Foreshock sequences and short-term earthquake predictability on East Pacific Rise transform faults

Jeffrey J. McGuire¹, Margaret S. Boettcher² & Thomas H. Jordan³

¹Department of Geology and Geophysics, Woods Hole Oceanographic Institution, and ²MIT-Woods Hole Oceanographic Institution Joint Program, Woods Hole, Massachusetts 02543-1541, USA
²Department of Earth Sciences, University of Southern California, Los Angeles, California 90089-7042, USA

East Pacific Rise transform faults are characterized by high slip rates (more than ten centimetres a year), predominately aseismic slip and maximum earthquake magnitudes of about 6.5. Using recordings from a hydroacoustic array deployed by the National Oceanic and Atmospheric Administration, we show here that East Pacific Rise transform faults also have a low number of aftershocks and high foreshock rates compared to continental strike-slip faults. The high ratio of foreshocks to aftershocks implies that such transform-fault seismicity cannot be explained by seismic triggering models in which there is no fundamental distinction between foreshocks, mainshocks and aftershocks. The foreshock sequences on East Pacific Rise transform faults can be used to predict (retrospectively) earthquakes of magnitude 5.4 or greater, in narrow spatial and temporal windows and with a high probability gain. The predictability of such transform earthquakes is consistent with a model in which slow slip transients trigger earthquakes, enrich their low-frequency radiation and accommodate much of the aseismic plate motion.

On average, before large earthquakes occur, local seismicity rates show a significant increase¹. In continental regions, where dense regional seismic networks provide the best data, most foreshock studies²⁴—⁵, though not all⁷, are consistent with the hypotheses that earthquake nucleation is independent of magnitude and that foreshocks result from a general triggering process in which there is no fundamental distinction between foreshocks, mainshocks and aftershocks. The inability to distinguish foreshocks from the statistical fluctuations in the continental background seismicity severely limits their usefulness in predicting large earthquakes⁶.

It is unclear, however, whether these statements apply to other tectonic environments, or how aseismic processes affect earthquake triggering. Aseismic slip transients with timescales of days to months have recently been observed in the subduction zones of Japan⁷–⁹ and Cascadia¹⁰, using continuously monitored GPS arrays. The possibility that aseismic slip triggers large earthquakes on subduction megathrusts is especially intriguing given the observation¹¹ that a slow slip transient occurred 15 minutes before the great 1960 Chilean megathrust earthquake, which had a moment magnitude \( M_w \) of 9.5, the largest ever recorded. Notably, subduction zones are observed to have higher foreshock rates than continental regions¹².

Another tectonic environment in which aseismic processes are thought to exert a strong influence on fault behaviour is mid-ocean ridge transform faults (RTFs). Studies over the last several decades¹³–¹⁵ have shown that on average most of the slip on RTFs, up to about 85% (ref. 15), is aseismic. Moreover, the seismic component of slip occurs in earthquakes that are relatively small \((M_w \leq 7.2)\) given the length of the faults¹⁶–¹⁸. Many of the larger RTF earthquakes are slow events with anomalous low-frequency radiation¹⁹,²⁰. Low-frequency spectral analyses¹⁹,²⁰ have indicated that slow RTF earthquakes are compound events comprising an ordinary rupture and a slow transient of comparable moment but much longer duration; in some cases, the slow component precedes, and presumably initiates, the main seismic component. Time-domain records of slow precursors to RTF earthquakes²¹,²² and episodes of coupled seismic slip observed on adjacent RTFs²⁰,²² support the inference of slow slip transients, but the subject remains controversial²³.

Hydroacoustic detection of foreshocks

Here we use data from a six-element hydroacoustic array deployed by the National Oceanic and Atmospheric Administration’s Pacific Marine Environmental Laboratory (NOAA-PMEL) to examine anomalous foreshock sequences on East Pacific Rise (EPR) transform faults (Fig. 1). The NOAA-PMEL arrays²⁴–²⁶ routinely locate EPR earthquakes with acoustic source level (ASL) magnitudes (see Methods) \( M_{\text{ASL}} \) below 3, reducing the detection threshold by 1.5–2.0 magnitude units below global seismicity catalogues (see Methods). A reconnaissance study indicated that foreshocks in the last hour before large events are significantly more common on EPR transform faults than on strike-slip faults in the North Atlantic, Northeast Pacific or Southern California²⁷.

Figure 2 displays stacks of the seismicity in space-time windows centred on nine mainshocks that occurred on the Discovery and Gofar transform faults between May 1996 and December 2001. This set of mainshocks comprised all \( M_w \geq 5.4 \) earthquakes on these two faults recorded by the NOAA-PMEL array in the Harvard Centroid Moment Tensor (CMT) catalogue²⁸ that did not follow within 1 week and 100 km of another mainshock. The longer window (Fig. 2a) shows low background seismicity tens of hours before the mainshocks and the subsequent aftershock decay. The shorter window (Fig. 2b) reveals an accelerating rate of seismicity close to the mainshock epicentres during the hour immediately preceding the mainshock origin times.

Earthquake triggering model

The anomalous nature of the RTF foreshock activity can be quantified in terms of the Epidemic Type Aftershock Sequence (ETAS) model of triggered seismicity²⁹,³⁰. ETAS is a marked point process model²⁹ in which all earthquake magnitudes above a lower cutoff \( m_s \) are independent samples of the Gutenberg–Richter (GR) probability distribution, \( P(m) = 10^{-(m-m_b)} \), where \( b \) is the slope of the distribution, and all earthquakes give birth to daughter events at
an average rate of $\phi(m, t) = \rho(m)\psi(t)$, where $t$ is time since the earthquake. This triggering rate is assumed to increase exponentially with magnitude, $\rho(m) = k \times 10^{\alpha(m-m_0)}$, where $\alpha$ is the triggering exponent, and to decay with time after a mother event according to the modified Omori law, $\psi(t) = \theta \frac{\beta}{(c + t)^{1+\theta}}$ (where $\theta > 0$). The constants $k$, $\theta$ and $c$ are parameters that vary among regions. Renormalization for a single mainshock of magnitude $m$ yields an average seismicity rate proportional to $\frac{N^0}{10^{\alpha m}}$, where $N^0$ is the expected number of aftershocks of any magnitude (type-II aftershocks); that is, not constrained to be smaller than $m_0$, and the constant $n = \int_0^\infty \rho(\mu)\,d\mu$ is the branching ratio, which equals both the average number of directly triggered aftershocks per event and the fraction of the earthquake population that is made up of triggered earthquakes. The aftershock rate decays with an effective Omori exponent $p = 1 + O(\theta)$.

In the ETAS model, the seismicity rate before a mainshock at $t = 0$ increases according to the inverse Omori law; that is, $\sim |t|^{-p'}$, where $p' = 1 + O(\theta)$, and the expected number of events of all magnitudes conditioned on the mainshock occurrence (type-II foreshocks) is independent of $m$. This conditional foreshock number can be approximated as the product of two factors: the probability that the mainshock is a triggered event, and the expected number of events in a cluster averaged over mainshock magnitude. The first is just the branching ratio $n$, and the second is the integral $\int_0^\infty N^0(\mu)\,d\mu$; therefore $N^0 \approx n^2/(1 - n)$. To include only earthquakes smaller than the mainshock (type-I foreshocks), we multiply the integrand by the probability that no event in a cluster exceeds $m$ and integrate over the appropriate magnitude range. If $k/(1 - n)$ is small and $m$ is large (conditions which apply to our data), then the extra probability factor is close to unity, and the results are $N^0 \approx N^0[1 - 10^{-(b-a)(m-m_0)}]$.

A similar modification to the aftershock number yields the foreshock/aftershock ratio:

$$\frac{N_f}{N_a} \approx n \frac{b}{b - \alpha} \left[ \frac{10^{(b-a)\Delta m_1} - 10^{(b-a)\Delta m_2}}{10^{\Delta m_1} - 10^{\Delta m_2}} \right].$$

Here we have generalized the formula to count foreshocks in the magnitude range from $m - \Delta m_1$ to $m - \Delta m_2$ and aftershocks from $m - \Delta m_1$ to $m - \Delta m_2$, where $0 \leq \Delta m_1 < \Delta m_2 \leq m - m_0$. This approximation, which applies to large mainshocks, differs conceptually from the expression recently used by ref. 3 to explain the foreshock/aftershock ratios from global and regional catalogues (see Methods).

**Anomalous foreshock activity**

Earthquake populations on RTFs are well described by a tapered GR distribution having a low-magnitude slope $b \simeq 1$ (refs 14, 15) similar to that of continental regions. The hydroacoustic catalogue for the EPR faults is consistent with this self-similar scaling, and its foreshock sequences decay according to Omori’s law with $p \simeq 1$, again similar to continental regions. However, global catalogues demonstrate that the aftershock productivity of large RTF earthquakes is lower than continental faults by approximately a factor of fifteen (Fig. 3). The low aftershock productivity combined with...
the poor detection thresholds of global catalogues makes it difficult to constrain the values of $n$ and $\alpha$ independently. A maximum-likelihood fit to the teleseismic RTF data yields a best-fit value of the triggering exponent $\alpha = 0.72$, and is consistent with the somewhat higher values found for California and Japan ($\alpha = 0.8-1.0$)\textsuperscript{13,34} (Fig. 3). Error bounds on the maximum-likelihood estimate are large, but aftershock counts using the hydroacoustic catalogue (points with horizontal bars in Fig. 3) also favour relatively high values of $\alpha$ and rule out values less than about 0.6 (see Supplementary Information).

The difference between oceanic and continental aftershocks primarily manifests itself in the intercept of the scaling relation, $k(1-n)$, which is offset by about a factor of fifteen (Fig. 3). The maximum-likelihood fit in Fig. 3 corresponds to a branching ratio, $n\approx 0.1$, compared to values approaching unity in continental seismic zones\textsuperscript{35}. As discussed in the Supplementary Information, the aftershock rate may be somewhat higher for the EPR faults, but we can say with a high degree of confidence that $n<0.3$. In other words, according to the ETAS model, most RTF earthquakes ($70\%$–$90\%$) would be primary events driven by as seismic plate-tectonic loading rather than aftershocks of previous earthquakes. Equation (1) with the maximum-likelihood estimate of $\alpha$ and $n$ predicts that the foreshock/aftershock ratio for RTFs should be about an order of magnitude lower than that observed in continents.

Instead, the EPR transform faults that are well recorded by the NOAA-PMEL array give values of $N_f/N_a$ that are an order of magnitude higher than observed in Southern California (Fig. 4). For both regions, we identified mainshocks as events in the Harvard CMT catalogue with $M_w\geq 5.4$ that did not follow within 1 week and 100 km of another mainshock (see Methods), and we compiled foreshock and aftershock statistics from the NOAA-PMEL and the Southern California Earthquake Data Center (SCEDC) catalogues. We counted all events with local magnitudes ($M_{ASL}$ or $M_L$) up to 2.8 units smaller than the mainshock $M_w$ in the 1-h interval before and the 5-h interval after the mainshock. Figure 4 compares the observed $N_f/N_a$ for spatial windows of various radii with the predictions of equation (1), corrected for the finite sampling intervals (see Methods). The SCEDC statistics satisfy an ETAS model with $\alpha = 0.8$–0.9, consistent with previous catalogue studies\textsuperscript{36,37}. However, foreshock rates from the NOAA-PMEL statistics are about two orders of magnitude greater than the ETAS predictions using the maximum-likelihood fit in Fig. 3. As shown in Fig. 4 and Supplementary Fig. S3, these results are robust with respect to the choice of windows and declustering procedures.

Therefore, we can reject the ETAS hypothesis that the clustering of foreshocks, mainshocks and aftershocks on RTFs can be described by the same seismic triggering mechanism. We infer that large earthquakes on EPR faults are preceded by an extended preparation process, possibly driven by subseismic transients (silent or quiet earthquakes), that can often be observed through foreshocks. This alternative hypothesis is consistent with the tightly localized distribution of the foreshocks about the mainshock in both space and time (Fig. 2 and Supplementary Fig. S4), which does not conform to the inverse-diffusive behaviour expected from the ETAS model\textsuperscript{4}.

The correspondence of slow slip with foreshocks was suggested as early as 1976 by Kanamori and Stewart\textsuperscript{40}, who noted a foreshock with a body-wave magnitude $m_b \approx 5$ about 500 s before the $M_w = 7$ slow earthquake on the Gibbs transform fault in the North Atlantic. More recently, McGuire et al. associated $m_b = 4.5$–5.0 foreshocks before the 1994 $M_w = 7.0$ Romanche\textsuperscript{20} and 1997 $M_w = 6.8$ Prince Edward Island\textsuperscript{21} earthquakes with slow precursors observed at low frequencies. Forsyth et al.\textsuperscript{22} suggested...
that a subseismic slip process was responsible for a swarm of contemporaneous seismicity on the Anakena and Raraku transform faults of the southern EPR recorded by an ocean-bottom seismometer array in 1995. These and other examples\textsuperscript{m} combined with the global aftershock depletion (Fig. 3) and the evidence for slow precursors to large earthquakes on RTFs worldwide\textsuperscript{m,19}, indicate that the aseismic, foreshock-generating process on EPR faults may be prevalent throughout the mid-ocean-ridge system, including the slower-slipping, colder RTFs in the Atlantic and Indian oceans.

**Short-term predictability of large earthquakes**

The high rate of proximate foreshocks suggests a naive scheme for short-term earthquake prediction—we simply assume that every event is a foreshock of an impending large earthquake. We can formalize this scheme into a well-posed prediction algorithm\textsuperscript{m}: whenever we observe any RTF event above some ASL magnitude threshold \( m_{0} \), within a specified RTF region, we issue an alert that an earthquake of moment magnitude greater than or equal to \( m_{r} \) will occur sometime during time window of length \( t_{P} \) immediately following the event and somewhere in a spatial window of radius \( r_{P} \) about the event's epicentre. Figure 5 illustrates this prediction algorithm for the parameter set \( m_{0} = 2.5 \) (ASL\textsuperscript{2}), \( m_{r} = 5.4 \) (ASL\textsuperscript{2}), \( t_{P} = 1 \) h, \( r_{P} = 15 \) km by applying it retrospectively to the two most active EPR transform faults, Discovery and Gofar. Of the nine candidate earthquakes that occurred during the catalogue interval May 1996–November 2001, six were located within the space-time prediction windows (Fig. 5) and thus constitute successful predictions. There were three failures-to-predict and about 1,400 false alarms.

Although the false-alarm rate is quite high, all alarms taken together occupy only about 0.15% of the total space-time volume of about 250 km \( \times \) 5.5 yr (see Supplementary Information). We can relate \( P[M|F] \), the probability of a mainshock \( M \) in the prediction window given the occurrence of a foreshock \( F \), to \( P(M) \), the probability of \( M \) in a random window of the same size, using the Bayes identity:

\[
P(M|F) = \frac{P(F|M) P(M)}{P(F)}
\]

\( P(M|F) \) is the fraction of mainshocks preceded by foreshocks, and \( P(F) \) equals the fraction of the space-time volume occupied by alerts. The ratio of these probabilities (the term in square brackets) is the probability gain factor \( g \) of the prediction algorithm\textsuperscript{m}. Our retrospective analysis of the Discovery and Gofar faults gives \( g = 450 \). This performance can be compared to prediction experiments in California, where even highly optimized algorithms with many adjustable parameters are thought to achieve probability gains of 10–20 or less\textsuperscript{m}.

The performance relative to random chance can be evaluated using Molchan's\textsuperscript{39} error diagram, which plots the failure-to-predict probability, \( 1 - P[M|F] \), against \( P(F) \) (Fig. 6). Large sets of random alarms should fall close to the line \( 1 - P(F) \), corresponding to no probability gain (\( g = 1 \)). Figure 6 shows that the Discovery–Gofar point lies well below the 99% confidence range for random chance; the probability of reproducing this performance with random alarms filling 0.15% of the space-time volume is less than one in ten million. Similar results were obtained by applying the algorithm retrospectively to the five active EPR transforms in the study area; nine of sixteen mainshocks were successfully predicted by alerts occupying 0.13% of the space-time volume, which gives \( g = 540 \) (Fig. 6, see Supplementary Discussion).

Our naive algorithm is far from optimal. For instance, raising the threshold magnitude \( m_{r} \) from 2.5 to 3.4 (that is, increasing the ASL from 207 to 215) reduces the number of alarms in the Gofar–Discovery catalogue to 407 without changing \( P(M|F) \), increasing \( g \) to about 1,000 (Fig. 6). More parameters could be added to improve the performance further.

**Discussion**

Mid-ocean ridges are far removed from urban centres on continents, so the direct societal value of short-term earthquake prediction on RTFs (assuming it could be operationally implemented using real-time, near-source monitoring) would be small. Nevertheless, the existence of short-term predictability in this tectonic environment—the main conclusion of this paper—is of considerable scientific interest, because it supports a physical linkage between foreshocks and mainshocks through stress changes driven by aseismic slip transients or some other fault preparation process
such as hydrothermal flow or nearby magmatic activity. To monitor RTFs geologically will require a new generation of ocean-bottom instrumentation, a considerable technological challenge. But collecting such data may well be the best way to test the hypothesis that three anomalous aspects of RTF seismicity—large slip deficits, high foreshock activity and slow earthquakes—can be explained by aseismic fault-slip transients. Given the importance of understanding the fundamental mechanics of earthquake predictability, overcoming the technological hurdles should be worth the effort.

Methods

We use the NOAA-PMEL hydroacoustic earthquake catalogue for the equatorial Pacific which began in 1996 (ref. 24). Their array (Fig. 1) records acoustic energy radiated into the water column by earthquakes and other sources (T-phases). T-phases propagate very efficiently in the low-velocity sound fixing and ranging (SOFAR) channel, and the array arrivals can be used to locate precisely where the energy entered the SOFAR channel. The standard errors in this source location are estimated to be 1-2 km, ±10 s for the southern/ northernmost faults (the Wilkes and Sirequitos faults) and are slightly smaller for the faults located in the centre of the array (the Discovery and Gofar faults). A propagation model is used to convert the magnitude of the pressure signal at the hydrophones into an ASL (measured in decibels) at the source location. We used a conversion between ASL and magnitude, Mw = 2.90 ± 0.16, found by Smith et al. (1997). The factor W is the scaling of theory breaks down for time value of (parameters c aftershock ratio of 0.134. The factor

Equation (1) is a more general form of models used in previous studies (12,40,41). Feltzer et al. (12) have demonstrated that the time-decay exponent p in the renormalized Omori law is not strictly a constant. When p<1/2, it varies from 1-θ for c < t<= c(t-1) → ∞ and t>> θ as t→∞. The theory breaks down for θ>>1, owing to the strong coupling between earthquake energy and seismicity rate, but the numerical simulations of ref. 4 indicate that the short-term value of θ increases approximately linearly from 1-θ at θ = 1/2 to 1 as p→∞. Hence, the term θ p -1 can be approximated by p = p -1 if t>1;

where the approximation does not hold. Equation (1) can thus be corrected for the finite foreshock and aftershock sampling intervals, d = 1 s and d = 1 h, by multiplying its right-hand side by the number of these intervals. This ratio varies from 0.97 to 0.94 over the plausible range of parameters c : 1 s = 1 min, θ = 0.2, so the correction is minor. Equation (1) is a more general form of models used in previous studies (12,40,41). Feltzer et al. (12) consider the case c = 0. Taking this limit in equation (1), we obtain

For the magnitude ranges of ref. 3 (ΔMw = 1.0, ΔMw = 0; ΔMw = 1.0, ΔMw = 0.4), the ratio in brackets represents their foreshock/ aftershock ratio of 0.134. The factor n(t) missing in their equation (6) probably lies between 1 and 2 for continental and subduction-zone seismicity (29), so their formula provides an adequate approximation in most regions. For RTFs, however, the scaling of the magnitude N1/N2 with the branching ratio n is numerically important (see Fig. 4). Received 3 September 2004; accepted 19 December 2005; doi:10.1038/nature03377.

Acknowledgements

We thank R. Dzika for answering questions about details of the hydroacoustic earthquake catalogues, D. Bohnenstiehl for suggestions on clarifying the manuscript, A. Helmstetter for her help in understanding ETAS, and V. Keilis-Borok, I. Zaliapin, for discussions about earthquake prediction algorithms. J.J.McG. was supported by the Frank and Lisa Hoch Fund. M.S.B. was supported by the Deep Ocean Exploration Institute at WHOI. This work was supported by the NSF, SCOR, and USGS.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.J.McG. (jigm@whoi.edu).

NATURE 3377—9/3/2005—V.BICKNELL—137936


7. Hirose, H., Hiraoka, K., Kimata, F., Fujii, N. & Miyazaki, S. A slow thrust slip event following the two large slip deficits, high foreshock activity and slow earthquakes—can be explained by aseismic fault-slip transients. Given the importance of understanding the fundamental mechanics of earthquake predictability, overcoming the technological hurdles should be worth the effort.

Equation (1) is a more general form of models used in previous studies (12,40,41). Feltzer et al. (12) have demonstrated that the time-decay exponent p in the renormalized Omori law is not strictly a constant. When p<1/2, it varies from 1-θ for c < t<= c(t-1) → ∞ and t>> θ as t→∞. The theory breaks down for θ>>1, owing to the strong coupling between earthquake energy and seismicity rate, but the numerical simulations of ref. 4 indicate that the short-term value of θ increases approximately linearly from 1-θ at θ = 1/2 to 1 as p→∞. Hence, the term θ p -1 can be approximated by p = p -1 if t>1;