Modeling earthquake source processes: from tectonics to dynamic rupture

Executive summary

Modeling earthquake source processes is a vibrant multi-physics, multi-scale, societally important endeavor that tightly links several geoscience disciplines – seismology, geodesy, geology, tectonophysics, hydrology - with numerical computing, data science, machine learning, applied mathematics, continuum mechanics, tribology, materials science, and engineering.

Due to significant recent advances in modeling approaches, observational capabilities, and laboratory experimentation, the field is positioned for rapid future progress towards physics-based, predictive modeling at the societal scales of interest, including scenarios of large destructive earthquakes, prediction of strong ground motion, physics-based estimates of long-term seismic hazard, and potential for induced seismicity. Such progress would be especially timely given considerable natural hazard in several regions of the United States, including California and Pacific Northwest, as well as rapidly accelerating, energy-related industrial activities throughout the world that may cause damaging earthquakes in regions ill-prepared for them.

This report highlights the wide range of relevant scales and processes, articulates the resulting multidisciplinary modeling challenges, emphasizes the need for integrative modeling, and suggests community initiatives that would catalyze future progress. The main cross-cutting themes are the potentially dominating role of fluids - both naturally occurring and added by anthropogenic activities - in the faulting processes in the crust; the effects of inelastic processes and structural complexity of the lithosphere; and the role of shear heating, chemical reactions, and thermomechanical coupling. Of particular importance for future progress is the identified need of rigorously capturing the effects of smaller-scale processes - which may be dominating in earthquake source problems due to extreme multi-scale localization of relevant structures - on the larger-scale phenomena of societal interest, such as destructive earthquakes and induced seismicity.

Future advances would be facilitated by catalyzing community initiatives, such as a systematic community-wide modeling effort to conquer the multi-scale nature of the problem through development of scale-appropriate constitutive laws and integration of earthquake source software solutions; a community effort to harness machine learning and data mining tools to improve observational and modeling inferences; capitalizing on energy-harvesting activities as ongoing field experiments; validation of the developed modeling by a controlled field experiment and well-instrumented laboratory experiments; and contributing physics-based modeling insight into the early warning system that is being implemented on the west coast of the United States. The needed improvements in imaging the earthquake source would rely on intensifying the existing disciplinary and community efforts on novel and dense observational networks, geological studies and fault-zone drilling, and laboratory experimentation.

The overarching modeling challenges relevant to geosciences as a whole include capturing effects of smaller-scale processes on larger-scale phenomena of interest in tractable and scientifically justifiable ways; using multiple types of observables of different precision to constrain multi-scale and multi-physics modeling; and significant software engineering developments needed to benefit from the evolving supercomputing infrastructure. Overall progress would be promoted by funding programs to support open-source code development in collaboration with computer scientists and software engineers; expanding access to supercomputing and storage resources; and educational programs for junior geoscience researchers to learn best modeling and programming practices.
This report is based on the presentations and discussions at the Workshop on “Modeling Earthquake Source Processes: from Tectonics to Dynamic Rupture” held on October 8-10, 2018 in Pasadena, California, and attended, in person and remotely, by ~140 members of the community (Appendix A).

The Workshop was facilitated by the Local Organizing Committee: Michael Gurnis (Chair), Men-Andrin Meier (co-Chair for Remote Participation), Jean-Philippe Avouac, Kim Baker-Gatchalian, Valère Lambert, Nadia Lapusta, Stacy Larochelle, Carolina Oseguera, Kavya Sudhir, and Zachary Ross.

The Workshop was sponsored by the National Science Foundation (NSF) and additionally supported by the Southern California Science Center (SCEC) and the Seismological Laboratory at the California Institute of Technology.

This final report is being submitted to the NSF and other federal agencies.

Preferred citation
Lapusta, N. et al., 2019, Modeling Earthquake Source Processes: from Tectonics to Dynamic Rupture, Report to the National Science Foundation.
Table of Content

I. Multi-physics, multi-scale problem with profound societal implications
   I.1 Integrative modeling of earthquake source processes: Urgency and promise
   I.2 Motivating problems important for societal resilience and sustainability
   I.3 Multiple coupled processes and scales: From tectonics to sub-mm fault cores

II. Multidisciplinary Modeling Challenges (MMCs)

MMC1. Developing constitutive laws for deformation/faulting: capturing small-scale processes
   MMC1.1 Shear resistance of granulated fault cores
   MMC1.2 Off-fault damage/plasticity accumulation and healing
   MMC1.3 Rheology at and below the brittle-ductile transition and fault loading
   MMC1.4 Coupling deformation/damage and fluid effects
   MMC1.5 Formulating scale-appropriate constitutive laws for fault response

MMC2. Building a coherent suite of numerical methodologies for multi-physics problems at larger scales
   MMC2.1 Dynamic rupture simulations: capturing a single earthquake event
   MMC2.2 Modeling sequences of earthquakes and slow slip/deformation
   MMC2.3 Extending simulations to fault networks
   MMC2.4 Modeling tectonics and earthquakes: variability in loading and geometry
   MMC2.5 Frameworks for coupling models at different scales

MMC3. Identifying relevant modeling ingredients by interpreting and improving a range of observations
   MMC3.1 Low-heat, low-stress operation of mature faults but not the rest of the crust
   MMC3.2 Spatio-temporal patterns of seismic/aseismic slip and distributed deformation
   MMC3.3 Properties of earthquake rupture events: rupture speeds, energy budget, and radiation
   MMC3.4 Magnitude-invariant stress drops over the entire range of earthquake magnitudes
   MMC3.5 Gutenberg-Richter law, Omori’s law for aftershocks, and other statistics on regional scales

MMC4. Exploring potential future behaviors: exploiting opportunities and identifying limitations
   MMC4.1 Predicting the source component of strong ground motions
   MMC4.2 Improving our understanding of potential extreme events
   MMC4.3 Capitalizing on fault-specific patterns of seismic and aseismic slip
   MCC4.4 Quantifying uncertainty and identifying key gaps in knowledge

III. Catalyzing community initiatives (IN)
   IN1. Community modeling ecosystem
   IN2. Machine learning and data mining for modeling and observations
   IN3. Modeling energy-harvesting activities as ongoing meso-scale field experiments, link to industry
   IN4. Validation by predictive modeling of a controlled large-scale field experiment
   IN5. Validation by modeling well-instrumented laboratory experiments
   IN6. Physics-based input into early warning algorithms
   IN7. Multidisciplinary summer schools

IV. Support for related community and disciplinary efforts
   IV.1 Observational networks: Seismic, space-based, borehole, fiber-optics, and sea-floor
   IV.2 Laboratory experiments for discovery and developing constitutive laws
   IV.3 Geological studies and drilling as the only direct window into the subsurface
   IV.4 Synergy and complementarity with CIG, CSDMS, SCEC, SZ4D, and other community efforts

Appendix A: Workshop participants
Appendix B: Multidisciplinary modeling challenges (MMCs), detailed descriptions of MMC1-MMC3
Appendix C: References
I. Multi-physics, multi-scale, multi-disciplinary problem with profound societal implications

I.1 Integrative modeling of earthquake source processes: Urgency and promise

Earthquake occurrence is a multi-physics, multi-scale, multidisciplinary problem with both profound societal implications and exciting fundamental science challenges. Integrative modeling of earthquake source processes is critically needed to deliver transformative science that capitalizes on recent progress in seismic, geodetic, geologic, laboratory, and numerical studies. The urgent need for accelerated progress is driven by both considerable natural seismic hazard in the US as well as rapidly increasing man-made seismicity due to energy-related activities. At the same time, there are new enabling opportunities due to data revolution and supercomputing.

We live on a restless planet enveloped by shifting tectonic plates – primary manifestations of Earth’s dynamic interior. These plates deform and interact, leading to great earthquakes that can claim many lives and leave whole cities in rubble (Figure I-1). What are the fundamental principles that control earthquake occurrence? How do we efficiently harness existing field, laboratory, and modeling approaches while creating game-changing capabilities for the future? How can all available data and physical understanding be used to forecast future scenarios of great earthquakes? Answering these questions is key to making society resilient to seismic hazard yet presents a challenging and fascinating scientific problem.

Further, our exploration of geo-resources rapidly expands, motivated by increasing energy demands and technological advances. While most induced earthquakes so far have been minor, the rapid increase in the observed seismicity brings into sharp focus the potential for destructive events in areas away from the plate boundaries that have not planned for seismic shaking. This further contributes to the urgent need to understand the underlying mechanisms that govern the earthquake source and occurrence of large, destructive events. At the same time, energy exploration provides a unique opportunity for promising scientific advances in understanding the deformation and failure of geomaterials, since water-injection wells and hydrothermal operations are, in fact, on-going - and quite expensive - field experiments that can be harnessed to advance and test our knowledge and our models.

Field observations and laboratory studies provide only parts of the puzzle. The modern recorded period is quite short compared to the complexity of the observed earthquake sequences and earthquake source behavior. Paleoseismic inferences are crucial for hazard assessment, but have non-unique interpretations while limited in their time span. Remote observations such as geodesy and seismology face challenges in interpreting the data in terms of the earthquake source behavior due to resolution issues, inadequate coverage, and inherent non-uniqueness of the underlying inverse problems. Laboratory studies can only consider relatively small samples, short durations, and hence experience difficulty in considering realistic fault structures and time scales. Observational and laboratory data can often be qualitatively consistent with several plausible earthquake source scenarios.

This reality motivates a concerted effort to develop predictive earthquake source models that integrate all available knowledge about the earthquake source while highlighting crucial gaps. Such models can eventually move society beyond making decisions based on short-term incomplete earthquake data sets, by incorporating the ingredients that the field – seismic, geodetic, geologic – data do not directly provide,
Figure 1-1. Most large earthquakes occur at tectonic plate boundaries around the world and in the United States. (a) Locations of great earthquakes from 1900 to 1975 (M > 7.9, green circles), 1976-2018 (M > 7.9, green-white mechanisms), and 2004-2018 (M > 8.0, blue mechanisms) around the Pacific and Indian Ocean subduction zones. Adapted from Lay (2015). (b) Past large earthquakes and microseismicity on the plate-boundary San Andreas Fault (SAF) and San Jacinto Fault (SJF) in Southern California. Adapted from Jiang and Lapusta (2016), data from references therein. (c) Tremor locations (red circles) and slow slip distribution (contoured at 1 cm interval) for the 2008 episodic slow slip and tremor episode in Northern Cascadia and its spatial relation to a potential coseismic slip scenario of a great earthquake (greyscale). The most recent Mw ~ 9 Cascadia earthquake occurred on 26 January 1700 as indicated by a tsunami in Japanese historical records (Satake et al., 2003). If the seismogenic portion of the megathrust has been fully locked since that time - and it has been quiescent recently - about 15 m of slip deficit has been accumulated. Adapted from Wang and Trehu (2016), data from references therein.
such as materials-science-based theories on how fault zone materials behave as well as measurable properties of geo-materials from the laboratory. The modeling provides a bridge between the relatively small laboratory scale and the relatively large scale of relevant observations, allowing us to uncover both the physical mechanisms and parameters relevant to natural faults (Figure I-2). Such comprehensive modeling can be accomplished on a range of temporal and spatial scales and can be continuously improved as the new observations and laboratory findings unfold. In fact, such modeling is needed even to properly interpret most laboratory experiments and field observations.

The integrative, predictive modeling of the earthquake source can significantly contribute to societal resilience and sustainability by transforming our understanding of seismic hazard and enabling safer exploration of georesources while enabling overall progress in fundamental geoscience problems (section I.2). Yet, the modeling faces significant challenges (section I.3), amplified by vast ranges of relevant spatial and temporal scales, coupled, non-linear physical and chemical processes, and the remote nature of most observations. Increasingly, we must face the possibility that seismic and interseismic events occurring at different times or on different fault systems are not all expressions of the same governing mechanism or principle, a common assumption used to overcome the short span of observations for any given geographical area, but can rather be dominated by a combination of coupled and evolving mechanisms with different manifestations in different rock types or tectonic environments.

Fortunately, the vast and rapidly increasing array of field observations and laboratory findings on multiple temporal and spatial scales also provides the needed inputs and constraints for the integrative modeling. With renewed systematic efforts towards formulating laboratory-based and theoretically sound constitutive relations (Multidisciplinary Modeling Challenge 1, MMC1 in section II) and creating a suite of modeling tools that use these constitutive relations to collectively capture the breadth of the scales and mechanisms involved (MMC2), we are well-positioned to build fundamentally based models of earthquake source processes that are constrained by the available observations (MMC3) and can be used in a predictive manner (MMC4). Rapid progress in this endeavor can be promoted by several new community initiatives (section III) and existing related activities (section IV).

I.2 Motivating problems important for societal resilience and sustainability

The main goal – and value added – of physics-based earthquake source modeling is to interpret a full range of field observations – seismic, geodetic, thermal etc. – in terms of models that have physically meaningful fault and bulk properties that can be evaluated, at least in principle, through lab, field, and smaller-scale numerical studies and hence can be updated as new lab/field/theoretical knowledge is developed (Figure I-2). The response of the models can then be interrogated for the full range of observationally and lab-constrained ingredients, providing insight into (i) the range of potential future behaviors, (ii) what additional knowledge can further constrain that range, and (iii) in the case of energy production and other human activity, strategies to minimize the induced seismic hazard.

Such integrative modeling can address the following pressing issues needed for societal resilience and sustainability:

- How can the combined field, laboratory, and theoretical knowledge of earthquake source processes improve our understanding and mitigation of seismic hazard?
- What are the plausible scenarios of likely and extreme future earthquake events on major faults in California and Pacific Northwest, as well as in other seismically active regions of the world?
- How do we utilize Earth's energy resources in sustainable ways - through hydrothermal energy, unconventional hydrocarbon recovery, and CO2 sequestration - without inducing destructive
Figure I-2. Multi-disciplinary iterative approach to modeling earthquake source processes. Earthquake source modeling strives to interpret a full range of observations in terms of models that have meaningful fault and bulk properties that can be evaluated, at least in principle, through lab, field, and smaller-scale numerical studies and hence can be updated as new knowledge is developed. This endeavor tightly links several geoscience disciplines – seismology, geodesy, geology, tectonophysics, hydrology - with numerical computing, data science, machine learning, applied mathematics, continuum mechanics, tribology, rock mechanics, materials science, and engineering. Images adapted from Xia et al. (2004), Mizoguchi et al. (2009), Barbot et al. (2012), and the XSEDE Stampede 2 cluster hosted at UT Austin.

- How do earthquakes initiate? Which environments, if any, have identifiable preparation processes or precursors for moderate and large earthquakes?
- When and how do small earthquakes grow into large ones? What is the earthquake energy budget?
- How much stops earthquakes? How extreme can they get?
- How much can we learn from microseismicity and slow deformation/slip between large destructive events?
- What controls the balance of seismic and aseismic slip/deformation?
- Which mechanisms dominate earthquake interaction and clustering?
- What are the feedbacks between longer-term evolving tectonic processes and earthquakes?
- How much of tectonic motion is taken up on known faults vs. unmapped faults or in the bulk?
- What additional knowledge – observational, laboratory, or theoretical – is most needed to constrain seismic hazard?
I.3 Multiple coupled processes and scales: From tectonics to sub-mm fault cores

The source of large destructive earthquakes are rapid shear ruptures, mostly on pre-existing faults in the Earth’s crust, that occur as the result of the overall deformation of the lithosphere driven by slow but inexorable tectonic motion. Many solid materials exhibit pronounced shear localization while deforming under confinement, and the Earth’s crust is no exception, having formed elaborate fault networks. Large earthquakes are especially common on the long-lived faults at plate boundaries (Figure I-1) that have existed for millions of years and presumably hosted hundreds to thousands of large earthquakes; such faults are often called mature. The mature faults are surrounded by other faults that help accommodate the plate motion, distributed depth-dependent zones of damage, and, at greater depth, continuous viscoplastic deformation of the hotter and hence more ductile lower crust and upper mantle. Many, if not most, smaller earthquakes and microseismicity occur within the damage zones and on the secondary faults. The rapid dynamic events co-exist with much slower, aseismic fault slip and deformation, with rates (or velocities) comparable to tectonic plate rates of the order of 1-100 millimeters per year. Recent observations have suggested a broader range of fault slip behaviors than previously recognized, from interseismic slow slip transients to tectonic tremor and low-frequency earthquakes¹. The smaller dynamic events and aseismic processes around and below the seismogenic faults redistribute fault loading locally, modify fault properties, and hence influence the nucleation, timing, and properties of the eventual large events. Furthermore, most of the subsurface is permeated with fluids, which are continuously added by metamorphic reactions at depth, and, increasingly, anthropogenic energy-exploration activities; the fluids are non-uniformly distributed due to vast anisotropic variations in permeability and strongly coupled to both fault and bulk deformation.

To create predictive models of earthquakes based on fundamental science, one needs to account for the multiple coupled scales and processes that govern the lithospheric deformation as a whole. The relevant ingredients (fault structure and shear resistance, bulk structure and rheology, fluid effects), their coupled, evolving, and heterogeneous nature, and their observational manifestations are introduced further in section II on Multidisciplinary Modeling Challenges. Here, we illustrate the scope of the problem by highlighting the multiple coupled spatial and temporal scales involved in modeling earthquake source processes.

The spatial scales of relevant, interactive lithospheric deformation processes span more than twelve orders of magnitude, from thousands of kilometers to microns (Figure I-3):
- 100-1000 km ($10^5$-$10^6$ m): spatial extents of the largest earthquake source ruptures and slow slip events;
- dimensions of the fault segments and systems²;
- 10-100 km ($10^4$-$10^5$ m): the depth below Earth’s surface directly relevant to most destructive seismogenic zones and their loading³; spatial extent of aseismic slip transients and perhaps large-scale earthquake nucleation in some cases⁴;
- 1-1000 m ($10^2$-$10^3$ m): typical source dimensions of microseismicity, e.g. relatively frequent Mw 2.0-3.0 events; the widths of damage zones around active faults with spatially and temporally varying

¹ (Dragert et al., 2001; Obara, 2002; Rogers and Dragert, 2003; Shelly et al., 2006, 2007; Schwartz and Rokosky, 2007; Brown et al., 2009; Outerbridge et al., 2010; Peng and Gomberg, 2010; Beroza and Ide, 2011; Bostock et al., 2015; Gomberg et al., 2016; Obara and Kato, 2016; Araki et al., 2017; McLaskey and Yamashita, 2017; Voss et al., 2017; Hawthorne and Bartlow, 2018)
² (Scholz, 2002; Lay et al., 2005; Gao et al., 2012; Bohnhoff et al., 2016; Perrin et al., 2016; Uphoff et al., 2017)
³ (Bürgmann and Dresen, 2008; Takeuchi and Fialko, 2012)
⁴ (Dodge et al., 1996; McGuire et al., 2005; Beroza and Ide, 2011; Kato et al., 2012)
is clear that the standard rate-segments, earthquake nucleation, interaction of seismic and aseismic slip, and other phenomena\(^1\). Yet it is clear that the standard rate-and-state formulation is an important but only first step in describing the

properties\(^1\); zones of rapid variations in slip rate and stress change at the tip of propagating dynamic ruptures\(^2\);

1-1000 mm \((10^3 - 10^0)\) m): the widths of the co-seismic near-fault layers affected by shear heating and fluid diffusion\(^3\);

0.1-10 mm \((10^{-1} - 10^{-2})\) m): widths of some actively shearing zones at the core of mature active faults - the weakest links that may dominate the resistance of faults to shear - that have been shown to accommodate kilometers of slip\(^4\);

1-10 \(\mu\)m \((10^{-6} - 10^{-5})\) m): the median size of rock particles in the gouge layer inside the fault-core shearing zone - with the particle sizes ranging from nanometers to several microns\(^5\) - the fluid-affected dynamics of which controls the strength of the fault core; the size of frictional contacts on macroscopically flat, microscopically rough experimental interfaces\(^6\).

The relevant temporal scales also span at least fourteen orders of magnitude, from thousands of years for short-term tectonic processes that provide fault loading to sub-second evolution of slip and fault properties at the tip of dynamic, wave-producing ruptures (Figure I-3):

100-10000 years \((10^2 - 10^{11})\) s): short-term tectonic processes; potential scale for significant variations in the seismic cycles; recurrence periods for the largest events\(^7\);

0.01-10 years \((10^{-5} - 10^0)\) s): postseismic effects such as afterslip, visco-elasto-plastic relaxation, and poroelastic effects\(^6\); potential duration of aseismic nucleation processes, duration of some slow slip events (aseismic transients)\(^6\); duration of significant healing processes on and off the fault\(^10\);

10-100 s \((10^{-1} - 10^2)\) s): duration of large seismic ruptures which propagate along faults with the average rupture speed close to the shear wave speed of rocks, about 3 km/s, and involve particle velocities of 1 m/s on average\(^11\);

0.001-1 s \((10^{-3} - 10^{-5})\) s): duration of the rapid variations in slip rate and stress change at the tip of propagating dynamic ruptures\(^12\).

Several physical and chemical ingredients have the potential to significantly affect and even dominate the earthquake source, some of them relevant to multiple spatial and temporal scales. Furthermore, the different ingredients can be highly coupled and varying in both time and space. The identification of rate-and-state frameworks for shear resistance of faults at low, aseismic slip rates has significantly advanced our understanding of the earthquake source, enabling modeling of both stick-slip and creeping fault segments, earthquake nucleation, interaction of seismic and aseismic slip, and other phenomena\(^13\). Yet it is clear that the standard rate-and-state formulation is an important but only first step in describing the

---

1 Schaff and Beroza, 2004; Ben-Zion, 2008; Cochran et al., 2009; Rockwell et al., 2009; Faulkner et al., 2010; Milliner et al. 2015
2 Archuleta, 1984; Day et al., 2005; Lapusta and Liu, 2009; Wollherr et al., 2018
3 Rice, 2006; Noda et al., 2009
4 Chester and Chester, 1998; Wibberley and Shimamoto, 2003
5 Engelder, 1974; Chester et al., 2005
6 Dieterich and Kilgore, 1994; Candela and Brodsky, 2016
7 Reid, 2010; Scholz, 2002; Herrendörfer et al., 2015
8 Bürgmann and Dresen, 2008; Barbot and Fialko, 2010; Wang et al., 2012; Hu et al., 2016; Sobolev and Mудасев, 2017; Faulkner et al., 2018
9 e.g., Dragert et al., 2001; Dieterich, 2007; Peng and Gomberg, 2010; Beroza and Ide, 2011; Bouchon et al., 2013; Brodsky and Lay, 2014; Li and Liu, 2016; Kaneko et al., 2017
10 e.g., Smith and Evans, 1984; Moore et al., 1994; Marone, 1998b; Brenquier et al., 2008; Kelly et al., 2013; Messen et al., 2013; Scuderi et al., 2014; van Dinther et al., 2014; Ingleby and Wright, 2017
11 Duputel et al, 2015; Denolle and Shearer, 2016; Ye et al., 2016, 2018; Meier et al., 2017
12 e.g., Archuleta, 1984; Day et al., 2005
13 Dieterich, 1979ab, 1981a; Rice and Ruina, 1983; Ruina, 1983; Tullis and Weeks, 1986; Blanpied et al., 1991, 1995; Marone, 1998ab; Ben-Zion and Rice, 1997; Rice et al., 2001; Rubin and Ampuero, 2005; Dieterich, 2007; Liu and Rice, 2007; Chen and Lapusta, 2009; Kaneko et al., 2010; Jiang and Fialko, 2016
Figure I-3. A wide range of temporal and spatial scales of earthquake source processes. Spatial scales range from the evolution of asperity contacts on the micron level, to the sub-millimeter scale of highly localized shear layers, to the meter scale of fault roughness and surrounding damage, to expansive fault networks covering 100s to 1000s of kilometers. Relevant temporal scales span the sub-second scale of rapid breakdown at the tip of the propagating earthquake rupture, to several-minute-long durations for large seismic ruptures, to years of postseismic deformation and aftershock sequences, to hundreds to thousands of years in tectonic loading. Adapted from: (left column) Plesch et al., 2007; Mitchell and Faulkner, 2009; Chester and Chester, 1998; Candela and Brodsky, 2016; Dieterich and Kilgore, 1994; (right column) van Dinther et al., 2013; Jolivet et al., 2015; Jiang and Lapusta, 2016; Day et al., 2005; Dunham et al., 2011a, and Great Southern California Shakeout (https://earthquake.usgs.gov/learn/topics/shakingsimulations/shakeout/).

<table>
<thead>
<tr>
<th>Length-scales of fault zone structure</th>
<th>Length and time-scales of earthquake processes</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 - 1000s km</td>
<td>100 - 10,000s years</td>
</tr>
<tr>
<td>1 - 1000 s km</td>
<td></td>
</tr>
<tr>
<td>1 - 100 s m</td>
<td></td>
</tr>
<tr>
<td>1 - 10 mm</td>
<td></td>
</tr>
<tr>
<td>10-100 km</td>
<td>1s - 10 min</td>
</tr>
<tr>
<td>10-100 km</td>
<td></td>
</tr>
<tr>
<td>1-1000 m</td>
<td>1 ms - 1 s</td>
</tr>
</tbody>
</table>

### Spatial Scales
- <50 cm to 100 m: Evolution of asperity contacts on the micron level, to the sub-millimeter scale of highly localized shear layers.
- 1 - 10 mm: Fault roughness and surrounding damage.
- 10 - 100 km: Expansive fault networks covering 100s to 1000s of kilometers.

### Temporal Scales
- Sub-second scale: Rapid breakdown at the tip of the propagating earthquake rupture.
- Several-minute-long: Large seismic ruptures.
- Years: Postseismic deformation and aftershock sequences.
- Hundreds to thousands of years: Tectonic loading.
Inelastic bulk processes during and between seismic events – including off-fault damage and its healing, dynamic shear localization and interseismic delocalization, and visco-elastic-plastic relaxation – affect both the dynamics of fault slip and how we observe it. Faults are non-planar and segmented on a range of scales, with both smaller-scale non-planarity (roughness) and larger-scale complexities (bends, step-overs) significantly affecting slip dynamics. Shear heating during rapid slips can completely modify the constitutive response of the shear zones, in part inducing substantial dynamic weakening. Shear heating is also important during transient postseismic and interseismic deformation, as it can substantially change the visco-elastic-plastic response and hence the fault loading from the continuous deformation below the fault. Fault and bulk properties are likely heterogeneous and evolving with time. Mature, large-displacement faults typically separate dissimilar rocks with different elastic properties, which results in dynamic rupture effects and potential rupture asymmetry.

Pore fluids, ubiquitous in nature, are additionally created on relevant times scales by dehydration reactions below the seismogenic zones and added to the subsurface at an increasing rate by anthropogenic activities (Figure I-4). Their presence, chemistry, and movement can dramatically alter the rheological properties of the bulk rocks, compressive stresses felt across faults, fault loading, and shear resistance of the slipping layers. For example, the newly discovered phenomena of episodic slow slip and tremor have been attributed to natural fluid effects (Figure I-5). It has been long known that increasing pore-fluid pressure promotes fault slip by reducing the effective stress, defined as the difference between the fault-normal stress and pore fluid pressure. It is now clear that such slip can take many forms including earthquakes, steady or transient creep, due to a combination of several physical mechanisms. Earthquakes also directly affect the hydrogeological structure: fault rock damage can create permeability pathways that heal after an earthquake and seismic waves can increase permeability even at great distance. This coupling can in turn affect future earthquakes. Hence a major challenge is to incorporate a fully coupled hydro-seismological modeling over the entire temporal and spatial scales of deformation.

1 (Andrews, 2005; Fialko et al., 2005; Ben-Zion and Shi, 2005; Rosakis et al., 2007; Bürgmann and Dresen, 2008; Bhat et al., 2010; Cochran et al., 2009; Heap et al., 2010; Faulkner et al., 2011; Takeuchi and Fialko, 2012; Gabriel et al., 2013; Shimamoto and Noda, 2014; Xu et al., 2015; Lambert and Barbot, 2016; Erickson et al., 2017; Allison and Dunham, 2018; Roten et al., 2017; Bürgmann, 2018)
2 (Duan and Oglesby, 2005a,b; Renard et al., 2006; Sagy et al., 2007; Candela et al., 2011; Brodsky et al., 2011; Dunham et al., 2011a,b; Fang and Dunham, 2013; Shi and Day, 2013; Bruhat et al., 2016; Harbord et al., 2017; Romanet et al., 2018; Tal and Hager, 2018; Ando and Kaneko, 2018; Ulrich et al., 2018)
3 (e.g., Sibson, 1973; Lachenbruch, 1980; Tsutsumi and Shimamoto, 1997; Andrews, 2002; Rice, 2006; Kitajima et al., 2010; Di Toro et al., 2011; Brantut et al., 2010; Goldsby and Tullis, 2011; Fulton et al., 2013; Noda and Lapusta, 2013; Tullis, 2015; Viesca and Garagash, 2015; Rubino et al., 2017)
4 (Hori et al., 2004; Tanikawa and Shimamoto, 2009; Ma et al., 2010; Meng et al., 2011; Simons et al., 2011; Lay et al., 2012; Yao et al., 2013; Ye et al., 2016a,b)
5 (Ben-Zion and Malin, 1991; Andrews and Ben-Zion, 1997; Harris and Day, 1997; Cochar and Rice, 2000; Ben-Zion, 2001; McGuire and Ben-Zion, 2005; Le Pichon et al., 2005; Rubin and Ampuero, 2007; Ampuero and Ben-Zion, 2008; Brietzke et al., 2009; Bulut et al., 2012; Allam et al., 2014; Yang et al., 2015; Shlomai and Finberg, 2016; Share and Ben-Zion, 2018)
6 (Fialko, 2004; Hauksson and Shearer, 2006)
7 (e.g., Hauksson and Shearer, 2006; Liu and Rice, 2007; Faccenda et al., 2012; Hyndman and Peacock, 2003)
8 (Elsworth, 2013; Keranen et al., 2014; Weingarten et al., 2015; Elsworth et al., 2016)
9 (Hubert and Rubey, 1959; Sibson, 1973; Lachenbruch, 1980; Segall and Rice, 1995; Caine et al., 1996; Rice, 2006; Dunham and Rice, 2006; Ge et al., 2009; Noda et al., 2009; Faulkner et al., 2010; Segall et al., 2010; Kitajima and Saffer, 2012; Sutherland et al., 2012; Bense et al., 2013; Townend et al., 2013; Guglielmi et al., 2015; Hirh and Beeler, 2015; Proctor and Hirth, 2015; Segall and Lu, 2015; Scuderi and Collettini, 2016; Beeler et al., 2016; Boulton et al., 2017; Townend et al., 2017)
10 (Scholz, 1998; Dragert et al., 2001; Liu and Rice, 2005; Guglielmi et al., 2015; Wei et al., 2015)
11 (Elkhoury et al., 2006; Manga et al., 2012; Xue et al., 2013; Weingarten and Ge, 2014)
12 (Sibson, 1992; Fulton and Brodsky, 2016; Xue et al., 2016)
Figure I-4. Seismicity potentially induced by anthropogenic fluid injections. (a) Induced seismicity can occur in conjunction with several industrial activities. Adapted from Grigoli et al. (2017). (b) M > 3.0 earthquakes (grey histogram) spatiotemporally associated (red histogram) and not associated (black line) with injection wells in the U.S. mid-continent from 1973 to 2014. Over the time period of the catalog, the number of nonassociated earthquakes per year has stayed approximately constant at 10 to 25 per year. Meanwhile, the number of associated earthquakes per year has risen from ~1 to 7 per year in the 1970s to 75 to 190 per year between 2011 and 2013 and >650 earthquakes in 2014. The U.S. mid-continent is defined to be between the dashed lines on the insert, following Ellsworth (2013). Red and white dots denote M ≥ 0.0 earthquakes that are and are not spatiotemporally associated with injection wells, respectively, in 1973-2014. Adapted from Weingarten et al. (2015). (c) A geothermal operation in California induced aseismic slip (colors) on a normal fault, potentially leading to a Mw 5.4 strike-slip event (contour lines), part of the the 2012 Brawley swarm. Adapted from Wei et al. (2015).
Figure I-5. Elevated pore fluid pressure on natural faults and transient slow slip. (a) (Top) The presence of excess fluids - and hence elevated pore fluid pressure - on the southwest Japan megathrust is indicated by elevated values of the seismic velocity ($vP/vS$) ratio (blue); the relocated hypocenters of regular earthquakes (black) and low-frequency earthquakes (LFEs, red) also shown. (Bottom) Schematic diagram of the interpretation in which the transient slow slip and LFEs occur on the plate interface, coincident with the zone of high pore fluid pressure, while the regular seismicity occurs primarily within the oceanic lower crust. Adapted from Shelly et al. (2006). (b) The presence of the excess fluids - and the associated transient slow slip events - may be due to dehydration reactions that are predicted to deliver fluids to the bottom of the seismogenic zone on some megathrusts but not others. (Left) Petrological phase equilibrium of the ocean crust. (Right) Calculated pressure-temperature paths and metamorphic conditions encountered by the oceanic crust subducted beneath Cascadia, southern Mexico, and southwest and northeast Japan. Shallow-dipping subduction zones, Cascadia (yellow), southern Mexico (green), and southwest Japan (orange), undergo a sequence of dehydration reactions near and downdip from the stability transition at $350^\circ$. Fluids released from such dehydration reactions can significantly increase the pore pressure. This elevated pore pressure is consistent with the findings of Shelly et al. (2006) (panels a) for the southwest Japan subduction zone. Adapted from Liu and Rice (2007).
Clearly, explicit inclusion of all scales and relevant processes in a single “master” model is not feasible at present or in the foreseeable future. Yet the most disparate scales and seemingly unrelated mechanisms can affect each other in profound ways. For example, slow slip in the interseismic period, including aseismic transients, redistributes stress conditions on the fault and potentially affects its structure and strength, which in turn could significantly affect the nucleation location and rupture propagation of the upcoming large events. In return, processes during dynamic rupture - such as enhanced dynamic weakening supported by ample laboratory evidence and some observational inferences - may control the level of stress on mature faults and stress changes transferred to the surrounding medium, modulating postseismic effects and interseismic fault behavior, including some microseismicity patterns. In another example, as the dynamic rupture propagates, the largest-scale energy transfer through dynamic waves, the tectonic and local fault loading conditions, and the local resistance of the actively shearing zone which can be sub-millimeter, all combine to govern its slip rate and rupture speed. Then, if the highly elevated stresses and strains at the rupture tip sufficiently damage the surrounding medium, changes are induced in both energy transfer (due to different elastic properties of the damaged medium), energy dissipation (due to damage creation), and shear resistance (e.g., through local changes in permeability and hence fluid pressure). Furthermore, the damage can be modulated and amplified by fault roughness and other geometric complexities of the fault, essentially resulting in dynamically induced heterogeneity. The overall interaction can then significantly affect the large-scale dynamics of the rupture - unless damage creation is relatively unimportant in some cases, e.g., at the seismogenic depths of mature faults.

The modeling efforts so far have examined the interaction of a subset of relevant processes over several of the relevant spatial and temporal scales, with the goal of determining the potential links, couplings, and relation of outcomes to observations. The overall effort has been quite productive and informative, resulting in a number of discoveries and collectively establishing the potential importance of various scales and processes (section II).

Given the successful and intensifying modeling, field, and laboratory efforts, the community is well-positioned for rapid future progress towards fundamentally based, predictive models at the societal scales of interest, including likely scenarios of large destructive earthquakes, prediction of strong ground motion, physics-based estimates of long-term seismic hazard, and potential for induced seismicity. Such progress will come from systematic modeling studies that combine all knowledge about relevant fault/bulk structure and mechanisms and use all available observations from the field and laboratory to determine the coupled effects of the relevant mechanisms, constrain parameter spaces, and develop abstractions suitable for modeling at the largest scales, as discussed in section II on the Multidisciplinary Modeling Challenges (MMCs). The progress will also require an increasing integration of modeling with laboratory and observational studies, which can be achieved through several new and existing community initiatives (sections III and IV).
II. Multidisciplinary modeling challenges (MMCs)

To study earthquake source processes, models combine the balance laws of continuum mechanics with representations of (a) the fault structure/geometry and its constitutive response (e.g., shear resistance), (b) the surrounding rock bulk and its constitutive response to deformation (i.e., rheology), and (c) tectonic loading and other inputs (e.g., fluid injection to study anthropogenic energy-related activities).

The balance laws used are typically classical and well-known, including the principles of mass and energy conservation and the principles of linear and angular momentum. The principle of linear momentum results in equations of motion in terms of the Cauchy stress tensor. The kinematic variables that describe deformation are also standard and most commonly include the infinitesimal strain tensor and relative displacements across interfaces often used to represent fault cores, with relative shear motion called slip and relative fault-normal motion called opening. Formulations involving finite strains and infinitesimal/finite rotations are also used.

Hence the main challenges in building realistic, predictive earthquake-source models are:

1. appropriately capturing the multiple coupled mechanisms involved in the fault and bulk response to loading over the multiple relevant spatial and temporal scales via theoretically sound, laboratory-based constitutive relations that connect the stress, deformation, pore fluid pressure, and other relevant variables to each other
   (MMC1, Combining field, lab, and modeling findings into scale-appropriate constitutive relations for deformation and faulting);

2. developing a coherent set of computational approaches that can resolve and interrogate the rich response of the resulting earthquake-source models in terms of earthquakes, slow slip/deformation, fluid effects, and other phenomena, again over multiple relevant temporal and spatial scales
   (MMC2, Building a coherent suite of numerical methodologies for multi-physics problems at larger scales);

3. determining the dominating mechanisms and relevant parameter regimes by comparing the model outcomes with the wide range of the field - seismic, paleoseismic, geodetic, geologic - and lab observations about the earthquake source
   (MMC3, Identifying relevant modeling ingredients by interpreting and improving a range of observations); and

4. using the resulting models predictively; how to do it appropriately, given the simplifications and uncertainty involved, is a challenging research problem in itself
   (MMC4, Forecasting potential future behaviors: exploiting opportunities and identifying limitations).

These challenges can only be met through multidisciplinary efforts that combine theoretical modeling, lab studies, and field observations.
MMC1. Developing constitutive laws for deformation/faulting: capturing small-scale processes

The geological, seismological, geodetic, laboratory, and modeling studies and their combinations have collectively significantly advanced our understanding of the coupled and evolving fault/bulk structure and their shear resistance and rheology. Much progress has come in the last two decades, due to insightful geological investigations of exhumed faults and active faults through trenching and drilling; advanced and highly instrumented laboratory investigations; dense observational networks and space-based monitoring; novel data processing techniques; and studies that combine the field/lab observations and theoretical modeling.

Capturing the uncovered complex multi-physics and multi-scale response of faults and the surrounding rocks to loading - with or without fluids, under the broad range of stress, temperature, and strain-rate regimes of interest - is the biggest challenge for realistic, predictive modeling of rock deformation, faulting, large earthquakes, and induced seismicity. To meet this challenge, the important next step is to combine the accumulating knowledge into theoretically sound, scale-appropriate constitutive relations with physically meaningful properties that can be extended as the new laboratory, theoretical, or field knowledge is developed. This is the focus of this section, with the associated modeling advances and challenges discussed in MMC2.

MMC1.1 Shear resistance of granular fault cores

The elaborate structure of seismogenic faults (Figure II-MMC1) includes highly localized, long-lived shear layers that accommodate most of the across-fault shear motion¹ (Figure II-MMC1.1). The resistance of such layers to shear - often called “fault friction” - is then a major, and potentially dominating, component of the overall earthquake source puzzle, highlighting the crucial importance of the sub-millimeter scale for the earthquake source problem. The shear resistance of such granular layers - and frictional surfaces - has indeed been actively studied in the laboratory and through modeling. Some of the earliest results highlighted the dependence of friction on slip velocity (also called slip rate), through both a direct effect and evolutionary effect with characteristic slip captured through a state variable, resulting in widely used formulations called “rate-and-state friction”². These findings enabled modeling a wide range of the observed earthquake source processes, from creeping and stick-slip fault regions due to steady-state velocity-strengthening and velocity-weakening friction, respectively, to aseismic earthquake nucleation on velocity-weakening faults due to the combination of the direct and evolutionary effects, to postseismic slip and aftershock-sequence decay (see MMC2.2).

Subsequent experimental studies and their analyses refined the findings and elucidated the dependence of shear resistance in granular layers on multiple evolving factors including slip rate, compressive normal stress, pore fluid pressure, mineral composition, dilatation/compaction, temperature, shear heating, and a range of pore fluid effects, including fluid generation due thermal decomposition³. One of the key findings

¹ (Chester and Logan, 1987; Chester et al., 1993; Wibberley and Shimamoto, 2003; Sutherland et al., 2012)
² (Dieterich, 1979ab, 1981a; Rice and Ruina, 1983; Ruina, 1983; Tullis and Weeks, 1986; Scholz, 1990; Blanpied et al., 1991, 1995; Beeler et al., 1994; Marone, 1998ab; Scholz, 1998)
³ (Marone et al., 1990; Spiers et al., 1990; Beeler et al., 1994, 2008; Chester, 1994; Dieterich and Kilgore, 1994; Lockner and Byerlee, 1994; Segall and Rice, 1995; Tsutsumi and Shimamoto, 1997; Marone 1998a,b; Bos et al., 2000; Bos and Spiers, 2001; Goldsby and Tullis, 2002, 2011; Di Toro et al., 2004, 2006, 2011; Han et al., 2007; Niemeijer et al., 2008, 2010b, 2012; Zhang and Spiers, 2005; Rice 2006; Schleicher et al., 2006, 2010; Collettini et al., 2009, 2011; Ikari et al., 2009; Samuelson et al., 2009; Brantut et al., 2010; Faulkner et al., 2011, 2018; Brown and Fialko, 2012; Kirkpatrick et al., 2013; Warr et al., 2014; Proctor et al., 2014; Carpenter et al., 2015, 2016;
Figure II-MCC1. Elaborate structure of seismogenic faults and the surrounding inelastic bulk. At the seismogenic depths, the depth-dependent fault structure includes localized, long-lived shear layers that accommodate most of the across-fault shear motion surrounded by damage zones of different types. The damage zones reach widths up to several kilometers in the top 3-5 km and can be much narrower through the seismogenic depths due to higher confinement. At the bottom of and/or below the seismogenic zone, the damage is suppressed but a broader ductile shear zone develops, potentially with localized strands of either remnants of the frictional shear layers or localized ductile shear zones. Some faults exhibit more complex structure at seismogenic depths, with multiple localized shear zones embedded inside the damage zone (right insert, adapted from Mitchell and Faulkner, 2009). The subduction zones have additional structural complexity, especially in the shallow portion; e.g., the Nankai accretionary wedge includes numerous splay faults branching off the principal plate interface (left insert, adapted from Strasser et al., 2009). Top panel by Dan Faulkner, with data from 1 Sibson, 1986; 2 Woodcock and Mort, 2008; 3 Dor et al., 2006; 4 Doan and Gary, 2009; 5 Mitchell et al., 2011; 6 Mitchell and Faulkner, 2009; 7 Anders and Wiltshire, 1990; 8 Vermilye and Scholz 1998; 9 Wilson et al., 2003; 10 Faulkner et al, 2001; 11 Cochran et al., 2009; 12 Kelly et al., 2013; 13 Wechsler et al., 2009; 14 Li et al., 2009; 15 Savage and Brodsky, 2011; 16 Heap and Faulkner, 2008; 17 Faulkner et al., 2006; Brenguer et al., 2008; 19 Polak et al., 2003; 20 Morrow et al., 2001; 21 Vaughn et al., 1986; 22 Smith and Evans 1984; 23 Brantley et al., 1990.

Scuderi et al., 2014; Bhattacharya et al., 2015; Platt et al., 2015; French et al., 2016; Niemeijer et al., 2016; Wojatschke et al., 2016; Lockner et al., 2017)
Figure II-MMC1.1. Resistance of localized shear layers: rate-and-state friction, with extensions, at low slip rates, enhanced dynamic weakening at high slip rates. (a) Kilometers of slip (i.e., relative shear motion) have been accommodated along a sub-millimeter shear layer (dashed red lines) on the exhumed Punchbowl fault in California. Adapted from Chester and Chester (1998). (b) Laboratory experiments that study the evolution of the shear resistance of such granular layers at low sliding rates show that a positive slip-rate jump is always accompanied by a positive, stabilizing direct effect of increasing shear resistance, followed by its evolution with characteristic slip. These features have been combined into widely used rate-and-state friction laws which are being extended to additional processes (MMC1.1; Appendix B; Figure V-MMC1.1). Adapted from Marone (1998). (c) Many laboratory studies indicate that the typical friction coefficients for sliding at low slip rates, relevant to plate motion and earthquake nucleation, are 0.6-0.8 for many (but not all) rock materials, consistent with Byerlee’s friction, but dramatic enhanced weakening occurs at high slip rates relevant for earthquake rupture. Adapted from Di Toro et al. (2011).
is that rapid shear, at the high slip rates expected co-seismically, can dramatically weaken the shearing layer, due to a number of mechanisms related to shear heating, fluid effects, and chemical alterations, making the response of the localized shear layer a potentially dominating factor in rupture dynamics and providing a promising explanation for the observed low-stress, low-heat operation of major, mature faults (section MMC 3.1). Another key finding is the important - and competing - effects of fluids, if present, in the shear resistance of the granular layers, which can be both stabilizing, e.g., through pore pressure decreases due to shear-induced inelastic dilatancy (so-called dilatant hardening) and destabilizing, e.g., through pore pressure increases due to shear heating. Such pore pressure effects have been invoked to explain large seismic slips and transient aseismic fault slip (MMC2.2, MMC3.2, MMC3.3).

Systematic untangling of all the discovered dependencies into a coherent, theoretically sound framework is well underway, based on rate-and-state friction formulations and their extensions, and will require continuing well-designed laboratory experimentation on realistic fault materials under realistic conditions and modeling at the scale of the shear layers1.

**MMC1.2 Off-fault damage/plasticity accumulation and healing**

The shear layers within the seismogenic depths are surrounded, in the fault-normal direction, by so-called damage zones of pervasive cracking, partially inherited from the fault formation and partially caused or at least maintained by fault slip and associated off-fault stressing; the damage zones are the result of inelastic deformation, with energy dissipation and evolution in effective elastic moduli and fluid transport properties. The width and properties of the damage zones are highly variable with depth, due to depth-increasing average confinement and temperature, and along strike, due to non-planar fault geometry (Figure II-MMC1). The renewal of damage during earthquake ruptures and healing in the interseismic period, as well as the associated fluid effects, can significantly affect the earthquake energy budget, slip dynamics, and fault strength evolution. Systematic investigation of the damage zones and their effects is an important frontier in earthquake source modeling (MMC2).

Laboratory experiments and theoretical modeling have significantly advanced our understanding of the evolution of small-scale damage (Figure II-MMC1.2). Studies under low, quasi-static loading rates, determined its interplay with elastic moduli reduction and failure process. Experiments under larger loading rates identified the change in damage patterns and increase in rock failure strength with strain rate. Other studies have focused on the healing processes, showing that over time, and with elevated temperature and fluids, microcracks heal by diffusional, dissolution-precipitation, and mechanical processes. The constitutive response of damaged rock volumes have been represented mainly by plastic

---

1 (Rudnicki and Rice, 1975; Daub and Carlson, 2008; Rice et al., 2014; Platt et al., 2014; 2015; Kothari and Elbanna, 2017; Ma and Elbanna, 2018)
2 (Li et al., 1997, 2006, 2014; Ben-Zion et al., 2003; Fialko et al., 2002; Cochran et al., 2009; Sammis et al., 2009; Sutherland et al., 2012; Qiu et al., 2017; Townend et al., 2017)
3 (Anders and Wittschoke, 1994; Cowie and Shipton, 1998; Vermilye and Scholz, 1998; Chester and Chester, 2000; Waldhauser and Ellsworth, 2002; Ben-Zion and Sammis, 2003; Shipton and Cowie, 2003; Wilson et al., 2003; Mitchell and Faulkner, 2009; Jolivet et al., 2009; Wechsler et al., 2009; Faulkner et al., 2011; Savage and Brodsky, 2011; Cochran et al., 2009; Allam and Ben-Zion, 2012; Kelly et al., 2013; Rowe et al., 2013; Lindsey et al., 2014)
4 (Cowie and Scholz, 1992; Harris and Day, 1997; Lyakhovsky et al., 1997; Ben-Zion and Andrews, 1998; Vidale and Li, 2003; Hamiel et al., 2004; Andrews, 2005; Rice et al., 2005; Peng and Ben-Zion, 2006; Brenguier et al., 2008; Dunham and Rice, 2008; Ben-Zion and Ampuero, 2009; Sammis et al., 2009; Bhat et al., 2012; Ngo et al., 2012; Gabriel et al., 2013; Kelly et al., 2013; Huang et al., 2016)
5 (Brace et al., 1966; Lockner et al., 1977; Hadley, 1976; Lockner et al., 1992; Katz and Reches, 2002; Oda et al., 2002; Takemura and Oda, 2004; Schubnel et al., 2006; Heap et al., 2010; Yang et al., 2015; Tal et al., 2016; Renard et al., 2018a,b)
6 (Grady, 1998; Doan and Gary, 2009; Zhang and Zhao, 2014; Aben et al., 2017)
7 (Smith and Evans, 1984; Hickman and Evans, 1987; Hickman and Evans, 1992; Richard et al., 2015; Brantut, 2015)
Figure II-MMC1.2: Studies of damage zones that surround natural faults. (a) Geological and seismological observations show variations in damage intensity as well as the associated reductions in elastic moduli with lateral distance from the fault core. Adapted from Aben et al. (2017). (b) Schematics of laboratory-determined failure strength of geological materials for a broad range of strain rates, indicating failure regimes and experimental apparatuses used. As the strain rates increase, the damage mechanisms - and hence the constitutive response - change. Adapted from Grady (1998), Zhang and Zhao (2014), and Aben et al. (2017). (c) Imaging three-dimensional microscale damage evolution towards shear faulting in a quartz-monzonite rock representative of crystalline rocks of the continental crust, using a triaxial rig transparent to X-rays and in situ synchrotron microtomography with a resolution of 6.5 μm. Microfractures nucleate and grow until the damage spans the whole sample. The dynamics and morphology of the microfractures can be quantified from the three-dimensional images, providing an invaluable input to micro-mechanical theories and modeling of damage and fault formation. Adapted from Renard et al. (2018a).
yielding or damage rheology formulations. The latter connect the evolution of elastic moduli and stresses with changes of crack density through damage variables and can be roughly categorized into micromechanics based formulations\(^1\) and thermodynamically based formulations\(^2\). Both approaches have been able to reproduce some of the macroscopic behaviors observed in rock experiments.

Yet realistic but tractable representations of the constitutive response of damage zones remain an open challenge, since they need to capture the effects of generation, interaction, and healing of numerous cracks on multiple scales, including several orders of magnitude between the spatial scales of small laboratory samples and typical field observations. Furthermore, the constitutive response is needed for a wide range of strain rates, from the quasi-static ones relevant to the interseismic periods to dynamic ones that act during dynamic rupture propagation. Therefore, conducting further experiments that elucidate the evolution and healing of damage at the laboratory scale, as well as modeling at the larger scale of the interacting cracks and/or damage-zone widths, is essential for formulating the appropriate constitutive relations for damage zones for use in simulations of large earthquake ruptures.

**MMC1.3 Rheology at and below the brittle-ductile transition and fault loading**

At the bottom or below the seismogenic zone, the response of rocks to loading changes from the elastic-brittle to ductile (or visco-elasto-plastic) due to increases in effective pressure and temperature\(^3\), in what is often called the brittle-ductile transition. The damage zones with fractures are suppressed at such depths and replaced with ductile shear and flow. Shear localization may still persist, as evidenced by a range of geologic and geophysical observations, including deep tectonic tremor, postseismic deformation, and exhumed faults\(^4\), possibly in the form of ductile shear zones that can result from a variety of mechanisms\(^5\) or perhaps at least partially in the form of persisting frictional, velocity-strengthening shear zones\(^6\) (MMC1.1) (Figure II-MMC1.3). Experimental data and geodetic observations suggest that power-law rheology relating stress, strain rate, and grain size - which is usually associated with dislocation creep, but may also apply to the grain-size sensitive diffusion creep - provides the best description for the ductile deformation\(^7\).

The ongoing shear deformation below the seismogenic portions of faults, either on the localized fault extensions or more distributed deformation zones in the bulk, significantly affects the loading of the shallower, seismogenic portions of the fault, and hence the stress conditions for earthquake nucleation and propagation. Indeed, given the immediate proximity of this deeper deformation to the seismogenic regions, it could be the dominating factor in fault loading. Much progress has been made in quantifying the evolution of the deeper deformation throughout the interseismic period and post-seismically, when the

---

\(^1\) (Brace and Bombolakis, 1963; Nemat-Nasser and Horii, 1982; Ashby and Sammis, 1990; Bhat et al., 2012; Kimberley et al., 2013)

\(^2\) (Lyakhovsky et al., 1997, 2011)

\(^3\) (Hobbs et al., 1986; Chester, 1995; Kohlstedt et al., 1995; Hirth and Kohlstedt, 1996, 2004; Niemeijer et al, 2002, 2008,2010a; Muhuri et al., 2003; Tenthorey et al, 2003; Melosh et al., 2018)

\(^4\) (Sibson 1977; Berthe et al. 1979; Poirier 1980; Hearn et al., 2002; Cole et al., 2007; Bürgmann and Dresen 2008; Shelly, 2010a,b; Bruhat et al., 2011)

\(^5\) (Yuen et al., 1978; Lister and Williams, 1979; Brodie, 1980; Brun and Cobbold, 1980; White et al., 1980; Tullis and Yund 1985; Rutter, 1999; Montesi and Hirth, 2003; Gueydan et al., 2003; Behr and Platt, 2011; Takeuchi and Fialko, 2012; Montesi, 2013; Mulyukova and Bercovici, 2018; Allison and Dunham, 2018; Beeler et al., 2018; Jaquet and Schmalholz, 2018; Nachlas et al., 2018)

\(^6\) (Hearn et al., 2002; Bruhat et al., 2011)

\(^7\) (Carter and Tsenn, 1987; Hirth and Kohlstedt,1995, 1996, 2004; De Bresser et al., 1998; Montesi and Hirth, 2003; Rybacki and Dresen 2004; Freed and Burgmann, 2004; Takeuchi and Fialko, 2013; Masuti et al., 2016)
Figure II-MMC1.3. Models for shear zones below the seismogenic depths that control the loading of the shallower seismogenic regions. (a) Conceptual model of a strike-slip fault in an elastic bulk with the seismogenic zone (gray) and creeping regions (yellow), in which the transition to aseismic behavior at depths occurs due to change from rate-weakening to rate-strengthening friction on the fault (Tse and Rice, 1986; Rice, 1993; Ben-Zion and Rice, 1997). Adapted from Jiang and Fialko (2016). (b-c) The transition also occurs in the bulk material, from elastic/brittle layers shallower to viscoelastic layers deeper. The fault-normal distribution of fault-parallel shear strain after an imposed earthquake slip on a shallower fault, computed for an experimentally derived, power-law creep rheology, depends on the water content (D, dry; W, wet). Adapted from Takeuchi and Fialko (2012). (d-f) Simulations of sequences of earthquakes and aseismic slip (SEAS, section II-MMC2) in a 2D quasi-dynamic model that combines depth-dependent friction properties from (a) with a power-law creep rheology like in (b) and full thermomechanical coupling. The viscoelastic creep can shut down the fault slow slip (d-e), which would exist otherwise (f). Adapted from Allison and Dunham (2018). (g) Extrapolation of lab-derived flow laws to conditions at the base of the seismogenic zone is, in part, supported by correlation of rock microstructures developed in the lab and found in the field. Adapted from Hirth et al. (2001). (h) Study of the postseismic response below the 2004 Mw 6.0 Parkfield earthquake concluded that substantial frictional afterslip on a deeper fault extension must have occurred; the preferred model combined both frictional afterslip and viscoelastic response. Adapted from Bruhat et al. (2011). Combined modeling, observational, and laboratory studies can constrain the properties of the fault and bulk regions below the seismogenic depths and determine how the deformation partitions between frictional and ductile shear.
deeper layers are loaded by the coseismic stress changes\textsuperscript{1}, although a transient response of a ductile shear zone governed by power-law creep to sudden stress perturbations is not easy to distinguish from that of fault slip governed by rate and state friction. Such temporal and spatial evolution of the deeper deformation can occur on various scales and, through stress transfer, have an important effect on the shallower earthquake-producing processes, including earthquake interaction. The structure and constitutive behavior of the fault extensions plays a dominating role in constraining the depth of large destructive earthquakes\textsuperscript{2}. As viscous deformation is strongly strain-rate strengthening, variations in strain rate produced by earthquakes in the seismogenic zone above may affect the dominant operative deformation mechanisms and response of the deeper substrate\textsuperscript{3}. Studies of exhumed faults indicate complex mineral structure of the bulk and potential operation of multiple deformation mechanisms\textsuperscript{4}. There are important questions about the role of fluids for both the ductile deformation and any remnants of frictional resistance at such depths\textsuperscript{5}.

Hence it is important to continue develop realistic and tractable representations of the deeper visco-elastic-plastic rheology as well as the structure and shear resistance of potential permanently localized fault extensions or “roots.”

**MMC1.4 Coupling deformation/damage and fluid effects**

Fluids permeate much of the subsurface, including aqueous pore fluids that reside in interstitial spaces within fault gouge or the bulk rock as well as water that is chemically bound in minerals or forms defects in silicates. Fluids have profound effects on the shear resistance of granular fault shear layers (MMC1.1; Appendix B) and constitutive response of deeper ductile rocks (MMC1.3). But they are also strongly coupled with the bulk rock deformation throughout most of the seismogenic zone, with even slight rock contractions or expansions changing pore fluid pressure and inducing fluid flow, and vice versa\textsuperscript{6}. Such coupling processes, often called poroelastic effects, occur both immediately adjacent to slipping zones during fast and slow slip\textsuperscript{7} and on the larger scale, e.g., during postseismic periods\textsuperscript{8}. Poroelastic effects are now recognized as a major component of the induced earthquake process\textsuperscript{9} (sections III-IN3, III-IN4). The permeability structure of fault zones is quite complex and expected to evolve due to the creation and healing of the damage zones\textsuperscript{10} (Figure II-MMC1.4). For example, the damage zones of Alpine fault in New Zealand form a hydraulically active system adjacent to the fault core that likely plays a key role in controlling the evolution of the effective stress, properties, and slip on the fault\textsuperscript{11}; the fault is also a prime target for modeling due to its near-periodic large earthquakes\textsuperscript{12}. Recent experiments reveal that the coupling between deformation and pore-fluid diffusion exerts critical control over the development of fracture patterns as well as the rate of faulting and fault slip\textsuperscript{13}. The evolution of permeability affects fluid

---

\textsuperscript{1} Hetland and Hager 2005; Fay and Humphreys 2006; Fialko, 2006; Barbot et al., 2009; Bruhat et al., 2011; Takeuchi and Fialko 2012, 2013; Hearn and Thatcher 2015; Lambert and Barbot, 2016; Sobolev and Muldashev, 2017; Allison and Dunham, 2018

\textsuperscript{2} Romanowicz and Ruff 2002, Leonard 2010, Denolle et al., 2016, Jiang and Lapusta, 2016; Weng and Yang, 2017

\textsuperscript{3} Rolandone et al., 2004; Verberne et al., 2017

\textsuperscript{4} Hamling et al., 2017; Bürgmann, 2018; Dascher-Cousineau et al., 2018; Melosh et al., 2018; Rowe et al. 2018

\textsuperscript{5} Kronenberg et al 1990; Scholz, 1990; Moore and Rymer, 2007; Hirth and Beeler, 2015; Mitchell et al., 2016; Noda and Takahashi, 2016; Khamrat et al., 2018

\textsuperscript{6} Wang, 2000

\textsuperscript{7} Lachenbruch, 1980; Rice, 2006; Dunham and Rice, 2008; Noda et al., 2009; Heimisson and Dunham, 2019

\textsuperscript{8} Sleep and Blanpied, 1992; Peltzer et al., 1998; Cocco and Rice, 2002; Jonsson et al., 2003; Filako, 2005

\textsuperscript{9} Segall, 1989; Segall and Lu, 2015; Goebel et al., 2017; Goebel and Brodsky, 2018

\textsuperscript{10} Wibberley and Shimamoto, 2003; Sutherland et al., 2012; Boulton et al., 2017; Townend et al., 2017; Faulkner et al., 2018; Janku-Capova et al., 2018

\textsuperscript{11} (e.g., Townend et al., 2017)

\textsuperscript{12} Berryman et al., 2012

\textsuperscript{13} French et al., 2016; French and Zhu, 2017
flow into or out of the actively shearing zones and hence the variations of pore pressure there, influencing fault strength (section MMC1.1). Rapid deformation of distant faults due to dynamic wave propagation may change fault-zone hydro-mechanical properties, causing delayed long-distance earthquake triggering. 

Hence it is important to both more consistently treat the rock bulk as a poroelastic - rather than just elastic - medium as well as develop tractable representations for evolution of poroelastic properties due to damage, dynamic wave propagation, and fluid injection.

Figure II-MMC1.4 Heterogeneous structure and hydromechanical properties (permeability) in the fault-normal direction. Sketch summary of the main elements of permeability structure across the Median Tectonic Line. (a) Summary of the structural zones; (b) summary permeability data for different confining pressures (stated at the bottom, with * denoting data from the deconfining path), for 20 MPa pore pressure, given the mapped distribution of fault rocks. Note that the fault-normal distance is logarithmic. Adapted from Wibberley and Shimamoto (2003). Mesoscale models of fault behavior are needed to understand the impact of this complexity and the associated fluid effects on larger scales.

1 (Brodsky and Prejean, 2005; Manga et al., 2012; Cox et al., 2015; O’Brien et al., 2016; Weaver et al., 2019)
MMC1.5 Formulating scale-appropriate constitutive laws for fault response

The shear deformation within the fault shear zones is often idealized as slip - or relative shear displacement - across a sharp interface, and the response of the surrounding medium is often idealized as elastic, both for reporting field or lab observations and for modeling purposes. For example, kinematic inversions of large earthquakes routinely report the final slip distribution, given the assumed fault geometry and elastic bulk structure. In dynamic earthquake source models with such assumptions, the entire constitutive response of the fault structure is then replaced by the constitutive response of interfaces embedded into an otherwise elastic bulk. Such or similar simplifications may be unavoidable for tractable earthquake source modeling at large spatial scales and with long simulated histories. Yet the constitutive response then has to be carefully constructed to capture, with some abstraction, the clearly relevant factors discussed in MMC1.1-1.4.

One aspect of the fault geometry that can intrinsically link the shear-layer resistance to slip and the inelastic deformation of the surrounding bulk is fault roughness or non-planarity. Furthermore, the actively shearing layers of seismogenic faults, often called fault cores, can vary greatly in width along strike, from the very localized, sub-millimeter deformation to much broader zones of the order of tens of meters with multiple shear surfaces. It is important to elucidate the factors that control the differences in structural complexity among faults; one promising aspect is the level of fault maturity. Deeper fault extensions of exhumed fault systems often consist of networks of thin discrete fault strands encompassed within zones of distributed, volumetric deformation, which may occur at variable slip rates throughout the earthquake cycle.

The explicit inclusion of structural complexities on the scale of 1-100 m or less is clearly not tractable in numerical models of earthquake processes at the larger, societally relevant scales. Hence, to achieve theoretically sound, laboratory-based, yet tractable earthquake source modeling, we need to formulate scale-appropriate constitutive relations by: (i) clearly defining the spatial scales that are subsumed by the notion of fault slip or other kinematic variables; (ii) identifying the relevant physical mechanisms for the constitutive response at those scales; and (iii) capturing their coupled effects in the appropriately formulated constitutive relations (section III-IN1), e.g., through smaller-scale modeling.

Note that while most of the geological investigations have been done in the more accessible continental faulting environments, a number of recent field studies have targeted exhumed subduction thrusts at a range of paleo-depths. The subduction environment differs in several ways, and much work has focused on how composition, fault structure, shear resistance, and bulk rheology change as materials and structures like seamounts and outer-rise faults are transported downward along the subducting plate interface (often called the megathrust). In an obvious difference, geodetic studies highlight spatial variations in the locking and creeping of many megathrusts along the trench, while continental faults are more commonly locked everywhere at the similar range of “seismogenic” depths, with some exceptions.

1 (Simons et al., 2011; Minson et al., 2014ab; Adams et al., 2016; Ye et al., 2016, 2018; Ragon et al., 2018)
2 (Brown and Scholz 1985, Power et al., 1987; Sagy et al., 2007, Dieterich and Smith, 2009; Brodsky et al., 2016; Candela and Brodsky, 2016)
3 (Chester and Logan, 1987; Chester et al., 1993; Wibberley and Shimamoto, 2003; Sutherland et al., 2012)
4 (Faulkner et al., 2003, 2008; Zoback et al., 2010; Savage and Brodsky, 2011; Rowe et al., 2013)
5 (Ben-Zion and Sammis, 2003; Bohnhoff et al., 2016; Perrin et al., 2016)
6 (Okubo and Aki, 1987; Rowe et al., 2011; Bürgmann, 2018; Melosh et al., 2018; Rowe et al. 2018)
7 (Fagereng and den Hartog, 2017; Rowe et al., 2011; Angiboust et al., 2015; and Behr et al., 2018)
8 (Song and Simons, 2003; Wang and Bilek, 2014; Wells et al., 2003; Fuller et al., 2006; Keren and Kirkpatrick, 2016; Saillard et al., 2017; Regalla et al., 2018)
9 (Wang and Dixon, 2004; Avouac, 2015)
Figure II-MMC1.5. Deriving scale-appropriate constitutive relations. (a) Earthquake source modeling at crustal scales aims to capture dynamic rupture and slow slip across the scale of 10-100 of kilometers. The modelers have been working to include structural complexities at such scales, but it would be quite difficult to capture structural complexity on much smaller scales, such as (b) fault roughness on scales comparable to slip, which would couple the shear resistance with the evolving damage and hydro-mechanical properties on such scales; (c) multi-shear-zone faults, and (d) variations in lithology, grain size, architecture, and structural fabric below seismogenic depths. Images adapted from Dieterich and Smith (2009), Mitchell and Faulkner (2009), and Melosh et al. (2018). The response of such structural complexities on the scale of 1 to 100 meters or less needs to be studied in smaller-scale models and turned into appropriate constitutive relationships.
believed to reflect mineralogical differences in the shallow fault zone\(^1\). Another important difference is in the occurrence of episodic slow slip events and tremor on many (but not all) megathrusts but only some continental faults\(^2\). At the same time, many aspects of the seismic/aseismic behaviors are universal and the inferred ranges of dynamic properties for earthquake ruptures are similar in both continental and subduction environments, in terms of their average stress drop and rupture velocity (MMC3), suggesting similarity in many generic aspects of the fault structure, nearby bulk, and mechanisms involved.

For more detailed descriptions of MMC1.1-1.5, refer to Section V in Appendix B.

**Key future goals**

**MMC1: Developing constitutive laws for deformation/faulting: capturing small-scale processes**

**MMC1.1 Shear resistance of granular fault cores: localization, dilation, shear heating, and healing**

- Conduct laboratory experiments and modeling to capture the evolving response of granular fault-core layers permeated by fluids under realistic conditions, including (i) particle distribution and mineralogy inferred for natural faults, (ii) strain rates spanning from near-locked to dynamic sliding conditions, and (iii) a range of temperature/confined/flow pressure conditions representative of natural faults for the shallow, mid-seismogenic, and below-seismogenic depths;
- Design laboratory experiments to enable insightful modeling;
- Determine dominant mechanisms controlling friction at high slip rates; in particular, assess the relative importance of thermal pressurization of pore fluids and dilatant hardening.

**MMC1.2 Off-fault damage/plasticity accumulation and healing**

- Elucidate constitutive relations for damage creation during earthquake ruptures: experiments and modeling of rapid (high-strain-rate) damage to ascertain the strain-rate dependence, damage patterns, and the associated elastic moduli evolutions under representative stress/temperature/flow conditions for the shallow, mid-seismogenic, and bottom-seismogenic depths;
- Investigate damage healing in the interseismic period: mechanical vs. chemical;
- Capture all findings in constitutive relations for damage zones on 0.1-100 meter scale needed for large-scale models: beyond the laboratory scale.

**MMC1.3 Rheology at and below the brittle-ductile transition and fault loading**

- Determine structure and constitutive response of deep localized fault “roots,” friction vs. flow: frictional slip interfaces presumably grade into ductile shear zones below the seismogenic zone, but the details of the combined structure and its constitutive response are poorly understood;
- Construct bulk flow rules for multi-mineral rocks and multiple concurrent deformation mechanisms: building on power-law rheologies;
- Investigate fluid effects: changes in flow laws and the role of effective stress in the ductile roots of fault zones.

**MMC1.4 Coupling deformation/damage and fluid effects**

- Incorporate fully coupled hydromechanical (e.g., poroelastic) bulk effects into models of earthquake source processes, which are potentially key to a number of earthquake effects including induced seismicity;

---

\(^1\) (e.g., Bürgmann et al., 2000; Jolivet et al., 2012, 2015; Rousset et al., 2016)
\(^2\) (e.g., Nadeau and Dolenc 2005; Shelly, 2017; Chen et al., 2018; Rousset et al., 2016; 2018)
- Determine constitutive relations for the evolution in hydro-mechanical properties due to near-fault damage and healing, which can have profound effects on the evolution of pore fluid pressure in the fault core and hence on the fault core shear resistance;
- Explore evolution in hydro-mechanical properties due oscillating dynamic stresses induced by wave propagation (shaking), potentially dominating effect for distant earthquake triggering.

**MMC1.5 Formulating scale-appropriate constitutive laws for fault response**
- Determine laws for effective shear resistance of localized shear layers and adjacent areas, due to the coupled evolution of localized friction of the spatially non-planar shear layer, nearby damage, and the resulting fluid flow;
- Determine laws for effective shear resistance of more complex fault cores, with multiple shear layers embedded into a damage layer, including near-surface “flower” structure;
- Distinguish between mature vs. immature faults through constitutive relations; potential candidates include the degree of local roughness, mineralogy, particle distributions, and width of the actively shearing granular layer.

**Cross-cutting themes in MMC1.1-1.5:**
- Systematic community-wide effort to produce coherent sets of constitutive laws for realistic fault materials under realistic fault conditions, using systematic sets of experiments, unified theoretical frameworks, and modeling at the lab scale to clarify and untangle various effects.
- Numerical experiments to construct constitutive laws on the intermediate scales of 0.1-100 m - larger than the typical lab samples but smaller than the spatial discretization in large-scale simulations - needed for tractable numerical modeling, with predictions for lab and field studies to test; such numerical experiments could capture the effects of intermediate-scale complexity and use lab-derived laws at lower scales.
MMC2. Building a coherent suite of numerical methodologies for multi-physics problems at larger scales

Continuum-mechanics-based models of earthquake and faulting processes are constructed using the constitutive response of faults and off-fault materials, as discussed in section MMC1, making them multi-physics, multi-scale, and nonlinear. The resulting equations are solved for the time evolution and spatial distributions of all relevant quantities - including fault slip, off-fault deformation, stresses, and other variables - which can only be done numerically in all but the simplest cases. These numerical solutions are known as earthquake source simulations.

Explicit inclusion of all scales and processes relevant to earthquake source processes (section I.3, MMC1 and MMC3) in a single “master” model is not possible now or in the foreseeable future, as the resulting simulations would be intractable. Addressing certain questions may require only a subset of the processes and scales involved and allow the modeler to employ justifiable approximations to simplify the problem and associated simulations in various ways. Indeed, an important goal of developing multi-scale, multi-physics modeling capabilities is to identify justifiable simplifications that can be used for simulations at larger temporal and spatial scales, at least in some common and/or important cases, such as mature continental faults and megathrusts that host most destructive earthquakes.

One of the choices is whether to conduct three-dimensional (3D) or two-dimensional (2D) simulations. Since one dimension is always needed to represent the off-fault stress transfer, the fault is represented as a 2D interface in 3D models but simplified to a one-dimensional (1D) representation (essentially, a straight or curved line) in 2D models. 2D simulations are indispensable as an exploration tool, allowing the modeler to consider wider ranges of processes, parameter spaces, and constraints from field and laboratory observations; however, 2D approaches are also limited in that they can qualitatively change the outcomes of certain problems, especially those involving fault roughness or heterogeneity. 3D models can provide a more realistic representation and facilitate more direct comparison with observations, but they are obviously numerically much more challenging, and hence could especially benefit from justifiable simplifications developed by smaller-scale and 2D models.

The goals and corresponding scope of earthquake source simulations generally fall into the following broad categories (MMC2.1-2.5). They all have significant accomplishments in addressing the motivating questions posed in Section I.2. With systematic future progress, they will evolve into a coherent set of complementary, physics-based numerical methodologies that, together, will lead to predictive, physics-based modeling that we seek to accomplish (MMC3, MMC4).

MMC2.1 Dynamic rupture simulations: capturing a single earthquake event

*Dynamic rupture simulations* focus on the coseismic phase of one earthquake rupture, addressing questions regarding how ruptures propagate and arrest for prescribed initial conditions on the fault and in the bulk, interact with frictional and structural heterogeneities, and radiate seismic waves. While such simulations clearly have to deal with only a subset of temporal scales, up to the typical duration of a large earthquake (of order 100 s), they face the full daunting range of spatial scales discussed in section I.3. Dynamic rupture simulations have been used to study large earthquake ruptures, typically by adjusting stress and frictional parameters to obtain a match between predicted and observed strong ground motion records and other available data. These simulations have provided quantitative constraints on coseismic
stress changes, rupture velocity, as well as dissipated and radiated energy\(^1\). While matching has typically been done by hand-tuning model parameters, recent efforts have focused on using dynamic rupture simulations in formal inversions\(^2\). Dynamic rupture simulations revealed the possibility of supershear propagation\(^3\), with its unique ground motion characteristics\(^4\), features later confirmed in the laboratory and inferred for many strike-slip earthquakes\(^5\) (Figure II-MMC2.1-1). Dynamic rupture simulations have also elucidated fundamental aspects of how ruptures navigate complex fault geometry and interact with fault roughness\(^6\) (Figure II-MMC2.1-2). The implications of proposed fault weakening processes, such as strongly rate-weakening friction and thermal pressurization, on rupture propagation have been quantified using dynamic rupture simulations, providing insight into why ruptures usually take the form of narrow slip pulses\(^7\).

Future progress in understanding and capturing rupture dynamics and strong ground motions will come from incorporating the full range of coupled mechanisms discussed in section MMC1, including complex fault geometries, off-fault inelastic processes, realistic shear-layer response, and fluid effects, which brings significant programming and computational challenges. One limitation of dynamic rupture simulations is the need to prescribe initial conditions on the fault which are not independently known and, in fact, likely depend on the physical assumptions (MMC3.1). A promising way forward in that regard is to incorporate physics-based inputs into the initial conditions motivated by longer-term simulations (MMC2.2-2.4). Dynamic rupture modeling will play an increasingly important role in quantifying the source contributions to strong ground motion (MMC 4.1).

**MMC2.2 Modeling sequences of earthquakes and slow slip/deformation**

Simulations of longer-term earthquake source behavior aim to capture fault and lithospheric deformation at the time scales of decades to thousands of years. These simulations often span multiple seismic events and interseismic periods. Most simulations of this type restrict attention to a single fault. They capture sequences of earthquakes and aseismic slip (SEAS), and are referred to as SEAS or earthquake cycle simulations, though the "cycle" terminology should not be understood to imply periodic event sequences. SEAS simulations encompass multiple instances of earthquake nucleation, coseismic rupture propagation, postseismic response, and aseismic processes in the interseismic periods, all occurring spontaneously in response to applied tectonic loading. SEAS simulations must therefore solve the quasi-static deformation problem. Most SEAS simulations are quasi-dynamic, a term that refers to a technique that replaces inertial dynamics during the coseismic phase with the radiation-damping approximation\(^8\).

SEAS simulations have provided insights into a broad range of slip phenomena, including earthquake nucleation, postseismic response from afterslip, patterns of seismic and aseismic slip, and aftershock sequences, including how the depth extent of earthquakes might be limited by the transition from velocity-weakening to velocity-strengthening friction that occurs at sufficiently high temperature\(^9\). SEAS

---

\(^1\) (Archuleta, 1984; Mikumo and Miyatake, 1995; Olsen et al., 1997; Mai and Beroza, 2002; Oglesby and Day, 2002; Ma et al., 2008)

\(^2\) (Peyrat and Olsen, 2004; Ruiz and Madariaga, 2011, 2013; Ruiz et al., 2017; Gallovic et al., 2018)

\(^3\) (Andrews, 1976a; Das and Aki, 1977b)

\(^4\) (Freund, 1979; Aagaard and Heaton, 2004; Dunham and Archuleta, 2004, 2005)

\(^5\) (Archuleta, 1984; Bouchon et al., 2001; Bouchon et al., 2002; Rosakis, 2002; Bouchon and Vallée, 2003; Ellsworth et al., 2004; Xia et al., 2004; Vallée et al., 2008)

\(^6\) (Harris and Day, 1993; Kame et al., 2003; Dunham et al., 2011b; Galvez et al., 2014, Douilly et al., 2015; Lozos et al., 2015, Withers et al., 2018a, Wolherr et al., 2018, Ando and Kaneko, 2018; Ulrich et al., 2018, 2019)

\(^7\) (Heaton, 1990; Zheng and Rice, 1998; Nielsen and Carlson, 2000; Noda et al., 2009; Gabriel et al., 2012)

\(^8\) (Rice, 1993)

\(^9\) (Tse and Rice, 1986; Shibazaki and Matsu‘ura, 1992; Dieterich, 1992, 1994; Ben-Zion and Rice, 1997; Lapusta and Rice, 2003; Kato and Tullis, 2003; Hori et al., 2004; Duan and Oglesby, 2005; Rubin and Ampuero, 2005;
Supershear rupture propagation discovered through modeling, confirmed in the lab, and inferred for many large strike-slip earthquakes. (a) Shear wave Mach fronts are a distinctive feature of supershear ruptures, shown here in the particle velocity field surrounding a steadily propagating supershear slip pulse (from an analytical solution). Adapted from Dunham and Archuleta (2005). Such propagation and Mach fronts have been observed in laboratory earthquake experiments (Figure I-2, bottom-right panel, adapted from Xia et al., 2004). (b-c) The transition from sub-Rayleigh to supershear rupture speed can generate both a supershear rupture and a trailing slip pulse propagating near the Rayleigh wave speed, both having distinctive ground motion characteristics matching those observed in the 2001 Mw 7.9 Denali Fault, Alaska, earthquake, as established in a 3D dynamic rupture simulation. Similar conclusions were reached in a laboratory earthquake study (Mello et al., 2013). Adapted from Dunham and Archuleta (2004). (d-e)
Figure II-MMC2.1-2. 3D dynamic rupture simulations with inelastic effects and complex fault geometry. (a-b) Strong ground motion (spectral accelerations at 3 s) obtained in a Mw 7.8 southern San Andreas dynamic rupture scenario for (a) linear medium and (b) medium that includes shallow soil nonlinearity. The dashed line shows the fault. Shallow soil nonlinearity significantly reduces ground motions. Adapted from Roten et al. (2018). (c-d) Dynamic rupture simulation of the 2016 Mw 7.8 Kaikoura earthquake on a complex fault network. To sustain dynamic rupture propagation in such a geometry, higher initial stresses at the bottom of the seismogenic regions were assumed as would be consistent with interseismic loading due to deeper creep. Adapted from Ulrich et al. (2018).
simulations have also provided insight into the mechanisms that cause episodic slow slip events and tremor\(^1\) (Figure II-MMC2.2). Simulations both within the continuum framework\(^2\) and with inherently-discrete models\(^3\) highlighted the role of strong fault heterogeneities in generating earthquake populations that satisfy the Gutenberg-Richter frequency-size statistics and Omori decay rates of aftershocks (section MMC3.5). Some SEAS simulations do account for wave-mediated inertial effects during dynamic rupture for multiple sequences of earthquakes, in both 2D\(^4\) and 3D\(^5\); the 3D simulations restrict attention to the highly simplified geometry of a single planar fault in an otherwise elastic bulk, focusing on relatively sophisticated representations of the fault resistance that includes shear heating effects, off-fault diffusion of both heat and fluids, and enhanced dynamic weakening. The inclusion of all wave-mediated inertial effects is essential to elucidate the full effects of enhanced dynamic weakening in earthquake sequence simulations, such as the formation of self-healing slip pulses and the resulting low-stress, low-heat fault operation\(^6\) (II-MMC3.1). Other SEAS simulations, mostly in 2D, have started to explore the effects of off-fault yielding, highlighting its role in the observationally inferred shallow slip deficit\(^7\); viscous flow at depth and its interplay with deeper fault slip\(^8\) (Figure II-MMC1.3); fault roughness\(^9\), and coupling to fluid flow\(^10\).

The important next challenge is to develop 2D and 3D simulations that can simulate SEAS (sequences of earthquakes and aseismic slip) with the combined effects of damage evolution, deep viscoelasticity, fluid flow, and fault geometrical complexity, for more realistic fault and bulk properties, with at least some implementations also resolving the full inertial effects during dynamic rupture. An important advantage of SEAS-like simulations over dynamic rupture simulations is the ability to investigate the feedbacks between rapid and slow processes and to capture the long-term evolution of stresses and other variables, resulting in the fault conditions before dynamic rupture consistent with the model ingredients. (Dynamic rupture simulations, in contrast, have to assume the fault initial conditions.) Such SEAS approaches can be constrained by a wide range of earthquake source observations (MMC3), and eventually provide physics-based insight into the potential for extreme events (MMC4.2) and long-term seismic hazard (MMC4.3).

**MMC2.3 Extending simulations to fault networks**

Most large earthquakes occur in the context of fault networks\(^11\) and many span several fault segments. This reality brings about the issue of earthquake interaction and highlights the need of simulating earthquake source processes over several fault segments or even regional-scale fault networks (Figure I-3, top-left panel\(^12\)). Both dynamic rupture (MMC2.1) and SEAS simulations (MMC2.2) have had notable successes in investigating earthquake propagation in nonplanar and multi-segment fault models, including step-overs and branched geometries\(^13\) (Figure II-MMC2.3), and these developments will

---

1 Shibazaki and Iio, 2003; Liu and Rice, 2005, 2007; Rubin, 2008; Segall et al., 2010; Shibazaki et al., 2010; Li and Liu, 2016; Luo and Ampuero, 2017
2 (e.g., Ben-Zion and Rice, 1997; Hillers et al., 2006; 2007)
3 (e.g. Ben-Zion and Rice, 1993; Ben-Zion, 1996; Zöller et al., 2005)
4 (Ben-Zion and Rice, 1997; Lapusta et al., 2000; Kaneko et al., 2011; Tal and Hager, 2018; Duan et al., 2019)
6 (Thomas et al., 2014)
7 (Kaneko and Fialko, 2011, Erickson et al., 2017)
8 (Lambert and Barbot, 2016; Allison and Dunham, 2018)
9 (Tal and Hager, 2018; Ozawa et al., 2019)
10 (McClure and Home, 2011; Heimisson and Dunham, 2018; Torberntsson et al., 2018)
11 (Plesch et al., 2007; Nicholson et al., 2014)
12 adapted from Plesch et al., 2007
13 (Harris et al., 1991, 2002; Harris and Day, 1999; Kame et al., 2003; Duan and Oglesby, 2006; Elliott et al., 2009; Galvez et al., 2014, Lozoz et al., 2015, Li and Liu, 2016; Withers et al., 2018a, Wollherr et al., 2018, Ando and Kaneko, 2018; Ulrich et al., 2018,2019)
Figure II-MMC2.2. 3D quasi-dynamic simulations of episodic slow slip events in the Cascadia subduction zone using non-planar geometry and rate-and-state friction laws. The modeled events capture the major characteristics revealed by GPS observations. (a) The 3-D northern Cascadia subduction model with a nonplanar geometry highlighted by a triangular mesh. PA: Port Angeles and SEA: Seattle. (b-c) Comparison between the computed moment rate (b) and observed detrended east component (c) of daily GPS time series of from 1997 to 2010 (13 years) at station ALBH (Pacific Northwest Geodetic Array: http://www.geodesy.cwu.edu/). (d-e) Simulated cumulative slip (d) and its percentage of the plate convergence (e) averaged for eight episodic slow slip events over 13 years. Adapted from Li and Liu (2016).
continue, aiming to include more realistic fault geometry and physics. Simulations with damage-rheology frameworks explored the coupled evolution of earthquake ruptures and fault geometry. Fully dynamic rupture simulations with realistic fault geometries can substantially advance our understanding of rupture interaction through fleeting but substantial stress changes brought by dynamic waves. SEAS simulations, even over a small number of interacting segments, can work to explore the full range of the proposed earthquake interaction mechanisms, determine their coupled effects, and explore the possibility of proposing simplified representations of the interactions at larger scales.

For the goal of simulating earthquake sources over regional networks for tens of thousands of years, several earthquake simulators have been developed. The term “simulators” typically refers to approaches that employ significant simplifications, compared to most SEAS simulations, in solution procedures and physical processes, in order to simulate earthquake sequences on complex, regional-scale 3D fault networks for tens of thousands of years or even longer. For example, earthquake simulators typically account only for the quasi-static stress transfer due to earthquake events, ignoring wave-mediated stress changes, aseismic slip/deformation, and fluid effects; employ approximate rule-based update schemes for earthquake progression instead of solutions of the governing continuum-mechanics equations; and use oversized numerical discretization. Such simplifications are currently necessary to permit computational simulation at the vast range of length and time scales involved. Earthquake simulators have matched a number of region-scale statistical relations (MMC3.5), including the Gutenberg-Richter scaling, and highlighted the importance of large-scale event interactions. The RSQSim simulator, that incorporates key properties of rate-and-state-based earthquake nucleation, also matches the Omori’s law statistics for aftershocks as well as can be tuned to reproduce long-term hazard forecasts. Since only earthquake simulators are currently capable of incorporating regional-scale fault networks, they can provide invaluable insight into the extent of earthquake interaction due to quasi-static stress changes and epistemic uncertainties associated with incomplete knowledge of fault network geometry and fault properties.

In studies of fault networks, the question of dominating mechanisms of earthquake interaction comes to the fore. Aftershock sequences and their properties, including the Omori’s law for their decay (section MMC3.5), have been explained by a number of mechanisms, including static stress changes imposed by mainshock coupled with rate-and-state effects; increased loading rate due to aseismic processes such as postseismic slip, relaxation of the viscoelastic lower crust, pore fluid motion and induced variations in fault strength, triggering due to dynamic stress changes, and evolution of irreversible damage in a viscoelastic damage rheology model. In addition to the static stress transfers due to co-seismic slip, two of the earthquake-interaction effects can be particularly important: loading of the seismogenic faults by the deeper aseismic shear slip or deformation (section MMC1.3) which concentrates the stress at the bottom of the seismogenic layer (Figures II-MMC2.1-2, III-MMC3.1-2) and dynamic triggering of nearby

---

1. (Ben-Zion et al., 1999; Lyakhovsky et al., 2001; Finzi et al., 2009; Kurzon et al., 2019)
2. (Dieterich and Richards-Dinger, 2010; Pollitz, 2012; Richards-Dinger and Dieterich, 2012; Tullis et al. 2012a, b; Sachs et al., 2012; Ward, 2012; Shaw et al., 2018)
3. (Dieterich and Richards-Dinger, 2010; Richards-Dinger and Dieterich, 2012; Tullis et al. 2012a, b; Shaw et al., 2018)
4. (e.g., Freed 2005)
5. (Dieterich, 1994; Gross and Bürgmann, 1998; Toda et al., 1998, 2005; Gomberg et al., 2005)
6. (e.g., Benioff, 1951; Perfettini and Avouac, 2004)
7. (e.g., Freed and Lin 2001)
8. (e.g., Nur and Booker, 1972; Bosl and Nur, 2002)
9. (e.g., Hill et al. 1993; Gomberg et al., 2003; Felzer and Brodsky, 2006)
10. (e.g., Ben-Zion and Lyakhovsky 2006)
and distant earthquakes by seismic waves, either directly or through modified hydromechanical and other fault properties\(^1\).

Future progress in capturing the interactive response of fault networks will come from more advanced dynamic rupture and SEAS simulations, more rigorous earthquake simulators, and coupled approaches (MMC2.5).

---

**Figure II-MMC2.3** 2D dynamic modeling of earthquake sequences on two non-planar fault strands, for the Altn Tagh Fault (ATF) near the Aksay bend in northeastern Tibet. (I) 2D fault geometry obtained from the fault surface trace. Two strands (SATF and NATF) coexist and are active. (II)-(VI) Examples of ruptures over 500 cycles. Events #167 and #456 are examples of jumping ruptures, which propagate onto both strands outside the bend. Other events mainly rupture one of the two strands, and the bend acts as a barrier to dynamic ruptures in these events. Adapted from Duan et al. (2019).

---

\(^1\) (Hill et al., 1993; Gomberg et al., 1994; Gomberg et al., 2001; Vidale and Li, 2003; Felzer and Brodsky, 2006; Nakata and Snieder, 2011; Manga et al., 2012; Xue et al., 2013; Hill and Prejean, 2014)
MMC2.4 Modeling tectonics and earthquakes: variability in loading and geometry

Tectonic earthquake-sequence simulations (TEC-SEAS) aim to expand the approaches developed to study long-term processes of geodynamic evolution and formation of fault systems to shorter time scales and physical processes relevant to individual earthquakes and their interseismic periods\(^1\). Advantages of TEC-SEAS simulations include spontaneous and self-consistent modeling of the evolution of hydro-thermo-mechanical structures in the bulk and fault zones\(^2\), realistic tectonic loading through slab pull\(^3\), seismicity in agreement with Gutenberg-Richter statistics, and seismic and aseismic fault growth\(^4\), facilitating comparison with geological observations (Figure II-MMCC2.4). These developments have led to insights on tectonic controls over long-term spatio-temporal seismicity patterns\(^5\), the early onset of visco-elastic relaxation\(^6\), the occurrence of slow slip\(^7\), and to identification of a secondary zone of uplift due to megathrust earthquakes\(^8\). The findings demonstrate the importance of modeling tectonic processes, or approximations thereof, in complex, geological settings driven by external, non-fault perpendicular forcing\(^9\).

TEC-SEAS simulations are rapidly advancing and have started to implement the physical processes relevant to fault zones over the scales of rapid seismic and aseismic slip (section II-MMC1), starting with representations of the standard rate-and-state friction (MMC1.1) while using a single constitutive formulation for on- and off-fault visco-elasto-plasticity\(^10\). Whereas conventional SEAS simulations (MMC2.2) implement friction in terms of slip velocity and stress tractions on fault interfaces, TEC-SEAS simulations employ finite-thickness shear zones, often localized to one grid spacing, with slip velocity and tractions replaced with invariants of plastic strain rate and invariants of the stress tensor, respectively. For numerical tractability, the modeling is currently 2D and sometimes resolves earthquake slip over years (Figure II-MMCC2.4), although some problem settings allow to resolve fault slip down to minutes\(^11\) and even milliseconds\(^12\).

Future developments will focus at crossing time and spatial scales in a more complete, accurate and computationally efficient manner in 2D and 3D models, improving unified constitutive relations and non-linear rheological implementations, and incorporating fluid flow and poro-elasticity. These developments present a promising avenue for developing SEAS-like simulations with realistic - and evolving - fault geometries, inelastic fault zones and bulk, hydro-thermo-mechanical coupling, and tectonic loading. Furthermore, these approaches open the opportunity to study the role of, and feedback with, geological and tectonic structures and processes.

---

\(^{1}\) (van Dinther et al., 2013ab, 2014; Lavier et al., 2013; Sobolev and Muldashev, 2017)

\(^{2}\) (van Dinther 2013b, Dal Zilio et al., 2018, van Zelst et al. 2019)

\(^{3}\) (D’Acquisto et al., 2018)

\(^{4}\) (Preuss et al., 2019)

\(^{5}\) (Herrendoerfer et al., 2015; Corbi et al., 2017; Brizzi et al., 2017; Dal Zilio et al. 2018)

\(^{6}\) (Sobolev and Muldashev, 2017)

\(^{7}\) (Tong and Lavier, 2018)

\(^{8}\) (van Dinther et al., 2019)

\(^{9}\) (Brizzi et al., 2017; D’Acquisto et al., 2018; van Zelst et al, 2019)

\(^{10}\) (van Dinther et al. 2014; Herrendoerfer et al., 2018; Tong and Lavier, 2018)

\(^{11}\) (Sobolev and Muldashev, 2017)

\(^{12}\) (Herrendoerfer et al., 2018, Preuss et al., 2019)
Modeling earthquake source processes: from tectonics to dynamic rupture

Figure II-MMC2.4. Across-scale modeling from tectonic to seismic processes. (a-b) Development of self-consistent long-term lithological/thermal structure (a) and seismicity (b) in a 2D geodynamic subduction-zone model. Rupture events with Gutenberg-Richter frequency-size distribution spontaneously occur. The inferred focal mechanisms show a broad pattern of different styles of faulting that are consistent with the local tectonic regime. Adapted from Dal Zilio et al. (2018). (c-d) Slip rate evolution (colors) during a megathrust event in a 2D first-generation TEC-SEAS model (c) and for a coupled approach in which the fault conditions from that simulation are used in a fully dynamic rupture model (d). Note similarities in the spontaneous rupture nucleation, propagation, and arrest, while time scales are very different. Rupture complexities result from bulk lithological complexities. Adapted from van Zelst et al. (2019).
MMC2.5 Frameworks for coupling models at different scales

A promising approach to rigorously bridge the many temporal and spatial scales involved in modeling earthquake source processes is to couple numerical computations at different scales. A number of SEAS studies (MMC2.2) have successfully used coupling between quasi-dynamic and fully-dynamic codes to study the transition from spontaneous nucleation to dynamic rupture propagation or to simulate long-term fault slip. Recently, coupling between a TEC-SEAS approach (MMC2.4) and dynamic rupture simulation was applied to study the impact of bulk structure and subduction dynamics on dynamic rupture propagation as well as the impact of dynamic inertia effects on fault slip and stress drop (Figure II-MMС2.4). This approach can be extended to simulations of fault networks, in which a simulation of interseismic loading of a fault network is coupled to dynamic rupture simulations only on specific, activated, segments. Machine learning can potentially also be integrated in these modeling paradigms to bridge the conundrum of scales (section IN.2). For example, machine learning may be used to learn the nonlinear relations between distributions of pre-event fault conditions in a long-term simulation and dynamics of the resulting seismic events, as well as to accelerate numerical calculations during the interseismic period.

For more detailed descriptions of MMC2.1-2.5, refer to Section V in Appendix B.

Key future goals
MMC2: Building a coherent suite of numerical methodologies for multi-physics problems at larger scales

MMC2.1 Dynamic rupture simulations: capturing a single earthquake event
- Develop dynamic rupture simulations with scale-consistent representations of all relevant processes, including fault geometry and evolving fault/bulk structure/properties (shear-layer resistance, damage zones, fluid effects, MMC1.1-1.4), to capture coupling between different physical mechanisms and evaluate their relative importance for rupture dynamics;
- Explore the best ways of selecting initial conditions, based on fault-resolved regional loading and insight from long-term simulations;
- Construct workflows for ensemble dynamic rupture simulations to predict strong ground motions.

MMC2.2 Sequences of earthquakes and slow slip/deformation
- Develop SEAS simulations that incorporate all or large combinations of relevant ingredients, including nonplanar faults, extended rate-and-state friction formulations including enhanced dynamic weakening (MMC1.1), off-fault damage and healing (MMC1.2), deeper viscoplastic regions (MMC1.3), fluid effects (MMC1.4), and wave-mediated stress transfers during co-seismic rupture;
- Employ the multi-physics simulations to evaluate the relative importance of different ingredients, for a range of parameter values, and formulate simplified representations of coupled effects;
- Use the developed simulations to constrain fault physics by reproducing multiple types of observations with the same modeling ingredients (as discussed in MMC3).

MMC2.3 Extensions of earthquake source simulations to fault networks
- Continue to develop multi-physics dynamic rupture simulations on complex geometries to (i) advance our understanding of the effects of fault geometry on fault slip, energy budget, and strong ground

1 (Hajarolavadi and Elbanna, 2017; Ma et al., 2018, van Zelst et al., 2019)
2 (Shibazaki and Matsu’ura, 1992; Schmitt et al., 2015)
3 (Kaneko et al., 2011)
4 (van Zelst et al., 2019)
motion, and (ii) study rupture interaction through fleeting but substantial dynamic stress changes carried by waves;

- Use SEAS modeling on single and several faults to study various proposed interaction mechanisms of earthquake rupture events, including static stress changes due to events themselves, static stress changes due to postseismic and interseismic creep/deformation, dynamic stress changes, and fluid effects including poroelastic stress changes; quantify their effects and relative importance at different distances and temporal scales; and determine dominating mechanisms that need to be included in regional-scale simulations;
- Develop methodologies for regional-scale earthquake simulators over fault networks that include simplified representations of multiple effects identified by more detailed studies; use them as research tools to understand which features of network geometries are key to include and how the model response changes for a range of parameters that are not well-constrained, such as properties of many secondary faults.

**MMC2.4 Modeling tectonics and earthquakes: variability in loading and geometry**

- Develop 2D and 3D TEC-SEAS simulations that can resolve relevant processes across tectonic and earthquake scales, hydro-thermo-mechanical coupling, unified fault and bulk rheologies, and inertial effects or approximations thereof in a physically consistent, accurate, and computationally efficient manner;
- Employ TEC-SEAS simulations to understand the spatiotemporal variability of fault locking and slip/deformation, including tectonic, rheological, hydrological, thermal and chemical controls, and their feedback mechanisms;
- Use these simulations to constrain and predict fault, crustal, lithospheric and mantle deformation and causative mechanisms thereof, understand differences between tectonic regimes, interpret geological observations across scales, verify upscaling of heterogeneous fault zones, and understand long-term seismicity behaviour of different regions.

**MMC2.5 Frameworks for coupling models at different scales**

- Create efficient and physically meaningful interfaces for coupling modeling approaches at different scales;
- Explore the feasibility of replacing explicit modeling components with machine learning outcomes;
- Establish workflows across communities working on different scales.

**Cross-cutting themes in MMC2.1-2.5**

- Using earthquake source models for systematic evaluation of all relevant physical mechanisms at the scales of interest, establishing the relative importance of different mechanisms, and identifying justifiable simplifications;
- Coupling earthquake source models at different scales with each other and/or machine learning;
- Need for supercomputing facilities, code development with computer scientists and software engineers, and code verification.

**MMC3. Identifying relevant modeling ingredients by interpreting and improving a range of observations**

Given the wide range of coupled and evolving ingredients in the earthquake source problem (sections I.3, MMC1, MMC2), the best strategy for identifying dominating features and property ranges, and hence
formulating relevant physics-based models, is to reproduce a wide range of observations with the same physical model. Fortunately, seismological, geodetic, geological, and laboratory studies provide a broad and increasing range of observables about the behavior of the earthquake source. While a specific observation may be explained by a variety of physical processes, aiming to reproduce a range of independent observations should help discriminate between the more and less relevant model ingredients and/or constrain the relevant parameter ranges. This process will be further facilitated by progress on developing scale-appropriate constitutive laws (MMC1) that may be able to capture the coupled effects of several clearly connected mechanisms with a manageable set of parameters.

At the same time, the observational inferences themselves may require improvement, as they are often based on non-unique inverse problems or interpret data based on oversimplified models that are no longer consistent with our current understanding. Hence the goal of modeling is not only to match observational inferences but also to verify and improve them based on our best understanding of earthquake source physics.

MMC3.1 Low-heat, low-stress operation of mature faults but not the rest of the crust

One of the important - and somewhat underused - constraints on the earthquake source modeling is the accumulating observational evidence on the low-heat, low-stress operation of mature faults but not the rest of the Earth’s crust. A number of observations suggest that well-developed, mature faults (Figure II-MMCl.1-1) that correspond to tectonic boundaries and host large earthquakes are generally “weak,” i.e. operate at low overall levels of shear prestress in comparison to ~100 MPa values that would be expected based on typical static friction coefficients of 0.6-0.8 for most rocks and effective normal stress given by overburden minus hydrostatic pore pressure of about 150 MPa at the representative seismic depth of 8 km. In what is known as the “heat paradox,” the outflow of heat observed for San Andreas Fault (SAF) implies that average shear stresses acting during sliding are of the order of 10 MPa or less1; other mature faults show similar behavior2. Such low stress values are supported by inferences of steep angles between the principal stress direction and fault trace3, analyses of the fault core obtained by drilling through shallow parts of faults that have experienced major recent events, including the great 2011 Mw 9.0 Tohoku event4, the geometry of thrust-belt wedges5, and existence of long-lived narrow shear zones that do not exhibit any evidence of melting6.

Two potential explanations for such low-stress, low-heat operation of mature faults are (i) chronic weakness, either due to much lower static friction coefficients than 0.6-0.8 or much higher pore pressure or (ii) enhanced dynamic weakening as widely observed in lab experiments at high slip rates (Figure II-MMCl.1; section MMC1, Appendix B). In fact, both explanations may be correct for different fault zones, with creeping mature fault segments having low friction coefficients7, at least some portions of subduction megathrusts having elevated pore fluid pressure8 (Figure I-5), and seismogenic segments of mature strike-slip faults potentially dominated by enhanced dynamic weakening9. Furthermore, the two possibilities may be enhancing each other, with somewhat chronically weak faults, such as subduction

2 (Nankali, 2011)
3 (e.g., Townend and Zoback, 2004; Zoback et al., 2007)
4 (Kano et al., 2006; Tanikawa and Shimamoto, 2009; Fulton et al., 2013)
5 (Suppe, 2007)
6 (Chester and Logan, 1987; Chester et al., 1993; Chester and Chester, 1998)
7 (Moore and Rymer, 2007; Lockner et al., 2011)
8 (Liu and Rice, 2007; Saffer and Tobin, 2011; Shelly et al., 2006)
9 (Sibson, 1973; Rice, 2006; Noda et al., 2009; Jiang and Lapusta, 2016, 2017)
Figure II-MMC3.1-1. Potential differences between mature vs. immature faults and implications for fault physics. (a) Evolution of the off-fault damage, fault segmentation, and planarity with increasing maturity, illustrated here for the case of faults growing along the fault strike. More mature fault segments that have experienced larger cumulative slip have more planar and less segmented structure as well as narrower damage zones due to long-term healing of the damage associated with the fault creation. This difference may be responsible for the inferred low-stress, low-heat operation of mature faults in contrast to other, less mature faults (section MMC3.1). Adapted from Perrin et al. (2016). (b-c) The effect of varying levels of local roughness on fault slip compared with a planar fault. The illustrated fault profiles (b) have the mean amplitude \( h \) of deviations from a planar surface at a length scale \( x \) given by \( h = xH \) where \( H = \text{rms slope} \). For natural fault surfaces, the exponent \( H \) is typically \( 0.5 \) to \( 1.0 \) (Power and Tullis, 1991; Sagy et al, 2007). The corresponding slip on the fault (c) is only modestly affected for the lower values of roughness potentially relevant to mature faults at depth, with progressively more severe effects for increasing roughness. Adapted from Dieterich and Smith (2009). Increasing roughness brings about increasing shear resistance in the form of roughness drag (Fang and Dunham, 2013) that may dominate the response of immature faults.
megathrusts, experiencing additional dynamic weakening\textsuperscript{1}. Earthquake source modeling can make testable predictions that might distinguish between the two hypotheses for the low-stress and low-heat operation of major faults; for example, the two hypotheses may result in markedly different microseismicity activity\textsuperscript{2} (Figure II.MMC3.1-2).

At the same time, many other, smaller-scale and/or less mature active faults appear to be operating at the expected high levels of stress\textsuperscript{3}. The difference in stress levels between mature and at least some immature faults is likely related to a combination of the different levels of enhanced dynamic weakening, fault roughness, the associated near-fault elastic deformation, and its effective contribution to shear resistance\textsuperscript{4}. Investigating these coupled effects through modeling will enable constraints on the degree of dynamic weakening and roughness on faults and potentially identify the differences between properties of more and less mature faults.

**MMC3.2 Spatio-temporal patterns of seismic/aseismic slip and distributed deformation**

In recent decades, revolutionary advances in geodetic and seismological observational systems have revealed a fascinating range of seismic and aseismic fault slip behaviors, from earthquakes to steady slip with rates up to plate rates, to postseismic slip in response to earthquake-induced stress changes, to spontaneous interseismic slow slip transients, to tectonic tremor and low-frequency earthquakes\textsuperscript{5} (Figure II-MMC3.2). Furthermore, the rocks around the fault experience distributed interseismic deformation (MMC1.3). The aseismic processes on, around, and below the seismogenic faults redistribute fault loading locally, modify fault properties, and hence influence the nucleation, timing, and properties of the eventual large events. Megathrust earthquakes are a case in point where changes in plate boundary behavior have been documented including slow slip event occurrence, slow slip periodicity, geodetic multi-year transients, and crustal stress\textsuperscript{6}. There is some evidence that regions of seismic and aseismic slip overlap, including seismic slips in previously creeping areas\textsuperscript{7}, as been observed in laboratory experiments\textsuperscript{8}. Patterns of seismic and aseismic slip are also triggered by fluid injections\textsuperscript{9}, and modeling the observations of such industrial activities as well as scientific field experiments presents a unique opportunity to improve and validate modeling (section IV-IN3).

Microseismicity in the interseismic period contains further important clues regarding the stress environment, structure, and physical properties of faults and the surrounding bulk, and their changes over

\[\textsuperscript{1} (Wibberley and Shimamoto, 2005; Faulkner et al., 2011; Cubas et al., 2015)\]
\[\textsuperscript{2} (Jiang and Lapusta, 2016, 2017)\]
\[\textsuperscript{3} (Zoback, 1992; Townend and Zoback, 2000; Zoback and Townend, 2001; Chery et al., 2004; Zoback et al., 2007; Hurd and Zoback, 2012; Schoenball and Ellsworth, 2017)\]
\[\textsuperscript{4} (Dietrich and Smith, 2009; Fang and Dunham, 2013)\]
\[\textsuperscript{5} (Bürgmann et al., 2000; Dragert et al., 2001; Obara, 2002; Rogers and Dragert, 2003; Dragert et al., 2001; Pritchard and Simons, 2006; Shelly et al., 2006, 2007; Schwartz and Rokosky, 2007; Chlieh et al., 2008, 2009; Rubinstein et al., 2008, 2010; Brown et al., 2009; Peng and Gomberg, 2010; Perfettini et al., 2010; Beroza and Ide, 2011; Loveless and Meade, 2011; Kaproth and Marone, 2013; Bostock et al., 2015; Gomberg et al., 2016; Obara and Kato, 2016; Bürgmann, 2018; Hawthorne and Bartlow, 2018)\]
\[\textsuperscript{6} (Suito et al., 2011; Hirose et al., 2012; Kato et al., 2012; Yoshida et al., 2012; Mavrommatis et al., 2014; Uchida et al., 2016; Vallee et al., 2017; Becker et al., 2018; linuma, 2018; Panet et al., 2018)\]
\[\textsuperscript{7} (Bürgmann et al., 2002; Pritchard and Simons, 2006; Chen and Lapusta, 2009; Johnson et al., 2012; Ito et al., 2013; Lin et al., 2013; Noda and Lapusta, 2013; Uchida and Matsuzawa, 2013; Perfettini and Avouac, 2014; Barnhart et al., 2016; Wech and Bartlow, 2014)\]
\[\textsuperscript{8} (Leeman et al., 2016; McLaskey and Yamashita, 2017)\]
\[\textsuperscript{9} (Ellsworth, 2013; Keranen et al., 2014; Guglielmi et al., 2015; Wei et al., 2015; Weingarten et al., 2015; Elsworth et al., 2016; De Barros et al., 2018)\]
Figure II-MMC3.1-2. Microseismicity quiescence and potential deeper extent of large earthquake ruptures in fully dynamic models of mature faults with enhanced dynamic weakening. The 3D models simulate sequences of earthquakes and aseismic slip (SEAS) with full accounting for wave-mediated inertial effects during dynamic ruptures, for a planar fault embedded in an elastic half-space, assuming roughness to be relatively unimportant and subsuming inelastic processes into the on-fault constitutive relation. (a) Two models with different depth extents of enhanced dynamic weakening (red hashed region) with respect to the rate-and-state velocity-weakening region (white) produce different predictions regarding the depth extent of large dynamic ruptures and presence of microseismicity. (b) Illustration of the simulated long-term fault behavior through snapshots of fault slip rates on a logarithmic scale. (c-d) The microseismicity that exists on the bottom part of the fault without enhanced dynamic weakening (M1) disappears when the enhanced dynamic weakening is extended into that part, promoting seismic slip in large earthquakes (M2). Such absence of microseismicity is observed on many large mature strike-slip fault segments that hosted large earthquakes (e.g., Figure I-1b) and may indicate the presence of enhanced dynamic weakening on such faults and deeper rupture propagation. Adapted from Jiang and Lapusta (2016). The current developments aim to incorporate the effects of fault roughness, off-fault damage, and ductile deeper response into such simulations.
Figure II-MMC3.2 Complex patterns of seismic and aseismic slip observed on natural faults and their potential relation to heterogeneous fault friction properties. (a-b) Periods before and after the 2016 Mw 7.8 Pedernales subduction earthquake from the central Ecuador subduction zone. (a) Long-term interseismic measurements constrain the background coupling distribution on the megathrust (shown in color and defined as the difference between the observed and long-term slip rate; coupling of 100% corresponds to a fully locked fault). Several locked sections of the fault slipped in spontaneous slow slip events (thick blue contours) accompanied by seismic swarms (light blue dots) and repeating earthquakes (blue squares). The 2016 Mw 7.8 Pedernales earthquake (black contours) ruptured a locked section between 20 and 30 km depth. (b) The earthquake was followed by afterslip up- and downdip of the rupture and triggered an SSE on a coupled segment about 100 km south of the mainshock rupture (blue contours). Adapted from Rolandone et al. (2018). (c-d) Patterns of seismic and aseismic fault slip on faults may be interpreted in terms of a patchwork of velocity-strengthening (yellow) and velocity-weakening (red) fault friction properties, as schematically illustrated for a partially coupled strike-slip fault (c) and subduction zone (d). Adapted from Bürgmann (2018). Earthquake source modeling has started to link the patterns of frictional properties and observations, leading to insights about the fault zone properties (MMC2.2-2.4; Figures II-MMC2.1, II-MMC2.2, II-MMC3.1-2). The next step is to create modeling approaches that combine the potentially relevant factors (MMC1) and are capable of resolving all these stages of fault slip and their effects on each other.
time or due to large earthquakes. In addition to the fascinating phenomena of low-frequency earthquakes and tremor that often accompany episodic slow slip events, small repeating earthquakes occur on a number of faults, mostly in the presence of fault creep, and their observations have been used to study various aspects of earthquake physics and mechanics, including fault creeping velocities, postseismic slip, earthquake interaction, recurrence, and stress drops. Furthermore, some large earthquakes are preceded by foreshocks, defined as smaller seismic events that occur within a certain distance in time and space to the main event. The physical mechanisms for foreshocks as well as their potential role in the nucleation of larger events are currently open questions, although a potential explanation is that foreshocks occur on seismic fault patches loaded by the surrounding aseismic (slow) slip due to large-scale earthquake nucleation.

SEAS models (MMC2.2) based on rate-and-state friction formulations (MMC1.1) and their extensions have succeeded in modeling many aspects of these phenomena. A surprisingly wide range of observations - including slip patterns and non-trivial properties of repeating earthquakes, tremor, episodic slow slip, and foreshock-like events - can be reproduced by using the standard rate-and-state formulations, with creeping velocity-strengthening areas and stick-slipping velocity-weakening patches, supplemented with heterogeneous fault friction properties (Figure II-MMC3.2) and, in some cases, with non-uniform/perturbed pore fluid pressure and/or nonplanar fault geometry. Expanding the standard formulation with a range of fluid-related and other mechanisms supported by laboratory experiments (MMC1.1) further enhances the ability of SEAS models to reproduce various phenomena, including episodic slow slip and overlap of seismic and aseismic slip. Most of these modeling efforts do not yet include distributed inelastic deformation (MMC1.2, MMC1.3) which can contribute to or even dominate some of the inferred interseismic phenomena typically attributed to fault slip, although there are promising beginnings that combine SEAS modeling and inelastic bulk effects. Note that one aspect of most explanations for the observed slow slip transients, needed nearly in all SEAS models, is significantly elevated pore fluid pressure, close to the overburden stresses expected to be 300-400 MPa at the relevant depths, with the effective normal stresses in 1-10 MPa range; laboratory experiments under such conditions are rare. Special conditions on at least some fault regions are revealed by the sensitivity of seismicity in some cases to stress perturbations of small amplitude, including those caused by tides, dynamic waves from distant events, and seasonal stress changes.

1 (Rubin et al., 1999; Waldhauser and Ellsworth, 2000; Hardebeck and Hauksson, 2001; Waldhauser, 2001; Hardebeck and Shearer, 2003; Waldhauser and Schaff, 2008; Hauksson et al., 2012; Oye et al., 2013; Uchida and Matsuzawa, 2013; Denolle and Shearer, 2016; Ross et al., 2017; Fan and McGuire, 2018; Wetzel et al. 2018)
2 (Ellsworth and Dietz, 1990; Vidale et al., 1994; Marone et al., 1995; Nadeau and Johnson, 1998; Beeler et al., 2001; Igarashi et al., 2003; Hickman et al., 2004; Imanishi et al., 2004; Matsubara et al., 2005; Allmann and Shearer, 2007; Chen et al., 2007; Dregier et al., 2007; Zoback et al., 2010; Rubinstein et al., 2012; Abercrombie, 2014)
3 (Jones and Molnar, 1976, 1979; Abercrombie and Mori, 1996; Dodge et al., 1995, 1996; Zanzerkia et al., 2003; McGuire et al., 2005; Bouchon et al., 2011, 2013; Kato et al., 2012; Brodsky and Lay, 2014)
4 (Hori et al., 2004; Kato, 2004; Liu and Rice, 2005, 2007; Hillers et al., 2006; Dieterich, 2007; Chen and Lapusta, 2009; Chen et al., 2010; Fang et al., 2010, 2011; Kaneko et al., 2010; Wei et al., 2013; Li and Liu, 2016; Veedu and Barbot, 2016; Luo and Ampuero, 2017; Cattania and Segall, 2018; Schaal and Lapusta, 2019)
5 (Shibazaki and Ito, 2003; Rubin, 2008; Segall et al., 2010; Shibazaki et al., 2010; Segall and Bradley, 2012; Noda and Lapusta, 2013; Jiang and Lapusta, 2016; Noda et al., 2017; Heimisson and Dunham, 2018; Lui and Lapusta, 2018)
6 (Deng et al., 1999; Hetland and Hager, 2005; Fay and Humphreys 2006; Bürgmann and Dresen, 2008; Takeuchi and Fialko, 2012; van Dinther et al., 2013b; Lambert and Barbot, 2016; Lavier et al., 2013; Lindsey et al., 2014a,b; Hearn and Thatcher, 2015; Milliner et al., 2015; Erickson et al., 2017; Sobolev and Muldashev, 2017; Allison and Dunham, 2018; Tong and Lavier, 2018)
7 (Kaneko and Fialko, 2011; Lambert and Barbot, 2016; Erickson et al., 2017; Allison and Dunham, 2018; Barbot, 2018)
8 (Cochran et al., 2004; Ader et al., 2014; Johnson et al., 2015; Rubinstein et al., 2008; Ader et al., 2014; Hill and Prejean, 2014; Beeler et al., 2016, 2018; Chen et al., 2018)
Narrowing down the dominating processes and constraining fault and bulk properties would require (i) more consistent sets of experiments that, in conjunction with modeling, explore the constitutive response of faults (MMC1) in the presence of fluids under a range of potentially realistic fault conditions, including substantial fluid overpressure, and (ii) developing and deploying more comprehensive multi-physics SEAS models (MMC 2) that could explore the various coupled effects, including both fault slip and distributed inelastic deformation.

**MMC3.3 Properties of earthquake rupture events: rupture speeds, energy budget, and radiation**

Similarly, detailed imaging of earthquake rupture has been enabled by rich sets of data from seismic, geodetic, and tsunami monitoring networks and advances in data processing techniques. The combined analyses of large earthquakes now routinely includes inversions for finite-fault slip history, sources of high-frequency radiation, and the associated parameters such as rupture speed, stress drop, and various scalings\(^1\). The general picture illuminated by these efforts is that slip generally starts from a small area of the fault compared with the eventual earthquake size; propagates along the fault with a speed (known as rupture speed) which is usually a significant fraction, 0.7-0.9, of the shear wave speed of the rock bulk - although some events can break out into a speed faster than the shear wave speed, producing a supershear rupture (MMC2.1); ultimately producing average total slip values that increase with the event size. The last finding is manifested in the magnitude invariance of static stress drop, a fascinating feature that extends to smaller events (MMC3.4).

Systematic analyses have revealed additional trends that can help constrain earthquake physics and earthquake source models. The relative strength of radiated energy appears to vary for different subduction zones around the world\(^2\) \((\text{Figure II-MMC3.3-1})\), potentially revealing more and less segmented interfaces; such apparent segmentation may relate to fault geometry, friction properties, areas of overpressure, etc. Depth-dependent strength of high-frequency radiation discovered in subduction zones\(^3\) can be related to a number of factors that are expected to change with depth, including increasing overburden pressure and temperature as well as the associated changes in the fault structure, roughness, and mineralogy. At the same time, systematic analysis of the moment rate functions reveals robust scaling attributes and linear growth for event stacks in varying magnitude bins\(^4\) \((\text{Figure II-MMC3.3-2, top})\). The breakdown work, which is the generalization of fracture energy that controls crack dynamics in basic fracture theories\(^5\), is inferred to increase with average slip or, equivalently, with the event size\(^6\) \((\text{Figure II-MMC3.3-2, bottom})\). This observation can be explained either by models that result in increasing weakening with slip, such as thermal pressurization of pore fluids\(^7\), or by off-fault energy dissipation in the damage zones that is expected to also increase with the earthquake size\(^8\), or a combination.

---

\(^1\) (Ishii et al., 2005; Bock et al., 2011; Simons et al., 2011; Yue et al., 2012; Baltay et al., 2014; Avouac et al., 2015; Duputel et al., 2015; Lay, 2015; Denolle and Shearer, 2016; Melgar et al., 2016; Hayes, 2017; Ye et al., 2016ab; Chounet and Vallée, 2018)

\(^2\) (Convers and Newman, 2011; Denolle and Shearer 2016; Ye et al., 2018; Chounet and Vallée, 2018)

\(^3\) (Simons et al., 2011; Meng et al., 2011; Kiser and Ishii, 2012; Lay et al., 2012; Yao et al., 2013; Ye et al., 2016b)

\(^4\) (Meier et al., 2017; Melgar and Hayes, 2017)

\(^5\) (Ida, 1972; Palmer and Rice, 1973; Rice, 1980; Freund, 1990; Bizzari and Cocco, 2006)

\(^6\) (Abercrombie and Rice, 2005; Rice, 2006; Viesca and Garagash, 2015)

\(^7\) (Rice, 2006)

\(^8\) (Andrews, 2005; Templeton and Rice, 2008; Ma and Andrews, 2010; Dunham et al., 2011a; Gabriel et al., 2013; Nielsen et al., 2016)
Figure II-MMC3.3-1 Inferred systematic variations in seismic radiation of natural large events. (a) Global variation of the Radiated Energy Enhancement Factor (REEF) for 119 major and great interplate thrust events ($M_w \geq 7.0$) from 1990 to 2017, color-coded by the REEF value in log scale. Larger values of REEF indicate more high-frequency radiated energy. The systematic regional patterns indicate that regional fault conditions influence dynamic rupture. (b) Estimates of the ratio of high-frequency energy to the total radiated energy for the large earthquakes between 1990-2015 show an increasing relationship between the contribution of high frequency energy and average source depth. Adapted from Ye et al. (2016, 2018). This is consistent with comparisons of coseismic slip distributions from finite fault inversions and the locations of high-frequency radiators (~1 Hz) determined by seismic back projection; such studies indicate that, while the majority of the slip occurring in a number of recent great earthquakes is located at the shallower regions closer to the trench, the high frequency radiation tends to originate from the deeper extensions of the rupture region (e.g., Simons et al., 2011). Earthquake source simulations can explore the various factors - fault geometry, friction properties, areas of overpressure - that can explain such observations, while also being consistent with the long-term seismic/aseismic slip patterns (e.g., Figure II-MMC3.2).
Figure II-MMC3.3-2. Inferred similarity in moment-rate functions and increase in breakdown energy with slip for earthquake ruptures of different magnitudes. (a-b) Average moment rate evolution for 119 major and great interplate thrust events (MW ≥ 7.0) from 1990 to 2017 (inversions from Ye et al., 2018). The median moment rate functions in different magnitude bins reveal a well-defined average behavior, on both absolute (a) and normalized (b) scales, despite the high variability between individual events. The typical normalized moment rate function is approximately triangular and rather symmetric, with a peak at 35-55% of the rupture duration. Adapted from Meier et al. (2017). Such observations provide a constraint on SEAS models. (c) The inferred trend of increasing breakdown energy of dynamic rupture with slip (also called fracture energy) can be explained by enhanced dynamic weakening due to thermal pressurization of pore fluids (Rice, 2006; Viesca and Garagash, 2015). Adapted from Viesca and Garagash (2015). Another candidate mechanism for the breakdown energy increase is the creation and renewal of fractures in the damage zone. Earthquake source modeling can study the relative importance of different dissipative processes.
Multi-physics modeling with improved constitutive laws, as discussed in MMC1 and MMC2, can explore the potential alternative explanations for the observed trends in the data. Other important aspects of large earthquakes for such modeling to investigate include (i) the factors controlling when and how large earthquakes break to the surface, especially in subduction zones with their potential for tsunami generation, given the complexity of the shallow structure (Figure II-MCC1, bottom left) and a number of potentially relevant physical mechanisms including wave reflections, enhanced dynamic weakening, and bimaterial effect; and (ii) conditions for supershear rupture propagation (MMC2.1; Figure II-MMC2.1-1) that has been observed for many large strike-slip earthquakes and in the laboratory, and can generate unusual and potentially much more damaging ground motion. The modeling can also help verify the observational inferences which are based on inverse methods with often non-unique solutions and propose physics-based procedures for the inversions.

MMC3.4 Magnitude-invariant stress drops over the entire range of earthquake magnitudes

Seismological inferences indicate that static stress drops - which represent the average difference between the initial and final shear stress for the earthquake source - are magnitude-invariant over at least seventeen orders of magnitude of earthquake moment, from seismicity of Mw -8 induced in laboratory experiments to great Mw 9 natural earthquakes (Figure II-MMC3.4). For any magnitude range, the stress drops are scattered around the mean values of ~1-10 MPa. The overall scale-invariance is a fascinating characteristic of the earthquake source, given: the potential differences in properties and stress levels between mature and immature faults (MMC3.1), the multitude of physical mechanisms that contribute to shear strength and its evolution before and during dynamic rupture (MMC1 and MMC2), and different locations of earthquakes, especially microseismicity, which can occur on mapped faults, off the mapped faults, and on locked and creeping segments. There is also no systematic dependence on depth for stress drops taken as a whole, although recent observations have found some increase of stress drops with depth in certain regions, and for depths greater than the Moho. In the frictional failure regime, the expected linear increase in the effective normal stress - from the near-zero values at the surface to 300 MPa or so at the bottom of the seismogenic zones - should lead to the proportional growth in stress drops, given similar rupture processes with similar friction evolution; clearly, this is not universally observed. Hence either the effective normal stress does not systematically increase with depth, e.g., due to fluid overpressure, or its increase is accompanied by systematic changes in the stress state, fault strength and its evolution, or event dynamics, or, most likely, the interconnected combination of all these factors. Some studies inferred positive scaling of stress drop with earthquake magnitude in certain regions.

1 (Ni et al., 2005; Stein and Okal, 2005; Lay et al., 2005; Ma and Beroza, 2008; Wendt et al., 2009; Fujiwara et al., 2011; Mori et al., 2011; Ma, 2012; Mitsui et al., 2012; Kosdon and Dunham, 2013; Noda and Lapusta, 2013; Xu et al., 2015; Jiang and Simons, 2016; Gabuchian et al., 2017; Lotto et al., 2017, 2018; Sun et al., 2017)
2 (Archuleta, 1984; Olsen et al., 1997; Rosakis, 2002; Bouchon and Vallée, 2003; Dunham and Archuleta, 2004, 2005; Xia et al., 2004; Bhat et al., 2007, 2010; Rosakis et al., 2007; Liu and Lapusta, 2008; Song et al., 2008; Bizzarri et al., 2010; Konca et al., 2010; Mello et al., 2010; 2014; Vallée and Dunham, 2012; Wang et al., 2012; Passeleu et al., 2013; Yue et al., 2013; Bruhat et al., 2016; Huang et al., 2016; Shlomai and Fineberg, 2016)
3 (Dettmer et al, 2014; Minson et al., 2014ab; Mai et al., 2017)
4 (Olsen et al., 1997; Peyrat et al., 2001; Ma et al., 2008; Ruiz and Madariaga, 2011, 2013; Gallovic et al., 2018)
5 (Ide and Beroza, 2001; Abercrombie and Rice, 2005; Beroza and Kanamori, 2007; Allmann and Shearer, 2009; Boettcher et al., 2009; Cocco et al, 2016)
6 (Oth, 2013; Uchide et al., 2014,Goebel et al., 2015; Trugman and Shearer, 2017; Wu et al., 2018)
7 (Vallee, 2013)
8 (Rice, 1992; Suppe 2014; Simpson 2018)
9 (Trugman and Shearer 2017; Wu et al, 2018)
Note that the uncertainties in estimating stress drops are likely considerable, especially for smaller-sized events, due to the simplified models\(^1\) used in typical (spectral-fitting) techniques to interpret the data and other factors\(^2\). For example, the underlying models assume that the earthquake source region is circular and without directivity. Increasingly better data coverage may enable the use of methods that can capture the elongated shapes and directivity of the earthquake source, significantly reducing the epistemic uncertainties associated with those factors\(^3\). Earthquake source modeling can help verify the procedures used, by creating synthetic seismograms and applying the seismological procedures to determine the stress drops precisely known from modeling and, potentially, identify improved approaches. At the same time, the significant scatter in the seismically estimated stress drops may be at least partially due to the variability in the earthquake sources, and modeling can examine the controlling factors for such variability.

Overall, the magnitude-invariance and near-depth-independence of stress drops provide significant constraints on modeling and as well as a ray of hope for the possibility of narrowing down - or coalescing - the multitude of potentially important physical processes to a few essential dependencies with a manageable number of parameters, while also reproducing the rich diversity of behaviors discussed in MMC3.1, MMC3.2, and MMC3.4. Challenges ahead are to establish, through modeling, which combinations of physical processes are consistent with the observed invariance of stress drops; improve observational constraints on stress drops; and clarify whether the depth dependence of stress drops observed in some studies is the rule or the exception.

Figure II-MMC3.4. Magnitude invariance of inferred stress drops over at least seventeen orders of earthquake moment magnitudes, from seismicity of Mw 8 induced in laboratory experiments to great Mw 9 natural earthquakes, with the median value of ~3 MPa. While most of the stress drops fall within the band of 1-10 MPa, many inferences are scattered outside that range. Adapted from Cocco et al. (2016). The magnitude invariance of stress drops is an important constraint on earthquake source models. The origin of the variability, including epistemic uncertainty (Kaneko and Shearer 2014, 2015; Wang and Day, 2017; Lin and Lapusta, 2018), can be investigated through earthquake source models.

\(^1\) (Brune, 1970; Madariaga, 1976)
\(^2\) (Noda et al., 2013, Kaneko and Shearer 2014, 2015, Adams et al., 2016; Ross and Ben Zion, 2016; Wang and Day, 2017; Lin and Lapusta, 2018)
\(^3\) (McGuire and Kaneko, 2018)
MMC3.5 Gutenberg-Richter law, Omori’s law for aftershocks, and other statistics on regional scales

Seismicity can be characterized by two fundamental statistical distributions: the frequency-magnitude distribution, describing the relative number of small and large earthquakes, and the interevent-time distribution, describing how events are spaced in time. One of the most well-known results in statistical seismology is the distribution of earthquake magnitudes, known as the Gutenberg-Richter law\(^1\), which states that earthquake magnitudes are distributed exponentially, with the number of earthquakes with magnitude at least equal to \( M \) given by \( \log N(M) = a - bM \), where the parameter \( a \) is related to the total earthquake rate, while \( b \) determines the relative number of small and large earthquakes. The interevent times between earthquakes on regional scales are dominated by clustering, the most prominent example of which are aftershock sequences with a power-law temporal decay known as Omori’s law\(^2\), or Poissonian (random) occurrence of the remaining, “background” seismicity\(^3\). Studies based on paleoseismic data indicate that quasi-periodic earthquakes of moderate and large magnitudes (Figure II-MMC3.5) may occur on certain types of faults, such as isolated, geometrically simple continental faults or oceanic strike-slip faults\(^4\); some of such conclusions have been disputed on a statistical basis\(^5\), although some of the data have been subjected to statistical tests\(^6\). On a related note, there is evidence that individual fault segments experience a higher fraction of “system-size” events (i.e. earthquakes that rupture the entire segment) than would be predicted by the Gutenberg-Richter scaling\(^7\), even though these claims have also been questioned\(^8\).

The earliest well-resolved SEAS simulations produced characteristic, near-periodic earthquake sequences, likely due to the simplicity and homogeneity of their seismogenic regions as well as their small size compared to the nucleation size\(^9\). Simulations with larger faults\(^10\) and introduction of fault roughness and complex fault geometries\(^11\) significantly complexify the response. Models of individual fault zones with strong heterogeneities within the continuum class\(^12\) and inherently-discrete frameworks\(^13\) reproduce various statistical properties of seismicity. In addition, earthquake simulators\(^14\) reproduce the observed statistics over regional fault networks, although their results may be influenced by their oversized numerical discretization that leads to numerically-controlled complexity even in simple models\(^15\). Further development of SEAS models to include realistic features on a single fault segment (MMC2.2) and to study how the behavior scales with increasingly more complex and realistic fault networks (MMC2.3, MMC2.4) will be essential to make continuing progress. Given that all potentially relevant fault and bulk processes (MMC1) cannot be resolved in sufficient detail at the scale of fault networks, progress in using smaller-scale models to develop scale-appropriate constitutive relations for

---

\(^1\) (Gutenberg and Richter, 1944)  
\(^2\) (Utsu et al. 1995)  
\(^3\) (Gardner and Knopoff 1974; van Stiphout et al., 2012; Saichev and Sornette, 2007)  
\(^4\) (McGuire 2008; Parsons 2008; Scharer et al. 2010; Berryman et al., 2012; Cochran et al., 2017; Howarth et al., 2018)  
\(^5\) (Kagan et al., 2012; Geller et al., 2015)  
\(^6\) (e.g., Scharer et al. 2010)  
\(^7\) (Schwartz and Coppersmith 1984; Wesnousky 1994; Parsons et. al., 2018)  
\(^8\) (Page et al., 2011; Morgan and Felzer, 2015)  
\(^9\) (Tse and Rice, 1986; Rice, 1993)  
\(^10\) (Shaw and Rice, 2000; Wu and Chen, 2014)  
\(^11\) (Fang and Dunham 2013; Tal and Hager, 2018)  
\(^12\) (Hillers et al., 2006, 2007)  
\(^13\) (Ben-Zion and Rice, 1993; Ben-Zion, 1996; Ben-Zion et al., 2003; Zöller et al., 2005, 2007)  
\(^14\) (Dieterich and Richards-Dinger, 2010; Pollitz, 2012; Richards-Dinger and Dieterich, 2012; Tullis et al. 2012a,b; Sachs et al., 2012; Ward, 2012; Shaw et al., 2018)  
\(^15\) (Rice, 1993; Lapusta et al., 2000)
Figure II-MMC3.5. Quasi-periodicity of large earthquakes on some mature strike-slip faults based on geologic (paleoseismic) records with statistical tests. (a-b) Estimates of recurrence intervals for (a) a 8000-year record of 24 surface-rupturing earthquakes on the Alpine Fault in New Zealand, shown as probability distributions, and (b) a 3000-year combined record of 29 surface-rupturing earthquakes for the San Andreas Fault in California, with the mean (square) and standard deviation. The vertical grey lines indicate the time since the last event. (c) Recurrence patterns of large earthquakes on three major transform faults: closest to periodic (with a coefficient of variation of 0.33) on the Alpine fault; quasi-periodic, based on statistical tests, on the San Andreas Fault; and random to clustered on the Dead Sea Transform in the Middle East (Marco et al., 1996; Kerry et al., 2011). Adapted from Berryman et al. (2012) and Scharer et al. (2010). While seismicity as a whole exhibits either clustering or random temporal behaviors, the largest earthquakes on some mature strike-slip faults have more regular, quasi-periodic behaviors, based on paleoseismic data which is the only source of long-term records of large earthquakes. Earthquake source simulations can explore the potential physical reasons behind the special nature of such faults, including their level of maturity (section MMC3.1), isolation, and long-term slip rate.
simulations at the fault-network scale becomes key. These advances can be combined with the idea of coupling models at different scales, including the use of machine learning (MMC2.5).

Two questions are of particular interest. The first one is the identification of the dominating processes for earthquake interaction, as in part expressed in Omori’s aftershock decay law, which can be explored by earthquake source simulations (MMC2). The second one is the properties of the largest events in the system and whether they could in fact be special, e.g. quasi-periodic, or different from the overall statistics in other ways. There are a number of physical reasons for the largest events to be special, due to their potential preferential occurrence on plate-boundary mature faults with low-heat, low-stress operation unlike other faults in the crust (MMC3.1).

For more detailed descriptions of MMC3.1-3.5, refer to Section V in Appendix B.

Key future goals

MMC3. Identifying relevant modeling ingredients by interpreting and improving a range of observations

MMC3.1 Low-heat, low-stress operation of mature faults but not the rest of the crust
- Employ SEAS models (MMC2) to investigate constraints that the observed low-heat, low-shear stress operation of mature faults - together with levels of microseismicity and other observations - puts on fault physical properties including pore pressure, potential enhanced dynamic weakening of the friction resistance, as well as degree of roughness and heterogeneity;
- Investigate the differences between the mature and immature faults in that regard that can be tested through observations;
- Use the results to inform dynamic rupture simulations and investigate the potential impact on strong ground motion.

MMC3.2 Spatio-temporal patterns of seismic/aseismic slip and distributed deformation
- Determine the dominating mechanisms among several potential explanations for the various observed patterns of seismic/aseismic slip and distributed deformation, and the associated range of fault properties, by using multi-physics SEAS and TEC-SEAS models (MMC2.2, MMC2.4) with the improved constitutive relations for fault slip and off-fault inelastic deformation (MMC1);
- Constrain models of faulting in the presence of fluids by laboratory experiments as well as by instrumenting and modeling the ongoing industrial activities that involve fluid injections, which are essentially meso-scale field experiments (see also III.IN3);
- Explore the implications of the determined fault properties on the nucleation, propagation, and arrest of large earthquake ruptures; in particular, (i) determine the conditions under which large earthquakes have identifiable precursors and/or can be triggered by slow slip transients, (ii) constrain the depth extent of large earthquake ruptures, and (iii) establish if/when earthquake ruptures can propagate through creeping regions.

MMC3.3 Properties of earthquake rupture events: rupture speeds, energy budget, and radiation
- Use SEAS models constrained by other observations as well as dynamic rupture models with the initial conditions informed by the SEAS models to investigate the origin of the observed properties of large events, including enhanced high frequency radiation from depth, systematic radiation patterns over the subduction zones, and increasing breakdown energy in larger events;
- Use the modeling to investigate the conditions for large earthquakes to break to the surface, especially in subduction zones, and to become supershear;
Develop procedures for checking and improving finite-fault inversions based on earthquake source simulations, including dynamic rupture inversions.

**MMC3.4 Magnitude-invariant stress drops over the entire range of earthquake sizes**
- Improve observational constraints on stress drops, in part, by improving the earthquake source models used; use earthquake source simulations to determine how much of the scatter could be physical and what the observed scatter tells us about fault heterogeneity;
- Systematically investigate the dependence of stress drops on depth with improved data and improved, simulation-informed methods, to clarify whether there is indeed no systematic depth dependence in most cases, or whether the overall data hides important local dependencies as suggested by some recent studies;
- Determine the combinations of fault physical mechanisms and properties - many of which are found to be scale- and depth-dependent - that result in the invariance of stress drop over the entire range of earthquake magnitudes and potentially with depth; a key challenge is to formulate multi-physics SEAS models that can reproduce a wide range of earthquake sizes in a numerically tractable way.

**MMC3.5 Gutenberg-Richter law, Omori’s law for aftershocks, and other statistics on regional scales**
- Use SEAS modeling on single and several faults to study various proposed interaction mechanisms of earthquake rupture events (MMC2.3) and determine dominating mechanisms that need to be included in regional-scale simulations;
- Use SEAS models developed for several segments as well as improved earthquake simulators developed for fault networks (MMC2.3) to investigate the properties of the largest events that occur on mature plate-boundary faults with low-heat, low-stress operation (MMC3.1) and whether they deviate from the overall properties of seismicity, e.g., by being quasi-periodic;
- Investigate the differences between the event statistics on mature vs. immature faults.

**Cross-cutting themes in MMC3.1-3.5**
- Using multiple types of observations to constrain multi-physics modeling; developing simplified representations of coupled physical processes that are consistent with observations, for use in larger-scale models;
- Elucidating similarities and differences in the structure and properties between mature vs. immature faults as well as continental vs. megathrust environments;
- Using modeling outcomes to verify and improve methods employed for observational inferences which are often based on highly simplified models.

**MMC4. Exploring potential future behaviors: exploiting opportunities and identifying limitations**

Sections MMC1-MMC3 describe the current efforts and future progress needed towards interpreting a full range of earthquake source observations in terms of physics-based models built on our best current understanding of relevant mechanisms and ingredients gained from laboratory, field, and theoretical studies.

The response of the developed models can then be interrogated for the full range of observationally and lab-constrained ingredients, providing insight into (i) the range of potential future behaviors, (ii) what additional knowledge can further constrain that range, and (iii) in the case of energy production and other
human activity, strategies to minimize the induced seismic hazard. The usefulness of such exploration depends on a number of factors, which include not only the range of available region-specific field data but also how successful the modeling effort is in narrowing down the relevant, scale-appropriate constitutive relations for the modeling in question. Hence a key challenge is to identify a tractable set of physical parameters that can be used to reproduce and forecast fault behavior on the regional scale. Once that set is established, one can use inverse and forward modeling to constrain these parameters from seismic, geodetic, and geologic observations.

Several opportunities for rapid progress in exploring potential future behaviors of the earthquake source are outlined below.

**MMC4.1 Predicting the source component of strong ground motions**

Dynamic rupture simulations (MMC2.1) have advanced to the point where small-scale geometric fault complexities can be introduced and properly resolved (Figure II-MMC4.1), rather than parameterized. For example, within the past decade, major insight has been gained into the important effects of fault surface roughness, described by self-similar power-law random fields, on both rupture propagation and the radiated waves. The resulting ground motion, simulated in media with statistical descriptions of small-scale velocity and density perturbations, is consistent with observed statistical variability\(^1\) as well as features of high-frequency excitation\(^2\). In addition, simulations have shown that the geometrical roughness of faults in various ways promotes - on deeper parts of the fault - and demotes - near the free surface - supershear rupture propagation\(^3\), creates frequency-dependent radiation patterns\(^4\), and produces power-law co-seismic surface slip fluctuations\(^5\). Similarly, simulations can now explicitly resolve small-scale heterogeneities in off-fault elastic properties that cause scattering of radiated waves\(^6\).

Ultimately, dynamic rupture simulations could replace, or at least further inform, kinematic (or ‘pseudo-dynamic’) rupture representations currently used in ground motion assessment\(^7\), which are based on more ad hoc rules to generate slip histories with correlations between kinematic source parameters (e.g., slip, rise time, rupture velocity). More work is needed to define workflows for introducing stochastic short-wavelength model components (stress, geometry, etc.) that result in ruptures having realistic ground motions. These workflows must be validated by comparison to ground motion data or perhaps empirically derived ground motion models that distill ground motion time series into intensity measures of engineering interest (e.g., spectral acceleration and peak ground velocity).

**MMC4.2 Improving our understanding of potential extreme events**

Paleoseismic investigations provide an essential and widely used source of information on past large seismic events\(^8\). Yet many great earthquakes of the last decades have revealed unexpected features. The great 2004 Mw 9.2 Sumatra-Andaman earthquake broke multiple lateral segments of the Sunda

---

\(^1\) (Withers et al., 2018)

\(^2\) (e.g., Dunham et al., 2011b; Shi and Day, 2013)

\(^3\) (Bruhat et al., 2016; Yao, 2017)

\(^4\) (e.g., Cho and Dunham, 2010; Wang et al., 2014)

\(^5\) (Yao, 2017)

\(^6\) (Bydlon and Dunham, 2015; Mai et al., 2017; Withers et al., 2018)

\(^7\) (Schmedes et al., 2013; Graves and Pitarka, 2016; Mai et al, 2017)

\(^8\) (Grant and Sieh, 1994; Dolan et al, 1995; Rubin et al., 1998; Field et al., 1999; Goldfinger et al., 2003; Klinger et al., 2003; Weldon et al., 2005; Kumar et al., 2006; Sieh et al., 2008; Hamiel et al., 2009; Scharer et al., 2010; Rubin et al., 2017)
II-MMC4.1 Simulations of dynamic rupture and ground motion from a 3D rough fault. (a) Model geometry and dimensions of a strike-slip rough fault, with self-similar fractal distribution ranging from 80 m to 80 km. (b-c) Snapshots of fault parallel velocity (m/s) at the surface using a 1D-layered model without (b) and with (c) small-scale media heterogeneity with a correlation length of 150 m. Adapted from Withers et al. (2018). Such simulations for a suite of model parameters can provide physics-based constraints on high-frequency components of strong ground motion.

Megathrust in a giant rupture, including segments that were thought to be aseismic\(^1\). During another great event, 2011 Mw 9.0 Tohoku earthquake in Japan, the near-trench area, commonly believed to be aseismic\(^2\), slipped 50 m or more\(^3\) and generated a monstrous, unexpectedly devastating tsunami\(^4\). The 2012 Mw 8.6 Indian Ocean earthquake was the largest recorded earthquake on strike-slip faults and it ruptured fracture zones in a diffuse plate boundary that were thought to be seismically inactive\(^5\). These events could not be anticipated based on the historical record, as the latter depicts only an incomplete time-limited portrait of the full range of possible seismic behaviors.

The examples show the potential extreme nature of earthquake source behaviors and pose the question whether similarly unexpected large events can occur in other regions such as Pacific Northwest or California in the United States. Let us consider the 2011 Mw 9.0 Tohoku earthquake in Japan, where its unexpected and unprecedented near-trench slip of 50-80 m came from the area that may have slipped

\(^1\) (Ni et al., 2005; Stein and Okal, 2005; Lay et al., 2005)
\(^2\) (Hyndman et al., 1997)
\(^3\) (Fujiwara et al., 2011; Simons et al., 2011)
\(^4\) (Mori et al., 2011)
\(^5\) (Meng et al., 2012)
stably in the past. Such bimodal behavior is not uncommon\(^1\) (MMC3.2). While many factors influence rupture propagation in the shallow portions of a megathrust\(^2\), one of the potential explanations consistent with other features of the regional behavior is enhanced dynamic weakening due to thermal pressurization of pore fluids\(^3\). If that is indeed the right answer - and more studies are needed to determine this - then earthquake source modeling can be used to study where else can this physical mechanism may allow creeping fault regions\(^4\) - which are currently considered more stable and safe - to participate in destructive events and produce large seismic slip. After large events, such creeping fault regions might remain locked for a while, but eventually they would accumulate enough stress to start creeping again, obscuring the evidence of their violent past. For example, does the creeping section of the San Andreas Fault in California provide a barrier between the southern and northern locked parts of the fault, as often assumed? Or can it rupture due to enhanced dynamic weakening, thereby producing a massive earthquake that would affect two large metropolitan areas, Los Angeles and San Francisco? How about the shallow portions of the Cascadia subduction zone in the Pacific Northwest? How deep can earthquake rupture propagate into the nominally stable deeper fault roots?

Earthquake source models, based on the best available knowledge from the laboratory and regional field data, can start to explore such questions by creating a range of physically admissible fault models and producing a range of outcomes for simulations with and without extreme scenarios, some of which may be testable with the available observations. At the very least, such modeling can determine the ranges of fault properties that could lead to extreme behaviors vs. not, and identify the knowledge gap that can then be filled, by either targeted remote observations (section IV.1) or drilling through the fault (section IV.2).

**MMC4.3 Capitalizing on fault-specific patterns of seismic and aseismic slip**

Some natural fault segments seem to exhibit relatively regular sequences of seismic events. One clear example is repeating earthquake sequences which tend to occur on creeping faults (section MMC3.2) and can be successfully simulated with the current SEAS models\(^5\). Paleoseismic evidence indicates that at least some large, mature faults, such as the Southern San Andreas Fault (SAF)\(^6\) and the Alpine fault in New Zealand\(^7\) also host relatively regular, perhaps quasi-periodic, earthquake ruptures (Figure II/MMC3.5). Such areas are a prime target for physics-based modeling. Consider the Parkfield segment of SAF that produces repeated Mw 6.0 events\(^8\). One of the models\(^9\) - based on the standard rate-and-state friction and planar fault, with some friction heterogeneity - can reproduce a number of the observed features, including the average recurrence interval, the switch in the nucleation locations from one edge of the fault to another, and the interseismic and postseismic records of the surface GPS stations. However, simulated variations in the recurrence interval of the Mw 6.0 events are significantly narrower than observed, suggesting that the model needs additional mechanisms to diversify the response: either more on-fault heterogeneity, or enhanced dynamic weakening, or nonplanar, rough fault geometry, or spatially variable off-fault damage, or a combination of them (MMC1). The 3D multi-physics SEAS approaches, currently under development (MMC2.2), will be able - in the near future - to investigate how

---

1. Bürgmann et al., 2002; Pritchard and Simons, 2006; Lin et al., 2013; Barnhart et al., 2016; Wech and Bartlow, 2014
2. Ma and Beroza, 2008; Kozdon and Dunham, 2013; Hubbard et al., 2016; Gabuchian et al., 2017; Lotto et al., 2017, 2018
3. Faulkner et al., 2011; Mitsui et al., 2012; Noda and Lapusta, 2013; Cubas et al., 2015
4. Harris, 2017
5. Kato, 2004, 2016; Dublanchet et al., 2013; Veedu and Barbot, 2016; Kaneko et al., 2017; Cattania and Segall, 2018; Lui and Lapusta, 2018
6. Scharer et al., 2010
7. Berryman et al., 2012; Cochran et al., 2017; Howarth et al., 2018
8. Bakun et al., 2005
9. Barbot et al., 2012
these factors affect the recurrence time of the Mw 6.0 events and make predictions about other observables that can be used to distinguish between the models, such as the average shear stress on the fault which should be low (MMC3.1) or the rates of occurrence of Mw 5 and smaller events.

The quasi-periodic response of some fault areas but not others (MMC 3.5) highlights the need to distinguish between different types of faults and tectonic environments. One potential distinction is between mature and immature faults, with mature faults being low-stress, more dynamically weakening, more localized, and/or less rough, with less off-fault damage (section MMC3.1). Another potential distinction is between the subduction-zone megathrusts and plate-boundary strike-slip faults; potential differences include more fluids in subduction zones, especially at depth where they are liberated by dehydration reactions; different mineralogy of faults; and complex structure of the hanging wall. Ranges of fault properties that lead to more vs. less regular behaviors can be investigated with earthquake source modeling.

Another type of area-specific phenomena that presents an opportunity for predictive modeling is a “seismic gap”, defined as a fault area known to produce significant earthquake ruptures that has not slipped in a long time compared to the nearby segments. Seismic gaps have been identified around the world at major plate boundaries\(^1\) based on a combination of historical seismic proxies (tsunamis deposits, turbidites, sag ponds), instrumented earthquakes, and paleoseismic events. These gaps may rupture in a piecemeal fashion, as in the Mentawai seismic gap in Sumatra\(^2\), or in a wholesale event involving multiple adjacent seismic gaps as in the 2004 Sumatra megaquake\(^3\).

It is generally understood that any perceived segmentation may not be permanent, but the result of a complex interaction of other factors, including the history of previous ruptures and their stress footprint\(^4\), and the dynamic interactions between fault segments\(^5\). Yet the very identification of the seismic gap implies a well-studied region, with earthquake ruptures and/or slow slip identified around the gap - or region of inaction over some period of time - and with some additional data, such as paleoseismic, that identifies the gap as potentially seismogenic (Figure II-MMC4.3). Long-term SEAS and TEC-SEAS simulations (MMC2.2, MMC2.4) can then attempt to match all these observations, to identify factors that lead to the gap not rupturing for a while as well as its potential future behaviors.

**MCC4.4 Quantifying uncertainty and identifying key gaps in knowledge**

Uncertainties clearly exist, both in observational and laboratory inferences and in the model outcomes. For instance, estimating subsurface fault slip evolution in space and time is typically an ill-conditioned inverse problem. Optimization-based inversion approaches commonly aim to find a preferred model subject to the balance of data fits and model features such as the spatial smoothness of fault slip, based on considerations of numerical solvability rather than physics. The a priori imposed nonphysical model constraints interact with the inferred models, modifying the outcome. Community inversion exercises and platforms have provided a clear demonstration and some measures of the variability of earthquake source models\(^6\). Given the progress in computational capabilities, producing ensembles of all plausible models that reasonably reproduce data with minimal non-physical model assumptions, e.g., in a

\(^1\) (McCann et al., 1979; Singh et al., 1981; Van Dissen and Berryman, 1996; Goldfinger et al., 2003; Cummins, 2007; Sieh et al., 2008; Heidarzadeh and Kijko, 2011; Hyodo and Hori, 2013; Protti et al., 2014)

\(^2\) (Konca et al., 2008; Salman et al., 2017)

\(^3\) (Stein and Okal, 2005)

\(^4\) (Kaneko et al., 2010; Barbot et al., 2012; Qiu et al., 2016; Michel et al., 2017)

\(^5\) (Harris and Day, 1999; Harris et al., 2002; Duan and Oglesby, 2006; Elliott et al., 2009)

\(^6\) (Mai and Thingbaijam, 2014; Mai et al., 2016)
II-MMC4.3 Assimilation of geological and geophysical data in SEAS simulations to understand the behavior of “seismic gaps”, i.e. fault areas known to produce significant earthquake ruptures that has not slipped in a long time compared to nearby segments. The numerical simulations (b) of sequences of earthquakes and aseismic slip (SEAS) at the Kathmandu, Nepal seismic gap (a) assimilate morphological gradients from interpretation of geological data (Hubbard et al., 2016) and produce sequences of partial and full ruptures of the Main Himalayan Thrust. Partial rupture events have characteristics and surface displacements similar to the 2015 Mw 7.8 Gorkha, Nepal earthquake. Adapted from Qiu et al. (2016).
probabilistic framework\textsuperscript{1}, becomes increasingly feasible and desirable. Quantifying the full uncertainty in the inferred earthquake characteristics is especially important for studying fault areas with limited observational constraints, such as the subduction zone forearc and the base of the seismogenic zone\textsuperscript{2}.

Better approaches to characterizing epistemic uncertainties of inferences based on inverse methods warrant more robust interpretation of geophysical signals. A common source of such uncertainties is the use of oversimplified forward models in inverting and interpreting data. Subsurface fault geometry and structure, for example, are often simplified ingredients in slip imaging. Resolving detailed earthquake source features through inversions will require the use of more realistic models of 3D fault geometry and heterogeneous structure, for both seismic and geodetic data\textsuperscript{3}. In addition to improving forward models, one can characterize and incorporate model prediction uncertainty by considering plausible errors in the modeling ingredients\textsuperscript{4}, so that the inferred finite-fault models reflect epistemic uncertainties, avoid overfitting data, and can be properly compared with earthquake source modeling.

Similarly, in earthquake source simulations, a major effort is needed to characterize and quantify how the variations in key model ingredients affect model predictions. Mutually exclusive model assumptions may explain limited data equally well. That is why modelers must explore the implications of their models beyond the available data so that their models can be put to the test in retrospect. Rather than determining a single preferred set of physical properties that enable the simulation to fit observations of interest, it would be quite informative to explore the range of physically plausible sets of physical properties and determine a range of models that fit observations. A good practice for exploring model uncertainties is to construct diagrams of the sensitivity of the model behavior on key parameters\textsuperscript{5} (Figure II-MM4.4). Then the collective future behaviors of the model, with all identified parameter sets, will provide some measure of uncertainty, given the type of model, and can be used to determine which properties need to be constrained the most. New techniques in machine learning or ensemble data assimilation may also help us to extract information on critical model states or parameters and facilitate the process of exploring model variability in numerical simulations\textsuperscript{6}. Understanding epistemic uncertainties would be particularly important when using models for hazard estimation. Overall, exploring uncertainty is essential to capturing a range of quantitatively and even qualitatively different future scenarios.

The synergy of field and laboratory observations and earthquake source modeling holds great promise for future progress. Better characterization of uncertainty in the observations and their interpretations encourages a comprehensive exploration of plausible earthquake source models. Reconciling earthquake source simulations with observations helps identify key gaps in our knowledge and motivate new observational techniques. Accelerating collection of high-quality data from existing and new data networks, more sophisticated data analysis and inversion methods, continuing insightful laboratory experimentation, and a coherent suite of more powerful earthquake source modeling tools would further reduce the epistemic uncertainties and transform our understanding of earthquake source processes.

\textsuperscript{1} (Minson et al., 2014a,b)
\textsuperscript{2} (e.g., Jiang and Simons, 2016; Yue et al., 2017)
\textsuperscript{3} (Hayes et al., 2018; Hsu et al., 2011; Tung and Masterlark, 2016)
\textsuperscript{4} (e.g., Duputel et al., 2014; Ragon et al., 2018)
\textsuperscript{5} (Kaneko et al., 2013)
\textsuperscript{6} (e.g., DeVries et al., 2018; van Dinther et al., 2019)
II-MMC4.4 Exploring uncertainty of earthquake source simulations by quantifying model sensitivity to key parameters. The geodetically inferred surface velocity and shallow fault creep along the central section of the North Anatolian Fault (NAF) is interpreted in terms of rate-and-state fault friction parameters using numerical simulations of SEAS. (a-b) Dependence of a root-mean-square (RMS) misfit (colors) between simulated surface velocity and InSAR data (a) on the friction properties assumed in the model and the size of the shallow velocity-strengthening layer (b). A dashed curve encircles models with comparably small RMS misfits. (c–e) Three specific examples of the simulated surface velocity, which are compared with the InSAR surface velocity. Adapted from Kaneko et al. (2013). Such explorations of all reasonable parameter values that allow models to fit observations can be used to obtain a range of suitable parameters and quantify the model uncertainty.
III. Catalyzing community initiatives (IN)

IN1. Community modeling ecosystem

As mentioned in section I, the main goal – and value added – of physics-based earthquake source modeling is to interpret a full range of field observations – seismic, geodetic, thermal and others – in terms of models with physically meaningful fault and bulk properties that can be evaluated, at least in principle, through lab, field, and smaller-scale numerical studies and hence can be updated as new knowledge is developed. Such models can then be used for predictive modeling and discovery. The earthquake source community is well underway towards this challenging goal as discussed in MMC1-MMC4.

Two sustained interrelated community modeling efforts will facilitate rapid future progress.

**Modeling collaboration to capture multi-scale constitutive response of fault structures**

A joint community effort is needed towards developing a systematic set of modeling tools, coupled with lab experiments and field studies, to conquer the multi-scale nature of the fault and rock constitutive response discussed in section MMC1.

The spatial scales of the underlying fault and bulk structure ranges from microns in the fault core to kilometers of nonplanar fault stretches (section I.3, Figure I-3). Its overall constitutive behavior is potentially significantly influenced by the weakest link - which is the shear strength of the localized shear zones - but intermixed with the effects of fault roughness on all scales, near-fault damage, and the associated pore fluid effects. The resulting constitutive response is needed for a wide range of time scales, from sub-second fault breakdown to thousands of years of tectonic loading and creep, and the related wide range of strain rates. What are the appropriate constitutive fault relations to adopt for simulations of crustal-scale fault networks having spatial discretization of the order of tens to hundreds of meters? Can we use the expressions developed from centimeter-scale laboratory samples directly and if so, then in what special cases?

We need to develop a set of modeling tools to interrogate the response of faults and surrounding rock materials on various relevant temporal and spatial scales, identify dominant physical mechanisms that can be scale-dependent, and formulate simplified joint constitutive representations of various coupled effects for use in larger-scale models. The most successful currently used constitutive relations for faulting (rate-and-state friction, damage rheologies, power-law creep) incorporate a measure of deformation rate (slip rate, strain rate) and other relevant state variables. The challenge then is identifying the relevant set of state variables and the smallest set of physically meaningful parameters that can be used to capture the response of a smaller-scale system (Figure II-MMC1.5) on a larger scale. In fact, whether such an upscaling is meaningful in all cases is an open fundamental question.

The components of this effort would include:

- Modeling at the scales of laboratory samples of traditional experiments, with features from microns to tens of millimeters, to untangle various coupled effects at the scale of localized shear layers for friction laws, bulk fracturing and healing for damage formulations, and bulk creeping processes in multiphase materials for viscoplastic response;
- Laboratory experiments under realistic fault conditions that are designed to be amenable to insightful modeling;
Modeling earthquake source processes: from tectonics to dynamic rupture

- Modeling at the scales of tens of millimeters to tens of meters, to connect the constitutive response at the typical laboratory scales to the smallest discretization sizes assumed in best-resolved large-scale models;
- Input from field scientists on the structure and properties of faults and bulk on such intermediate scales;
- Analytical theories on how to extract a robust larger-scale constitutive response from such efforts.

Some of the modeling tools created for larger-scale simulations can be adapted for this purpose. A number of current efforts are directed towards capturing a range of physical mechanisms at the crustal scale, and the numerical challenges involved will perhaps be more manageable at the meso-scale discussed here. Validation of the numerical approaches using larger-scale experimental setups (IV-IN5) would be especially fruitful in this regard.

The challenging issue of appropriately capturing smaller-scale processes in larger-scale models is a science-wide problem of significant current interest in other areas of geosciences as well as in materials science, applied mathematics, and various branches of engineering. So some approaches (section V-MMC1) - and perhaps experts - can be borrowed from those fields.

**Modeling environment with open-source tools, software optimization, and supercomputing access**

The existing and developing earthquake source modeling approaches (section MMC1, MMC2) are being actively used to reproduce observations (MMC3), investigate competing views on earthquake mechanics, and reveal unexpected features (MMC4). To facilitate this process of discovery and integration of the vast range of laboratory and field data, it is important to create and maintain a modeling environment with open-source tools, software optimization, and supercomputing access. Such an environment can also promote the integration of capabilities of different approaches, best modeling practices, as well as code verification exercises that have been quite successful in both making sure the codes work as advertised as well as pushing the modeling community forward to tackle new challenges.

Openly accessible community codes enable a wider range of scientists to address their research priorities through modeling as well as facilitates development of better combined capabilities. Examples of successful community codes include SpecFEM, PyLith, and SeisSol. However, releasing and maintaining software as an open-source project has to be accompanied by its optimization, documentation, installation support, tutorials, and usage examples. This requires a level of programming support that most geoscientists, even modelers, do not have. It is important to establish sustained sources of support for such modelling environment building efforts.

The use of the existing and newly developed earthquake source modeling codes at the societal level of interest, especially in 3D, relies on accessible and efficient use of supercomputing. For example, tremendous progress has been made recently in explicitly modelling physics-based scenario earthquakes using 3-D simulation methodologies. Observational constraints, e.g., models of fault geometry and variations in bulk wave speeds, can now be routinely included, and tailored geosoftware and emerging “modeling engines” (e.g. ExaHyPE) enable efficient use of modern high-performance infrastructure. With

---

1 (Harris et al., 2018)
2 (Komatitsch and Tromp, 1999; Aagaard et al., 2013; Uphoff et al., 2017)
3 (e.g., Galvez et al., 2014, Lozos et al., 2015, Withers et al., 2018a, Wollherr et al., 2018, Ando and Kaneko, 2018; Ulrich et al., 2018, 2019)
4 (e.g. Pegasus Workflow, Deelmann et al., 2006; Asagi Geoinformation Server, Rettenberger et al., 2016)
5 (Rannabauer et al., 2018)
the performance of recently developed earthquake modeling tools reaching multiple $10^{15}$ floating point operations per second (PFLOPS), the calculation of a multitude of scenarios is now feasible. These capabilities have been developed due to links between earthquake source modelers and scientists from scientific computing, applied mathematics, and engineering\(^1\), which allowed to address key challenges related to computational science, such as efficient local time-stepping schemes suitable for modern hardware, dynamic load balancing, asynchronous input and output overlapping with computation, adaptive mesh refinement, and others.

Future progress will rely on establishing routine interaction with computational scientists to allow the earthquake source codes to efficiently utilize novel hardware architectures as well as on building community resources of open-source modeling tools, processed observations, validation exercises, precompiled models for urgent computing, and full “modeling setups” (as established, for example, in the NEMO project for ocean modeling\(^2\)).

**IN2. Machine learning and data mining for modeling and observations**

A better understanding of complex earthquake behavior will be enabled through improved seismological, geodetic, and geologic observations from data sets that are increasingly rich in content. Progress in understanding will also come through modeling earthquake behavior in simulations that are increasingly realistic. The techniques of machine learning and data mining from the rapidly evolving field of data science have begun to play an important role in these efforts, and that role is sure to expand dramatically in the near future.

The impact of data mining on seismology has been transformative. Through template matching – the systematic search of continuous seismic data for matches to known earthquake waveform signatures\(^3\) – seismologists have determined that tremor is comprised of the superposition of low frequency earthquakes\(^4\), highlighted the time-dependent effects of static and dynamic triggering on seismicity\(^5\), documented the relationship between mainshock slip and aftershock activity\(^6\), and disentangled the effects of afterslip and pore fluid diffusion on the spatio-temporal evolution of aftershock activity\(^7\). Uninformed similarity search, in which the template waveform signature is not known a priori, has also proven effective and led to improved understanding of seismicity\(^8\).

The impact of machine learning has the potential to be at least as profound\(^9\). Seismology has large, labeled existing data sets that are ideal for training deep neural networks (Figure III-IN2). Machine learning has already proven effective in data denoising\(^10\) as well as earthquake detection, location, and characterization\(^11\). The result of these efforts will be an end-to-end workflow from input waveforms to output earthquake catalogs that are far more complete and feature much more reliable source characteristics. The examples above are from seismology, but with sufficient data, similar benefits should be realizable from other (geologic and geodetic) measurements of earthquake phenomena as well.

---

1. Cui et al., 2013; Heinecke et al., 2014; Ichimura et al., 2015; Roten et al., 2016; Uphoff et al., 2017
2. [www.nemo-ocean.eu](http://www.nemo-ocean.eu)
3. (Gibbons and Ringdal, 2006)
4. (Shelly et al., 2007)
5. (Meng and Peng, 2014)
6. (Huang et al., 2017)
7. (Ross et al., 2017)
8. (Aguiar and Beroza, 2014; Yoon et al., 2015; Skoumal et al, 2016, 2018; Yoon et al., 2017)
9. (Kong et al., 2018)
10. (Zhu et al., 2019)
11. (Mousavi et al., 2016; 2019; Perol et al., 2018; Zhang et al., 2018; Ross et al., 2018, 2019; Zhu and Beroza 2019)
Figure III-IN2. Applications of machine learning techniques to earthquake source data and modeling. (a) Illustration of a convolutional network training on seismic waveform data in order to develop a generalized seismic phase detection algorithm. Unlike state-of-the-art data mining approaches based on template matching, such algorithms can detect events without pre-existing templates. Adapted from Ross et al. (2018). (b-d) Performance of an artificial neural network (ANN) in reproducing the deformation pattern for a viscoelastic case. (b) Viscoelastic code solution for a point source at 6 km depth for a Maxwell rheology of $10^{19}$ Pa·s at $t = 100$ years after an earthquake. (c) ANN prediction for comparison. (d) Residuals between the true viscoelastic solution and the ANN prediction; note that they are small. Once trained, the ANN runs more than 4 orders of magnitude faster than the full visco-elastic code. Adapted from DeVries et al. (2017).

Machine learning can play an equally important role in addressing some of the challenges in earthquake source modeling identified in this report. For example, a trained deep neural network can effectively model the visco-elastic response to imposed stress changes thousands of times faster than the full physics-based calculation\(^1\) (Figure III-IN2). Machine learning has been used to predict the timing of laboratory earthquakes\(^2\), infer laboratory friction from tremor-like signals\(^3\), compare slow and fast earthquakes\(^4\), and explore the relationship between a mainshock and its aftershocks\(^5\).

In earthquake source modeling, machine learning can potentially be used to replace explicit physics-based modeling at certain scales with inferences of trained neural networks (section MMC2.5). One of the most challenging problems in earthquake science is its multi-scale nature (section I.3) and the need to appropriately capture the response of smaller scales at the larger scales of interest. This challenge may be addressed through the operation of deep learning techniques on forward simulations, to encapsulate hidden variables that are representative of small scales and propagate them to the larger scales. Recent progress in the engineering mechanics community in applying neural networks to emulate material

---

1. DeVries et al., 2017
2. Rouet-Leduc et al., 2017
3. Hulbert et al., 2018
4. Rouet-Leduc et al., 2018
5. DeVries et al., 2018
response with history dependence\textsuperscript{1} suggests that this may be a promising endeavor. For example, neural networks, trained on an appropriate set of simulations, may be able to eliminate the need for explicitly formulated scale-appropriate constitutive relations discussed in MMC1.5, by directly establishing a mapping between the characteristic inputs, such as slip/strain rate, temperature etc, and outcomes, such as the corresponding shear resistance or damage evolution. In a similar fashion, machine learning techniques may be able to replace certain phases of large-scale earthquake source simulations, such as determining the post-seismic visco-elastic response of the deeper layers below the seismogenic faults or replacing explicit dynamic rupture modeling in a long-term simulation of a fault network (MMC2.5).

Developing appropriate techniques for such approaches would require collaborations with computer scientists and applied mathematicians, as well as suitable educational programs for geoscience students. These efforts can be promoted by targeted funding programs.

IN3. Modeling energy-harvesting activities as ongoing meso-scale field experiments, link to industry

The need for more and cleaner energy is driving a wide range of activities involving pumping fluids in or out of the subsurface. These activities include geothermal energy production, CO\textsubscript{2} sequestration, gas storage, conventional and non-conventional oil and gas production, and industrial wastewater disposal (Figure I-4). These activities impart pore pressure and stress changes and can result in shear faulting or hydraulic fracturing. The management of such operations and the mitigation of associated hazards, induced earthquakes in particular, has been a major challenge\textsuperscript{2}. Addressing this challenge is particularly important in the context of public pressure to reduce CO\textsubscript{2} emissions.

Meeting the 2°C target of the International Panel on Climate Change would indeed require a very substantial effort of CO\textsubscript{2} sequestration and storage as well as resorting, at a large scale, to carbon-neutral sources of energy such as geothermal energy. Induced seismicity is an unfortunate impediment to the development of geothermal energy production outside the limited areas of natural hydrothermal activity. To harvest heat efficiently from low permeability crystalline basement, the permeability needs to be enhanced through hydraulic stimulation or fracturing. In the context of Enhanced Geothermal System (EGS), induced seismicity is thus unwanted on larger scales but needed on smaller scales\textsuperscript{3}. Today, hydraulic stimulation of crystalline basement for EGS reaches depths exceeding 6 km, well into the range where large earthquakes nucleate\textsuperscript{4}. In addition, the energy demand during the transition to a low-carbon economy will most likely require a sustained production of gas, in particular from unconventional reservoirs, as it is the fossil fuel with the least carbon impact. This method of production requires hydraulic fracturing of the source rocks along horizontal wells and produces large volumes of water, which need to be disposed. These activities would benefit greatly from improved models of earthquake source processes induced by deep well injection or production.

It has long been recognized that injection and extraction of fluids into and out of the subsurface can cause earthquakes\textsuperscript{5}. Examples of significant man-made earthquakes have become abundant. Most recently, a magnitude 5.4 earthquake in South Korea in 2017 was potentially caused by hydraulic stimulation for

\textsuperscript{1} (Wang and Sun, 2017, 2018a,b)
\textsuperscript{2} (Majer et al., 2007; Zoback and Gorelick, 2012; Ellsworth, 2013; Keranen et al., 2014; Weingarten et al., 2015; Elsworth et al., 2016; Nakai et al., 2017)
\textsuperscript{3} (Majer et al., 2007)
\textsuperscript{4} (Kwiatek et al., 2018)
\textsuperscript{5} (e.g., Evans, 1966; Rothé, 1970; Segall et al., 1994; Ellsworth, 2013; Bourne et al., 2014)
geothermal heat production\(^1\). In Basel, Switzerland, earthquakes with magnitude up to 3.4 were triggered in 2006 during hydraulic stimulation for an Enhanced Geothermal System project\(^2\). The earthquake caused insignificant damage but public concern led to the immediate project suspension. In the central U.S., where seismicity is naturally low, seismic activity significantly increased recently (Figure I-4), peaking in 2015 to a level higher than in California, mostly driven by massive wastewater disposal over a broad region\(^3\).

Our ability to control or forecast induced seismicity is currently limited. Improvements would require significantly advancing our understanding of fluid effects in faulting in general (MMC1.4) and in relation to induced seismicity in particular, and advanced earthquake source modeling of the type discussed in this report (section MMC2).

At the same time, every Enhanced Geothermal System project - and other similar activities - is a potential field experiment, which can provide a wealth of information about: the state and heterogeneity of stress in the crust; the hydro-mechanical coupling and distribution of poroelastic properties; and the respective role of pore pressure diffusion, poro-elastic stress, and slow slip (Figure III-IN3) in triggering earthquakes.

Multiple types of data can be acquired, using enhanced monitoring of microseismicity with automatized machine learning algorithms\(^4\); new sensing technologies that allow dense coverage such as MEMs or fiber optics (section IV.1); and remote sensing, in particular inSAR and GNSS to monitor surface deformation and infer pressure changes and deformation of the subsurface\(^5\). These data sets can be supplemented by in-situ and laboratory experiments\(^6\) designed to investigate the coupling between fluid flow and fault slip, leading to enhanced modeling capabilities for coupling fluid flow with rock deformation and fault slip constrained by these observations.

IN4. Validation by predictive modeling of a controlled large-scale field experiment

Earthquake science has historically been handicapped by a lack of controlled experiments. We glean what information we can from natural, unpredictable occurrences that are seldom properly instrumented and result in data that often has fundamental gaps. In particular, questions related to aseismic earthquake initiation processes as well as damage and heat generation during dynamic rupture propagation would benefit from instrumentation being placed prior to an anticipated event to record the entire proceedings.

The current surge in human-induced seismicity has made the problem even more pressing\(^7\) (section III-IN3). Industrial activities are producing earthquakes in seismically unprecedented numbers. The situation has resulted in one of the largest, unintended human interventions on the natural world in history with surprisingly little controlled, scientific experimentation. While monitoring these ongoing activities could significantly advance our science, the proprietary nature of industrial data and the competing economic interests make public study of the current situation challenging. The knowledge gap is creating a perilous situation where scientists are being called upon to advise without an adequate foundation.

---

\(^1\) (Grigoli et al., 2018; Kim et al., 2018)  
\(^2\) (Deichmann and Giardini, 2009)  
\(^3\) (Elsworth et al., 2016)  
\(^4\) (Mousavi et al., 2016; 2019; Perol et al., 2018; Zhang et al., 2018; Ross et al., 2018, 2019; Zhu and Beroza 2019)  
\(^5\) (Fialko and Simons, 2000; Jha et al., 2015; Shirzaei et al., 2016)  
\(^6\) (Guglielmi et al., 2015)  
\(^7\) (Ellsworth, 2013; Rubinstein and Mahani, 2015)
Figure III-IN3. Potential mechanisms for induced seismicity due to anthropogenic fluid injections, including initially aseismic slip and poroelastic effects. (a-b) Underground fluid injection experiments on natural faults indicate that pore pressure increase (blue line in (b)) induces aseismic slip on the fault first (black line in (b)), followed by microseismic activity off the fault (red line in (b)). The slip was measured in-situ by a specially designed displacement sensor (a). The experiment took place at a depth of 282 m below the earth’s surface, in cretaceous limestone of the southeast France sedimentary basin. Adapted from Guglielmi et al. (2015). (c-d) The evolving view of how fluid injection can make earthquakes. Historically (c), injection was recognized to activate faults in the small region directly connected to the added water, shown in blue. More recently (d), an additional halo of squeezed rock, shown in red, surrounds the pressurized fluid due to poroelastic effects and can activate more distant faults. Adapted from Goebel and Brodsky (2018).
The most ambitious previous earthquake controlled experiment was in Rangely, Colorado, nearly 50 years ago\(^1\). The experiment approximated the solid Earth as a rigid body and only considered Coulomb failure. As a result, there was no geodetic instrumentation and no consideration of poroelasticity, which is now recognized as a major component of the induced earthquake process\(^2\) (MMC 1.4, Figure III-IN3). In the intervening half-century, experiments have focused on the permeability enhancement potential of water injection and its accompanying seismicity\(^3\), or small-scale slow slip in response to fluid injection\(^4\) but have not usually been designed to attack the basic science problems of earthquake initiation and propagation.

The recent SEISMS workshop outlined a plan for an active earthquake experiment\(^5\) and identified some potential locations. The community consensus was that an active earthquake experiment is required to address three major goals of the field: (1) a predictive understanding of the processes that produce human-induced seismicity, (2) knowledge of the sequence of events leading to earthquake nucleation, and (3) understanding of the controls on the magnitude of earthquakes. Successfully producing a moderate-sized earthquake by fluid injection, as opposed to slow slip or a cluster of microseismicity, will be a feat on its own, and would require a fault-monitoring campaign and physics-based modeling. The resulting dynamic event will allow the community to validate our models of dynamic rupture (MMC2.1), evolving off-fault damage (MMC1.2), and rupture interaction with fluids (MMC1.4). Achieving these goals will require full utilization of recent advances in field and laboratory instrumentation (sections IV.1-IV.3) including distributed acoustic systems, satellite-based geodetic methods combined with borehole instrumentation, and testing of recovered fault rock.

IN5. Validation by modeling well-instrumented laboratory experiments

When detailed measurements are made, laboratory experiments can be used as a quantitative benchmark to test the assumptions of numerical models. While laboratory experiments form the basis of many of the constitutive relationships described previously (MMC1) and can serve as a qualitative guide or source of ideas and inspiration for modeling studies, some laboratory experiments produce results that can be used to test earthquake source models. These experiments—typically conducted on larger and more compliant samples—can be compared to models of earthquake nucleation, dynamic rupture propagation and termination, and slow creeping fronts. A well instrumented experiment at this larger scale is a useful stepping stone between smaller-scale lab experiments, used to constrain friction constitutive equations, and larger scale earthquakes occurring on natural faults. Furthermore, the combination of numerical modeling and laboratory experiments can act synergistically. The laboratory experiments can provide physical constraints for the numerical models, while the modeling facilitates researchers to better understand, generalize, and scale up the experimental results.

The experiments suitable for such exercises are typically conducted either on meter-scale rock samples\(^6\) (III-IN5) or on analog materials composed of glassy polymers\(^7\). Conducting well-controlled experiments in much larger rock samples is extremely difficult due to the sample weight, issues with alignment, and loading. The polymers have elastic moduli that are about 30 times lower than rock, and all instability scales are proportional to them. As a result, dynamic rupture processes that would require tens of meters

---

\(^1\) (Raleigh et al., 1976)
\(^2\) (Segall, 1989; Segall and Lu, 2015; Goebel et al., 2017; Goebel and Brodsky, 2018)
\(^3\) (Doetsch et al., 2018; Dorbach et al., 2009; Eaton et al., 2018; Kneafsey et al. 2018; Rutledge and Phillips, 2003)
\(^4\) (Guglielmi et al., 2015)
\(^5\) (Savage et al., 2017)
\(^6\) (Dieterich, 1981b; McLaskey et al., 2014; 2015; Yamashita et al. 2015)
\(^7\) (Rosakis, 2002; Rubinstein et al., 2004; Xia et al, 2004; Rubino et al., 2017)
of rock to develop can be observed on sub-meter polymer samples. The complexity of such experiments requires either arrays of point sensors\(^1\) or a full-field imaging technique\(^2\). The meter-scale rock experiments have been pioneered at the USGS, Menlo Park (Figure III-IN5) and inspired more recent large-scale rock experiments at NIED Japan\(^3\) and Cornell University\(^4\). Laboratory experiments of this type, as well as similar ones conducted on smaller plastic samples, have been used as a tool to study modes of dynamic rupture propagation\(^5\), earthquake initiation processes\(^6\), and rupture arrest\(^7\). Studies of stick-slip cycles on decimeter scale samples in the laboratory have compared and contrasted the spatio-temporal variations of acoustic emissions with the patterns that are sometimes seen throughout the seismic cycle for natural earthquakes, for example in terms of roughness or stress/time-dependence of the frequency-magnitude relationship\(^8\).

Modelling and laboratory experiments can work synergistically. Laboratory experiments can be an effective way of validating certain aspects of numerical modeling, but numerical modeling can also create additional insights to the laboratory experiments. Examples of the synergy in combining experimental and numerical approaches include the measurement of the radiation from dynamic ruptures\(^9\), understanding of nucleation processes\(^10\), arrest of slip\(^11\), thrust rupture interaction with the free surface\(^12\), and confirmation of rate-and-state effects and enhanced dynamic weakening acting in concert to create dynamic friction evolution during spontaneous rupture propagation\(^13\).

Analog polymer-based experimental setups and 3D printed materials\(^14\) can allow to recreate in the precisely desired and quantified fault geometries interacting with multiple physical processes, to study and validate the existing approaches to their modeling their coupled effects. Such setups has already been used to study the effects of branched geometries and damage on rupture propagation\(^15\) and are being extended to the studies of fluid injection\(^16\) and rock gouge shear layers. The ability to do precise measurements for the kinematic field as well as monitoring the rupture propagation will lead to unprecedented opportunities in understanding frictional resistance, off-fault damage, rupture characteristics, and the resulting features of seismic radiation.

Key questions that can be addressed by combined larger-scale laboratory experiments and numerical models include: How is dynamic rupture stopped? Candidates include barriers set up by geometrical heterogeneity (jogs, bumps, step-overs), rheological heterogeneity (fault sections with velocity strengthening frictional behavior), or by stress distribution (fault sections with low stress that are unfavorable for rupture propagation). How do earthquakes initiate? Theoretical and numerical studies

---

1. Okubo and Dieterich, 1984; Ohnaka and Kuwahara, 1990; Yamashita et al., 2018
2. (e.g. Rosakis et al., 2006; Nielsen et al., 2010; Rubino et al., 2017)
3. (Fukuyama et al., 2014)
4. (Ke et al., 2018)
5. (e.g. Xia et al., 2004, Lu et al., 2007; 2010a)
6. (e.g. Okubo and Dieterich, 1984; Ohnaka and Kuwahara, 1990; Latour et al., 2013; McLaskey and Kilgore, 2013; McLaskey et al., 2014; 2015; Schaal and Lapusta, 2019)
7. (Ke et al., 2018)
8. (e.g. Goebel et al., 2013; Riviere et al., 2018)
9. (Svetlizky et al, 2016)
10. (Ben-David et al., 2010; Kaneko et al., 2016)
11. (Kammer et al., 2015, Bayart et al., 2016)
12. (Gabuchian et al., 2017)
13. (Rubino et al., 2017)
14. (e.g. Hanaor et al., 2016; Vanorio and Kanitpanyacharoen, 2015)
15. (Biegel et al., 2010; Bhat et al., 2010)
16. (Gori et al., 2019)
Figure III-IN5. Validation of earthquake source modeling through well-instrumented larger-scale laboratory experiments. (a-c) While most laboratory experiments on rocks operate on 1-10 cm samples to study realistic fault conditions, some experimental setups feature well-instrumented meter-scale samples (2-m granite sample is shown) that allow for slow applied loading and measure the induced slow fault slip (c) as well as record multiple dynamic stick-slip events (insert in (a)). The interpretation of all available measurements (b) shows an initial zone of expanding aseismic slip, with intervening small-scale dynamic events (circles), which eventually grows into dynamic rupture (dashed regions), illustrating the concept of aseismic earthquake nucleation. Adapted from McLaskey and Kilgore (2013), McLaskey et al. (2014). SEAS models based on rate-and-state friction show qualitatively similar features (Ampuero and Rubin, 2008; Schaal and Lapusta, 2019). (d-e) Laboratory earthquake experiments on analog materials (d) capitalize on their much smaller instability scales to reproduce various dynamic rupture phenomena, including supershear rupture propagation, self-healing slip pulses, propagation along bimaterial interfaces and others (Rosakis et al., 2007; Rubino et al., 2017; Gori et al., 2018); such experiments are currently being developed to include aseismic nucleation due to fluid injection (Gori et al., 2019). Digital image correlation techniques (DIC) allow to produce full-field evolving maps of displacements, strains, and stresses (e). Adapted from Rubino et al. (2017). Both types of experiments can be used to validate SEAS and dynamic rupture models.
have shown that nucleation is sensitive to details of the friction constitutive formulation\(^1\), loading rate\(^2\), and heterogeneity\(^3\). How do aseismic and seismic slip interact? Slow fronts and aseismic creep have been observed in some laboratory experiments\(^4\) and recent models of the interactions between aseismic slip and seismicity (sections MMC2.2, 3.2) show that this can be a promising collaboration. How can pore fluids in fault zones affect and trigger earthquakes? Larger-scale, well instrumented laboratory experiments in conjunction with modeling can shed light on the interactions between subsurface fluids and earthquakes.

IN6. Physics-based input into early warning algorithms

An Earthquake Early Warning (EEW) system for the west coast of the US has been in development since 2006 and has recently gone online\(^5\). EEW algorithms detect and characterize ongoing ruptures in real-time and provide alerts of impending ground motion at target sites before they are reached by strong ground motion (Figure III-IN6). Depending mainly on earthquake magnitude, distance to the target site, and tectonic situation, such systems can provide seconds to tens of seconds of warning time before strong shaking. The current EEW algorithms have mostly been developed and calibrated using the existing seismic and geodetic data sets from tectonically active regions. Owing to the relative scarcity of observations for the most crucial cases - namely short-distance observations of large earthquake events - future improvements of EEW systems may rely heavily on modeled synthetic data sets at all stages, from event detection and characterization all the way to ground motion prediction and alerting.

Input from physics-based earthquake source modeling has potential to significantly improve the quality for EEW alerts through i) improved rupture predictability, ii) more accurate and precise real-time ground motion predictions, and iii) infrastructure upgrades, such as modeling-informed expansions of observational networks.

Empirical studies of rupture growth suggest that early hopes of strong magnitude predictability of earthquake rupture\(^6\) may have been overly optimistic. It does not seem to be possible to use rupture onset observations to predict that final size of large-magnitude events. Instead, on average, small and large earthquakes seem to follow a universal growth behavior and differ only in the sense that smaller earthquakes stop rupturing sooner\(^7\) (Figure III-IN6). This sets an upper bound for possible warning times that EEW systems can provide.

The warning times could be improved, however, if rupture predictability could be increased, e.g. by considering regional differences in rupture evolution patterns, or by considering fault maturity models\(^8\) (Figures II-MMC3.1-1, II-MMC3.5, sections MMC3.1, MMC3.5, MMC4.3). On a mature and hence relatively smooth fault, an ongoing rupture is potentially less likely to stop. If a rupture initiation of a certain size is detected on a known mature fault, it may therefore make sense to follow a different alerting strategy then when an initiation is discovered where no mature faults are present. Quantifying the effect of rupture smoothness and fault segmentation on the probabilities of rupture propagation is challenging, however, and real-world observations are limited to a small number of recent large earthquakes and to a body of paleoseismic observations. Exploiting the information contained in these data sets will require

---

1 (Ampuero and Rubin, 2008)
2 (Kaneko and Lapusta, 2008)
3 (Ray and Viesca, 2017)
4 (Rubinstein et al., 2004)
5 (Kohler et al., 2018)
6 (e.g. Olson and Allen, 2005)
7 (Meier et al, 2016)
8 (e.g., Böse and Heaton, 2010, Perrin et al., 2016)
modeling efforts with state-of-the-art dynamic earthquake source simulations that can accurately describe rupture behavior on non-planar faults with realistic structures (sections MMC2, MMC4.3).

Alternatively, rupture predictability could potentially be increased by leveraging prior information, e.g. from pre-event seismicity rates\(^1\) or from real-time geodetic observations on accelerating aseismic deformation. Machine learning models, furthermore, may be able to identify diagnostic signal attributes that may enhance rupture predictability, as was recently reported in a laboratory earthquake study\(^2\). Deep learning algorithms could be trained to determine alerting levels directly from large synthetic rupture data sets, such as the CyberShake compilation\(^3\).

The timeliness and correctness of EEW alerts could further be improved by more accurate and precise ground motion prediction models (section MMC4.1). The research domain of ground motion prediction is increasingly moving from simple empirical studies to sophisticated full-waveform modeling experiments that take into account both rupture complexity and structural complexity along the propagation path and near the surface\(^4\). Such efforts promise to help reducing the epistemic uncertainty in ground motion prediction, which would significantly improve the quality of EEW alerts, by lowering the rates of strong ground motion over- and under-predictions.

**IN7. Multidisciplinary summer schools**

Multidisciplinary collaborative efforts in both research and education are crucial to tackle complex grand challenges of earthquake source processes that are described in previous sections. To facilitate discussion and collaboration between different disciplines and foster efforts to educate a greater number of students from a broader range of disciplines and more diverse perspectives, community summer schools are necessary. The target disciplines include geophysics, geology, rock mechanics, computer science, data science, applied mathematics, materials science, physics, and engineering. The multidisciplinary summer schools will be designed to bring different disciplines together and discuss thematic grand challenges as well as disciplinary advances. They will provide opportunities for early-career scientists to interact and communicate with each other as well as with more senior colleagues in their and other disciplines. Each summer school will explore specific topics in earthquake source processes by integrating lectures, training courses, tutorials, projects, and fieldwork, depending on the topics. The primary target participants of the summer schools will be graduate students and postdocs, but researchers at all levels from early-career to more senior levels will join as lecturers and facilitators. Workshops before or after the summer school may be also held to discuss specific topics or technical developments. Great examples of summer schools are Cooperative Institute for Dynamic Earth Research (CIDER) summer program, series of SCEC-ERI-DPRI International Summer School on Earthquake Science, and Earthquakes summer school in Cargese, France. Regular, biannual summer schools in the U.S. would expand interactions across different disciplines to promote innovative multidisciplinary collaboration in earthquake source processes.

\(^1\) (e.g., Yin et al., 2018)  
\(^2\) (Rouet-Leduc et al., 2017)  
\(^3\) (Graves et al., 2011)  
\(^4\) (Rodgers et al., 2018; Withers et al., 2018)
Figure III-IN6. Earthquake Early Warning (EEW) recently implemented on the west coast of the United States. (a) Warnings rely on fast detection of an earthquake’s pressure (P) waves, which travel faster than the more damaging shear (S) waves. After using P waves to estimate the earthquake size and location, the system relays electronic warnings ahead of the oncoming S waves. Adapted from Voosen (2018). (b) The warning time would be longer if potentially large, damaging earthquakes could be identified based on short initial recordings. Unfortunately, observations show that small and large earthquakes seem to follow a universal growth behavior (illustrated here by median evolution of peak absolute ground displacement recorded at short distances (<25km) from earthquakes of a wide range of magnitudes and differ only in the sense that smaller earthquakes stop rupturing sooner. Adapted from Meier et al. (2016). Physics-based earthquake source simulations may suggest more refined strategies based, for example, on whether earthquake initiation is detected on mature vs. immature faults.
IV. Support for related community and disciplinary efforts

IV.1 Observational networks: Seismic, space-based, borehole, fiber-optics, and sea-floor

Earthquake source processes occur at a wide range of spatial and time scales (section I). Improving in-situ measurements of earthquakes is critical to understand the physics, but the observations are typically made remotely, in the far field. Instrumentation, and thus observation, tends to be band-limited and spatially aliased. Historically, measurements of slow, quasi-static deformation were made using space-based (GPS and satellite radar imaging) technologies with continuous spatial coverage but poor temporal resolution, while dynamic measurements were made using seismic networks with good temporal sampling but relatively poor spatial resolution. Today and in the immediate future, the combination of observational networks, including novel solutions, is increasingly able to encompass the broad temporal and spatial scales of earthquake source phenomena.

Permanent observatories
Near-field measurement of large earthquakes, for instance those expected to occur on shallow onshore fault systems, allow to infer rupture velocity, rupture width, some aspects of stress evolution and energy budget, and near-source inelastic response. The growing capabilities and continuous recordings of permanent seismic networks, such as the ones maintained by the California Integrated Seismic Networks, the National Research Institute for Earth Science and Disaster Resilience (NIED, Japan) (Figure IV-1), and Geonet (New Zealand), enable continuous monitoring of seismic activity and detailed recordings of large earthquakes\(^1\). GPS networks and repeated satellite imagery have enabled inferences of inter-, co-, and post-seismic deformation\(^2\). Repeated optical imagery has quantified the style and amount of inelastic off-fault deformation\(^3\). The resulting data and its inferences are crucial to understanding earthquake source processes (section II-MMC3) and these efforts should actively continue.

Large-N seismic observatories
Dense instrumentation of earthquake regions (Figure IV-1) has become feasible with the development of affordable short-period geophones, referred to as nodes, that contain sensor, digitizer, and battery. These nodes have become popular in academic research thanks to the collaborations between academic and industry partners\(^4\) and their low cost. They have allowed for the dense temporary instrumentations of fault zones, with impressive resolution of tens of meters in some cases, compared with kilometers of tens of kilometers for more traditional networks. Such dense measurements have been use to reveal a detailed image of the near-source anisotropic elastic and inelastic structure\(^5\), fault-zone seismicity\(^6\), and presence of seismicity in the lower crust\(^7\). Recently, a community initiative lead to the deployment of over 390 instruments (nodes, short period, broadbands) to monitor the induced seismic activity in Oklahoma\(^8\). Such efforts should be expanded in the future, with some installations made into permanent fault observatories, as they have the potential to capture unique information about the earthquake source invaluable to earthquake source modeling, including dynamics of microseismicity that reveals fault structures and potential changes of their properties; potential preparation processes of earthquakes.

---

\(^1\) (Bouchon et al., 2000; Kaneko et al., 2017; Ross et al., 2017; Sun et al., 2017)
\(^2\) (Ruiz et al., 2014; Yamagiwa et al., 2015)
\(^3\) (Milliner et al., 2015; Klinger et al., 2018)
\(^4\) (Chang et al., 2013; Hollis et al, 2013)
\(^5\) (Ben-Zion et al., 2015; Hillers et al., 2016)
\(^6\) (Meng and Ben-Zion, 2017)
\(^7\) (Inbal et al., 2015)
\(^8\) (Sweet et al., 2018)
Figure IV-1. Future advances in earthquake source modeling require data both traditional and novel observational networks, such as dense rupture observatories, Distributed Acoustic Sensing (DAS), and seafloor geodesy. (a) The development of affordable and highly portable instruments has allowed for the deployment of temporary dense networks for detailed imaging and provided the potential for quick deployment immediately following significant events, as illustrated by the instrumentation of the San Jacinto fault using 1000 nodes with 10-30 meter spacing. Such networks can be developed into permanent rupture observatories. Adapted from Ben-Zion et al. (2015). (b) Widespread tens-of-kilometer long telecommunication fiber networks around the world, especially in urban areas and on ocean floors, can be converted into dense seismic arrays with meter spacing using DAS. For example, 20 km (red line) of the existing fiber cables in the JPL Deep Space Network near Goldstone, CA (red and black lines) have be turned into over 4000 sensors. Adapted from Yu et al. (2019). (c) Permanent seismic, geodetic, and other networks provide invaluable continuous monitoring. New developments include seafloor-geodesy networks of unprecedented sizes that can provide novel insights on both seismic and aseismic fault deformation near trenches of subduction zones, which are currently least constrained.
illuminated either through microseismicity changes or temporal changes in volume properties; and, during large dynamic ruptures, the ability to infer the evolution of fault break-down at the rupture tip and other propagation details that cannot be deciphered from stand-alone measurements of traditional seismic networks.

Broad-scope observational networks
In the laboratory, a range of instruments, including strain meters, thermometers, seismometers, measure various coupled fields, allowing for in-depth fundamental understanding of the underlying processes. Such comprehensive approach to measuring various fields should be replicated in natural observatories. An example is the WISSARD array located on the Whillans Ice Plain, which combines surface GPS and seismometers, borehole seismometers-geophones, and distributed temperature probes. Ice Streams such as Whillans are areas of fast moving ice atop bedrock and Whillans is one that slips seismically as modulated by tides, with relatively short recurrence intervals that allow for repeated measurements. Access to the near-source region is possible by drilling through the ice. In another example, the Rustrel experiment is an in-situ laboratory, 300 meters underground, whereby monitored injection of fluids triggered seismic and aseismic slip on a pre-existing and reactivated fault in France. The hydromechanical and seismological measurements showed unprecedented and direct evidence of an aseismic motion preceding seismic failures, providing unique field constraints on earthquake course models; the results have been successfully modeled using rate-and-state friction laws (II-MMC1.1) and hydromechanical coupling.

Distributed Acoustic Sensing
Distributed Acoustic Sensing (DAS) is an emerging technology that converts every meter of an optical fiber into a seismic strain sensor, by transmitting a laser pulse through the fiber and tracking the movement of glass defects through interferometry of back-scattered lights. In the last few years, several pilot DAS networks, such as the Stanford Array using the fiber network on the Stanford campus, the Goldstone Array using the fibers deployed by JPL, the Pasadena Array using fibers owned by the City of Pasadena, and a pilot experiment in the SAFOD borehole all delivered high quality seismic data at 1-10m spacing. Further expansion of such networks will allow sampling regional seismic wavefields at high spatial resolution and provide invaluable constraints on rupture processes of intermediate-magnitude earthquakes within the network and block-by-block strong ground motion distribution in urban areas. DAS networks will also increase our earthquake detection capability, and allow to map and monitor faults in urban areas that may pose great threats. The comparatively low installation and operational cost of extended fiber networks, as compared to seismometers or accelerometers, would make it possible to instrument large faults, such as the San Andreas Fault in southern California, with an array of multiple fibers running parallel and across the fault.

Sea-floor geodetic observatory
Seafloor geodesy has been fundamental in determining the locking and unlocking mechanisms of fault slip at the trench of subduction zone megathrusts. Seafloor displacements are otherwise inferred through

---

1 (McLaskey et al., 2014; 2015; Rubino et al., 2017)
2 (Bindschadler et al., 2013; Winberry et al., 2013; Lipovsky and Dunham, 2016)
3 (De Barros et al., 2018)
4 (Guglielmi et al., 2015)
5 (Martin and Biondi, 2018)
6 (Ellsworth et al., 2018)
7 (Lindsey et al., 2017; Yu et al., 2019)
8 (Li and Zhan, 2018)
9 (Bürgmann and Chadwell, 2014; Yokota et al., 2016)
tsunami modeling, which is a non-unique and regularized inversion\(^1\). Repeated seafloor bathymetry measurements have enabled the validation of the amount an shallow slip near the trench of about 62 m during the M9.0 2011 Tohoku-oki earthquake\(^2\). Additionally, seafloor bathymetry provides ground truth discriminant between fault slip and submarine landslide as the source of an observed tsunami. New technologies are enabling cost-effective long-term seafloor horizontal GPS measurement. A successful example of geodetic and seismic offshore observatory is that of the Guerrero Seismic Gap, Mexico\(^3\) and other dense networks are planned (Figure IV-1). In the US, earthquake-prone off-shore regions of Cascadia and Alaska would benefit from such observatories.

### IV.2 Laboratory experiments for developing constitutive laws and discovery

Laboratory experimental studies have allowed us to elucidate a number of physico-chemical processes related to the earthquake source – including fault friction (MMC1.1), damage evolution (MMC1.2), deeper ductile behavior (MMC1.3), and fluid effects (MMC1.4) - as well as to explore earthquake source dynamics including earthquake nucleation and dynamic rupture propagation (III-IN5). By characterizing the deformation and faulting processes of rocks at controlled conditions simulating in-situ stress, temperature, strain rate, strain, pore fluid pressure, and pore fluid chemistry, experimental findings motivated the development of rate-and-state friction, damage theories, and power-law creep formulations, which have been widely used for modeling field observations (MMC2, MMC3). However, key challenges remain (MMC1) and advanced experimental studies will play an indispensable role in addressing them.

Yet, a complete description of the physics of fault slip and off-fault deformation is still lacking (MMC1). For slip rates slow compared to the seismic regime, the experimentally developed rate-and-state constitutive formulations have been constrained for many fault materials and partially justified through fundamental considerations (section V-MMC1.1 in Appendix B). While the rate-and-state framework - of using the rate of deformation and additional physically meaningful state variables to capture the experimental response - is likely robust, the exact functional forms of the formulations and their parameters are not yet known due to lack of complete underlying theory and discovery of additional processes relevant for natural faults. Revealing the micro-mechanisms that control different types of deformation is necessary to establish the physics-based formulations for the temperature, pore fluid pressure, and strain-rate effects. New techniques using X-ray source and synchrotron beamline and improved microscale imaging and accurate stress measurements on small samples at extreme conditions will provide the deformation mechanisms that can be used to build these formulations\(^4\); currently, these techniques are largely used to explore quasi-static damage in whole rock samples but they can be extended to studies of well-slipped granular layers similar to those at the core of natural faults. At the same time, dedicated experiments under hydrothermal conditions are needed to rigorously develop and test the constitutive formulations over a broader range of real-Earth conditions. Future experiments will also build on the recent successes in determining the coupling between deformation and pore-fluid diffusion which exerts critical control over the development of fracture patterns as well as the rate of faulting and fault slip\(^5\).

Similarly, for the deformation in the brittle-ductile transition region and below (MMC1.3), the mechanism boundaries among frictional sliding, crystal-plastic creep and pressure solution are predicted in microphysical models. However, the ductile components are based on steady-state flow laws, e.g.

---

\(^{1}\) (e.g., Saito et al, 2011, Jiang and Simons, 2016)
\(^{2}\) (Fujiwara et al., 2011; Sun et al, 2017)
\(^{3}\) (Cruz-Atienza et al, 2018)
\(^{4}\) (e.g., Burnley and Zhang, 2008; Renard et al., 2017)
\(^{5}\) (French et al., 2016; French and Zhu, 2017)
dislocation creep\(^1\) or pressure solution creep\(^2\). It is important to establish how fluctuations in strain rate affect flow stress in the ductile regime.

At high slip velocity, controlled experiments over significant slips have only been possible in rotary-shear apparatuses under low normal stress conditions, largely unconfined, and without fluids\(^3\), while spontaneous high-slip rate experiments are limited to short durations and small displacement\(^4\). Still, these recent high-velocity experiments have vastly improved our understanding of thermal effects on slip at seismic rates and the resulting enhanced dynamic weakening (MMC1.1; Figure II-MMC1.1). However, the thermo-mechanical processes in the experiments are generally not well characterized and difficult to extrapolate to natural co-seismic slip\(^5\). Future efforts will focus on the development of laboratory equipment that can explore large, confined, high-velocity slip over a more realistic range of simulated crustal conditions, e.g., in the presence of fluids, and with extensive instrumentation. This advance will provide key, currently missing, knowledge needed for formulating constitutive laws for coseismic slip that link the high-velocity response to the low-velocity rate-and-state formulations.

Another frontier is in conducting experiments for high-strain-rate off-fault inelastic processes during co-seismic rupture and for interseismic healing (MMC1.2). The existing damage theories are largely based on experiments under low strain rates applicable to the interseismic periods, yet the damage-related properties of rocks significantly depend on the strain rate (MMC 1.2; Figure II-MMC1.2). Dynamic damage in the rocks surrounding the fault dissipates energy, modifies elastic properties that affects wave propagation (and hence dynamic stress exchange between portions of the rupture), and produces dilatant deformation that can significantly affect the hydromechanical properties (and hence fluid-related effects). Experiments producing this off-fault damage in the laboratory can measure the work required to create it, in addition to capturing the changes of physical properties, such as modulus and permeability. Additionally, experiments to determine the rates of interseismic damage recovery near the Earth’s surface and at deeper hydrothermal conditions are necessary for modeling of interseismic periods. These important future strides will build on the promising recent advances in experiments designed to replicate high-strain-rate damage at a range of conditions as well as damage recovery and fracture healing\(^6\).

Finally, well-instrumented laboratory experiments that produce spontaneous earthquake source or fracturing phenomena will be further developed and can be used for discovery and to validate earthquake source modeling (section III-IN5). Such studies illustrate the coupling (and feedbacks) of a range of co-seismic processes as well as provide constraints on scale-dependent behavior\(^7\). Such experiments to date are limited to bare, smooth surfaces. More realistic conditions will be created by the introduction of granulated fault gouge and controlled fault roughness.

Overall, the future focus will be on multi-physics, scale-appropriate constitutive formulations for fault slip and deformation, developed from and validated by experimental studies, as discussed in section III-IN1 on “Community modeling ecosystem: Modeling collaboration to capture multi-scale constitutive response of faults.” Such theoretically and physically sound constitutive descriptions will allow extrapolation to a range of conditions and across a range of scales\(^8\). The novel advanced experiments discussed here will

---

1. (Aharonov and Scholz, 2019)
2. (Chen and Spiers, 2016)
3. (Di Toro et al., 2011)
4. (Passelègue et al., 2013; Lockner et al., 2017)
5. (Nielsen et al., 2008)
6. (Brantut, 2015 and references therein)
7. (Rosakis, 2002; Stanchitz et al., 2006, Passelègue et al., 2013; McLaskey et al., 2014; Yamashita et al., 2015; Rubino et al., 2017; Brantut, 2018; Renard et al., 2018)
8. (Chen and Spiers, 2016, Aharonov and Scholz, 2018, van den Ende et al., 2018)
inevitably lead to new discoveries. Among the requirements to maintain fundamental research in experimental studies are investments in new technologies and equipment development\(^1\), and cross-discipline collaborations with materials science and engineering\(^2\).

### IV.3 Geological studies and drilling as the only direct window into the subsurface

Geological studies have significantly contributed to our understanding of the earthquake source processes, by illuminating the complex depth-dependent structure of the fault zones (MMC1), establishing paleoseismological constraints on timing and magnitude of previous earthquakes, and sampling the in situ conditions within active faults through drilling. These activities need to actively continue to identify the range of structures that exist and how these evolve through the seismic cycle, from rupture nucleation, to its dynamic propagation and arrest, to post-seismic effects and interseismic healing. Continued geological investigations would provide key constraints on the validity of both experimental programs to determine the physical and mechanical properties of fault rocks as well as modeling studies of natural fault behaviour.

**Studies of exhumed faults**

Studies of exhumed faults have provided the basis for much of our understanding of the range of the complex structures and properties of faults, as well as their spatial heterogeneity. Because exhumed faults record deformation on large spatial scales, over large durations, and at natural strain rates, virtually all of our working knowledge of seismic fault zone architecture and deformation mechanisms come from field studies\(^3\). While natural faults display a high degree of complexity, systematic studies have been able to identify common and key aspects of faults (MMC1), including constraints on the thickness of principal slip zone, important for heat production and an energy sink during seismic faulting\(^4\); earthquake source parameters from studies of pseudotachylyte\(^5\); the state of stress surrounding faults\(^6\); operative deformation mechanisms in faults\(^7\); fluid flow properties of fault zones\(^8\); distribution of damage surrounding faults that can affect rupture characteristics and fluid flow\(^9\); and how fault zone structure is indicative of the mechanics of faults from creep to seismic slip\(^10\). These studies highlight the importance of field studies on exhumed faults for understanding the faulting process, but it also illustrate the need for future studies on a range of scales in order to provide insights and constraints from nature itself.

**Paleoseismological constraints on timing and magnitude of previous earthquakes**

Painstaking field observations underpinned by accurate dating have yielded valuable time series of earthquake activity on many recognised faults\(^11\). Not only do these studies give a record of surface-
rupturing events on faults that may have very low current rates of seismicity (Figure III-MMC3.5), but they also give valuable information on near surface processes. They provide a unique constraint on area-specific SEAS modeling of potential scenarios of future large events (II-MMC4). At the same time, since paleoseismic inferences provide a distribution of potential rupture times, they often have non-unique interpretations in terms of which parts of the fault ruptured together. Collecting more paleoseismic data on key plate-boundary faults would enable more systematic interpretation of the existing data and invaluable ground truth for area-specific modeling of earthquake sequences.

Unique importance of drilling into active fault zones
The studies of exhumed faults, active faults through remote observations, and numerical modeling efforts have collectively created several potential pictures of how an active fault looks at representative seismogenic depths. Consider a mature plate-boundary fault with kilometers of relative displacement. Is its structure mostly a highly localized layer of fault gouge surrounded by a largely healed damage zone, as often hypothesized1 based on studies of some mature exhumed faults2 (Figures II-MMC1-1 and II-MMC3-1)? Or is it a broad zone of damage with multiple fault strands and shear deformation distributed among them3 (Figures II-MMC1-5 and II-MMC3-1)? For which faults one or the other picture applies at depth? The fault response and relevant parameters of large earthquake ruptures in these two cases could be entirely different (section II-MMC3-1). To verify our indirect inferences on the structure of active faults, what materials comprise them, what are the temperature, pore pressure, and stress conditions at the seismogenic depths of 3-10 km and deeper, we must go to the source. This requires drilling which, while expensive, is a highly developed field thanks to the global petroleum industry. Over the past 2+ decades, a number of successful scientific drilling investigations of earthquake faults have been carried out around the world: SAFOD4, Nojima5, Chelungpu6, Wenchuan7, Tohoku8, Alpine Fault9, and Koyna10. While limited to the relatively near surface (a few kilometers at most, and often much shallower), these studies have provided unique constraints on the conditions under which earthquakes occur. At the same time, observations from drilling projects are almost entirely limited to either the interseismic or postseismic period, and major gaps in our knowledge remain that are vital for constraining theory and models of the process. We have few physical samples from active faults, few detailed measurements of in-situ stress, temperature, or pore pressure, almost no near-field measurements of deformation leading up to or during faulting, and no observations that constrain current nucleation or friction models. A number of earthquake-physics can be addressed by a comprehensive fault zone drilling experiment11. One approach would be to capitalize on the “accidental” experiment of induced seismicity (section III-IN3) to develop a targeted field site where earthquakes are being driven by industrial processes. Alternatively, a dedicated test site could be developed as a deep underground laboratory (section III-IN4). Furthermore, in addition to providing physical measurements in the near source, drilling provides real fault materials that are largely unaffected by alteration during exhumation. This provides additional valuable opportunities of studying the nature and structure of these sampled fault rocks as well as performing laboratory experiments on them under simulated crustal conditions.

1 (Rice, 2006)
2 (Chester and Logan, 1987; Chester et al., 1993; Chester et al., 2005)
3 (Savage and Brodsky, 2011; Rabinowitz et al., 2015; Keren and Kirkpatrick, 2016; Rowe et al., 2018)
4 (Bradbury et al., 2011; Hickman and Zoback, 2004; Zoback et al., 2010)
5 (Ando, 2001; Ikeda et al., 2001; Ohtani et al., 2001)
6 (Ma et al., 2006; Wu et al., 2008)
7 (Li et al., 2013; Xue et al., 2013)
8 (Chester et al., 2013; Fulton et al., 2013; Lin et al., 2013)
9 (Sutherland et al., 2012; Townend et al., 2013; Toy et al., 2015; Sutherland et al., 2017)
10 (Gupta et al., 2015, 2017)
11 (Savage et al., 2017)
IV.4 Synergy and complementarity with CIG, CSDMS, SCEC, SZ4D, and other community efforts

A better understanding of earthquake source processes requires bringing together diverse communities of scientists. A number of community organizations in existence or in an active planning phase address complementary problems and provide important lessons for how such science can successfully proceed.

The *Computational Infrastructure for Geodynamics* (CIG\(^1\)) supports development of reusable, well-documented, open-source code in the solid Earth sciences, enabling collaboration between a community of domain scientists and computational and applied math experts. The goals of CIG are to empower Earth scientists by providing access to cutting-edge computing approaches and resources, through co-sponsorship, training workshops, and programming “hackathons”. Particularly successful initiatives include work on bringing best-practice code development and documentation to the community and supporting benchmark efforts. While spanning a much broader range of geoscience topics, from geodynamo to magma migration, CIG has a partial focus on crustal dynamics and seismology, and hosts a number of relevant codes, e.g., PyLith\(^2\) for rupture propagation and visco-elastic small strain crustal deformation and SPECFEM\(^3\) and SW4\(^4\) for wave propagation. Opportunities exist to further broaden the involvement across a wider user base in earthquake source modeling.

The *Community Surface Dynamics Modeling System* (CSDMS\(^5\)) is an effort to bring together geomorphologists and lithospheric deformation modelers to advance our understanding of the interplay between landscape evolution, tectonics, and climate for planetary evolution, including hazards, for example via the feedbacks on fault forcing driven by sediment transport. The mass redistribution due to surface processes has been shown to have a direct control on slip rates and geometrical evolution in normal fault systems, for example. The CSDMS community also provides a successful model for a code-coupling framework in which interoperability has been enforced by requiring a common code interface. Like CIG, CSDMS supports a number of important community building and workforce development efforts on the interface of geomorphology and lithospheric deformation, helping to educate students and PIs, and learning from each other across disciplines.

Besides their successes, CSDMS and CIG offer cautionary tales for advancing cross-scale modeling as well. In particular, early efforts at coupling codes by prescribing results from large-scale computations as boundary conditions for small-scale codes with more complete physics led only to limited applications. CSDMS framework is so far limited to serial execution, and many problems demand massively parallel, well-load-balanced implementations. A recent CIG and CSDMS sponsored workshop indeed included the recommendation of the development of new, tightly integrated, adaptive- and variable-resolution methods using advanced computational libraries to solve the most challenging lithospheric deformation problems in parallel on modern machines\(^6\).

The *Southern California Earthquake Center* (SCEC\(^7\)) has a geographically defined science objective, to understand the earthquake source processes in light of their effect on seismic hazard and risk for southern California. SCEC is driven by the idea of supporting a virtual and in-person “collaboratory” that

---

1. https://geodynamics.org/cig/about/
5. https://csdms.colorado.edu/wiki/Main_Page
7. https://www.scec.org/
brings together geophysicists, geologists, engineers, and emergency managers to work on community models. These community models are in various stages of completeness, and strive to capture fault geometry, seismic velocity structure, geodetic deformation, temperature, stress, rheology, and geology. Those frameworks are openly debated, developed, and shared, facilitating an integration of structural constraints and dynamic modeling for the study of the San Andreas fault system. SCEC has made much headway in bridging the gap between science and application, and outreach and training efforts have broadened participation and diversity of the workforce. Key achievements of SCEC include the establishment of high-performance computing workflows, a reinvigoration of the thorough study of earthquake predictability, and the support of benchmarking and model building efforts. Many SCEC initiatives can be key partners in advancing earthquake source science, e.g., by providing observational constraints in the form of community models, including the evaluation of fault system interaction models for the scaling and evolution of stress and seismicity heterogeneity.

Understanding the Processes that Underlie Subduction Zone Hazards in 4D (SZ4D) is a new initiative that seeks to advance our understanding of arc volcanoes and megathrusts through the establishment of a science funding program, new, large-scale infrastructure for observational efforts, and the Modelling Collaboratory for Subduction Zone Science (MCS). Such a collaboratory is meant to be a new kind of community center, driven by the development of numerical models that are capable of capturing geological and geophysical constraints for advancing cross-scale earthquake and volcano modeling, including for physics-based, decadal hazard assessment (Figure IV-2). Combining best practices from earlier community efforts, the MCS seeks to create and deploy “Lego blocks” of efficient and robust codes for the study of fault systems in general, with focus on testing interactions across spatio-temporal scales as well as provide regionally specific block assemblies for different subduction zones worldwide. The MCS may form a platform for cross-megathrust study as well as integration of cross-disciplinary observatory data, enabling and supporting new discourse between a range of stakeholders. The commonality of interests with studies of the earthquake source includes shared tools for testing physical ingredients, such as candidate fault constitutive laws, in data-rich environments for adjoint, inverse, and optimal experimental design problems within the megathrust context.

In Europe, research and training networks such as SPICE and QUEST established excellence in computational seismology focusing specifically on the assembly of open software, tools, and algorithms. Recent efforts aim at deep integration of supercomputing centres and Earth science modeling, e.g. by establishing a Centre of Excellence for Exascale in Solid Earth (ChEESE) coordinated by the Barcelona Supercomputing Centre, to harness European institutions in charge of operational monitoring networks, tier-0 supercomputing centers, academia, hardware developers and third-parties from SMEs, Industry and public-governance to prepare flagship codes and enable services for exascale supercomputing in the area of solid earth. The VERCE project explored service-oriented HPC architecture and data-intensive platforms integrated within the European Plate Observing System (EPOS) and computing infrastructures (GRID, HPC and CLOUD). The Faulty2SHA ESC working group successfully establishes exchange between field geologists, fault modelers and seismic hazard practitioners by thematic meetings and workshops, training days, special sessions at international conferences and tools development.

---

2 http://sz4dmcsc.org
3 http://www.spice-rtn.org/
4 http://www.quest-itn.org/
5 www.cheese-coe.eu
6 (http://www.verce.eu, Atkinson et al., 2015)
7 (https://fault2sha.net)
All of these community initiatives will benefit from the new science and method development that is outlined in this report on the study of the earthquake source, and the communities can leverage each other’s investments and efforts to advance earthquake systems science.

Figure IV-2. Modeling collaboratory for subduction zone science. The goal of this modeling framework is to capture physico-chemical processes bridging convection, fractionation, tectonics, megathrust dynamics, and volcano dynamics while assimilating comprehensive dataset for improved hazard assessment. Adapted from Becker (2018), with images from Tong and Lavier (2016), W. Behr (pers. com, 11/2017), Naif et al. (2015), Schmalzle et al. (2014), Proctor and Hirth (2015) and Gomberg et a. (2010). Targeted progress in modeling earthquake source processes described in this report will benefit this and other community initiatives.
Appendix A: Workshop participants

Participants (on-site and remote) at the Workshop on “Modeling Earthquake Source Processes: from Tectonics to Dynamic Rupture” held on October 8-10, 2018 in Pasadena, California

Brad Aagaard United States Geological Survey
Rachel Abercrombie Boston University
Niloufar Abolfathian University of Southern California
Lise Alalouf McGill University
Kali Allison Stanford University
Jean-Paul Ampuero California Institute of Technology, Université Côte d’Azur
Khurram Aslam University of Memphis
Jean-Philippe Avouac California Institute of Technology
Han Bao University of California Los Angeles
Sylvain Barbot University of Southern California
Thorsten Becker University of Texas at Austin
Nicholas Beeler United States Geological Survey
Yehuda Ben-Zion University of Southern California
Gregory Beroza Stanford University
Harsha Bhat École Normale Supérieure
James Biemiller University of Texas Institute for Geophysics
Nicolas Brantut University College London
Emily Brodsky University of California Santa Cruz
Lucile Bruhat École Normale Supérieure
Brennan Brunsvik University of Louisiana
Roland Bürgmann University of California Berkeley
Camilla Cattania Stanford University
Kejie Chen California Institute of Technology
Xiang Chen The Chinese University of Hong Kong
Yifang Cheng University of Southern California
Eunseo Choi The University of Memphis
Elizabeth Cochran United States Geological Survey
Fabio Corbi Roma Tre University
Luca Dal Zilio Swiss Federal Institute of Technology in Zurich
Marine Denolle Harvard University
Martijn van den Ende Utrecht University
Ylona van Dinther Swiss Federal Institute of Technology in Zurich / Utrecht University
Giulio Di Toro
University of Padua

Jim Dieterich
University of California Riverside

Roby Douilly
University of California Riverside

Ben Duan
Texas A&M University

Eric Dunham
Stanford University

Ahmed Elbanna
University of Illinois Urbana Champaign

William Ellsworth
Stanford University

Brittany Erickson
Portland State University

Wenyuan Fan
Woods Hole Oceanographic Institution

Daniel Faulkner
University of Liverpool

Behrooz Ferdowsi
Princeton University

Yuri Fialko
Scripps Institute of Oceanography, University of California San Diego

Jay Fineberg
The Hebrew University of Jerusalem

William Frank
University of Southern California

Jeff Freymueller
Michigan State University

Alice-Agnès Gabriel
Ludwig Maximilian University of Munich

Percy Galvez
King Abdullah University of Science and Technology

Dara Goldberg
Scripps Institute of Oceanography, University of California San Diego

David Goldsby
University of Pennsylvania

Joan Gomberg
United States Geological Survey

Ashley Griffith
Ohio State University

Michael Gurnis
California Institute of Technology

Jennifer Haase
Scripps Institute of Oceanography, University of California San Diego

Egil Hauksson
California Institute of Technology

Steve Hickman
United States Geological Survey

Greg Hirth
Brown University

Heidi Houston
University of Southern California

Yihe Huang
University of Michigan

Ken Hudnut
United States Geological Survey

Benjamin Idini
California Institute of Technology

David Jackson
University of California Los Angeles

Tamara Jeppsson
Texas A&M University

Yunzhong Jia
Nanyang Technological University

Junle Jiang
Cornell University

Taka Kanaya
University of Maryland

Yoshihiro Kaneko
GNS Science

Zoheir Khademian
Colorado School of Mines
Modeling earthquake source processes: from tectonics to dynamic rupture

Carlos Reinoza Gómez
Centro de Investigación Científica y de Educación Superior de Ensenada

Arthur Rodgers
Lawrence Livermore National Laboratory

Ares Rosakis
California Institute of Technology

Zachary Ross
California Institute of Technology

Christie Rowe
McGill University

Kenny Ryan
United States Geological Survey

Valerie Sahakian
University of Oregon

Charles Sammis
University of Southern California

Hamid Sana
Institute of Rock Structure and Mechanics, Czech Academy of Sciences

Swasti Saxena
University of Nevada Reno

Paul Segall
Stanford University

Daya Shanker
Indian Institute of Technology Roorkee

Farrokh Sheibani
Massachusetts Institute of Technology

Zhichao Shen
California Institute of Technology

Pengcheng Shi
University of Rhode Island

Toshihiko Shimamoto
Shimamoto Earth and Environment Ltd.

Mark Simons
California Institute of Technology

Hiroki Sone
University of Wisconsin-Madison

Oliver Stephenson
California Institute of Technology

Yuval Tal
California Institute of Technology

Amanda Thomas
University of Oregon

Harold Tobin
University of Wisconsin-Madison, University of Washington

Xinyue Tong
University of Texas at Austin

John Townend
Victoria University of Wellington

Daniel Trugman
Los Alamos National Laboratory

Victor Tsai
California Institute of Technology

Terry Tullis
Brown University

Martijn van den Ende
Utrecht University

Ylona van Dinther
Swiss Federal Institute of Technology in Zurich / Utrecht University

John Vidale
University of Southern California

Robert Viesca
Tufts University

Juan Carlos Villegas Lanza
Instituto Geofísico del Perú

Robert Walker
University of Southern California

Yongfei Wang
University of California San Diego, San Diego State University

Kyle Withers
United States Geological Survey

Yuqing Xie
University of California Los Angeles
<table>
<thead>
<tr>
<th>Name</th>
<th>Last Name</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhuo</td>
<td>Yang</td>
<td>Harvard University</td>
</tr>
<tr>
<td>Hongfeng</td>
<td>Yang</td>
<td>Chinese University of Hong Kong</td>
</tr>
<tr>
<td>Jiuxun</td>
<td>Yin</td>
<td>Harvard University</td>
</tr>
<tr>
<td>Zhongwen</td>
<td>Zhan</td>
<td>California Institute of Technology</td>
</tr>
<tr>
<td>Wenlu</td>
<td>Zhu</td>
<td>University of Maryland</td>
</tr>
<tr>
<td>Paolo</td>
<td>Zimmaro</td>
<td>University of California Los Angeles</td>
</tr>
</tbody>
</table>