Supporting Information for "Break of slope in earthquake-size distribution reveals creep rate along the San Andreas fault system"

Inessa Vorobieva,[1,2] Peter Shebalin,[1,2] Clément Narteau,[2]

Corresponding author: I. Vorobieva, Institute of Earthquake Prediction Theory and Mathematical Geophysics, 84/32 Profsouznaya, Moscow 117997, Russia. (vorobiev@mitp.ru)

[1]Institute of Earthquake Prediction
Theory and Mathematical Geophysics,
84/32 Profsouznaya, Moscow 117997,
Russia.

[2]Équipe de Dynamique des Fluides
Géologiques, Institut de Physique du Globe
de Paris, Sorbonne Paris Cité, Univ. Paris
Diderot, UMR 7154 CNRS, 1 rue Jussieu,
75238 Paris, Cedex 05, France.

Contents of this file

- Text S1 to S7:

  Text S1. Multiscale analysis of spatial and temporal variations of the magnitude of completeness.

  Text S2. Determination of b-value over a truncated range of magnitude.

  Text S3. Statistical test for the GR law along the SAF.

  Text S4. Variation of the creep rate and earthquake-size distribution with depth.

  Text S5. Variation of the seismic activity with respect to creep rate.

  Text S6. Relation between creep rate and the break of slope of the earthquake-size distribution using constant sample size.

  Text S7. Influence of large earthquakes and their aftershocks on long-term earthquake-size distribution.

- Table S1.

- Figures S1 to S11.

Text S1. Multiscale analysis of spatial and temporal variations of the magnitude of completeness

We expand the multiscale method proposed by *Vorobieva et al.* [2013] to estimate the completeness magnitude M_c of earthquake catalogues over space and time.

The Gutenberg–Richter law describing the earthquake frequency–magnitude distribution (FMD) might not hold over the entire magnitude range, and a scaling relation should be obtained by adapting the dimension of the studied space-time volume to the magnitude range. Based on empirical relations in seismotectonics, we associate ranges of larger magnitudes with increasing space-time volume for data selection. Then, for each point in space and time, we document the earthquake FMD at all length scales within the corresponding earthquake magnitude ranges. High resolution of the M_c-value is achieved through the determination of the smallest space–time-magnitude scale in which the Gutenberg–Richter law is verified. The multiscale procedure isolates the magnitude range that meets the best local seismicity and record capacity.

Here we analyse linear seismic zones in which individual event may be characterised by one spatial coordinate and one time coordinate. The earthquake magnitude scale is decomposed into a set of ranges $\mathcal{M}_i = [M_i, M_i + W_M]$ using an overlapping sliding window of constant width W_M and step Δm. Each earthquake magnitude range $\mathcal{M}_i$ is associated with a characteristic length scale

$$R_i = R_0 \times 10^{pM_i} \tag{1}$$

and a characteristic time scale

$$T_i = T_0 \times 10^{pM_i}. \tag{2}$$

For a given magnitude range $\mathcal{M}_i$ and a given point in space and time along the fault, we select all the $[M_i, M_i + W_{\mathrm{M}}]$ events occurring within a distance of R_i over a time period T_i preceding the point in time under investigation.

The further spatial-temporal analysis of the completeness magnitude M_{c} is similar to the spatial analysis detailed in *Vorobieva et al.* [2013]. Analysing all the the magnitude ranges $\mathcal{M}_i$ at a given point in space and time, the M_{c}-value corresponds to the lower limit of the smallest range $\mathcal{M}_i$ in which the FMD satisfies the GR law. In practice, we start with the smallest magnitude–space–time domain and successively test if the catalogue is complete according to the following conditions:

1. The number of events in the magnitude range $\mathcal{M}_i$ within the space-time domain is larger than a constant number N_{c}.

2. The FMD within the magnitude range $\mathcal{M}_i$ satisfies the GR law.

A local M_{c}-value is given by the smallest M_i-value for which these two conditions are met. M_{c}-value maps are obtained by repeating the procedure at every point in space and time.

We compute the M_{c}-value in the period 1984-2014 using the NCEDC double-difference earthquake catalogue [*Waldhauser and Schaff*, 2008; *Waldhauser*, 2009]) along two linear fault zones: the San Andreas fault from the Peninsula to the Cholame-Carrizo sections; the Hayward-Calaveras faults (Fig. S1a). The M_{c} maps shown in Figs. S1b,c are constructed with the set of parameters shown in Tab. S1. The M_{c}-value is smaller than 1.5 in both space-time regions under consideration.

Text S2. Determination of b-value over a truncated range of magnitude

Considering that the FMD follows the GR law, the b-value may be obtained with the maximum likelihood point estimator method. For a continuous (i.e., exact) magnitude distribution and an infinite maximum magnitude, *Aki* [1965] shows that

$$b = \frac{\log_{10}(e)}{\langle M \rangle - M_*}, \tag{3}$$

where $\langle M \rangle$ is the mean magnitude of the sample and M_* a lower limit for which the catalogue is complete. Nevertheless, for a finite maximum magnitude and/or grouped magnitude data, Eq. 3 gives a biased estimation of the b-value due to biased values of $\langle M \rangle$ and M_*. To overcome these issues, *Bender* [1983] proposes a new set of maximum likelihood formulas. Based on these formulas, *Vorobieva et al.* [2013] develop an iterative procedure to compute the b-value on a truncated magnitude range $[M_{\min}, M_{\max}]$. The procedure is verified on synthetic datasets with known b-values ([*Vorobieva et al.*, 2013], table 1)

Let

$$W_{\mathrm{M}} = M_{\max} - M_{\min} = K\Delta m. \tag{4}$$

The $[M_{\min} - \Delta m/2, M_{\max} + \Delta m/2]$ magnitude range is decomposed into $K + 1$ bins of constant width Δm. The central value of each magnitude bin is

$$M_k = M_{\min} + k\Delta m, \tag{5}$$

with $k = \{0, 1, 2, \ldots, K\}$, so that $M_0 = M_{\min}$ and $M_K = M_{\max}$.

For each bin, n_k and $\langle m \rangle_k$ are the number and the mean magnitude of $M \in [M_k - \Delta m/2, M_k + \Delta m/2[$ earthquakes, respectively. Similarly, N_k and $\langle M \rangle_k$ are the number

and the mean magnitude of $M \in [M_k - \Delta m/2, \infty[$ earthquakes, respectively. At the beginning of the iterative procedure, the initial b-value for the entire magnitude range is

$$b_0 = \frac{\log_{10}(N_1) - \log_{10}(N_K)}{M_K - M_1}. \tag{6}$$

For a given b_j-value, an iteration consists of the following steps:

1. Estimation of $\langle M \rangle_K$ taking $M_* = M_{\max} - \Delta m/2$ and $b = b_j$ in Eq. 3.

2. Estimation of $\langle m \rangle_k$ and $\langle M \rangle_k$ from $k = K - 1$ to $k = 1$ using successively for each bin k two recursive formulas

$$\langle m \rangle_k = \frac{\int_{M_k - \Delta m/2}^{M_k + \Delta m/2} \mu 10^{a - b_j \mu} \mathrm{d}\mu}{\int_{M_k - \Delta m/2}^{M_k + \Delta m/2} 10^{a - b_j \mu} \mathrm{d}\mu} = M_k + \frac{\log_{10}(e)}{b_j} - \frac{\Delta m}{10^{b_j \Delta m} - 1} - \frac{\Delta m}{2} \tag{7}$$

and

$$\langle M \rangle_k = \frac{\langle M \rangle_{k+1} N_{k+1} + \langle m \rangle_k n_k}{N_k} \tag{8}$$

3. Estimation of b_{j+1} taking $\langle M \rangle = \langle M \rangle_1$ and $M_* = M_1 - \Delta m/2$ in Eq. 3.

The iterative process stops when $|b_j - b_{j-1}| < \varepsilon$. In a vast majority of cases, it takes less than 20 iterations to converge to the stationary b-value using $\varepsilon = 10^{-3}$. Nevertheless, if the iterative procedure does not converge after 100 iterations, we consider that the FMD does not follow a GR law within the magnitude range $[M_1,\ M_K]$, i. e. $[M_{\min} + \Delta m,\ M_{\max}]$.

It is commonly considered that right part of the FMD is not critical for b-value estimates, so that the truncation of FMD in the range of larger magnitudes does not influence the result. However, when the FMD differs significantly from an exponential distribution, the cut-off in large magnitudes may significantly change b-value estimates. This is the case in the creeping section of the San Andreas fault, where FMD does not follow the GR law

and where the differences in b-value reach 0.2 when increasing the large magnitude cut-off (Fig. S2).

Text S3. Statistical test for the GR law along the SAF

Studying the FMDs along the San Andreas faults system, we find significant differences in b-value estimates depending on intervals of magnitude. Based on these observations, we conclude that FMD does not always follow the GR law. The null hypothesis H_0 is that the true size distribution of earthquakes is exponential. We use Pearson's chi-squared test to evaluate how likely the observed difference with the exponential distribution arose by chance. The chi-squared test is based on the analysis of non-cumulative distribution.

The value of the test-statistic X^2 estimates the difference between an observed and a theoretical distribution

$$X^2 = N \sum_{i=1}^{k} \frac{(\frac{n_i}{N} - p_i)^2}{p_i} \tag{9}$$

here n_i is non-cumulative number of earthquakes in the i-th magnitude bin, $N = \sum_{i=1}^{k} n_i$ total number of observed events, k the number of magnitude bins, and p_i the theoretical probability that an event falls to the i-th magnitude bin. The observed FMD is constructed in 40 km long segments along the San Andreas fault system, and the theoretical exponential FMD is constructed using the best-fit values of the parameters of the GR law obtained with maximum likelihood point estimator in the entire range of magnitudes $M \geq 1.5$. The X^2-value follows the χ^2_{k-1} distribution with $k-1$ degrees of freedom. Using this distribution and a given level of confidence, the test relies on the probability P to have a value larger than the observed X^2-value.

As shown in Fig. S3, the null hypothesis in the creeping section of San Andreas is rejected with very high confidence level, as the probability $P(\chi^2_{k-1} \geq X^2)$ is less than 10^{-6}. On the other hand, the the GR distribution is not rejected in the non-creeping segments.

We construct the earthquake size distributions for all locked/low creeping sections, and fast/moderate creeping sections of SAF, Calaveras and Hayward fault systems (Fig. S4). Earthquake-size distributions incorporate a large number of earthquakes, confirming at a high confidence level our main conclusions:

(1) locked seismogenic zones generally obey the GR law, whereas creep processes break it down, promoting instead band-limited scaling functions;

(2) creeping and locked segments obey the same scaling relation over lower magnitude ranges, but that there is a deficit of large events along creeping segments.

Text S4. Variation of the creep rate and earthquake-size distribution with depth

The creep rate is measured at the surface, and its changes in depth can only be inferred by modelling. Most theories suggest that creep rapidly decreases with depth and that the locked depth depends on the creep rate [e.g., *Templeton et al.*, 2008; *Weldon II et al.*, 2013]. In addition, below the brittle-ductile transition (> 15 km), continuous deformation occur and modify the scaling behaviour of the earthquake-size distribution [*Spada et al.*, 2013].

To concentrate only on the along-strike variation of creep-rate at the surface, we repeat the calculation of the $b_{M2.8}$ and $b_{M1.5}$-values using only the shallower events with a focal

depth $h \leq 7$ km (Fig. S5a). All the results are similar to those obtained across the entire seismogenic depth, except along the 270-290 km segment, where there is a fast transition in creep-rate from 27 mm/year to 0 mm/year (compare Fig. S5b and Fig. 1b of the main manuscript). The earthquake-size distribution in this zone of transition changes rapidly with depth: for $h \leq 7$ km events, there is a clear break of slope with a steeper slope in larger magnitude ranges as observed along fast creeping segments (Fig. S5c); for $h > 7$ km events, a single scaling behaviour is observed as along the locked segments of the San-Andreas fault (Fig. S5d). Such an observation is in good agreement with the inversion of *Murray et al.* [2001] who showed that the creep rate decreases rapidly with depth, reaching a minimum in the middle part of the seismogenic zone, at a depth of about 5 - 8 km. It also cofirms that the shape of the earthquake-size distribution may depend on the type of deformation.

The amount of recorded earthquakes with $M \geq 2.8$ along the Hayward-Calaveras faults within the upper 7 km is insufficient to determine $b_{M2.8}$-value except in the 103-147 km segment. The $b_{M1.5}$ and $b_{M2.8}$-values for this shallow seismicity are similar to ones within the entire seismogenic depth (Fig. S6).

Using shallow events with focal depth $h \leq 7$ km, the dependence of $b_{M1.5}$ and $b_{M2.8}$-values on creep rate is shown in Fig. S7.

Text S5. Variation of the seismic activity with respect to creep rate

A significant growth of the number of small earthquakes in the magnitude interval [1.5, 2.7] is observed along the fast and moderate creeping segments of the San Andreas fault where creep rate exceeds 12-15 mm/y. The number of earthquakes is an order of

magnitude larger than in the aftershock zones of strong earthquakes. Simultaneously the number of larger earthquakes in the magnitude interval [2.8, 4.0] strongly drops along the fast creeping segments where creep rate exceeds 18-20 mm/y (Fig. S8a). Similar variations of seismic activity are not observed along the Hayward-Calaveras faults because the creep rate is significantly smaller than along the San Andreas fault and does not exceed 10 mm/y (Fig. S8b).

Text S6. Relation between creep rate and the break of slope of the earthquake-size distribution using constant sample size

In this section we use constant sample size of earthquakes along the San-Andreas fault instead of constant 40 km segment. We select a fixed number of earthquakes in the magnitude interval [2.8, 4.0], and, using these earthquakes we obtain segments along the fault. Then, using these segments, we study size distribution of earthquakes in two successive magnitude intervals [1.5, 2.7] and [2.8, 4.0] (Figs. S9a and S10a).

The length of the segment containing 50 (100) earthquakes varies strongly along San Andreas depending on the seismic activity. It changes from 1.04 to 86 km for 50 events sample (Fig. S9b), and from 2.64 to 107 km for 100 events sample (Fig. S10b). Despite stronger variations of the $b_{M2.8}$-values than in constant segment length of 40 km, their average values correlate with the creep rate, while $b_{M1.5}$-values are more stable and look independent on the creep rate.

Text S7. Influence of large earthquakes and their aftershocks on long-term earthquake-size distribution

In order to test the influence of large earthquakes and their aftershocks on the earthquake-size distribution, we exclude one-year of seismicity after Loma Prieta and after Parkfield earthquakes. Thus, we remove most of their aftershocks and possible long-range interaction with other segments along the SAF. Fig. S11 shows the corresponding $b_{M1.5}$ and $b_{M2.8}$-values using the same methodology as in the main manuscript. The changes in b-values are only observed along the Loma Prieta zone, where the $b_{M2.8}$-value cannot be determined due to insufficient statistics and where the $b_{M1.5}$-value is increased by 0.2. The visual analysis of the earthquake-size distribution in the Loma Prieta zone does not show a breakdown for larger events (Fig. S11). This result shows that aftershocks of large earthquakes do not alter our conclusion about the relation between the break of slope in earthquake-size distribution and creep rate.

References

Aki, K. (1965), Maximum likelihood estimate of b in the formula $\log N = a - bM$ and its confidence level, *Bull. Earthquake Res. Inst.*, *43*, 237–239.

Bender, B. (1983), Maximum likelihood estimation of b-values for magnitude grouped data, *Bull. Seism. Soc. Am.*, *73*, 831–851.

Murray, J., P. Segall, and P. Cervelli (2001), Inversion of GPS data for spatially variable slip-rate on the San Andreas fault near Parkfield, *Geophys. Res. Lett.*, *28*, 359–362, doi: 10.1029/2000GL011933.

Spada, M. and Tormann, T. and Wiemer, S. and Enescu, B (2013), Generic dependence of the frequency-size distribution of earthquakes on depth and its relation to the strength profile of the crust, *Geophysical research letters*, *40*(4), 709–714, doi:

10.1029/2012GL054198.

Templeton, D., R. Nadeau, and R. Burgmann (2008), Behavior of repeating earthquake sequences in Central California and implications for subsurface fault creep, *Bull. Seism. Soc. Am.*, *98*(1), 52–65, doi: 10.1785/0120070026.

Vorobieva, I., C. Narteau, P. Shebalin, F. Beauducel, A. Nercessian, V. Clourad, and M.-P. Bouin (2013), Multiscale Mapping of Completeness Magnitude of Earthquake Catalogs, *Bull. Seism. Soc. Am.*, *103*, 2188–2202, doi: 10.1785/0120120132.

Waldhauser, F. (2009), Near-real-time double-difference event location using long-term seismic archives, with application to Northern California, *Bull. Seism. Soc. Am.*, *99*, 2736–2848, doi:10.1785/0120080294.

Waldhauser, F., and D. Schaff (2008), Large-scale relocation of two decades of Northern California seismicity using cross-correlation and double-difference methods, *J. Geophys. Res.*, *113*, B08,311, doi:10.1029/2007JB005479.

Weldon II, R. J., D. A. Schmidt, A. J. L., E. M. Weldon, and T. E. Dawson (2013), Appendix D – Compilation of Creep Rate Data for California Faults and Calculation of Moment Reduction due to Creep, *USGS Open-File Report 20131165, CGS Special Report 228, and Southern California Earthquake Center Publication 1792*, *228*, http://pubs.usgs.gov/of/2013/1165/.

Table S1. Units and values of the parameters of the multiscale method for the construction of the M_c-value maps shown in Figs. S1b,c.

Parameters		Units	Value
R_0	Length scale	km	12
T_0	Time scale	year	0.6
p	Magnitude scaling		0.4
δx	space step	km	2
δt	time step	year	0.2
N_c	Number of events		50
W_M	Magnitude range		1.0
Δm	Magnitude step		0.1

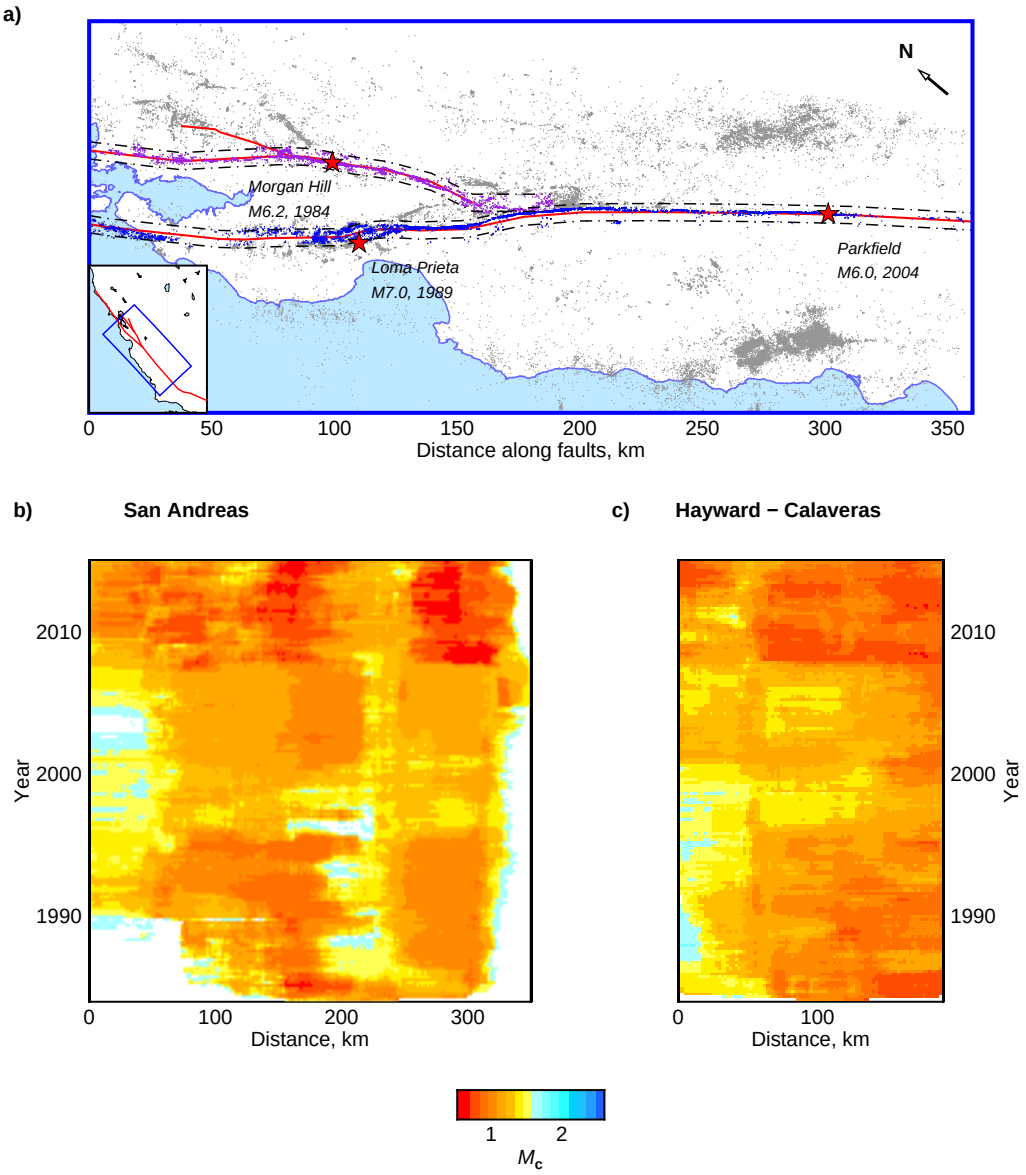

Figure S1. Variations of the completeness magnitude M_c in space and time along two fault zones in California. (a) Recorded seismicity along the San Andreas fault system (grey dots). Earthquakes along the San Andreas fault from the Peninsula to the Cholame-Carrizo sections and along the Hayward-Calaveras faults are shown in blue and purple, respectively. (b) M_c-value along the San Andreas fault. (c) M_c-value along the Hayward-Calaveras faults. Maps are constructed using the multiscale method of [*Vorobieva et al.*, 2013] (see text) and the NCEDC double-difference earthquake catalogue [*Waldhauser and Schaff*, 2008; *Waldhauser*, 2009].

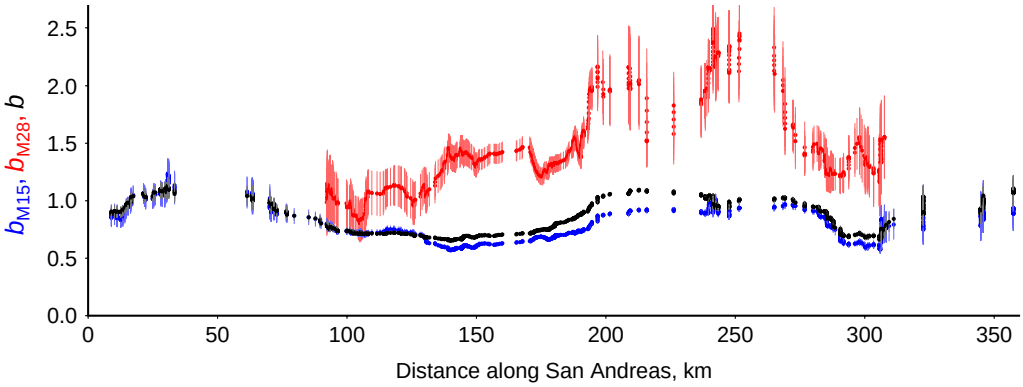

Figure S2. Dependence of b-value estimates on the large magnitude cut-off. The difference between b-values determined in the magnitude intervals [1.5, 2.7[(blue) and $[1.5, M_{\max}^{\text{obs}}]$ (black) reaches 0.2 in the segments where FMDs do not follow the GR law. Then, the FMD have significantly different slopes in the range of small (blue) and large magnitudes (red). The error bars are one standard deviation obtained from 500 bootstrap samples.

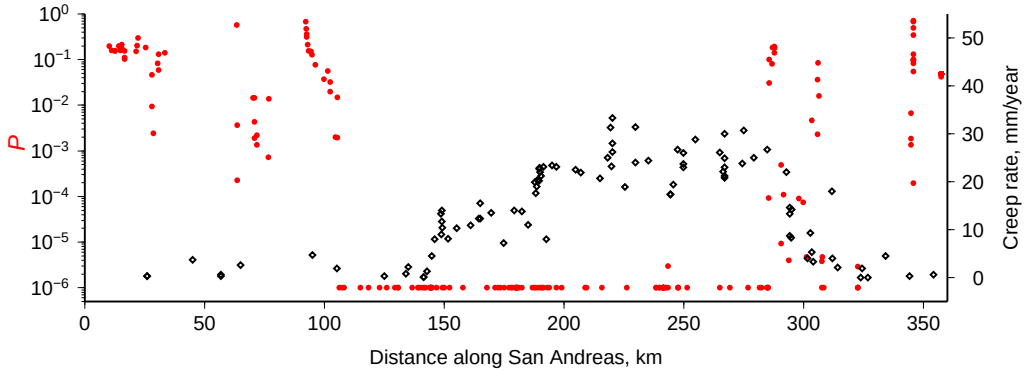

Figure S3. Pearson's χ^2 statistical test on the earthquake-size distribution along the San-Andreas fault. The null hypothesis H_0 is that the earthquake-size distribution follow the GR law (i.e., an exponential distribution). Using the GR law, red dots show the probability P to observe the X^2-value larger than the observed one (see Eq. 9). For drawing purposes the values of $P < 10^{-6}$ are shown at $P = 10^{-6}$. Black diamonds show the surface creep rate measurements [*Weldon II et al.*, 2013]. The null-hypothesis can be rejected with high confidence level along the creeping section of the San Andreas fault. Along the non-creeping segments the GR law cannot be rejected.

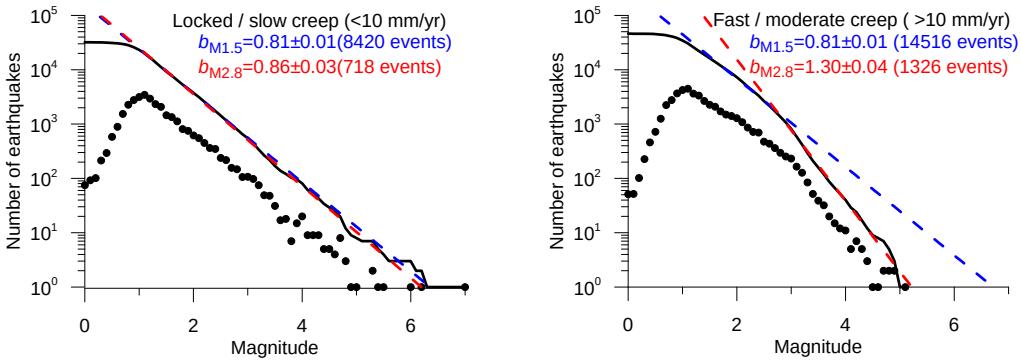

Figure S4. Earthquake size distributions in the locked/low creeping sections (left) and in the fast/moderate creeping sections (right) along the San Andreas, Calaveras and Hayward fault systems. The b-values are determined using a truncated GR law in the [1.5, 2.7] (blue) and [2.8, 4.0] (red) magnitude ranges; the number of events in each magnitude range is given in parenthesis. The difference between $b_{M1.5}$ and $b_{M2.8}$-values in the creeping section exceeds 10 times the accuracy (standard deviation) of b-value estimates.

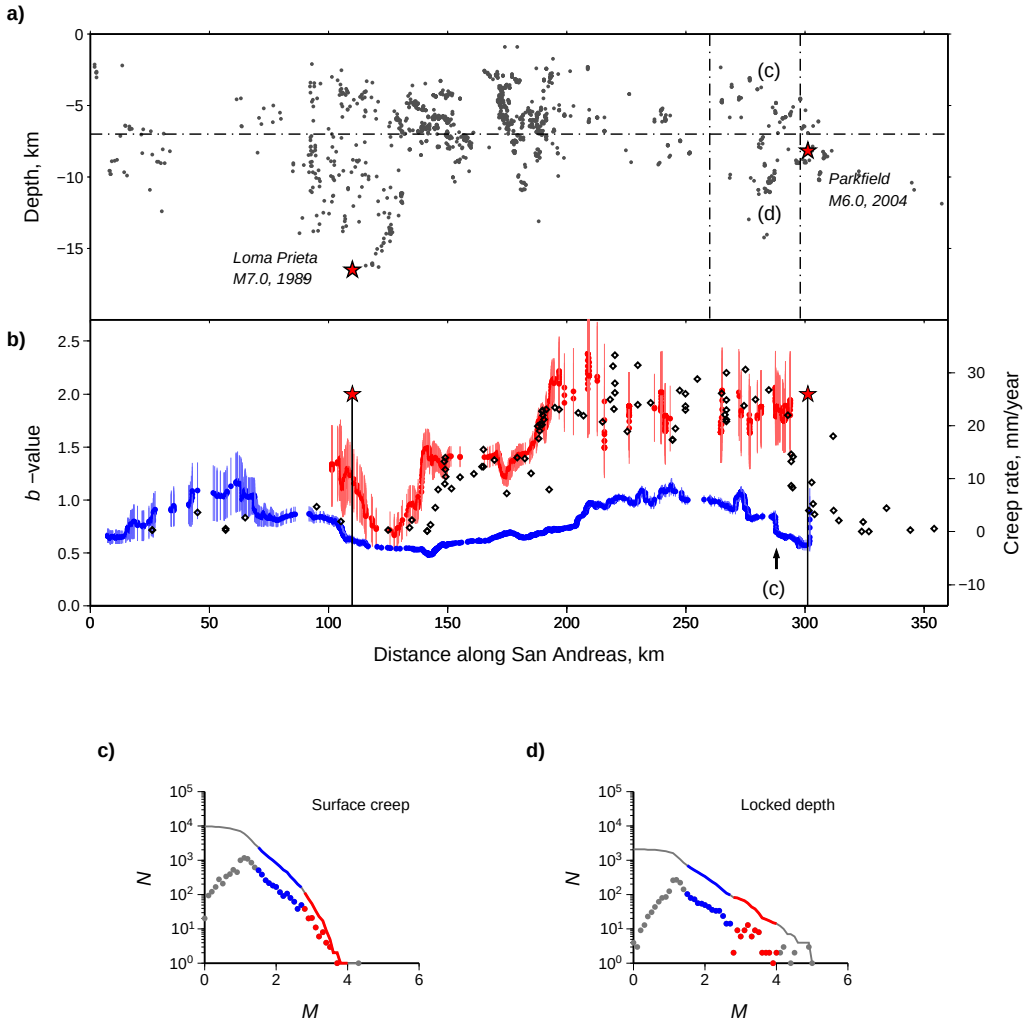

Figure S5. Creep rate and break of slope in earthquake-size distributions along the San Andreas fault for earthquakes with focal depth $h \leq 7$ km. (a) Distribution of $M \geq 2.8$ earthquake hypocenters close to the San-Andreas fault (see Fig. 1 of the main text). Vertical dashed lines limit the transition zone between fast creeping and locked segments. (b) Surface creep rate measurements (black diamonds) and b-values determined in two consecutive magnitude ranges: [1.5, 2.7] (blue dots), and [2.8, 4.0] (red dots) for earthquakes with the focal depth $h \leq 7$ km; b-values are assigned to the mean position of earthquakes in the current sliding-window. Errorbars are standard deviations obtained by bootstrap resampling. (c) FMD in the transition zone (vertical dashed lines) for earthquakes with a focal depth $h \leq 7$ km. (d) FMD in the transition zone for earthquakes with a focal depth $h > 7$ km.

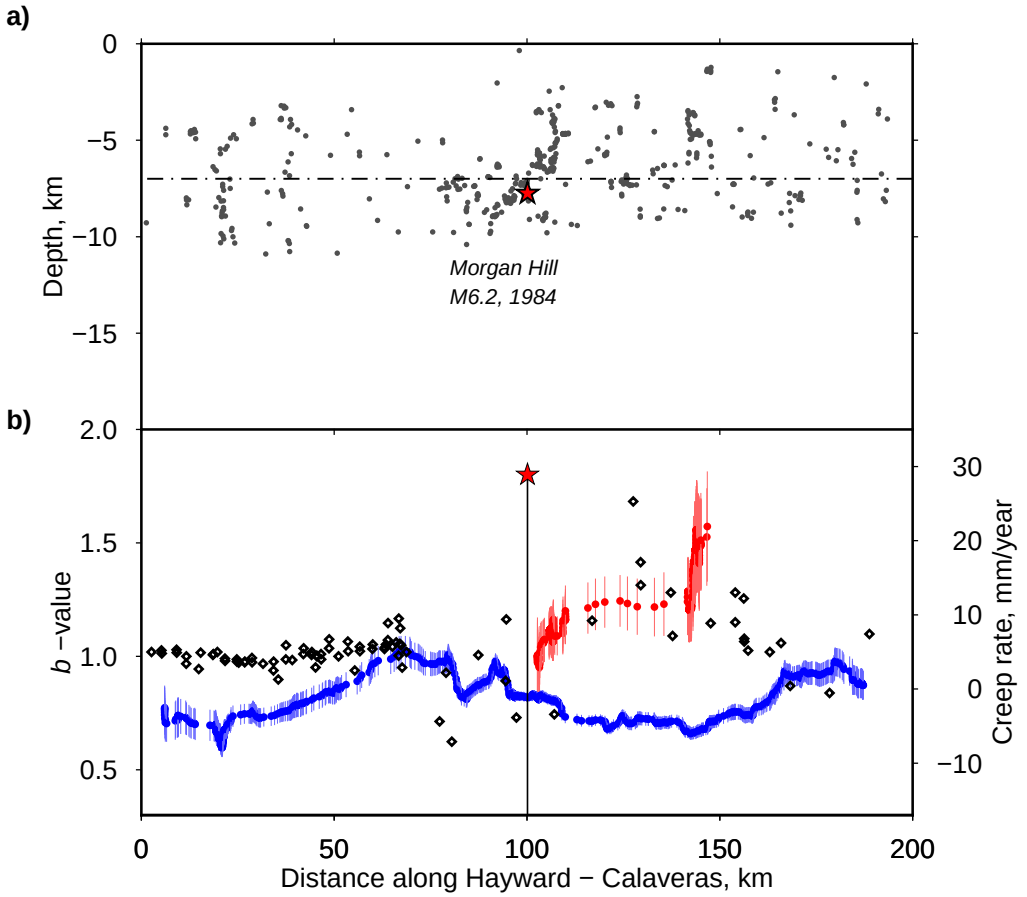

Figure S6. Creep rate and break of slope in earthquake-size distributions along the Hayward-Calaveras faults for earthquakes with focal depth $h \leq 7$ km. (a) Distribution of $M \geq 2.8$ earthquake hypocenters close to the Hayward-Calaveras faultf (see Fig. 2 of the main text). b) Surface creep rate measurements (black diamonds) and b-values in two magnitude intervals [1.5, 2.7] (blue dots) and [2.8, 4.0] (red dots); b-values are assigned to the mean position of earthquakes in the current sliding-window. Errorbars are standard deviations obtained by bootstrap resampling.

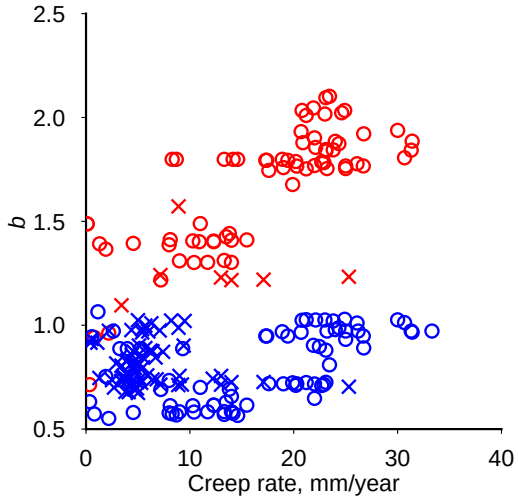

Figure S7. Dependence of the slope of earthquake-size distribution on creep rate for earthquakes with focal depth $h \leq 7$ km. b-values in two consecutive magnitude ranges, [1.5, 2.7] (blue) and [2.8, 4.0] (red), along the SAF (circles) and the Hayward-Calaveras faults (crosses).

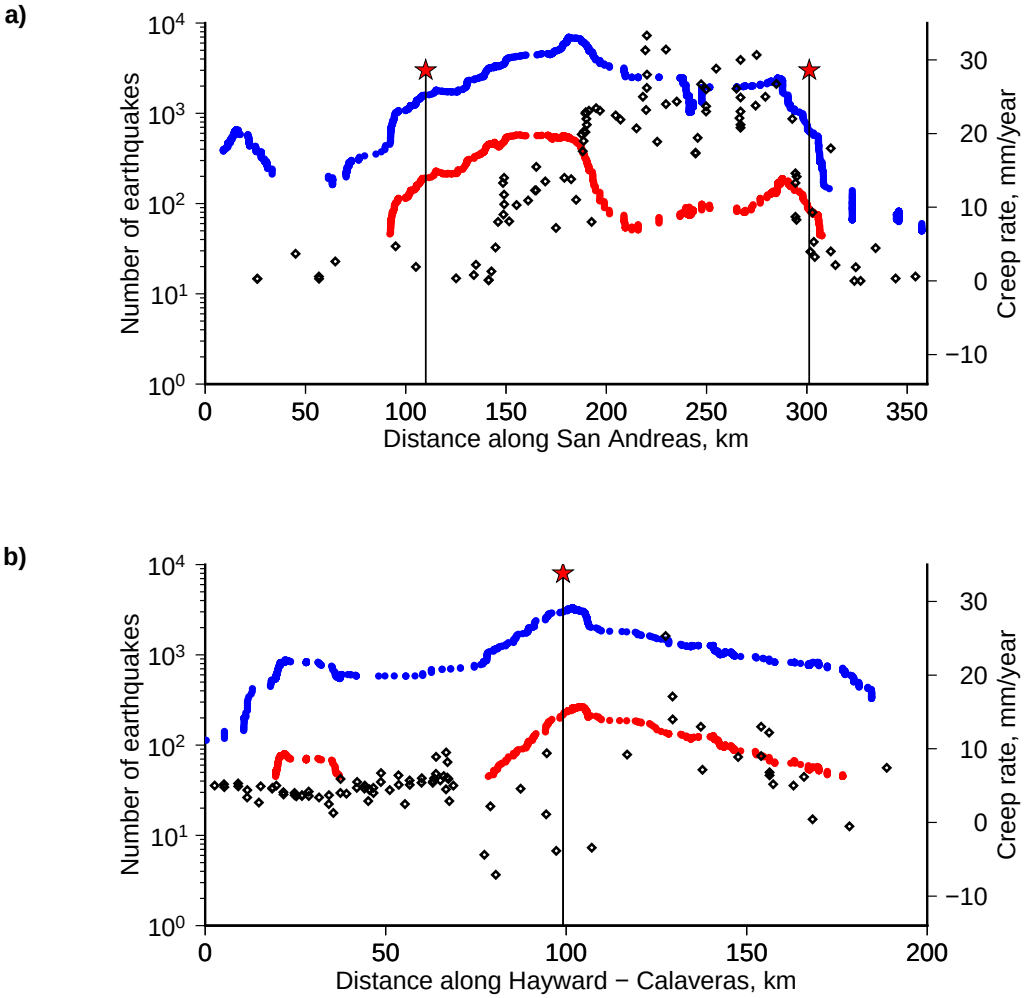

Figure S8. Variation of the seismic activity with the creep rate. (a) Number of earthquakes along the San Andreas faults in 40 km sliding segments in the magnitude intervals [1.5, 2.7] (blue dots), and [2.8, 4.0] (red dots). Black diamonds show surface creep rate measurements [*Weldon II et al.*, 2013]. Locations of Loma Prieta and Parkfield earthquakes are marked by vertical bars and red stars. b) Number of earthquakes along the Hayward-Calaveras faults in 40 km sliding segments in the magnitude intervals [1.5, 2.7] (blue dots), and [2.8, 4.0] (red dots). Black diamonds show surface creep rate measurements [*Weldon II et al.*, 2013]. Locations of the Morgan Hill earthquake is marked by vertical bar and red star.

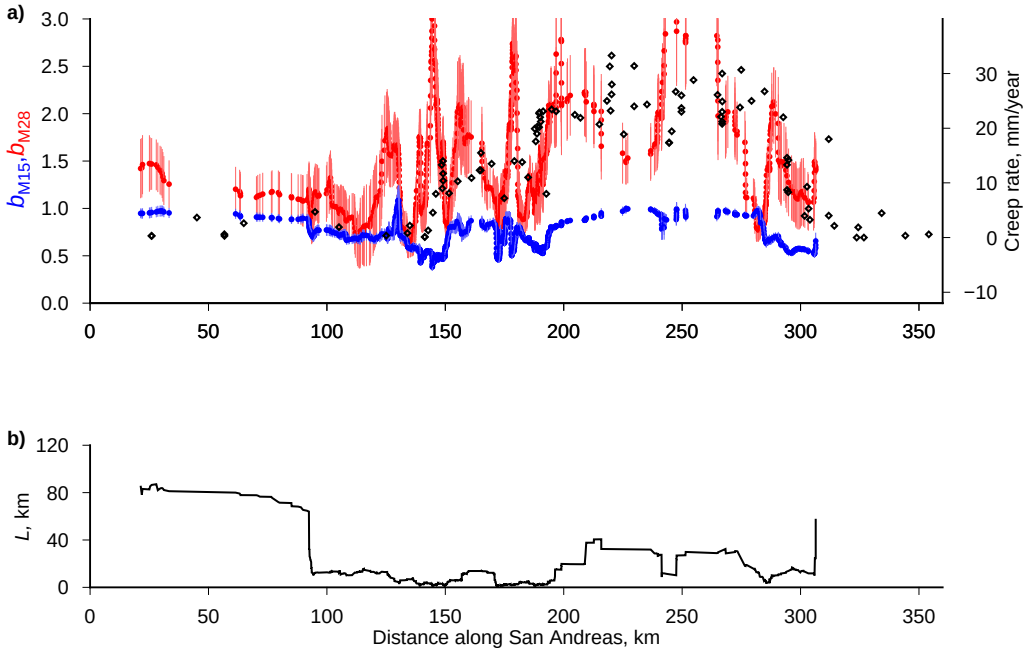

Figure S9. Creep rate and break of slope in earthquake-size distributions along the San Andreas fault using a constant sample size of 50 earthquakes. a) Surface creep rate measurements (black diamonds) and b-values in two [1.5, 2.7] (blue dots) and [2.8, 4.0] (red dots); b-values are assigned to the mean position of earthquakes in the current window. Errorbars are standard deviations obtained by bootstrap resampling. b) Variation of the length of segment containing 50 earthquakes in the magnitude interval [2.8, 4] along the San Andreas fault.

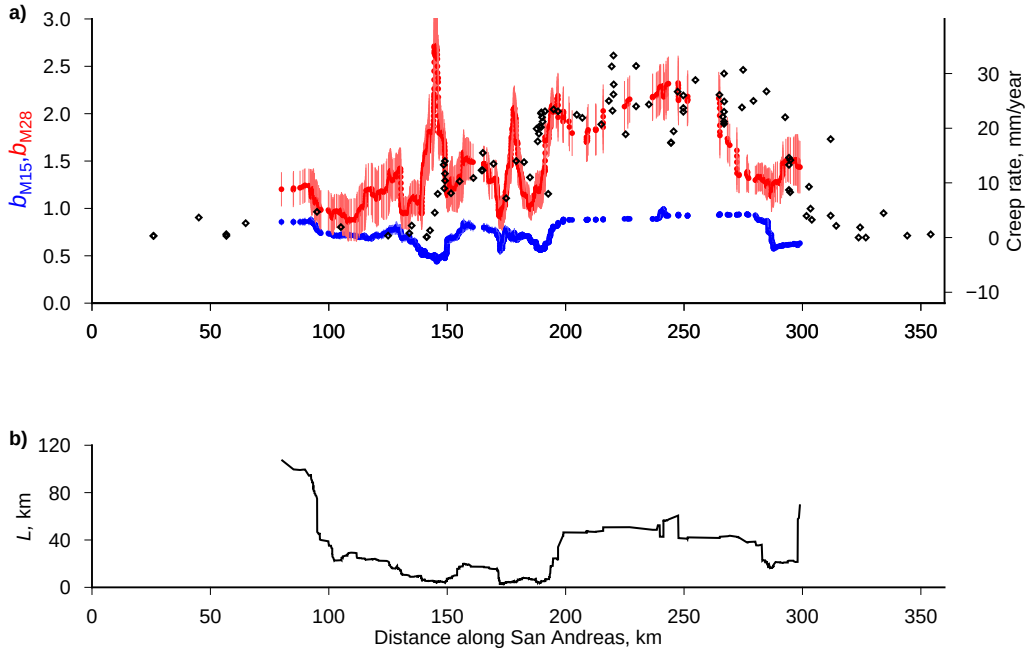

Figure S10. Creep rate and break of slope in earthquake-size distributions along the San Andreas fault using a constant sample size of 100 earthquakes. a) Surface creep rate measurements (black diamonds) and b-values in two [1.5, 2.7] (blue dots) and [2.8, 4.0] (red dots); b-values are assigned to the mean position of earthquakes in the current window. Errorbars are standard deviations obtained by bootstrap resampling. b) Variation of the length of segment containing 100 earthquakes in the magnitude interval [2.8, 4] along the San Andreas fault.

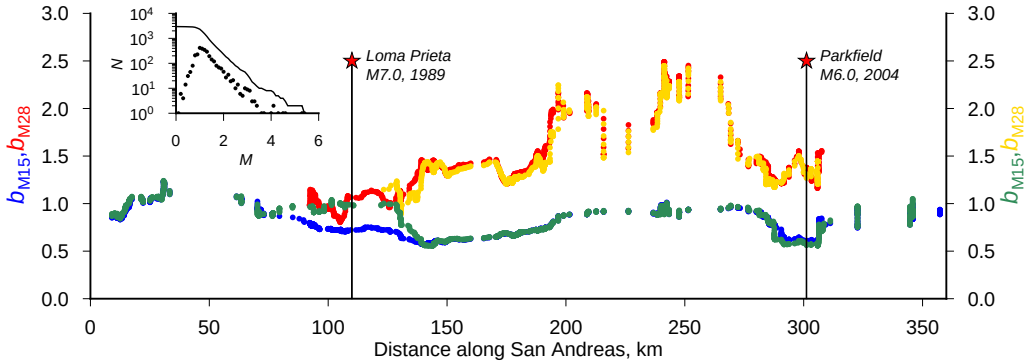

Figure S11. Influence of large earthquakes and their aftershocks on the slope of the earthquake-size distribution. b-values in two consecutive magnitude ranges, [1.5, 2.7] and [2.8, 4.0], are determined for the entire catalog (blue and red dots, respectively) and for a catalog in which we have excluded one-year of seismicity after the Loma Prieta and Parkfield earthquakes (green and yellow dots, respectively). Inset: the earthquake-size distribution in the Loma Prieta zone (80-130 km) without the first year of seismicity following the Loma Prieta and Parkfield earthquakes.