AGU PUBLICATIONS

Geophysical Research Letters

RESEARCH LETTER
10.1002/2017GL073504

Key Points:
- Global seismic moment release increased since December 2004
- Global seismic moment release during 1918–2014 cannot be fit under assumption of independence between earthquake times and seismic moments
- Rise of great earthquakes during the 1960s and after December 2004 might be caused by a higher corner moment

Supporting Information:
- Supporting Information S1

Correspondence to:
I. Zaliapin,
zal@unr.edu

Citation:
Zaliapin, I., and C. Kreemer (2017), Systematic fluctuations in the global seismic moment release, *Geophys. Res. Lett.*, *44*, 4820–4828, doi:10.1002/2017GL073504.

Received 20 MAR 2017
Accepted 1 MAY 2017
Accepted article online 8 MAY 2017
Published online 29 MAY 2017

Systematic fluctuations in the global seismic moment release

Ilya Zaliapin[1] and **Corné Kreemer**[2]

[1]Department of Mathematics and Statistics, University of Nevada, Reno, Nevada, USA, [2]Nevada Bureau of Mines and Geology, and Seismological Laboratory, University of Nevada, Reno, Nevada, USA

Abstract We revisit the significance of the increased number of great earthquakes since December 2004. An analysis of the global seismic moment release during 1918–2014 inferred from the International Seismological Centre catalog rejects a null hypothesis of independence between earthquake seismic moments and occurrence times. Our results suggest the existence of temporal variation in the seismic moment distribution on a global scale with decreased moment release during 1975–2004 and a transition to a regime with increased moment release during the 1960s and after December 2004. We use complementary likelihood and regression analyses based on nonparametric resampling and parametric Monte Carlo simulations to construct tests powerful enough to reject the null hypothesis of independence.

1. Introduction

The global rate of great earthquakes exhibits an increase since the Sumatra earthquake of 26 December 2004 [e.g., *Lay*, 2015]. The question of whether this increase is an expected statistical variation of a time-independent process or an evidence of genuine temporal fluctuations has received substantial attention [*Ammon et al.*, 2011; *Bufe and Perkins*, 2011; *Michael*, 2011; *Daub et al.*, 2012; *Beroza*, 2012; *Shearer and Stark*, 2012; *Parsons and Geist*, 2012, 2014; *Ben-Naim et al.*, 2013]. A conventional way of addressing this problem is to consider the number of events above a threshold and to test possible deviations from a stationary Poisson process. We argue that this approach might not be powerful enough. Instead, to reveal if the observed change is a significant deviation from a long-term expectation, it is important to (i) analyze moment release, not only the number of events; (ii) consider large enough range of examined seismic moments—say, all events with magnitude above $M_w = 7$; and (iii) consider sequential seismic moment release, not the statistical distribution of all moments collected during the examined time interval. Our analysis combines these aspects and indicates that at variance with the majority of studies referenced above, a time-independent model cannot describe the global moment release. In particular, the apparent increase of moment release during the last decade is a statistically significant feature of the global seismic process.

2. Data and Methods

2.1. Earthquake Catalog Declustering

We work with version 4 of the global International Seismological Centre-Global Earthquake Model (ISC-GEM) for the period of 1918–2014 [*Storchak et al.*, 2013, 2015]. We change the moment-magnitude of the 2004 Great Sumatra earthquake to $M_w = 9.3$ in order to be consistent with *Stein and Okal* [2005] and *Tsai et al.* [2005]. Since we ultimately would like to invoke plate tectonic considerations to explain significant temporal variations in seismic moment release, we only consider events with depth $z \leq 70$ km.

Aftershock sequences typically produce space–time concentrations of large-magnitude events, which can influence analysis of time-moment dependencies. To eliminate clustering related to aftershock sequences, and focus on potential global variations of seismic moment release, we decluster the examined catalog using the approach of *Zaliapin et al.* [2008] and *Zaliapin and Ben-Zion* [2013a] summarized in Text S1 in the supporting information.

2.2. Seismic Moment

We convert moment magnitudes M_w to scalar seismic moment M via the relation used for the ISC-GEM catalog:

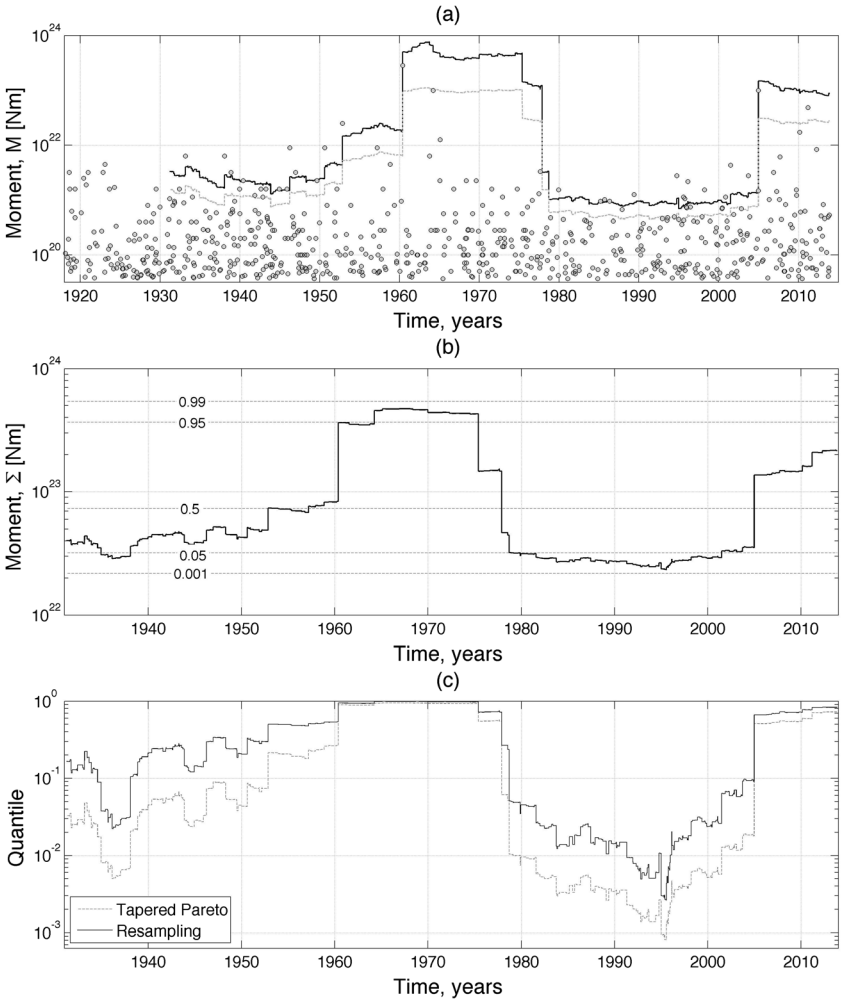

Figure 1. Time fluctuations of the global seismic moment release: an illustration. The analysis uses 692 declustered earthquakes with $M_w \geq 7$ ($M \geq 3.55 \times 10^{19}$ Nm) during 1918–2014 from version 4 of the ISC-GEM catalog [*Storchak et al.*, 2013]. (a) Examined events (gray circles) and the estimated corner moment M_c of tapered Pareto distribution (solid line) as a function of time. The estimation is done in a sliding window of 100 events. A lower 95% confidence boundary is shown by dashed line; the upper limit is unconstrained. (b) The observed moment release $\Sigma_{100}(t)$ in a sliding window of $m = 100$ events (solid line) versus empirical quantiles (dashed lines) obtained using 10^5 resamplings of the observed moments. The level of quantiles in indicated in the figure. (c) Quantile of the observed moment $\Sigma_{100}(t)$ released in a sliding window of $m = 100$ events with respect to a resampled process (solid line) and tapered Pareto process (dashed line).

$$M = 10^{1.5Mw+9.05} \tag{1}$$

The sequence of examined moments that correspond to $M_w \geq 7$ ($M \geq 3.55 \times 10^{19}$ Nm) is shown in Figure 1a. The annual number $N(M)$ of earthquakes with moment equal to or above M is conventionally described by a power law [*Kagan*, 2002]:

$$N(M) \propto M^{-\beta}, \beta \approx 2/3 \tag{2}$$

A proper normalization leads to the Pareto distribution

$$1 - F(M) = \text{Prob}[\text{seismic moment} > M] = \left(\frac{M}{M_{\min}}\right)^{-\beta}, M \geq M_{\min}. \tag{3}$$

The Pareto distribution (3) is a good approximation to the small-to-intermediate earthquake moments in a global catalog [*Kagan*, 2002]. At the same time, the distribution of the largest moments is currently unsettled due to the scarcity of data [e.g., *Bell et al.*, 2013; *Zöller*, 2013]. It is, however, a common observation that the

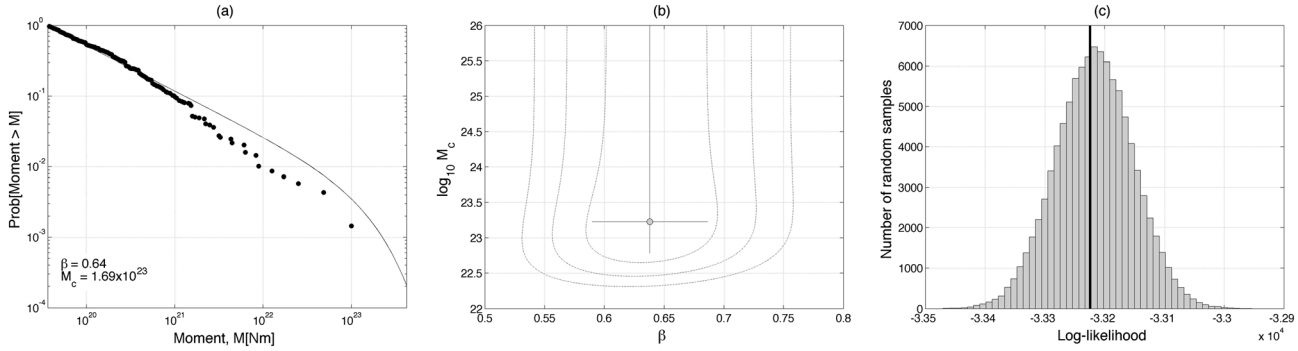

Figure 2. Tapered Pareto approximation of equation (4) to the observed seismic moment for events with $M_w \geq 7$ during 1918–2014 after declustering (692 earthquakes). (a) Estimated survival function Prob[moment $>$ M] (line) and the empirical distribution of the observed moments (circles). The distribution parameters estimated by the maximum likelihood method are indicated in the figure. The largest event (1960 Chile, $M = 2.8 \times 10^{23}$) corresponds to zero survival function value, and hence is not shown. (b) Likelihood profile (dashed lines), the maximum likelihood estimation (circle), and 95% marginal confidence intervals (lines) for the tapered Pareto parameters (β, M_c). (c) Log-likelihood for the observed moment sample (vertical black line) versus the distribution of the log likelihood for 10^5 random samples from a tapered Pareto distribution with the same parameters.

number of the largest earthquakes (e.g., $M > 10^{21}$ Nm, or $M_w > 8$) in the available global catalogs is smaller than that predicted by equation (3), as illustrated in Figure 2a [see also *Kagan*, 2002; *Zaliapin et al.*, 2005]. A theoretical support to this observation comes from the fact that Pareto law (3) with $\beta < 1$ has infinite mathematical expectation; this implies an infinite total moment release [e.g., *Knopoff and Kagan*, 1977; *Zaliapin et al.*, 2005] that is physically implausible. To avoid this problem, the tail of the distribution (3) should be modified to decay faster than M^{-1}. A common choice is the tapered Pareto distribution, which uses an exponential truncation of the power tail:

$$1 - F(M) = \left(\frac{M}{M_{\min}}\right)^{-\beta} \exp\left(\frac{M_{\min} - M}{M_c}\right), M \geq M_{\min}. \quad (4)$$

The power exponent β and corner moment M_c of the tapered Pareto distribution (4) can be estimated using the maximum likelihood approach [*Kagan and Schoenberg*, 2001; *Kagan*, 2002; *Vere-Jones et al.*, 2001], which we use in this study. We emphasize however, that the main conclusions of our work are independent of a specific form of the tail of the moment distribution, as we discuss below.

2.3. Examined Data Subsets and Tapered Pareto Parameters

We refer to the results of *Michael* [2014] on the catalog completeness and consider two subsets of the examined data: (i) $M_w \geq 7$ during 1918–2014. This subset contains 881 original events and 692 events after declustering. The tapered Pareto parameter estimates are $\beta = 0.64$ and $M_c = 1.69 \times 10^{23}$ Nm (Figure 2). (ii) $M_w \geq 6.6$ during 1955–2014. This subset contains 1387 original events and 983 events after declustering. It corresponds to the tapered Pareto estimates $\beta = 0.63$ and $M_c = 2.53 \times 10^{23}$ Nm (Figure S4 in the supporting information).

2.4. Resampling Analysis

Consider the observed seismic moment release as a marked point process $\{t_i, M_i\}$, using only times and seismic moments of earthquakes, and disregarding the event locations. The null hypothesis tested in this work is that the earthquake moments are independent of the occurrence times. Under this null, given the seismic moments M_i and occurrence times t_i of N observed earthquakes, the observed moment release $\{t_i, M_i\}$ is a realization of a marked point process $R(t, M)$ that assigns the probability $[N!]^{-1}$ to every trajectory of the form $\{t_i, M_{\pi(i)}\}$ for any permutation π on $\{1,\ldots,N\}$. Accordingly, the statistical properties of the observed realization $\{t_i, M_i\}$ should be comparable to those of the other trajectories $\{t_i, M_{\pi(i)}\}$. However, if the observed trajectory shows properties that cannot be reproduced by process $R(t, M)$, this may indicate that the initial assumption is violated—the times and seismic moments in the examined catalog are interrelated. We simulate sufficiently large collections of trajectories from process R by reshuffling the moments M_i of the observed events.

2.5. Monte Carlo Analysis

Alternatively, a process that might have produced the observations $\{t_i, M_i\}$ can be generated by Monte Carlo simulations of the tapered Pareto model of equation (4). Specifically, given times t_i, we consider a collection of seismic moments M_i that are independent of the times and each other and are drawn from a tapered Pareto distribution with parameters matching that for the examined earthquakes. The distribution parameters are obtained using the maximum likelihood estimation for the observed seismic moments.

3. Motivation: Apparent Fluctuations of the Moment Release

This section presents observations that provide intuition and motivation for our later formal analyses. Figure 1a shows time-dependent estimation of the corner moment M_c in the examined catalog with $M_w \geq 7$ during 1918–2014 in a moving window of 100 events, with the estimated value shown at the time of the last window event. This plot may suggest that the corner moment varies in time, with peaks ($M_c \geq 10^{23}$ Nm) during the 1960s and after December 2004 and a trough ($M_c \approx 10^{21}$ Nm) during 1980–2004. The respective estimation of the power index β is shown in Figure S2. Similar results are obtained for $M_w \geq 6.6$ during 1955–2014 (Figure S3). Next, Figure 1b shows the total seismic moment $\Sigma_m(t)$ released in a sliding window of $m = 100$ consecutive events. If seismic moments are independent of times, then $\Sigma_m(t)$ has a stationary distribution that can be approximated by resampling—drawing a large number B of independent random samples of size m from the population $\{M_i\}$ of the observed moments. Such resampling quantiles estimated with $B = 10^5$ are shown by dashed lines in Figure 1b. Notably, the observed moment spends over 20 years (1934–1936 and 1980–2001) under the 0.05 quantile and approaches the 0.001 quantile around 1995. Figure 1c provides an alternate view of the same data: it rescales the y axis to show the quantile of $\Sigma_{100}(t)$ with respect to the resampling distribution (solid line). The panel also shows quantiles of $\Sigma_{100}(t)$ with respect to a tapered Pareto distribution with parameters estimated from the examined events (listed in section 2.3). This plot emphasizes the apparent moment deficiency during the 1990s.

Of course, a mere observation of fluctuations in the estimated corner moment (Figure 1a) or cumulative moment release (Figure 1b) does not imply the existence of actual temporal changes. For instance, choosing too short a window would produce artificial variations of the moment due to sample variability. Furthermore, the quantile values shown in Figures 1b and 1c cannot be interpreted as the P values of a level-exceedance test. A formal significance analysis has to account for dependence among consecutive values of Σ_{100} and dig into deeper level-exceedance properties of a stationary moment release process. We only use this sliding window analysis for visual and intuitive support to our more formal techniques that are introduced below.

4. Results

4.1. Sample-Wide Likelihood Analysis

Figure 2a illustrates the fit of the tapered Pareto distribution of equation (4) in a declustered catalog with $M_w \geq 7$ during 1918–2014, using the maximum likelihood approach [*Vere-Jones et al.*, 2001]. The likelihood profile and marginal confidence intervals of the tapered Pareto parameters are shown in Figure 2b. The likelihood estimation provides finite margins of error for the power law index β, which is reflected in closed likelihood domain for β at every value of the corner moment M_c. At the same time, the estimation of the corner moment is unconstrained from above (open likelihood profile), which indicates that the data are consistent with any large enough value of M_c. This is a well-known property of corner moment estimation [e.g., *Kagan*, 2002; *Vere-Jones et al.*, 2001].

The tapered Pareto distribution provides a close approximation to the moments up to $M = 10^{21}$ Nm, while the larger moments visually deviate downward from the estimated distribution (Figure 2a). This deviation, however, is not statistically significant, as is illustrated in our next experiment. Namely, we generate 10^5 random samples of size $N = 692$ (the same as in the examined catalog) from the tapered Pareto distribution with parameters $\beta = 0.64$, $M_c = 1.69 \times 10^{23}$ Nm estimated from the observations. For every sample, we repeat the analysis done for the observed earthquakes—estimate the tapered Pareto parameters (that may deviate from those used to generate the sample) and compute the likelihood of the synthetic sample with respect to the estimated tapered Pareto distribution. Recall that the likelihood of a sample $\{X_i\}$, $i = 1,\ldots,N$ with respect to a probability density function $f(x)$ is defined as $L = \Pi_{i=1}^{N} f(X_i)$. Figure 2c shows that the observed likelihood

(vertical line) falls in the middle of the synthetic likelihood distribution, suggesting that the deviations of the observed moment sample from its estimated distribution are comparable to those in a random sample from the tapered Pareto distribution. This suggests that while our fitted distribution overestimates the probability of $M > 10^{21}$ Nm, from the sample-wide likelihood point of view, such overestimation does not imply that the tapered Pareto distribution provides an unreasonable fit to the observed data. Similar results are obtained for events with $M_w \geq 6.6$ during 1955–2014; see Figure S4.

The analysis of Figure 2 cannot be used to conclude that the examined moment sample is actually coming from a tapered Pareto distribution with fixed parameters. To show this, we simulate a mixture of two tapered Pareto distributions with common lowest moment $M_{\min} = 3.55 \times 10^{19}$ Nm ($M_w = 7.0$), power index $\beta = 0.64$, and different corner moments of $M_c = 8 \times 10^{23}$ and 3×10^{21} Nm, respectively. We then apply the above likelihood analysis to this mixture (see Figure S5). The analysis fails to detect the existence of two populations—similar to the results for the observed earthquakes, the likelihood of our mixed sample is not statistically different from that of a sample from a single tapered Pareto distribution with estimated parameters $\beta = 0.64$ and $M_c = 1.45 \times 10^{23}$.

This failure to detect a mixture of two distributions is related to the unnecessary generality of an alternate model—the above likelihood analysis compares the tapered Pareto distribution with estimated parameters against all possible deviations from this model. To improve the power of inference, we consider in the next section a specific alternate model: a tapered Pareto distribution with time-varying parameters.

4.2. Likelihood Analysis With a Time-Dependent Alternative

In this section, we compare the goodness-of-fit for two alternative models of the seismic moment release based on the tapered Pareto distribution (4). Model A assumes fixed parameters over the entire time interval; in Model B the parameters may vary, being constant within given subintervals. The models are compared using the Akaike information criterion (AIC) [*Akaike*, 1974] that combines the goodness-of-fit and complexity of a model in a quantity

$$\mathrm{AIC} = 2[\text{no.of model parameters} - \ln(\text{model likelihood})]. \tag{5}$$

A model with the smallest AIC value provides the best fit to the data. In particular, the relative likelihood of two models with $\mathrm{AIC}_1 < \mathrm{AIC}_2$ is $\exp[(\mathrm{AIC}_1 - \mathrm{AIC}_2)/2]$, which provides an intuitive interpretation for the difference $\Delta = \mathrm{AIC}_1 - \mathrm{AIC}_2$.

Model A here has two parameters—M_c and β; Model B with k subintervals has $(3k - 1)$ parameters: two parameters for the tapered Pareto distribution in each of k subintervals and $(k - 1)$ interval boundaries. Visual inspection of Figure 1 may suggest the following set of intervals for Model B: [1918, 1950], [1950, 1975], [1975, 2000], and [2000, 2014]. With this choice of subintervals and for $M_w \geq 7$ during 1918–2014 we obtain $\Delta = \mathrm{AIC}_B - \mathrm{AIC}_A = -24$ ($\mathrm{AIC}_A = 66{,}453$; $\mathrm{AIC}_B = 66{,}429$) and for $M_w \geq 6.6$ during 1955–2014 (leaving only three intervals for model B) we obtain $\Delta = \mathrm{AIC}_B - \mathrm{AIC}_A = -13$ ($\mathrm{AIC}_A = 91{,}770$; $\mathrm{AIC}_B = 91{,}757$). The negative values of Δ show that Model B, which allows fluctuations of the tapered Pareto parameters with time, provides a much closer fit to the data than a time-independent Model A.

To ensure that this analysis is not biased by our selection of subintervals, we perform a random-boundary experiment. Figure 3 compares the distribution of Δ values obtained by selecting random boundaries that partition the interval [1918, 2014] into four subintervals, with a condition that each interval is longer than 5 years. The random boundary model is applied 10^4 times to the observed catalog (Figure 3a), to 10^3 resampled catalogs $\{t_i, M_{\pi(i)}\}$ (Figure 3b), and to 10^3 synthetic catalogs with moments drawn independently from a tapered Pareto distribution with time-independent parameters (Figure 3c). This analysis shows that (i) as expected, synthetic sequences (Figure 3c) ultimately prefer time-independent parameterization ($\Delta > 0$), (ii) resampled sequences (Figure 3b) strongly prefer time-independent parameterization, and (iii) the examined sequence of observed moments is often better parameterized by a time-dependent model ($\Delta < 0$), with Δ values extending to -30, which is never achieved in our simulation experiments. Moreover, the entire observed distribution of Δ is shifted toward negative vales with respect to the synthetic distributions; this means that even for those random subintervals that prefer time-independent parameterization ($\Delta > 0$), the time-independent model fit to the observed seismicity has lower value of Δ (less preference to a time-independent model) compared to that in synthetic or resampled sequences. This suggests that the

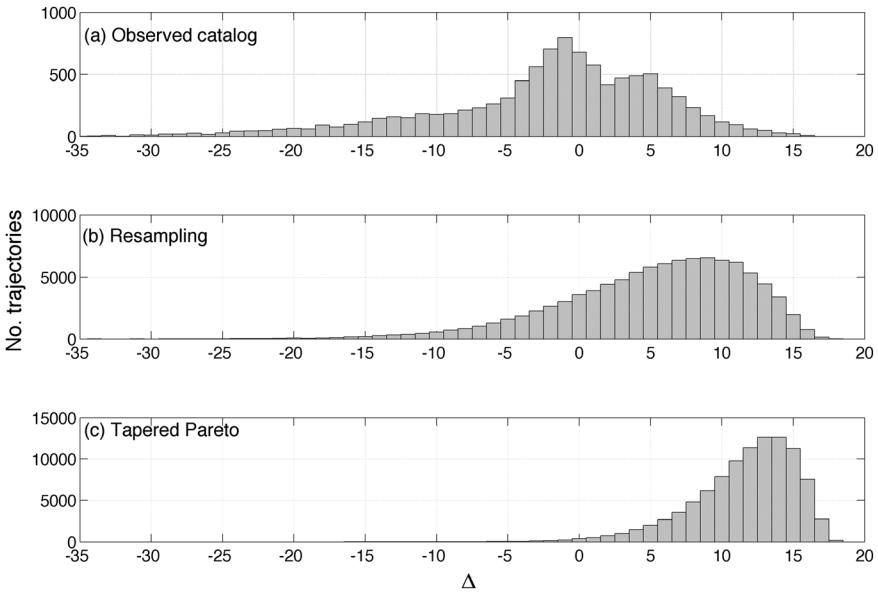

Figure 3. Likelihood analysis with a time-dependent alternative (section 4.2). The figure shows the distribution of $\Delta = \mathrm{AIC}_{\text{time-dependent}} - \mathrm{AIC}_{\text{time-inependent}}$. The time-dependent model uses four random subintervals. Positive values of Δ indicate preference to a time-independent model. (a) Observed sequence of 692 moments in a declustered catalog for $M_w \geq 7$ during 1918–2014; 10^4 sets of random boundaries. (b) Resampling: 10^3 reshuffled catalogs $\{t_i, M_{\pi(i)}\}$; 100 sets of random boundaries for each catalog. (c) Synthetic sequence of 692 moments from tapered Pareto distribution with $\beta = 0.64$ and $M_c = 1.69 \times 10^{23}$ Nm: 10^3 synthetic catalogs; 100 sets of random boundaries for each catalog.

observed catalog cannot be considered a realization of a process with independent times and seismic moments. Finally, the distribution of Δ in a random moment catalog (Figure 3c) is significantly different from that in a resampled catalog (Figure 3b), which provides further support to our claim that the observed sequence of moments cannot be modeled by a fixed-parameter tapered Pareto distribution. Similar results are obtained with different number of random subintervals and for earthquakes with $M_w \geq 6.6$ during 1955–2014 (Figure S6).

4.3. Regression Analysis

Here we test independence of times and seismic moments in a regression framework. For any fixed integer m, let $S_{m,j} = M_{jm-m+1} + \ldots + M_{jm}$ be the sequence (indexed by j) of total seismic moments released in nonoverlapping groups of m consecutive events, and $u_{m,j} = t_{jm}$ be the time of the last event in jth group. We introduce two related quantities: $E_{m,j}$ is an effective moment magnitude that corresponds to the seismic moment $S_{m,j}$ according to equation (1), and $r_{m,j}$ is the rank (position in the ordered sequence) of the moment $S_{m,j}$ among $\{S_{m,j}\}_j$. If the sequence of moments M_i is independent of occurrence times t_i, then the sequences $S_{m,j}$, $E_{m,j}$, and $r_{m,j}$ are independent of $u_{m,j}$ for any fixed m. Consider a regression model

$$\mu \equiv E[r_m] = \alpha + \alpha_1 z + \alpha_2 z^2 + \alpha_3 z^3 + \alpha_4 z^4, \tag{6}$$

where $z = u_m - \mathrm{mean}(u_m)$. Using the ranked values r_m instead of S_m as the response variable reduces the model estimation artifacts related to asymmetric, heavy-tailed distribution of the seismic moment values. In the absence of dependence between r_m and u_m, the coefficients α_i, $i = 1,2,3,4$ in (6) are not significant. To assess finite sample properties of the model estimation for the examined data, we fix $m = 20$ and fit the model (6) to 10^5 independent resampled catalogs $\{t_i, M_{\pi(i)}\}$. The resulting distribution of each of the P values for α_i is uniform (according to Kolmogorov-Smirnov test), with statistically significant correlations among the estimated coefficients. For instance, under the assumption of independent uniform distribution of the four P values, 5% (1%) of the experiments will have the minimal P value below 0.013 (0.0026), according to the Bonferroni correction. The respective value in our simulations is 3.4% (0.7%), which reflects the correlation among the regression coefficients.

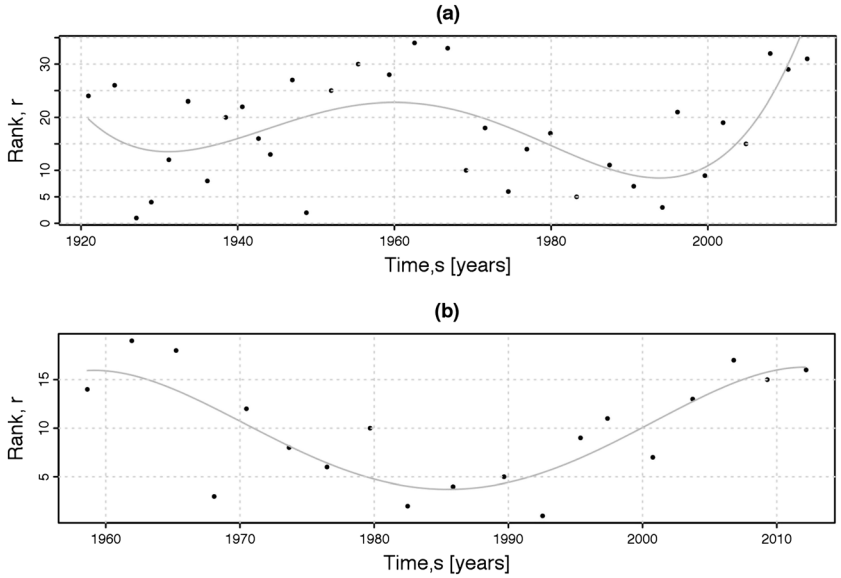

Figure 4. Regression analysis. The rank r_m of the total moment S_m released in nonoverlapping groups of m consecutive events (black dots), approximated by a time-dependent polynomial of fourth degree (gray line). Declustered catalog. (a) $M_w \geq 7$ during 1918–2014, $m = 20$. (b) $M_w \geq 6.6$ during 1955–2014, $m = 50$.

The estimation of the regression model for the observed declustered seismicity with $M_w \geq 7$ during 1918–2014 and $m = 20$ is illustrated in Figure 4a. The coefficients α_i have the P values of 7×10^{-2}, 6×10^{-3}, 7×10^{-2}, and 4×10^{-3}, respectively. Comparing this with the resampled results, we find that (i) only 1% of the simulated catalogs had the minimal P value below that observed in the real catalog and (ii) only 0.3% (0.6%) of the simulated catalogs had all P values below 0.07 (0.1), which is the case for the observed catalog. Furthermore, the goodness-of-fit in the observed catalog is much better than that in simulations: the value of the AIC in the observed data is below 99.4% of the AIC values in the resampling experiments. This emphasizes the existence of time-dependent pattern of moment release in the observed data that creates a possibility for a good time-dependent fit and that is not reproduced by resampled catalogs. Very similar results, which strongly support time-moment dependence, are obtained with (i) alternative choices of m, (ii) using $S_{m,j}$ or $E_{m,j}$ as the response variable, and (iii) with $M_w \geq 6.6$ during 1955–2014 (Figure 4b). The regression order of four is selected here as the minimal order capable of depicting two apparent peaks of moment release. Text S2 describes additional regression experiments that do not require preselecting the regression order; they are in a strong agreement with the results shown in this section.

5. Discussion and Conclusions

The apparent increase of global observed seismic moment release since the 2004 great Sumatra earthquake has been debated as either an expected statistical fluctuation of a time-independent process or as an indication of time-dependent changes. Our analysis suggests that the observed moment release is time-dependent. Specifically, we combine likelihood goodness-of-fit analysis (section 4.2 and Figures 3 and S6) and regression analysis (section 4.3 and Figure 4) to see if a model with time-independent parameters can fit the observed sequence of seismic moments. We find that this in not the case—a time-dependent model is strongly preferred by the considered approaches; this corroborates a visual impression from sliding window seismic moment analysis (section 3 and Figures 1 and S3).

Our results have at least two distinct interpretations. First, they might suggest that the seismic moment during the 1960s and after December 2004 is drawn from a significantly different distribution compared to the one estimated in a time-independent model for the whole catalog. Alternatively, the consecutive values of the seismic moment might be correlated (and may or may not come from the same distribution). Finding the actual mechanism of moment fluctuations, statistical and/or physical, is a matter of separate research. An important conclusion of this study is that independently of a particular mechanism, the observed raise of the

great earthquakes after December 2004 is a statistically significant feature of the global moment release process.

Shearer and Stark [2012] recognized diminished moment release in the decades before December 2004, although they dismissed that as indicative of time-dependent behavior [see also *Michael*, 2011; *Daub et al.*, 2012; *Beroza*, 2012; *Parsons and Geist*, 2012, 2014; *Ben-Naim et al.*, 2013]. This dismissal is explained by the focus on the rate of events in time, rather than consecutive moment release. In general, studying the rate of events above a threshold $M_{\min}$ may decrease the power of respective inference. Indeed, selecting a high $M_{\min}$ leads to irresolute analysis of small number of examined events, while selecting a low $M_{\min}$ masks the changes in the large-moment domain by numerous small events. Considering the temporal variation in moment release, as is done in this study, might remove this problem.

Testing dependence between times and seismic moment is a challenging problem, with many conventional tests being not powerful enough to properly reject the independence null, because of a too general alternative and an intricate (nonlinear and nonmonotone) form of time-moment dependence. This is the case for the conventional Pearson, Kendall, Spearman, and Hoeffding independence tests [*Hollander et al.*, 2014], as well as much more powerful distance correlation test [*Székely et al.*, 2007]. We do not show these negative results here, restricting this work to the approaches that have sufficient power to reject the independence null.

The period of significantly lower moment release followed immediately the decade with the highest moment release during the 1960s. *Bufe and Perkins* [2005] already showed that this pattern is unlikely random and that it suggests that the global variation in seismic moment release has physical underpinning. The most likely candidate for a multidecadal global low following the 1960s high could be stress change related to viscoelastic relaxation. Indeed, such a mechanism has been used to explain an apparent decrease in seismicity rate in southern California after the Great Chile and Alaska earthquakes in the 1960s [*Selva and Marzocchi*, 2005]. Although that work has been extended globally [*Marzocchi and Selva*, 2008], it has only been based on seismicity rate variations. Any rigorous test of the role of viscoelastic relaxation on the seismicity rate of all or only the largest events has to take into account geometrical considerations; e,g, the largest earthquakes can only occur at a limited number of locations which may or may not be optimally orientated to be susceptible to relaxation-induced stresses. Still, our finding that the global temporal fluctuation of seismic moment actually involves a significant change in the corner magnitude may provide new insight behind any physical mechanism.

Acknowledgments
We are grateful to Andy Michael, Editor Andrew Newman, and an anonymous reviewer whose comments helped to substantially improve the initial version of the paper. This research was supported by the Southern California Earthquake Center (contribution no. 6265). SCEC is funded by NSF Cooperative Agreement EAR-1033462 and USGS Cooperative Agreement G12AC20038. The ISC-GEM catalog is available at http://www.isc.ac.uk/iscgem/.

References

Akaike, H. (1974), A new look at the statistical model identification, *IEEE Trans. Autom. Control*, *19*(6), 716–723, doi:10.1109/TAC.1974.1100705.

Ammon, C. J., R. C. Aster, T. Lay, and D. W. Simpson (2011), The Tohoku earthquake and a 110-year spatiotemporal record of global seismic strain release, *Seismol. Res. Lett.*, *82*, 454.

Bell, A. F., M. Naylor, and I. G. Main (2013), Convergence of the frequency-size distribution of global earthquakes, *Geophys. Res. Lett.*, *40*, 2585–2589, doi:10.1002/grl.50416.

Ben-Naim, E., E. G. Daub, and P. A. Johnson (2013), Recurrence statistics of great earthquakes, *Geophys. Res. Lett.*, *40*, 3021–3025, doi:10.1002/grl.50605.

Beroza, G. C. (2012), How many great earthquakes should we expect?, *Proc. Natl. Ac. Sci. U.S.A.*, *109*(3), 651–652.

Burnham, K. P., and D. R. Anderson (2003), *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, pp. 1–454, Springer-Verlag, New York.

Bufe, C. G., and D. M. Perkins (2005), Evidence for a global seismicmoment release sequence, *Bull. Seismol. Soc. Am.*, *95*(3), 833–843, doi:10.1785/0120040110.

Bufe, C. G., and D. Perkins (2011), The 2011 Tohoku earthquake: Resumption of temporal clustering of Earth's megaquakes, *Seismol. Res. Lett.*, *82*, 455.

Daub, E. G., E. Ben-Naim, R. A. Guyer, and P. A. Johnson (2012), Are megaquakes clustered?, *Geophys. Res. Lett.*, *39*, L06308, doi:10.1029/2012GL051465.

Gu, C., A. Y. Schumann, M. Baiesi, and J. Davidsen (2013), Triggering cascades and statistical properties of aftershocks, *J. Geophys. Res. Solid Earth*, *118*, 4278–4295, doi:10.1002/jgrb.50306.

Hollander, M., D. A. Wolfe, and E. Chicken (2014), *Nonparametric Statistical Methods*, 3rd ed., 844 pp., John Wiley, Hoboken, N. J.

Kagan, Y. Y. (2002), Seismic moment distribution revisited: I. Statistical results, *Geophys. J. Intl.*, *148*(3), 520–541.

Kagan, Y. Y., and F. Schoenberg (2001), Estimation of the upper cutoff parameter for the tapered Pareto distribution, *J. Appl. Prob.*, *158-175*.

Knopoff, L., and Y. Y. Kagan (1977), Analysis of the theory of extremes as applied to earthquake problems, *J. Geophys. Res.*, *82*, 5647–5657, doi:10.1029/JB082i036p05647.

Lay, T. (2015), The surge of great earthquakes from 2004 to 2014, *Earth Planet. Sci. Lett.*, *409*, 133–146.

Marzocchi, W., and J. Selva (2008), Long-term influence of giant earthquakes: Backward empirical evidence and forward test, *Bull. Seismol. Soc. Am.*, *98*(3), 1102–1112, doi:10.1785/0120070203.

Michael, A. J. (2011), Random variability explains apparent global clustering of large earthquakes, *Geophys. Res. Lett.*, *38*, L21301, doi:10.1029/2011GL049443.
Michael, A. J. (2014), How complete is the ISC-GEM global earthquake catalog?, *Bull. Seismol. Soc. Am.*, *104*(4), 1829–1837, doi:10.1785/0120130227.
Parsons, T., and E. L. Geist (2012), Were global $M \geq 8.3$ earthquake time intervals random between 1900 and 2011?, *Bull. Seismol. Soc. Am.*, *102*(4), 1583–1592.
Parsons, T., and E. L. Geist (2014), The 2010–2014.3 global earthquake rate increase, *Geophys. Res. Lett.*, *41*, 4479–4485, doi:10.1002/2014GL060513.
Schoenball, M., N. C. Davatzes, and J. M. G. Glen (2015), Differentiating induced and natural seismicity using space-time-magnitude statistics applied to the Coso geothermal field, *Geophys. Res. Lett.*, *42*, 6221–6228, doi:10.1002/2015GL064772.
Shearer, P. M., and P. B. Stark (2012), Global risk of big earthquakes has not recently increased, *Proc. Natl. Ac. Sci. U.S.A.*, *109*(3), 717–721.
Selva, J., and W. Marzocchi (2005), Variations of southern California seismicity: Empirical evidence and possible physical causes, *J. Geophys. Res.*, *110*, B11306, doi:10.1029/2004JB003494.
Stein, S., and E. A. Okal (2005), Speed and size of the Sumatra earthquake, *Nature*, *434*, 581–582, doi:10.1038/434581a.
Storchak, D. A., D. Di Giacomo, I. Bondár, E. R. Engdahl, J. Harris, W. H. K. Lee, A. Villaseñor, and P. Bormann (2013), Public release of the ISC-GEM global instrumental earthquake catalogue (1900–2009), *Seismol. Res. Lett.*, *84*(5), 810–815, doi:10.1785/0220130034.
Storchak, D. A., D. Di Giacomo, E. R. Engdahl, J. Harris, I. Bondár, W. H. K. Lee, P. Bormann, and A. Villaseñor (2015), The ISC-GEM global instrumental earthquake catalogue (1900–2009): Introduction, *Phys. Earth Planet. Inter.*, *239*, 48–63, doi:10.1016/j.pepi.2014.06.009.
Székely, G. J., M. L. Rizzo, and N. K. Bakirov (2007), Measuring and testing dependence by correlation of distances, *The Annals of Statistics*, *35*(6), 2769–2794.
Tsai, V. C., M. Nettles, G. Ekström, and A. Dziewonski (2005), Multiple CMT source analysis of the 2004 Sumatra earthquake, *Geophys. Res. Lett.*, *32*, L17304, doi:10.1029/2005GL023813.
Vere-Jones, D., R. Robinson, and W. Yang (2001), Remarks on the accelerated moment release model: Problems of model formulation, simulation and estimation, *Geophys. J. Intl.*, *144*(3), 517–531.
Zaliapin, I., and Y. Ben-Zion (2013a), Earthquake clusters in southern California I: Identification and stability, *J. Geophys. Res. Solid Earth*, *118*, 2847–2864, doi:10.1002/jgrb.50179.
Zaliapin, I., and Y. Ben-Zion (2013b), Earthquake clusters in southern California II: Classification and relation to physical properties of the crust, *J. Geophys. Res. Solid Earth*, *118*, 2865–2877, doi:10.1002/jgrb.50178.
Zaliapin, I., and Y. Ben-Zion (2016a), Discriminating characteristics of tectonic and human-induced seismicity, *Bull. Seismol. Soc. Am.*, *106*(3), 846–859, doi:10.1785/0120150211.
Zaliapin, I., and Y. Ben-Zion (2016b) Earthquake declustering via a nearest-neighbor approach in space-time-magnitude domain, Abstract S31E-07 presented at 2016 Fall Meeting, AGU, San Francisco, Calif., 12-16 Dec.
Zaliapin, I., A. Gabrielov, V. Keilis-Borok, and H. Wong (2008), Clustering analysis of seismicity and aftershock identification, *Phys. Rev. Lett.*, *101*(1), 018501.
Zaliapin, I., Y. Kagan, and F. Schoenberg (2005), Approximating the distribution of Pareto sums, *Pure. Appl. Geophys.*, *162*, 1187–1228.
Zhang, Q., and P. M. Shearer (2016), A new method to identify earthquake swarms applied to seismicity near the San Jacinto fault, California, *Geophys. J. Intl.*, *205*(2), 995–1005.
Zöller, G. (2013), Convergence of the frequency-magnitude distribution of global earthquakes: Maybe in 200 years, *Geophys. Res. Lett.*, *40*, 3873–3877, doi:10.1002/grl.50779.