Introduction to special section: Stress transfer, earthquake triggering, and time-dependent seismic hazard


Link to publication record in Ulster University Research Portal

Published in:
Journal of Geophysical Research: Solid Earth

Publication Status:
Published (in print/issue): 01/05/2005

DOI:
10.1029/2005JB003692

Document Version
Publisher's PDF, also known as Version of record

General rights
Copyright for the publications made accessible via Ulster University’s Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Ulster University’s institutional repository that provides access to Ulster’s research outputs. Every effort has been made to ensure that content in the Research Portal does not infringe any person’s rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact pure-support@ulster.ac.uk.
Introduction to special section: Stress transfer, earthquake triggering, and time-dependent seismic hazard

Sandy Steacy
Geophysics Research Group, School of Environmental Sciences, University of Ulster, Coleraine, Northern Ireland, UK

Joan Gomberg
U.S. Geological Survey, Memphis, Tennessee, USA

Massimo Cocco
Seismology and Tectonophysics Department, Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy

Received 15 February 2005; revised 29 March 2005; accepted 31 March 2005; published 13 May 2005.

[1] In this introduction, we review much of the recent work related to stress transfer, earthquake triggering, and time-dependent seismic hazard in order to provide context for the special section on these subjects. Considerable advances have been made in the past decade, and we focus on our understanding of stress transfer at various temporal and spatial scales, review recent studies of the role of fluids in earthquake triggering, describe evidence for the connection between volcanism and earthquake triggering, examine observational evidence for triggering at all scales, and finally discuss the link between earthquake triggering and time-dependent seismic hazard. We conclude by speculating on future areas of research in the next decade.

Citation: Steacy, S., J. Gomberg, and M. Cocco (2005), Introduction to special section: Stress transfer, earthquake triggering, and time-dependent seismic hazard, J. Geophys. Res., 110, B05S01, doi:10.1029/2005JB003692.

1. Introduction

[2] Herein we introduce the topic of this special section on “stress transfer, earthquake triggering, and time-dependent seismic hazard,” providing some definitions and background on the subject. We focus on literature published since about 1998, and for earlier work we refer readers to the excellent reviews of Harris [1998], Stein [1999], and King and Cocco [2001]. The references cited herein by no means compose a complete listing on the subject, as our primary purpose is to provide context for the papers in the special section. “Earthquake triggering” refers to the causal relationship between changes in the characteristics of, or processes acting on, a fault and in the potential for rupture to nucleate on it. Most often the considered changes are the stresses acting on the fault surface, and we discuss triggering in this context but note that other changes may be relevant such as those in the strain or displacement field, or perhaps even some chemical property. “Stress transfer” refers to the process by which a stress perturbation is communicated to the fault. As the sources of stress changes and processes of stress transfer vary temporally and spatially, hazard assessments also must account for these variations. Understanding and quantifying this is the realm of “time-dependent seismic hazard.”

[3] Most of our knowledge of earthquake triggering comes from observations of measurable changes in earthquake occurrence rates following a triggering event. Changes in earthquake rates are usually evident in changes in the seismicity rate of some population of small earthquakes (e.g., aftershocks), or as the delayed or early recurrence of a single or several large earthquakes. However, despite abundant observations of earthquake triggering, many questions remain unanswered about the physical mechanisms that may explain them. We highlight the most intriguing and recent of these observations, and the questions and models proposed to explain them. Our summary begins with discussion of the stress transfer models appropriate to different temporal and spatial scales; often considered separately as static, dynamic, and visco-elastic processes, but noting that in reality they occur as part of a continuum. We then discuss current thinking about the roles of fluids and volcanic processes in earthquake triggering. We present an overview of the observational evidence for triggering at all scales next. We finish with a few paragraphs describing the state of time-dependent seismic hazard research and implementation, and some concluding speculation about where future research may lead us.

2. Stress Transfer Model

[4] Static stress changes (or “triggers”) generally refer to changes that effectively occur instantaneously and
permanently. Most commonly this refers to the time-independent deformations caused by the net slippage on a fault resulting from an earthquake. Dynamic stress triggers are those caused by the passage of seismic waves, which notably are oscillatory and transient. Dynamic and static stress changes cannot be distinguished either observationally or theoretically at short times and distances from an earthquake, and both attenuate approximately as some inverse power of the distance. Static stress changes attenuate more rapidly, roughly inversely to the cube of the distance from the causative fault, while dynamic stress changes attenuate more slowly and therefore dominate at large distances. Coseismic stress perturbations also cause viscoelastic relaxation of the lower crust and the upper mantle, which contributes to time-dependent stress transfer over long times, which can lead to either stress enhancement or decrease. Other natural processes may modify the temporal evolution and spatial pattern of stresses acting on a fault, such as the growth or melting of glaciers [Grollimund and Zoback, 2001] and tides [see Tanaka et al., 2002, and references therein], as well as human activities such as mining, hydrocarbon and geothermal energy production, and reservoir construction [Gupta and Chadha, 1995; Trij u and Fehler, 1998; Richardson and Jordan, 2002]. These too may promote or inhibit the earthquake failure, but we leave them to future summaries.

[5] As the abundance of phenomenological studies has confirmed the reality of static stress change triggering (summarized below), studies increasingly focus on questions related to the underlying physical processes. Among these are the following: How can very small stress changes lead to measurable changes in failure rates? Why is failure often delayed from the occurrence of the stress change? Is there a static stress change threshold that must be exceeded for triggering to occur? Most attempts to answer these have focused on frictional failure models, particularly the failure rate change formulation of Dieterich [1994], based on rate-state friction theory. This formulation and generalizations of it have been examined from a purely theoretical perspective [e.g., in Gomberg et al., 2000, 2001; Ziv and Rubin, 2003; Gomberg et al., 2005b], and predictions compared with observations [e.g., Gross, 2001; Toda et al., 2002; Toda and Stein, 2003]. These models show the efficacy of small static stress changes to cause faults to fail early (or late), yet not immediately after the application of the stress change and also predict basic observations like aftershock rates that decay inversely with time. The delayed occurrence of triggered seismic events is explained in these models as the effect of the frictional response of the fault to the applied stress perturbations. Most recently, it has become possible to verify observationally details of predicted aftershock decay rates, such as an apparent deficit of aftershocks in the initial seconds to minutes after the main shock. Interestingly, most frictional models only consider static stress change triggering as altering the time of failure of faults for which failure was inevitable anyhow. However, studies like those of Voisin et al. [2004] consider the potential for static stress changes (and dynamic stresses) to move a stably sliding frictional fault into a regime in which it exhibits stick-slip behavior. This perhaps is one way in which triggering is not simply an acceleration of failures but creation of new ones.

[6] Phenomenological studies have also probed looking just at the Coulomb static stress changes, often referred to as the Coulomb failure function (see Cocco and Rice [2002] for further details):

$$\Delta \text{CFF} = \Delta \tau + \mu (\Delta \sigma_n + \Delta P),$$  

where $\Delta \tau$ are the shear stress changes calculated along the slip direction on the assumed fault plane, $\Delta \sigma_n$ are the normal stress changes (positive for extension), $\mu$ is the friction coefficient and $\Delta P$ indicates the pore pressure changes. Recent studies have examined the importance of the various stress components, $\mu$, and $\Delta P$ (discussed in later paragraphs). Additionally, questions concerning triggering thresholds and correlations between the prevalence of triggering and faulting type, environmental and geological characteristics have been examined. In all these, generally the likelihood of causality is based on some estimate of the correlation between static stress changes and observed changes in failure rates. Studies have shown that in some cases shear stress changes correlate better with triggered seismicity and in others, normal stress changes do, perhaps implying coefficients of friction that vary from fault to fault [e.g., Parsons et al., 1999; Perfettini et al., 1999; ten Brink and Lin, 2004]. Most studies have examined strike-slip systems, and although fewer in number, the degree of correlations (or lack of) measured in triggering studies of predominantly thrust faulting in compressional environments do not differ in any clear or systematic way [Ma et al., 2005].

[7] While most studies suggest that triggering requires a minimum stress change, the variation in this threshold spans an order of magnitude or more, and at least two studies infer that triggering is not a threshold process [Ziv and Rubin, 2000; Ogata, 2005]. A variety of phenomenological studies and theoretical models seem to suggest that triggering is facilitated by low ambient confining stresses or high pore pressures [e.g., Gross, 2001; Streit and Cox, 2001; Toda and Stein, 2002; Woessner et al., 2004]. Toda et al. [2005] hypothesize that since frictional models predict that triggering manifests as a change in seismicity rate, areas with the most apparent triggered activity should be those with relatively high background seismicity rates. They find this to be generally true in their analysis of seismicity in southern California. Their results also imply that triggering primarily represents clock-advanced failure, rather than creation of new earthquakes (which are less likely to correlate with background seismicity).

[8] Here we note that perhaps the most novel work in static stress triggering has been in the development of approaches to translate quantitatively static stress changes into changes in failure probabilities. We save this discussion for the end.

[9] The spatially widespread, immediate changes in seismicity rates that followed the 1992 $M_w = 7.3$ Landers, California, earthquake removed most doubts about the potential for dynamic triggering [Hill et al., 1993]. Similar observations made following the 2002 $M_w = 7.9$ Denali, Alaska, earthquake demonstrated that the post-Landers phenomena were not likely freak events [Eberhart-Phillips et al., 2003; Gomberg et al., 2004; Husen et al., 2004a, 2004b; Prejean et al., 2004; Pankow et al., 2004]. In
addition to these observations and other studies documenting episodes of triggered seismicity at distances well beyond the traditional aftershock zone, which typically equals several multiples of the main shock fault dimension [e.g., Brodsky et al., 2000; Hough, 2001; Mohamad et al., 2000; Hough, 2005], there have been a number of studies that have shown observationally that dynamic triggering is not only a far-field phenomenon [e.g., Harris et al., 2002; Gomberg et al., 2003]. Indeed, logically this should not be surprising as triggered faults have no knowledge of the distance traveled by the triggering dynamic deformations, and as noted in our introductory remarks, in the near field of an earthquake both static and dynamic stress changes both occur [Fischer and Horalek, 2005].

[10] Two hypotheses have been posed and tested to distinguish observationally dynamic and static stress change triggering in the near field. The first hypothesizes that dynamic stresses only promote failure so that stress shadows, or areas in which negative static stress changes correlate positively with decreases in failure rates, can only be due to static stress changes [Toda and Stein, 2003]. Harris and Simpson [1993] first coined the phrase “stress shadow” to describe the apparent deficit of moderate earthquakes following large events in the San Francisco Bay region in 1868, 1906, and 1989, and the existence of a stress shadow following the 1906 earthquake has become widely accepted [Kenner and Segall, 1999; Reasenberg et al., 2003; Pollitz et al., 2004]. Stress shadows also have been inferred in other studies [e.g., Parsons et al., 1999; Stein, 1999; Wyss and Wiemer, 2000; Ogata et al., 2003; Toda and Stein, 2003; Woessner et al., 2004]. However, Felzer and Brodsky [2005] have applied a new method for measuring rate changes and conclude that stress shadows either do not exist or cannot be distinguished from other processes (e.g., decreasing rates from aftershock sequences of small earthquakes). Their results are consistent with those of Marsan [2003], who fit a nonstationary Poisson probability function to seismicity before and after three large California earthquakes and found rate decreases were very rarely observed in the 100 days following each main shock. Thus it seems that the existence and/or persistence of stress shadows remains an open question. The second hypothesis is that dynamically triggered rate changes should be more likely in areas of strong focusing of seismic waves, as expected for directivity associated with unilaterally propagating ruptures. A correlation between seismicity rate changes and rupture directivity has been noted in cases of triggering at remote distances [Mohamad et al., 2000; Gomberg et al., 2001; Kilb et al., 2002; Kilb, 2003; Gomberg et al., 2004] and in the near field [Gomberg et al., 2003].

[11] Considerable progress has been made in the development and validation of theoretical models of the triggering potential of both static and dynamic loads [Gomberg et al., 2001; Antonioli et al., 2002, 2004; Voisin et al., 2000, 2004]. Like static stress change models, most of these focus on frictional faults or other types of accelerating failure models (e.g., subcritical crack growth). When considering triggering as clock advancing of inevitable failures, these types of models do not appear capable of generating sufficiently delayed failures to explain the extended duration of aftershock sequences [Gomberg et al., 2001; Belardinelli et al., 2003]. Most other models of dynamic triggering involve fluids in various ways. A particularly novel one proposes that fluids mobilized by shaking may unplug fractures and cause a redistribution of pore pressures, possibly reducing the effective stress on some faults thereby promoting failure [Brodsky et al., 2003].

[12] A causal relationship between dynamic deformations and changes in failure rates has been established, and evidence thus far suggests triggering requires that some minimum dynamic deformation amplitude be exceeded [Gomberg et al., 2004; Brodsky and Prejean, 2005]. However, this threshold clearly varies spatially [e.g., Gomberg et al., 2004] and perhaps temporally as well. Moreover, other characteristics of the dynamic deformation field, of the environment surrounding a fault, and of the fault itself undoubtedly also affect triggering. Theoretical models have examined the dependence on frequency, among other things [Voisin, 2001, 2002; Beeler and Lockner, 2003; Perfettini et al., 2003], but to date, no consensus has emerged. The aforementioned model involving fracture unloading proposed by Brodsky et al. [2003] implies low-frequency waves more effectively trigger, with some observational support provided in an analysis of data from Long Valley, California, an active geothermal area [Brodsky and Prejean, 2005]. No clear environmental factor alone seems to be a reliable predictor of the potential for dynamic triggering, although a sufficient number of observations from diverse locations show that such triggering is more likely in geothermal or volcanic areas [Brodsky et al., 2000; Power et al., 2001; Prejean et al., 2004; Brodsky and Prejean, 2005] but certainly can occur elsewhere [Hough, 2001; Mohamad et al., 2000; Gomberg et al., 2003, 2004; Hough, 2005].

[13] The recognition of dynamic triggering of earthquakes has motivated several laboratory studies appropriate to seismogenic conditions, although the effects of cyclic loading have been studied extensively in engineering fields [see Boettcher and Marone, 2004, and references therein]. The results of laboratory experiments on simulated creeping faults subjected to a constant stressing and oscillating normal stresses show that the latter may cause both strengthening and weakening, with weakening occurring for sufficiently large-amplitude high-frequency oscillations [Richardson and Marone, 1999; Perfettini et al., 2001; Boeticher and Marone, 2004]. The laboratory experiments of Beeler and Lockner [2003] simulate the stick-slip response of faults loaded at constant slip rate modulated sinusoidally and show a response that at low frequencies is Coulomb-like (failure occurs at peak stress, implying a threshold amplitude that decreases inversely with frequency). At periods shorter than the timescales of frictional nucleation processes becomes essentially frequency-independent, with a threshold amplitude that is larger than a Coulomb model would predict but still smaller than that at low frequencies.

[14] Perhaps the strongest evidence for viscoelastic triggering comes from the 1999 \( M_w = 7.1 \) Hector Mine, California, earthquake. This event occurred 20 km away and a little over 7 years after the 1992 \( M_w = 7.3 \) Landers, California, earthquake. The close proximity in space and
time of these two events suggested a causal link. However, observations of static stress transfer were ambiguous at best [Harris and Simpson, 2002], with some models predicting positive Coulomb static stress changes at the Hector Mine hypocenter [Parsons and Dreger, 2000] and others negative changes [U.S. Geological Survey, Southern California Earthquake Center, and California Division of Mines and Geology, 2000]. Other explanations for a possible link between the two events included triggering by aftershocks [Felzer et al., 2002; Harris and Simpson, 2002], poroelastic recovery [Masterlark and Wang, 2002], and dynamic stress changes [Kilb, 2003]. Freed and Lin [2001] used a three-dimensional (3-D) viscoelastic model to simulate flow in the lower crust and upper mantle due to the Landers earthquake and found positive postseismic stress increases of 1–2 bars at the Hector Mine hypocenter; they hypothesized that this viscoelastic stress change triggered the event and explained its time delay. In similar studies, Zeng [2001] modeled viscoelastic flow in the lower crust and found a stress increase >1 bar at the hypocenter, while Pollitz and Sacks [2002] suggested a stress increase of approximately 0.7 bar. A detailed 3-D finite element model of the interaction between these events also is given by Cianetti et al. [2005].

[15] Viscoelastic stress transfer has been examined in studies of other fault systems, with a number of authors comparing its possible effects with those from static stress changes. Hearn et al. [2002] found that viscoelastic stress transfer due to the 1999 $M_w = 7.4$ Izmit, Turkey, earthquake increased the stress load on the $M_w = 7.1$ Duzce, Turkey, hypocenter by about 70%; they also modeled a viscoelastic stress increase in the Marmara Sea area to the west but noted that their results were highly model-dependent. Freed and Lin [2002] proposed that viscoelastic stress changes on the San Andreas fault due to four $M > 6$ events in the Mojave Desert in southern California would more than double the static stress change by 2020. In a model of the 1906 San Francisco stress shadow, Parsons [2002a] found that including viscoelastic effects increased the duration of the stress shadow and lowered earthquake probabilities with respect to models using deep dislocation slip. Pollitz et al. [2004] included a simple viscoelastic coupling in a model of strain accumulation in the San Francisco Bay area and concluded that virtually all moderate to large events were consistent with triggering and that the region recovered from the stress shadow in about 1980, consistent with the results of Parsons [2002a].

[16] Viscoelastic stress transfer has also been proposed to explain event correlations over large times and distances. Rydelek and Sacks [2001] proposed that it explained the southward migration of $M \geq 5.6$ events along the San Jacinto fault in southern California from the large 1857 San Andreas event, and Chery et al. [2001] suggested that it explained three $M > 8$ events that occurred up to 400 km apart in Mongolia between 1905 and 1952. Similarly, Pollitz et al. [2003] concluded that five $M > 7$ events and a number of smaller ones in Mongolia were related by coseismic and viscoelastic stress changes and proposed that continental faults could experience significant stress transfer over timescales of decades and length scales of hundreds of kilometers. This view was supported by To et al. [2004], who suggested that the 2001 $M_w = 7.6$ Bhuj, India, earthquake was triggered by the 1819 $M_w = 7.7$ Rann of Kachchh, India, event, 100 km away.

3. Fluids and Earthquake Triggering

[17] The importance of fluids in promoting earthquake failures has been suggested by many different authors [Nur and Booker, 1972]. Geological observations suggest that fault zones contain a core with a thin slipping plane [Chester et al., 1993; Chester and Chester, 1998; Sibson, 2003] embedded in a highly fractured damage zone, possibly fluid-saturated [see also Ben Zion and Sammis, 2003]. The response of a poroelastic medium to a sudden applied stress changes should take into account the spatial and temporal evolution of pore pressure. Sudden pore pressure changes, $\Delta P$, modify the coseismic stress redistribution because they affect the Coulomb Failure Function (equation (1)). These pore pressure changes represent the undrained response of the medium since they occur on a timescale that is sufficiently short that diffusive transport (i.e., fluid flow) cannot occur. Beeler et al. [2000] and Cocco and Rice [2002] have discussed the proper way to correctly include undrained pore pressure changes in Coulomb stress modeling and obtain a general expression for the effective normal stress perturbations. These authors compared two alternative pore pressure models: the constant apparent friction model (widely used in the literature, in which pore pressure is proportional to normal stress changes) and a second, more general, isotropic, poroelastic model (in which pore pressure perturbations are proportional to the mean, or volumetric, stress changes). These pore pressure models have been used and compared to explain the spatial pattern of aftershocks [Nostro et al., 2005].

[18] At longer timescales the drained response of the medium will modify the pore pressure evolution promoting fluid flow [Nur and Booker, 1972]. The temporal evolution of pore pressure depends on the permeability and the thickness of the slipping zone, where fluids may flow. Diffusive processes of pore pressure relaxation in fractured and saturated rocks have been proposed to explain both earthquakes [Noir et al., 1997; Bosl and Nur, 2002] and induced seismicity [see Shapiro et al., 2003; Paroditis et al., 2005, and references therein]. Moreover, poroelastic rebound has been proposed to explain postearthquake ground deformation [Felzer et al., 1998]. A clear correlation between postearthquake surface deformation and pore pressure transients has been recently found in south Iceland [Jonsson et al., 2003].

[19] The temporal migration of aftershock locations as well as of induced seismicity is currently modeled through the Darcy law [Carslaw and Jaeger, 1959], which is solved to obtain the time-dependent pore pressure relaxation,

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial x_i} \left[D_{ij} \frac{\partial p}{\partial x_j}\right],$$

(2)

where $D_{ij}$ are the components of the hydraulic diffusivity tensor and $p$ is the pressure. The hydraulic diffusivity tensor ($D$) can be related to permeability tensor ($K$) using the relation

$$D = \frac{K}{\varepsilon^2 C_T},$$

(3)
where ε and ν are the rock porosity and the fluid viscosity, respectively, and $C_T$ is the coefficient of the isothermal compressibility. This pore pressure evolution can be generated either by the coseismic stress changes caused by a previous nearby earthquake (as in the case of aftershocks or natural seismicity) or by fluid injection in boreholes. The spatiotemporal pattern of seismicity is recognized as a signature of fluid diffusion processes (see Hainzl and Ogata [2005] or Paroditis et al. [2005]). The analysis of seismicity pattern in terms of Darcy flow allows the estimate of hydraulic diffusivity and sometimes permeability [see Shapiro et al., 2003; Antonioli et al., 2005]. These estimates of hydraulic diffusivity obtained for geothermal and tectonic areas range between $10^{-2}$ and $10^2$ m$^2$/s [see Talwani et al., 1999; Shapiro et al., 1999]. Moreover, the hydraulic diffusivity tensor should be anisotropic, particularly for tectonic areas where this anisotropy arises from the permeability anisotropy of damage zones.

[26] The hydraulic diffusivity and the permeability often are considered constant, because they are estimated as average values for a volume. However, the propagation of a dynamic rupture within the damage zone might change porosity and increase permeability. Recently, Miller et al. [2004] proposed that seismicity on the hanging wall of normal faults is promoted by a pressure pulse originating (coseismically) from a known deep source of trapped high-pressure fluids and propagating into the damage region created by the earthquake. They suggest that high-pressure CO$_2$ may be trapped in the anhydrite/dolomite sedimentary layer. Antonioli et al. [2005] suggest that such a mechanism might explain the 1997 Colfiorito earthquake sequence.

4. Volcanism and Earthquake Triggering

[21] Interactions between earthquakes and volcanic eruptions have been studied in several regions worldwide. The results of these studies clearly show that large fault systems and volcanic sources (dikes, sill or magma chambers) are mechanically coupled [Hill et al., 2002]. Volcanic processes may promote earthquakes on surrounding faults by increasing the elastic stress and similarly a volcanic system can be perturbed by small stress changes induced by a large neighboring earthquake [Nostro et al., 1998; Hill et al., 2002; N. Feuillet et al., Stress interaction between seismic and volcanic activity at Mount Etna, submitted to Geophysical Journal International, 2005, hereinafter referred to as Feuillet et al., submitted manuscript, 2005]. In Italy, Nostro et al. [1998] have shown that large normal faulting earthquakes along the Apennines could have promoted Vesuvius eruptions by compressing its magmatic chamber at depth and by opening suitably oriented dikes at the surface. Feuillet et al. (submitted manuscript, 2005) investigated the two-way coupling between large-magnitude historical earthquakes in eastern Sicily and eruptions at the Etna volcano by modeling the stress transfer in a three-dimensional, elastic, half-space. Comparison between eruptive sequences and historical seismicity shows that the large earthquakes that struck eastern Sicily occurred after long periods of flank activity at Etna. Toda et al. [2002] have shown that dike opening at Izu Island (Japan) promoted $M \geq 6$ strike-slip earthquakes tens of kilometers away from the volcanoes, and an evident seismicity rate increase. Cayol et al. [2000] found that earthquakes produced far from the Kilauea rift zone (Hawaii) were promoted by Coulomb stress changes caused by the dilation of the volcano rift-dike during its large 1983 flank eruption. Feuillet et al. [2004] investigated the spatial correlation between seismicity during an earthquake swarm and Coulomb stress changes caused by magma intrusion in a local volcanic source in an extensional tectonic stress field at the Alban Hills Volcano (Italy). Moreover, fluid triggering has been reported in volcanic areas [see Waite and Smith, 2002; Prejean et al., 2003]. Magma bodies produce fluids and heat existing fluid compartments causing pore pressure changes and promoting fluid flow.

[23] Hill et al. [2002] present an interesting review of earthquake-volcano interactions pointing out the role of static and dynamic stress changes as well as that of viscoelastic relaxation. Most of the studies cited above have investigated short- or intermediate-range interactions (at distances within 100 km from volcanoes). However, one of the clearest observations of seismicity rate increase caused by dynamic triggering occurs in volcanic systems. The observations collected after the 1992 Landers, California, earthquake contributed to the evidence of earthquake-volcano interactions, but at the same time emphasized their complexity. Several other recent earthquakes allowed the collection of observations of dynamic triggering of seismicity in volcanic systems (see above). Despite numerous observations and the numerical modeling of earthquake-volcano interactions through stress transfer, many other studies are needed to statistically corroborate this mutual relationship (to advance this goal, new instrumental data and complete catalogues are required) and to determine the physics and the chemistry of the triggering processes.

5. Observational Evidence of Earthquake Triggering at All Scales

[21] Early research on earthquake triggering investigated earthquake interaction primarily on two distinct scales, main shock–main shock and main shock–aftershock. More recent work suggests that triggering can occur at all scales and hence triggered events can be of any size. In this section we summarize recent observational evidence of earthquake triggering as well as new techniques for quantifying stress changes and their consequences, and outline some unresolved problems.

[24] The now classic example of main shock triggering is the 1999 Izmit, Turkey, earthquake ($M_w = 7.4$) that occurred along a section of the North Anatolian fault. This section had been identified by Stein et al. [1997] as being highly loaded due to both secular loading and coseismic stress transfer from previous large earthquakes ($M > 6.7$) along the North Anatolian fault and by Nalbant et al. [1998], who computed a large coseismic stress increase due to regional events ($M > 6$) in northwest Turkey. Modeling studies show that stress redistribution from the Izmit earthquake may have led to positive Coulomb stress changes on the fault both to the west, near Istanbul [Hubert-Ferrari et al., 2000], and to the east [Barka, 1999; Utkucu et al., 2003]; the latter experienced the $M_w = 7.1$ Duzce earthquake approximately 3 months later, and it is now believed that the seismic hazard is very high in the Istanbul region [Parsons, 2004].
[25] Large event stress interaction has been recently inferred in a number of other regions. Doser and Robinson [2002] found that at least six of seven strike slip events ($M \geq 5.9$) in a $150 \times 150$ km region in New Zealand occurred on faults experiencing positive Coulomb stress. Anderson and Ji [2003] showed that stress transfer due to the $M_c = 6.7$ Nenana Mountain, Alaska, earthquake may have promoted the occurrence of the $M_c = 7.9$ Denali event 10 days later. Nalbant et al. [2002] computed both secular and coseismic stresses along the East Anatolian fault and concluded that 9 of 10 events ($M > 6.7$) occurred along portions of the fault experienced positive Coulomb stress.

[26] Evidence for large event interaction has also been inferred for normal faults [Troyise et al., 1999; Cocco et al., 2000; Payne et al., 2004; Nostro et al., 2005] and between subduction and normal faulting earthquakes [Gardi et al., 2000; Mikumo et al., 2002; Robinson, 2003; Lin and Stein, 2004]. Additionally, it has been suggested that thrust and strike slip faults can interact as can thrust faults in certain circumstances [Lin and Stein, 2004]. Stress triggering of large earthquakes has also been inferred from studies of global catalogs [Parsons, 2002b; Wan et al., 2004; McKernon and Main, 2005] and has been interpreted in terms of “anomalous” or intermittent diffusion in contrast to the homogeneous (or nonintermittent) diffusion expected for processes like viscoelastic relaxation [Marsan et al., 2000; Huc and Main, 2003]. A spatial correspondence between aftershocks and positive Coulomb stress has been documented by a number of researchers. In Greece, both the 2001 $M = 6.4$ Skyros and 2003 $M = 6.2$ Lefkada earthquakes triggered off-fault aftershocks [Karakostas et al., 2003, 2004] as did the 1999 $M = 7.6$ Chi-Chi, Taiwan, event [Wang and Chen, 2001; Ma et al., 2005]. Additionally, large off-fault aftershocks of the 1998 $M = 8.1$ Antarctica earthquake are generally consistent with a model of Coulomb static stress triggering although inferences are not conclusive because of considerable uncertainties in the event’s rupture process [Toda and Stein, 2000]. In some cases the assumed relation between positive Coulomb static stress changes and the spatial distribution of aftershocks has been used to make inferences about specific rupture characteristics. Seebert and Armbruster [2000] used the observed aftershock distribution of the Landers earthquake to constrain its slip pattern, and Gahalaut et al. [2003] used the aftershock distribution of the 1993 $M_w = 6.2$ Killari, India earthquake to distinguish between six possible fault plane solutions.

[27] Various techniques have been used to quantify the correspondence between triggered seismicity and stress. For aftershocks, a widely used measure is the percentage of events for which the stress resolved onto the nodal plane(s) of the aftershocks is positive, although difficulties may arise due to nodal plane uncertainties and the fact that the apparent correspondence increases with increasing coefficient of friction since in a completely random stress field, the Coulomb stresses on the two nodal planes become less correlated as $\mu$ increases, increasing the likelihood that one plane will experience positive stress [Steacy et al., 2004]. Suggested solutions to these problems include bootstrap sampling of both pre and post main shock events [Anderson and Johnson, 1999] and computing the spatial correlation between the aftershocks and the stress map [Steacy et al., 2004].

[28] Rate changes can be computed by simply counting the number of events per time period in any box of interest and comparing with background seismicity, generally with some degree of smoothing [Toda and Stein, 2002] or by more sophisticated methods that attempt to account for the statistical properties of seismicity [Matthews and Reasenberg, 1988; Reasenberg and Simpson, 1992; Marsan and Nalbant, 2005]. Toda and Stein [2003] calculated rate changes with a smoothed 7 day running window to show that seismicity increases induced by the 26 March 1997 Kagoshima, Japan, earthquake were abruptly truncated by the occurrence of the 13 May event 4 km away. Woessner et al. [2004] obtained similar results by calculating rate changes with models based on nested modified Omori laws.

[29] Another approach to computing rate changes is the epidemic-type aftershock sequence (ETAS) model in which a statistical point process that depends on the occurrence times and magnitudes of previous events is used to calculate the probability of the occurrence of an event at any elapsed time [see Ogata, 2005, and references therein]. Such models have been used to show possible precursory quiescence [Ogata, 2001; Ogata et al., 2003; Ogata, 2005], to investigate the relative role of fluids and stress transfer in earthquake swarms [Hainzl and Ogata, 2005] and to support the idea that earthquake triggering occurs at all scales and hence there is no mechanistic differences between foreshocks, main shocks, and aftershocks [Helmstetter and Sornette, 2003].

[30] Of recent interest has been the question of the extent to which aftershocks modify the main shock induced stress field. Steacy et al. [2004] found that the correlation between Coulomb stress changes from the Landers earthquake and the spatial distribution of aftershocks was not improved by inclusion of the $M_c = 6.4$ Big Bear aftershock. However, Harris and Simpson [2002] proposed that the occurrence of the 1999 $M_c = 7.1$ Hector Mine earthquake might be explained by the stress changes due to the 1992 $M_c = 5.2$ Pisgah Crater aftershock. Similarly, Felzer et al. [2002] suggested that the Hector Mine earthquake was triggered by a chain of Landers aftershocks and, through a series of Monte Carlo simulations, concluded that secondary after- shock triggering is very common.

[31] Other evidence for the importance of small-scale earthquake triggering comes from Felzer et al. [2003], who studied nine aftershock sequences in California and found that 35–40% of all aftershocks were secondary, that the percentage of secondary aftershocks increased with time elapsed from the main shock and that these events were not constrained to occur in regions of the positive, main shock-induced Coulomb stress changes. On the basis of an analysis of the southern California catalog spanning 28 years, Helmstetter [2004] and Helmstetter et al. [2005] found that although individual large earthquakes trigger more events than do small ones, small earthquakes are as important for overall triggering because they occur more often. D. Marsan (The role of small earthquakes in redistributing crustal elastic stress, submitted to Geophysical Journal International, 2005) also used southern California data and concluded that aftershock stress perturbations are very important locally (i.e., aftershocks trigger other
aftershocks) but do not, however, significantly affect regional-scale Coulomb stress maps because the fractal clustering of seismicity means that the events only occupy a small percentage of the region.

6. Earthquake Triggering and Probabilities

[33] Our understanding of stress transfer and triggering has advanced sufficiently that it has begun to be put into practical use, in ways that affect public policy and business decisions. To date, this has happened in the form of public dissemination of probabilities of recurrence of large earthquakes over periods of tens of years [e.g., see Working Group on California Earthquake Probabilities, 2003], of aftershocks over periods of days [e.g., Wiemer, 2000], and in some places as time-dependent maps of earthquake shaking [e.g., Cramer et al., 2000].

[34] Stress transfer leads to changes in the probability of earthquake occurrence. The challenge then is to quantify the probability of failure under tectonic loading alone and then to assess how probabilities change when loading is perturbed. We summarize the state of this challenge by noting some of the approaches developed to characterize recurrence probabilities and changes to them for both large earthquakes and aftershocks.

[35] Most approaches to estimating failure probabilities require knowledge of recurrence patterns, or more specifically, observations that constrain probability models (i.e., probability density functions, time-dependent behaviors, etc.) with measures of their uncertainties. Biasi et al. [2002] describe a method for assessing the accuracy of paleoseismic data and how they may affect probability estimates. Kagan [2005] provides statistical models that quantify the degree to which various field or catalog measurements constrain estimates of earthquake recurrence. Wesson et al. [2003] describe a new approach employing Bayesian inference with seismic intensity or instrumental earthquake locations, together with geologic and seismologic data, to make quantitative estimates of the probabilities that specific past earthquakes are associated with specific faults.

[36] While many studies focus on single faults, several studies have considered the interactions among many faults throughout their seismic cycles. These rely on physical models of rupture and stress transfer, and while these may not capture all the complexity of real-world observations they benefit from the ability to sample many seismic cycles that observational data generally lack. Examples of long-term evolution of seismicity on regional-scale fault networks are given by Ward [2000], Rundle et al. [2001], King and Bowman [2003], Robinson [2004], Kagan et al. [2005], and others. Recurrence statistics may be derived from these model outputs, and used in estimating probabilities of future events either on individual faults or in some region. Turning observational or theoretical measures of earthquake recurrence into failure probabilities requires assumption of some statistical model that describes these in aggregate, in ways that incorporate our understanding of the seismic cycle and stress transfer. Many statistical distributions have been employed to characterize recurrence under ambient conditions, with one of the most recent being the Brownian passage time model [Matthews et al., 2002] that attempts to represent quantitatively temporal fluctuations in loading and relaxation during the interseismic period. Probabilities also vary spatially, and methods to quantify these variations rely largely on seismicity catalogs as proxies for failure probabilities [Stock and Smith, 2002; Wyss and Matsumura, 2002; Faenza et al., 2003].

[37] Initial attempts to include the effects of a nearby earthquake on the probability of failure of a fault considered only the permanent, or static stress change cast as a clock advance, equivalent to an effective shortening the mean recurrence time or to an advancing of the elapsed time [Working Group on California Earthquake Probabilities, 1990]. Stein et al. [1997] were perhaps the first group to suggest a strategy to incorporate time-dependent stress transfer processes, particularly those related to rate-state friction, into earthquake probabilities. Most recently, Hardebeck [2004] suggested a somewhat more general approach, which Gomberg et al. [2005a] generalize even further. These approaches are based on the idea that a fault fails with some known distribution of recurrence times, which may be perturbed predictably according to specified models of stress transfer and rupture nucleation. Huc and Main [2003] attempt to quantify earthquake occurrence correlations in time and space using global catalogs, find behaviors reminiscent of near-critical point systems in physics, and suggest analytic descriptions that may be used to calculate conditional probabilities. In addition to stress transfer associated with perturbing earthquakes, Mazzotti and Adams [2004] show that slow slip events also may modify earthquake failure probabilities. They found that episodic slow slip in Cascadia might modulate the conditional probability dramatically (30–100 times), particularly within the first few weeks following a slip event. Most recently, Nadeau and Dolenc [2005] document episodic tremor and possible aseismic slip beneath the San Andreas fault in California. The mechanisms underlying these phenomena and their impact on hazard assessments are important current research topics.

[38] Establishing the credibility of probability estimates requires studies of their robustness. Parsons [2005] examines theoretically the sensitivity of estimates of single-fault probability changes due to stress transfer to various input parameters. Most importantly, he shows that we should be confident in these estimates only when a perturbing event is very close to the fault in question and/or the tectonic stressing rate is low (i.e., the perturbing stress change is large relative to the ambient stressing rate) and when the fault is well characterized. This is consistent with the results of Kagan et al. [2005], who find that for southern California seismicity the impact of stress transfer from large earthquakes on recurrence of other faults is weak at best, although they question whether this may be an observational limitation rather than a physical consequence. After comparing empirically based probability estimates for faults in the San Francisco Bay area with more model-based ones, Reasenberg et al. [2003] emphasized the importance of considering a variety of probability models, combining estimates with appropriate weighting. In addition to sensitivity testing and comparing the outputs of various probability models, forecasts need to be tested against observations not used in the derivation of the
models. Kagan and Jackson [2000] present approaches to long-term and short-term forecasting of moderate and larger earthquakes and for testing their forecasting effectiveness, and other such tests are given by Console [2001] and Papadimitriou et al. [2001]. Cramer et al. [2000] compare time-dependent and -independent ground motion maps for California and examine what the differences are most sensitive to. The effect of introducing time dependence becomes more significant as uncertainties in model inputs shrink and as the mean recurrence period increases.

[38] Recent work on the specific case of aftershock probabilities in a main shock–aftershock sequence has assigned differing degrees of significance to the calculation of the main shock induced stress fields. McCloskey et al. [2003] proposed that Coulomb stress patterns could be used to assess the likely spatial distribution of aftershocks and suggested that the first step was to determine, in advance, the best planes onto which Coulomb stresses should be resolved; this work has been extended by Steacy et al. [2005]. Chan and Ma [2004] and Steacy et al. [2004] also suggested that spatial aftershock distributions might be estimated in near real time, and they showed that meaningful calculations of Coulomb stress could be made as soon as the earthquake rupture geometry was well constrained, independent of the details of the slip on the rupture plane. Toda et al. [2005] looked at the spatial distribution of aftershocks, but they suggested that Coulomb stress changes should be combined with background seismicity rates to constrain areas likely, and unlikely, to experience aftershocks.

[39] Other authors have suggested that spatially varying aftershock probabilities can be computed without calculating Coulomb stress changes. Wiemer [2000] and Wiemer et al. [2002] proposed that observations of the initial portions of aftershock sequences could be used to directly estimate the probabilities of later events. To do this, they calculated a, b, and p (the latter for the modified Omori law) seismicity parameters on a regular spatial grid and computed the time-dependent earthquake probabilities at each grid node. They then translated these results into maps of peak ground acceleration and probability of exceedance. Using a different method, Felzer et al. [2003] suggested that forecasts of late aftershocks based on earlier events outperformed similar forecasts based on the main shock induced stress perturbation.

[40] While our focus here is on the scientific aspects of time-dependent seismic hazard, attempts also are being made to improve how developments are conveyed to the public and other users. One example is the extensive and varied documentation describing earthquake probabilities estimated for the San Francisco Bay area (available at http://quake.wr.usgs.gov/research/seismology/wg02/). Another approach actually involves the public in the hazard calculations directly. This is exemplified in the “Open Seismic Hazard Assessment” online tools, which may be used to test a wide variety of earthquake probability and other models separately or in concert [Field et al., 2003].

7. Speculations on the Future

[41] Indeed, much progress has been made in the last decade on the topics of stress transfer, earthquake triggering, and time-dependent seismic hazard. We conclude our introduction to the special section by highlighting some of the most interesting, yet challenging questions that the next decade’s research may answer given emerging modeling capabilities, data types, and perhaps most importantly, ideas. Examples of very basic questions concern the existence of stress shadows, the nature of triggering thresholds (e.g., do they exist, are they binary?), mechanisms explaining delays between stress changes and failure, the significance of aseismic slip and harmonic tremor, and the physical processes and environmental conditions that control the potential for triggering, including the proper way to account for secular loading in different tectonic settings.

[42] While compelling correlations between static stress changes and aftershock distributions provided the earliest evidence of triggering [e.g., Smith and Van deLindt, 1969] the last decade’s research has revealed the probable complexity of triggering processes in the near field. In the next decade we surely will see new near-field observational constraints from high-dynamic range, broad bandwidth recordings of the deformation field generated by triggering earthquakes and acting on potentially triggered faults. These will come from Earthscope deployments (see http://earthscope.org), dense seismic and geodetic networks in the United States (e.g., see http://www.amss.org), Japan (e.g., http://www.bosai.go.jp/center/index_e.html), and elsewhere. Improved instrumentation and monitoring networks should continue to reveal details of triggering phenomena with ever higher fidelity, such as definitive observations of earthquakes (or absence of) triggered dynamically within the wave train radiated from the triggering event. We also will benefit simply from the passage of time, as the Earth provides more examples of major and great earthquakes. Another, relatively underutilized, potential source of new observations is the laboratory. While abundant studies of frictional sliding and to a lesser degree, stick-slip behavior under constant loading have been conducted, few have focused on triggering under the wide range of loading conditions that occur in nature and mechanisms of stress transfer (see Beeler and Lockner [2003] and Boettcher and Marone [2004] for examples of nonconstant loading studies and references to constant loading studies). Experiments under conditions more appropriate to seismogenic depths also are needed and should be possible as technologies advance.

[43] The last decade has seen greater emphasis on understanding the physics underlying earthquake triggering and the time dependence of earthquake failure, rather than simply compiling observational evidence. As noted above, the testing and application of frictional failure models have reached quite mature states, particularly when considering triggering as clock advancing inevitable earthquakes. However, these and most other models still fail to explain delayed dynamic triggering, and little work has been done to understand if and how triggering may promote earthquake occurrence on faults that otherwise would be inactive. Just as triggering loads vary tremendously in their temporal and spatial characteristics, we need to ask whether a single mechanism (and model) of stress transfer leading to failure exists or whether many mechanisms operate.

[44] Recent advances in computing have led to models that account for more and more of the complexities that
exist in nature. The trend of modeling development suggests that within years it will be possible to conduct computer simulations that model processes that operate on timescales of a single rupture (seconds) to many hundreds of years, and spatial scales that span nucleation patch dimensions to regional fault networks. Such models will allow us to test our ideas of fault interaction in a wide variety of settings and over long timescales. The challenge, of course, will be validating the extent to which they represent the essential physics of the earthquake process and hence the confidence with which model results can be used to inform earthquake probabilities.

[45] Acknowledgments. We thank Francis Albarede for making the special section possible, the authors and reviewers of the manuscripts within it for all their hard work, Agnes Ganivet for keeping an eye on deadlines, and Jeanette Panning for heroic efforts to keep everything on track. Ruth Harris, Bob Simpson, Karen Felzer, the AE Stephanie Prejean, and the foremen provided detailed and constructive reviews that improved this introduction.

References


Hill, D. P., et al. (1993), Seismicity remotely triggered by the magnitude 7.3 Landers, California, earthquake, Science, 260, 1617–1623.


Hough, S. E. (2005), Remotely triggered earthquakes following moderate mainshocks (or, why California is not falling into the ocean), Seismol. Res. Lett., 76(1), 58–66.


