d$ (Eq. 7) makes both the clustered and background modes more prominent (third row). Figs. S7 and S8 present nearest-neighbor diagrams for simulations with case 1 incorporating additional variations of parameters affecting the evolving temperature and creep velocity on the fault.

4. Discussion

We develop an earthquake simulator (EQsim) for a large segmented fault zone with simple effective physics involving static/dynamic friction and temperature-dependent creep that produces diverse observations that can be compared with earthquake catalog data. A key innovation of the EQsim relative to earlier versions of the model (e.g., Ben-Zion and Rice, 1993; Ben-Zion, 1996) is the calculation of frictional heat due to seismic slip on the fault, and coupling the associated temperature changes to values of the coefficients governing creep on the fault (Eqs. 5-7). The coupling of temperature changes and creep coefficients produces transient variations between seismic and aseismic motions following the occurrence of earthquakes on the fault (Fig. 3). The feedback between temperature and creep facilitates the generation of post-seismic stable slip and spatio-temporal clustering of earthquakes, including realistic aftershock sequences and nearest-neighbor diagrams of seismicity resembling observations (Figs. 6, 8, 11, S4, S7, S8). The temperature changes on the fault are controlled, in addition to the local stress level and seismic slip, by assumed width of the slip zone, and cooling distance from the fault where the temperature is assumed fixed. Modifying these parameters, along with Arrhenius activation energy that is reduced during earthquakes and then recovers logarithmically with time produces changes in the aseismic slip on the fault, earthquake clustering, and properties of the nearest-neighbor diagrams (Fig. 11, S7, S8). Changing the difference between the static and kinetic frictions significantly affects the frequency-magnitude event statistics and slip patterns generated by individual earthquakes (e.g., Figs. 4, 9, 10, S5).

The presented results demonstrate that the EQsim can produce by changing a few controlling parameters a wide variety of fault slip phenomena and earthquake statistics that can be compared individually, and more importantly as a correlated set of model outcomes, with observations. The simulated results include relatively smooth and fractal-like slip distributions of earthquakes, temperature changes and transient aseismic slip following large events, and seismicity patterns that may vary between the characteristic and Gutenberg-Richter frequency-size event statistics, aftershock sequences following the Omori-Utsu law, magnitude difference between mainshocks and largest aftershocks
consistent on average with the Bâth law, and nearest-neighbor diagrams with background and clustered modes. To the best of our knowledge, no other earthquake simulator produces such diverse patterns of fault slip and seismicity corresponding overall to observations. The EQsim presented in this paper captures evidently in simple effective ways essential ingredients that govern the dynamics of seismicity on large individual faults.

The results on aftershocks and nearest-neighbor diagrams (Figs. 6, 8, 11) illustrate the importance of aseismic slip for generating clustered earthquakes on a fault. Transient aseismic slip is enhanced in this study by frictional heat produced by seismic slip, but additional factors not included in our model such as clay minerals (e.g., Byerlee, 1978; Moore and Luckner, 2011), fluid effects (e.g., Sibson, 1973; Lockner and Byerlee, 1994), and various factors that change the normal stress on the fault (e.g., Ben-Zion, 2001; Barbot, 2024) can also enhance the generation of transient aseismic slip. The nearest-neighbor diagrams simulated in this study include the background and clustered modes of seismicity, but the shapes of these modes are somewhat different than those associated with observed earthquakes (e.g., Zaliapin and Ben-Zion, 2013a, 2016). Additional model ingredients such as evolution of fluids with slip on the fault (e.g., Miller et al., 1999; Zhu et al., 2020) and the incorporation of fault-network in the simulations (e.g. Dieterich and Richards-Dinger, 2010; Zhao et al., 2024) may produce more realistic nearest neighbor diagrams of earthquakes.

As mentioned in the introduction, numerical simulations are essential for studying seismicity patterns on time scales larger than single large earthquake cycles, and they can be used to clarify key ingredients responsible for generating specific patterns and potential processes leading to large earthquakes (e.g., Ward, 1996; Ben-Zion and Lyakhovsky, 2002; Cattania and Segall, 2021; Im and Avouac, 2023). Earthquake simulators may be used to conduct numerical experiments for conditions not available for laboratory and field investigations (e.g., size of spatial and temporal domains) and simulation results may suggest informative combinations of signals and patterns for data analysis (e.g. Ben-Zion et al., 2003; Pollitz, 2009; Rundle et al., 2011; Zigone et al., 2015; Milner et al., 2022). In addition to addressing basic science problems, numerical simulations can be used to forecast evolving probabilities of earthquakes in different space-time-magnitude domains (Dieterich and Richards-Dinger, 2010; Zöller and Ben-Zion, 2014; Field, 2019; Shaw, 2023), with significant practical relevance for aiding the development of seismic hazard maps.

Continuing work with the EQsim presented in this paper may involve using different realizations of frictional and creep properties to improve the understanding of interactions between brittle and creeping fault sections (e.g., Zigone et al., 2015), the behavior of faults at different evolutionary stages (e.g., Hillers et al., 2007), and informative patterns in relation to large earthquakes (e.g., Eneva and Ben-Zion, 1997a,b; Lu and Vere-Jones, 2001; Enescu et al., 2006). In particular, evolving cluster characteristics (e.g., ratios of single earthquakes to families with many events) in relation to evolving stress heterogeneities may provide important insight on approaching times of large earthquakes. Future simulations with different (e.g. periodic/episodic) transient loadings in addition to the long term tectonic stressing can help clarifying evolving susceptibility to triggering mechanisms at different stages of large earthquake cycles (e.g., Ben-Zion, 2012; Brinkman et al. 2015). Model realizations with loading perturbations representing fluid extraction and injection (e.g., superposed episodic changes of normal stress in portions of the model) can help clarifying clustering features that can be used to discriminate
between tectonic and anthropogenic seismicity (Zaliapin and Ben-Zion, 2016; Hsu et al., 2024). Some such work will be done in follow up studies.

Data Availability
The results of this paper are based on numerical simulations using the developed EQsim. Most of the presented results use the data of simulated earthquake catalogs. The simulation code and example synthetic catalogs will be placed in open depositaries and made available to the public during the process of finalizing the paper.

Acknowledgements
The developed earthquake simulator benefitted from early contributions of Mathias Holschneider and Iain Bailey. The study was supported by the National Science Foundation (grants and OAC-2311206) and the Statewide California Earthquake Center (based on NSF Cooperative Agreement EAR-2225216 and USGS Cooperative Agreement G24AC00072-00).
References


73. Trugman, D., and Y. Ben-Zion (2024), A unified potency magnitude earthquake catalog for the western US, preprint.


Patterns of earthquakes and aseismic slip on a heterogeneous strike-slip fault with static/kinetic friction and temperature-dependent creep

Xiaoyu Zhou¹, Yehuda Ben-Zion¹,²

¹. University of Southern California, Department of Earth Science
². Statewide California Earthquake Center

Key Points

- We develop an earthquake simulator for failure events on a large heterogeneous strike-slip fault in an elastic half-space.

- The simulator includes a feedback mechanism between frictional heat generated by earthquake slip and subsequent creep slip on the fault.

- The feedback facilitates the generation of spatio-temporal clustering of earthquakes on the fault including realistic aftershock sequences.
Abstract

An earthquake simulator is developed to study the dynamics of seismicity and seismic/aseismic slip partitioning on a heterogeneous strike-slip fault using a generalized model of a discrete fault governed by static/dynamic friction and creep in an elastic half-space. Previous versions of the simulator were shown to produce various realistic seismicity patterns (e.g., frequency-magnitude event statistics, hypocenter and slip distributions, temporal occurrence) using friction levels and creep properties that vary in space but are fixed in time. The new simulator incorporates frictional heat generation by earthquake slip leading to temperature rises, subsequent diffusion cooling into the half space, and time-dependent creep on the fault. The model assumes a power law dependence of creep velocity on the local shear stress, with temperature-dependent coefficients based on the Arrhenius equation. Temperature rises due to seismic slip produce increased aseismic slip, which can lead to further stress concentrations, aftershocks, and heat generation in a feedback loop. The partitioning of seismic/aseismic slip and space-time evolution of seismicity are strongly affected by the temperature changes on the fault. The results are also affected significantly by the difference between the static and kinetic friction levels. The model produces realistic spatio-temporal distribution of seismicity, transient aseismic slip patterns, foreshock-mainshock-aftershock sequences, and a bimodal distribution of earthquakes with background and clustered events similar to observations. The simulator (EQsim) may be used to clarify relations between fault properties and different features of seismicity and aseismic slip, and to improve the understanding of failure patterns preceding large earthquakes.

Plain language summary

The physics governing the dynamics of seismicity is currently not fully understood because of the inherent complexity of the problem and the lack of data covering many large earthquake cycles. Large catalogs of seismic and aseismic events generated for different sets of model parameters can aid the understanding of seismicity patterns. Here we develop a simulator of earthquakes and aseismic slip events on a large heterogeneous strike-slip fault in a 3D elastic solid that includes temperature changes on the fault from frictional heat generated by earthquake slip. The simulator generates a wide variety of slip and seismicity patterns including transient aseismic slip following earthquakes, different forms of frequency-size event statistics, and spatio-temporal clustering of earthquakes on the fault including realistic aftershock sequences. The results produced by the simulator may be used to study correlations between different signals and clarify processes leading to large earthquakes.

1. Introduction

Earthquakes are a prime example of a complex natural process with far-from-equilibrium strongly nonlinear dynamics that is not well understood, with substantial societal and economic relevance for large populations worldwide (e.g., Ben-Zion et al., 2022). The lack of quantitative data on timescales capturing multiple large earthquake cycles is a fundamental impediment for progress in the field. Numerical simulations provide the only path for overcoming the lack of data and elucidating spatio-temporal patterns and long-term dynamics that extend the knowledge beyond sporadic case
studies and regional statistical laws. Simulation results that are constrained by a wide variety of observations can allow us to develop possible causal relations between physical variables and responses of faults, clarify processes leading to large events, and enhance the general understanding of earthquake and fault dynamics on various space-time scales. Laboratory results from friction, fracture, and other experiments provide fundamental insights on small-scale failure processes. However, it is unknown how to scale-up microscopic processes on small-scale laboratory samples under the experimental conditions to the complex faulting environment in the crust with strong geometrical and material heterogeneities, non-steady-state loadings over wide ranges of periods and amplitudes, etc. It is therefore essential to develop and use earthquake simulators that are constrained with the available observations of seismicity. The simulators should be constrained with diverse observations to reduce the non-uniqueness inherent to any one type of results (e.g., frequency-size or temporal statistics).

Various types of earthquake simulators have been developed in the last few decades, focusing among other topics on understanding spatio-temporal seismicity patterns, clarifying relations between seismic and aseismic slip on faults, developing scaling relations, and simulating earthquake catalogs to aid probabilistic forecasting of large earthquakes. Example simulators include block-spring models and cellular automata with nearest-neighbor stress interactions (e.g., Burridge and Knopoff, 1967; Carlson and Langer, 1989; Rundle and Klein, 1993), discrete models of heterogeneous large faults with long-range stress interactions and slip patches that can fail independently of their neighbors (e.g., Ben-Zion and Rice, 1993; Robinson and Benites, 1995; Zöller and Ben-Zion, 2014; Zielke et al., 2015), continuum models of faults where patches can only fail as part of a coordinated failure with their neighbors (e.g., Rice and Ben-Zion 1996; Lapusta et al., 2000; Hillers et al., 2006; Erickson et al., 2020; Abdelmeguid and Elbanna, 2022), hybrid discrete-continuum models of fault networks (e.g., Ward, 1996; Pollitz, 2009; Richards-Dinger and Dieterich, 2012; Shaw et al., 2018; Im and Avouac, 2023), and hybrid models accounting for coupled evolution of earthquakes and faults (e.g., Lyakhovsky et al., 1997; 2001; Ben-Zion et al., 1999). For a more detailed overview of earthquake simulators, see Rundle et al. (2003), Ben-Zion (2008), Tullis et al. (2012) and Zielke and Mai (2023).

Ben-Zion and Rice (1993, 1995) and Ben-Zion (1996) showed that discrete fault models, described further in section 2, with a fundamental segmentation length represented by the fault discreteness and long-range stress interactions, provide powerful tools for studying multiple aspects of earthquake patterns including frequency-size and temporal statistics, spatial distribution of events along strike and with depth, and more. However, these models did not account for aftershock sequences and other aspects of clustering that are observed features of seismicity (e.g. Utsu, 2002; Zaliapin and Ben-Zion, 2022). Statistical analyses of regional and global earthquake catalogs demonstrate that earthquakes can be generally described as a superposition of background and triggered events (e.g., Kagan and Jackson, 1991; Ogata, 1999; Gu et al., 2013). This concept is at the core of the stochastic epidemic type aftershock sequence model (ETAS) that provides the standard statistical description of seismicity (Ogata, 1998, 1999).

Following Zaliapin et al. (2008) and Zaliapin and Ben-Zion (2013a,b), a robust data-driven way to separate between background and triggered events is to use a proximity measure between pairs of earthquakes defined (Baiesi and Paczuski, 2004) as
\[ \eta_{ij} = (t_j - t_i)^{10^{-qbm_i}} \cdot (r_{ij})^d \cdot 10^{-(1-q)bm_i}, \quad t_j > t_i \]  

(1a)

where \( t_j - t_i \) is the time between events \( i \) and \( j \), \( r_{ij} \) is the spatial distance between the events, \( m_i \) is the magnitude of the earlier event \( i \), and \( d_f \) is the fractal dimension of the hypocenter (or epicenter) distribution. In data analysis, it is convenient to use rescaled proximities \( R \) and \( T \) defined as

\[ R_{ij} = (r_{ij})^d \cdot 10^{-(1-q)bm_i}, \quad T_{ij} = t_{ij} \cdot 10^{-qbm_i}. \]  

(1b)

Identifying for each earthquake \( j \) a parent event \( i \) that has the closest proximity \( \eta \) among all earlier events in the catalog based on Eq. (1), and plotting the rescaled nearest-neighbor proximities in 2D diagrams of Log \( R \) vs. Log \( T \) leads generally to a bi-modal distribution that facilitates the classification of seismicity to background and clustered (triggered) events (Zaliapin and Ben-Zion, 2013a,b; Ruhl et al., 2016; Martinez-Garzón et al., 2019; Goebel et al., 2019). For observed examples of such bi-modal distributions, see Figure 4 of Zaliapin and Ben-Zion (2013a) and additional observational results in the papers cited above. The main goal of this work is to develop a deterministic earthquake simulator incorporating a few basic mechanical processes that can be used to study observed aspects of seismicity and aseismic slip, including clustering in relation to the assumed fault properties.

The developed simulator (EQsim) generalizes the discrete fault model of Ben-Zion (1996) with quenched (space-dependent) friction and creep properties by adding time-dependency to the creep coefficients. The time-dependency of creep coefficients and aseismic slip on the fault is implemented by calculating frictional heat from each seismic slip, converting the heat to a temperature increase in the slip zone, and coupling the evolving temperature to changes of creep coefficients using the Arrhenius equation. Between seismic slip episodes the temperature at each slip patch is reduced through fault-normal heat conduction to the bulk. The evolution of stress, displacement, and temperature field is calculated in a quasi-static manner. The dynamics of the system is governed by the distributions of the frictional and creep properties, parameters determining the temperature changes due to seismic slip, and the long-range stress transfer due to seismic and aseismic fault motion. The thermal feedback loop between seismic displacements and fault creep properties produces post-seismic slip that leads to clustering of earthquakes on the fault. A parameter-space study is used to develop correlations between fault properties and different aspects of seismicity. For a range of parameters, the simulations produce a bimodal distribution of earthquakes between background and clustered events similar to observations.

In the next section, we describe the EQsim and the relevant physical parameters that affect the results. Section 3 illustrates different types of simulation output, with an emphasis on earthquake clustering in space-time and the bimodal distribution of earthquakes between background and clustered events, which to our knowledge have not been simulated so far by any other physical model. The ability of the model to simulate a wide range of observations provides a path to understand processes and parameters that produce different types of seismicity. The results and possible continuations of the research are discussed in the final section 4.
2. Model

The model consists of a vertical cellular fault zone embedded in a 3-D elastic half-space (Fig. 1). The half-space has a quasi-brittle upper layer that can fail by a combination of frictional and creep processes, over a lower crust and upper mantle sections where constant sliding is imposed at a plate rate $v_{\text{plate}} = 35 \text{ mm/yr}$. The model assumes that the first order slip process in a large fault that is considerably narrower than the along-strike and depth dimensions zone can be mapped onto an ideal planar fault represented by a collection of cellular slip patches on a 2-D computational grid in a half-space (Ben-Zion and Rice 1993, 1995; Ben-Zion, 1996). At each grid position $(x, z; y=0)$, time-dependent frictional and creep deformations are calculated to represent the corresponding deformation of a volumetric fault zone section centered around line $(x + \Delta x/2, z + \Delta z/2)$ with a cross section $(\Delta x, \Delta z)$, where $\Delta x$ and $\Delta z$ denote length and depth of numerical cells along the fault, respectively.

Figure 1. (a) A schematic representation of a heterogeneous strike-slip fault embedded in a 3D elastic half-space. Constant plate motion of 35 mm/yr is applied on both sides and below the computational grid, which has dimensions of 70 km by 17.5 km. Each cell in the grid has creep velocity based on local shear stress, activation energy, and temperature. (b) (Left) Initial Strength profile of the fault. Cells within
7.5 km on both sides are considered boundary cells where the creep velocity gradually increases to plate motion. Thermal effect is also turned off within the boundary region by setting activation energy to zero. (Right) Failure envelope as the minimum between static frictional strength and creep strength. The brittle-ductile transition depth can vary based on the chosen activation energy value.

The stress level at fault location \((i, j)\) and time \(t\), generated by boundary conditions and deforming fault patches, is computed from the slip distribution along the fault using a discretized boundary integral

\[
\tau(i, j, t) = \sum_{k, l} K(i, j; k, l)[V_p t - u(k, l, t)],
\]

where \(K(i, j; k, l)\) is stress transfer for a unit slip over a rectangular patch (Chinnery, 1961). The source dislocation term in the square bracket from Eq. (2) is the slip deficit of cell \((k, l)\) at time \(t\) with respect to the plate motion. The total slip \(u(k, l, t)\) at each cell location \((k, l)\) is a combination of aseismic slip from creep process operating between earthquakes and seismic slip generated by earthquake brittle failures.

Following Ben-Zion and Rice (1993, 1995), the brittle process is governed by a spatial distribution of static frictional strength \(\tau_s\), dynamic strength \(\tau_d\), and arrest stress \(\tau_a\). The model assumes that the \(\tau_s\) increases linearly with depth \(z\) and is given by

\[
\tau_s(z) = \tau_0 + f_s(\rho_s - \rho_w)gz,
\]

where \(\tau_0\) is cohesion, \(f_s\) is the static coefficient of friction assumed to have a uniform value of 0.75 along the fault, \(\rho_s = 2600 \text{ kg/m}^3\) and \(\rho_w = 1000 \text{ kg/m}^3\) are densities of rock and water, respectively, and \(g\) is the gravitational acceleration. Eq. (3) produces an effective normal stress gradient of 18 MPa/km. The stress at any fault location (cell) changes both gradually due to far-field tectonic loading and creep deformations at all fault locations, and abruptly due to brittle failures along the fault. When the stress level \(\tau(x, z)\) at the center of a cell reaches the static strength \(\tau_s\), brittle failure occurs. The strength drops to a dynamic level \(\tau_d < \tau_s\), and the cell slips as to reduce the local stress to an lower arrest level \(\tau_a < \tau_d\) to account for a dynamic overshoot (e.g., Madariaga, 1976).

The static strength, dynamic strength, and arrest stress are related as

\[
\varepsilon = \frac{\tau_s - \tau_a}{\tau_s - \tau_a},
\]

where \(\varepsilon\) is a strength-change parameter. Positive values of \(\varepsilon\) correspond to dynamic weakening during failure appropriate for brittle deformation in the upper crust, while negative values correspond to dynamic strengthening appropriate for regions sustaining ductile deformation (Mehta et al. 2006; Ben-Zion et al. 2011). Analytical and numerical results show that the limit value \(\varepsilon = 0\) represents a critical dynamic regime with scale-invariant response functions during failure (Fisher et al., 1997; Dahmen et al., 2009; Ben-Zion, 2012). Stress drops during earthquakes at failing cells increase the
stresses elsewhere along the fault, according to the long range stress transfer function $K(i, j; k, l)$ in Eq.(2), which may trigger more brittle slip instabilities. Failed cells in a composite event have a lower brittle failure strength $\tau_d$ until the end of the event, and are more likely to fail again. A composite event is finished when all cells have local stress below their strength level. Then, a variable evolutionary time step is used to advance the fields forward so that only one cell (hypocenter) reaches the static strength at the beginning of the next brittle event. The evolutionary time step has a maximum value of 3 days to ensure accurate calculations of all time-dependent physical values including stress, slip, temperature, and creep velocity.

The creep process is governed by a power law dependence of creep-velocity $V_c$ on the local stress and coefficients that depend exponentially on the local temperature and activation energy following the Arrhenius function,

$$v_c(x, z, t) = A \exp(-\frac{E(x, z)}{RT(x, z)}) \tau(x, z, t)^3.$$  \hfill (5)

Here $A$ is a constant, $E$ is activation energy that has a baseline uniform initial value with random fluctuations along the fault. Grid points are selected randomly with a probability of 0.4 to have activation energy $E \pm r_n$, where $r_n = 10 \text{ kJ/mol}$. $T$ is the local temperature, and $R$ is the universal gas constant. The creep parameters are chosen to generate a pine-tree shape strength profile as shown in Fig. 1, with a brittle-ductile transition depth near 10 km. The strength profile is calculated as the minimum between static frictional strength $\tau_s$, and creep strength $\tau_c$ by setting $V_c = V_p$ in Eq. (5). To avoid ongoing events near the boundaries of the model, values of the Arrhenius constant $A$ at fault locations within $x_{BD} = 7.5 \text{ km}$ on both sides of the computational grid are tuned so that the side boundaries mimic creeping behavior of the deep part of the fault.

A key aspect of the model is the mechanical coupling between seismic slip and aseismic slip, which is embedded in the thermal feedback loop connecting stress, slip, and creep velocity. When cells fail during an earthquake, frictional heat generated by the seismic slip increases the local temperature, causing local creep velocity to increase rapidly based on Eq. (5). Cells with accelerated creep velocities then increase the stress at the neighboring cells, which can trigger subsequent brittle failures. Combining the heat generation with a simple 1D cooling in the fault-normal direction, the model captures the basic interaction between seismic slip and aseismic creep by updating at each time step the temperature term in the Arrhenius equation. The fault-normal cooling is done using

$$T(x, z, t) = T_0(z) + (\frac{\tau_d}{\rho c_w^l} + \Delta T) \cdot \exp[-\frac{a(t_{now} - t_{eq})}{w_h^l}].$$  \hfill (6)

The first term $T_0$ is a background temperature profile with a surface value set to be 278.15 degrees $K$ and a 20 $K/km$ temperature depth gradient. The pre-exponential factor accounts for temperature increase in a slip zone of width $w_h$ with mass density $\rho$ and specific heat $c$, generated by seismic slip $d$ under stress $\tau$, and the residual temperature change $\Delta T$ from previous earthquakes. The exponential
function accounts for cooling by conduction normal to the fault, ignoring heat exchange along the fault (between cells), where \( \alpha \) is the thermal diffusivity, \( L \) is the distance into the half-space where the temperature field remains constant, \( t_{EQ} \) is the time of the seismic slip, and \( t_{Now} \) is a subsequent time under consideration. Eq. (6) is calculated using a finite difference method with numerical parameters chosen to meet the Courant-Friedrichs-Lewy condition to ensure numerical stability. The results were benchmarked against the analytical integral solution of Cardwell et al. (1978) for step increase of heat rate in a fault zone of a given width, with temperature decay from the fault zone to a uniform surrounding in a half space.

The model also implements a local co-seismic drop in the activation energy consistent with laboratory results for damaged and intact rocks (e.g., Atkinson, 1984; Chester, 1994; Kirby and Kronenberg, 1987; Tenthorey and Cox, 2006), and a subsequent logarithmic recovery of the activation energy after an earthquake accounting for rock healing. This is implemented using (Zöller et al., 2005),

\[
E(x, z, t) = E_d + C \cdot (E_{max} - E_d) \log \left(1 + \frac{t_{Now} - t_{EQ}}{t_{Heal}}\right),
\]

where \( E_d \) is the activation energy of damaged rock, \( E_{max} \) is the activation energy of intact rock, \( C \) is a constant, and \( t_{Heal} \) is a time interval for healing of the activation energy. After a model earthquake leading to lower activation energy, failed cells creep faster for a period of time based on the prescribed healing time \( t_{Heal} \), which decreases linearly with depth. Cells participating in model earthquakes may also have repeating brittle failures if the local stress reaches the dynamic friction. By changing the slip zone width \( w_h \), cooling distance \( L \) Eq. (6), and co-seismic activation energy drop \( dE = E_{max} - E_d \) the model can generate different seismic and aseismic slip patterns including aftershock sequences and bimodal distribution of clustered and background seismicity. Table 1 summarizes the geometrical and material parameters used in most simulations.

### Table 1: Baseline parameter values used in model simulations.

<table>
<thead>
<tr>
<th>Geometrical and Material Parameters</th>
<th>Symbols</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance along strike</td>
<td>( X )</td>
<td>70 km</td>
</tr>
<tr>
<td>Number of cells along strike</td>
<td>( nL )</td>
<td>256</td>
</tr>
<tr>
<td>Depth</td>
<td>( Z )</td>
<td>17.5 km</td>
</tr>
<tr>
<td>Number of cells along depth</td>
<td>( nD )</td>
<td>64</td>
</tr>
<tr>
<td>Length of numerical cell</td>
<td>( \Delta x )</td>
<td>70 km/( nL )</td>
</tr>
<tr>
<td>Depth of numerical cell</td>
<td>( \Delta z )</td>
<td>17.5 km/( nD )</td>
</tr>
<tr>
<td>Plate velocity</td>
<td>( v_p )</td>
<td>35 mm/yr</td>
</tr>
<tr>
<td>Parameter</td>
<td>Value</td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>---------------------</td>
<td></td>
</tr>
<tr>
<td>Slip velocity</td>
<td>( v_s ) = 6 km/s</td>
<td></td>
</tr>
<tr>
<td>Rigidity of the half-space</td>
<td>( \mu ) = 30 GPa</td>
<td></td>
</tr>
<tr>
<td>Strength change parameter</td>
<td>( \varepsilon ) = 0.8</td>
<td></td>
</tr>
<tr>
<td>Cohesion</td>
<td>( \tau_0 ) = 10 MPa</td>
<td></td>
</tr>
<tr>
<td>Static friction coefficient</td>
<td>( f_s ) = [0.7,0.8], uniform distribution</td>
<td></td>
</tr>
<tr>
<td>Static stress gradient</td>
<td>( d \tau ) = 18 MPa/km</td>
<td></td>
</tr>
<tr>
<td>Earthquake stress drop</td>
<td>( \tau_s - \tau_a ) = [2,6] MPa, uniform distribution</td>
<td></td>
</tr>
<tr>
<td>Arrhenius amplitude</td>
<td>( A ) = 6.31e-17</td>
<td></td>
</tr>
<tr>
<td>Activation energy of intact rock</td>
<td>( E_{max} ) = 130 kJ/mol</td>
<td></td>
</tr>
<tr>
<td>Activation energy of damaged rock</td>
<td>( E_d ) = 0.8E</td>
<td></td>
</tr>
<tr>
<td>Activation energy fluctuation</td>
<td>( r_n(E) ) = ( \pm 12 ) kJ/mol</td>
<td></td>
</tr>
<tr>
<td>Maximum healing time of activation energy</td>
<td>( t_{Healmax} ) = 1 year</td>
<td></td>
</tr>
<tr>
<td>Minimum healing time of activation energy</td>
<td>( t_{Healmin} ) = 1 day</td>
<td></td>
</tr>
<tr>
<td>Specific heat</td>
<td>( c_p ) = 790 J/kgK</td>
<td></td>
</tr>
<tr>
<td>Diffusivity</td>
<td>( \kappa ) = 0.01 cm^2/s</td>
<td></td>
</tr>
<tr>
<td>Slip Zone Width</td>
<td>( w_h ) = 5 cm</td>
<td></td>
</tr>
<tr>
<td>Cooling distance</td>
<td>( L ) = 100 m</td>
<td></td>
</tr>
<tr>
<td>Surface temperature</td>
<td>( T_0 ) = 278.15 K</td>
<td></td>
</tr>
<tr>
<td>Temperature gradient</td>
<td>( dT ) = 20 K/km</td>
<td></td>
</tr>
<tr>
<td>Creep boundary, horizontal</td>
<td>( x_{BD} ) = 7.5 km</td>
<td></td>
</tr>
</tbody>
</table>

3. Results

The simulations use a model with a cellular computational grid in an elastic half-space (Fig. 1a), where evolving stress, slip, and temperature fields are generated in response to ongoing loading imposed as velocity boundary conditions on the surrounding fault regions. A representative failure envelope is shown in Fig. 1b as the minimum between \( \tau_s(z) \) and \( \tau_{\text{creep}}(x,z) = v_c(x,z) = v_{\text{plate}} \), where \( \tau_s(z) \) is the frictional failure stress given by Eq. (3) and \( \tau_{\text{creep}} \) is creep failure stress given by Eq. (5). Near the edges of the modeled fault section (\( x \leq x_{BD} = 7.5 \text{ km} \) & \( x \geq 70 \text{ km} - x_{BD} \)), the temperature
dependence of creep is turned off by setting the activation energy in Eq. (4) to zero. The Arrhenius amplitude is adjusted so that the creep rate mimics the values at the deeper part of the fault as the edges of the model are approached. The depth of the brittle-ductile transition controlling the depth distribution of the seismicity depends on the adopted value for activation energy. The presented results focus on the main portion of the computational grid ($x > 7.5 \ km$ & $x < 62.5 \ km$) not including the boundary edges.

The simulations cover several cases where we vary only a subset of parameters affecting the degree of heterogeneity of arrest stress $\tau_a$ during earthquake failures, the strength-change parameter $\varepsilon$, and evolving temperature on the faults (slip zone width $w_r$, cooling distance $L$). All other properties are kept fixed at the values listed in Table 1. The failure envelope used in the simulations corresponds to the intersection between the frictional (blue) and creep (red with random fluctuations between green and black) strength lines in Fig. 1b. We consider primarily two distributions of $\tau_a$ representing different types of heterogeneous fault structures. The first case (Fig. 2a) is associated with a uniform distribution with random fluctuations following $\tau_s - \tau_a = 4 \ MPa \pm 2 \ MPa$. The second case (Fig. 2b) represents a fault with relatively low stress drop segments separated by high stress drop barriers. Here, cells at free surface ($z = 0$) are selected randomly with a probability of 0.2 to initiate one-cell-thick segment barriers, which are extended to depth by 2-D random walks. Stress drop values within the segment barriers are taken from a uniform distribution between [8, 9] MPa. The rest of the fault has stress values between [2, 4] MPa. We consider several cases of strength-change parameter $\varepsilon$ and several values for slip zone width and cooling distances as described below.

Figure 2. Model realizations with different distributions of arrest stresses leading to different earthquake stress drops. (a) Case 1 with a uniform random distribution of stress drop values between [2, 6] MPa. (b)
Assumed stress drops representing a fault with quasi-vertical barriers having larger stress drops than the rest of the fault.

Figures 3a and 3b illustrate several model processes for case 1 with \( \varepsilon = 0.8 \), slip zone width \( \omega = 5 \text{ cm} \) and cooling distance \( L = 100 \text{ m} \) near middle along-strike positions that participate in relatively large events (\( x = 35.4 \text{ km} \) and \( x = 32.7 \text{ km} \)) and at two different depths (\( z = 4.5 \text{ km} \) and \( z = 10.3 \text{ km} \)). The dotted vertical lines in all panels indicate times of earthquakes with magnitudes larger than 6. The magnitudes are calculated from the seismic potency values (slip times the rupture area) of the events using the quadratic potency-magnitude scaling relation of Trugman and Ben-Zion (2024), which updates the earlier empirical quadratic relation of Ben-Zion and Zhu (2002) based on considerable more data. The obtained earthquake magnitudes are referred to as potency magnitudes and are denoted by \( M_P \). The top panels show shear stress vs. time, with overall linear increasing values due to the tectonic loading, supposed fluctuations generated by earthquakes on the fault, and abrupt drops at times of earthquake ruptures that include the examined cells. The abrupt drops and other changes during model earthquakes are shown better in the zoomed-in plots on the right.

Figure 3(a). Time evolution of physical parameters at a cell located at 35.4 km along strike and 4.5 km deep. Blue vertical lines indicate the time of all the system size (\( M_P > 6 \)) earthquakes. The panels from top
to bottom show (i) stress, (ii) activation energy (black) and temperature change (red), (iii) creep velocity plotted with a log scale, and (iv) cumulative seismic slip (black) and cumulative aseismic slip (red).

The second panels from top show the activation energy (left) and temperature change (right) vs. model. Earthquakes spanning a given cell trigger abrupt decrease in activation energy and corresponding increase in temperature. The activation energy drop is followed by a logarithmic healing (Eq. 7). The time interval of healing in Fig. 3a is longer than that in Fig. 3b, as the healing period is prescribed to decrease with normal stress and hence depth. The amount of co-seismic temperature change is determined by shear stress and seismic slip (Eq. 6), so larger temperature changes are generated at deeper parts of the fault and by larger earthquakes. The third set of panels show the creep velocity calculated using (Eq. 5) and the evolving quantities shown in the first two panels. The creep velocity is dominated by the local temperature and activation energy as illustrated clearly in the zoomed-in plots. The bottom panels in Figures 3a and 3b show the cumulative seismic (left) and aseismic (right) slip at the examined cells. Every step increase in the seismic slip is followed by increasing creep slip velocity and aseismic slip. The ratio between cumulative seismic to aseismic slip decreases with depth as a result of the increased ambient temperature.

Figure 3(b). Corresponding results to (a) for a cell located at 32.7 km along strike and 10.3 km deep.
Figures 4a-d present results from a 100-year simulation for case 1 with $\varepsilon = 0.8$, slip zone width $w_h = 5\,\text{cm}$ and cooling distance $L = 100\,\text{m}$. The model hypocenters form clusters and gaps from the surface to roughly 14 km depth (Fig. 4a). Here and in the other results, the hypocenter locations are placed randomly within the cell where they occur. In this model realization, events with potency magnitudes $Mp \geq 6.0$ (red stars) have hypocenters in the depth range 4.5-7 km. During the 100-year simulation, 80 percent of the hypocenters, including those of the events with $Mp > 6$, fall in the top 10 km that define the main seismogenic zone. The seismicity below the nominal brittle-ductile transition boundary, at about $12 \pm 2\,\text{km}$ based on the failure envelope (Fig. 1b), is generated by relatively-deep brittle patches based on the random fluctuations of the activation energy, and to some extent also transient deepening of the brittle-ductile transition due to stress concentration near the bottom of the large events. The magnitude vs. time plot (Fig. 4c) exhibits temporal clustering after most but not all moderate to large events ($Mp \geq 5.5$), which is consistent with observational results (e.g., Ben-Zion, 2008, Fig. 4). Events with hypocenters near model boundaries ($x \leq x_{BD} = 7.5\,\text{km}$ & $x \geq 70\,\text{km} - x_{BD}$) are excluded from the catalog, which accounts for aftershock sequences without a mainshock (e.g. around $t = 3050\,\text{yr}$ in Fig. 4c). The frequency-magnitude distribution of the 100-year catalog (Fig. 4d) follows overall the characteristic earthquake behavior. The results consist of Gutenberg-Richter statistics $\log_{10} n(M) = a - bM$ for the low magnitude events ($Mp < 4$) and enhanced rate of larger events. The deviations from Gutenberg-Richter statistics are also clear in the limited density of earthquakes in the magnitude range 4-6 (Fig. 4c).

Figure 4. Simulation results for case 1 with uniform random distribution of arrest stress and $\varepsilon = 0.8$. (a) Locations of earthquakes (circles) during 100-year simulation with $Mp \geq 6.0$ events denoted by red stars. (b) Depth distribution of seismicity from the 100-year simulation. (c) Potency magnitude of events vs. time. The boxes and stars indicate $Mp \geq 6.0$ events with and without clear aftershock sequences,
respectively. (d) Frequency-size distribution of simulated events in the magnitude range from 2.0 to 6.4. The reference red line has a slope of 1.6.

The $b$-value for the small events is approximately 1.6, which is larger than the typical $b$-value of about 1 of observed data and previous simulations with earlier versions of the model without creep (e.g. Ben-Zion and Rice, 1993; 1995) and with time-independent creep (e.g., Ben-Zion, 1996; Ben-Zion et al., 2003). (We note that the differences from previous results would be larger if we were converting earthquake potencies to magnitudes with the scaling relation of Ben-Zion and Zhu (2002), as was done before, rather than the updated scaling relation of Trugman and Ben-Zion (2024) based on considerable more data.) The higher $b$-value in the current simulations is related to the enhanced postseismic creep generated by the feedback between earthquake-induced fault heating and creep coefficients, which reduces the seismic slip (model earthquakes) on the fault. This is supported by the observation that the frequency-magnitude distribution of low magnitude events in the top 5 km of the fault with minimal creep is characterized by $b$~1 (Fig. 5). As shown in previous studies (e.g. Fisher et al., 1997; Ben-Zion, 2008) and below, simulations with $\varepsilon$ close to zero, representing fault failure processes operating near criticality, produce frequency-size event statistics over the entire seismogenic zone with $b$ value of about 1.

![Figure 5](image-url)

Figure 5. Frequency-size statistics of earthquakes in (a) the top 5 km of the fault and (b) the deeper seismogenic zone producing the case 1 results over the entire fault shown in Fig. 4d.

Figures 6a-c illustrate results associated with an aftershock sequence of a mainshock with a $M_p = 6.32$ and hypocenter depth of 7.5 km in the simulation for case 1 discussed above (Fig. 4). The mainshock ruptured the entire seismogenic zone and was followed by an aftershock sequence consisting of roughly 10,000 events over 120 days. The aftershock locations and times are plotted in Fig. 6a on the background of the mainshock slip that exceeded 0.5 m in some places. The early aftershocks started around the hypocenter and expanded quickly to cover first the shallow part of the seismogenic zone and then greater depth. Temperature changes along the rupture zone led to aseismic slip that triggered additional aftershocks and produced an extended sequence. Some of the aftershocks occurred below the nominal brittle-ductile transition because of the rapid stress/strain transfer at the bottom of the
seismogenic zone. The aftershock decay rate follows generally the Omori-Utsu law (Utsu et al., 1995), with fluctuations that reflect in part secondary aftershock sequences as around day 15 (red arrow) when an $Mp = 5.93$ event occurs (Fig. 6b). The initial aftershocks and aseismic slip trigger early on an aftershock with $Mp 5.0$ followed closely by a larger $Mp = 5.93$ event (Fig. 6c).

Figure 6. An aftershock sequence generated by case 1 with $\varepsilon = 0.8$ that lasts around 130 days. (a) Red to yellow colors show slip values of the mainshock with a hypocenter indicated by the red star. Locations of aftershocks are plotted with green and blue circles with color indicating the time of the aftershocks. (b) Number of aftershocks plotted using a 3-day time window. The fitting curve (red) uses the Omori-Utsu parameters $k = 10,000$, $c = 1$ hr, and $p = 1$. Fitting curves with $c$ values of 1 min and 1 day are given in the supplementary Fig. S1. (c) Potency magnitude vs. time of the entire aftershock sequence with an additional 20 events preceding the mainshock plotted for illustration. The magnitude 6.32 mainshock is highlighted as the red star.

Figures 7a-d present corresponding results to those of Fig. 4 for case 2 with quasi-random vertical barriers, $\varepsilon = 0.8$, slip zone width $w_h = 5$ cm and cooling distance $L = 100$ m. The quasi-random vertical barriers increase somewhat the range of Gutenberg-Richter event statistics (Fig. 7d) but the $b$-value of the simulated low magnitude event statistics is again about 2. Similar results for case 3 of $\tau_a$ distribution on the fault (Fig. S2) including several circular asperties on the fault are presented in the supplementary information (Fig. S3-4). Figures 8a-c illustrate various aspects of a mainshock-aftershock sequence in the simulation for case 2. In this case, the slip variations along the fault reflect the locations of the barriers as
illustrated in Fig. 8a for a mainshock with $M_p = 6.30$. The vertical barriers at roughly $x = 15 \text{ km}$ and $x = 35 \text{ km}$ significantly reduce the slip beyond them and lead to a wider range of moderate to large events, which accounts for the expanded range of Gutenberg-Richter event statistics shown in Fig. 8d. The aftershock decay rates in these and other cases follow overall the Omori-Utsu law (Fig. 8b). The stress drop variations in the vertical barriers produce larger gaps in the magnitudes of aftershocks in Fig. 8c compared to the results for the more uniform case 1 (Fig. 5c). The magnitude difference $\Delta M_p$ between the mainshocks and largest aftershocks in all simulated aftershock sequences for cases 1-3 (Figs. 4, 7, S2) range between 0.4 and 2 with an average value of 1.15, consistent with the empirical Båth law (Båth, 1965).

Figure 7. Results for case 2 with quasi-vertical stress drop barriers and $\varepsilon = 0.8$. (a) Locations of earthquakes (circles) during 100-year simulation with $M_p \geq 6.0$ events denoted by red stars. (b) Depth distribution of seismicity from the 100-year simulation. (c) Potency magnitude ($M_p$) of events vs. time. The boxes and stars indicate $M_p \geq 5.5$ events with and without clear aftershock sequences, respectively. (d) Frequency-magnitude distribution of simulated events in the magnitude range from 2.0 to 6.4. The reference red line has a slope of $b = 1.6$. 
Figure 8. Results for an aftershock sequence generated by case 2. (a) An aftershock sequence of about 130 days on the background of the slip value of the mainshock with a hypocenter marked by the red star. Locations of aftershocks are plotted with a color scale indicating their time. (b) Number of aftershocks in 3-day time windows and the Omori-Utsu fitting curve (red) with $k = 4,000$, $c = 1\, \text{hr}$, and $p = 1$. (c) Potency magnitude vs. time of the entire aftershock sequence with an additional 20 events preceding the mainshock. The magnitude 6.30 mainshock is highlighted as the red star.

In the simulations presented so far with $\varepsilon = 0.8$ (Eq. 4), the rupture areas are overall connected and various earthquake quantities follow standard scaling relations for crack-like ruptures. However, $\varepsilon$ values near 0 on a finite-fault leads to a critical behavior associated with scale-invariant slip patterns, power-law event statistics, and different scaling relations of various earthquake quantities (e.g., Fisher et al., 1997; Ben-Zion, 2012). The critical behavior is illustrated in Fig. 9 with results for case 1 but with $\varepsilon = 0.1$. This model realization produces more events than with $\varepsilon = 0.8$ (Fig. 4), and the frequency-size statistics follow the Gutenberg-Richter distribution with a $b$-value of 1 (other than the finite size bending at the upper magnitude range). In addition, the slip distributions of large events are considerably more heterogeneous (Figs. 10 and S5) and are associated with pulse-like ruptures and different scaling relations than those characterizing cases with relatively large $\varepsilon$ values (Fisher et al., 1997; Dahmen and Ben-Zion, 2009). Simulations for case 2 with quasi-vertical barriers and $\varepsilon = 0.1$ produce results similar to those shown in Fig. 9 with a $b$-value of 1 (Fig. S6).
Figure 9. Simulation results for case 1 with a strength change parameter $\varepsilon = 0.1$. (a) Locations of earthquakes (circles) during 100-year simulation with $M_p \geq 5.7$ events denoted by red stars. (b) Depth distribution of seismicity from the 100-year simulation. (c) Potency magnitude of events vs. time. The boxes and stars indicate $M_p \geq 5.7$ events with and without clear aftershock sequences, respectively. (d) Frequency-size distribution of the simulated events in the magnitude range from 3.0 to 6.4. The reference red line has a slope of $b = 1$. 

$M_p = 6.03$ 

$M_p = 5.82$
Figure 10. Slip distributions of events generated by case 1 uniform model with different strength change parameters. (a) \( \varepsilon = 0.8 \) and \( M_p = 6.03 \) (b) \( \varepsilon = 0.1 \) and \( M_p = 5.82 \). The stars denote the hypocenter locations.

Figure 11. Nearest-Neighbor (NN) diagrams based on synthetic catalogs for case 1 (uniform arrest stress distribution) with \( \varepsilon = 0.8 \) and three sets of parameters controlling the fault temperature and creep velocity. Each of the three realizations has only one changing parameter. The rows from top to bottom represent (i) the baseline case 1 with a cooling distance \( L = 100 \) \( \text{m} \), slip zone width \( w_h = 5 \) \( \text{cm} \), and activation energy of damaged rock \( E_d = 0.8E_{\text{max}} \); (ii) a case with slip zone width \( w_h = 100 \) \( \text{cm} \); (iii) a case with activation energy \( E_d = 0.9E_{\text{max}} \). The columns from left to right show NN diagrams for (i) 10-year catalog including at least one \( M_p \geq 6.0 \) mainshock-aftershock sequence; (ii) 10-year catalog with no \( M_p \geq 6.0 \) mainshock-aftershock sequence; (iii) 50-year catalog containing multiple sequences. The 50-year window is chosen to include multiple earthquake cycles to produce representative NN results for the examined case.

Figure 11 presents nearest-neighbor diagrams of earthquake proximities (Eqs. 1a and 1b) using simulated results of three variations of case 1. Each row has results for one changing parameter related
to the evolving temperature and creep velocity on the fault. At each column from left to right, nearest-neighbor diagrams are generated using synthetic results of 10-year catalog containing one $M_p \geq 6$ mainshock-aftershock sequence, 10-year catalog without a large aftershock sequence, and 50-year catalog with multiple earthquake cycles. The results in the middle column without a large aftershock sequence are shifted to larger rescaled times and distances (weaker clustering) relative to the results in the left column with a large aftershock sequence, and resemble more the background mode of seismicity (e.g., Zaliapin and Ben-Zion, 2013a). Increasing the shear zone width $w_h$ (second row) reduced the temperature changes generated on the fault by earthquake slip, and hence the rate and amount of post-seismic creep. A lower co-seismic temperature change reduces the clustering of seismicity and changes the results in the NN diagram toward a background mode. Decreasing the co-seismic change of activation energy $dE = E_{\text{max}} - E_d$ (Eq. 7) makes both the clustered and background modes more prominent (third row). Figs. S7 and S8 present nearest-neighbor diagrams for simulations with case 1 incorporating additional variations of parameters affecting the evolving temperature and creep velocity on the fault.

4. Discussion
We develop an earthquake simulator (EQsim) for a large segmented fault zone with simple effective physics involving static/dynamic friction and temperature-dependent creep that produces diverse observations that can be compared with earthquake catalog data. A key innovation of the EQsim relative to earlier versions of the model (e.g., Ben-Zion and Rice, 1993; Ben-Zion, 1996) is the calculation of frictional heat due to seismic slip on the fault, and coupling the associated temperature changes to values of the coefficients governing creep on the fault (Eqs. 5-7). The coupling of temperature changes and creep coefficients produces transient variations between seismic and aseismic motions following the occurrence of earthquakes on the fault (Fig. 3). The feedback between temperature and creep facilitates the generation of post-seismic stable slip and spatio-temporal clustering of earthquakes, including realistic aftershock sequences and nearest-neighbor diagrams of seismicity resembling observations (Figs. 6, 8, 11, S4, S7, S8). The temperature changes on the fault are controlled, in addition to the local stress level and seismic slip, by assumed width of the slip zone, and cooling distance from the fault where the temperature is assumed fixed. Modifying these parameters, along with Arrhenius activation energy that is reduced during earthquakes and then recovers logarithmically with time produces changes in the aseismic slip on the fault, earthquake clustering, and properties of the nearest-neighbor diagrams (Fig. 11, S7, S8). Changing the difference between the static and kinetic frictions significantly affects the frequency-magnitude event statistics and slip patterns generated by individual earthquakes (e.g., Figs. 4, 9, 10, S5).

The presented results demonstrate that the EQsim can produce by changing a few controlling parameters a wide variety of fault slip phenomena and earthquake statistics that can be compared individually, and more importantly as a correlated set of model outcomes, with observations. The simulated results include relatively smooth and fractal-like slip distributions of earthquakes, temperature changes and transient aseismic slip following large events, and seismicity patterns that may vary between the characteristic and Gutenberg-Richter frequency-size event statistics, aftershock sequences following the Omori-Utsu law, magnitude difference between mainshocks and largest aftershocks
The results on aftershocks and nearest-neighbor diagrams (Figs. 6, 8, 11) illustrate the importance of aseismic slip for generating clustered earthquakes on a fault. Transient aseismic slip is enhanced in this study by frictional heat produced by seismic slip, but additional factors not included in our model such as clay minerals (e.g., Byerlee, 1978; Moore and Luckner, 2011), fluid effects (e.g., Sibson, 1973; Lockner and Byerlee, 1994), and various factors that change the normal stress on the fault (e.g., Ben-Zion, 2001; Barbot, 2024) can also enhance the generation of transient aseismic slip. The nearest-neighbor diagrams simulated in this study include the background and clustered modes of seismicity, but the shapes of these modes are somewhat different than those associated with observed earthquakes (e.g., Zaliapin and Ben-Zion, 2013a, 2016). Additional model ingredients such as evolution of fluids with slip on the fault (e.g., Miller et al., 1999; Zhu et al., 2020) and the incorporation of fault-network in the simulations (e.g. Dieterich and Richards-Dinger, 2010; Zhao et al., 2024) may produce more realistic nearest neighbor diagrams of earthquakes.

As mentioned in the introduction, numerical simulations are essential for studying seismicity patterns on time scales larger than single large earthquake cycles, and they can be used to clarify key ingredients responsible for generating specific patterns and potential processes leading to large earthquakes (e.g., Ward, 1996; Ben-Zion and Lyakhovsky, 2002; Cattania and Segall, 2021; Im and Avouac, 2023). Earthquake simulators may be used to conduct numerical experiments for conditions not available for laboratory and field investigations (e.g., size of spatial and temporal domains) and simulation results may suggest informative combinations of signals and patterns for data analysis (e.g., Ben-Zion et al., 2003; Pollitz, 2009; Rundle et al., 2011; Zigone et al., 2015; Milner et al., 2022). In addition to addressing basic science problems, numerical simulations can be used to forecast evolving probabilities of earthquakes in different space-time-magnitude domains (Dieterich and Richards-Dinger, 2010; Zöller and Ben-Zion, 2014; Field, 2019; Shaw, 2023), with significant practical relevance for aiding the development of seismic hazard maps.

Continuing work with the EQsim presented in this paper may involve using different realizations of frictional and creep properties to improve the understanding of interactions between brittle and creeping fault sections (e.g., Zigone et al., 2015), the behavior of faults at different evolutionary stages (e.g., Hillers et al., 2007), and informative patterns in relation to large earthquakes (e.g., Eneva and Ben-Zion, 1997a,b; Lu and Vere-Jones, 2001; Enescu et al., 2006). In particular, evolving cluster characteristics (e.g., ratios of single earthquakes to families with many events) in relation to evolving stress heterogeneities may provide important insight on approaching times of large earthquakes. Future simulations with different (e.g. periodic/episodic) transient loadings in addition to the long term tectonic stressing can help clarifying evolving susceptibility to triggering mechanisms at different stages of large earthquake cycles (e.g., Ben-Zion, 2012; Brinkman et al. 2015). Model realizations with loading perturbations representing fluid extraction and injection (e.g., superposed episodic changes of normal stress in portions of the model) can help clarifying clustering features that can be used to discriminate
between tectonic and anthropogenic seismicity (Zaliapin and Ben-Zion, 2016; Hsu et al., 2024). Some such work will be done in follow up studies.

Data Availability
The results of this paper are based on numerical simulations using the developed EQsim. Most of the presented results use the data of simulated earthquake catalogs. The simulation code and example synthetic catalogs will be placed in open depositaries and made available to the public during the process of finalizing the paper.

Acknowledgements
The developed earthquake simulator benefitted from early contributions of Mathias Holschneider and Iain Bailey. The study was supported by the National Science Foundation (grants and OAC-2311206) and the Statewide California Earthquake Center (based on NSF Cooperative Agreement EAR-2225216 and USGS Cooperative Agreement G24AC00072-00).
References


73. Trugman, D., and Y. Ben-Zion (2024), A unified potency magnitude earthquake catalog for the western US, preprint.


Journal of Geophysical Research
Supporting Information for

Patterns of earthquakes and aseismic slip on a heterogeneous strike-slip fault with static/kinetic friction and temperature-dependent creep

Xiaoyu Zhou¹ and Yehuda Ben-Zion¹²

¹ Department of Earth Sciences, University of Southern California, Los Angeles, CA, USA
² Statewide California Earthquake Center, University of Southern California, Los Angeles, CA, USA

Introduction
The supporting information includes additional figures that further illustrate the simulation results described in the main text for additional parameters.

Contents of this file
Figures S1 to S8

Fig. S1: Aftershock decay rates fitted with different values of the Omori-Utsu parameter c.

Fig. S2: Model realization with distribution of stress drops on the fault including circular asperities.

Fig. S3: Simulation results for the model shown in Fig. S2

Fig. S4: An aftershock sequence for the model shown in Fig. S2.

Fig. S5: Slip distributions for earthquakes generated with different strength-changing parameters.

Fig. S6: Simulation results for case 2 of the main text with a strength-changing parameter of 0.1

Fig. S7: Nearest-neighbor diagrams for different parameters controlling the fault temperature.

Fig. S8: Nearest-neighbor diagrams for additional parameters controlling the fault temperature.
Figure S1: Number of aftershocks plotted using a 3-day time window. The fitting curve (red) uses $k = 10,000$, $p = 1$, and (a) $c = 1$ day or (b) $c = 1$ min.
Figure S2: Case 3 model realization with assumed stress drops representing a fault with circular stress drop asperities of different sizes. Inside the asperity regions, stress drop values follow a uniform distribution between [6, 9] MPa. The rest of the fault has a uniform distribution of stress drop between [2, 4] MPa.
Figure S3: Results for case 3 with circular stress drop asperities and $\varepsilon = 0.8$. (a) Locations of earthquakes (circles) during 100-year simulation with $M_p \geq 6.0$ events denoted by red stars. (b) Depth distribution of seismicity from the 100-year simulation. (c) Potency magnitude ($M_p$) of events vs. time. The boxes and stars indicate $M_p \geq 5.5$ events with and without clear aftershock sequences, respectively. (d) Frequency-magnitude distribution of simulated events in the magnitude range from 2.0 to 6.4. The reference red line has a slope of $b = 1.6$. 
Figure S4: An aftershock sequence generated by case 3 with $\varepsilon = 0.8$ that lasts around 110 days. (a) Red to yellow colors show slip values of the mainshock with a hypocenter indicated by the red star. Locations of aftershocks are plotted with green and blue circles with color indicating the time of the aftershock. (b) Number of aftershocks plotted using a 3-day time window. The fitting curve (red) uses the Omori-Utsu parameters $k = 1, 200$, $c = 1$ hr, and $p = 1$. (c) Potency magnitude vs. time of the entire aftershock sequence with an additional 20 events preceding the mainshock plotted for illustration. The magnitude 6.21 mainshock is highlighted as the red star.
Figure S5: Slip distributions of events generated for case 1 (uniform stress drops) with different strength change parameters. (a) $\varepsilon = 0.8$; (b) $\varepsilon = 0.1$. 
Figure S6. Simulation results for case 2 of arrest stress values with a strength change parameter $\varepsilon = 0.1$. (a) Locations of earthquakes (circles) during 100-year simulation with $Mp \geq 5.7$ events denoted by red stars. (b) Depth distribution of seismicity from the 100-year simulation. (c) Potency magnitude ($Mp$) of events vs. time. The boxes and stars indicate $Mp \geq 5.7$ events with and without clear aftershock sequences, respectively. (d) Frequency magnitude distribution of simulated events in the magnitude range from 2.0 to 6.4. Reference line (red) has a slope of $b = 1$. 
Figure S7: Nearest-Neighbor (NN) diagrams based on synthetic 50-year catalogs for case 1 (uniform arrest stress distribution) with $\varepsilon = 0.8$ and two sets of parameters controlling the fault temperature. In the horizontal direction, NN diagrams show results of models with decreasing cooling distance $L$. In the vertical direction, NN diagrams show NN results of models with increasing slip zone width $w_h$. 
Figure S8: Nearest-Neighbor (NN) diagrams based on synthetic 50-year catalogs for case 1 (uniform arrest stress distribution) with $ε = 0.8$ and two sets of parameters controlling the creep velocity.