Reports on Progress in **Physics**

REVIEW

The complex dynamics of earthquake fault systems: new approaches to forecasting and nowcasting of earthquakes

To cite this article: John B Rundle *et al* 2021 *Rep. Prog. Phys.* **84** 076801

Reports on Progress in Physics

The Complex Dynamics of Earthquake Fault Systems: New Approaches to Forecasting and Nowcasting of Earthquakes

by

John B. Rundle[1,2,3] Seth Stein[4], Andrea Donnellan[5], Donald L Turcotte[2], William Klein[6] and Cameron Saylor[1]

[1]Department of Physics and Astronomy
[2]Department of Earth & Planetary Sciences
University of California
Davis, CA 95616

[3]Santa Fe Institute
1399 Hyde Park Rd
Santa Fe, NM 87501

[4]Department of Earth and Planetary Sciences
and Institute for Policy Research
Northwestern University
Evanston, IL 60208

[5]Jet Propulsion Laboratory, California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109

[6]Department of Physics
Boston University
Boston, MA 02215

Abstract

Charles Richter's observation that "only fools and charlatans predict earthquakes," reflects the fact that despite more than 100 years of effort, seismologists remain unable to do so with reliable and accurate results. Meaningful prediction involves specifying the location, time, and size of an earthquake before it occurs to greater precision than expected purely by chance from the known statistics of earthquakes in an area. In this context, "forecasting" implies a prediction with a specification of a probability of the time, location, and magnitude. Two general approaches have been used. In one, the rate of motion accumulating across faults and the amount of slip in past earthquakes is used to infer where and when future earthquakes will occur and the shaking that would be expected. Because the intervals between earthquakes are highly variable, these long-term forecasts are accurate to no better than a hundred years. They are thus valuable for earthquake hazard mitigation, given the long lives of structures, but have clear limitations. The second approach is to identify potentially observable changes in the earth that precede earthquakes. Various precursors have been suggested, and may have been real in certain cases, but none have yet proved to be a general feature preceding all earthquakes or to stand out convincingly from the normal variability of the earth's behavior. However, new types of data, models, and computational power may provide avenues for progress using machine learning that were not previously available. At present, it is unclear whether deterministic earthquake prediction is possible. The frustrations of this search have led to the observation that (echoing Yogi Berra) "it is difficult to predict earthquakes, especially before they happen." However, because success would be of enormous societal benefit, the search for methods of earthquake prediction and forecasting will likely continue. In this review, we note that the focus is on anticipating the earthquake rupture before it occurs, rather than characterizing it rapidly just after it occurs. The latter is the domain of earthquake early warning, which we do not treat in detail here, although we include a short discussion in the machine learning section at the end.

1. Introduction

In order to provide an impartial test of earthquake prediction the United States Geological Survey initiated the Parkfield (California) Earthquake Prediction Experiment in 1985 (Bakun and Lindh, 1985). Earthquakes on this section of the San Andreas fault had occurred in 1847, 1881, 1901, 1922, 1934, and 1966. It was expected that the next earthquake in the sequence would occur in 1988 ± 5 years. An extensive array of instrumentation was deployed. The expected earthquake finally occurred on September 28, 2004. No precursory observations outside the expected background levels were observed (Bakun et al., 2005). The earthquake had not been predicted. This result has been interpreted to imply the infeasibility of deterministic short-term prediction of earthquakes on a consistent basis.

Successful near-term predictions of future earthquakes, which have happened on occasion, are very limited. A notable exception was the M = 7.3 Haicheng earthquake in northeast China that occurred on February 4, 1975. This prediction resulted in the evacuation of the city which saved many lives. It was reported that the prediction was based on foreshocks, groundwater anomalies and animal behavior. It should be noted, however, that no prediction was made prior to the occurrence of the M = 7.8 Tangshan earthquake in China on July 28, 1976. Reports suggest the death toll in this case was as high as 600,000.

It seems surprising that it is not possible to make reliable short-term predictions of the occurrence of a major earthquake (Kanamori, 2003; Keilis-Borok, 2002; Mogi, 1985; Scholz, 2019; Turcotte, 1991). Based on analog laboratory experiments, precursory micro cracking expressed as small earthquakes should occur, and precursory strain would also be expected. Foreshocks occur prior to about 25% of major earthquakes, but it is difficult to distinguish foreshocks from background seismicity since they are all “just earthquakes”.

An important recent development in this area was the Regional Earthquake Likelihood Models (RELM) test of earthquake forecasts in California. Forecasts had to be submitted prior to the start of the evaluation period, so this was a true prospective evaluation. Six participants submitted forecasts for 7,700 cells. Two of the forecasts showed promise, these being the Pattern Informatics (PI) and Epidemic Type Aftershock Sequence (ETAS) forecasts. We discuss this competition, among many other topics, below. But first we describe the history of earthquake prediction studies and the search for reliable precursors.

2. History of Earthquake Precursor Studies

In the 1960s and 1970s, well-funded government earthquake prediction programs began in the US, China, Japan, and the USSR. These programs relied on two approaches. One, based on laboratory experiments showing changes in the physical properties of rocks prior to fracture, involved searching for precursors or observable behavior that precedes earthquakes. A second was based on the idea of the seismic cycle, in which strain accumulates over time following a large earthquake. Hence areas on major faults that had not had recent earthquakes could be considered “seismic gaps” likely to have large earthquakes.

The idea that earthquake prediction was about to become reality was promoted heavily in the media. US Geological Survey director William Pecora announced in 1969 “We are predicting another massive earthquake certainly within the next 30 years and most likely in the next decade or so" on the San Andreas fault. California senator Alan Cranston, prediction’s leading political supporter, told reporters that “we have the technology to develop a reliable prediction system already at hand.” Although the President's science advisor questioned the need for an expensive program given the low death rate from earthquakes, lobbying prevailed and funding poured into the US program and similar programs in other countries.

To date this search has proved generally unsuccessful. As a result, it is unclear whether earthquake prediction is even possible. In one hypothesis, all earthquakes start off as tiny earthquakes, which happen frequently. However, only a few cascade via a failure process into large earthquakes. This hypothesis draws on ideas from nonlinear dynamics or chaos theory, in which small perturbations can grow to have unpredictable large consequences. These ideas were posed in terms of the possibility that the flap of a butterfly’s wings in Brazil might set off a tornado in Texas, or in general that minuscule disturbances do not affect the overall frequency of storms but can modify when they occur (Lorenz, 1995). In this view, because there is nothing special about those tiny earthquakes that happen to grow into large ones, the interval between large earthquakes is highly variable, and no observable precursors should occur before them. If so, earthquake prediction is either impossible or nearly so.

Support for this view comes from the failure to observe a compelling pattern of precursory behavior before earthquakes (Geller, 1997). Various possible precursors have been suggested—and some may have been real in certain cases—but none have yet proved to be a general feature preceding all earthquakes, or to stand out convincingly from the normal range of the earth's variable behavior. In many previous cases, it was not realized that a successful prediction scheme must allow not only for successful predictions, but also failures-to-predict and false alarms. Although it is tempting to note a precursory pattern after an earthquake based on a small set of data and to suggest that the earthquake might have been predicted, rigorous tests with large sets of data are needed to tell whether a possible precursory behavior is real, and whether it correlates with earthquakes more frequently than expected purely by chance. In addition, after-the-fact searches for precursors have the advantage that one knows where to look. Most crucially, any such pattern needs to be tested by predicting future earthquakes. That is why the RELM test of earthquake forecasts (discussed below) was a significant advance.

One class of precursors involves foreshocks, which are smaller earthquakes that occur before a main shock, actually a semantic definition. Many earthquakes, in hindsight, have followed periods of anomalous seismicity. In some cases, there is a flurry of microseismicity, which are very small earthquakes similar to the cracking that precedes the snap of a bent stick. In other cases, there is no preceding seismicity of any significance. However, faults often show periods of either elevated ("activation") or nonexistent ("quiescent") microseismicity that are not followed by a large earthquake. Alternatively, the level of microseismicity before a large event can be unremarkable, occurring at a normal low level. The lack of a pattern highlights the problem with possible earthquake precursors. To date, no changes that might be associated with an upcoming earthquake are consistently distinguishable from the normal variations in seismicity that are not followed by a large earthquake.

Another class of possible precursors involves changes in the properties of rock within a fault zone preceding a large earthquake. It has been suggested that as a region experiences a buildup of elastic stress and strain, microcracks may form and fill with water, lowering the strength of the rock and eventually leading to an earthquake. This effect has been advocated based on data showing changes in the level of radon gas, presumably reflecting the development of microcracks that allow radon to escape. For example, the radon detected in groundwater rose steadily in the months before the moment magnitude Mw 6.9, 1995 Kobe earthquake, increased further two weeks before the earthquake, and then returned to a background level.

A variety of similar observations have been reported. In some cases, the ratio of *P*- and *S*-wave speeds in the region of an earthquake has been reported to have decreased by as much as 10% before an earthquake. Such observations would be consistent with laboratory experiments and would reflect cracks opening in the rock (lowering wave speeds) due to increasing stress and later filling (increasing wave speeds). However, this phenomenon has not been substantiated as a general phenomenon. Similar difficulties beset reports of a decrease in the electrical resistivity of the ground before some earthquakes, consistent with large-scale microcracking. Changes in the amount and composition of groundwater have also been observed. For example, a geyser in Calistoga, California, changed its period between eruptions before the Mw 6.9 1989 Loma Prieta and Mw 5.7 1975 Oroville, California, earthquakes.

Efforts have also been made to identify ground deformation immediately preceding earthquakes. The most famous of these studies was the report in 1975 of 30–45 cm of uplift along the San Andreas fault near Palmdale, California. This highly publicized “Palmdale Bulge” was interpreted as evidence of an impending large earthquake and was a factor in the US government’s decision to launch the National Earthquake Hazards Reduction Program aimed at studying and predicting earthquakes. US Geological Survey director Vincent McKelvey expressed his view that “a great earthquake” would occur “in the area presently affected by the … ‘Palmdale Bulge’... possibly within the next decade” that might cause up to 12,000 deaths, 48,000 serious injuries, 40,000 damaged buildings, and up to $25 billion in damage. The California Seismic Safety Commission stated that “the uplift should be considered a possible threat to public safety” and urged immediate action to prepare for a possible disaster. News media joined the cry.

In the end, the earthquake did not occur, and reanalysis of the data implied that the bulge had been an artifact of errors involved in referring the vertical motions to sea level via a traverse across the San Gabriel mountains. It was realized that the apparent bulging of the ground was produced by a combination of three systematic measurement errors. They were necessarily systematic in space and in time. The culprits were (i) atmospheric refraction errors that made hills look too small, steadily declining as sight lengths were reduced from one levelling survey to the next, which made the hills appear to rise, (ii) the misinterpretation of subsidence due to water withdrawal. Saugus was subsiding, as opposed to the areas surrounding Saugus uplifting! (iii) the inclusion of a bad rod in some of the leveling surveys. These discoveries were led by (i) Bill Strange of the NGS, (ii) Robert Reilinger (then at Cornell) and (iii) David Jackson at UCLA. (Rundle and McNutt, 1981; M. Bevis, personal communication, 2020).

Hence the Bulge changed to “the Palmdale soufflé – flattened almost entirely by careful analysis of data” (Hough, 2007). Subsequent studies elsewhere, using newer and more accurate techniques including the Global Positioning System satellites, satellite radar interferometry, and borehole strainmeters have not yet convincingly detected precursory ground deformation.

An often-reported precursor that is even harder to quantify is anomalous animal behavior. What the animals are sensing (high-frequency noise, electromagnetic fields, gas emissions) is unclear. Moreover, because it is hard to distinguish “anomalous” behaviors from the usual range of animal behaviors, most such observations have been “postdictions,” coming after rather than before an earthquake.

Chinese scientists have attempted to predict earthquakes using precursors. Chinese sources report a successful prediction in which the city of Haicheng was evacuated in 1975, prior to a magnitude 7.4 earthquake that damaged more than 90% of the houses. The prediction is said to have been based on precursors, including ground deformation, changes in the electromagnetic field and groundwater levels, anomalous animal behavior, and significant foreshocks. However, in the following year, the magnitude 7.8 Tangshan earthquake occurred not too far away without precursors. In minutes, 250,000 people died, and another 500,000 people were injured. In the following month, an earthquake warning in the Kwangtung province caused people to sleep in tents for two months, but no earthquake occurred. Similarly, no anomalous behavior was identified before the magnitude 7.9 Wenchuan earthquake in 2008. Because foreign scientists have not yet

been able to assess the Chinese data and the record of predictions, including both false positives (predictions without earthquakes) and false negatives (earthquakes without predictions), it is difficult to evaluate the program.

Despite tantalizing suggestions, at present there is still an absence of reliable precursors. Most researchers thus feel that although earthquake prediction would be seismology's greatest triumph, it is either far away or will never happen. However, because success would be of enormous societal benefit, the search for methods of earthquake prediction continues. To further this point, we now consider the famous Parkfield earthquake prediction experiment.

2.1. The Search for Earthquake Precursors

2.1.1 The Parkfield Earthquake Prediction Experiment. Even with the dates of previous major earthquakes, it is difficult to predict when the next one will occur, as illustrated by the segment of the San Andreas fault near Parkfield, California, a town of about 20 people whose motto is "Be here when it happens." Earthquakes of magnitude 5–6 occurred in 1857, 1881, 1901, 1922, 1934, and 1966. The average recurrence interval is 22 years, and a linear fit to these dates made 1988 ± 5 years the likely date of the next event. In 1985, it was predicted at the 95% confidence level that the next Parkfield earthquake would occur before 1993, which was the USA's first official earthquake prediction (Bakun et al., 2005).

Seismometers, strainmeters, creepmeters, GPS receivers, tiltmeters, water level gauges, electromagnetic sensors, and video cameras were set up to monitor what would happen before and during the earthquake. The US National Earthquake Prediction Evaluation Council endorsed the highly publicized $20 million "Parkfield" project. The *Economist* magazine commented, "Parkfield is geophysics' Waterloo. If the earthquake comes without warnings of any kind, earthquakes are unpredictable, and science is defeated. There will be no excuses left, for never has an ambush been more carefully laid."

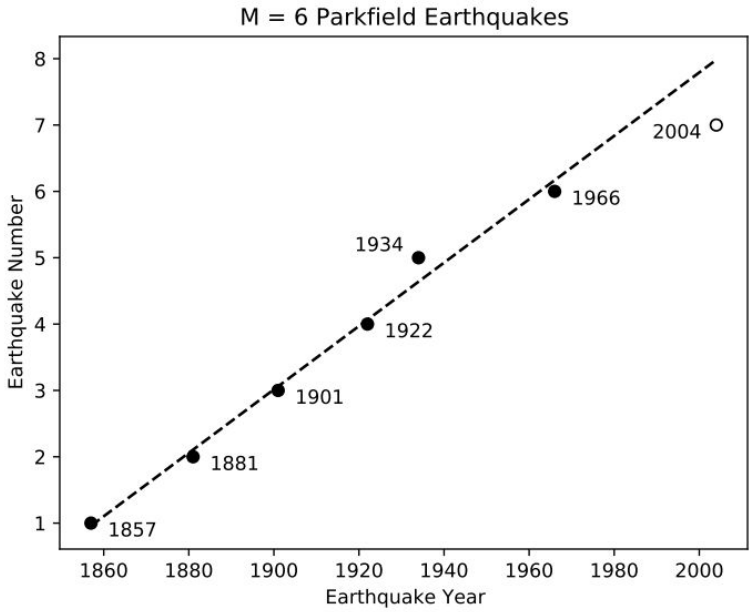

Figure 1. The Parkfield, CA earthquake that was predicted to occur within 5 years of 1988 did not occur until 2004. Black dots show when the earthquake occurred, and the best-fitting line indicates when they should have occurred at intervals of 22 years.

Exactly that happened. The earthquake did not occur by 1993, leading *Science* magazine to conclude, “Seismologists’ first official earthquake forecast has failed, ushering in an era of heightened uncertainty and more modest ambitions” (Kerr, 1993). A likely explanation was that the uncertainty in the repeat time had been underestimated by discounting the fact that the 1934 earthquake did not fit the pattern well (Figure 1) (Savage, 1993).

An earthquake eventually occurred near Parkfield on September 28, 2004, eleven years after the end of the prediction window, with no detectable precursors that could have led to a short-term prediction (Kerr, 2004). It is unclear whether the 2004 event should be regarded as the predicted earthquake coming too late, or just the next earthquake on that part of the fault.

For that matter, we do not know whether the fact that earthquakes occurred about 22 years apart reflects an important aspect of the physics of this particular part of the San Andreas, or just an apparent pattern that arose by chance given that we have a short history and many segments of the San Andreas of similar length. After all, flipping a coin enough times will give some impressive-looking patterns of heads or tails. With only a short set of data, we could easily interpret significance to what was actually a random fluctuation and thus be “Fooled by Randomness” (e.g., Taleb, 2004). It is possible the 1983 Mw 6.4 Coalinga earthquake (Kanamori, 1983) was the “missing” Parkfield event suggesting that earthquake forecasting should be based on regional spatial and temporal scales (Tiampo et al., 2005) rather than fault based. As is usual with such questions, only time will tell.

2.1.2 Load Unload Response Ratio (LURR). LURR is a method that was developed in China and is widely used by the Institute of Earthquake Forecasting of the China Earthquake Administration in the official yearly earthquake forecast of China, which is required by law in that country (Yin et al., 1995; 2004; Yuan et al., 2010; Zhang et al., 2008) . However, this method has not been widely researched or used in other countries. The basic idea is that tidal stresses on the earth can be used as a diagnostic for the state of increasing stress prior to a major earthquake. Tidal stresses are cyclic, so in principle there should exist an observable asymmetry in the response to the daily tidal stressing. Based on similar observations of acoustic emissions from stressed rock samples in the laboratory, it might be expected that microseismicity would be higher in the increasing stress part of the cycle, and lower in the de-stressing part (Zhang et al., 2008). However, data from actual field experiments are controversial (Smith and Sammis, 2004).

2.1.3 Accelerating Moment Release (AMR). AMR was a method that was based on the hypothesis that, prior to fractures in the laboratory and the field, there should be a precursory period of accelerating seismicity or slip, otherwise known as a “runup to failure” (Varnes, 1989; Bufe and Varnes, 1993; Bufe et al., 1994; Jaume and Sykes, 1999; Bowman and King, 2001; Sornette and Sammis, 1995). In these papers, it was proposed that the accelerating slip is characterized by a power law in time, therefore suggesting the possible existence of critical point

dynamics. While based on a reasonable physical hypothesis of a cascading sequence of material failure, the phenomenology has so far not been observed in natural fault systems (Guilheim et al., 2013).

3. Basic Equations of Earthquake Science

3.1 Observational laws. The oldest observational laws of earthquake science are the Gutenberg-Richter (GR) magnitude-frequency relation and the Omori-Utsu relation for aftershock decay. The GR relation states that the number of earthquakes *M* larger than a value *m* in a given region and over a fixed time interval is given by:

$$N(M \geq m) = 10^{a} 10^{-bm} \tag{1}$$

Here a and b are constants that depend on the time interval and the geographic region under consideration. Note further that equation (1) is not a density function, rather it is a survivor function or exceedance function.

The magnitude m was originally defined in terms of a local magnitude developed by C.F. Richter, but the magnitude m is now most commonly determined by the seismic moment *W*:

$$1.5\, m = \log_{10} W - 9.0 \tag{2}$$

Expression (2) is in S.I notation. The quantity *W* is found from matching observational seismic timeseries obtained from seismometers to a model involving a pair of oriented and opposed double-couples (dipoles), thus giving rise to a quadripolar radiation pattern. Its scalar value is given by:

$$\mathrm{W} = \mu\ \Delta u\ A \tag{3}$$

where μ is the elastic modulus of rigidity, Δu is the average slip (discontinuity in displacement) across the fault, and *A* is the slipped area. Combining equations (1) and (2), we find that in fact, the GR law is a scaling relation (power law):

$$N = 10^{a}\ W^{\left\{-\frac{2b}{3}\right\}} \tag{4}$$

where *m* is given by (2).

The remaining equation is the Omori-Utsu law (e.g., Scholz, 2019), which was proposed by Omori following the 1891 Nobi, Japan earthquake, surface wave magnitude Ms = 8.0, and expresses the temporal decay of the number of aftershocks following a main shock:

$$\frac{dN}{dt} = \frac{K}{(c+t)^{p}}$$

(5)

In (5), *p*, *K* and *c* are constants, to be found by fitting equation (4) to the observational data, and *t* is the time elapsed since the main shock. In its original form, the constant *c* was not present, it was later added by Utsu to regularize the equation at *t* = 0. An example of Omori-Utsu decay is shown in Figure 3 below.

3.2 Elasticity. Models to develop and test earthquake nowcast and forecast methodologies are based on the known processes of brittle fracture in rocks, typically modeled as a shear fracture in an elastic material. The equations of linear elasticity are used to describe the physics of the process. Most of the models used for nowcasting, forecasting and prediction are either statistical or elastostatic, where seismic radiation is neglected. The motivation for this approach is the focus on the slow processes leading up to the rupture, rather than on details of the rupture dynamics.

To understand the basic methods, let us define a stress tensor in *d* = 3:

$$\sigma(\boldsymbol{x},t) = \sigma_{ij}(\boldsymbol{x},t) \tag{6}$$

and a strain tensor:

$$\varepsilon(\boldsymbol{x},t) = \varepsilon_{ij}(\boldsymbol{x},t) = \frac{1}{2}\left(\frac{\partial u_i(\boldsymbol{x},t)}{\partial x_j} + \frac{\partial u_j(\boldsymbol{x},t)}{\partial x_i}\right) \tag{7}$$

where the (infinitesimal) displacement in the elastic medium at location $\boldsymbol{x}$ and time t is $\boldsymbol{u}(\boldsymbol{x},t) = u_i(\boldsymbol{x},t)$.

To relate the stress tensor to the strain tensor, the simplest assumption is to use the constitutive law for isotropic linear elasticity:

$$\sigma_{ij}(\boldsymbol{x},t) = \delta_{ij}\,\lambda\,\varepsilon_{kk}(\boldsymbol{x},t) + 2\,\mu\,\varepsilon_{ij}(\boldsymbol{x},t) \tag{8}$$

where δ_{ij} is the Kronecker delta, μ and λ are the Lamé constants of linear elasticity, and repeated indices are summed.

The equation of elastic equilibrium can be stated in the form:

$$\nabla \cdot \boldsymbol{\sigma}(\boldsymbol{x},t) = \boldsymbol{f}\,(\boldsymbol{x},t) \tag{9}$$

where $\boldsymbol{f}(\boldsymbol{x}, t)$ is a body force. It has been shown that the body force appropriate to shear slip on a fault can be found by the following method.

For a fault element at position $\boldsymbol{x}'$, we wish to find the displacement and stress at position $\boldsymbol{x}$, i.e., we wish to find the Green's functions. To do so, we let:

$$\boldsymbol{u}(\boldsymbol{x}-\boldsymbol{x}',t) = \hat{\boldsymbol{u}}(t)\,\delta(\boldsymbol{x}-\boldsymbol{x}') \tag{10}$$

where $\hat{\boldsymbol{u}}$ is a unit vector, and $\delta(\boldsymbol{x})$ is the Dirac delta function. We then compute the strain according to equation (7), followed by taking the divergence of that strain tensor. The result is solutions (Green's functions) of the form:

$$u_i(\boldsymbol{x},t) = \int G_{ikl}(\boldsymbol{x}-\boldsymbol{x}')\,\Delta u_k(\boldsymbol{x}',t)\,n_l\,dA \tag{11}$$

$$\sigma_{ij}(\boldsymbol{x},t) = \int T_{ijkl}(\boldsymbol{x}-\boldsymbol{x}')\,\Delta u_k(\boldsymbol{x}',t)\,n_l\,dA$$

Here $G_{ikl}(\boldsymbol{x}-\boldsymbol{x}')$ is the displacement Green's function, and $T_{ijkl}(\boldsymbol{x}-\boldsymbol{x}')$ is the stress Green's function. As before, $\Delta u_k(\boldsymbol{x}')$ is the displacement discontinuity across the fault in kth direction at $\boldsymbol{x}'$, and n_l is the unit normal to the fault, with dA being the element of fault area. Note that the detailed construction of these Green's functions for point and rectangular faults can be found in many papers, with the most widely used version being found in the paper by Okada (1992).

In many of the simple models used to describe a single planar fault, such as the slider block models discussed below, a spatial coarse graining is used to subdivide a fault into a partition of squares of area ΔA, each square representing a slider block. Then, for example, we can write the force or stress on a slider block (= element of area) schematically as:

$$\sigma_i(t) = \sum_j T_{ij} s_j(t) \tag{12}$$

where the T_{ij} are combinations of spring constants (discussed below).

4. Complexity and Earthquake Fault Systems

Most researchers have now set their sights on probabilistic earthquake forecasting, rather than deterministic earthquake prediction using precursors. The focus is presently on determining whether forecasts covering months to years are feasible, in addition perhaps to decadal time scales.

To understand the causes of the problems noted above, we now briefly turn to an analysis of the structure and dynamics of earthquake faults systems, and how these may influence our ability to make reliable earthquake forecasts. We begin with a discussion of the idea of earthquake cycles and supercycles. In the context of complex systems, earthquake cycles can be related to the idea of limit cycles in complex systems. In this case a limit cycle is a repetitive behavior—a

simple example would be a sine wave. Of course, limit cycles are one manifestation of the dynamics of complex systems, others being chaotic and fixed-point dynamics.

4.1 Earthquake Cycles and Supercycles. Since the Mw 7.9 San Francisco earthquake of April 18, 1906, the dominant paradigm in earthquake seismology has been the earthquake cycle, in which strain accumulates between large earthquakes due to motion between the two sides of a locked fault. That strain is released by slip on the fault when an earthquake occurs (Reid, 1910). Over time, this process should conceptually give rise to approximately periodic earthquakes and a steady accumulation of cumulative displacement across the fault.

However, long records of large earthquakes using paleoseismic records—geological data spanning thousands of years or more—often show more complex behavior, as reviewed by Salditch et al. (2020). The earthquakes occurred in supercycles, sequences of temporal clusters of seismicity, cumulative displacement, and cumulative strain release separated by intervals of lower levels of activity.

Supercycles pose a challenge for earthquake forecasting because such long-term variability is difficult to reconcile with commonly used models of earthquake recurrence (Stein and Wysession, 2009). In the Poisson model earthquakes occur randomly in time and the probability of a large earthquake is constant with time, so the fault has no memory of when the previous large earthquake occurred. In a seismic cycle or renewal model, the probability is quasi-periodic, dropping to zero after a large earthquake, then increasing with time, so the probability of a large earthquake depends only on the time since the past one, and the fault has only "short-term memory."

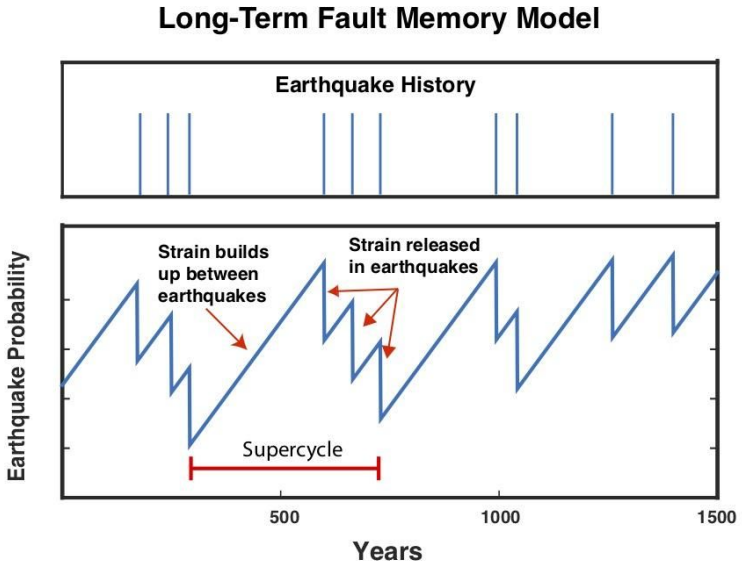

Figure 2. Long-Term Fault Memory model. (Top) Simulated earthquake history. (Bottom) Earthquake probability versus time. (Salditch et al., 2020)

This situation suggests that faults have "long-term memory," such that the occurrence of large earthquakes depends on earthquake history over multiple previous earthquake cycles (Figure 2). Faults having long-term memory would have important consequences for earthquake forecasting. Weldon et al. (2004) point out that:

"...resetting of the clock during each earthquake not only is conceptually important but also forms the practical basis for all earthquake forecasting because earthquake recurrence is statistically modeled as a renewal process... In a renewal process, intervals between earthquakes must be unrelated so their variability can be expressed by (and conditional probabilities calculated from) independent random variables. Thus, if the next earthquake depends upon the strain history prior to that earthquake cycle, both our understanding of Earth and our forecasts of earthquake hazard must be modified... there can be little doubt that the simple renewal model of an elastic rebound driven seismic cycle will need to be expanded to accommodate variations that span multiple seismic cycles."

A simple model for supercycles, Long-Term Fault Memory (LTFM), extends the standard earthquake cycle model. It assumes that the probability of a large earthquake reflects the accumulated strain rather than elapsed time. The probability increases as strain accumulates over time until an earthquake happens, after which it decreases, but not necessarily to zero. Hence, the probability of an earthquake depends on the earthquake history over multiple prior cycles.

LTFM is a stochastic process, a Markov chain with states at discrete times corresponding to values of accumulated strain, reflected in the probability $P(t)$. The probability that an earthquake occurs at time t, conditional on the history of strain accumulation and release at prior times, depends only on the most recent level of strain at time $t-1$. Given $P(t)$, the probability does not otherwise depend on time, so the history prior to t is fully captured by $P(t-1)$.

LTFM can also be posed using the classic probability model of drawing balls from an urn (Stein and Stein, 2013). If some balls are labeled "E" for earthquake and others are labeled "N" for no earthquake, the probability of an earthquake is that of drawing an E-ball, the ratio of the number of E-balls to the total number of balls. If after drawing a ball, we replace it, the probability of an earthquake is constant or time-independent in successive draws, because one happening does not change the probability of another happening.

Thus, an earthquake is never "overdue" because one has not happened recently, and the fact that one happened recently does not make another less likely. LTFM corresponds to an alternative sampling such that the fraction of E-balls and the probability of another event change with time. We add E-balls after a draw when an earthquake does not occur and remove E-balls when one occurs. Thus, the probability of an earthquake increases with time until one happens, after which it decreases and then grows again. Earthquakes are not independent, because one happening changes the probability of another.

Viewing supercycles as a result of long-term fault memory fits into a general framework in the literature of complex dynamical systems. Clustered events, described as “bursts,” are observed in many disparate systems, from the firing system of a single neuron to an outgoing mobile phone sequence (Karsai et al., 2012; Rundle and Donnellan, 2020, discussed below). Such systems display “…*a bursty, intermittent nature, characterized by short timeframes of intense activity followed by long times of no or reduced activity*,” (Goh and Barabasi, 2008). The system’s state depends on its history, so it has long-term memory (Beran et al., 2013).

LTFM simulations over timescales corresponding to the duration of paleoseismic records find that the distribution of earthquake recurrence times can appear strongly periodic, weakly periodic, Poissonian, or bursty. Thus, a given paleoseismic window may not capture long-term trends in seismicity. This effect is significant for earthquake hazard assessment because whether an earthquake history is assumed to contain clusters can be more important than the probability density function chosen to describe the recurrence times. In such cases, probability estimates of the next earthquake will depend crucially on whether the cluster is treated as ongoing or finished.

4.2 Interactions and Scales. Complex nonlinear systems are characterized by many interacting agents, each agent having some type of nonlinear behavior, as well as interactions with other agents. They have a multiplicity of scales in space and time, form coherent space-time structures by means of their internal dynamics, have nonlinear threshold dynamics, and are typically driven and dissipative (e.g., Rundle et al, 2003). Examples of these types of systems include markets and the economy, evolutionary, biological and neural systems, the internet, flocking of birds and schooling of fish, earthquakes, and many more (Rundle et al., 2019). None of these systems evolve according to a central plan. Rather, their dynamics are guided by a few basic bottom-up principles rather than a top-down organizational structure.

In the example of earthquakes, these faults are embedded in complex geomaterials, and are driven by slowly accumulating tectonic forces, or, in the case of induced seismicity, by injection of fracking fluids. Rocks make up the earth's crust, and are disordered solids having a wide range of scales, both structurally and dynamically as they deform (Turcotte and Shcherbakov, 2006; Turcotte et al., 2003).

On the microscopic (micron) scale, dislocations and lattice defects within grains represent important contributors to solid deformation. On larger scales (millimeter), grain dynamics including shearing, microcrack formation, and changes in the porosity matrix contribute. On still larger scales (centimeters to meters and larger), macroscopic fracturing in tension and shear, asperity de-pinning, and other mechanical processes lead to observable macroscopic deformation. On the largest (tectonic) scales, the self-similarity is also manifested as a wide array of earthquake faults on all scales, from the local to the tectonic plate scale of thousands of km (Scholz, 2019).

Observations of rock masses over this range of spatial scales indicate that the failure modes of these systems, such as fracture and other forms of catastrophic failure demonstrate scale invariant deformation, or power law behavior, characteristic of complex non-linear systems. These are observed in both laboratory settings in acoustic emission experiments, as well as in large scale field settings associated with tectonic faults (Gutenberg-Richter magnitude-frequency relation; Omori relation for aftershocks). One important reason for this behavior is that driven

threshold systems of rock masses in which defects interact with long range interactions display near mean field dynamics and ergodic behavior (Rundle and Klein, 1989; Rundle et al., 1996; Klein et al., 1997; 2000a, b, c; 2007; Tiampo et al., 2002a). This result, which was first proposed based on simulations and theory, was subsequently observed in field observations on the tectonic scale (Tiampo et al., 2002b).

In both laboratory and field scale settings, a wide variety of timescales are also observed (Scholz, 2019). These include the source-process time scale of seconds to minutes on which earthquakes occur, as well as the loading time scales of tens to hundreds to thousands of years on which earthquakes recur in active tectonic regions. Other phenomena, to be discussed below, such as small earthquake swarms and other thermal and physical processes, operate on time scales as short as days, and as long as months to years (Scholz, 2019).

Modeling these types of processes requires consideration of fully interacting fields of dislocations, defects, damage, and other material disorder. In much of the previous work over the last decades on these types of systems, disordered fields were assumed to be non-interacting, allowing classical solid-solid mixture theories to be employed (e.g., Hashin and Shtrickman, 1962). With respect to earthquake faults, it was only emphasized within the last few decades that earthquake faults interact by means of transmission of tectonic stress, mediated by the presence of the brittle-elastic rocks within which the faults are embedded.

With the development of new high-performance computing hardware and algorithms, together with new theoretical methods based on statistical field theories, we can now model a wide variety of fully interacting disordered systems. One interesting example of such a macroscopic model is the interacting earthquake fault system model “Virtual California” (Rundle, 1988; Heien and Sachs, 2012), used in understanding the physics of interacting earthquake fault systems. We will briefly consider and review this type of tectonic/macroscopic model in a later section, inasmuch as it allows the construction of simulation testbeds to carry out experiments on the dynamical timescales and spatial scales of interest.

An interesting new development is associated with earthquakes in regions where oil and gas are being mined, termed induced seismicity. These earthquake events are the result of new fracking technology that has transformed previously relatively non-seismic regions—such as the US state of Oklahoma and the Groningen region of the Netherlands—into zones of frequent and damaging seismic activity (Luginbuhl et al., 2018c).

In association with this new induced seismicity, an important new model that can be considered is the invasion percolation (“IP”) model. IP was developed by Wilkinson and Willemsen (1983) and Wilkinson and Barsony (1984) at Schlumberger-Doll Research to describe the infiltration of a fluid-filled (“oil” or “gas”) porous rock by another invading fluid (“water”). The model has been studied by (Roux and Guyon, 1989; Knackstedt et al., 2000; Ebrahimi 2010; Norris et al., 2014; Rundle et al., 2019) primarily for applications of extracting oil and gas from reservoirs, and also in the context of the computation of scaling exponents. Laboratory examples of IP have also been observed (Roux and Wilkinson, 1988).

Until now, most of the research on this model has been concerned with understanding the scaling exponents and universality class of the clusters produced by the model (Roux and Guyon,

1989; Paczuski et al., 1996; Knackstedt et al., 2000). Direct application to flow in rocks has been discussed by Wettstein et al. (2012).

Yet the physics of the model can be applied to a number of other processes, for example the progressive development of slip on a shear fracture or fault. Notable among the physical processes of IP is the concept of *bursts*. These can be defined as rapid changes in the configuration of the percolation lattice and correspond physically to the formation of a sub-lattice having greater permeability or conductivity than the site(s) from which the sub-lattice originates.

The multiplicity of these spatial and temporal scales, together with the power-law scaling observed in the Gutenberg-Richter and Omori statistical laws, lend support to the basic view that earthquake fault systems are examples of general complex dynamical systems, in many ways similar to systems seen elsewhere. Examples of other types of physical systems that display similar behaviors include stream networks, vascular networks, spin systems approaching criticality, and optimization problems. Examples of systems from the social sciences displaying similar dynamics include queuing problems, and social science network problems in biology and economics (Ebrahimi, 2010).

5. Nucleation and Phase Transitions

5.1. Nucleation and Fracture. The idea that earthquakes are a shear fracture has allowed progress to be made using ideas from statistical physics. Fracture can be viewed as a catastrophic event that begins with nucleation, a first order phase transition. Griffith (1921) was the first to recognize that there is a similarity between nucleation of droplets and/or bubbles in liquid and gases as proposed by Gibbs (1878), and fracture. For example, we note that the the Griffith (1921) model of an fracture or crack is found by writing the free energy (Rundle and Klein, 1989):

$$F = -B\, l^2 + 2\gamma\, l \tag{13}$$

where B is a bulk free energy and 2γ is a surface free energy. Or in other words, B is the elastic strain energy lost when a crack of length l is introduced into the elastic material, and 2γ is the energy required to separate the crack surfaces. Instability occurs and the crack extends when the crack length l exceeds a critical value l_c determined by the extremum of F:

$$\frac{dF}{dl} = 0 \;\Rightarrow\; l_c = \frac{\gamma}{B} \tag{14}$$

In general, nucleation is usually modeled as a competition between a bulk free energy (B), and a surface free energy (2γ). The bulk free energy tends to lower the overall energy at the expense of the surface free energy. In the case of thermal and magnetic phase transitions, the surface free energy is also called a surface tension. Since the material damage that precedes fracture has a stochastic component, whether it is annealed or quenched, the relation between

damage and failure is statistical. This makes the methods of statistical mechanics relevant and the analysis of the relation between damage and catastrophic failure in simple models an important component for elucidating general principles. Several excellent articles and texts in physics, materials science and earth science communities document these ideas and serve as good references on progress in these fields (Alava et al., 2006; Kelton and Greer, 2010; Ben-Zion, 2008).

Earthquake seismicity has also been viewed as an example of accumulating material damage leading to failure on a major fault, and has been described by statistical field theories. For example, one can decompose a strain field $E_{ij}(\boldsymbol{x},t)$ into an elastic $\varepsilon_{ij}(\boldsymbol{x},t)$and damage $\alpha_{ij}(\boldsymbol{x},t)$component:

$$E_{ij}(\boldsymbol{x},t) \equiv \varepsilon_{ij}(\boldsymbol{x},t) + \alpha_{ij}(\boldsymbol{x},t) \tag{15}$$

One can then write a Ginzburg-Landau type field theory for the free energy for the energy in terms of the strain and damage fields, and then find the Euler-Lagrange equation by a functional derivative. The result are equations that modify the elastic moduli in the constitutive laws by factors such as $\mu \; -> \; \mu(1-\alpha^2)$, so that as damage accumulates (α increases), the rigidity of the material decreases, and large displacements and fractures become inevitable.

Earthquake nucleation has therefore been described as an example of nucleation near a classical spinodal, or limit of stability (Klein et al., 2000a, b, c). In this view, earthquake faults can enter a relatively long period of metastability, ending with an eventual slip event, an earthquake. Unlike classical nucleation, spinodal nucleation occurs when the range of interaction is long. In this physical picture, the slip on the fault, or alternately the deficit in slip relative to the far-field plate tectonic displacement, is the order parameter. Scaling of event sizes is observed in spinodal nucleation, but not in classical nucleation.

Other views of earthquakes have emphasized the similarity to second order phase transitions. Several authors view fracture and earthquakes as a second order critical point (Sornette and Virieux, 1992; Carlson and Langer, 1989), rather than as a nucleation event (Rundle and Klein, 1989; Klein et al., 1997). Recall that second order transitions, while they do show scaling of event sizes, are in fact equilibrium transitions, whereas nucleation is a non-equilibrium transition. The heat generated by frictional sliding of the fault surfaces is considered to be the analog of the latent heat in liquid-gas or magnetic first order phase transitions.

Shekhawat et al. (2013) used a two-dimensional model of a fuse network to study the effect of system size on the nature and probability of a fracture event. A fuse network is a model in which an electric current is passed through an array of connected electrical fuses, which can burn out or fail if the current is too large. The model was used as an analog for fracture of materials. They argued that there were different regimes of fracture and established a phase diagram in which the nature of the event crosses over from a fracture process that resembled a percolation transition (a second order transition) to one that resembled the nucleation of a classical crack, as

the system size increased. Experimental support for the idea that fracture is a phase transition can be seen in several investigations as described below.

Laboratory experiments can elucidate the relation between material damage and the onset of a catastrophic failure event that could lead to material degradation. The latter, for example, seems to be characteristic of the foreshocks that sometimes seem to precede major earthquakes. Although there have been significant advances in locating and characterizing this type of precursory damage in materials (Li et al., 2012; Hefferan et al., 2010; Guyer et al., 1995; Guyer and Johnson, 2009) there has been little progress in relating the type and distribution of damage to the onset of a major catastrophic failure such as is observed in major earthquakes.

Fracture under pressure in fiber boards has been studied by Garcimartin et al. (1997) who recorded the acoustic emissions generated by the formation of micro cracks that preceded the catastrophic fracture event. Noting the power law size distribution of these events the authors conclude that the fracture could be viewed as a critical phenomenon. Although there have been significant insights obtained from studies such as the ones cited above, a general framework that can unify these results is still lacking and many questions remain.

Damage can also initiate nucleation via the heterogeneous nucleation process. A great deal of work has been done to understand heterogeneous nucleation in special cases such as nucleation on surfaces (Klein et al., 1997; 2000a, b, c; 2007; 2009; Kelton and Greer, 2010; Muller et al., 2000; Koster et al., 1990) and aerosols (Flossman et al., 1985; Hamill et al., 1977; Hegg and Baker, 2009). As with fracture, an overall framework is lacking. The role of defects such as vacancies or dislocations in crystal-crystal transitions is not understood (Kelton and Greer, 2010), and neither is the effect of quench rates in multi-component systems (Gunton et al., 1983).

The fact that the state of the fields of nucleation and fracture are similar is not surprising. They are in many ways the same phenomenon. As noted, Griffith (1921) was the first to understand that the formation of a classical crack in a brittle material was a nucleation event. Rundle and Klein (1989) adapted a field theoretic approach used to study nucleation near the spinodal in long range interaction systems (Unger and Klein, 1984). Their model was applied to nucleation in metals. They obtained a theoretical description of the process zone associated with acoustic emissions produced by molecular bonds breaking ahead of the advancing crack opening (e.g., Broberg, 1999).

5.2 Nucleation and Failure Cascades. The idea of spinodal nucleation as a process leading to earthquakes is associated with the idea that earthquakes are part of a cascading process, where earthquakes that begin with small slipping areas progressively overcoming “pinned” sites to grow into large events. Pinned sites are called “asperities” in the literature (Scholz, 2019). Models for this type of process are often characterized by the question of “why do earthquakes stop?” A model for the cascade process was proposed by Rundle et al. (1998) based on the idea that slip events extend by means of a fractional Brownian walk through a random field via a series of burst events. More recent work has related this type of Brownian walk to bond percolation theory (Rundle et al., 2019).

With respect to earthquakes as a kind of generalized phase transition and a cascade to failure, Varotsos et al. (2001; 2002; 2011, 2013; 2014; 2020) and Sarlis et al. (2018) have proposed that earthquakes represent a dynamical phase transition associated with an order parameter k_1. That

parameter is defined as the variance of a time series of seismic electric signals. Furthermore, they define an entropy in natural time, and show that this quantity exhibits a series of critical fluctuations leading up to major earthquakes, both in simulation models, and in nature (Varotsos et al., 2011). These ideas depend on a definition of "natural time", that is discussed in more detail below (Varotsos et al., 2001).

Other work on similar ideas has been presented by Chen et al. (2008a, b). They proposed an alternative variant of the sandpile model with random internal connections to demonstrate the state of intermittent criticality or nucleation. The modified sandpile model (long-range connective sandpile model) has characteristics of power-law frequency-size distribution. The model shows reductions in the scaling exponents before large avalanches that mimics the possible reduction of Gutenberg-Richter b-values in real seismicity (Lee et al., 2008). Lee et al. (2012) also consider failure in a fiber-bundle model to address the problem of precursory damage. The study observes a clearly defined nucleation phase followed by a catastrophic rupture (Lee et al., 2012).

6. Earthquake Data

Earthquake data that are available for the study of dynamical processes fall into several categories. The first is seismicity data, that includes hypocentral data from catalogs, which list the location of initial slip, the magnitude of the eventual earthquake, and the origin time. Other data are measures of surface deformation, including Global Navigation Satellite System (GNSS) data, previously referred to as Global Positioning System (GPS) data. Another form of surface deformation data arises from radar satellites or aircraft that illuminate the earth and can be processed into interferometric synthetic aperture radar (InSAR) products. Stereo photogrammetric pairs can also be used to determine deformation from large events. These are the primary types of data that we discuss, although still other types of data include chemical, thermal, and electromagnetic (Donnellan, et al., 2019).

6.1 Earthquake Seismicity. Earthquake data are organized and available in online catalogs maintained by organizations such as the United States Geological Survey. Catalog data include the origin time, magnitude, latitude, longitude, depth and other descriptive information on the location where the earthquake rupture first begins (the hypocenter). Magnitudes can be of several types, but the most used is the moment magnitude scale, based on the seismic moment of the event (Hanks and Kanamori, 1979). The seismic moment is a measure of the mechanical and thermal energy release of the system as the earthquake occurs and is typically computed by fitting models of the source to waveforms observed on seismograms.

The data show that earthquakes of all magnitudes are known to cluster strongly in space and time (e.g., Scholz, 2019; Reasenberg, 1985). As noted, such burst phenomena are widely observed in many areas of science (Bahar et al., 2015; Mantegna and Stanley, 2004; Paczuski et al., 1996). One can introduce a definition of seismic bursts that encompasses both seismic swarms and aftershock sequences, with applications to other types of clustered events as we describe below. An example of aftershock sequences within 600 km of Los Angeles in association with several major earthquakes is shown in Figure 3. It can be seen that the activity following the main shock subsides to the background within several months. This is an example of Omori's law of aftershock occurrence (e.g., Scholz 2019).

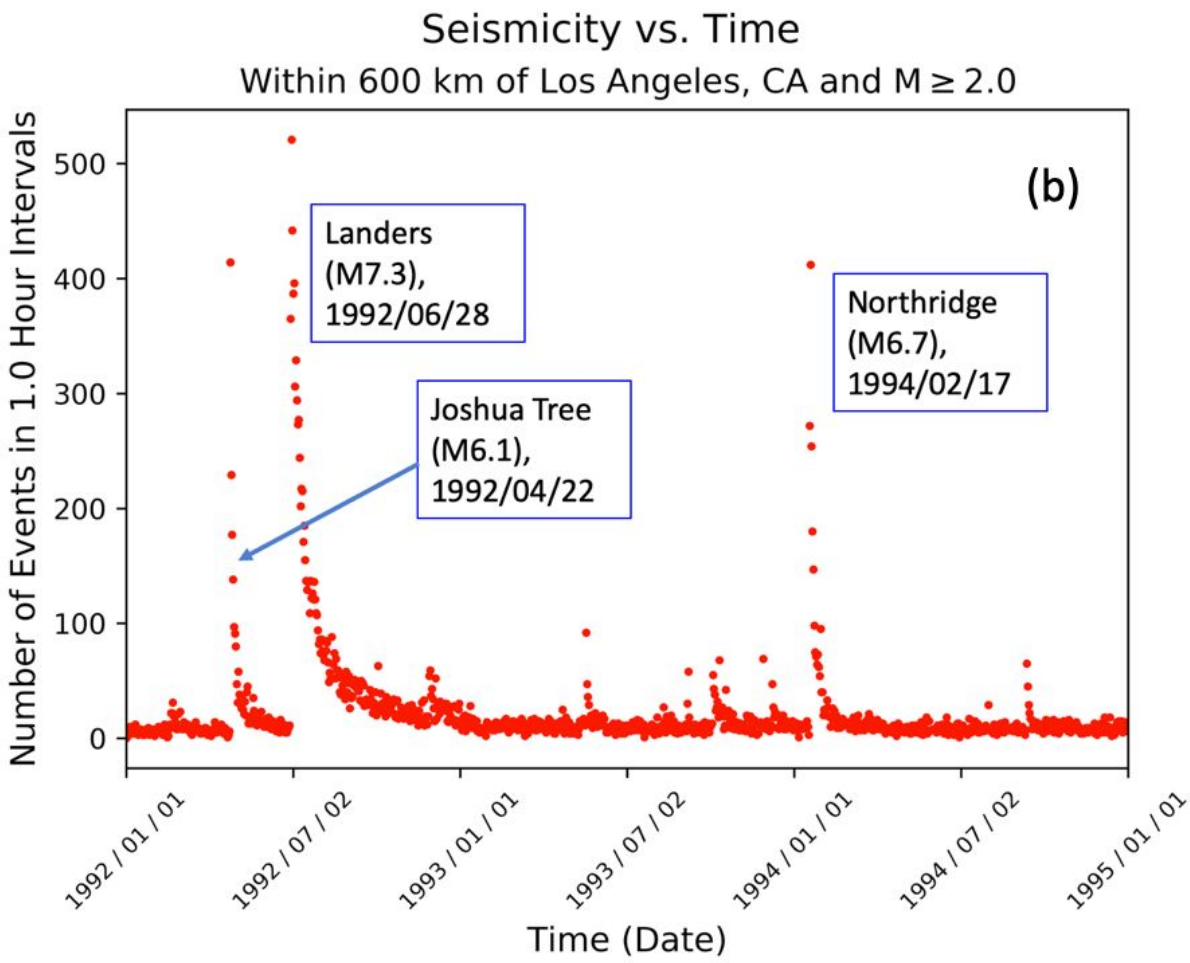

Figure 3. Seismicity having magnitudes M≥2 within 600 km of Los Angeles from 1992-1995.

6.2 Global Navigation Satellite Systems (GNSS). GNSS data, of which Global Positioning System (GPS) is one of the earliest and most familiar examples, is another type of data being analyzed for use in earthquake forecasting and nowcasting. Significant work has been done in the development of cost effective and efficient GNSS-based data systems to quickly and efficiently estimate a number of vital earthquake-related parameters (e.g. Hudnut et al., 1994; Tsuji et al, 1995).

GNSS is also useful for tracking crustal deformation associated with the earthquake cycle (e.g. Hager et al, 1991; Sagiya et al, 2000). GNSS data an also be used to illuminate many of the processes present in postseismic deformation, and thereby to contribute understanding to earthquake physics (Melbourne et al., 2002). GNSS can even be used to track tsunami waves that arise as a result of great submarine earthquakes for communities nearest the earthquake's epicentre and as they propagate to distant coastlines around the world through the effects of ionospheric gravity waves (LaBrecque et al., 2019). In short, GNSS measurement of crustal deformation can be used to measure tectonic deformation prior to earthquakes, coseismic offsets

from earthquakes with decreasing latency, and postseismic motions, all of which inform models of how the Earth's crust accumulates strain, then fractures, and finally recovers.

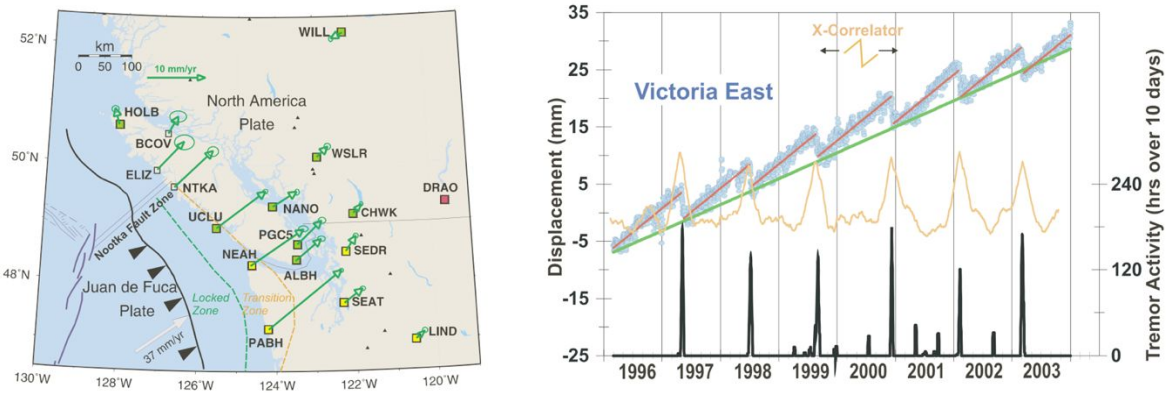

Figure 4. Role of GNSS observations in the analysis of Episodic Tremor and Slip. At left is the region of the Cascadia subduction zone off the Pacific Northwest, showing a map of stations at which GNSS observations are routinely monitored. Of note are stations ALBH at the southern end of Vancouver island, and station DRAO in the stable continental interior. At right above is shown the displacement of ALBH with respect to DRAO over the years 1995 - 2004. Displacement is shown as the blue circles and red lines that represent best fits to the data. The green line shows the steady aseismic trend. Bottom right is a record of the regional small earthquake bursts that accompany the slip events. The gold line in the middle represents the correlation of the slip data with a detrended sawtooth curve, illustrating the repetitive nature of the events. Figures and data are from Dragert et al. (2004).

Another important application of GNSS is the observation and analysis of Episodic Tremor and Slip, a phenomenon that was discovered by Dragert et al. (2001, 2004; Brudzinski and Allen, 2007) along the Cascadia subduction zone along the Pacific Northwest coast of California, Oregon, Washington and British Columbia (Figure 4). These events occur at relatively shallow depths of 30 km or so on the plate interface and are associated with brief episodes of slip and bursts of small earthquakes. ETS has been observed elsewhere in the world as well, including such locations as Costa Rica (Walter et al., 2007) and central Japan (Obara and Sekine, 2009)

Measurement of crustal deformation to inform earthquake fault behavior dates back to the early 1980s (e.g., Davis et al., 1989). By the early 1990s GNSS had been used to identify additional contributions to the Pacific - North American plate motion from faults beyond the San Andreas fault proper (Feigl et al., 1993). On a more local to regional scale, GNSS crustal deformation measurements combined with modeling identified the geometry of faults near the Ventura basin and were used to estimate the earthquake potential of the faults as capable of producing a M6.4 earthquake (Donnellan et al, 1993a, b). In early 1994 the M = 6.7 Northridge earthquake occurred (Jones et al., 1994), demonstrating the value of applying measurement of crustal deformation to earthquake hazard assessment.

The success of GNSS for understanding earthquakes led to the deployment of continuous GNSS networks in California (Blewitt et al., 1993; Bock et al., 1993; Hudnut et al., 2001), the western US (Herring et al, 2016), Japan (Tsukahara, 1997), and globally (Larson et al, 1997). By the early 2010s GPS networks were relied on for understanding crustal deformation and fault activity. Surface deformation was incorporated into the most recent Uniform California Earthquake Rupture Forecast version 3 (UCERF-3) led by the United States Geological Survey (USGS) with input from the California Geological Survey (CGS) and research community (Field et al, 2013; Field et al, 2014). These long-term forecasts provide an assessment of earthquake hazard and intermediate fault behavior.

Retrospective real-time analysis of GNSS networks have shown that the moments and slip displacement patterns of large magnitude earthquakes can be calculated within 3-5 minutes (Ruhl et al, 2017; Melgar et al, 2020). Furthermore, algorithms now exist to use these earthquake source models to assess the likelihood of tsunamis and to predict the extent, inundation and runup of tsunami waves. Recently for example, a joint NOAA/NASA effort has further demonstrated the consistent estimates of tsunami energy using GNSS for improved early warning (Titov et al., 2016).

An important application of real-time GNSS data is for tsunami early warning, as a result of great submarine earthquakes. The 2004 M = 9.2 Sumatra-Andaman event (Ammon et al., 2005; Ishii et al., 2005; Lay et al., 2005; Stein & Okal, 2005, Subarya et al., 2006) resulted in over 250,000 casualties, the majority of them on the nearby Sumatra mainland, with inundation heights of up to 30 m (Paris et al. 2007). Improvements in earthquake forecasting can be expected to yield significant benefits in tsunami warning as well. As another example, the M = 8.8 2010 Maule earthquake in Chile (Lay et al., 2010; Delouis et al. 2010) resulted in 124 tsunami related fatalities and wave heights up to 15-30 m in the near-source coast (Fritz et al., 2011).

Still another example is the 2011 M = 9.0 Tohoku-oki earthquake in Japan (Simons et al., 2011; Lay & Kanamori, 2011), which generated a tsunami with inundation amplitudes as high as 40 m. That event resulted in over 15,000 casualties (Mori et al., 2012) and was the first case of a large tsunami impinging upon a heavily developed and industrialized coastline in modern times. In addition to the tragic loss of life, the economic collapse of the near-source coastline, which spans nearly 400 km, was almost complete (Hayashi, 2012).

Retrospective analysis in simulated real-time mode of high-rate (1 Hz) GNSS (primarily GPS) data was collected during the 2011 Tohoku-oki event on the Japanese mainland from a network of more than 1000 stations. Those data convincingly demonstrated that tsunami warnings in coastal regions immediately adjacent to large events could be effectively issued without regard for magnitude or faulting type (Melgar et al., 2013; Song et al., 2012; Xu and Song, 2013).

By 2020, there will be over 160 GNSS satellites including those of GPS, European Galileo, Russian GLONASS, Chinese BeiDou, Japanese QZSS, Indian IRNSS and other satellite constellation broadcasting over 400 signals across the L-band, nearly double the number today at any location. The expanded GNSS constellation will improve the accuracy of the system and will likely provide future advancements in early warning capabilities.

In summary, the augmentation of existing monitoring networks with real-time GNSS would enable more accurate and timely determination of the magnitude for large earthquakes (> ~M =

8), identification of the location, geometry, and extent of fault rupture and the orientation of ground displacement as input to earthquake forecasts and improved tsunami forecasting and real-time prediction models. Increased access and use of real-time GNSS data from existing and modernized networks would avoid or minimize underestimating the likelihood of devastating earthquakes and tsunamis (Goldberg et al. 2018, 2019; Donnellan et al., 2019; LaBrecque et al., 2019).

6.3 Interferometric Synthetic Aperture Radar (InSAR). InSAR is a satellite-based radar technology that produces images of deformation of the earth's surface following dynamic events. Since the 1992 M = 7.3 Landers earthquake (Massonnet et al., 1993), InSAR measurements and data scenes have progressed from innovative promise to a relatively routine capability (e.g. Glowacka et al, 2010; Brooks et al, 2007; Ryder and Burgmann, 2008; Johanson and Burgmann, 2010; Tong et al, 2010; Wisely and Schmidt, 2010; Wei et al, 2015; Xue et al, 2015; Dreger et al, 2015), although the number of interferograms available is still limited by the lack of satellites or airborne vehicle platforms.

Examples of both satellite-based imagers (Figure 5, left, [1]) from data obtained by the European Space Agency (ESA) Sentinel-1 radar satellite (e.g., Fielding et al., 2014), and airborne imagers for the magnitude Mw 6.0 West Napa earthquake show a consistent pattern. Together with data from the NASA UAVSAR instrument, these data show that the earthquake occurred on the near vertical, NNW striking West Napa right-lateral strike slip fault with average slip of approximately ~0.5 m. UAVSAR is the Uninhabited Aerial Vehicle Synthetic Aperture Radar aircraft.

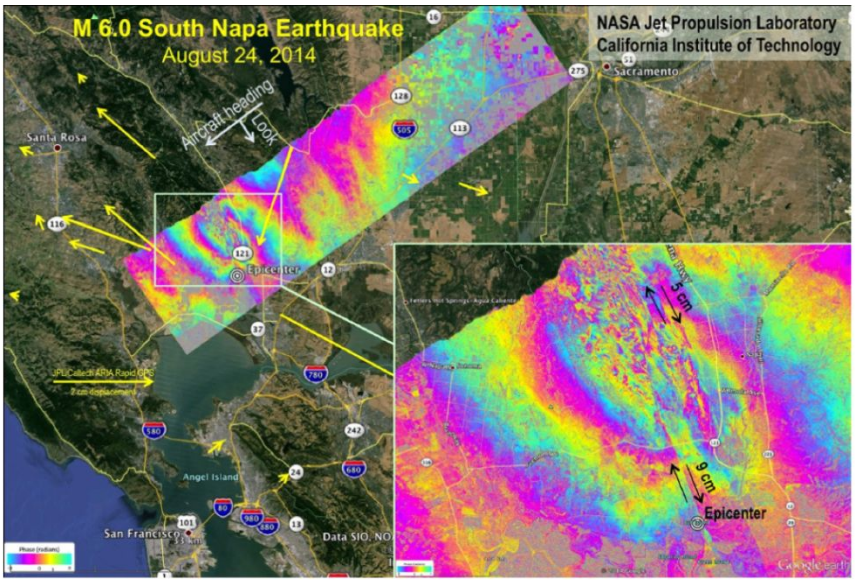

Figure 5. Interferogram of the South Napa earthquake of August 24, 2014 captured by the NASA Uninhabited Aerial Vehicle Synthetic Aperture Radar instrument. L-band fringes represent displacements of 24 cm along the line of sight to the instrument. Credit: NASA/JPL

Ref: https://nisar.jpl.nasa.gov/mission/get-to-know-sar/interferometry

Earthquakes in the Los Angeles basin can produce major damage and loss of life, examples include the 1933 M = 6.4 Long Beach earthquake and the 1993 M = 6.7 Northridge earthquake. The seismic moment of aftershocks from the earthquake represented only about 24% of the total deformation, as revealed by the geodetic measurements (Donnellan et al, 1998).

More interesting was the M = 5.1 La Habra earthquake of 28 March 2014 that occurred between the Puente Hills thrust fault and the Whittier Narrows faults (Donnellan et al., 2015). Deformation associated with the La Habra earthquake was captured by the UAVSAR vehicle (Figure 2 in Donnellan et al., 2015). The UAVSAR data were collected in the Los Angeles basin since 2009 as a part of an experiment to 1) forecast earthquakes in California (Rundle et al., 2002, 2003; Holliday et al., 2007) and 2) validate the forecasts with a systematic program of observation via alternate data acquisition methods. These forecasts indicated a high probability of an earthquake at the southern boundary of the transverse ranges in the area where the 2008 Chino Hills M = 5.5 earthquake occurred.

The surface deformation signature of the La Habra earthquake revealed by the UAVSAR instrument was subtle but illustrative of the power of the data. Features including cracked pavement, broken curbs, and damaged structural foundations were found to be associated with the UAVSAR images and were found via field investigations that might otherwise not have associated the damage with the fault (Donnellan et al., 2015). Simple analyses of the historical seismicity near Los Angeles indicate that the probability of a major earthquake of magnitude M = 6.1 to M = 6.7 is high at the present time within a circle of radius 100 km of Los Angeles (Donnellan et al., 2015).

As a final example, UAVSAR captured slip from the 2020 M = 7.2 El Mayor-Cucapah earthquake that ruptured in Baja Mexico to just north of the US-Mexico border, but triggered slip on an extensive network of faults even farther to the north (Rymer et al, 2010; Wei et al, 2011; Donnellan et al, 2014). The M = 5.7 Ocotillo aftershock occurred about 2.5 months after the mainshock, northwest of the rupture and just south of the Elsinore fault, suggesting that the two faults systems could be connected and possibly loading the Elsinore fault from the 2010 event.

In addition to these data, NISAR (the planned NASA-ISRO SAR mission), is another space-based radar instrument that is planned to launch in the 2022 timeframe. NISAR promises to generate vast quantities of new imaging data for crustal deformation research. This mission, which is a collaborative effort between NASA and the Indian Space Research Organization ISRO, is designed to operate both an L-band and an S-band radar in order to obtain data at two wavelengths for a minimum of 3 years. NISAR is planned to measure at least two components of the point-to-point vector displacements with a sampling interval of 12 days or shorter over at least 80% of 12-day or shorter intervals. The maximum gap in temporal sampling is expected to be 60 days over pre-specified regions of Earth's land surface. Accuracy is expected to be 3.5 (1+L1/2) mm or better, over length scales 0.1 km < L < 50 km, with resolution of 100 meters, over at least 70% of the specified regions.

7. Models for Earthquake Failure

Models for earthquake failure have been proposed as two basic types, statistical models and physical models. The statistical models assume some form of a probability distribution, and then

attempt to define the parameters in terms of observables. The idea here is to determine expressions for the probability of earthquake occurrence based on the assumed statistical distribution. The physical models begin with a description based on stress and strain, and use some form of dynamics to produce catalogs of computational simulation data that can then be analyzed statistically for failure probabilities. Dynamics are important as they relate to the underlying tectonic forces and stresses to the observed displacements at the earth's surface, and to the patterns of earthquake events that occur in space and time. Here we briefly summarize several of these models and approaches.

7.1 Statistical Distributions for Models. The most widely used statistical distribution has been a Poisson model for earthquake occurrence, since large earthquakes are known to recur on major faults. Thus in this view, the physics is considered to be similar to nuclear decay processes or cars arriving at a store. The Poisson model for an earthquake to occur within a future random time *t* is:

$$P_m(\mathrm{T} \leq t) = 1 - e^{-\upsilon_m t} \tag{16}$$

Here υ_m is the rate of occurrence of an earthquake of a given magnitude *m* or larger. An interesting property of the Poisson distribution is that it has no memory of past events. This is easily shown by computing the conditional probability that at earthquake occurs within a time Δt after *t*, given that it has not occurred before *t*:

$$P_m(\mathrm{T}|\mathrm{T} \geq t) = \frac{P_m(t+\Delta t) - P_m(t)}{1 - P_m(t)}$$

$$= 1 - e^{-\upsilon_m \Delta t} \tag{17}$$

As can be seen, the final expression does not depend on the time *t*, only on the future time interval Δt.

A generalization of the Poisson model is the Weibull (1952) model, which is often used in failure analysis for engineering materials:

$$P_m(\mathrm{T} \leq t) = 1 - e^{-\left(\frac{t}{\tau}\right)^{\beta}} \tag{18}$$

Parameters include the nominal failure time τ and the exponent β. As we show below, the Weibull model can be used to develop forecast models that can be tested with statistical test protocols.

Other commonly used models included the Brownian Passage time model, more commonly called the inverse Gaussian model. Here the expression for the probability is more complex, in

that it describes the time a Brownian process takes to reach a fixed time *t*. Another probability model often used is the log-normal distribution, which describes multiplicative random processes.

7.2 Simple Physical Models. Simple models for the earthquake sliding process have been developed and compared to experiments and observations where applicable. The first model was introduced by Burridge and Knopoff (1967 "BK"), a dynamic model using sliding frictional blocks and massive blocks connected by springs. The first cellular automaton slider block model was introduced by Rundle and Jackson (1977; “RJ”) using massless blocks.

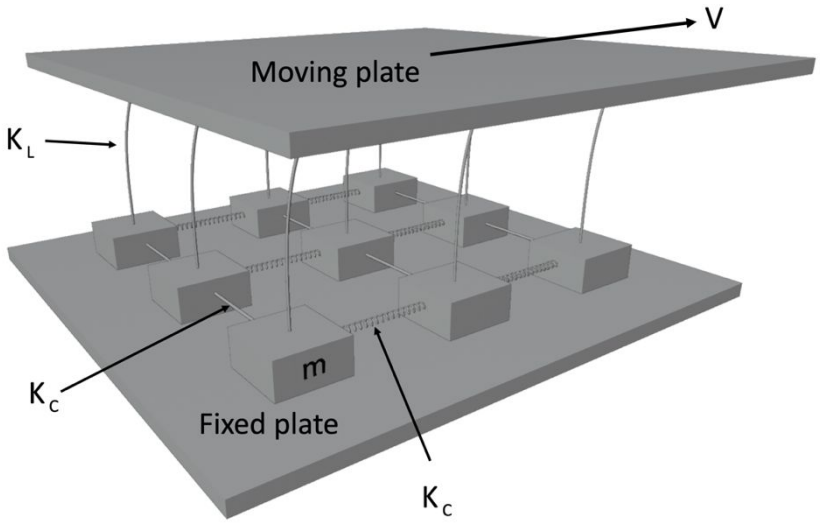

Figure 6. Schematic image of a slider block model. The small blocks slide with friction on the fixed plate and are loaded by the moving plate. Small blocks interact by means of the coupling springs.

In both models, each block is connected to neighbors by coupling springs of strength K_C, thereby allowing the blocks to interact. In addition, each block is connected to a slowly moving and persistently advancing loader plate by a spring of constant K_L that serves to increase the stress on all the blocks. These models were introduced to the physics community by Carlson and Langer (1989; "CL") and by Olami, Feder and Christensen (1992; “OFC”).

An example of this type of model in *d* = 2 is shown in Figure 6**.** In these models, each block or site is assigned a failure threshold and a residual stress. The system is initialized by assigning a stress at random to each site. Stress on each site increases linearly with time between slip events due to the action of a “loader plate”, representing the increase of tectonic stress on the fault. Each site is visited and if the stress is larger than the threshold, the site fails. Once a site fails, it can trigger sliding at other sites to which it is connected by coupling springs (the interactions). As a result, some of the stress is lost, while the remainder is transferred to neighbors. Once the cascade of failing blocks ends, the cycle begins again.

More specifically, the force or stress on each slider block can be represented by a generalization of equation (11):

$$\sigma_i = \; - \; \left\{ K_C \sum_{\substack{i \neq j \\ j < R}} s_j(t) + \; K_L \left[s_i(t) - Vt \right] \right\}$$

(19)

where $s_i(t)$ is the slip at time *t* on block i, V is the velocity of the loader plate. In the original dynamic model of Burridge and Knopoff (1967), the time-dependent positions of the blocks were found by solving coupled equations for blocks of mass m:

$$m \frac{d^2 s_i}{dt^2} = \sigma_i(t) - \sigma_i^F$$

(20)

Here σ_i^F is the frictional resistance to sliding on block i. In the original BK models, the frictional force had the form:

$$\sigma_i^F = \sigma_o^F - \; \theta \frac{vi(t)}{vi(t) + \; v_C}$$

(21)

where σ_o^F , θ, and v_C are constants, and

$$v_i(t) = \frac{ds_i}{dt}$$

(22)

On the other hand, with the massless slider block RJ models, the elastic stress was computed using expression (19), and a slip value was computed by an update rule such as:

$$s_i(t+1) = s_i(t) + \Delta s_i \; \Theta\left[\sigma_i(t) - \; \sigma_i^F \right]$$

(23)

Here $\Theta[\; * \;]$ is the Heaviside step function, and Δs and σ_i^F are constants. For the special case of *d* = 2, and each block is connected to 4 neighbor blocks by coupling springs, the jump in slip Δs is often expressed in terms of a "stress drop" term, $\Delta\sigma = \sigma_i^F$:

$$\Delta s_i \; = \frac{\Delta\sigma}{K_L + 4K_C}$$

(24)

so that slip of the block reduces the stress σ_i to zero.

In this model, which is typically initiated with random initial conditions, a block fails when the persistently advancing plate loads enough stress onto a block so that the failure threshold is reached. At that point, the block fails. Slip of a single block can induce a cascade or avalanche of failing blocks by virtue of the coupling springs, with each slipping block continuing the cascade as it transfers stress to its neighbors. It is worth noting that the OFC model is the same model but formulated in terms of stress variables rather than displacement or slip variables.

In a variety of papers, (Rundle and Klein, 1992; Rundle et al., 1995; Klein et al., 2000a, b, c; Serino et al., 2011) have adapted this model in several ways. First, the stress transfer range *R* is assumed to be large, R→∞, to model elastic forces, which can be shown to have an infinite range of interaction.

In the large *R* limit the system demonstrates Gutenburg-Richter (GR) scaling of small to medium size earthquake events (Klein et al., 2000a ,b, c). The dynamics also show rare large events that do not scale, and have the properties of nucleation events (Klein et al., 2000a, b, c; Klein et al., 2009). The scaling of the events is generally the same as the scaling of precursors to the main fracture event in the chipboard fracture experiment (Garcimartin et al., 1997).

Another variation on this model is the Traveling Density Wave model (Rundle et al., 1996) in which the frictional or pinning force is derived from a potential in the form of a traveling harmonic wave consisting of sines and cosines. The idea is that the population of pinning points can be decomposed into a Fourier series. Numerical simulations show that populations with identical phases produce one repetitive, large earthquake, often called a "characteristic earthquake" in the literature (Scholz, 2019). Populations of pinning points having random phases, by contrast, produce a scaling distribution of earthquakes. Thus the randomness of the phases is a kind of control that can be used to obtain a diversity of dynamics.

Another adaptation was to introduce damage as described above (Rundle and Klein, 1992; Rundle et al., 1995; Klein et al., 2007). Damage was introduced by modeling micro cracks as sites at random that dissipate any stress that was transferred to it under shear. Serino et al. (2011) were able to show that the addition of this form of damage introduced an exponential cutoff to the GR scaling in the model.

Applying this model to single faults in the Southern California fault system, Serino et al. (2011) showed that the fault scaling could be fit by an exponentially damped power law. Since faults in a fault system could have different amounts of damage, to obtain the GR scaling for a fault system Serino et al (2011) superimposed the ensemble of single fault scaling laws to obtain a different exponent that had a larger range of power law fit and whose value depended on the number of faults with a given level of damage, thus explaining how different fault systems can have slightly different GR scaling exponents.

This result is similar to what is expected for systems with damage. Since damage will not be uniformly distributed in a sample, scaling laws (for e.g. fracture bursts) must be a superposition of scaling from different regions of the material. This paradigm can explain the observed differences in the scaling laws for different materials seen in Garcimartin et al. (1997) and demonstrates how simple models can lead to important insights into the behavior of real materials. Under tensile stress, damage can result in regions that can no longer support loads leading to stress concentration (Shekhawat et al., 2013; Kanninen and Popelar, 1985).

Another type of model was recently proposed by Rundle et al. (2019) to model fault rupture in the presence of pore fluids. This model is a type of invasion percolation model, first proposed by Wilkinson and Willemsen (1983). The model is intended to represent the occurrence of burst-like dynamics such as are seen in earthquake swarms and aftershocks, based on a type of constrained Leath invasion percolation (CLIP) model.

Interpreting the percolation sites as units of energy release, Rundle et al. (2019) showed that the model reproduces the observed natural scaling of earthquakes with the correct scaling exponent in the limit that the occupation probability equals the critical bond percolation probability for the onset of connection across the grid. Comparing these results to observed scaling of earthquakes in several geological regimes, they find good quantitative agreement, in which the Gutenberg-Richter *b*-value (scaling exponent) is b=1 at $p = p_{occ}$, and b >1 at $p< p_{occ}$.

7.3 Topologically Realistic Earthquake Simulators. Earthquake simulators are a type of model in which earthquake faults are represented as topologically realistic dislocation surfaces subject to slow long-term loading at tectonic rates, and upon which frictional models are used to prescribe the physics of stick-slip motion. Virtual Quake (formerly Virtual California, Rundle, 1988) is such a model, and its history and use is described in Sachs et al. (2012). Virtual Quake has been used to compute earthquake probabilities by simulating a long history of synthetic earthquakes, then using the statistics in the simulated catalogs to compute probabilities of future events (Rundle, 1988; Rundle et al., 2006; Van Aalsburg et al., 2007; Yikilmaz et al., 2010). Other simulators based on similar principles now exist as well, including RSQSim, ViscoSim, and AllCal (Tullis et al., 2012).

The Virtual Quake model includes stress accumulation and release as well as stress interactions between faults in the model, including the San Andreas fault and other adjacent faults. The model is based on a set of mapped faults with estimated slip rates, a prescribed plate tectonic motion, earthquakes on all faults, and purely elastic interactions (Rundle, 1988; Rundle et al., 2001, 2002, 2004). Earthquake activity data and slip rates on these model faults are obtained from geologic databases.

Too implement the Virtual Quake model, one first defines a set of fault surfaces, and applies a coarse graining algorithm to partition the faults into smaller areas. Once these partitions are defined, the resulting model can be treated essentially as a slider block model. One then assigns properties including coefficients of friction and long term slip velocities. Since the faults are embedded in a d = 3 medium, the failure threshold is then defined by:

$$\sigma_i^F = \mu_{S,i}\,\sigma_{N,i} \tag{25}$$

where $\mu_{S,i}$ is a coefficient of static friction on the i^{th} coarse-grained fault partition, and $\sigma_{N,i}$ is the local normal stress, which is composed both of a time-varying dynamical element, and the gravitational overburden.

The time-varying stress fields are given by equation (11), with the addition of a loading term:

$$\sigma(\boldsymbol{x}_i,t) = \sigma_i(t) = -\left\{\sum_{\substack{j \neq i \\ j < R}} T_{ij}\, s_j(t) + T_{ii}\left[s_i(t) - V_i t\right]\right\} \tag{26}$$

where the T_{ij} are the spring constants K_C and K_L as in (19), and are found by integrating the stress Green's functions in (11) over the elementary coarse grained fault elements. The quantity V_i is the long-term rate of offset across the fault element located at $\boldsymbol{x}_i$.

Other simulators use a similar approach as far as the fault interactions are concerned, but typically differ in their use of friction laws. The RSQSim simulator of Richards-Dinger et al. (2012) uses a significantly modified version of rate-and-state friction, which is a model in which the coefficient of friction $\mu(\theta(t), V(t))$ depends linearly on a state variable $\theta(t)$ together with a logaritmic dependence on the slipping velocity of the fault interface $V(t)$ (Dieterich, 1979, 1992, 1994). A criticism of this model is that if the fault is locked so that $V = 0$, a logarithmic singularity appears. Later modifications of the model addressed ways to remove this problem. The state variable is taken to represent the contact time of pinning points on the fault, and evolves in response to the driving velocity $V(t)$. The AllCal simulator of Ward (2012) uses a set of scales to prescribe rupture and healing properties. The ViscoSim simulator of Pollitz (2012) uses a more conventional Coulomb type friction law with a dynamic overshoot parameter.

At the present time, Virtual California is also the only code that has developed a useful crustal deformation component to accompany the slip history simulator. Both surface deformation and interferograms are routinely computed. A typical example of interferograms from simulated earthquakes are shown in (Figure 7). The crustal deformation component of the model is useful when comparing model displacements to observed displacements from GNSS or InSAR measurements.

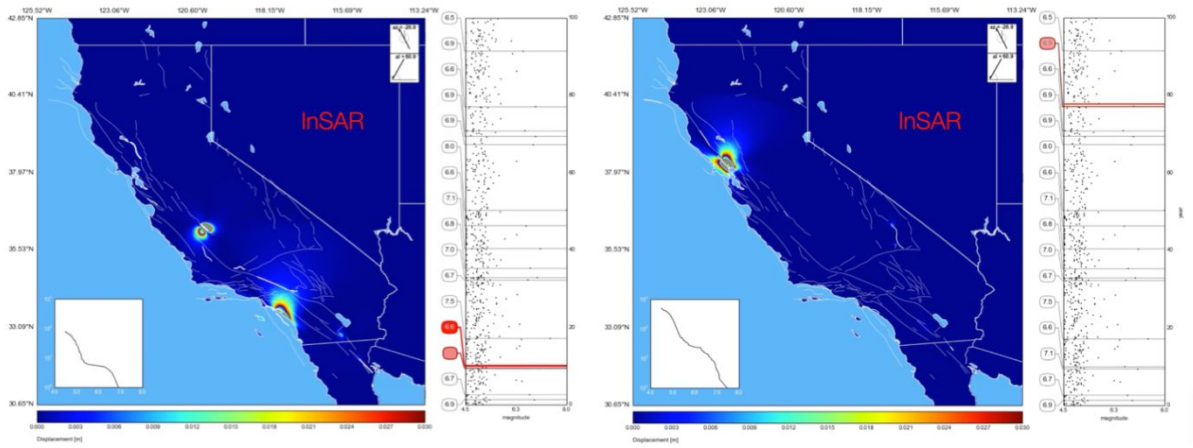

Figure 7. Two frames from a Virtual Quake simulation of 200 years of earthquakes in California, with events represented as InSAR L-band interferograms. The assumed look angle from the satellite to the ground is shown at upper right. At lower left is the Gutenberg-Richter magnitude-frequency diagram that is built over the course of the simulation. In the right panel of each figure, events larger than M6.5 are indicated by bubbles with horizontal line indicators, all other events as points. Simulation courtesy of M. Sachs (Rundle Group). Statistics of events from simulations can be used to identify potential earthquake patterns in space and time, and to forecast future events after comparing the patterns with the known history of large earthquakes in the region.

It should also be noted that the earthquake simulators described above all use some form of Boundary Element (BE) method to compute the stress transfer and kinematic Green's functions. With boundary elements, one defines a series of rectangular fault elements and uses the tabulations in Okada (1992) to compute the stress transfer coefficients and the surface displacement Green's functions. The same method can be used to compute the gravity change Green's functions using the Okubo (1992) tabulations. The advantage of the BE approach is that it is an order N method, where N is the number of boundary elements, so that computation time grows only with the number of fault elements. However, the disadvantage is that stresses are only computed on the boundary elements themselves, so stresses at arbitrary locations within the earth are not computed. Also, displacements are generally only computed at the free (earth's) surface.

Virtual Quake is the only earthquake simulator code whose source code, documentation and user manual are freely available either through the NSF-sponsored site Computational Infrastructure for Geodynamics (CIG: https://geodynamics.org/cig/software/vq/), or via the community site GitHub (https://github.com/geodynamics/vq/issues).

7.4 Statistical Forecast Models. A current model that is frequently used for the forecast of future earthquake aftershocks is the Epidemic Type Aftershock Sequence (ETAS) model (Ogata, 1988; 1998; 2004; 2011; Helmstetter and Sornette, 2003). The original earthquake is the parent

and the parent produces offspring which are first order aftershocks. The offspring can then become parents generating offspring that are second order aftershocks, and so forth. The frequency-magnitude statistics of the offspring are given by the Gutenberg-Richter law and the time dependence by Omori's law.

The key parameter in the ETAS model is the branching ratio n which is the number of offspring earthquakes generated on average by a parent earthquake. If n is greater than one the number of earthquakes grows without bound and is thus unrealistic. If n is near one a large fraction of earthquakes can be aftershocks. Smaller values of n require more random background earthquakes. A typical value is n = 0.8 but it is very difficult to separate background random earthquakes from ETAS aftershocks in terms of observations.

The key equations in the ETAS are then the Gutenberg-Richter magnitude-frequency law (1), the Omori law of aftershock decay (5), and an earthquake productivity relation that specifies the probability that an earthquake will give rise to daughter earthquakes:

$$K(m) = A\ 10^{\alpha(m - m_c)} \tag{27}$$

Here $K(m)$ is the rate of production of magnitude m earthquakes above the completeness threshold m_c of the catalog. Combining the rate (27) with the GR (1) and Omori (5) laws, one obtains a time-dependent rate of earthquake occurrence that can then be inserted into the Poisson probability law, thus yielding a probability for an earthquake of magnitude m to occur.

An alternative model is the Branching Aftershock Seismicity Sequence (BASS) model (Holliday et al., 2007). Both ETAS and BASS models consider multiple generations of aftershocks. The ETAS model uses the productivity relation (27) based on an average parent-offspring ratio. The ETAS model requires two parameters, *A* and α, for this relation without physical justification.

On the other hand, the BASS model utilizes the modified form of Bath's law specified by observations, in which the average difference in magnitude between a mainshock and its largest aftershock obtained from the aftershock GR distribution, Δ_m is used. This term is introduced into the Gutenberg-Richter law (1), and replaces the productivity relation (27). The BASS model is the self-similar limit of ETAS. The arbitrary productivity relation in ETAS is replaced by Bath's law that on average the largest aftershock is a fixed magnitude difference Δ_m (~ 1.2) less than the main shock. A major advantage of the BASS modal is that the two unconstrained parameters in the ETAS productivity relation are replaced by Δ_m and the *b*-value scaling exponent in Gutenberg-Richter scaling, both directly constrained by observations.

In the BASS model, the fraction of main shocks that have foreshocks is independent of mainshock magnitude, while in ETAS the dependence is assumed to be exponential, an assumption not confirmed by observations. However, in a number of studies on the statistical variability of Δ_m (Vere-Jones, 1969; Console et al., 2003; Helmstetter and Sornette, 2003) it has been argued that this law might be an artifact caused by the different criteria that seismologists apply to define mainshocks and aftershocks. Another criticism of the BASS model is that it can produce infinite numbers of aftershocks. However, this problem is easily removed by a physically

acceptable limit on the upper magnitudes of aftershocks. This is equivalent to an inverse Bath's law that an aftershock cannot be $\Delta_m \sim 3$ bigger than a mainshock.

7.5 Invasion Percolation. The process of energy extraction by fracking has led to many analyses of induced seismicity, which is observed to proceed by sudden clusters or bursts of activity. Here we consider new *invasion percolation* ("IP") models. IP was a model developed by Wilkinson and Willemsen (1983) and Wilkinson and Barsony (1984) at Schlumberger-Doll Research to describe the infiltration of a fluid-filled ("oil" or "gas") porous rock by another invading fluid ("water"). An interesting recent review of the literature in this area has been given by Ebrahimi (2010). In fact, remarkably little research has been carried out on this model despite the broad applicability of the physical processes. Understanding the processes of cluster formation in models like IP leads to new ideas, discussed below, for analysis of earthquake clusters or bursts in nowcasting models.

Until now, most of the research on this model has been concerned with analyzing the scaling exponents and universality class properties of the clusters produced by the model (Roux and Guyon, 1989; Paczuski et al., 1996; Knackstedt et al., 2000). Discussion of direct application to flow in rocks has been detailed in Wettstein et al. (2012). Laboratory examples of IP have also been observed (Roux and Wilkinson, 1988).

Invasion percolation is a very simple model in which a lattice of sites is specified with bonds between them. The sites are taken to represent larger "pores" in the rock, with the bonds representing small capillaries (flow paths) between the pores. To further specify the model, a series of uniformly distributed random numbers are generated and assigned to each of the bonds in the lattice. These random numbers are taken to represent the tendency for the invading or wetting fluid to pass through the capillaries. Resistance to the invading fluid is via capillary forces rather than viscous fluid forces.

In the "standard" model for invasion percolation, the invading fluid is introduced along one side (say, the left-hand side) of a square lattice. The dynamics proceeds by locating the bond with the lowest value of probability p, and then marking that bond and the connecting site as having been "invaded". In the next step, the bonds leading out from all invaded sites are then examined, the lowest value of probability p is found, and the bond and its connected site are marked as "invaded".

This process is repeated until the invaded fluid encounters the opposite boundary (say, the right-hand side of the lattice). By definition, there is one connected cluster of sites extending from the left side to the right side. This cluster is essentially the same as the "infinite" or "spanning" cluster of random site percolation (Stauffer and Aharony, 1994).

8. Prediction, Forecasting, Nowcasting, Testing

As described above, prediction in the context of earthquakes can be regarded as the precise specification of time, location, and magnitude of an impending earthquake. The process of producing a prediction generally involves a search for hypothesized precursory phenomena. As discussed earlier, many studies have found that reliable earthquake prediction, with associated estimates of successful predictions, false alarms, and failures to predict, is extremely difficult, if not impossible.

On the other hand, forecasting is the specification of the probability of a future earthquake, usually within some level of confidence. To make a forecast, one must assume a probability law governing the system, an assumption that is subject to debate.

Finally, a nowcast is the computation of the current state of risk of the system, often by proxy data (Rundle et al., 2016b; 2018; 2019). Although a nowcast implies a level of near-future hazard, it is not explicitly stated. In this review, we will not further consider the earthquake prediction problem, rather we focus on forecasting and nowcasting.

8.1 Long-Term Forecasting. Long-term earthquake forecasts have often been proposed for future time periods of decades. Most long-term earthquake forecasts are posed as the probability that an earthquake will occur in a given future time window. These estimates involve two primary choices, the data that are used to describe when and where previous earthquakes occurred, and the models that are used to forecast when future earthquakes will happen. The effect of these choices can be illustrated with simple examples as we discuss below.

8.2 Medium-Term Forecasting. Several general methods of medium-term forecasting have been proposed for time periods of months to years. Typically, these are based on the use of simulations of earthquakes, as in the following.

- Using statistical simulations such as ETAS or BASS, the observed background rate of small earthquake activity is used to drive the model, producing a statistical ensemble of possible "future" events. These simulated events, which generally follow a Gutenberg-Richter magnitude frequency relation, an Omori law of aftershock decay, and a subsidiary relation such as productivity or Bath's law, are then used to compute a probability of a future large earthquake. Note that these three equations are temporal statistics. Additional assumptions must be invoked to forecast locations of possible future events.

- Topologically realistic earthquake simulations can be used to produce large catalogs of "realistic" earthquakes using simulators such as the Virtual Quake or RSQSim models. These catalogs are then searched to find “past” sequences of events that match as well as possible the recent actual sequence observed in the given geographic region of interest. Then one uses the ensembles of the simulated "future" events to estimate the probability of a large earthquake (Van Aalsburg et al., 2007).

8.3 Nowcasting. Nowcasting is a much more recent idea that is based on similar approaches in finance and weather/climate research. The idea is to use proxy data to estimate the current hazard of the system, without assuming any type of probability model. The proxy data are used to track the changing state of the system through time. It implies, but does not explicitly state, a current and near-future level of hazard. Examples of proxy data for nowcasts of large earthquake hazards that are useful include small earthquake numbers from earthquake catalogs, GNSS and InSAR data, and other records of ongoing activity and crustal deformation.

Earthquake nowcasting (Rundle et al., 2016, 2018, 2019a, b; Pasari and Mehta, 2018; Pasari, 2019, 2020; Pasari and Sharma, 2020) can be used to define the current state of risk from large earthquakes. These methods have begun to be applied to India (Pasari, 2019), Japan (K. Nanjo

personal comm., 2020) and Greece (G. Chouliaras, personal comm. 2019). We discuss these in more detail below.

8.4 Testing. Forecasts (in particular) must be tested to determine the accuracy of the forecasts. While testing of forecasts is relatively new in the earthquake community, valuable methods can be found from similar forecasts made in the weather and climate community (see, for example, https://www.cawcr.gov.au/projects/verification/). In general, the testing schema that can be used for backtesting are methods that have been developed in the meteorological validation and verification community (Joliffe and Stephenson, 2003; Casati et al., 2008).

For our purposes, these tests are of two types: Reliability/Attributes (R/A) diagrams (Murphy, 1973; Hsu and Murphy, 1986; Murphy, 1988), and Receiver Operating Characteristic (ROC) tests (Green and Swets, 1966; Kharin and Zwiers, 2003; Joliffe and Stephenson, 2003). The R/A test is conditioned on the forecasts (i.e., given the forecast probability Y, what actually happened?). The ROC test is conditioned on the observations (i.e., given that X occurred, what was the forecast?). These tests have a long provenance and have properties that are well understood.

Tests such as Receiver Operating Characteristic (ROC) and Reliability/Attributes diagrams are used as standard methods for evaluating the accuracy and precision of forecasts and nowcasts. In many applications of machine learning, for example, use and analysis of ROC diagrams is standard, along with skill scores, correlation coefficients and other statistics derived from the underlying contingency table, which is more typically called the confusion matrix in machine learning applications.

In the following, we describe several examples of new methods and analyses, and follow with a discussion. We note that the selection of these forecast methods is not exhaustive, rather the discussion should be regarded as illustrative of the broad classes of proposed methods.

9. Examples of Proposed Earthquake Forecast and Nowcast Methods

9.1 Earthquake Cycle Models and Long-Term Forecasts. We first consider the Pacific Northwest of the United States. The earthquake (and associated tsunami) hazard in the Pacific Northwest that is primarily due to subduction of the Juan de Fuca plate beneath North America. This tectonic plate boundary is known to produce massive earthquakes that had magnitudes as large as M9, the last such event being the event of January 28, 1700 CE (Goldfinger et al., 2012, 2013). That event is recorded in geological tsunami deposits along the Pacific Northwest coast, a series of drowned forests, and corresponding as well to an "orphan tsunami" that struck the eastern coast of Japan 10 hours later.

Note in particular that this subduction zone along the Pacific Northwest was the site at which the phenomenon of Episodic Tremor and Slip (ETS) was first identified (Dragert et al., 2004) as discussed above. ETS corresponds to slow slip events accompanied by bursts of small magnitude seismicity at regular intervals of approximately 13 to 16 months. It is not known at this time whether these events signal an accumulation or a release of accumulating stress. We discuss forecasting of these events in the section on machine learning below.

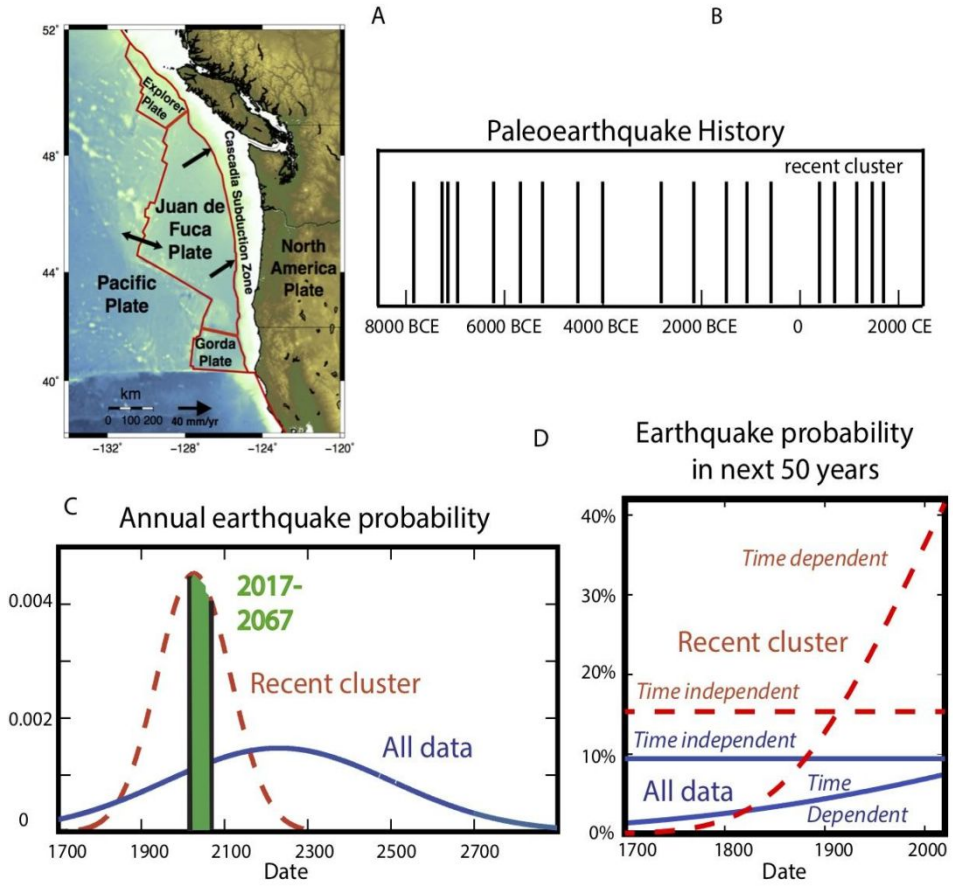

Figure 8. a) Geometry of the Cascadia subduction zone. b) Paleoearthquake history from turbidite deposits. c) Probabilities of an earthquake in the next year as a function of time assuming a Gaussian distribution of recurrence times with mean and standard deviation corresponding to the recent cluster (red/dashed lines) or the entire paleoearthquake record (blue/solid lines). Shaded area under the curves corresponds to the probability in next 50 years. d) Conditional probability of an earthquake in the next 50 years, given that last was in 1700, depending on whether we are still in the recent cluster and whether earthquake recurrence is described by a time-independent or time-independent process (Stein et al., 2017).

Although no large earthquakes have occurred along the plate interface for hundreds of years, a record of large paleoearthquakes has been compiled from subsidence data on land and using

turbidites, which are offshore deposits recording slope failure. This record (Figure 8), spanning 10,000 years, is among the world's longest (Goldfinger et al., 2012, 2013).

Other types of precursors have been studied in several contexts. For example, decadal-scale geodetic precursors have been studied by Mavrommatis et al. (2015) and Ito et al. (2013). These events were interpreted as a preparation process for the March 11, 2011 M9.1 Tohoku earthquake that eventually occurred. The geodetic events appeared to show an acceleration in the rate of recurrence of the mainshock as the rupture time approached. These studies in Japan may have a direct bearing on the similar events in the Pacific Northwest of the United States.

The recurrence intervals, differences between the dates of successive paleo-earthquakes, are key to estimating when the next may occur. The 18 intervals have a mean of 530 years and a standard deviation of 271 years. However, earthquakes seem to have happened in clusters of events, separated by 700 to 1000 year gaps. The recent cluster covering 1500 years has a mean of 326 years, and standard deviation of 88 years. Earthquakes within a cluster occur more frequently and regularly than in the full record. Hence when to expect the next earthquake depends on whether we assume that we are in the recent cluster, or that the cluster is over

Despite years of effort, seismologists have not found an optimal, compelling way to describe earthquake probabilities (Stark and Freedman, 2003; Parsons, 2008; Matthews et al., 2002; Kagan et al., 2012). Shimazaki and Nakata (1980) proposed that great earthquake cycles occur with either time-predictable or slip-predictable dynamics. In the former, if slip in past earthquakes is known, extrapolation of current rates of offset is used to predict the time of the next event. The slip-predictable model is the reverse, if the time of the next event is known, the slip can be inferred. Neither of these approaches has been found to be satisfactory for anticipating future earthquake recurrence (Rubinstein et al., 2012).

Although many methods for understanding long-term earthquake recurrence are used, all fall into two basic classes. In one, large earthquake recurrence is described by a time-independent (Poisson) process. This has no “memory,” so a future earthquake is equally likely immediately after the past one and much later. The probability of an earthquake in the next t years is approximately t/τ, where τ is the assumed mean recurrence time. Because this probability is constant, an earthquake cannot be “overdue.” Using the entire paleo-earthquake record, the chance of an earthquake in the next 50 years is 50/530=0.094 or 9.4%. Alternatively, assuming that we are still in the recent cluster gives a probability about twice as large: 50/326=0.15 or 15%.

However, seismological instincts favor earthquake cycle models, in which strain builds up slowly after an earthquake to produce the next one. In this case, the probability of a large earthquake is small immediately after one occurs and grows with time. In such time-dependent models, the recurrence interval is described by a probability density function. The simplest uses the familiar Gaussian distribution. The "bell curves" in Figure 8c show probabilities of an earthquake in the next year, which peak at dates corresponding to the assumed mean recurrence. Assuming we are in the recent cluster, the probability is high, because the 317 years since 1700CE is about the mean recurrence of 326 years. The probability is lower assuming that we are not in the cluster, because the mean recurrence for the entire record is 530 years, so we are not as far into the cycle.

To find the probability of an earthquake within 50 years, we integrate under a bell curve from a start time to 50 years in the future and include the condition that the earthquake has not happened by the start time. The resulting curves (Figure 8d), giving the conditional probabilities, are small shortly after 1700CE and increase with time. Using the entire record, the chance of an earthquake in 50 years after 2020 is 0.074 or 7.4%. However, assuming that we are still in the recent cluster gives a probability about 6 times larger: 0.41 or 41%. The higher probability results from the smaller mean recurrence time and standard deviation.

Figure 8d also shows flat lines starting at 1700CE, corresponding to time-independent models. If the time-dependent model predicts higher probability than the time-independent model, an earthquake can be considered "overdue", which occurs if we are in the cluster.

Comparing these cases shows how earthquake probability estimates depend on the probability model chosen and the data used to choose the model parameters. Other plausible choices are possible. Various probability density functions can be used. The data can be treated in more complex ways: considering different subsets, assigning different magnitudes to different paleo events, and assuming that different events broke different parts of the subduction zone. Each choice yields a different probability estimate. Thus, although it is often said that "the probability of an earthquake is N%," any estimate involves specifying the assumptions made. Different plausible assumptions yield different probabilities.

One of the more frequent methods proposed to forecast earthquakes involved changes in the Gutenberg-Richter *b*-value (see equation 1), which can be viewed as the ratio of small earthquakes to large earthquakes. The idea is that if the b-value declines in value, a large earthquake becomes relatively more probable (Haberman, 1986; 1991; Weimer and Wyss, 2000). While some of the changes in b-value may be real, an important part of that story was the appearance of many later studies of apparent seismicity changes that were at least artificial due to changes in the seismic network detection threshold over time (so, for example, as the seismic networks improved, more low magnitudes events were detected and included in the catalogs).

9.2 Pattern Informatics Method for Earthquake Forecasting. This method was developed by Tiampo et al. (2002b), Rundle et al. (2002) and Holliday et al. (2007) as means of forecasting the locations of future large earthquakes. It was tested retrospectively and prospectively in the RELM competition as described below.

The approach divides the seismogenic region to be studied into a grid of square boxes or pixels whose size is related to the magnitude of the earthquakes to be forecast. The rates of seismicity in each box are studied to quantify anomalous behavior. The basic idea is that any seismicity precursors represent changes, either a local increase or decrease of seismic activity, so our method identifies the locations in which these changes are most significant during a predefined change interval. The subsequent forecast interval is the time window during which the forecast is valid.

The Pattern Informatics (PI) method starts by constructing a spatial coarse-graining of a region such as California, i.e., a partition of grid boxes centered on the locations $\boldsymbol{x}_i$. A state vector $\psi_i(\boldsymbol{x}_i,t)$ is then constructed that, each component of which represents the number of small

earthquakes (larger than a catalog completeness level m_C) over a time interval $\Delta t = t - t_b$. One then computes the change in state vector at each location $\boldsymbol{x}_i$:

$$\Delta\psi_i(\boldsymbol{x}_i, t_1 \rightarrow t_2) = \psi_i(\boldsymbol{x}_i, t_2) - \psi_i(\boldsymbol{x}_i, t_1) \tag{28}$$

Defining an average $\langle \bullet \rangle$ over all values in the catalog $t_b \leq t$, we finally compute the Pattern Informatics (PI) value at each location $\boldsymbol{x}_i$:

$$PI(x_i, t_1 \rightarrow t_2) = \left\langle |\Delta\psi_i(\boldsymbol{x}_i, t_1 \rightarrow t_2)| \right\rangle^2 \tag{29}$$

In Figure 9 below, the Log_{10}(PI) is plotted for the Southern California region for a 5 year period corresponding to the Relative Earthquake Likelihood Model competition (Field, 2007).

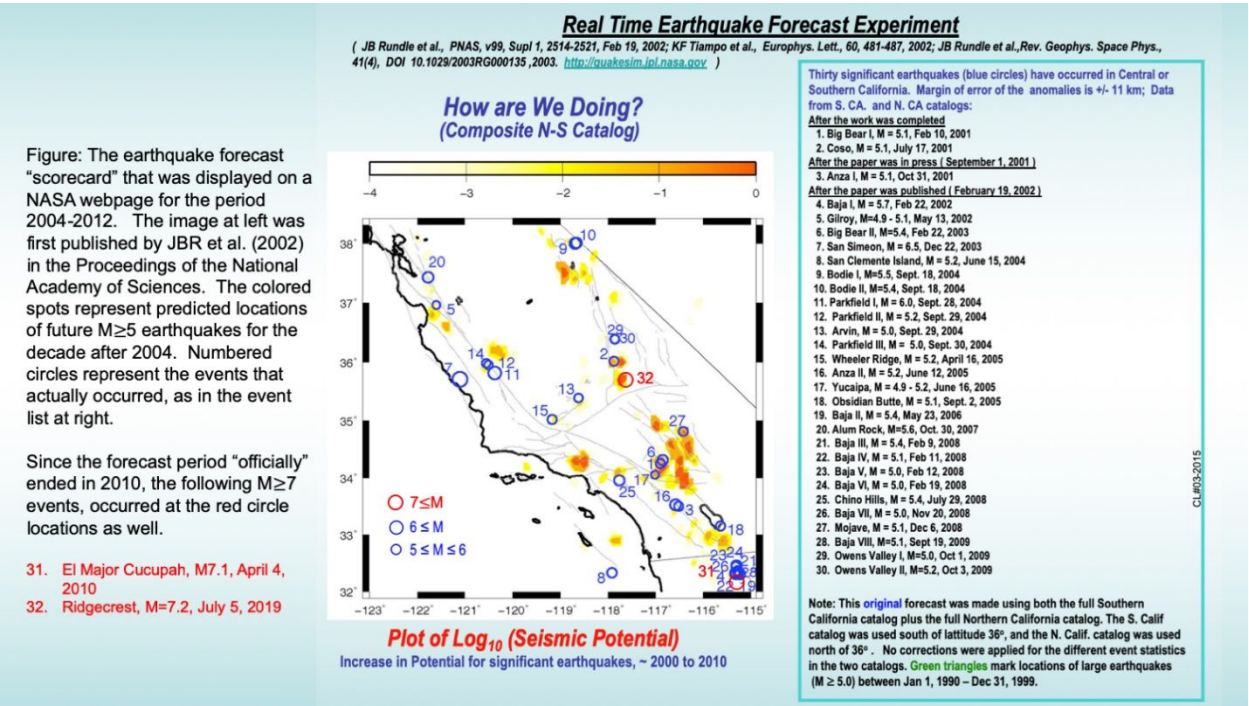

Figure 9. Earthquake forecast "scorecard" produced and displayed on a NASA web site at the request of NASA program managers. The original map with colored patches was published in 2002 in the Proceedings of the National Academy of Sciences. The small blue circles represent earthquakes that occurred over the subsequent 8 years. The red circles represent the recent M = 7.1 El Mayor-Cucapah, Mexico earthquake on April 4, 2010 (circle 31), and the M = 7.2 Ridgecrest, CA earthquake of July 5, 2019, that occurred well after the evaluation period ended. The results show that the original colored patches calculated in 2001 for the most part identified the location of the future events successfully, with few false positives, as far as 18 years into the future.

9.3 Medium-Term Forecasts: The RELM Test. The Regional Earthquake Likelihood Model (RELM) test of earthquake forecasts in California (Field, 2007) was the first competitive evaluation of forecasts of future earthquake occurrence, carried out in a completely prospective manner. A set of rules was first specified for all entrants into the competition, a common data set was defined, and a future testing period was identified. According to the terms of the competition, once a forecast for the future period of 5 years was submitted, no further changes or adjustments were allowed for the period during which the competition was held.

Participants submitted expected probabilities of occurrence of M ≥ 4.95 earthquakes in 0.1◦ × 0.1◦ cells in a region defined as greater California for the period 1 January 2006 to 31 December 2010, a 5-year interval. Probabilities were submitted for 7682 geographic grid boxes or cells in California and adjacent regions.

During this period, 31 M ≥ 4.95 earthquakes occurred in the test region. These "large" earthquakes occurred in 22 of the predefined test cells. Seismic activity during this period was dominated by earthquakes associated with the M = 7.2, 4 April 2010 El Mayor-Cucapah earthquake in northern Mexico. This earthquake occurred in the test region, and 16 of the other 30 earthquakes in the test region could be associated with it. Nine complete forecasts were submitted by six participants.

After the close of the test period, the data were analyzed by Lee et al. (2011). The results were presented in a very simple way that allowed the reader to evaluate which forecast is the most “successful” in terms of the locations of future earthquakes. Lee et al. (2011) suggest ways in which the results can be used to improve future forecasts. The six participants in the competition used very different methods to construct their forecasts. Of note were two of the forecasts, one that used the ETAS method, and one that used a method called "Pattern Informatics" (Lee et al., 2011; Rundle et al., 2002; Tiampo et al., 2002a,b), discussed below.

The results are shown in Table 1, which considers only the forecasts of whether a test earthquake was expected to occur in the cells in which earthquakes actually occurred. These probabilities are given in Table 1 and are the probabilities that a M 4.95 will occur in the i^{th} cell during the test period. The probability λ_i is normalized so that the sum of the probabilities over all cells is 22, the number of cells in which earthquakes actually occurred.

A perfect forecast would have $\lambda_i = 1$ in each of these cells and $\lambda_i = 0$ in all other cells. Seven submissions of probabilities are given in Table 1. The details of the way in which the submitted probabilities λ_i were used to obtain the normalized probabilities are given in Lee et al. (2011), along with further details of the submitted forecasts. It is also of interest to compare the submitted forecast probabilities with random (no skill) values, which is the (non-normalized) constant value $\lambda_i = 2.86 \times 10^{-3} = 22/7682$.

In Table 1, the competitors are identified as (1) Bird and Liu (B and L), (2) Ebel et al. (Ebel), (3) Helmstetter et al. (Helm.), (4) Holliday et al. (Holl.), (5) Ward combined (W-C), (6) Ward geodetic (W-G), (7) Wiemer and Schorlemmer (W and S). The 31 earthquakes are identified as (A-V). The highest (best) probabilities are designated with gray bars. However, other groups (Schorlemmer et al., 2010; Zechar et al., 2013) using analysis methods different from those of Lee et al. obtained different results.

Cell ID	EQ ID	B and L	Ebel	Helm.	Holl.	W-C	W-G	W and S
A	1, 7, 8, 16, 24	1.99E-02	2.20E-02	1.17E-01	3.32E-02	1.87E-02	1.28E-02	1.24E-01
B	2	1.41E-02	3.40E-02	7.20E-02	3.32E-02	1.08E-03	1.86E-03	4.99E-02
C	3	7.40E-03	6.59E-03	7.41E-03	3.32E-02	8.93E-04	1.54E-03	7.91E-03
D	4	3.54E-02	3.29E-02	6.97E-02	3.32E-02	9.50E-04	1.64E-03	3.59E-02
E	5	7.23E-03	1.10E-03	2.29E-03	9.72E-05	9.25E-04	1.59E-03	1.58E-07
F	6	9.37E-03	2.85E-02	3.07E-02	3.32E-02	5.29E-03	8.12E-03	4.55E-02
G	9,10	9.11E-03	5.49E-03	2.55E-02	3.32E-02	2.25E-02	1.27E-02	2.38E-02
H	11	3.42E-04	5.49E-03	9.15E-04	1.62E-04	3.77E-04	6.49E-04	2.06E-04
I	12	2.14E-03	1.10E-03	3.65E-03	2.05E-04	1.14E-03	1.96E-03	9.89E-03
J	13	1.68E-03	8.78E-03	1.11E-02	3.32E-02	8.11E-03	5.12E-03	1.13E-02
K	14	3.12E-02	2.20E-02	3.30E-02	3.32E-02	1.93E-02	1.17E-02	5.90E-02
L	15	2.07E-03	5.49E-03	6.93E-03	3.32E-03	4.80E-03	5.45E-03	2.64E-03
M	17, 18	1.74E-03	2.20E-03	5.78E-03	3.32E-02	3.88E-03	4.61E-03	5.38E-04
N	19	5.83E-02	6.59E-03	1.49E-02	3.32E-02	1.65E-02	1.23E-02	7.44E-03
O	20	1.25E-02	1.43E-02	9.45E-03	3.32E-02	9.30E-04	1.60E-03	1.62E-02
P	21	6.48E-03	3.29E-02	2.71E-02	3.32E-02	9.03E-04	1.55E-03	7.46E-03
Q	22, 25, 28	2.88E-02	2.20E-02	2.84E-02	3.32E-02	1.66E-02	1.30E-02	5.23E-02
R	23, 26	3.06E-02	1.54E-02	1.43E-02	1.73E-04	1.78E-02	1.38E-02	1.58E-02
S	27	2.13E-02	5.49E-03	1.26E-02	3.32E-02	9.55E-03	7.93E-03	1.19E-02
T	29	1.83E-02	1.32E-02	2.43E-02	3.32E-02	6.35E-03	3.90E-03	4.99E-02
U	30	1.26E-02	3.07E-02	1.03E-01	3.32E-03	1.61E-02	5.47E-03	5.16E-02
V	31	6.76E-03	1.54E-02	5.55E-03	3.32E-02	1.54E-02	1.43E-02	2.64E-03

Table 1. From Lee et al. (2011). Normalized probabilities of occurrence of an earthquake with M > 4.95 for the 22 cells in which earthquakes occurred during the test period. The association of cell id's (A - V) with the earthquake id's (1-31) from Table 1 is illustrated in Fig. 3. Seven submitted forecasts are given: (1) Bird and Liu (B and L), (2) Ebel et al. (Ebel), (3) Helmstetter et al. (Helm.), (4) Holliday et al. (Holl.), (5) Ward combined (W-C), (6) Ward geodetic (W-G), (7) Wiemer and Schorlemmer (W and S). The highest (best) probabilities are highlighted in gray.

From Table 1, the two most successful forecasts were the Holliday et al. forecast, and the Weimer-Schorlemmer forecast. The Holliday forecast led all forecasts with 8 of the highest probabilities, while the W-S forecast had 6 of the highest probabilities. Recall that the Holliday forecast used the Pattern Informatics (PI) method, whereas the W-S forecast used a method based on the Gutenberg-Richter law.

Note that the Pattern Informatics method is similar to another method named Relative Intensity (RI) method, which posits that future large earthquakes will occur at locations having the largest number of small earthquakes over a defined period of time. The PI method differs from the RI method in that for the PI method, it is assumed that the future large earthquakes will occur at locations where the change in the number of small earthquakes is the largest (Holliday et al., 2006).

Forecast testing experiments have been conducted not only in California but also in other seismically active regions, such as Italy (Taroni et al., 2018) and Japan (Nanjo et al., 2012; Ogata et al., 2013). Other authors have extended the methods such as the PI and RI methods to ensemble approaches. Cheong et al. (2014) has proposed a model based on what they term a fusion-fission process of sticking points, or asperities, on plate interfaces. These asperities are shown to coalesce in predictable statistical ways prior to major earthquakes such as the 1999 Chi-Chi, Taiwan, earthquake. Chang et al. (2020) have proposed an ensemble model, the "PI

Soup of Groups" model, for Italian earthquakes. This PISOG method was found to reduce the inherent noise in the PI method prior to the 2009 L'Aquila and 2016 Nocia, Italy, earthquakes.

9.4 Medium-Term Forecasts Based on Earthquake Simulations using the ETAS and BASS Models. In this method, statistical earthquake simulations are carried out with the intention of producing "realistic" simulated catalogs of earthquakes. As described above, these statistical simulations are based on the Gutenberg-Richter relation, the Omori relation, and either an earthquake productivity relation (ETAS), or the Båth's law for maximum aftershock magnitude (BASS). These stochastic catalogs can then be data-mined for space-time patterns of small earthquake activity that may precede large earthquakes. Similar patterns of activity would then be searched in observed earthquake activity with the idea of using these as the inputs to forecast probabilities.

Yoder et al. (2015) introduced a method of estimating the near-field (near the rupture boundary and immediately following the mainshock) spatial density and temporal rate of aftershocks based on a fractal dimension $D > 0$ model of mainshock and aftershock events. From this model, the ETAS parameter space can be tightly constrained to facilitate accurate estimates of aftershock rates and probabilities, based on earthquake scaling relations, with minimal operator input and data fitting.

The model was then compared to six $6 < M < 9$ recent earthquakes, followed by a discussion of the implications of this model both with respect to earthquake physics and as they relate to seismic hazard assessment. Note that this model is particularly well suited to automated, web-deployed, and rapid response seismic hazard applications (Yoder et al., 2015).

The model discussed in Yoder et al. (2015) is based on the BASS-ETAS model for simulating earthquake seismicity, specifically including and focusing on aftershocks. The model is initiated with a ''seed'' catalog of one or more earthquakes. This seed catalog can be based on background seismicity—suggesting regional earthquake forecast applications, or it can include only one or a few specific earthquakes—suggesting local short-term aftershock hazard applications.

Given a seed catalog, each member earthquake produces aftershocks. Each of these aftershocks is treated as an independent earthquake which produces its own aftershocks, which produce aftershocks, and so forth. In contemporary ETAS, earthquakes are treated as dimension 0 point-like objects, located at the event epicenter.

The model is parameterized so that the expected intensity of each recursive generation is weaker than its parent event(s), so eventually the process dies out. Synthetic catalogs of discrete events (earthquakes) can be generated.

In other instances, including the model presented in Yoder et al. (2015), it is sufficient to simply calculate local productivity rates, from which earthquake probability fields and foreshock statistics can be calculated directly, as a function of position and time.

This method was deployed for the April 24, 2015 $M = 7.8$ Nepal earthquake and its May 12, 2015 $M = 7.3$ aftershock. The forecast was prepared and presented at several telecons that were conducted in response to the $M = 7.8$ mainshock. On the morning (in the United States) of May 11, 2015, at 9:00 am PDT, a telecom was held to discuss NASA data needs, at which the

BASS-ETAS aftershock was presented (D. Green, personal communication, 2015). The aftershock forecast correctly predicted the location of the M = 7.3 aftershock that occurred a few hours later on May 12, 2015 Nepal Standard Time.

9.5 Forecasts Based on Earthquake Simulations, Virtual Quake and the Uniform California Earthquake Rupture Forecast (UCERF). As in the BASS-ETAS simulation approach, catalogs of realistic earthquake ruptures can be produced using the topologically realistic earthquake simulators such as the Virtual Quake model. These catalogs can be used to produce forecasts, by matching patterns of simulated activity to observed earthquake histories, and then using the optimally matched catalogs to project activity forward in time. In general, an ensemble approach is used in which future earthquake probabilities are computed by considering groups of simulations.

As a group, topologically realistic earthquake simulators are high performance computational simulations that include 1) realistic and detailed fault geometries, 2) loading of the faults so that the earthquake rate averaged over the previous thousands of years is at the observed rate, 3) interactions between faults mediated by elastic or perhaps viscoelastic stress transfer, and 4) friction laws motivated by laboratory or field observations. They may also be extended to include seismic wave simulation codes and can and will be used to provide the initial conditions for the wave codes.

A project led by T. Tullis (Brown University) compared the results of various existing simulators to determine how well their statistics agreed with each other, and how well they agreed with the observations of observed earthquake statistics (Tullis, 2012). The project involved the creation of an archive of long-term synthetic catalogs of earthquakes in California using these simulators. There were four simulations in the archive:

1. Virtual California (“VirtCal”, now called Virtual Quake). A summary of this simulation has been given by Rundle et al. (1994) and more recently Sachs et al. (2012)
2. RSQSim. A summary of this simulation approach has been discussed by Richards-Dinger and Dieterich (2010).
3. All Cal. A summary has been given by Ward (2000).
4. ViscoSim. A summary has been given by Pollitz (2011).

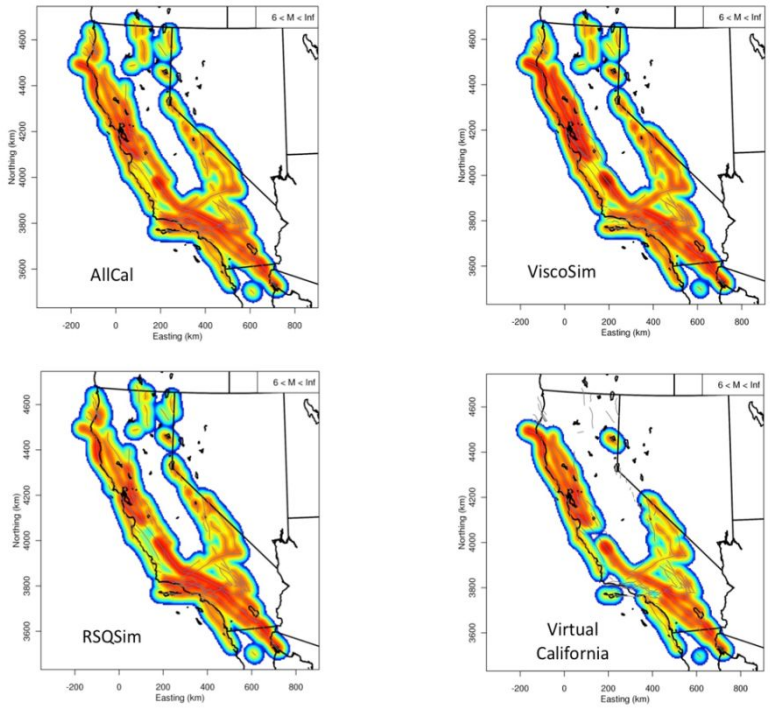

Figure 10. Time independent probabilities of earthquake occurrence obtained by the four simulations. (after Tullis et al., 2012). We note that the Uniform California Earthquake Rupture Forecast project (UCERF) has adopted this type of methodology for the forecasting of California earthquakes. It is in development by other countries around the world as well. As of this time, the reasons for differences in the four results have not been identified.

These simulations were used to generate time-independent (Poisson) probabilities of occurrence for earthquakes in California. Simulations have been used to generate synthetic catalogs having durations of tens of thousands of years on the different fault boundary elements in the models. These catalogs were then used to compute rates of earthquake occurrence for specified magnitude bands. As in other rate-based models, the rates were then smeared over several tens of km radius to generate earthquake rate maps similar to those computed in the UCERF models. Forecast probability (rate) maps of occurrence for earthquakes having $M \geq 6$ per year per km^2 are given in Figure 10 above.

Another example of the use of VQ-type simulations in forecasting is based on interval statistics on the fault segments. In Rundle et al. (2005), it was shown that in 40,000 years of simulations for the San Francisco section of the San Andreas fault system, 395 events occurred having an average recurrence interval of 101 years. Using the time intervals between successive events, estimated probabilities were constructed for earthquakes in the San Francisco area.

To use these data in forecasting real earthquakes, time t is measured forward from the time of occurrence of the last great earthquake. The time is the time of the last great earthquake. For

San Francisco, this is now 114 years ago. The waiting time is measured forward from the present. Given a best-fitting probability distribution, as well as the current waiting time since the last major earthquake, one can use the simulation data to compute a conditional probability for the next major earthquake on the San Andreas fault near San Francisco (Wesnousky et al., 1984).

That is, given the time of the last large earthquake and a forecast time interval, one can estimate the probability of the next such large earthquake from the simulation data. As Rundle et al. (2005) showed, the Weibull distribution, which is commonly used in failure and reliability analysis, provides an excellent fit to both the distribution of interval times as well as the statistics of the waiting times. In both simulations and in our Weibull fit, the median waiting times systematically decrease with increases in the time since the last great earthquake.

9.6 Natural Time. Natural time is an idea originally proposed by (Varotsos et al., 2001; 2002; 2013; 2014; 2020; Sarlis et al., 2018). Natural time is just the count or number of small earthquakes between large earthquakes. It can be contrasted with the clock time, which is the days-weeks-months-years between large earthquakes. They developed the idea to analyze seismic electric signals prior to large earthquakes. Seismic electric signals (SES) are low frequency (≤1Hz) transient changes of the electric field of the Earth that have been found to precede earthquakes with lead time ranging from several hours to a few months and their analysis enables the estimation of the epicentral area. As discussed above, they introduce an event count, the natural time of a series of events, and proposed that its variance was an order parameter for the second order seismic phase transition, an idea discussed above. They found that the variance k_1 of the seismic electric signals approached a critical value for large earthquakes in Greece (Varotsos et al., 2001; 2002).

Subsequently, they found that these ideas applied to events in Japan as well (Varotsos et al., 2013). Furthermore, Varotsos et al. (2014) applied detrended fluctuation analysis to the seismic electric time series, and they find that the earthquake magnitude time series exhibits several minima as the time of the earthquake approached.

Sarlis et al. (2018) define an entropy in terms of natural time and find that it exhibits a minimum in the Olami-Feder-Christensen (Olami et al., 1992: OFC) model for earthquakes, a cellular automaton slider block model. With this idea, they examined the seismicity leading up to the March 11, 2011 M9.3 Tohoku earthquake. They observed similar fluctuations before that event. Further details can be found in those papers.

The natural time concept was adopted by Holliday et al. (2006). As in the preceding, the idea is to use counts of small earthquake events as a measure of “time” between large events. The motivation for this idea is the Gutenberg-Richter magnitude-frequency relation, which specifies that on average, a statistically average number of small earthquakes is associated with a larger earthquake of a given size.

For example, in California the seismic *b*-value, which characterizes this relation, is observed to be b=0.85, implying that there are 357 magnitude M > 3 earthquakes for every M > 6 earthquake. Note that this relationship is stated as an exceedance (survivor distribution function, sdf) rather than a probability density function (pdf). Thus, the natural time interval between M > 6 earthquakes is on average a count of 357 M > 3 events.

9.7 Statistical Forecasts Using Natural Time. As an example of how natural time may be used in forecasting, we consider the Natural Time Weibull (NTW) forecast method (Holliday et al 2016; Rundle et al, 2016). The Natural Time Weibull model is based on the idea that progress through the seismic cycle of small earthquakes in a region influences the probability for the next major earthquake. It is therefore of interest to consider the statistical structure of the seismic cycle, building on the idea of nowcasting discussed below. It is a Bayesian forecast model since it explicitly assumes a conditional probability distribution, the Weibull (1952) model, to project the current activity forward in time. This idea can be extended by combining statistical data of this type with data from GNSS, InSAR, and other types of data. Examples can be found in Rundle et al., (2016a)

To illustrate how to compute the NTW probabilty, we define a large seismically active area A that includes dozens of large earthquakes of M > 6. From the Gutenberg-Richter relation, we find the number N of small earthquakes that are expected for each large earthquake. We identify the most recent of these large earthquakes and compute the number $n(t)$ of small earthquakes 6> M ≥ 3 that have occurred since the last large event.

Using the Weibull distribution (equation 18), we compute the probability that a large earthquake will occur following an additional Δn small earthquakes, conditioned on the fact that no large earthquake has occurred prior to the current count of small earthquakes $n(t)$:

$$P(n+\Delta n|n) \; = 1. \; - \; exp\left\{ -\left(\frac{n+\Delta n}{N}\right)^{\beta} + \left(\frac{n}{N}\right)^{\beta} \right\} \tag{30}$$

To estimate the future clock time (as opposed to the natural time) at which a major earthquake might occur, we assume that the current average (or “Poisson”) rate ν of small earthquakes continues to hold for short intervals into the future. Thus, to convert natural time to clock time, we set:

$$\Delta n \; \approx \; \nu \Delta t \tag{31}$$

The best fitting value for the exponent β is found by retrospective testing to be $\beta \approx 1.4$ (Rundle et al., 2012).

9.8 Nowcasting Methods. As described previously, Rundle et al (2016b, 2018, 2019) and Luginbuhl et al. (2018a,b, 2019) applied the idea of nowcasting to seismically active regions to determine the current state of the fault system, and its current level of progress through the earthquake cycle. In the implementation of this idea, they used the global catalog of earthquakes, using "small" earthquakes to determine the level of hazard from "large" earthquakes in the region.

In the past, this determination of the state of a regional fault system has focused on trying to estimate the state of stress in the earth, its relation to the failure strength of the active faults in a region, and the rate of accumulation of tectonic stress (Scholz, 2019). Determining the values of these parameters would allow researchers to estimate the proximity to failure of the faults in the

region. This would be an answer to the question of “how far along is the region in the earthquake cycle?".

The nowcasting method is based on the idea of an earthquake cycle as discussed above. A specific region and a specific large earthquake magnitude of interest are defined, ensuring that there is enough data to span at least ~20 or more large earthquake cycles in the region, which holds for the example below. An "earthquake potential score" (EPS) is then defined as the cumulative probability distribution P(n < n(t)) for the current count n(t) for the small earthquakes in the region. Physically, the EPS corresponds to an estimate of the level of progress through the earthquake cycle in the defined region at the current time.

An example application of this method is shown in Figure 11, which shows the EPS for the region surrounding Los Angeles within a circle of radius 150 km, for earthquakes of magnitude M ≥ 6 (Rundle et al., 2016b; 2018). The last such earthquake was the Northridge, CA earthquake of January 17, 1994. The green vertical bars represent a histogram of the number of small earthquakes between large earthquakes M ≥ 6 in a region 4000 km x 4000 km surrounding Los Angeles. The solid red line is the corresponding cumulative distribution function (CDF). The thin dashed lines represent the 68% confidence bound on the CDF. The red dot represents the number of small earthquakes that have occurred in the region since the Northridge event.

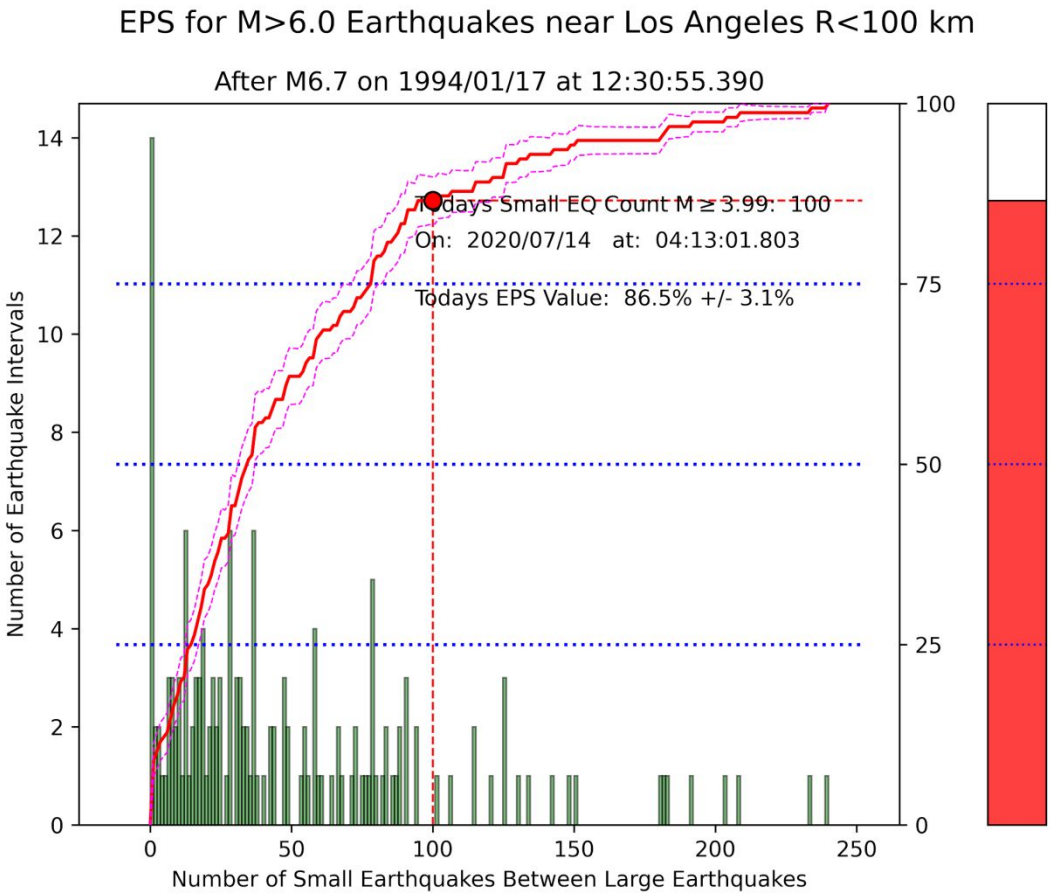

Figure 11. Earthquake Potential Score (EPS) determined by the nowcasting method. Green vertical bars represent the histogram of number of small earthquakes M ≥ 3.99 within a region of 4000 km^2 around Los Angeles between large earthquakes of M ≥ 6. Solid red curve is the cumulative distribution function corresponding to the histogram. Dashed lines represent the 68% confidence interval. Red dot represents the number of small earthquakes within 100 km of Los Angeles, indicating that the EPS for this region around Los Angeles is 86.5%, or alternatively, that this region has progressed 86.5% through the typical earthquake cycle of M ≥ 6 events in the Los Angeles region. "Thermometer" on the right is a visual representation of the EPS.

In addition to this method, Rundle et al. (2019) have shown that Shannon information entropy can also be used to define the EPS. Shannon information entropy was developed to characterize the information content transmitted between a source and a receiver by means of a communication channel (Shannon, 1948; Cover and Thomas, 1991; Stone, 2015).

Nowcasting methods are also being applied to earthquakes in India (Pasari, 2019; Pasari and Mehta 2018; Pasari and Sharma, 2020), Greece (G. Chouliaras, personal comm., 2020), and Japan (K. Nanjo, personal comm., 2020). In the case of India, Pasari and Mehta (2018) showed that EPS values for events having magnitudes M ≥ 6 in a 300 km circular area in New Delhi, Chandigarh, Dehradun and Shimla reach about 0.56, 0.87, 0.85 and 0.88, respectively. For events M ≥ 6 in a 250 km circular area around Dhaka and Kohlkata are 72% and 40% respectively (Pasari 2019). Results for other regions in India are listed in Pasari and Sharma (2020a, b).

In a recent paper, Perez-Oregon et al. (2020) have shown how to transform nowcasting models into forecasting models for two model systems, one being the slider block model of Olami-Feder Christensen, and the other being a system in which the events obey a log-normal distribution. These are toy models as described above but may be applicable to real data. The forecast methods are tested by means of the Receiver Operating Characteristic method as briefly described above in the section on testing and found to produce high quality results.

9.9 Machine Learning. Machine learning is a generic term that includes a variety of supervised and unsupervised methods to extract patterns and other types of information from data (e.g., Geron, 2019; Kong et al., 2019; Trugman, 2017). Methods of interest in this area include not only cluster analysis (Unsupervised learning), but also optimization (supervised learning), a form of regression. Here we describe several very new approaches that make use of machine learning for the purpose of anticipating future earthquake activity and relating that activity to the underlying physics (Burkov, 2019).

9.9.1 Seismic Bursts and Radial Localization. As discussed in Rundle and Donnellan (2020), seismic bursts are sequences of small earthquakes strongly clustered in space and time and include seismic swarms and aftershock sequences. A readily observable property of these events is the radius of gyration of the event locations which was found by Rundle and Donnellan (2020) to connect the bursts to the temporal occurrence of the largest M ≥ 7 earthquakes in California since 1984.

Their definition of a seismic burst is the occurrence of an unusual sequence of earthquakes closely clustered in space and time (i.e., Hill and Prejean, 2007; Peresan and Gentili, 2018; Zaliapin and Ben-Zion, 2016a, b). They define two general types of bursts, Type I and Type II:

- A Type I seismic burst is a mainshock-aftershock sequence, in which the initiating event has the largest magnitude in the sequence and is typically followed by a power-law Omori decay of occurrence of smaller events (Omori, 1900; Scholz, 2019).
- A Type II seismic burst is defined as a sequence of similar magnitude events in which the largest magnitude event is not the initiating event, and in which there is not typically a subsequent power-law decay.

The earthquakes defining the bursts are small, usually of magnitudes characterizing the catalog completeness level. For the Southern California region, they consider small earthquakes of magnitudes M ≥ 3.3. This magnitude threshold was chosen as a value high enough to ensure completeness of the catalog data used. The catalog containing these events is downloaded from the US Geological Survey earthquake search database.

Rundle and Donnellan (2020) investigated the Southern California region contained within a 600 km circle surrounding Los Angeles, California. They also consider time series beginning after 1984/1/1, after the data became most reliable in terms of catalog completeness, with accurate locations. The region is arbitrary in terms of method but requires a complete catalog to be adequately applied and tested. If small earthquakes are missing from the catalog, the clusters so defined will likewise not be correctly defined—they will have missing events. Or potentially important clusters will not be present at all.

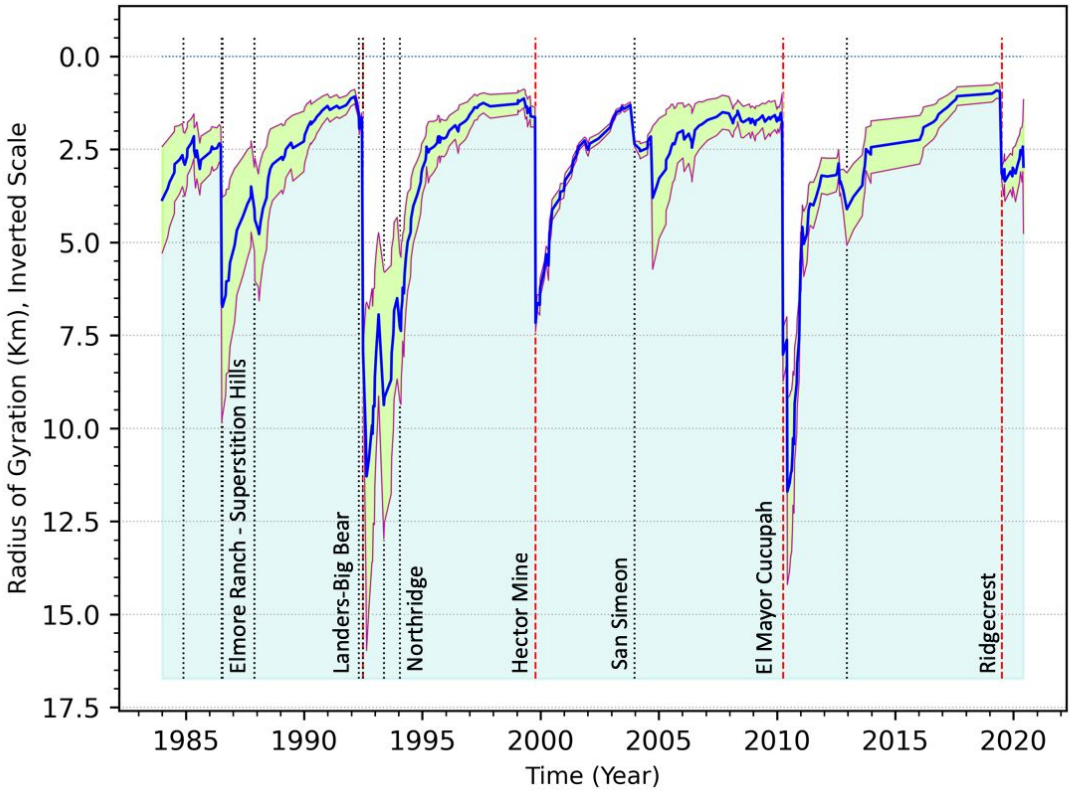

Figure 12. Change in the Radius of Gyration of small earthquake bursts in time computed using an exponential moving average (EMA). Region considered is a circle of radius 400 km around Los Angeles. Small earthquakes had minimum magnitude of M ≥ 3.29. Vertical red dashed lines are the 4 large earthquakes with magnitudes M>7 earthquakes since 1985: Landers (June 28, 1992 M = 7.3); Hector Mine (October 16, 1999 M = 7.1); El Mayor-Cucapah (April 4, 2010 M = 7.2); Ridgecrest (July 5, 2019 M = 7.1). Vertical black dotted lines are the times at which events having 6 ≤ M < 7 occurred.

In the Southern California earthquake catalog, hundreds of these potentially coherent space-time structures were identified in a region defined by a circle of radius 600 km around Los Angeles. The horizontal radius of gyration, R_G, was computed for each cluster, the radius of gyration R_G being defined as:

$$R_G^2 = \frac{1}{N_C}\sum_i \ (X_i - X_{CM})^2 + (Y_i - Y_{CM})^2 \tag{32}$$

where X_{CM}, Y_{CM} are the longitude and latitude of the center of mass (centroid) of the burst (cluster), X_i, Y_i are the longitude and latitude of the epicenters of the small earthquakes that make up the burst, and N_C is the number of small earthquakes in the burst.,

The bursts are then filtered to identify those bursts with large numbers of events closely clustered in space, which Rundle and Donnellan (2020) call "compact" bursts. The basic assumption is that these compact bursts reflect the dynamics associated with large earthquakes.

Once the burst catalog is filtered, an exponential moving average is applied to construct a time series for the Southern California region. The R_G of these bursts systematically decreases prior to large earthquakes, in a process that we might term "radial localization." The R_G then rapidly increases during an aftershock sequence, and a new cycle of "radial localization" then begins.

These time series display cycles of recharge and discharge reminiscent of seismic stress accumulation and release in the elastic rebound process as described in their Figure 2. The complex burst dynamics observed are evidently a property of the region, rather than being associated with individual faults. This new method improves earthquake nowcasting for evaluating the current state of hazard in a seismically active region. An example of this phenomenon is shown in the figure.

As Figure 12 shows, the minimum radius of gyration (minimum R_G) prior to M ≥ 7 mainshocks in California is typically 1 to 2 km. Achievement of each of these R_G values was followed within 1 to 3 years by an M ≥ 7 earthquake, the only exception being the M = 6.5 December 22, 2003 San Simeon earthquake.

However, the time series recovered from that event and soon evolved towards the minimum RG again. It is found that no M ≥ 7 earthquakes are observed at R_G values greater than the ensemble mean value of 2.5 Km. For that reason, R_G = 2.5 km can be considered to define a "low risk" threshold for M ≥ 7 earthquakes.

In more recent work, we have shown that similar time series illustrating apparent stress accumulation and release for the seismic cycle in Southern California can be computed using Principal Component Analysis and timeseries prediction and machine learning methods. (Rundle et al., 2020).

9.9.2 Supervised Machine Learning Using Decision Tree Classification. Recently Rouet-Leduc et al. (2017) and Hulbert et al. (2018) have shown that machine learning might be useful in earthquake forecasting applications, using laboratory experiments on sheared rock samples of model faults. The experiments were stick-slip events in sheared model faults and were intended to model the physics of earthquake failure in a simple way. They find that the acoustic signals produced during the shearing events can be identified using machine learning methods. Specifically, their goal was to predict the time of failure of the sheared rock by analysis of the acoustic emission signals.

To carry out this program, they constructed labeled feature vectors using 100 of the potentially relevant statistical properties of the acoustic signals. These properties were the mean, variance, kurtosis and other higher moments of the acoustic emission signals in a boxcar moving time window, each window overlapping the previous one by 90%. The labels for the training set represented the time remaining until the next major slip event. They then used a random forest classifier (Burkov, 2019), each time window using 1000 decision trees, to predict the time remaining to the next major slip event. Random forest classifiers are easily programmed using the scikit-learn library of Python codes. They found that the classifier performed well, with a coefficient of determination R^2 = 0.89, much better than a naïve model based on the periodicity of events, for which R^2 = 0.3.

Rouet-Leduc et al. (2019) then applied this method to the seismic sequences observed as episodic tremor and slip in the Cascadia subduction zone off the coast of the Pacific Northwest of the United States. In these events, quiescent periods of roughly 14 months are punctuated by sudden rapid slip along with the emission of small seismic signals during a short time window. In a similar way to the laboratory experiments, they find that their method can predict the displacement rates along the fault with a Pearson correlation coefficient of 0.66.

9.9.3 Supervised Learning Using Convolutional Neural Nets. It has been suggested by studies of models of earthquake faults and fault systems that the existence of Gutenburg-Richter(GR) scaling implies that earthquakes are associated with fluctuations at a critical point (Klein et al., 2007; Klein et al., 2000a, b, c; Rundle and Jackson, 1977; Olami et al., 1992; Rundle et al., 2003) and the concepts of scaling and renormalization group would imply that the physics of the events is the same on all scales (Stanley, 1971; Wilson, 1983 Serino et al., 2011). This would imply that there is no information in small events that can be used to characterize large events. The question then naturally arises as to whether machine learning algorithms would be effective in forecasting large events.

To investigate the efficiency of machine learning in systems with GR scaling Pun et al. (2020) study the predictability of event sizes in the Olami-Feder-Christensen (Olami et al., 1992) model, which simulates an earthquake fault, at different proximities to the critical point using a convolutional neural network. The distribution of event sizes satisfies a power law with a cutoff for large events. They find that predictability decreases as criticality is approached and that prediction

is possible only for large, non-scaling events. Their results suggest that earthquake faults that satisfy Gutenberg-Richter scaling are difficult to forecast.

Pun et al. (2020) address the question of predictability near and at criticality by applying machine learning to the OFC model. Previous work (Pepke and Carlson, 1994) showed that predictability in the OFC model decreases as the conservative limit (a critical point) is approached. Pun et al. (2020) find consistent results and investigate the predictability of events near another critical point in the OFC model: the recently observed noise transition critical point (Matin et al., 2020). By using a convolutional neural network (CNN), they find that the event sizes are more difficult to forecast as the critical point is approached and that only large events that do not satisfy power-law scaling can be successfully predicted.

Pun et al. (2020) found evidence that events whose size distribution satisfies a power law lack distinguishable features that allow the machine to predict their size. This lack of distinguishable features is related to the difficulty of distinguishing between the fluctuations and the background at critical points (Coniglio and Klein, 1980). For the large nonscaling events, there exists features that allow the machine to successfully predict the event sizes. Similar conclusions are found for the dissipation (Christensen and Olami, 1992) transition. Their results suggest that large nonscaling events are qualitatively different from the smaller scaling events. This conclusion agrees with the conjecture (Bak and Chen, 1989) that prediction is not possible at a true critical point, where there is no deviation from a power law for large events.

10. Earthquake Early Warning

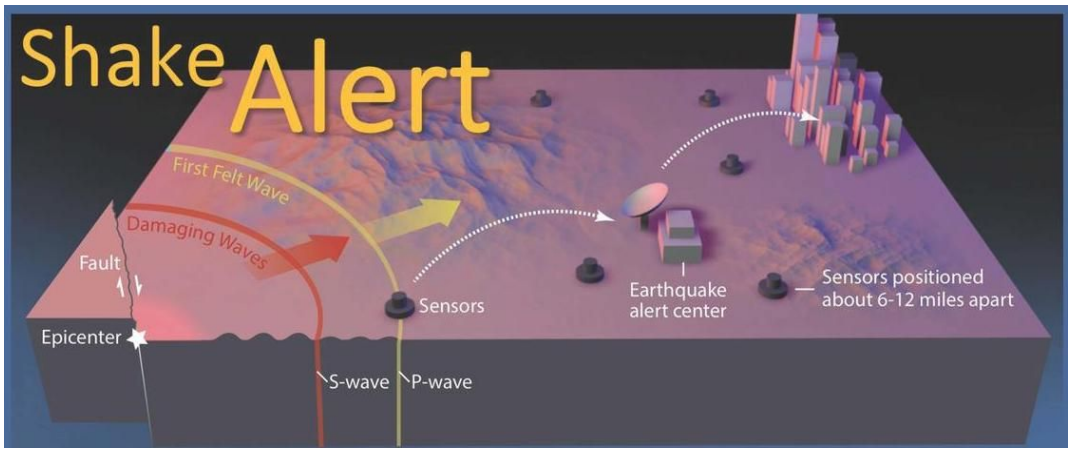

Figure 13. Schematic diagram of an earthquake early warning system (USGS).

As described in the introduction, the focus of this review was intended to be on the anticipation of earthquake ruptures. However, once the rupture has begun modern technology makes it possible to alert distant locations of impending strong shaking so that preparations can be made, an emerging technology of Earthquake Early Warning (EEW).

In an EEW system, local seismic networks automatically and immediately locate an earthquake and determine its size (Allen, 2013). P-waves recorded at seismic stations near the epicenter are used to rapidly estimate whether shaking due to later-arriving S waves larger S-waves will reach

a level deemed hazardous. If so, a warning can be sent ahead of the S-wave arrivals to more distant locations where weaker shaking is expected. The warning is relayed via Internet or mobile phones (Figure 13).

Although the warning times are short, a few tens of seconds at most, they could in principle be useful in seismically active areas along the west coast of the US. People could take cover and medical procedures can be immediately stopped. Generators could come online rapidly. Automatic systems could slow or stop trains, elevators, and airport takeoffs and landing, and shut down or secure sensitive facilities such as power plants.

Interesting questions are being studied about what information can be given to different potential users. In general, users who are willing and able to take action quickly at low levels of shaking will get more warning time in exchange for more false alarms, with earthquakes that do not produce high levels of shaking. How useful the warnings will be is unclear (Minson et al., 2018; Wald, 2020). Because strong shaking decays rapidly with distance from an earthquake, areas with the greatest shaking will receive little to no warning; areas with moderate shaking will likely receive a short, ~10 s, warning; and areas with light shaking areas will most likely receive a significant warning, 10 s or more.

Other nations with significant seismic activity already deploy EEW systems. These nations and economies include Japan, Taiwan, and Mexico.

11. Conclusion

In this paper, we have discussed recent developments in earthquake forecasting and nowcasting. We have briefly summarized some of the history of earthquake prediction studies and their search for precursory phenomena. In fact, the Parkfield earthquake prediction experiment, and the eventual M6.0 September 28, 2004 earthquake, led to a fundamental reappraisal of the possibility of deterministic, short-term, earthquake prediction. As discussed above, most researchers are currently pessimistic about the prospect. Much of the early enthusiasm was based on the few successful predictions that had been made, such as the M7.3, February 4, 1975, Haicheng, China prediction, while ignoring the many false alarms and failures-to-predict.

It is now recognized that earthquake forecasts must be evaluated using standard statistical tests that do not assume in advance any properties of the unknown statistical interevent distribution. Evaluation tests of forecast methods must consider not only successful forecasts, but also false alarms and failures-to-predict. These can be properly included using confusion matrices, also called contingency tables (Geron, 2019), leading to Receiver Operating Characteristic (ROC) curves, Reliability diagrams, and Skill Scores.

One of the more interesting tests was the Regional Earthquake Likelihood Models (RELM) contest. As discussed above, RELM was a true prospective test of medium-term earthquake forecasting methods, in which a prescribed data set was used. The contest was for 5 years into the future, and the results were subsequently made freely available for analysis by anyone (Schorlemmer and Gerstenberger, 2007; Lee et al., 2011). While not definitive, the test demonstrated that locations of moderate earthquakes can be determined to some degree in advance of occurrence. The conclusion generally showed that larger earthquakes are most likely

to occur at locations where the most small earthquakes occur. Fundamentally, this result follows from the Gutenberg-Richter magnitude-frequency relation (Scholz, 2019), that large earthquakes tend to be accompanied by a given number of small earthquakes specified by the *b*-value scaling exponent.

In the future, machine learning (ML) will play an increasingly important role in the development of earthquake forecasting models. ML methods have numerous advantages over previous approaches, in that they can utilize a diversity of data types, they are less subject to perceived human biases, and they can be evaluated objectively. The key to these methods is the feature engineering step, wherein the data are arranged in feature vectors, and appropriate labels are chosen. Some methods, such as decision trees, random forests, and convolutional neural networks, may have advantages over other methods. Disadvantages may appear in the form of long computing times, or difficulty of code construction, but the existence of standard open access libraries such as Scikit-Learn, Tensorflow and Keras have considerably simplified this step.

We therefore expect that the development of earthquake forecasting methods utilizing machine learning represents the future of this field, given the increasing rate of accumulation of high-quality data, and the steadily increasing compute power available.

11. Acknowledgements

We first thank the editors of Reports on Progress in Physics for the invitation to write this review article. Research by JBR and CS has been supported by a grant from the US Department of Energy to the University of California, Davis, number DoE grant number DE- SC0017324, and by a grant from the National Aeronautics and Space Administration to the University of California, Davis, number NNX17AI32G. Contributions by SS have been supported by Northwestern's Institute for Policy Research, Buffet Institute for Global studies, and Institute for Complex Systems. Research by Andrea Donnellan was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. Contributions by WK were supported by Boston University. The authors would also like to acknowledge helpful comments by M. Bevis and 2 anonymous referees.

12. References

Alava, M. J., Nukala, P.K.V.V. and Zapperi, S. Statistical models of fracture. *Advances in Physics*, 55, 349 - 476 (2006).

Allen, R. Seismic hazards; seconds count: *Nature*, v. 502, no. 7469 (2013).

Ammon, C., Ji, C., Thio, H., Robinson, D., Ni, S., Hjörleifsdóttir, V., Kanamori, H., Lay, T., Das, S., Helmberger, D., Ichinose, G., Polet, J. and Wald, D. Rupture Process of the 2004 Sumatra-Andaman Earthquake. *Science*, 308, 1133-1139 (2005).

Bahar Halpern, K., Tanami, S., Landen, S., Chapal, M., Szlak, L., Hutzler, A., Nizhberg, A. and Itzkovitz, S. Bursty gene expression in the intact mammalian liver. *Molecular Cell*, 58(1), 147 (2015).

Bak, P and Chen, K. The physics of fractals. *Physica D*, 38, 5 (1989).

Bakun, W.H. and Lindh, A.G. The Parkfield, California, Earthquake Prediction Experiment. *Science*, 229, 619-624 (1985).

Bakun, W.H., Aagaard, B., Dost, B., Ellsworth, W.L., Hardebeck. J.L., Harris, R.A., Ji, C., Johnston, M.J.S., Langbein, J., Lienkaemper, J.J., Michael, A.J., Murray, J.R., Nadeau, K.M., Reasenberg, P.A., Reichle, M.S., Roeloffs, E.A., Shakal, A., Simpson, R.W. and Waldhauser, F. Implications for prediction and hazard assessment from the 2004 Parkfield earthquake. *Nature*, 437, 969-974 (2005).

Ben-Zion, Y. Collective Behavior of Earthquakes and Faults: Continuum-Discreet Transitions, Progressive Evolutionary Changes and Different Dynamic Regimes. *Reviews of Geophysics*, 46(4) (2008).

Beran, J., Feng, Y., Ghosh, S. and Kulik, R. *Long-Memory Processes*. Springer-Verlag, Berlin, Heidelberg (2013).

Blewitt, G., Heflin, M.B., Hurst, K.J., Jefferson, D.C., Webb, F.H. and Zumberge, J.F. Absolute far-field displacements from the 28 June 1992 Landers earthquake sequence. *Nature*, 361, 340-342 (1993).

Bock, Y., Agnew, D.C., Fang, P., Genrich, J.F., Hager, B.H., Herring, T.A., Hudnut, K.W., King, R.W., Larsen, S., Minster, J.-B., Stark, K., Wdowinski, S. and Wyatt, F.K. Detection of crustal deformation from the Landers earthquake sequence using continuous geodetic measurements. *Nature*, 361, 337-340 (1993).

Bowman, D.D. and King, G.C.P. Accelerating seismicity and stress accumulation before large earthquakes. *Geophys. Res. Lett.*, 28, 4039-4042 (2001).

Broberg, K.B. *Cracks and Fracture*. Academic Press, New York (1999).

Brooks, B.A., Merrifield, M.A., Foster, J., Werner, C.L., Gomez, F., Bevis, M. and Gill, S. Space geodetic determination of spatial variability in relative sea level change, Los Angeles basin. *Geophys. Rev. Lett.*, 34(1) (2007).

Brudzinski, Michael R., and Richard M. Allen. "Segmentation in episodic tremor and slip all along Cascadia." Geology 35.10 (2007): 907-910.

Bufe, C.G., and Varnes, D.J. Predictive Modeling of the Seismic Cycle of the Greater San Francisco Bay Region. *J. Geophys. Res.*, 98, 9871–9883 (1993).

Bufe, C.G., Nishenko, S.P. and Varnes, D.J. Seismicity trends and potential for large earthquakes in the Alaska-Aleutian region. *PAGEOPH*, 142**,** 83–99 (1994).

Burkov, A. *The Hundred-Page Machine Learning Book*. Andriy Burkov, Quebec City, Canada (2019). ISBN-13: 978-1999579500 .ISBN-10: 199957950X

Burridge, R.O. and Knopoff, L. Model and theoretical seismicity. *Bull. Seism. Soc. Am.*, 57, 341-371 (1967).

Carlson, J.M. and Langer, J.S. Mechanical model of an earthquake fault. *Phys. Rev. A, Gen.Phys.*, 40(11), 6470-6484 (1989).

Casati, B., Wilson, L.J., Stephenson, D.B., Nurmi, P., Ghelli, A., Pocernich, M., Damrath, U., Ebert, E.E., Brown, B.G. and Mason, S. Forecast verification: current status and future directions. *Meteorol. Appl.*, 15, 3-18 (2008).

Chang, L.-Y., Chen, C.-C., Telesca, L., Li, H.-C. and Cheong, S.A. Pattern Informatics and the Soup-of-Groups Model of Earthquakes: A Case Study of Italian Seismicity. *Pure Appl. Geophys.*, 177, 4089-4096 (2020).

Cheong, S.A., Tan, T.L., Chen, C.-C., Chang, W.-L., Liu, Z. Chew, L.Y., Sloot, P.M.A. and Johnson, N.F. Short-term forecasting of Taiwanese earthquakes using a universal model of fusion-fission processes. *Sci. Rep.,* 4, 3624 (2014).

Chen, C.-C., Lee, Y.-T. and Chiao, L.-Y. Intermittent criticality in the long-range connective sandpile (LRCS) model. *Physics Letters A*, 372, 4340-4343 (2008a).

Chen, C.-C., Lee, Y.-T. and Chang, Y.-F. A relationship between Hurst exponents of slip and waiting time data of earthquakes. *Physica A,* 387, 4643-4648 (2008b).

Christensen K., and Olami, Z. Scaling, phase transitions, and nonuniversality in a self-organized critical cellular-automaton model. *Phys. Rev. A*, 46, 1829 (1992).

Coniglio, A. C. and Klein, W. Correlated site-bond percolation and Ising critical droplets. *J. Phys. A*, 13, 2775 (1980).

Console, R., Lombardi, A.M., Murru, M. and Rhoades, D. Bath's law and the self-similarity of earthquakes. *J. Geophys. Res.*, 108(B2), 2128 (2003).

Cover, T.M. and Thomas, J.A. Elements of Information Theory, John Wiley, NY (1991).

Davis, T.L., Namson, J. and Yerkes, R.F. A cross-section of the Los Angeles area: Seismically active fold and thrust belt, the 1987 Whittier Narrows earthquake, and earthquake hazard. *J. Geophys. Res.*, 94, 9644-9664 (1989).

Delouis, B., Nocquet, J.-M., and Vallée, M. Slip distribution of the February 27, 2010 Mw = 8.8 Maule Earthquake, central Chile, from static and high-rate GPS, InSAR, and broadband teleseismic data. *Geophys. Res. Lett.*, 37(17) (2010).

Dieterich, J. H., Modeling of rock friction: 1. Experimental results and constitutive equations, *Journal of Geophysical Research*: Solid Earth 84.B5 (1979): 2161-2168.

Dieterich, J. H., Earthquake nucleation on faults with rate-and state-dependent strength, *Tectonophysics* 211.1-4 (1992): 115-134.

Dieterich, J.H., A constitutive law for rate of earthquake production and its application to earthquake clustering, *Journal of Geophysical Research: Solid Earth* 99.B2 (1994): 2601-2618.

Donnellan, A., Hager, B.H. and King, R.W. Discrepancy between geological and geodetic deformation rates in the Ventura Basin. *Nature*, *366*(6453), 333-336 (1993a).

Donnellan, A., Hager, B.H., King, R.W. and Herring, T.A. Geodetic measurement of deformation in the Ventura Basin region, southern California. *Journal of Geophysical Research: Solid Earth*. *98*(B12), 21727-21739 (1993b).

Donnellan, A., Parker, J., Hensley, S., Pierce, M., Wang, J. and Rundle, J. UAVSAR observations of triggered slip on the Imperial, Superstition Hills, and East Elmore Ranch Faults associated with the 2010 M 7.2 El Mayor-Cucapah earthquake. *Geochem., Geophys., Geosys.*, *15*(3), 815-829 (2014).

Donnellan, A., Grant-Ludwig, L., Parker, J.W., Rundle, J.B., Wang, J., Pierce, M., Blewitt, G. and Hensley, S. Potential for a large earthquake near Los Angeles inferred from the 2014 La Habra earthquake. *Earth and Space Science*, 2(9), 378-385 (2015).

Donnellan, A., Pierce, M., Wang, J., Ben-Zion, Y., Lou, Y., Padgett, C., Parker, J., Hawkins, B., Granat, R., Glasscoe, M., Rundle, J.B. and Grant-Ludwig, L. The Quakes Concept for Observing and Mitigating Natural Disasters. *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium* (2019).

Dragert, H., K. Wang, and T. S. James, A Silent Slip Event on the Deeper Cascadia Subduction Interface, Science, 292, 1525–1528, 2001.

Dragert, H., K. Wang, and G. Rogers. "Geodetic and seismic signatures of episodic tremor and slip in the northern Cascadia subduction zone." *Earth, planets and space,* 56.12, 1143-1150 (2004).

Dreger, D.S., Huang, M.H., Rodgers, A., Taira, T., and Wooddell, K. Kinematic finite-source model for 24 August 2014 South Napa, California, earthquake from joint inversion of seismic, GPS and InSAR data. *Seism. Res. Lett*., 86(2A), 327-334 (2015).

Ebrahimi, F. Invasion percolation: A computational algorithm for complex phenomena, *CiSE*, 12, 84-93 (2010).

Feigl, K.L., Agnew, D.C., Bock, Y., Dong, D., Donnellan, A., Hager, B.H., Herring, T.A., Jackson, D.D., Jordan, T.H., King, R.W. and Larsen, S. Space geodetic measurement of crustal deformation in central and southern California, 1984–1992. *J. Geophys. Res.: Solid Earth*. 98(B12), 21677-21712 (1993).

Field, Edward H., Overview of the working group for the development of regional earthquake likelihood models (RELM), Seismological Research Letters 78.1 (2007): 7-16.

Field, E.H., Biasi, G.P., Bird, P., Dawson, T.E., Felzer, K.R., Jackson, D.D., Johnson, K.M., Jordan, T.H., Madden, C., Michael, A.J., Milner, K.R., Page, M.T., Parsons, T., Powers, P.M., Shaw, B.E., Thatcher, W.R., Weldon, R.J., II, and Zeng, Y., 2013, Uniform California earthquake rupture forecast, version 3 (UCERF3)—The time-independent model: U.S. Geological Survey Open-File Report 2013–1165. *California Geological Survey Special Report 228,* and *Southern California Earthquake Center Publication* 1792, 97 (2013). http://pubs.usgs.gov/of/2013/1165/.

Field, E. H., Arrowsmith, J R., Biasi, G. P., Bird, P., Dawson, T. E., Felzer, K. R., Jackson, D. D., Johnson, K. M., Jordan, T. H., Madden, C., Michael, A. J., Milner, K. R., Page, M. T. Uniform California Earthquake Rupture Forecast, Version 3 (UCERF3) --The Time-Independent Model. *Bull. Seism. Soc. Am.*, Vol. 104, No. 3, 1122-1180 (2014).

Fielding, E.J., Milillo, P., Burgmann, R., Yun, S.H. and Samsonov, S.V. Coseismic and postseismic deformation from the August 2014 Mw 6.0 South Napa earthquake measured

with InSAR time series. American Geophysical Union Fall Meeting 2014, American Geophysical Union (2014).

Flossman, A.I., Hall, W.D. and Pruppacher, H.R. Theoretical Study of the Wet Removal of Atmospheric Pollutants. Part I: The Redistribution of Aerosols Captured through Nucleation and Impaction Scavenging by Growing Cloud Drops. *J. Atmos. Sci.*, 42(6), 583-606 (1985).

Fritz, H.M., Petroff, C.M., Catalán, P.A., Cienfuegos, R., Winckler, P., Kalligeris, N., Weiss, R., Barrientos, S.E., Meneses, G., Valderas-Bermejo, C., Ebeling, C., Papadopoulos, A., Contreras, M., Almar, R., Dominguez, J.C. and Synolakis, C.E. Field Survey of the 27 February 2010 Chile Tsunami. *Pure Appl. Geophys.*, 168, 1989–2010 (2011).

Garcimartin, A., Guarino, A., Bellon, L. and Ciliberto, S. Statistical properties of fracture precursors. *Phys. Rev. Lett.,* 79, 3202-3205 (1997).

Geller, R.J. Earthquake prediction: a critical review. *Geophys. J. Int.*, 131, 425-450 (1997).

Geron, A. *Hands-On Machine Learning with Scikit-Learn, Keras and Tensorflow*. O'Reilly, Boston (2019).

Gibbs, J.W. On the Equilibrium of Heterogeneous Substances. *Transactions of the Connecticut Academy of Arts and Sciences*, 3, 108–248; 343–524 (1878).

Glowacka, E., Sarychikhina, O., Suarez, F., Nava, F.A. and Mellors, R. Anthropogenic subsidence in the Mexicali valley, Baja California and slip on the Saltillo fault. *Eviron. Earth Sci.*, 59, 1515-1524 (2010).

Goh, K.I. and Barabasi, A.L. Burstiness and memory in complex systems. *Europhys. Lett.* 81(4) (2008).

Goldberg, D. E., Melgar, D., Bock, Y. and Allen, R.M. Geodetic Observations of Weak Determinism in Rupture Evolution of Large Earthquakes. *J. Geophys. Res.,* 123 (2018).

Goldberg, D. E., Melgar, D. and Bock, Y. Seismogeodetic P-wave amplitude: No evidence for strong determinism. *Geophys. Res. Lett.,* 46, 11118–11126 (2019).

Goldfinger, C., Ikeda, Y. and Yeats, R.S. Superquakes, supercycles, and global earthquake clustering: Recent research and recent quakes reveal surprises in major fault systems. *Earth magazine* (2013).

Goldfinger, C., Nelson, C., Morey, A., Johnson, J., Patton, J., Karabanov, E., Gutiérrez-Pastor, J., Eriksson, A., Gràcia, E., Dunhill, G., Enkin, R., Dallimore, A. and Vallier, T. Turbidite Event History: Methods and Implications for Holocene Paleoseismicity of the Cascadia Subduction Zone. *U.S. Geological Survey Professional Paper* 1661-F (2012).

Green, D.M. and Swets, J.M. Signal detection theory and psychophysic*.* New York: John Wiley and Sons (1966).

Griffith, A.A. The Phenomena of Rupture and Flow in Solids. *Philos. Trans. of the Royal Society Series A*, 221(582-593) 163-169 (1921).

Guilhem, A., Bürgmann, R., Freed, A.M. and Ali, S.T. Testing the accelerating moment release (AMR) hypothesis in areas of high stress. *Geophys. J. Int.*, 195, 785–798 (2013).

Gunton, J. D., San Miguel, M. and Sahni, P. The dynamics of first order phase transitions. In: Domb, C. and Lebowitz, J.L. *Phase Transitions and Critical Phenomena*, Vol. 8, Academic Press, New York (1983).

Guyer, R., McCall, K. and Boitnott, G. Hysteresis, discrete memory and nonlinear wave propagation in rock: A new paradigm, *Phys. Rev. Lett.*, 74, 3491-3494 (1995).

Guyer, R. and Johnson, P. *Nonlinear Mesoscopic Elasticity: The Complex Behavior of Granular Media including Rocks and Soil*. Wiley-VCH, Berlin, 410 (2009).

Hager, B.H., King, R.W. and Murray, M.H. Measurement of crustal deformation using the Global Positioning System. *Annual Review of Earth and Planetary Sciences*. 19(1), 351-382 (1991).

Hamill, P., Toon, O.B. and Kiang, C.S. Microphysical Processes Affecting Aerosol Particles. *J. Atmos. Sci.*, 34, 1104-1119 (1977).

Hanks, T.C. and Kanamori, H. A moment magnitude scale. *J. of Geophys. Res. Solid Earth*. 84(B5), 2348-2350 (1979).

Hashin, Z. and Shtrickman, S. On some variational principles in anisotropic and nonhomogeneous elasticity. *J. Mech. Phys. Solids*, 10, 335-342 (1962).

Hayashi, T. Japan's Post-Disaster Economic Reconstruction: From Kobe to Tohoku. *Asian Econ. J.*, 26(3), 189-210 (2012).

Hefferan, C. M., Li, S.F., Lind, J. and Suter, R. Tests of microstructure reconstruction by forward modeling of high energy X-ray diffraction microscopy data. *Powder Diffraction*, 25(2), 132 (2010).

Hegg, D.A. and Baker, M.B. Nucleation in the atmosphere. *Rep. Prog. Phys.*, 72(5) (2009).

Heien, E.M. and Sachs, M.K. Understanding long term earthquake behavior through simulation. *Computing in Science and Engineering*, 14(5), 10-20 (2012).

Helmstetter, A. and Sornette, D. Predictability in the Epidemic-Type Aftershock Sequence model of interacting triggered seismicity. *J. Geophys. Res.*, 108(B10) (2003).

Herring, T.A., Melbourne, T.I., Murray, M.H., Floyd, M.A., Szeliga, W.M., King, R.W., Phillips, D.A., Puskas, C.M., Santillan, M. and Wang, L. Plate Boundary Observatory and related networks: GPS data analysis methods and geodetic products. *Reviews of Geophysics*, *54*(4), 759-808 (2016).

Hill, D. P. and Prejean, S.G. Dynamic triggering. Treatise on Geophysics, vol. 4, Earthquake Seismology, edited by H. Kanamori. Elsevier, Amsterdam, 257-291 (2007).

Holliday, JR, Rundle, J.B., Tiampo, K.F., Klein, W. and Donnellan, A. Systematic procedural analysis of the pattern informatics method for forecasting large (M>5) earthquakes in southern California. *Pure. Appl. Geophys.*, 163(11-12), 2433-2454 (2006).

Holliday, J.R., Chen, C.C., Tiampo, K.F., Rundle, J.B., Turcotte, D.L. and Donnellan, A. A RELM earthquake forecast based on Pattern Informatics. *Seism. Res. Lett.*, 78, 87-93 (2007).

Holliday, J.R., Graves, W.R., Rundle, J.B. and Turcotte, D.L. Computing earthquake probabilities on global scales. *Pure. Appl. Geophys.*, 173, 739-748 (2016).

Hough, S.E. *Richter's Scale: Measure of an Earthquake, Measure of a Man*. Princeton University Press, Princeton (2007).

Hsu, W.-R. and Murphy, A.H. The attributes diagram: A geometrical framework for assessing the quality of probability forecasts. *Int. J. Forecasting*, 2, 285-293 (1986).

Hudnut, K.W., Bock, Y., Cline, M., Fang, P., Feng, Y., Freymueller, J., Ge, X., Gross, W.K., Jackson, D., Kim, M., King, N.E., Langbein, J., Larsen, S.C., Lisowski, M., Shen, Z.-K., Svarc, J. and Zhang, J. Co-seismic displacements of the 1992 landers earthquake sequence. *Bull. of the Seism. Soc. of Am.* 84(3), 625-645 (1994).

Hudnut, K.W., Bock, Y., Galetzka, J.E., Webb, F.H. and Young, W.H. The southern California integrated GPS network (SCIGN). In *The 10th FIG International Symposium on Deformation Measurements*. Orange California, USA, 19-22 (2001).

Hulbert, C., Rouet-Leduc, B., McBrearty, I.W., and Johnson, P.A. Patterns in Seismic Energy and Earthquake Hazard in Northern Chile. American Geophysical Union, Fall Meeting 2018 (2018).

Ishii, M., Shearer, P.M., Houston, H. and Vidale, J.E. Extent, duration and speed of the 2004 Sumatra–Andaman earthquake imaged by the Hi-Net array. *Nature*, 435**,** 933–936 (2005).

Ito, Yoshihiro, et al. Episodic slow slip events in the Japan subduction zone before the 2011 Tohoku-Oki earthquake. *Tectonophys*. 600 14-26 (2013).

Jaumé S.C. and Sykes L.R. Evolving Towards a Critical Point: A Review of Accelerating Seismic Moment/Energy Release Prior to Large and Great Earthquakes. In: Wyss, M., Shimazaki, K. and Ito, A. Seismicity Patterns, their Statistical Significance and Physical Meaning. *PAGEOPH Topical Volumes*, Birkhäuser, Basel (1999).

Johanson, I.A. and Burgmann, R. Coseismic and postseismic slip from the 2003 San Simeon earthquake and their effects on backthrust slip and the 2004 Parkfield earthquake. *J. Geophys. Res*., 15 (2010).

Joliffe, I.T. and Stephenson, D.B. *Forecast Verification: A Practitioners' Guide in Atmospheric Science*, John Wiley & Sons, New Jersey (2003).

Jones, L., K. Aki, M. Celebi, A. Donnellan, J. Hall, R. Harris, E. Hauksson, T. Heaton, S. Hough, K. Hudnut, K. Hutton, M. Johnston, W. Joyner, H. Kanamori, G. Marshall, A. Michael, J. Mori, M. Murray, D. Ponti, P. Reasenberg, D. Schwartz, L. Seeber, A. Shakal, R. Simpson, H. Thio, M. Todorovska, M. Trifunic, D. Wald, and M. L. Zobak. The Magnitude 6.7 Northridge California, Earthquake of January 17, 1994. *Science*, 266, 389–397 (1994).

Kagan, Y.Y., Jackson, D.D., and Geller, R.J. Characteristic earthquake model, 1884- 2011, R.I.P. *Seismol. Res. Lett.*, 83, 951–953 (2012).

Kanamori, H. Earthquake Prediction: An Overview. 1205-1216. In: Lee, W.H.K., Kanamori, H., Jennings, P.C. and Kisslinger, C. *International Handbook of Earthquake and Engineering Seismology*. Academic Press, Amsterdam (2003).

Kanninen, M.F. and Popelar, C.H. *Advanced Fracture Mechanics*. Oxford University Press, New York (1985).

Karsai, M., Kimmo, K., Barabási, A.-L. and Kertész, J. Universal features of correlated bursty behavior. *Scientific Reports*, 2, 397 (2012).

Keilis-Borok, V. Earthquake Prediction: State-of-the-Art and Emerging Possibilities. *Ann. Rev. Earth Planet. Sci.*, 30, 1-33 (2002).

Kelton K.F. and A.L. Greer. *Nucleation in Condensed Matter: Applications in Materials and Biology*. Elsevier, Boston (2010).

Kerr, R.A. Parkfield quakes skip a beat. *Science,* 259*,* 1120-1122 (1993).

Kerr, R.A. Parkfield keeps secrets after a long-awaited quake. *Science,* 306*,* 206-207 (2004).

Kharin, V.V. and Zwiers, F.W. On the ROC score of probability forecasts*. J. Climate*, 16, 4145-4150 (2003).

Klein, W., Rundle, J.B. and Ferguson, C.D. Critical phenomena and metastability in models of earthquake faults, *Phys. Rev. Lett.*, 78, 3793 (1997).

Klein, W., Ferguson, C.D., Anghel, M., Rundle, J.B. and Sa' Martins, J.S. Cluster analysis in earthquake fault models with long-range interactions, *AGU Monograph*. In: Rundle, J.B., Turcotte, D.L. and Klein, W. *Geocomplexity and the Physics of Earthquakes* (2000a).

Klein, W., Gould, H., Tobochnik, J., Alexander, F.J., Anghel, M. and Johnson, G. Clusters and fluctuations at mean-field critical points and spinodals, *Phys. Rev. Lett.*, 85, 1270 (2000b).

Klein, W., Anghel, M., Ferguson, C.D., Rundle, J.B. and Sa' Martins, J.S. Statistical Analysis of a Model for Earthquake Faults with Long Range Stress Transfer. In: Rundle, J.B. Turcotte, D.L. and Klein, W. *GeoComplexity and the Physics of Earthquakes*, American Geophysical Union, Washington D.C. (2000c).

Klein, W., Gould, H., Gulbahce, N., Rundle, J.B., Tiampo, K.F. Structure of fluctuations near mean-field critical points and spinodals and its implication for physical processes. *Phys. Rev. E*, 75, Art. 031114 (2007).

Klein, W., Xia, J., Ferguson, C.D., Gould, H., Tiampo, K.F. and Rundle, J.B. Models of Earthquake Faults: Ergodicity and Forecasting, *Int. J. Mod. Phys. B*, 23, 5553-5564 (2009).

Knackstedt, M., Sahimi, M. and Sheppard, A.P. Invasion percolation with long-range correlations: First-order phase transition and nonuniversal scaling properties, *Phys. Rev. E*, 61, 4920-4934 (2000).

Kong, Q., Trugman, D. T., Ross, Z. E., Bianco, M. J., Meade, B. J. and Gerstoft, P. Machine learning in seismology: Turning data into insights. *Seism. Res. Lett.*, 90(1), 3-14 (2019).

Koster, U., Punge-Witteler, B. and Steinbrink, G. Surface Crystallization of Metal-Metalloid Glasses. *Key Engineering Materials*, 40-41, 53-62 (1990).

LaBrecque, J., Rundle, J.B. and Bawden, G. Global Navigation Satellite System Enhancement for Tsunami Early Warning Systems. *Global Assessment Report on Disaster Risk Reduction, UN Office for Disaster Risk Reduction, Geneva, Switzerland* (2019).

Larson, K.M., Freymueller, J.T. and Philipsen, S. Global plate velocities from the Global Positioning System. *J. of Geophys. Res.: Solid Earth*, *102*(B5), 9961-9981 (1997).

Lay, T., Kanamori, H., Ammon, C., Nettles, M., Ward, S., Aster, R., Beck, S., Bilek, S., Brudzinski, M., Butler, R., Deshon, H., Ekström, G., Satake, K. and Sipkin, S. The Great Sumatra-Andaman Earthquake of 26 December 2004. *Science*, 308, 1127-1133 (2005).

Lay, T., Ammon, C. J., Kanamori, H., Koper, K. D., Sufri, O., and Hutko, A. R. Teleseismic inversion for rupture process of the 27 February 2010 Chile (M_w 8.8) earthquake, *Geophys. Res. Lett.*, 37 (2010).

Lay, T. and Kanamori, H. Insights from the great 2011 Japan earthquake. *Phys. Today*, 64(12), 33 (2011).

Lee, Y.-T., Chen, C.-C., Chang, Y.-F. and Chiao, L.-Y. Precursory phenomena associated with large avalanches in the long-range connective sandpile (LRCS) model. *Physica A*, 387, 5263-5270 (2008).

Lee, Y.T., Turcotte, D.L., Holliday, J.R., Sachs, M.K., Rundle, J.B., Chen, C.C., and Tiampo, K.F. Results of the Regional Earthquake Likelihood Models (RELM) test of earthquake forecasts in California. *Proc. Nat. Acad. Sci.*, 108, 16533-16538 (2011).

Lee, Y.-T., Turcotte, D.L., Rundle, J.B. and Chen, C-C. A statistical damage model with implications for precursory seismicity. *Acta Geophysica*, 60, 638-663 (2012).

Li, S. F., Lind, J., Hefferan, C.M., Pokharel, R., Lienert, U., Rollett, A.D. and Suter, R.M. Three-dimensional plastic response in polycrystalline copper via near-field high-energy Xray diffraction microscopy. *J. Appl. Crystallogr.*, 45, 1098 (2012).

Lorenz, E., *The Essence of Chaos*. University of Washington Press, Seattle (1995).

Luginbuhl, M., Rundle, J.B., Hawkins, A. and Turcotte, D.L. Nowcasting earthquakes: a comparison of induced earthquakes in Oklahoma and at the Geysers, California. *Pure Appl. Geophys.*, 175(1), 49-65 (2018a).

Luginbuhl, M., Rundle, J.B. and Turcotte, D.L. Natural time and nowcasting earthquakes: are large global earthquakes temporally clustered? *Pure Appl. Geophys.*, 175, 661-670 (2018b).

Luginbuhl, M., Rundle, J.B. and Turcotte, D.L. Natural time and nowcasting induced seismicity at the Groningen gas field in the Netherlands. *Geophys. J. Int.*, 215, 753-759 (2018c).

Luginbuhl, M., Rundle, J.B. and Turcotte, D.L. Natural Time and Nowcasting Earthquakes: Are Large Global Earthquakes Temporally Clustered? In: Williams, C., Peng, Z., Zhang, Y., Fukuyama, E., Goebel, T., Yoder, M. (eds) Earthquakes and Multi-hazards Around the Pacific Rim, Vol. II. PAGEOPH Topical Volumes. Birkhäuser, Cham. (2019).

Mantegna, R.N. and Stanley, H.E. *An Introduction to Econophysics, Correlations and Complexity in Finance*. Cambridge University Press, Cambridge (2004).

Massonnet, D., Rossi, M., Carmona, C., Adragna, F., Peltzer, G., Feigi, K., and Rabaute, T. The displacement field of the Landers earthquake mapped by radar interferometry. *Nature*, 364, 138-142 (1993).

Matin, S., Pun, C.-K., Gould, H. and Klein, W. Effective ergodicity breaking phase transition in a driven-dissipative system. *Phys. Rev. E*, 101 (2020).

Matthews, M.V., Ellsworth, W.L., and Reasenberg, P.A. A Brownian model for recurrent earthquakes. *Bull. Seismol. Soc. Am.*, 92, 2233–2250 (2002).

Mavrommatis, A.P., Segall, P. and Johnson, K.M. A decadal-scale deformation transient prior to the 2011 Mw 9.0 Tohoku-oki earthquake. *Geophys. Res. Lett.*, 41(13), 4486-4494 (2015).

Melbourne, T.I., Webb, F.H., Stock, J.M. and Reigber, C. Rapid postseismic transients in subduction zones from continuous GPS. *J. Geophys. Res.: Solid Earth*, 107(B10) (2002).

Melgar, D. and Bock, Y. Near-field tsunami models with rapid earthquake source inversions from land and ocean-based observations: The potential for forecast and warning. *J. Geophys. Res*., 118, 1-17 (2013).

Melgar, D., Crowell, B.W., Geng, J. and Allen, R.M. Earthquake magnitude calculation without saturation from the scaling of peak ground displacement: GPS PGD SCALING. *Geophys. Res. Lett.*, 42(13) (2020).

Minson, S.E., Meier, M., Baltay, A.S., Hanks, T.C., Cochran, E.S. The limits of earthquake early warning: Timeliness of ground motion estimates. *Science Advances* 4(3) (2018).

Mogi, K. *Earthquake Prediction*. Academic Press, Tokyo (1985).

Mori, N., Takahashi, T. and the 2011 Tohoku earthquake tsunami joint survey group. Nationwide Post Event Survey and Analysis of the 2011 Tohoku Earthquake Tsunami. *Coast. Eng. J.*, 54(1), 1250001-1-1250001-27 (2012).

Muller, R., Zanotto, E.D. and Fokin, V.M. Surface Crystallization of Silicate Glasses: Nucleation Sites and Kinetics. *J. Non-Cryst. Solids*, 274, 208-231 (2000).

Murphy, A.H. A new vector partition of the probability score. *J. Appl. Meteor.*, 12, 595-600 (1973).

Murphy, A.H. Skill scores based on the mean square error and their relationships to the correlation coefficient. *Mon. Wea. Rev*., 116, 2417-2424 (1988).

Nanjo, K.Z., Tsuruoka, H., Yokoi, S., Ogata, Y., Falcone, G., Hirata, N., Ishigaki, Y., Jordan, T.H., Kasahara, K., Obara, K., Schorlemmer, D., Shiomi, K. and Zhuang, J. Predictability study on the aftershock sequence following the 2011 Tohoku-Oki, Japan, earthquake: first results. *Geophys. J. Int.*, 191(2), 653-658 (2012).

Norris, J.Q., Turcotte, D.L. and Rundle, J.B. Loopless nontrapping invasion-percolation model for fracking. *Phys. Rev. E*., **89** (2014).

Obara, Kazushige, and Shutaro Sekine. "Characteristic activity and migration of episodic tremor and slow-slip events in central Japan." Earth, planets and space 61.7 (2009): 853-862.

Ogata, Y. Statistical models for earthquake occurrences and residual analysis for point processes. *J. Am. Stat. Assoc.,* 83(401), 9–27 (1988).

Ogata, Y. Space-time point-process models for earthquake occurrences. *Ann. Inst. Stat. Math*, 50(2), 379–402 (1998).

Ogata, Y. Space-time model for regional seismicity and detection of crustal stress changes. *J. Geophys. Res. Solid Earth*, 109(B3) (2004).

Ogata, Y. Significant improvements of the space-time ETAS model for forecasting of accurate baseline seismicity. *Earth Planets Space*, 63(3), 217–229 (2011).

Ogata, Y., Katsura, K., Falcone, G., Nanjo, K. and Zhuang, J. Comprehensive and Topical Evaluations of Earthquake Forecasts in Terms of Number, Time, Space and Magnitude. *Bull. Seismic. Soc. Am.*, 103(3), 1692-1708 (2013).

Okada, Y. Internal deformation due to shear and tensile faults in a half space. *Bull. Seism. Soc. Am.*, 82, 1018-1040 (1992).

Okubo, S. Gravity and potential changes due to shear and tensile faults in a half space. *J. Geophys. Res*., 97, 7137-7144 (1992).

Olami, Z., Feder, H.J.S. and Christensen, K. Self Organized Criticality in a Continuous Non-Conservative Cellular Automaton Modeling Earthquakes. *Phys. Rev. Lett.*, 68, 1244-1247 (1992).

Omori, F. Note on the aftershocks of the Hokkaido-earthquake of March 22nd, 1894, *Publ. Earthquake Invest. Comm. in Foreign Languages*, 4 (1900).

Paczuski, M., Maslov, S. and Bak, P. Avalanche dynamics in evolution, growth, and depinning models. *Phys. Rev E*, 53, 414-443 (1996).

Paris, R., Lavigne, F., Wassmer, P. and Sartohadi, J. Coastal sedimentation associated with the December 26, 2004 tsunami in Lhok Nga, west Banda Aceh (Sumatra, Indonesia). *Mar. Geol.*, 238(1-4), 93-106 (2007).

Parsons, T. Earthquake recurrence on the south Hayward fault is most consistent with a time dependent, renewal process. *Geophys. Res. Lett.*, 35 (2008).

Pasari, S. Nowcasting earthquakes in the Bay-of-Bengal region. *Pure Appl. Geophys.* 23, 537-559 (2019).

Pasari, S. Stochastic Modeling of Earthquake Interevent Counts (Natural Times) in Northwest Himalaya and Adjoining Regions. In: Bhattacharyya, S., Kumar, J. and Ghoshal, K. Mathematical Modeling and Computational Tools, *Springer Proceedings in Mathematics & Statistics*, 320, 495-501, Springer, Singapore (2020).

Pasari, S., and Mehta, A. 2018. Nowcasting earthquakes in the northwest Himalaya and surrounding regions. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-5, 855–859 (2018).

Pasari, S., and Sharma, Y. Contemporary Earthquake Hazards in the West-Northwest Himalaya: A Statistical perspective through Natural Times. *Seismol. Res. Lett.* (in print) (2020).

Pepke, S. L. and Carlson, J.M. Predictability of self-organizing systems. *Phys. Rev. E*, 50, 236 (1994).

Peresan, A. and Gentili, S. Seismic cluster analysis in Northeastern Italy by the nearest-neighbor approach. *Phys. Earth Planet Interiors*, 274, 87-104 (2018).

Perez-Oregon, Jennifer, Fernando Angulo-Brown, and Nicholas Vassiliou Sarlis. "Nowcasting Avalanches as Earthquakes and the Predictability of Strong Avalanches in the Olami-Feder-Christensen Model." Entropy 22.11 (2020): 1228.

Pollitz, F. F. ViscoSim earthquake simulator. *Seismol. Res. Lett*., 83, 979–982 (2012).

Pun, C.K., Matin, S., Klein, W. and Gould, H. Prediction in a Driven Dissipative System Displaying a Continuous Phase Transition using Machine Learning. *Phys. Rev. E*, 101 (2020).

Reasenberg, P. Second-order moment of central California seismicity, 1969–1982. *J. Geophys. Res.*, 90(B7), 5479–5495 (1985).

Reid, H.F. The mechanics of the earthquake, the California earthquake of April 18, 1906, *Report of the State Investigation Commission*, *Vol.* 2. Carnegie Institution of Washington, Washington, D.C. (1910).

Richards-Dinger, K., and Dieterich, J.H. RSQSim earthquake simulator, *Seismol. Res. Lett*., 83, 983–990 (2012).

Rouet-Leduc, B., Hulbert, C., Lubbers, N., Barros, K., Humphreys, C.J. and Johnson, P.A. Machine Learning Predicts Laboratory Earthquakes. *Geophys. Res. Lett.,* 44(18) (2017).

Rouet-Leduc, B., Hulbert, C. and Johnson, P.A. Continuous chatter of the Cascadia subduction zone revealed by machine learning. *Nature Geoscience*, 12, 75-79 (2019).f

Roux, J.-N. and Wilkinson, D. Resistance jumps in mercury injection in porous media. *Phys. Rev. A*, 37, 3921-3926 (1988).

Roux, S. and Guyon, E. Temporal development of invasion percolation, *J Phys. A: Math. Gen.*, 22, 3693-3705 (1989).

Rubinstein, J.L., Ellsworth, W.L., Chen, K.H., and Uchida, N. Fixed recurrence and slip models better predict earthquake behavior than the time- and slip-predictable models: 1. Repeating earthquakes. *J. Geophys. Res.*, 117 (2012).

Ruhl, C.J., Melgar, D., Grapenthin, R. and Allen, R.M. The value of real-time GNSS to earthquake early warning. *Geophys. Res. Lett.*, 44(16), 8311-8319 (2017).

Rundle, J. B. and Jackson, D.D. Numerical Simulation of Earthquake Sequences, *Bull. Seismol. Soc. Am.*, 67, 1363-1377 (1977).

Rundle, John B., and Marcia McNutt. "Southern California uplift—Is it or isn't it?." *Eos, Transactions American Geophysical Union*, 62.10, 97-98 (1981)

Rundle, J.B. A physical model for earthquakes: 2. Application to southern California, *J. Geophys. Res.*, 93, 6255 - 6274 (1988).

Rundle, J. B. and Klein, W. Nonclassical nucleation and growth of cohesive tensile cracks. *Phys. Rev. Lett., 63*, 171–174 (1989).

Rundle, J. B. and Klein, W. Nonlinear dynamical models for earthquakes and frictional sliding: An overview. In: The 33rd U.S. Symposium on Rock Mechanics (USRMS), 3-5 June, Santa Fe, New Mexico. American Rock Mechanics Association (1992).

Rundle, J.B., Klein, W., Gross, S. and Turcotte, D.L. Boltzmann fluctuations in numerical simulations of nonequilibrium threshold systems. *Phys. Rev. Lett.*, 75, 1658-1661 (1995).

Rundle, J.B., Klein, W. and Gross, S. Dynamics of a traveling density wave model for earthquakes. *Phys. Rev. Lett.*, 76, 4285-4288 (1996).

Rundle, J.B., Preston, E., McGinnis, S., and W. Klein. Why earthquakes stop: Growth and arrest in stochastic fields. *Phys. Rev. Lett.*, 80, 5698 (1998).

Rundle, P.B., Rundle, J.B., Tiampo, K.F., S. Martins, J.S., McGinnis, S., and Klein, W. Nonlinear network dynamics on earthquake fault systems, *Phys. Rev. Lett.,* 87 (2001).

Rundle, J.B., Rundle, P.B., Klein, W., Martins, J.S.S., Tiampo, K.F., Donnellan, A. and Kellogg, L.H. GEM plate boundary simulations for the Plate Boundary Observatory: Understanding the physics of earthquakes on complex fault systems. *Pure Appl. Geophys.*, 159, 2357-2381 (2002).

Rundle, J. B., Turcotte, D. L., Shcherbakov, R., Klein, W., & Sammis, C. Statistical physics approach to understanding the multiscale dynamics of earthquake fault systems. *Rev. Geophys.*, 41, 1019 (2003).

Rundle, J.B., Rundle, P.B., Donnellan, A. and Fox, G.C. Gutenberg-Richter statistics in topologically realistic system-level earthquake stress-evolution simulations. *Earth Planets Space*, 56, 761-771 (2004).

Rundle, P.B., Rundle, J.B., Tiampo, K.F., Donnellan, A. and Turcotte, D.L. Virtual California: Fault model, frictional parameters, applications. *PAGEOPH*, 163, 1819-1846 (2006).

Rundle, John B., J.R. Holliday, W. R. Graves, D. L. Turcotte, K. F. Tiampo, and W. Klein, Probabilities for large events in driven threshold systems, *Phys. Rev. E*, 86.2, 021106 (2012).

Rundle, J.B., Holliday, J.R., Graves, W. R., Rundle, P.B., Rundle, D.E., Hannon, A., Graves, C., Jeremic, B., Kunnath, S., Feltstykket, R., Mayeda, K., Turcotte, D.L. and Donnellan, A. A Practioner's Guide to Operational Real Time Earthquake Forecasting. In: Anderson, R. and Ferriz, H. *Applied Geology in California*, Star Publishing, California, 983-1003 (2016a).

Rundle, J.B., Donnellan, A., Grant Ludwig, L, Gong, G., Turcotte, D.L. and Luginbuhl, M. Nowcasting earthquakes. *Earth and Space Science*, 3, 480-486 (2016b).

Rundle, J.B., Luginbuhl, M., Giguere, A., and Turcotte, D.L. Natural time, nowcasting and the physics of earthquakes: Estimation of risk to global megacities. *Pure Appl. Geophys.*, 175, 647-660 (2018).

Rundle, J.B., Giguere, A., Turcotte, D.L., Crutchfield, J.P. and Donnellan, A. Global seismic nowcasting with Shannon information entropy. *Earth and Space Science*, 6, 456-472 (2019).

Rundle, J.B and Donnellan, A. Nowcasting earthquakes in southern California with machine learning: Bursts, swarms and aftershocks may be related to levels of regional tectonic stress. *Earth and Space Science* (in press) (2020).

Ryder, I. and Burgmann, R. Spatial variations in slip deficit on the central San Andreas fault from InSAR. *Geophys. J. Int*, 175, 837-852 (2008).

Rymer, M.J., Treiman, J.A., Kendrick, K.J., Lienkaemper, J.J., Weldon, R.J., Bilham, R., Wei, M., Fielding, E.J., Hernandez, J.L., Olson, B.P. and Irvine, P.J. Triggered surface slips in southern California associated with the 2010 El Mayor-Cucapah, Baja California, Mexico, earthquake. *US Geological Survey Open-File Report*, *1333* (2010).

Sachs, M. K., Yikilmaz, M.B., Heien, E.M., Rundle, J.B., Turcotte, D.L. and Kellogg, L.H. Virtual California earthquake simulator. *Seismol. Res. Lett*., 83, 973–978 (2012).

Sagiya, T., Miyazaki, S.I. and Tada, T. Continuous GPS array and present-day crustal deformation of Japan. *Pure and Applied Geophysics*. *157*(11-12), 2303-2322 (2000).

Salditch, L., Stein, S., Neely, J., Spencer, B.D., Brooks, E.M., Agnon, A. and Liu, M. Earthquake supercycles and long-term fault memory. *Tectonophys.*, 774 (2020).

Sarlis, N.V., Skordas, E.S. and Varotsos, P.A. A remarkable change of the entropy of seismicity in natural time under time reversal before the super-giant M9 Tohoku earthquake on 11 March 2011. *EPL*, 124 (2018).

Savage, J. C. The Parkfield prediction fallacy. *Bull. Seism. Soc. Am., 83*, 1-6 (1993).

Scholz, C.H. The Mechanics of Earthquakes and Faulting, Cambridge University Press, Cambridge (2019).

Schorlemmer, D. and Gerstenberger, M. C. RELM Testing Center. *Seismol. Res. Lett.*, 78(1), 30-36 (2007).

Schorlemmer, D., Zechar, J.D., Werner, M.J., Field, E.H., Jackson, D.D. and Jordan, T.H. First Results of the Regional Earthquake Likelihood Models Experiment. *Pure Appl. Geophys.* 167, 859-876 (2010).

Serino, C., Tiampo, K.F. and Klein, W. A New Approach to Gutenburg-Richter Scaling. *Phys. Rev. Lett.*, 106 (2011).

Shannon, C.E. A Mathematical Theory of Communication, *Bell System Technical Journal*, 27, 379–423 & 623–656 (1948).

Shekhawat, A., Zapperi, S. and Sethna, J.P. From Damage Percolation to Crack Nucleation Through Finite Size Criticality. *Phys. Rev. Lett.*, 110 (2013).

Shimazaki, K., and Nakata, T. Time-predictable recurrence model for large earthquakes. *Geophys. Res. Lett.*, *7*, 279–282 (1980).

Simons, M., Minson, S., Sladen, A., Ortega Culaciati, F., Jiang, J., Owen, S., Meng, L., Ampuero, J.P., Wei, S., Chu, R., Helmberger, D., Kanamori, H., Hetland, E., Moore, A. and Webb, F. The 2011 Magnitude 9.0 Tohoku-Oki Earthquake: Mosaicking the Megathrust from Seconds to Centuries. *Science*, 332, 1421-1425 (2011).

Smith, S.W. and Sammis, C.G. Revisiting the Tidal Activation of Seismicity With A Damage Mechanics and Friction Point of View. In: Donnellan, A., Mora, P., Matsu'ura, M. and Yin, X. Computational Earthquake Science Part II. *PAGEOPH Topical Volumes*, Birkhäuser, Basel (2004).

Song, Y. T., Fukumori, I., Shum, C.K. and Yi, Y. Merging tsunamis of the 2011 Tohoku-Oki earthquake detected over the open ocean. *Geophys. Res. Lett.*, 39(5) (2012).

Sornette, D. and Virieux, J. Linking short-timescale deformation to long-timescale tectonics. *Nature*, 357**,** 401–404 (1992).

Sornette, D. and Sammis, C. Complex critical exponents from Renormalization Group theory of earthquakes: Implications for earthquake prediction. *J. Phys. I*, 5(5), 607-619 (1995).

Stanley, R.P. Theory and Application of Plane Partitions: Part 1 and Part 2. *Studies in Appl. Math.*, 50(2-3) (1971).

Stark, P.B., and Freedman, D. What is the chance of an earthquake? In: Mulargia, F. and Geller, R.J. *Earthquake Science and Seismic Risk Reduction*, Kluwer, Dordrecht, Netherlands (2003).

Stauffer, D and Aharony, A. *Introduction to Percolation Theory*. Taylor and Francis, London (1994).

Stein, S. *Disaster Deferred: how new science is changing our view of earthquake hazards in the Midwest*. Columbia University Press, New York (2010).

Stein, S. and Okal, E. Speed and size of the Sumatra earthquake. *Nature*, 434, 581–582 (2005).

Stein, S. and Wysession, M. An introduction to seismology, earthquakes, and earth structure, John Wiley & Sons, New Jersey (2009).

Stein, S. and Stein, J.L., 2013. Shallow versus deep uncertainties in natural hazard assessments. *EOS*, 94(14), 133-140 (2013).

Stein, S., Salditch, L., Brooks, E., Spencer, B. and Campbell, M. Is the coast toast? Exploring Cascadia earthquake probabilities. *GSA Today*, 27(11), 6-7 (2017).

Stone, J.V. Information Theory, A Tutorial Introduction. Sebtel Press (2015).

Subarya, C., Chlieh, M., Prawirodirdjo, L. Avouac, J.-P., Bock, Y., Sieh, K., Meltzner, A.J., Natawidjaja, D.H. and McCaffrey, R. Plate-boundary deformation associated with the great Sumatra–Andaman earthquake. *Nature*, 440**,** 46–51 (2006).

Taleb, N.N. *Fooled by Randomness: The Hidden Role of Chance in Life and in the Markets*, Random House, New York (2004).

Taroni, M., Marzocchi, W., Schorlemmer, D., Werner, M.J., Wiemer, S., Zechar, J.D., Heiniger, L. and Euchner, F. Prospective CSEP Evaluation of 1-Day, 3-Month, and 5-Yr Earthquake Forecasts for Italy. *Seism. Res. Lett.*, 89(4), 1251-1261 (2018).

Tiampo, K.F., Rundle, J.B., Gross, S.J., McGinnis, S. and Klein, W. Eigenpatterns in southern California seismicity. *J. Geophys. Res*., 107(B12), 2354 (2002a).

Tiampo, K.F., Rundle, J.B., McGinnis, S.A. and Klein, W. Pattern dynamics and forecast methods in seismically active regions. *Pure Appl. Geophys.*, 159, 2429-2467 (2002b).

K.F. Tiampo, J.B. Rundle, W. Klein, Seismicity rate and stress change – stress shadows determined using the Pattern Informatics technique, pp. 271-276, *Proc. 4th ACES Workshop*, ed. by XC Yin, P. Mora, and Y Zhang, Beijing, China (2005). ISBN 0-9750394-1-5

Titov, V., Song, Y.T., Tang, L., Bernard, E.N., Bar-Sever, Y. and Wei, Y. Consistent estimates of tsunami energy show promise for improved early warning, *Pure Appl. Geophys*., 173, 3863–3880 (2016).

Tong, X., Sandwell, D.T. and Fialko, Y. Coseismic slip model of the 2008 Wenchuan earthquake derived from joint inversion of interferometric synthetic aperture radar, GPS, and field data. *J. Geophys. Res.*, 115 (2010).

Trugman, D. T. Deviant Earthquakes: Data-driven Constraints on the Variability in Earthquake Source Properties and Seismic Hazard (Doctoral dissertation, University of California San Diego, San Diego, CA) (2017).

Tsukahara, K. Recent Progress in the Study of Crustal Movement and Geoid Determination by the Dense GPS Network over the Japanese Islands. *Gravity, Geoid and Marine Geodesy*. Springer, Berlin, Heidelberg, 1-8 (1997).

Tsuji, H., Hatanaka, Y., Sagiya, T. and Hashimoto, M. Coseismic crustal deformation from the 1994 Hokkaido-Toho-Oki Earthquake Monitored by a nationwide continuous GPS array in Japan. *Geophys. Res. Lett.* 22(13), 1669-1672 (1995).

Tullis, T.E., Richards-Dinger, K., Barall, M., Dieterich, J.H., Field, E.H., Heien, E.M., Kellogg, L.H., Pollitz, F.F., Rundle, J.B., Sachs, M.K., Turcotte, D.L., Ward, S.N. and Yikilmaz, M.B. Generic Earthquake Simulator. *Seismol. Res. Lett.*, 83, 959-963 (2012).

Turcotte D.L. Earthquake Prediction. *Ann. Rev. Earth Planet. Sci*., 19, 263-281 (1991).

Turcotte, D.L., Newman, W. and Shcherbakov, R. Micro and macroscopic models of rock fracture. *Geophys. J. Int.*, 152, 718-728 (2003).

Turcotte, D.L. and Shcherbakov, R. Can damage mechanics explain temporal scaling laws in brittle fracture and seismicity? *Pure Appl. Geophys*., 163, 1031-1045 (2006).

Unger, C. and Klein, W. Nucleation theory near the classical spinodal. *Phys. Rev B*, 29, 2698-2708 (1984).

USGS ShakeAlert—An Earthquake Early Warning System for the United States West Coast, Fact Sheet 2014–3083, version 1.2 February 2017

Van Aalsburg, J., Grant, L.B., Yakovlev, G., Rundle, P.B., Rundle, J.B., Turcotte, D.L. and Donnellan, A. A feasibility study of data assimilation in numerical simulations of earthquake fault systems. *Phys. Earth. Planet. Int.*, 163, 149-162 (2007).

Varnes, D.J. (1989), Predicting Earthquakes by Analyzing Accelerating Precursory Seismic Activity. *Pure Appl. Geophys*., 4, 661–686 (1989).

Varotsos, P., Sarlis, N.V. and Skordas, E.S. Spatiotemporal complexity aspects on the interrelation between Seismic Electric Signals and seismicity, *Practica of Athens Academy*, 76, 294-321 (2001).

Varotsos, P., Sarlis, N.V. and Skordas, E.S. Long-range correlations in the electric signals that precede rupture. *Phys. Rev. E*, 66 (2002).

Varotsos, P., Sarlis, N.V. and Skordas, E.S. *Natural Time Analysis: The new view of time. Precursory Seismic Electric Signals, Earthquakes and other Complex Time-Series.* Springer-Verlag, Berlin Heidelberg (2011).

Varotsos, P., Sarlis, N.V., Skordas, E.S. and Lazaridou, M.S. Seismic Electric Signals: An additional fact showing their physical interconnection with seismicity. *Tectonophys.*, 589, 116-125 (2013).

Varotsos, P., Sarlis, N.V. and Skordas, E.S. Study of the temporal correlations in the magnitude time series before major earthquakes in Japan. *J. Geophys. Res. Space Phys.*, 119, 9192-9206 (2014).

Varotsos, P.A., Skordas, E.S. and Sarlis, N.V. Fluctuations of the entropy change under time reversal: Further investigations on identifying the occurrence time of an impending major earthquake. *EPL*, 130 (2020).

Vere-Jones, D. A note on the statistical interpretation of Bath's law. *Bull. Seism. Soc. Am.*, 59, 1535-1541 (1969).

Wald, D.J. (2020) Practical limitations of earthquake early warning. Earthquake Spectra 36(3) 1412–1447

Ward, S. N. ALLCAL earthquake simulator. *Seismol. Res. Lett*., 83, 964–972 (2012).

Walter, Jacob I., et al. "Persistent tremor within the northern Costa Rica seismogenic zone." Geophysical Research Letters 38.1 (2011).

Wei, M., Sandwell, D., Fialko, Y. and Bilham, R. Slip on faults in the Imperial Valley triggered by the 4 April 2010 Mw 7.2 El Mayor-Cucapah earthquake revealed by InSAR. *Geophys. Res. Lett.*, *38*(1) (2011).

Wei, S.J., Avouac, J.-P., Hudnut, K.W., Donnellan, A., Parker, J.W., Graves, R.W., Helmberger, D., Fielding, E., Liu, Z., Cappa, F. and Eneva, M. The 2012 Brawley swarm triggered by injection-induced aseismic slip. *Earth Planet. Sci. Lett*., 422, 115-125 (2015).

Weibull, W. Statistical design of fatigue experiments. *J. Appl. Mech*. 19, 109-113 (1952).

Weldon, R., Scharer, K., Fumal, T. and Biasi, G. Wrightwood and the earthquake cycle: what a long recurrence record tells us about how faults work. *GSA Today* 14(9), 4–10 (2004).

Wesnousky, S.G., Scholz, C.H., Shimazaki, K. and Matsuda, T. Integrations of geological seismological data for the analysis of seismic hazard: A case study of Japan. *Seis. Soc. Am. Bull.* 74**,** 687-705 (1984).

Wettstein, S.J., Wittel, F.K., Araújo, N. A.M., Lanyon, B. and Herrmann, H.J. From invasion percolation to flow in rock fracture networks. *Physica A*, 391(1-2), 264-277 (2012).

Wiemer, S. and M. Wyss (2000) Minimum magnitude of completeness of earthquake catalogs: Examples from Alaska, the Western United States and Japan, Bull. Seism. Soc. Amer., 90, 859-869.

Wilkinson, D. and Willemsen, J.F. Invasion percolation, a new form of percolation theory. *J. Phys. A: Math. Gen.,* 16, 3365-3376 (1983).

Wilkinson, D. and Barsony, M. Monte Carlo study of invasion percolation clusters in two and three dimensions. *J. Phys. A.: Math Gen*., 17, L129-L135 (1984).

Wilson, Kenneth G. "The renormalization group and critical phenomena." Reviews of Modern Physics 55.3 (1983): 583.

Wisely, B.A. and Schmidt, D. Deciphering vertical deformation and poroelastic parameters in a tectonically active fault-bound aquifer using InSAR and well level data, San Bernardino basin, California. *Geophys. J. Int.*, 181, 1185-1200 (2010).

Xu, Z. and Song, Y.T. Combining the all-source Green's functions and the GPS-derived source for fast tsunami prediction – illustrated by the March 2011 Japan tsunami. *J. Atmos. Oceanic Tech.*, 30(7), 1542-1554 (2013).

Xue, L., Schwartz, S., Liu, Z. and Feng, L.J. Interseismic megathrust coupling beneath the Nicoya Peninsula, Costa Rica, from the joint inversion of InSAR and GPS data. *J. Geophys. Res.*, 120, 3703-3722 (2015).

Yikilmaz, M.B., Turcotte, D.L., Yakovlev, G., Rundle, J.B. and Kellogg, L.H. Virtual California earthquake simulations: Simple models and their application to an observed sequence of earthquakes. *Geophys. J. Int.*, 180, 734-742 (2010).

Yin, X., Chen, X., Song, Z. and Yin, C. A new approach to earthquake prediction: The Load/Unload Response Ratio (LURR) theory. *PAGEOPH*, 145**,** 701–715 (1995).

Yin, X.-C., Yu, H.-Z., Kukshenko, V., Xu, Z.-Y., Wu, Z., Li, M., Peng, K., Elizarov, S. and Li, Q. Load-Unload Response Ratio (LURR), Accelerating Moment/Energy Release (AM/ER) and state vector saltation as precursors to failure of rock specimens. *PAGEOPH*, 161, 2405–2416 (2004).

Yoder, M. R., Rundle, J. B. and Glasscoe, M. T. Near-fieldetas constraints and applications to seismic hazard assessment. *Pure Appl. Geophys*.,172(8), 2277–2293 (2015).

Yuan, S., Yin, X.-C. and Liang, N. Load-unload response ratio and its application to estimate future seismicity of Qiandao Lake region. *Procedia Eng.,* 4, 333-339 (2010).

Zaliapin, I. and Ben-Zion, Y. A global classification and characterization of earthquake clusters. *Geophys. J. Int.*, 207(1), 608-634 (2016a).

Zaliapin, I. and Ben-Zion, Y. Discriminating characteristics of Tectonic and Human-Induced Seismicity. *Bull. Seism. Soc. Am.*, 106(3), 846-859 (2016b).

Zechar, J.D., Schorlemmer, D., Werner, M.J., Gerstenberger, M.C., Rhoades, D.A. and Jordan, T.H. Regional Earthquake Likelihood Models I: First-Order Results. *Bull. Seism. Soc. Am.*, 103(2A), 787-798 (2013).

Zhang, Y., Wu, Y., Yin, X.C., Peng, K., Zhang, L., Yu, A., and Zhang, X., Comparison between LURR and state vector analysis before strong earthquakes in southern California since 1980. In: Tiampo, K.F., Weatherly, D.K. and Weinstein, S.A. *Earthquakes: Simulations, Sources and Tsunamis*, *PAGEOPH*, Birkhäuser Basel, 737-748 (2008).