Seismic and Aseismic Moment Budget and Implication for the Seismic Potential of the Parkfield Segment of the San Andreas Fault.

Sylvain Michel¹, Jean-Philippe Avouac², Romain Jolivet³ and Lifeng Wang⁴

¹University of Cambridge, Department of Earth Sciences, Cambridge, UK

²California Institute of Technology, Department of Geology and Planetary Sciences, Pasadena, CA, USA

³Laboratoire de Géologie, Département de géosciences, École Normale Supérieure, CNRS, UMR 8538, PSL Research University, 75005 Paris, France

⁴State Key Laboratory of Earthquake Dynamics, Institute of Geology, China Earthquake Administration, Beijing, China

Sylvain Michel also at 2

Abstract

This study explores methods to assess the seismic potential of a fault based on geodetic measurements, geological information of fault slip rate and seismicity data. The methods are applied to the Parkfield section along the San Andreas Fault (SAF) at the transition zone between the SAF creeping segment in the North and the locked section of Cholame to the south, where $M_w \sim 6$ earthquakes occurred every 24.5 years on average since the M7.7 Fort Tejon earthquake of 1857. We compare the moment released by the known earthquakes and associated postseismic deformation with the moment deficit accumulated during the interseismic period derived from geodetic measurement of interseismic strain. We find that the recurring M6 earthquakes are insufficient to balance the slip budget. We discuss and evaluate various possible scenarios which might account for the residual moment deficit and implications of the possible magnitude and return

period of $M_w > 6$ earthquakes on that fault segment. The most likely explanation is that this fault segment hosts $M6.5$ to $M7.5$ earthquakes, with a return period of 140 to 300 years. Such events could happen as independent earthquakes in conjunction with ruptures of the Carrizo plain segment of the SAF. We show how the results from our analysis can be formally incorporated in probabilistic seismic hazard assessment assuming various magnitude-frequency distribution and renewal time models.

Introduction

Crustal deformation is mostly taken up by slip localized on a limited number of large faults. This paradigm holds in particular in California [Meade and Hager, 2005] where the San Andreas Fault and its peripheral faults form the main fault system. Since earthquakes represent increments of fault slip and that deformation of the upper crust is considered to be mostly seismic, these faults are also assumed to host the largest crustal earthquakes. The relationship between seismicity, faults and geodetic strain has long been conceptualized by the elastic rebound theory of Reid [1910] which states that, on the long term average, elastic strain accumulating around a fault should be balanced by elastic strain released during earthquakes. It is clear, however, that within the seismogenic depth range, slip can be either seismic or aseismic and that the slip rate on a fault and the partitioning of seismic and aseismic slip are the primary factors determining the seismic hazard associated with a particular fault [e.g., Avouac, 2015]. The long-term slip rate on a fault can be determined from geological and morphotectonic studies. Once this information is known, the partitioning of seismic and aseismic slip can in principle be derived from seismicity but would require catalogues long

51 enough to be representative of the long-term seismicity. Such catalogues are generally not
52 available. Another approach is based on the assumption that the partitioning of seismic and
53 aseismic slip is determined by spatial variations of fault frictional properties, assumed constant
54 with time. In that case, geodetic measurements of interseismic strain can be used to reveal
55 locked asperities, where friction is presumably rate-weakening, and estimate the accumulation
56 rate of moment deficit building up in the interseismic period (between major earthquakes). This
57 moment needs then to be balanced by the moment released by the large earthquakes and
58 transient aseismic slip. Such slip budget offers ways to estimate the most probable magnitude
59 and frequency of the larger earthquakes on a particular fault. This approach has been recently
60 applied to the Himalayan arc, the Sumatra subduction zone and the longitudinal valley fault in
61 Taiwan [Ader *et al.*, 2012; Thomas *et al.*, 2014; Stevens and Avouac, 2016].

62 Here, we test and refine this approach on the Parkfield segment of the San Andreas fault. This
63 segment lies at the transition zone between the locked segment of the San Andreas Fault (SAF)
64 to the south and the creeping zone in the north (Figure 1). It has experienced six $M_w \sim 6$
65 earthquakes since the $M_w \sim 7.7$ Fort Tejon event of 1857 [Sieh, 1978; Bakun *et al.*, 2005].
66 Owing to its quasi-periodic behavior and occurrence of the latest $M_w 6.0$ earthquake in 2004,
67 this section of the SAF has been intensively studied. These studies have yielded good
68 constraints on co- and postseismic deformation related to the 2004 earthquake [Langbein *et al.*,
69 2005, 2006; Johanson *et al.*, 2006; Murray and Langbein, 2006; Barbot *et al.*, 2009; Bruhat *et*
70 *al.*, 2011; Wang *et al.*, 2012, 2014], transient slow-slip events [Murray and Segall, 2005], and
71 interseismic loading [Murray *et al.*, 2001; Murray and Langbein, 2006; Johnson, 2013; Wang
72 *et al.*, 2014; Jolivet *et al.*, 2015; Tong *et al.*, 2015]. Several studies have noted that strain build-
73 up in the interseismic period does not seem to be balanced by the strain released by $M \sim 6$ events,
74 given their rate since 1857 [Segall and Harris, 1987; Murray and Segall, 2002; Murray and
75 Langbein, 2006; Toké and Arrowsmith, 2006]. This finding has implications for seismic hazard,

as it implies that $M \sim 6$ events should be more frequent, on the long-term average, than has been observed since 1857, or that occasional larger events should occur. The amount of data and the frequent $M_w \sim 6$ events makes the Parkfield segment of the SAF a particularly appropriate test case to assess the possibility of constraining seismic hazard based on a moment budget approach.

Hereafter, we first describe the methodology used in this study. We then apply it to the Parkfield segment of the SAF and test the sensitivity of our results to various assumptions about the contribution of aftershocks and postseismic deformation. We conclude that $M_w > 6$ earthquakes are required to close the budget and assess the impact of this result on seismic hazard.

Method

We base our approach on the assumption that the rate of moment deficit accumulating in the interseismic period is, on average over the long-term, equal to the rate of moment released by seismic and transient aseismic slip. Our objective is to derive a probabilistic estimate of the magnitude of the largest possible earthquake along a fault segment, together with an associated recurrence time for such an earthquake. To do so, we calculate the moment released by observed seismicity and afterslip, and divide by the duration of the catalogue to get the average moment release rate. We compare this estimate with the moment deficit rate derived from models of interseismic strain. If the moment deficit rate is larger than the observed moment release rate, the observed maximum magnitude earthquake might not be the most extreme event that can occur along the fault segment. We thus explore the space of magnitude and frequency of maximum magnitude earthquakes to find which events can balance the moment budget and be plausible considering the current statistical distribution of earthquakes. We account for

100 aftershocks, background seismicity and postseismic slip for each maximum magnitude
 101 earthquake tested. This method allows to assess seismic hazard considering uncertainties on
 102 the seismic and geodetic data and accounting for our understanding of the behavior of a fault
 103 segment. In the following, we detail the method. A flow chart describing the approach, step by
 104 step, is available in the supplementary (Table S.1). Table 1 lists the parameters used in this
 105 study.

106 The rate of moment deficit accumulation, $\dot{m}_0$ (in N.m.yr⁻¹), can be written

$$\dot{m}_0 = \int_{Fault} \mu D dA, \quad (1)$$

107 where μ and A are the shear modulus and the fault area, and D is the slip deficit rate. The slip
 108 deficit rate can be expressed as $D = V_{plate} * \chi$ where V_{plate} is the long-term plate rate and χ is
 109 the interseismic coupling. The interseismic coupling is defined as the ratio between the deficit
 110 of slip and the long-term slip of the fault and is given by

$$\chi = 1 - \frac{S}{V_{plate}}, \quad (2)$$

111 where S is the creep rate observed during the interseismic period. χ is 0 for a fault patch
 112 creeping at the long-term slip rate and 1 for a fully locked patch.

113 The amount of moment released seismically can be estimated from earthquake catalogues, for
 114 example an historical catalogue. The average total seismic moment released per year $\dot{m}_S$ is
 115 given by

$$\dot{m}_S = \frac{\sum_{j=1}^N m_{S,j}}{t_{hist}}, \quad (3)$$

116 where N is the total number of events in the catalogue, $m_{S,j}$ is the seismic moment of each
 117 earthquake j and t_{hist} is the time period covered by the catalogue. The observed seismicity
 118 might be seen as one particular realization of a stochastic process over a certain period of time.
 119 It might not be representative of the long-term average seismic moment rate if the period of

time covered by the data is short compared to the return period of the largest possible earthquake.

The moment released by the known seismicity most often does not balance the moment deficit due to interseismic coupling. Many causes can lead to a deficit of seismicity ($\dot{m}_0 > \dot{m}_S$): (1) the largest possible earthquake is not present in the catalogue because of its too short time span; (2) the largest possible earthquake is present in the catalogue but the duration of the catalogue is longer than the average return period of such an event; (3) the undetected seismicity contributes significantly to the moment budget; (4) transient aseismic slip such as afterslip or slow slip events contribute significantly to the moment budget; (5) a fraction of interseismic strain is anelastic and aseismic and is therefore not to be released seismically; (6) a large earthquake with its epicenter outside the study area may have extended into the area of interest and released a fraction of the moment deficit. In this case, the catalogue does not capture the event. In the context of this study, this could have happened during the M~7.7 1857 Fort Tejon mainshock or possibly as a foreshock. This earthquake ruptured the San Andreas Fault south of Parkfield (Figure 1) and might have ruptured the Parkfield segment as well [Sieh, 1978].

Alternatively, $\dot{m}_S$, the rate of moment released seismically, can exceed $\dot{m}_0$, the rate of moment build up: (1) the largest possible event is in the catalogue but the period of time covered by the catalogue is shorter than the average return period of such an event; (2) such events have occurred more frequently over this period of time than over the long term average; (3) interseismic strain is not stationary in time and the period covered by the geodetic data corresponds to a loading rate that is less than the average over the long term. In any case, the comparison between $\dot{m}_S$ and $\dot{m}_0$ provides information on the magnitude and average return period of the largest earthquake needed to balance the slip budget on the long term.

The next step consists in calculating the probability of a seismicity model to balance the moment budget and be consistent with the known seismicity. Key parameters of the seismicity model are the magnitude and return period of the largest earthquake.

The probability that the largest event is of magnitude M_{max} and has on average a return period of τ_{max} can be written as the product of two probabilities,

$$P(M_{max}, \tau_{max}) = P_{Budget}(M_{max}, \tau_{max}) * P_{Hist}(M_{max}, \tau_{max}) . \quad (4)$$

$P_{Budget}(M_{max}, \tau_{max})$ is the probability that an earthquake of magnitude M_{max} and its associated aftershocks and aseismic afterslip release a moment equal to the deficit of moment accumulated over the return period τ_{max} (i.e. the probability that an earthquake of magnitude M_{max} balances the budget). $P_{Hist}(M_{max}, \tau_{max})$ is the probability that an event of magnitude M_{max} and return period τ_{max} is the maximum possible earthquake based on the historical seismicity.

We calculate P_{Budget} based on the assumption that the maximum-magnitude earthquake is followed by aftershocks and that aseismic afterslip releases a moment proportional to the moment released seismically. For a given mainshock moment magnitude and recurrence time, we compare the moment released by the mainshock and postseismic relaxation (aftershocks and aseismic afterslip) to the estimated rate of moment deficit building up on the fault. Interseismic models of fault coupling derived using a Bayesian approach directly provide the Probability Density Function (PDF) of the rate of moment accumulation [Wang *et al.*, 2014; Jolivet *et al.*, 2015].

We assume that, on average over the long-term, seismicity follows the empirical Gutenberg-Richter (GR) law [Gutenberg and Richter, 1944]

$$\log_{10}(N_M) = a - b M, \quad (5)$$

where a and b are constants relating respectively to the total number of earthquakes recorded per year and the relative distribution between small and large earthquakes, and N_M is the cumulative number of earthquakes per year over magnitude M . We assume that the law applies to a catalogue of independent events (with aftershocks removed through ‘declustering’) as well as to dependent events (a catalogue including aftershocks) and that the b-value is the same in both cases (for a same given area). These assumptions are for examples consistent with the earthquake statistics observed in Southern California [Marsan and Lengline, 2008].

The number of events per year with magnitudes in the range $[M - \frac{\partial M}{2}, M + \frac{\partial M}{2}]$, is then

$$r = 10^{N_{M-\frac{\partial M}{2}}} - 10^{N_{M+\frac{\partial M}{2}}}. \quad (6)$$

The moment release rate of earthquakes with moment between 0 and m_a , over a period of time, should then converge (in the limit of infinite time) towards

$$\dot{m}_{Sa} = \int_0^{m_a} r \, dm. \quad (7)$$

We assume that earthquakes in the study area are bounded by a maximum event of moment m_{max} and magnitude M_{max} . The largest aftershock has often a magnitude of about 1 unit less than the magnitude of the mainshock [Båth, 1965] which might imply a bi-modal earthquake distribution. Assuming that the background seismicity does not reach magnitudes larger than the largest aftershock, the return period of the main event and that of the maximum aftershock is the same and the GR relationship runs through $M = M_{max} - 1$ and $\tau = \tau_{max}$.

However, the Båth law is not a physical law and is probably linked to a statistical finite size effect (the number of aftershocks above a certain magnitude is finite and this finite number determines the difference of magnitude between the mainshock and the largest aftershocks) [Helmstetter and Sornette, 2003]. It is possible that at the limit of infinite time, although each

single cluster of earthquakes could follow the Båth law, the total GR distribution would apply up to the maximum value of the distribution $M = M_{max}$ and $\tau = \tau_{max}$.

In the following, we test both cases and assume that these hypothesis bracket the contribution to the total seismic moment release of earthquakes smaller than the maximum earthquake.

Additionally, we assume that the moment released by aseismic afterslip is a proportion α of the moment released seismically. Assuming that the moment release rate of seismic and aseismic transient slip events balances the rate of moment deficit accumulation on the long run, we get

$$\dot{m}_0 = (\dot{m}_{sm} + \dot{m}_{sa}) (1 + \alpha), \quad (8)$$

where $\dot{m}_{sm}$ is the moment release rate of the largest earthquake (the subscript ‘S’ stands for seismic and ‘m’ for mainshock), and $\dot{m}_{sa}$ the moment release rate of aftershocks and background seismicity (the subscript ‘S’ stands for seismic and the subscript ‘a’ stands for aftershock).

With the assumptions listed above, it is possible to calculate the probability P_{Budget} of closing the moment budget for a given magnitude and return period of the maximum earthquake. Without taking the Båth law into account, this probability is highest along a straight line in the Gutenberg-Richter plot corresponding to the return period of the largest event [Molnar, 1979; Ader et al., 2012],

$$\tau_{max}(m_{max}) = \frac{1}{1-2b/3} \frac{(1+\alpha) m_{max}}{\dot{m}_0} \quad \text{if } b < 3/2, \quad (9)$$

where m_{max} is the moment released by the largest mainshock. With the Båth law, the contribution of smaller events become nearly negligible,

$$\tau_{max}(m_{max}) = \frac{10^{-\frac{3B}{2}}}{1-2b/3} \frac{(1+\alpha) m_{max}}{\dot{m}_0}, \quad (10)$$

with B the difference in magnitude between the largest event and its largest aftershock. The probability density distribution P_{Budget} depends on the uncertainties on $\dot{m}_0$.

To illustrate the procedure, let us consider a magnitude of 7 with a return time of τ_{max} . The mainshock would release a moment of $3.5 \cdot 10^{19}$ N.m. We add to this moment the contribution of aftershocks, background seismicity and aseismic afterslip. The resulting moment is divided by τ_{max} to estimate the average moment release rate for the events of magnitude 7 and return period τ_{max} . We compare this moment release rate to the rate of accumulation of moment deficit in the interseismic period calculated from interseismic slip rate models.

Without any other constraint, the maximum possible magnitude on a fault that is accumulating moment deficit at a rate $\dot{m}_0$ is actually unbounded: a small and frequent largest earthquake would balance the moment budget as well as a larger infrequent one. Seismicity observations can be used to tighten the space of possible solutions and incorporated through the calculation of $P_{Hist}(M_{max}, \tau_{max})$. This probability represents the probability of the largest earthquake having a magnitude M_{max} and a return period τ_{max} given the known seismicity. To calculate this probability, we consider two approaches:

- Approach 1: We assume that the magnitude and frequency of the maximum-magnitude earthquake falls on the GR law derived from a declustered catalog of local seismicity. This hypothesis is questionable as it is not proven that the GR law applies up to the largest possible event at a local scale;
- Approach 2: The possible magnitude and frequency of the maximum-magnitude earthquake must be consistent with the observed largest event over the observation period (it has to be larger than or equal to the known largest event, and the return period of the larger event cannot be significantly smaller than the observation period)

In both cases, a model of inter-event time distribution is required. In the results presented here, we consider that independent events follow a Poisson process. This is probably a reasonable assumption for events that would rupture only a fraction of the studied fault segment, as two successive events would generally not occur at the same location, so that rebuilding stresses might not be required. For larger events, it might be more appropriate to assume a renewal-time model. In the supplement, we show the results obtained using a Brownian Passage Time model [Matthews *et al.*, 2002; Field and Jordan, 2015], which are only marginally different from the ones obtained with the Poisson model.

In the case of a Poisson model, the probability of n events occurring during the time period t is

$$P_{poisson}(n, t, \tau) = \frac{(t/\tau)^n}{n!} e^{-t/\tau}, \quad (11)$$

where τ is the average recurrence time of the Poisson process.

Following the first approach, the probability of (M_{max}, τ_{max}) being the magnitude and return period of the largest event depends on the magnitude of the largest observed earthquake M_{hist} and on the return period $\tau_{GR}(M_{max})$ predicted by the GR law derived from a declustered catalog:

$$\text{if } M_{max} < M_{hist}, \quad P_{Hist}(M_{max}, \tau_{max}) = 0, \quad (12)$$

$$\begin{aligned} \text{if } M_{max} > M_{hist}, \quad P_{Hist}(M_{max}, \tau_{max}) &\propto P_{poisson}(n = 1, \tau_{max}, \tau_{GR}(M_{max})) = \\ &(\tau_{max}/\tau_{GR}(M_{max}))e^{-\tau_{max}/\tau_{GR}(M_{max})}. \end{aligned} \quad (13)$$

We use the sign $\propto$ to indicate proportionality, as the PDF is normalized.

Following the second approach, the probability of (M_{max}, τ_{max}) being the magnitude and return period of the largest event depends on the magnitude of the largest observed earthquake

M_{hist} and on time period covered by the catalog t_{hist} . It can be defined as the probability to have no earthquakes of magnitude over M_{hist} occurring during the time period of the catalog:

$$\text{if } M_{max} < M_{hist}, P_{Hist}(M_{max}, \tau_{max}) = 0, \quad (14)$$

$$\text{if } M_{max} > M_{hist}, P_{Hist}(M_{max}, \tau_{max}) \propto P_{poisson}(n = 0, \tau_{hist}, \tau_{max}) = e^{-\frac{t_{hist}}{\tau_{max}}}. \quad (15)$$

The probability drops rapidly to zero as τ_{max} becomes smaller than t_{hist} . It becomes uniform quickly as τ_{max} gets larger than t_{hist} .

Figure 2 shows a schematic representation of the probability $P_{Budget} * P_{Hist}$ in both cases. In the first case, the probability that an earthquake is the maximum possible earthquake in view of the observed seismicity and also closes the moment budget is highest at the intersection between the GR law (equation 4) and the line representing the return period of the largest earthquake required to close the moment budget (equation 9 or 10). In the second case, the probability that an earthquake is the maximum possible earthquake in view of the observed seismicity and also closes the moment budget is highest along the line representing the budget closure condition (equation 9 or 10). Any large magnitude ($M_{max} > M_{hist}$) and very infrequent maximum earthquake that closes the moment budget is considered acceptable as long as its return period is long compared to the observation period ($\tau_{max} > t_{hist}$).

In principle, we could also use phenomenological scaling laws, or physical constraints [Scholz, 1982] to limit the range of possible earthquake magnitude on a particular fault segment. For instance, we could impose the co-seismic stress drop to be between 0.1MPa and 100MPa, as generally observed [e.g., Kanamori and Brodsky, 2004]. This would constrain the maximum possible moment given the size of the locked area. We find such constraints to be too loose to

be useful and are therefore not included as an a priori in our analysis. We use them a posteriori to validate our assessment qualitatively.

When it comes to seismic hazard, the two methods should not yield much different outcomes if the hazard is calculated over a period of time t similar in duration to the earthquake catalog t_{hist} . They could however differ significantly for $t \gg t_{hist}$. To assess the impact of choosing between approaches 1 and 2, we calculate the probability $P_{Hazard}(M > M_{Test}, t)$ of an independent event over a period t exceeding a magnitude M_{Test} for different values of t . The probability P_{Hazard} of an independent event with magnitude $> M_{Test}$ during a period of time t is

$$P_{Hazard}(M > M_{Test}, t, \tau_M) = 1 - P_{poisson}(n = 0, t, \tau_M) = 1 - e^{-\frac{t}{\tau_M}}, \quad (16)$$

where τ_M is the average recurrence time of independent earthquakes with magnitude larger than M_{Test} which can be calculated from the Gutenberg-Richter law ($\tau_M = 10^{-(a-b M_{Test})}$, with a and b being determined from a ‘declustered’ catalog). The probability P_{Hazard} calculated this way does not account for aftershocks.

To construct Figure 3 we first choose a value of t . We then grid the magnitude-frequency space and test systematically all magnitudes of the maximum possible earthquake and all the return periods within a search area. For each sample tested, we represent the probability $P_{Hazard}(M > M_{Test}, t)$ with a curve shaded according to the probability $P_{Budget} * P_{Hist}$. We can thus apprehend visually the most likely earthquake scenarios that might happen during the time period t (Figure 3). A mean probability of all the possibilities tested, which we weight with the product $P_{Budget} * P_{Hist}$, can be obtained from equation

$$\bar{P}_{Hazard}(M > M_{Test}, t) = \frac{\sum_{k=1}^N P_{Hazard,k} * P_{Budget,k} * P_{Hist,k}}{\sum_{k=1}^N P_{Budget,k} * P_{Hist,k}}, \quad (17)$$

where N is the number of samples tested, k is the sample index, $P_{Hazard,k}$ is the probability of an independent event with magnitude $>M_{Test}$ during the time period t for the sample k , and $P_{Budget,k} * P_{Hist,k}$ is the probability of sample k given the seismicity observations and the moment budget closure condition.

To illustrate the procedure, let us first choose a sample of maximum magnitude, M_{max} , and its time recurrence, τ_{max} . We assume that this event belongs to a distribution of independent events that follows the GR law. Choosing $M_{Test} = 4$ (for example), we can calculate the probability of having an independent earthquake over magnitude 4 during a time period t , knowing that those events have a recurrence time given by the GR law (equation 16). Applying this to the full range of $M_{Test} \in [0, M_{max}]$, P_{Hazard} will be represented by a line in Figure 3. This P_{Hazard} is associated with a specific M_{max} and τ_{max} , and corresponds to a specific $P_{Budget}(M_{max}, \tau_{max}) * P_{Hist}(M_{max}, \tau_{max})$. If we test each possible M_{max} and τ_{max} , and plot their P_{Hazard} in the same representation, the shade of each line indicating $P_{Budget} * P_{Hist}$, we then obtain Figure 3. We could then calculate the average P_{Hazard} of all those lines, but it would not account for the probability to close the moment budget and be plausible considering the observed seismicity. We then weight each P_{Hazard} obtained for a particular choice of M_{max} and τ_{max} by its related $P_{Budget} * P_{Hist}$, and use this to calculate the weighted average $\bar{P}_{Hazard}$ (equation 17) represented by the white dashed line in Figure 3.

We sample the probabilities using the Hasting-Metropolis Monte Carlo Markov Chain (MCMC) procedure [Metropolis et al., 1953; Hasting, 1970]. We calculate for each sample k , with a given magnitude and frequency of the largest earthquake, the probability that it is realistic knowing the data $P_{Hist,k}$ and the probability that it closes the budget $P_{Budget,k}$. The final probability $P_{G,k}$ (i.e. the posteriori probability) of a given sample would thus be

$$P_{G,k} \propto U_M U_{Freq} P_{Budget,k} P_{Hist,k}, \quad (18)$$

where U_M and U_{Freq} are the uniform laws chosen for the MCMC sampling for the magnitude and frequency, respectively (i.e. the a priori probability).

Application to the Parkfield segment of the San Andreas Fault

Moment budget

Interseismic strain around the Parkfield segment of the San Andreas fault has been investigated in a number of recent studies [Johnson, 2013; Wang *et al.*, 2014; Jolivet *et al.*, 2015; Tong *et al.*, 2015]. Here we rely on the studies of Wang *et al.* [2014] and Jolivet *et al.* [2015], as they provide a probabilistic description of slip rate from Bayesian inversions which can be directly used as an input in our study. We consider the same fault geometry and focus on the same fault area as Wang *et al.* [2014] (Figure 4 and 5). To assess the moment budget, we use interseismic models from Wang *et al.* [2014] and Jolivet *et al.* [2015] which were both derived from inversion of interseismic displacements measured at the surface assuming a fault embedded in an elastic medium. Wang *et al.* [2014] assumed a homogeneous elastic half-space and Jolivet *et al.* [2015] considered depth variations of elastic moduli. We convert slip potency (the integral of slip over fault area) to moment according to the elastic structure assumed in these studies.

Wang *et al.* [2014] present three different Bayesian inversions of GPS data from 14 stations covering the period from 1999 to 2004 including the mainshock of 2004 and 5 days of postseismic relaxation. They use a 70km long, ~19 km deep fault subdivided into 180 patches, each 1.4 km long along strike and 1 km wide. A steady slip rate is assumed at depth greater than 19 km. We considered their favored model (referred to as MW hereafter), which was obtained from a joint inversion of the interseismic, co-seismic and 5-days postseismic data,

designed to maximize the consistency between the three slip distributions. The slip distributions are therefore complementing each other (Figure 4): the co-seismic rupture is restricted to the area that was locked during the interseismic period; the inversions assume that postseismic slip is driven by the co-seismic stress change and thus yields a ring of afterslip surrounding the co-seismic rupture. The steady slip rate on the deeper extension of the fault is estimated to 32.1 mm/yr in this model. The Bayesian framework provides uncertainties on all quantities determined from the inversion.

Jolivet *et al.* [2015] (MJ) also used a Bayesian approach and derived an interseismic model from GPS and InSAR data covering the period from 2006 to 2010. The modeled fault zone extends over 200 km from the Cholame plain to north of San Juan Baptista (SJB). The patch size varies depending on the resolution from 4 km at the surface to 25 km at 20 km depth. The long-term slip rate varies along strike. In the Parkfield area considered in this study, it varies from 31.1 mm/yr to 36.6 mm/yr. The model shows a gradual northward decrease of interseismic coupling from the locked zone in the southeast to the creeping zone in the northwest (Figure 4). This model shows a deficit of slip extending to depth greater than 19km, thus deeper than the seismogenic depth. Given the absence of seismicity at such depth and the fact that temporal variations of strain rates in this depth range are probably primarily due to viscoelastic relaxation as was observed following the 2004 earthquake [Bruhat *et al.*, 2011], we consider slip deficit accumulation only in the 0-19km depth range of our fault model.

Both models yield deep slip rates in agreement with the geological long term slip rate on the SAF which is estimated to 33.9 ± 2.9 mm/yr in the Carrizo plain south of the Parkfield segment [Sieh and Jahns, 1984]. These rates are also consistent with those ($34.9\text{--}36.0 \pm 0.5$ mm/yr) derived from elastic block modeling of regional tectonics [Meade and Hager, 2005; Tong *et al.*, 2014]. They are, however, larger than the local estimate of $26.2 +6.4/-4.3$ mm/yr of Toké *et al.* [Toké *et al.*, 2011]. Both models show high interseismic coupling in the area that ruptured

in 2004. They differ significantly in part because they use different interseismic observations but also because of different methods and a priori assumptions. In addition to enforcing consistency between the three phases of the earthquake cycle, the inversion used to derive MW was regularized via spatial smoothing. By contrast MJ was obtained without any constraint on the smoothness of the slip rate distribution nor on the relationship between interseismic and co-seismic slip. Note that the long-term slip-rate on the continuation of the fault at the depth is constant in MW but varies along strike in MJ. For MJ, fault patches are assigned the long-term slip rate of the deeper creeping patches. Those located astride different deep creeping patches are divided in sub-patches. The shear modulus used to convert slip to moment is 30GPa for MW and varies between 20.5GPa and 66.2GPa for MJ [Jolivet *et al.*, 2015].

The Bayesian approach used in Jolivet *et al.* [2015] and Wang *et al.* [2014] provides thousands of scenarios, which were tested against geodetic data. For both studies, we calculate the moment deficit rate of every scenario using equation 1 and 2 to derive the PDF of the moment deficit rate. The rate of accumulation of moment deficit for each model is indicated in Table 2 and the corresponding PDFs are shown in Figure 6.

We calculated and represented in Figure 7 the moment deficit rate on a completely locked 70 x 20 km section loaded at 32.1 mm/yr (full black line) and a section with the coupling pattern from the MW model (black dashed line). Based on interseismic model MW the moment deficit builds up at a rate of $6.90 \pm 0.64 \cdot 10^{17}$ N.m/yr. The PDF associated with this estimate is shown in Figure 6. Some of this deficit was released by aseismic afterslip following the 2004 event which released about $3.7 \cdot 10^{18}$ N.m Bruhat *et al.* [2011]. We do not use the afterslip model of Wang *et al.* [2014], as their model covers only 5 days after the mainshock. The dotted line in Figure 7 shows the remaining deficit of moment which amounts to $5.39 \cdot 10^{17}$ N.m/yr.

Based on interseismic model MJ, the moment deficit accumulates in the interseismic period at a rate as large as $8.82 \pm 1.10 \cdot 10^{17}$ N.m/yr (Figures 6). The difference with the rate obtained with MW is partly due to the different pattern of coupling and to the difference of deep creep rates (31.1 and 36.5 mm/yr for MJ and 32.1 mm/yr for MW). The rate is reduced to $7.31 \cdot 10^{17}$ N.m/yr if aseismic afterslip is subtracted (Figure 8).

The moment released by seismicity is estimated from the Northern California earthquake catalog [NCEDC, 2014] from 1984 to 2015 (up to 5/2/2015). We consider earthquakes located within 5 km of the fault (Figure 5). Figure 7 shows the cumulative moment released by earthquakes of magnitude under M taken directly from the catalog (gray dashed line). This estimate might not be representative of the interseismic cycle. The inset in Figure 5 shows minor temporal fluctuations except for the strong increase of seismicity associated to the 2004 aftershocks. The moment released by seismicity over the 31 years covered by the catalog, representative of the 2004 event (87%), its aftershocks (7%) and the background seismicity (6%), accounts for only 5% of the deficit of moment that has accumulated over this time period according to the interseismic model MW, or 7% if the postseismic moment released is taken into account (Figure 7). It represents an even lower fraction (4%) of the deficit of moment calculated based on the interseismic model MJ (5% if taking into account the aseismic moment released by afterslip; Figure 8). As mentioned above, the deficit could suggest that the return period of the maximum magnitude earthquake is over-estimated or that larger magnitude events are needed on this fault segment. We now consider the first possibility.

Six $M_w \sim 6$ earthquakes occurred at Parkfield between 1857 and 2004, yielding an average return period of 24.5 years. The catalog covers 31 years and includes the 2004 event and its aftershocks. Some corrections might thus be needed to represent the long-term average seismicity. Let us assume that the 2004 Parkfield earthquake is characteristic of the sequence of M_w 6 earthquakes that have occurred since 1857. We assume that such an event and its

associated aftershocks, return every 24.5 ± 9.2 year on average, where the uncertainty is the 1-sigma standard deviation of the time intervals. We consider the moment released by such an event to be equal to the mean moment estimated for 2004 from various publications [Johanson *et al.*, 2006; Langbein *et al.*, 2006; Liu *et al.*, 2006; Murray and Langbein, 2006; Barbot *et al.*, 2009; Bruhat *et al.*, 2011; NCEDC, 2014; Wang *et al.*, 2014]. We use the standard deviation of these estimates as an estimate of the uncertainty at the 67% confidence level. The average moment released is then $1.48 \cdot 10^{18} \pm 4.7 \cdot 10^{17}$ N.m (M_w 6.08).

We estimate the moment released by aftershocks by comparing the seismicity rate over the 2004-2008 period with the 1984-2004 period. We also assume that each event triggered as much aseismic afterslip as the 2004 event. According to Bruhat *et al.*, 2011, afterslip following the 2004 event released about $1.57 \cdot 10^{18}$ N.m assuming viscous flow under 19km depth, or about $3.7 \cdot 10^{18}$ N.m if viscous relaxation is excluded, which is more than twice the coseismic moment release. Hereafter we assume that afterslip released up to 200% of coseismic moment, which gives us an upper-bound on the total moment released by this sequence of earthquake. The moment released by afterslip is subtracted from the deficit of moment accumulating in the interseismic period (Figure 7 and 8). The values corresponding to this moment budget estimate are reported in Table 2 and Table 3, and illustrated in Figure 6, 7 and 8. Clearly, the moment released by 2004-like events returning every 24.5 ± 9.2 year on average falls still way short of balancing the moment budget (12.1% of MW's model).

Returning $M_w \sim 6$ earthquakes at Parkfield do not necessarily release the same moment as the 2004 event. The 1966 Parkfield earthquake is considered to be relatively similar to the 2004 one in terms of its moment release and rupture area [Bakun *et al.*, 2005]. According to the reported damages [Toppozada *et al.*, 2002], the events of 1901 and 1922 may have been stronger (M_w 6.4 and M_w 6.3 respectively). Moreover, several $M_w > 5$ earthquakes that occurred between 1857 and 2015 are considered to be independent from the $M_w \sim 6$ earthquake

($M5.5$ in 1877 and $M5.8$ in 1908), or direct aftershocks of the 1901 and 1922 earthquakes. Figures 7 and 8 show that if we now consider the historical catalog of Toppozada et al. [2002], the moment released by seismicity is higher but still too small to balance the moment budget (Table 3). Assuming that the interseismic period covered by geodetic data is representative of the long-term average, we derive that seismicity and afterslip released at most 16.1% (MW) of the deficit of moment accumulated since 1857. The epicenter of the 1901 $M6.4$ earthquake is actually located slightly north of our study area. The damage distribution indicates that the rupture propagated within the Parkfield area, but it is not sure that the rupture area was confined to the fault area consider in our study. By assuming that the moment of this earthquake was entirely released by slip on the fault segment considered in this study we probably tend to overestimate the seismic moment release.

Considering this time only for the moment deficit rate from the seismogenic zone (~ 15 km depth) and reducing the recurrence time of the $M6$ to 17 years as estimated by [Wang et al., 2015] does not close the moment budget either. The moment deficit rate is reduced to $3.86 \pm 0.52 \cdot 10^{17}$ N.m/yr based on the interseismic model MW cut at 15km depth with afterslip subtracted, and the seismic moment released is increased to $9.32 \cdot 10^{16}$ N.m/yr. Giving this extreme scenario, the percentage of seismic moment released compared to the moment deficit rate is still low (24%).

Some coseismic models of the 2004 $M6$ earthquake indicate slip up to 30km south of the hypocenter [Wang et al., 2014, 2015] and a similar rupture extent to that proposed for the 1966 earthquake [Murray and Langbein, 2006]. However, most models indicate that the 2004 rupture did not go further than 10 km south of the hypocenter [Johanson et al., 2006; Liu et al., 2006; Allmann and Shearer, 2007; Barbot et al., 2009; Bruhat et al., 2011]. By limiting the extent of our area of study to 10 km south of the 2004 hypocenter and by cutting the interseismic models at an even more shallower seismogenic zone (10km depth), we obtain a

moment deficit rate of $2.93 \cdot 10^{17}$ N.m/yr for MW and $2.99 \cdot 10^{17}$ N.m/yr for MJ. The moment deficit rate PDF of the model MJ would then slightly overlap with the moment released by earthquakes and postseismic slip. The probability to close the moment budget is nevertheless 2%. On the other hand, the MW model would still not overlap.

This analysis shows that balancing the moment budget on the Parkfield segment of the San Andreas fault probably requires more frequent or larger earthquakes than what the instrumental and historical data suggest. All the seismic moment release rate discussed in this section are available in Table 3.

Note that Tong *et al.* [2015] integral method yields a moment deficit rate around $5.8 \cdot 10^{17}$ N.m/yr, which is slightly smaller than the best fitting value of MW but within the uncertainty range. We consider that the range of rates of moment deficit explored by using the Bayesian inversions of MJ and MW models of interseismic coupling provides a reasonable estimate of the range of possible values given the uncertainties on the geodetic data and choice of the modeling technique.

Maximum-Magnitude Earthquake Evaluation

We now explore the range of possible magnitudes and frequencies for the largest possible earthquake needed to balance the moment budget. We first compute the probability of closing the moment budget for a given magnitude and frequency for the largest possible earthquake. We next calculate the probability of the possible earthquake models based on either approach 1 (seismicity follows the Gutenberg-Richter law up to the largest possible earthquake) or 2 (larger earthquakes can fall off the Gutenberg-Richter distribution indicated by instrumental and historical data).

We use the MCMC method to sample these probabilities. Each sample is taken from a uniform PDF between 6.4 and 9 for magnitudes, and between 10^{-6} and 10^4 yrs⁻¹ for frequencies. We calculate for each sample k the value proportional to the probability $P_{G,k}$ as described in the method section. For each estimation, we run the sampler for 900000 steps, rejecting the first 500 steps to ensure accurate sampling of the PDF. The magnitude-frequency space is then binned into cells of size 0.01 for the moment magnitude axis and logarithmically binned into cells of size 0.01 concerning the frequency axis. The number of samples in each cell divided by the total number of samples approximates the probability $P_{G,k}$ as defined by the MCMC method.

Moment Budget closure, P_{Budget}

We calculate the probability $P_{Budget,k}$ of closing the moment budget based on the uncertainties on the parameters of interseismic models MJ or MW. We estimate the moment released by seismicity for a tested magnitude M_k and recurrence time τ_k of the largest earthquake and estimate the probability that it balances the moment budget using interseismic models MJ or MW. The probabilities on the interseismic models are taken from the distribution of their moment deficit using a bin of 10^{16} N.m.

The moment released by aftershocks and independent events with magnitude smaller than the largest earthquake is integrated by assuming that their magnitude-frequency distribution follows the GR law. We consider two cases. In the first case, the aftershocks and independent events follow the GR law up to the largest earthquake included. In the second case, we account for Båth law; the largest aftershock is 1 unit of magnitude unit lower than the mainshock, as it is observed in the 2004 event. We suppose also that the b -value is constant and equal to the one estimated from the modified 1984-2015 catalog. We use the maximum curvature method

[Wiemer and Wyss, 2000] and Bender's formula [Aki, 1965; Utsu, 1966; Bender, 1983] to evaluate the magnitude of completeness and the b -value. The estimated magnitude of completeness is increased by 0.2 to compensate a methodological bias of the method [Woessner and Wiemer, 2005]. The b -value, estimated at 0.91 ± 0.06 , is then used to account for the magnitude-frequency distribution of both aftershocks and independent events. The a -value is easily calculated from equation (5).

Moment released by aseismic afterslip is accounted for by supposing that, for all earthquakes, it amounts to a fixed percentage of the seismic moment. This percentage might in fact depend on earthquake magnitude but the available data are insufficient to derive a reliable law [Lin *et al.*, 2013]. We considered a value of 25%, close to the mean value observed from various case studies [Avouac, 2015], and a more extreme value of 200%, similar to that estimated for the 2004 earthquake. The value of 200% is an overestimation of Bruhat *et al.*, [2011].

P_{Hist}: First Approach

In the Method section, we proposed two different approaches regarding how the frequency-magnitude of the largest possible earthquake might be evaluated based on the known seismicity. The first case, the most restrictive, assumes that the distribution of a declustered catalog follows the GR law up to the largest event. Some declustered catalogs are available for California [Duttilleul *et al.*, 2015] but they contain less than 200 earthquakes within 5 km from the fault segment considered in our study. This small number does not permit any reliable b -value estimation [Woessner and Wiemer, 2005]. Therefore, we will rely on a non-declustered catalog.

Assuming a postseismic moment release equivalent to 200% of the mainshock seismic moment, not accounting for Båth's law and selecting MW as the interseismic model represents

the most “optimistic” situation where the magnitude of the largest earthquake needed to close the moment budget is minimum. In this scenario, P_G (Equation 18) peaks around $\sim M7.6$, corresponding to a seismic moment of $2.8 \cdot 10^{20}$ N.m, with a recurrence time period of ~ 3300 years (Figure 9). The release of such a large moment on a 70km long segment seems however improbable, however, especially in view of the limited locked zone in the interseismic period. The average slip in such an event, supposing that the whole fault ruptures (20x70 km) and taking a 30GPa shear modulus, would be of ~ 7 m. If we assume a rupture restricted to the locked portion of this fault segment, which represents only 1/3 of the fault area, the average slip would need to be as great as 21m. This seems highly improbable.

We note in Figure 9 that the magnitude-frequency distribution of instrumental earthquakes does not align well with the historical seismicity. The earthquake rate is lower than expected if we assume the GR law extends up to larger magnitudes. The misalignment could be due to various causes. One is that the historical catalog essentially consists of independent events while the instrumental catalog contains aftershocks. The two catalogs could belong to the same Gutenberg-Richter distribution if the b -value of the declustered catalogs was lower than when the aftershocks are included as some studies suggest [Frohlich and Davis, 1993]. Extrapolating the Gutenberg-Richter law to the largest possible magnitude might be incorrect as a number of studies have argued that larger magnitude events are more frequent than would be expected from extrapolating the b -value estimated from lower magnitude seismicity [Schwartz and Coppersmith, 1984]. Another possibility is that the seismicity rate between 1984 and 2015 would have been particularly low compared to a longer term average [Page and Felzer, 2015]. In both case, the method used here to estimate P_{Hist} would overestimate the magnitude of the largest magnitude earthquake.

With the first approach, using the marginal cumulative probabilities, we can assess the probability that the largest event does not exceed a certain magnitude and the frequency of the

largest event is inferior to a certain value (Figure 10.a and c). The probability of closing the moment budget on the Parkfield segment of the San Andreas Fault with the magnitude of the maximum earthquake not exceeding 6.4 is only 0.3%. Therefore, provided that the GR law is valid up to magnitude $M_{6.4}$, which seems a reasonable assumption, the analysis requires larger earthquakes. Figure 10 indicates that there is a 95% chance that the maximum event does not exceed a $M_{8.63}$. There is also a 95% chance that the return period of the largest magnitude event does not exceed ~112000 years. It seems very improbable that the Parkfield segment could host an earthquake of such a magnitude as it would require 253m of slip in a 19x70km area. The probability of balancing the moment budget with a maximum earthquake not exceeding 7.4 is about 50%, indicating that even with such a large maximum earthquake, balancing the moment budget remains unlikely if the GR law applies. Note that if the largest earthquakes involve ruptures that extend beyond the Parkfield segment, our analysis only provides constraints on the moment released within the Parkfield segment. In that case, the assumption that this earthquake belongs to a GR distribution defined by earthquakes on the Parkfield segment is certainly incorrect and the second approach is then more relevant.

P_{Hist} : Second approach

Our second approach only assumes that the largest possible earthquake needs to be less frequent than the largest known earthquake. The probability of its return period is estimated based on the historical catalog (only the duration of the period covered matters) assuming a Poisson process. If we consider the interseismic model MW, discard the Båth law and assume that postseismic deformation releases 200% of the coseismic moment we get a more reasonable result than with approach 1 (Figure 11). A $M_{6.7}$ every 140 years is then sufficient to balance the moment budget and is acceptable in view of the known seismicity. This point aligns well with the frequency-magnitude distribution of historical events but is off the distribution of

instrumental events, essentially because of the break in slope around $M \sim 4$. This is our favored scenario. Note, however, that the method does not exclude $M > 6.7$ events.

Figure 12 shows that the PDF of the magnitude-frequency of the largest event varies depending on the choice of the interseismic model (MJ or MW), on whether or not Båth's law is accounted for, and on the assumption that either 25% or 200% of the mainshocks seismic moment is released by postseismic effects. The Båth law has a significant impact. If it is taken into account, aftershocks and independent events amount to only ~8% of the moment released by the largest earthquake compared to ~156% if the Båth law is discarded. Consequently, when the Båth law is not accounted for, the aftershocks and independent events sequence decrease drastically the maximum-magnitude needed for budget closure. Similarly, postseismic deformation amplifies the moment released by the largest earthquake, its aftershocks and the associated independent events. The higher the ratio between the postseismic and coseismic effect is, the lower the magnitude or the frequency of the largest event needed to balance the slip budget.

The estimate of the b-value could also impact the moment budget balance and the implication of the magnitude-frequency of the largest earthquake. By only taking the background seismicity time period (1983-2004 and 2008-2015), the largest event drops down to a $M4.9$ and the b-value is equal to ~0.90 which is quite close to the b-value estimate used above. To further investigate the b-value influence, we use our favored scenario: model MW, Båth law discarded and postseismic deformation equivalent to 200% of the coseismic moment release. This scenario should be the most sensitive to changes in b-value considering that there is no gap between the largest and next largest magnitude earthquakes and considering as well the high post-coseismic ratio. Applying the background b-value to our favored model, a $M6.71$ is then needed every 140 years to close the budget instead of a $M6.70$. The changes added by the Parkfield 2004 event and its aftershocks are therefore minimal. Additionally, if we raise our

initial b-value by its standard deviation ($\sigma \sim 0.06$), i.e. 0.97, the largest magnitude needed would drop down to 6.67 if it had to happen every 140 years. Changes in the b-value inside a reasonable range have thus a very limited impact on the maximum-magnitude event which turns to negligible if we assume Båth's law.

The two extreme cases illustrated by Figure 12.a and f enables to appreciate the range of magnitude and frequency for an earthquake to close the moment budget. For a 140-year recurrence time, the largest earthquake required would range between $M6.7$ for the first case and $M7.3$ for the second. This underlines the sensitivity to the hypothesis regarding the contribution to the moment budget of aftershocks and postseismic deformation. As an example, if we suppose that there is a magnitude 7 occurring every 140 years and we assume Båth's law, the postseismic moment released would be between 162 ± 34 % (MW) and 236 ± 59 % (MJ).

Following the second approach, the marginal cumulative probabilities reflect primarily the assumptions that the a priori probabilities are uniform, between 6.3 and 9 for the magnitude of the largest earthquake, and 10^4 and 10^{-6} for its frequency (Figure 10.b and d). It shows that the historical catalogue does not put any constrain on the magnitude of the maximum event for return period's much longer than the catalogue. The magnitude for which 95% of the maximum non-cumulative probability is reached can instead give us an indication of the magnitude and frequency needed to close the moment budget (Figure S2). A $M7.34$ is sufficient to close the moment budget at a 95% confidence level, as well as a time recurrence of 2000 years. Higher magnitudes and recurrence times than those estimates are equally plausible but would not be necessary.

Scaling laws linking stress drop, $\Delta\sigma$, area of rupture, A , and moment released during an earthquake, m , can give us an indication on the maximum magnitude through the equation $\Delta\sigma \approx C m A^{-3/2}$, where C is a geometric constant close to unity [Scholz, 1982]. Earthquake

stress drops generally vary between 0.1MPa and 100MPa [Kanamori and Brodsky, 2004]. Consequently, the rupture of the full fault (70 x 19km) would be equivalent to a maximum magnitude of between M6.39 and M8.38. For events that rupture half of the segment length and for which the down dip extent is 13km (35 x 13km), the maximum magnitude would vary between M5.92 and M7.92. For comparison, Parkfield 2004 stress drop is ~0.61MPa [Wang et al., 2014]. However, such range of stress drop and the corresponding magnitude is way too large to provide any constraints to our problem, hence we decided not to include this criterion in our analysis.

Implications for Seismic Hazard

For each tested value of the magnitude and frequency of the largest event, one can determine a seismic hazard quantity, such as for example, the probability P_{Hazard} that an independent earthquake (not an aftershock) would exceed a given magnitude M_{Test} over a given period of time t . Each curve can then be weighted according to the probability of the scenario. The weighted mean $\bar{P}_{Hazard}$ then represents the overall hazard resulting from all possible scenarios. As explained in the section of the first approach, the number of independent events taken from the declustered catalog is too low to estimate the local b-value. So, we use the b-value derived from a non-declustered catalog.

Figures 13.a and 13.b show the probability of having an earthquake larger than a given magnitude over a period of 30 years and 200 years based on the probability distribution of the magnitude-frequency of the largest earthquake derived with the second approach. The two panels correspond to different a priori assumptions regarding the maximum possible magnitude of the largest earthquake. If we assume that the largest earthquake cannot exceed 7.5, the probability of a $M>6$ event over 30 years is 43% or 6% for a $M>7$ (Figure 13.a). If we assume that the largest earthquake can be as large as magnitude 9, a certainly outrageous value, the

probability is reduced to 23% for a $M > 6$ and 3% for a $M > 7$ (Figure 13.b). The earthquakes with magnitudes between 7.5 and 9 with a high P_G contribute to the seismic moment release so that smaller magnitude earthquakes don't need to be as frequent as in the previous case, leading to a lower $\bar{P}_{Hazard}$. Similarly, Figures 13.c and 13.d show the hazard curves resulting from the probabilities of the magnitude-frequency of the largest earthquake derived with the second approach described above for a period of time of 200 years. If we assume that the largest earthquake cannot exceed 7.5, the probability of a $M > 6$ event over 200 years is 96%, and 30% for a $M > 7$ (Figure 13.c). If we assume that the largest earthquake can be as large as magnitude 9, the probability is reduced to 68% for $M > 6$ and to 15% for $M > 7$. These tests illustrate the mild sensitivity of the method to the a priori assumption on the maximum value of the largest possible magnitude.

The first approach, assuming a GR Richter distribution up to the largest event, yields very similar results (Figure S4) showing that these curves are not very sensitive to the choice of method, essentially because the periods chosen (of 30 and 200 years) are small compared with the return period of the maximum earthquake.

Discussion

Our assessment of the moment budget on the Parkfield segment of the SAF suggests that the sequence of $M \sim 6$ events returning every ~ 24 years as observed since 1857 accounts for only about 12.1% of the deficit of moment that is accumulating in the interseismic period. This conclusion would imply that this fault segment must hold larger, less frequent, earthquakes. Before discussing this possibility, we examine various factors that could be responsible for the moment deficit. The hypothesis of a larger earthquake is indeed not the only scenario that can

explain residual moment deficit. Our analysis relies on a number of assumptions that might be questioned.

First, we assume that interseismic coupling is stationary and use models of interseismic coupling derived from decades of geodetic data (GPS and InSAR). The time period covered by these data might seem in fact short if the moment deficit is primarily balanced by events with a return period on the order of centuries. Furthermore, significant changes in interseismic coupling have been reported on subduction megathrusts [Meltzner *et al.*, 2015; Tsang *et al.*, 2015; Yokota and Koketsu, 2015]. The stationary hypothesis is supported locally by the consistency between the model MJ proposed by Jolivet *et al.*, (2015) based on InSAR and GPS data from 2006 to 2010, and the model MW proposed by Wang *et al.*, (2014) based on GPS data ranges from 1999 to 2004. However, according to Barbot *et al.* [2013], the southern end of the Parkfield segment underwent some changes before and after the 2004 *M*₆ event. The boundary between the Parkfield and Cholame segments appears to creep faster following the 2004 earthquake than previously, which would indicate a variation in time of the coupling pattern. Moreover, except for an aseismic transient slip coincident with three *M*~4.5 earthquakes in the mid-1990s [Gwyther *et al.*, 1996; Langbein *et al.*, 1999; Nadeau and McEvilly, 1999; Gao *et al.*, 2000; Roeloffs, 2001; Murray and Segall, 2005], no strong changes in the interseismic slip rate seems to have been detected between 1966 and 2004 [Murray and Langbein, 2006; Barbot *et al.*, 2013]. This suggests no modification of the locked section pattern or presence of other creep events during this time. However, if an event equivalent to the one in the mid-1990s (~*M*_{5.6}) occurred once every 24.5 years (the period of the *M*₆ interseismic cycle), the moment released would be insufficient to close the moment budget of any of the models used in this study. It would cover 2.4% of the smallest residual moment deficit, calculated from the model of Wang *et al.*, (2014). Temporal variations of coupling and

slow events seem thus much too small to be able to account for the deficit of moment reported in our study.

Second, we assumed that all of the geodetic deformation is elastic strain. The overall consistency of geodetic strain rates with geologic slip rates supports this assumption for the San Andreas Fault. However, there are a number of segments where significant discrepancies have been reported. Such discrepancies are expected from stress redistribution among faults during the seismic cycle [Ben-Zion *et al.*, 1993; Loveless and Meade, 2011; Tong *et al.*, 2014] or could be due to distributed anelastic strain during the interseismic period [Oskin *et al.*, 2008]. In that regard, it is interesting to note that the deep creep rates that we used to estimate interseismic loading (31.1 and 36.5 mm/yr for MJ and 32.1 mm/yr for MW) are larger than the local estimate of $26.2 \pm 6.4/-4.3$ mm/yr of Toké *et al.* [Toké *et al.*, 2011]. This discrepancy could reflect either distributed anelastic strain or temporal variations in interseismic loading.

Our analysis thus suggests that the Parkfield segment of the SAF needs to release more seismic moment on the long-term average than was released by the $M \sim 6$ events that have occurred repeatedly since after the Fort Tejon earthquake of 1857. This imbalance points to occasional larger events. This fault segment could probably host significantly larger events. We note that the zone of high coupling derived from modeling interseismic geodetic data seems larger than the area that ruptured in 2004 (Figure 4a). If we assume an average stress drop similar to the 0.61 MPa observed during the 2004 earthquake [Wang *et al.*, 2014], which was at the lower end of commonly observed values, and a complete rupture of the locked area of Figure 4b (equal to 632 km^2 for $\chi > 0.7$), we could expect a $M \sim 6.3$ event. Assuming an average stress drop of 1.5 MPa, larger than in 2004, this fault segment could produce a $M 6.6$ earthquake. In fact, the locked area of Figure 4a, appears to be made of two locked patches separated by a zone of lower interseismic coupling. The partially creeping zone could act as a barrier to the propagation of seismic rupture. This pattern, analog to that investigated by Kaneko *et al.*, 2010,

is favorable to generate a bimodal seismic behavior with moderate earthquakes rupturing one or the other asperity and larger infrequent earthquakes rupturing the whole patch. Another possibility is that the Parkfield segment would occasionally rupture together with large events on the Carrizo plain segment of the SAF. In that case, the interaction between the two fault segments could lead to a larger moment release on the Parkfield segment than in the case of rupture confined to it (a longer rupture length requires larger slip if the stress drop is similar).

There are actually hints that the $M7.7$ 1857 Fort Tejon earthquake may have ruptured the Parkfield segment. Paleoseismic studies [Sieh, 1978; Lienkaemper, 2001] indicate 2 to 6 m of slip at sites near highway 46 due to this earthquake. The northernmost observation attributed to the 1857 earthquake is a 4.5 m offset, 1.7 km north of highway 46 [Lienkaemper, 2001], whereas a 3.5 m offset has also been measured 1.3 km south of it [Sieh, 1978]. Such large surface slip measurements show that the Parkfield segment can occasionally slip much more than during the 2004 earthquake (i.e. typically ~ 20 cm).

Assuming a 3.8m slip (as averaged over the measurements along the Cholame section [Sieh, 1978; Zielke *et al.*, 2012]) with a 10 km along dip width rupture that reaches 22 km North-west of the highway 46, the moment released within the limits of the model of Wang *et al* (2014) is of $4.38 \cdot 10^{19}$ N.m, equivalent to a $M7.03$ earthquake. The moment includes the aftershocks and postseismic effects since no distinction between them and the mainshock can be estimated from paleoseismic data. Assuming aftershocks contribute between 8% and 156% (depending on whether we assume Bath's law or not) and that the postseismic impact is about 200% of the coseismic released, then the seismic rupture alone would account between a $M6.72$ and $M6.47$.

In the literature, the recurrence time of 1857-like earthquakes, if there are any characteristic earthquakes [Zielke *et al.*, 2012], is suggested to be roughly between 140 and 255 years [Sieh, 1978; Zielke *et al.*, 2010]. In this case, the return of co-seismic ruptures similar to that of 1857

would contribute between $2.88 \cdot 10^{17}$ N.m/yr and $1.58 \cdot 10^{17}$ N.m/yr to the moment release rate.

This would reduce the moment deficit rate, i.e. between 42% and 23% for MW (between 61%

and 33% of the moment deficit residual rate), and 21% and 11% for MJ (between 24% and

13% of the moment deficit residual rate).

Finally, it has been proposed that enhanced fault weakening could occasionally allow large

ruptures to propagate in creeping areas [Noda and Lapusta, 2013]. In such case, the Parkfield

segment could also occasionally release much more seismic moment than during a $M \sim 6$ event

confined to the locked area of the Parkfield segment.

Conclusion

In principle, the rate of accumulation of moment deficit on a fault, determined from modeling

of interseismic strain measurements, can be used to constrain the frequency and magnitude of

the largest possible earthquake by requiring the moment budget to balance over the seismic

cycle. Testing this approach on the Parkfield segment shows that a $M \sim 6$ earthquake returning

about every 24 years, as has been observed over the last 150 years, falls short of balancing the

rate of moment deficit accumulation. The conclusion seems robust in view of the uncertainties

on the data and model assumptions.

The most likely explanation is that this fault segment hosts larger earthquakes, in the range

$M6.5$ to $M7.5$, with a return period of 140 to 300 years. Such episodes of seismic moment

release could happen as independent earthquakes rupturing the locked asperities revealed by

interseismic coupling models or in conjunction with rupture of the Carrizo plain segment of

the SAF, as potentially happened in 1857. The methodology presented in this study allows the

incorporation of formal geodetic constraints on the rate of moment deficit accumulation into

probabilistic seismic hazard assessments.

768

769 Data and Resources

770 The ANSS catalog was searched using [http://www.quake.geo.berkeley.edu/anss/catalog-](http://www.quake.geo.berkeley.edu/anss/catalog-search.html)
771 [search.html](http://www.quake.geo.berkeley.edu/anss/catalog-search.html) (last accessed February 2015)

772 Acknowledgment

773 We would like to thank Adriano Gualandi for insightful discussions and for providing the
774 recurrence time equation taking account Bath's law (Equation 10). We thank two anonymous
775 reviewers for their detailed and constructive criticisms which helped improve our study
776 substantially. We thank Alex Copley, James Jackson and Andy Howell for discussions. Some
777 of the figures were created using the Generic Mapping Tool (GMT) software [Wessel and
778 Smith, 1998].

779

780 References

- 781 Ader, T., J.-P. Avouac, J. Liu-Zeng, H. Lyon-Caen, L. Bollinger, J. Galetzka, J. Genrich, M.
782 Thomas, K. Chanard, S. Sapkota et al. (2012), Convergence rate across the Nepal
783 Himalaya and interseismic coupling on the Main Himalayan Thrust: Implications for
784 seismic hazard, *J. Geophys. Res.*, *117*(B04403), 1–16, doi:10.1029/2011JB009071.
- 785 Aki, K. (1965), Maximum Likelihood Estimate of b in the Formula $\log N = a - bM$ and its
786 Confidence Limits, *東京大学地震研究所彙報*, *43*(2), 237–239.
- 787 Allmann, B. P., and P. M. Shearer (2007), A High-Frequency Secondary Event, *Science*, *318*,
788 1279–1283, doi:10.1126/science.1146537.
- 789 Avouac, J.-P. (2015), From Geodetic Imaging of Seismic and Aseismic Fault Slip to

790 Dynamic Modeling of the Seismic Cycle, *Annu. Rev. Earth Planet. Sci.*, *43*, 233-271,
791 doi:10.1146/annurev-earth-060614-105302.

792 Bakun, W. H., B. Aagaard, B. Dost, W. L. Ellsworth, J. L. Hardebeck, R. A. Harris, C. Ji, M.
793 J. S. Johnston, J. Langbein, J. J. Lienkaemper, et al. (2005), Implications for prediction
794 and hazard assessment from the 2004 Parkfield earthquake, *Nature*, *437*(7061), 969–
795 974, doi:10.1038/nature04067.

796 Barbot, S., Y. Fialko, and Y. Bock (2009), Postseismic deformation due to the Mw 6.0 2004
797 Parkfield earthquake: Stress-driven creep on a fault with spatially variable rate-and-state
798 friction parameters, *J. Geophys. Res.*, *114*(B07405), 1–26, doi:10.1029/2008JB005748.

799 Barbot, S., P. Agram, and M. De Michele (2013), Change of apparent segmentation of the
800 San Andreas fault around Parkfield from space geodetic observations across multiple
801 periods, *J. Geophys. Res. Solid Earth*, *118*, 6311–6327, doi:10.1002/2013JB010442.

802 Båth, M. (1965), Lateral inhomogeneities of the upper mantle, *Tectonophysics*, *2*(6), 483–
803 514, doi:10.1016/0040-1951(65)90003-X.

804 Bender, B. (1983), Maximum likelihood estimation of b values for magnitude grouped data,
805 *Bull. Seismol. Soc. Am.*, *73*(3), 831–851.

806 Bruhat, L., S. Barbot, and J. P. Avouac (2011), Evidence for postseismic deformation of the
807 lower crust following the 2004 Mw6.0 Parkfield earthquake, *J. Geophys. Res.*,
808 *116*(B08401), 1–10, doi:10.1029/2010JB008073.

809 Duttilleul, P., C. W. Johnson, R. Bürgmann, Y. Wan, and Z. Shen (2015), Multifrequential
810 periodogram analysis of earthquake occurrence: An alternative approach to the Schuster
811 spectrum, with two examples in central California, *J. Geophys. Res. Solid Earth*, *120*,
812 8494–8515, doi:10.1002/2015JB012467.

813 Field, E. H., and T. H. Jordan (2015), Time-Dependent Renewal-Model Probabilities When
814 Date of Last Earthquake is Unknown, *Bull. Seismol. Soc. Am.*, *105*(1), 459–463,
815 doi:10.1785/0120140096.

816 Frohlich, C., and S. D. Davis (1993), Teleseismic b values; Or, much ado about 1.0, *J.*
817 *Geophys. Res.*, *98*(B1), 631, doi:10.1029/92JB01891.

818 Gao, S. S., P. G. Silver, and A. T. Linde (2000), Analysis of deformation data at Parkfield,
819 California: Detection of a long-term strain transient, *J. Geophys. Res.*, *105*(B2), 2955–
820 2967, doi:10.1029/1999JB900383.

821 Gutenberg, B., and C. F. Richter (1944), Frequence of earthquakes in California, *Bull.*
822 *Seismol. Soc. Am.*, *34*, 185–188.

823 Gwyther, R. L., M. T. Gladwin, M. Mee, and R. H. G. Hart (1996), Anomalous shear strain at
824 Parkfield during 1993-1994, *Geophys. Res. Lett.*, *23*(18), 2425–2428.

825 Hasting, W. K. (1970), Monte Carlo sampling methods using Markov chains and their
826 applications, *Biometrika*, *57*, 97.

827 Helmstetter, A., and D. Sornette (2003), Bath’s law derived from the Gutenberg-Richter law
828 and from aftershock properties, *Geophys. Res. Lett.*, *30*(20), 4,
829 doi:2069\r10.1029/2003gl018186.

830 Johanson, I. A., E. J. Fielding, F. Rolandone, and R. Bürgmann (2006), Coseismic and
831 postseismic slip of the 2004 Parkfield earthquake from space-geodetic data, *Bull.*
832 *Seismol. Soc. Am.*, *96*(4 B), 269–282, doi:10.1785/0120050818.

833 Johnson, K. M. (2013), Is stress accumulating on the creeping section of the San Andreas
834 fault?, *Geophys. Res. Lett.*, *40*(23), 6101–6105, doi:10.1002/2013GL058184.

835 Jolivet, R., M. Simons, P. S. Agram, Z. Duputel, and Z.-K. Shen (2015), Aseismic slip and

836 seismogenic coupling along the central San Andreas Fault, *Geophys. Res. Lett.*, *42*, 297–
837 306, doi:10.1002/2014GL062222.

838 Kanamori, H., and E. E. Brodsky (2004), The physics of earthquakes, *Reports Prog. Phys.*,
839 *67*, 1429–1496, doi:10.1088/0034-4885/67/8/R03.

840 Kaneko, Y., J.-P. Avouac, and N. Lapusta (2010), Towards inferring earthquake patterns
841 from geodetic observations of interseismic coupling, *Nat. Geosci.*, *3*(5), 363–369,
842 doi:10.1038/ngeo843.

843 Langbein, J., R. L. Gwyther, R. H. G. Hart, and M. T. Gladwyn (1999), Slip-rate increase at
844 Parkfield in 1993 detected by high-precision EDM and borehole tensor strainmeters,
845 *Geophys. J. Int.*, *26*(16), 2529–2532.

846 Langbein, J., R. Borchardt, D. Dreger, J. Fletcher, J. L. Hardebeck, M. Hellweg, C. Ji, M.
847 Johnston, J. R. Murray, and R. Nadeau (2005), Preliminary Report on the 28 September
848 2004, M 6.0 Parkfield, California Earthquake, *Seismol. Res. Lett.*, *76*(1), 10–26,
849 doi:10.1785/gssrl.76.1.10.

850 Langbein, J., J. R. Murray, and H. A. Snyder (2006), Coseismic and initial postseismic
851 deformation from the 2004 Parkfield, California, earthquake, observed by global
852 positioning system, electronic distance meter, creepmeters, and borehole strainmeters,
853 *Bull. Seismol. Soc. Am.*, *96*(4 B), 304–320, doi:10.1785/0120050823.

854 Lienkaemper, J. J. (2001), 1857 slip on the San Andreas fault Southeast of Cholame,
855 California, *Bull. Seismol. Soc. Am.*, *91*(6), 1659–1672, doi:10.1785/0120000043.

856 Lin, Y. N. N., A. Sladen, F. Ortega-Culaciati, M. Simons, J.-P. Avouac, E. J. Fielding, B. A.
857 Brooks, M. Bevis, J. Genrich, A. Rietbrock et al. (2013), Coseismic and postseismic slip
858 associated with the 2010 Maule Earthquake, Chile: Characterizing the Arauco Peninsula

859 barrier effect, *J. Geophys. Res. Solid Earth*, 118, doi:10.1002/jgrb.50207.

860 Liu, P., S. Custodio, and R. J. Archuleta (2006), Kinematic Inversion of the 2004 M 6.0
861 Parkfield Earthquake Including an Approximation to Site Effects, *Bull. Seismol. Soc.*
862 *Am.*, 96(4B), S143–S158, doi:10.1785/0120050826.

863 Marsan, D., and O. Lengline (2008), Extending Earthquakes’ Reach Through Cascading,
864 *Science*, 319(5866), 1076–1079, doi:10.1126/science.1148783.

865 Matthews, M. V, W. L. Ellsworth, and P. A. Reasenberg (2002), A Brownian Model for
866 Recurrent Earthquakes, *Bull. Seismol. Soc. Am.*, 92(6), 2233–2250.

867 Meade, B. J., and B. H. Hager (2005), Block models of crustal motion in southern California
868 constrained by GPS measurements, *J. Geophys. Res.*, 110(B03403),
869 doi:10.1029/2004JB003209.

870 Meltzner, A. J., K. Sieh, H.-W. Chiang, C.-C. Wu, L. L. H. Tsang, C.-C. Shen, E. M. Hill, B.
871 W. Suwargadi, D. H. Natawidjaja, B. Philibosian et al. (2015), Time-varying
872 interseismic strain rates and similar seismic ruptures on the Nias-Simeulue patch of the
873 Sunda megathrust, *Quat. Sci. Rev.*, 122, 258–281, doi:10.1016/j.quascirev.2015.06.003.

874 Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953),
875 Equation of State Calculations by Fast Computing Machines, *J. Chem. Phys.*, 21(6),
876 1087–1092, doi:doi:10.1063/1.1699114.

877 Molnar, P. (1979), Earthquake Recurrence Intervals and Plate Tectonics, *Bull. Seismol. Soc.*
878 *Am.*, 69(1), 115–133.

879 Murray, J., and J. Langbein (2006), Slip on the San Andreas fault at Parkfield, California,
880 over two earthquake cycles, and the implications for seismic hazard, *Bull. Seismol. Soc.*
881 *Am.*, 96(4 B), 283–303, doi:10.1785/0120050820.

882 Murray, J., and P. Segall (2002), Testing time-predictable earthquake recurrence by direct
883 measurement of strain accumulation and release, *Nature*, 419, 287–291,
884 doi:10.1038/nature00984.

885 Murray, J. R., and P. Segall (2005), Spatiotemporal evolution of a transient slip event on the
886 San Andreas fault near Parkfield, California, *J. Geophys. Res.*, 110(B09407),
887 doi:10.1029/2005JB003651.

888 Murray, J. R., P. Segall, P. Cervelli, W. Prescott, and J. Svarc (2001), Inversion of GPS data
889 for spatially variable slip-rate on the San Andreas Fault near Parkfield, CA, *Geophys.*
890 *Res. Lett.*, 28(2), 359–362, doi:10.1029/2000GL011933.

891 Nadeau, R. M., and T. V. McEvilly (1999), Fault Slip Rates at Depth from Recurrence
892 Intervals of Repeating Microearthquakes, *Science*, 285(5428), 718–721,
893 doi:10.1126/science.285.5428.718.

894 NCEDC (2014), *Northern California Earthquake Data Center*, Northern California
895 Earthquake Data Center. UC Berkeley Seismological Laboratory. Dataset.
896 doi:10.7932/NCEDC.

897 Noda, H., and N. Lapusta (2013), Stable creeping fault segments can become destructive as a
898 result of dynamic weakening., *Nature*, 493(7433), 518–521, doi:10.1038/nature11703.

899 Oskin, M., L. Perg, E. Shelef, M. Strane, E. Gurney, B. Singer, and X. Zhang (2008),
900 Elevated shear zone loading rate during an earthquake cluster in eastern California,
901 *Geology*, 36(6), 507–510, doi:10.1130/G24814A.1.

902 Page, M., and K. Felzer (2015), Southern San Andreas Fault Seismicity is Consistent with the
903 Gutenberg-Richter Magnitude-Frequency Distribution, *Bull. Seismol. Soc. Am.*, 105(4),
904 1–11, doi:10.1785/0120140340.

905 Reid, H. F. (1910), *The California Earthquake of April 18, 1906, Report of the State*
 906 *Earthquake Investigation Commission, The Mechanism of the Earthquake*, Washington,
 907 DC.

908 Roeloffs, E. A. (2001), Creep rate changes at Parkfield, California 1966–1999: Seasonal,
 909 precipitation induced, and tectonic, *J. Geophys. Res.*, *106*(B8), 16,525–16,547,
 910 doi:10.1029/2001JB000352.

911 Scholz, C. H. (1982), Scaling laws for large earthquakes: consequences for physical models,
 912 *Bull. Seism. Soc. Am*, *72*(1).

913 Schwartz, D. P., and K. J. Coppersmith (1984), Fault behavior and characteristic earthquakes:
 914 Examples from the Wasatch and San Andreas Fault Zones, *J. Geophys. Res.*, *89*(B7),
 915 5681–5698, doi:10.1029/JB089iB07p05681.

916 Segall, P., and R. Harris (1987), Earthquake Deformation Cycle on the San Andreas Fault
 917 Near Parkfield, California, *J. Geophys. Res.*, *92*(B10), 10511–10525,
 918 doi:10.1029/JB092iB10p10511.

919 Sieh, K. E. (1978), Slip along the San Andreas fault associated with the great 1857
 920 earthquake, *Bull. Seismol. Soc. Am.*, *68*(5), 1421–1448.

921 Sieh, K. E., and R. H. Jahns (1984), Holocene activity of the San Andreas fault at Wallace
 922 Creek, California Holocene activity of the San Andreas fault at Wallace Creek,
 923 California, *Geol. Soc. Am. Bull.*, (8), 883–896, doi:10.1130/0016-7606(1984)95<883.

924 Stevens, V. L., and J.-P. Avouac (2016), Millenary Mw >9.0 earthquakes required by
 925 geodetic strain in the Himalaya, *Geophys. Res. Lett.*, *43*, 1118–1123,
 926 doi:10.1002/2015GL067336.

927 Thomas, M. Y., J. P. Avouac, J. Champenois, J. C. Lee, and L. C. Kuo (2014),

928 Spatiotemporal evolution of seismic and aseismic slip on the Longitudinal Valley Fault,
929 Taiwan, *J. Geophys. Res. Solid Earth*, *119*, doi:10.1002/2013JB010603.

930 Toké, N. a., and J. R. Arrowsmith (2006), Reassessment of a slip budget along the Parkfield
931 segment of the San Andreas fault, *Bull. Seismol. Soc. Am.*, *96*(4 B), 339–348,
932 doi:10.1785/0120050829.

933 Toké, N. a., J. R. Arrowsmith, M. J. Rymer, A. Landgraf, D. E. Haddad, M. Busch, J. Cohan,
934 and A. Hannah (2011), Late Holocene slip rate of the San Andreas fault and its
935 accommodation by creep and moderate-magnitude earthquakes at Parkfield, California,
936 *Geology*, *39*(3), 243–246, doi:10.1130/G31498.1.

937 Tong, X., B. Smith-Konter, and D. T. Sandwell (2014), Is there a discrepancy between
938 geological and geodetic slip rates along the San Andreas Fault System?, *J. Geophys.*
939 *Res. Solid Earth*, *119*, 2518–2538, doi:10.1002/2013JB010765.Received.

940 Tong, X., D. T. Sandwell, and B. Smith-Konter (2015), An integral method to estimate the
941 moment accumulation rate on the Creeping Section of the San Andreas Fault, *Geophys.*
942 *J. Int.*, *203*(1), 48–62, doi:10.1093/gji/ggv269.

943 Topozada, T. R., D. M. Branum, M. S. Reichle, and C. L. Hallstrom (2002), San Andreas
944 Fault Zone , California : $M \geq 5.5$ Earthquake History, *Bull. Seismol. Soc. Am.*, *92*(7),
945 2555–2601.

946 Tsang, L. L. H., A. J. Meltzner, E. M. Hill, J. T. Freymueller, and K. Sieh (2015), A
947 paleogeodetic record of variable interseismic rates and megathrust coupling at Simeulue
948 Island , Sumatra, *Geophys. Res. Lett.*, *42*, 10585–10594, doi:10.1002/2015GL066366.1.

949 Utsu, T. (1966), A Statistical Significance Test of the Difference in b-value between Two
950 Earthquake Groups, *J. Phys. Earth*, *14*(2), 37–40, doi:10.4294/jpe1952.14.37.

951 Wang, L., S. Hainzl, G. Zöller, and M. Holschneider (2012), Stress- and aftershock-
 952 constrained joint inversions for coseismic and postseismic slip applied to the 2004 M6.0
 953 Parkfield earthquake, *J. Geophys. Res.*, *117*(B07406), doi:10.1029/2011JB009017.

954 Wang, L., S. Hainzl, and G. Zöller (2014), Assessment of stress coupling among the inter-,
 955 co-and post-seismic phases related to the 2004 M6 parkfield earthquake, *Geophys. J.*
 956 *Int.*, *197*(3), 1858–1868, doi:10.1093/gji/ggu102.

957 Wang, L., S. Hainzl, and P. M. Mai (2015), Quantifying slip balance in the earthquake cycle:
 958 Coseismic slip model constrained by interseismic coupling, *J. Geophys. Res. Solid*
 959 *Earth*, *120*, doi:10.1002/2015JB011987.

960 Wiemer, S., and M. Wyss (2000), Minimum magnitude of completeness in earthquake
 961 catalogs: Examples from Alaska, the Western United States, and Japan, *Bull. Seismol.*
 962 *Soc. Am.*, *90*(4), 859–869, doi:10.1785/0119990114.

963 Woessner, J., and S. Wiemer (2005), Assessing the quality of earthquake catalogues:
 964 Estimating the magnitude of completeness and its uncertainty, *Bull. Seismol. Soc. Am.*,
 965 *95*(2), 684–698, doi:10.1785/0120040007.

966 Yokota, Y., and K. Koketsu (2015), A very long-term transient event preceding the 2011
 967 Tohoku earthquake, *Nat. Commun.*, *6*, 5934, doi:10.1038/ncomms6934.

968 Zielke, O., J. R. Arrowsmith, L. G. Ludwig, and S. O. Akçiz (2010), Slip in the 1857 and
 969 Earlier Large Earthquakes Along the Carrizo Plain, San Andreas Fault, *Science*, *327*,
 970 1119–1122, doi:10.1126/science.1182781

971 Zielke, O., J. R. Arrowsmith, L. G. Ludwig, and S. O. Akciz (2012), High-resolution
 972 topography-derived offsets along the 1857 Fort Tejon earthquake rupture trace, San
 973 Andreas fault, *Bull. Seismol. Soc. Am.*, *102*(3), 1135–1154, doi:10.1785/0120110230.

974

975

976

977

978 **Sylvain MICHEL**

979 Bullard Laboratories, Department of Earth Sciences, University of Cambridge, Madingley

980 Road, Cambridge CB3 0EZ, UK

981 Sylvain_michel@live.fr

982

983 **Jean-Philippe AVOUAC**

984 Department of Geological and Planetary Science, California Institute of Technology, 1200 E.

985 California Boulevard, Pasadena, California 91125, USA

986 avouac@gps.caltech.edu

987

988 **Romain JOLIVET**

989 Laboratoire de Géologie, CNRS UMR 8538, Département de Géosciences, PSL Research

990 University, École Normale Supérieure, 24 rue Lhomond, 75005 Paris, France.

991 romain.jolivet@ens.fr

992

993 **Lifeng WANG**

994 State Key Laboratory of Earthquake Dynamics, Institute of Geology, China Earthquake

995 Administration, Beijing, China

996 wanglf@ies.ac.cn

997

998 **Table 1.** List of variables used in this study

<i>Variables</i>	<i>Symbol</i>	<i>Comments</i>
Moment (N.m)	m	$\dot{m}_0$: rate of moment deficit accumulation
		$\dot{m}_S$: average total moment release rate of seismicity
		$\dot{m}_{Sa}$: moment release rate of aftershocks and background seismicity
		$\dot{m}_{Sm}$: moment release rate of the largest earthquake
Magnitude	M	M_{max} : magnitude of the largest event
		M_{hist} : magnitude of the largest observed event (in catalogue)
		M_{Test} : magnitude tested for the probability to have an earthquake with $M > M_{Test}$ during a certain time period (P_{Hazard})
Shear Modulus (Pa)	μ	
Slip Deficit Rate (N.m.yr ⁻¹)	D	
Fault Area (m ²)	A	
Long Term Plate Rate (mm.yr ⁻¹)	V_{plate}	
Interseismic Coupling	χ	
Slip Rate in the Interseismic Period (mm.yr ⁻¹)	S	
Return Period of Events (yrs)	τ	τ_{max} : return period of the largest event
		τ_{GR} : return period predicted by the observed GR law for $M > M_{hist}$
		τ_M : average recurrence time of independent events with $M > M_{Test}$ predicted by a given GR law
Total Number of Earthquakes Recorded per Year (from the GR law)	a	
Relative Distribution between Small and Large Earthquakes (from the GR law)	b	
Cumulative Number of Earthquakes per Year over Magnitude M	N	
Number of Events per Year in a Magnitude Range	r	
Percentage of aseismic afterslip moment release compared to the moment released seismically	α	
Time Period (yrs)	t	t_{hist} : time period covered by the earthquake catalogue
Number of events during the time period t	n	

999

1000

1001

1002 **Table 2.** Parkfield moment deficit rate

<i>Interseismic models</i>	<i>Moment deficit rate (N.m/yr)</i>
MW [Wang <i>et al.</i> , 2014]	6.90 +/- 0.64 10 ¹⁷
MJ uncut [Jolivet <i>et al.</i> , 2015]	1.40 +/- 0.28 10 ¹⁸
MJ cut at 19km depth	8.82 +/- 1.10 10 ¹⁷
MW cut at 15km depth	6.04 +/- 0.52 10 ¹⁷

1003

1004

1005 **Table 3.** Parkfield seismic and aseismic moment release

<i>Catalog used</i>	<i>Seismic Moment rate (N.m/yr)</i>
Initial Catalog (31 years) [NCEDC, 2014]	3.70 10 ¹⁶
Modified catalog : M6.08+Aftershocks every 24.5 years	6.54 10 ¹⁶
Toppozada's catalogue [Toppozada <i>et al.</i> , 2002]	8.67 10 ¹⁶
Modified catalog : M6.08+Aftershocks every 17 years	9.32 10 ¹⁶

1006

Figure 1. Setting of the Parkfield segment of the San Andreas Fault. This segment lies at the transition between the $M7.7$ Fort Tejon earthquake rupture of 1857 (thin arrow along the SAF), to the South, and the creeping segment of the SAF to the North. The white stars indicate the latest epicenters of $M6+$ earthquakes in the region corresponding to the 2004 $M6$ Parkfield, 1983 $M6.3$ Coalinga and 2003 $M6.6$ San Simeon earthquakes.

Figure 2. Schematic representation of the two methods used in this study to estimate magnitude and frequency of the largest earthquake on the Parkfield segment of the San Andreas Fault. In both cases it is assumed that, over the long-term average, the moment released by earthquakes and aseismic afterslip balances the deficit of moment accumulating in the interseismic period due to partial fault locking. The black dash-dotted line represents the magnitude-frequency distribution of the largest earthquake according to that condition. The continuous grey line shows the magnitude-frequency distribution of observed earthquakes based on instrumental/and or historical

seismicity and after declustering. They are assumed to follow the Gutenberg-Richter law. The star is the maximum earthquake in the catalog and the dashed grey line shows extrapolation to larger magnitudes. (a) Here we assume that, over the long-term average, the Gutenberg-Richter law applies up to the largest possible earthquake. The most probable maximum-magnitude earthquake should then lie at the intersection between the moment budget closure line and the Gutenberg-Richter distribution of the declustered catalogue. (b) Here we only assume that the maximum earthquake exceeds the largest observed earthquake and that the probability of not observing this maximum earthquake over the historical period can be calculated assuming a Poisson law. Any large magnitude is possible but the frequency (hence the probability of occurrence over a given period) drop because of the condition of closure of the moment budget.

Figure 3. The probability of an earthquake exceeding a certain magnitude in a period of time of t years can be calculated based on our analysis. To do so we calculate for each possible truncated frequency-magnitude distribution the probability of exceeding a given magnitude and multiply this probability by the probability of the model based on $P_{Budget} * P_{Hist}$. The white dashed line represents $\bar{P}_{Hazard}$ from considering all possible scenarios.

Figure 4. Presentation of the interseismic models used in this study. (a) Interseismic model of Jolivet *et al.* (2015), referred to as model MJ in this study, which was derived from GPS and INSAR data between 2006 and 2010. (b) Interseismic model of Wang *et al.* (2014), referred to as model MW. This model was derived from geodetic data from 1999 to 2004. It was obtained from a joint inversion of interseismic, coseismic and postseismic slip with the constraints that co-seismic slip occurs in an interseismically locked area and that post-seismic slip (afterslip) is driven by co-seismic stress increase. The corresponding coseismic and afterslip models are shown in panels (c) and (d). The white star represents the hypocenter of the 2004 Parkfield earthquake.

Figure 5. (a) Map of the Parkfield segment of the San Andreas Fault. Dots show seismicity from the Northern California earthquake catalogue of the Northern California Earthquake Data Center [NCEDC, 2014] between 1984 and 2015. Our study is based on the earthquakes less than 5 km away from fault segment considered here (within the black rectangle). Epicenters of the six $M \sim 6$ earthquakes since 1857 are from Toppozada et al. [2002] (black and white stars). The town of Parkfield is also indicated by a white dot. (b) Parkfield seismicity rate from 1984 to 2015 for earthquakes with magnitude over 2. The vertical black dashed line corresponds to the 2004 $M6$ earthquake.

Figure 6. Probability Distribution Function of the rate of moment deficit accumulation on the Parkfield segment of the San Andreas Fault calculated based on interseismic models MW [Wang *et al.*, 2014] and MJ [Jolivet *et al.*, 2015]. The uncertainties are calculated based on the uncertainties of the interseismic slip models estimated in the original studies. Other sources of uncertainties (e.g. those related to the elastic properties of the subsurface) are ignored. The three histograms corresponds respectively, from left to right, to the model MW, MJ cut at 19km depth and MJ uncut, extending down to 40km. The vertical solid line is the seismic moment that would be released annually assuming that over the long term an event similar to the 2004 earthquake occurs every 24.5 years and is assigned a moment magnitude of $M6.08$, which represent some average value from various studies (see text for discussion). The vertical dashed line is the same but takes into account the postseismic moment released after a $M6$ mainshock similar to the 2004 event. The shaded patches associated with the vertical lines are the 1-sigma standard deviation from the seismic moment average of values taken from various studies.

Figure 7 (a) Representation of the budget of seismic and aseismic slip on the Parkfield segment of the SAF based on interseismic model MW [Wang *et al.*, 2014]. The gray dashed curve is the cumulated moment released per year by earthquakes with magnitude less than the value in abscissa. This curve is based on seismicity within 5km of the Parkfield fault segment over 31 years, from 1984 to 2015. The catalogue includes the 2004 earthquake, which is assigned a moment magnitude 5.97, and aftershocks. The black solid curve is the cumulated moment that would be released annually assuming that over the long term an event similar to the 2004 earthquake occurs every 24.5 years and is assigned a moment magnitude of $M6.08$, which represent some average value from various studies (see text for discussion). The white dot is the moment released assuming a $M6.08$ every 24.5yr and neglecting the contribution of aftershocks and background seismicity to the seismic moment release rate. The black dots show the moment released by historical seismicity from 1857 to 2015 according on the catalogue of Topozada *et al.* [2002]. The thin horizontal solid black line is the moment deficit accumulating per year if the modeled fault area was completely locked (for a loading rate of 32.1mm/yr). The thin horizontal dashed black line shows the moment deficit rate based on the interseismic model of MW. The thin horizontal dotted black line takes into account the postseismic moment released after a $M6$ mainshock similar to the 2004 event. The shading represent the uncertainty associated to this estimate. (b) Zoom of the black box which enables the appreciation of the large impact of the mainshock magnitude change and the weak impact of the background seismicity and aftershock's on the total amount of moment released by earthquakes.

Figure 8: Representation of the budget of seismic and aseismic slip on the Parkfield segment of the SAF based on interseismic model MJ [Jolivet *et al.*, 2015] with postseismic moment release taken into account. The solid black curve is the cumulated moment that would be released annually assuming that over the long term an event similar to the 2004 earthquake occurs every 24.5 years and is assigned a moment magnitude of $M_{6.08}$, which represent some average value from various studies (see text for discussion). The black dots show the moment released by historical seismicity from 1857 to 2015 according on the catalogue of Toppozada *et al.* [2002]. The dotted lines are the moment deficit accumulating per year based on the interseismic models (MW, MJ cut at 19km depth, MJ uncut, from bottom to top, respectively), and taking into account the postseismic moment released after a mainshock similar to the 2004 event.. The shading represent the uncertainty associated to the estimate of MJ cut at 19km depth.

Figure 9: Maximum-magnitude earthquake probability assuming the case where large earthquakes should follow the GR law of a declustered catalog. We use here the model MW, do not account for Bath's law, and suppose that the ratio between postseismic and coseismic moment release is equal to 200%. The number of events from the declustered ANSS catalog being too low in the area selected, we apply the method on a undeclustered ANSS catalog. The results are thus biased and must be interpreted accordingly. The black curve represents the magnitude-frequency distribution of the Parkfield area using the ANSS catalog between 1984 and 2015. The gray curve is the modified magnitude-frequency distribution where the $M6$ and its aftershocks are fixed to occur every 24.5 years. The GR law (black dashed line) is taken from this distribution. The stars represent the historical data [Toppozada *et al.*, 2002] and the black line represents a $M7$ earthquake with a recurrence time between 140 and 250.

Figure 10: Marginal cumulative probabilities for a maximum magnitude earthquake to close the moment budget and be plausible considering the data and the two approaches explained in the Method section. (a) and (c) are the marginal cumulative probabilities for the magnitude and frequency, respectively, for our first approach. It shows the probability of the largest earthquake to not exceed a certain magnitude (a) and the probability of the frequency of the largest event to not exceed a certain value (c). (b) and (d) are the equivalent of (a) and (c) for our second approach.

Figure 11: Maximum-magnitude earthquake probability assuming the case where large earthquakes should have a recurrence time lower than the largest earthquake currently observed. We use here the model MW, do not account for Bath's law, and suppose that the ratio between postseismic and coseismic moment release is equal to 200%. The black curve represents the magnitude-frequency distribution of the Parkfield area using the ANSS catalog between 1984 and 2015. The gray curve is the modified magnitude-frequency distribution where the $M6$ and its aftershocks are fixed to occur every 24.5 years. The stars represent the historical data [Toppozada *et al.*, 2002] and the black line represents a $M7$ earthquake with a recurrence time between 140 and 250. The black circle defines the location of a $M6.7$ event occurring every 140 years, our favored scenario.

Figure 12. Probability distribution to have a maximum-magnitude event that might exist and close the moment budget. The black curve represents the original magnitude-frequency distribution of the Parkfield area using the ANSS catalog between 1984 and 2015. The gray curve depicts the modified magnitude-frequency distribution of

earthquakes in the Parkfield region with a $M6.08$ and its aftershocks every 24.5 years. The stars represent the historical data [Toppozada *et al.*, 2002] and the black line represents a $M7$ earthquake with a recurrence time between 140 and 250 years. The probabilities computed depend on the model chosen, either MW or MJ, on the ratio between the postseismic and the coseismic moment release, and on whether the Bâth law is applied or not.

Figure 13. Probability to have earthquakes over a certain magnitude in a period of time of t years considering the probability distribution of a maximum-magnitude earthquake to exist and close the moment budget. We test our favored scenario using model MW, with the Bath law not accounted for and with a postseismic moment release equivalent to 200% of the coseismic one. The white dashed line corresponds to $\bar{P}_{\text{Hazard}}$. (a) and (b) are the results for $t=30$ years, and (c) and (d) are for $t=200$ years. (a) and (c) suppose a maximum-magnitude possible of $M7.5$, and (b) and (d) suppose that the maximum-magnitude possible is a $M9$.

Figure 1

[Click here to download Figure fig 1.jpg](#)

Figure 2

[Click here to download Figure fig 2.jpg](#)

Figure 3

[Click here to download Figure fig 3.jpg](#)

Figure 4

[Click here to download Figure fig 4.jpg](#)

(a)

Figure 6

Figure 7

[Click here to download Figure Fig 7_new.jpg](#)

Figure 8

Figure 9

[Click here to download Figure Fig 9.jpg](#)

Figure 11

Figure 12

[Click here to download Figure Fig 12_new.jpg](#)

Figure 13

[Click here to download Figure fig 13.jpg](#)

Seismic and Aseismic Moment Budget and Implication for the Seismic Potential of the Parkfield Segment of the San Andreas Fault.

Sylvain Michel, Jean-Philippe Avouac, Romain Jolivet and Lifeng Wang

Description of the supplement

We present in the supplement the flowchart of the methodology and we show the results of our methodology obtained using a Brownian Passage Time model [*Matthews et al.*, 2002; *Field and Jordan*, 2015], which are only marginally different from the ones obtained with the Poisson model.

List of Table Captions

Table S1: Flow chart of methodology.

List of Figure Captions

Figure S1: Maximum-magnitude earthquake probability assuming the case where large earthquakes should have a recurrence time lower than the largest earthquake currently observed, using the BPT model. We use here the model MW, do not account for Bath's law, and suppose that the ratio between postseismic and coseismic moment release is equal to 200%. The black curve represents the magnitude-frequency distribution of Parkfield's area using the ANSS catalog between 1984 and 2015. The gray curve is the modified magnitude-frequency distribution where the M6 and its aftershocks are fixed to occur every 24.5 years. The stars incarnate the historical data [Toppozada et al., 2002] and the black line represents a M7 earthquake with a recurrence time between 140 and 250. The black circle defines the location of a M6.7 event occurring every 140 years, our favored scenario.

Figure S2: (a) and (b) are the marginal non-cumulative probability normalized by their maximum value for the magnitude and frequency, respectively, for our second approach. They give an indication about the magnitude and

frequency of the largest event needed to close the moment budget, using the BPT (gray line) and Poisson model (black dotted line).

Figure S3: Probability to have earthquakes over of a certain magnitude in a period of time of $t=30$ years considering the probability distribution of a maximum-magnitude earthquake to exist and close the moment budget for the BPT model. We test our favored scenario using model MW, with the Bath law not accounted for and with a postseismic moment release equivalent to 200% of the coseismic one. The white dashed line corresponds to $\bar{P}_{Hazard}$ for the BPT model and the black dashed line corresponds to $\bar{P}_{Hazard}$ for the Poisson model.

Figure S4: Probability to have earthquakes over a certain magnitude in a period of time of t years considering the probability distribution of a maximum-magnitude earthquake to exist and close the moment budget for approach 1. We test our favored scenario using model MW, with the Bath law not accounted for and with a postseismic moment release equivalent to 200% of the coseismic one. The white dashed line corresponds to $\bar{P}_{Hazard}$ for approach1, and the black dashed line is for approach2. (a) and (b) are the results for $t=30$ years, and (c) and (d) are for $t=200$ years. (a) and (c) suppose a maximum-magnitude possible of M7.5, and (b) and (d) suppose that the maximum-magnitude possible is a M9.

In the methodology section of the main text, we base all our probability calculations on Poisson's model. However, other options are possible and we show here an example using a specific renewal model, the Brownian Passage Time (BPT) model [Matthews *et al.*, 2002; Field and Jordan, 2015]. The probability density function (PDF) of a renewal model, $f(t)$, describes the probability to have the first earthquake at time t since the last event. The PDF of the BPT model is

$$f(t; \tau, \alpha_{BPT}) = \left(\frac{\tau}{2\pi\alpha_{BPT}^2 t^3} \right)^{1/2} \exp \left\{ -\frac{(t - \tau)^2}{2\tau\alpha_{BPT}^2 t} \right\}, \quad (S1)$$

where τ is the earthquake mean recurrence time and α_{BPT} is defined by the ratio between the mean and the standard deviation of the recurrence time, σ .

Let $F(t)$ represent the cumulative distribution function (CDF) of $f(t)$ and T the time since the last earthquake. The probability of having no earthquakes in a time interval ΔT knowing that no earthquakes happened until T can be rewritten as the probability to have no earthquake until $T + \Delta T$, knowing there has been no earthquakes up to T :

$$\begin{aligned} P(T + \Delta T < t \mid T < t) &= \frac{P(T + \Delta T < t)P(T < t \mid T + \Delta T < t)}{P(T < t)} \\ &= \frac{P(T + \Delta T < t)}{P(T < t)} \\ &= \frac{1 - P(T + \Delta T \geq t)}{1 - P(T \geq t)} \\ &= \frac{1 - F(T + \Delta T)}{1 - F(T)}. \end{aligned} \quad (S2)$$

Note that compared to the Poisson model, which just needed τ and ΔT to be set, the BPT model needs additionally two more variable: α_{BPT} and T . In our calculation, we fix the α_{BPT} of all events equal to the α_{BPT} of Parkfield's $M6$ ($\alpha_{BPT}=0.37$). The PDF of the time of the last event for renewal

models is given by: $p(T) = \frac{1-F(T)}{\tau}$ [Field and Jordan, 2015]. The probability P_{Hist} of our first approach for renewal models is then:

$$\begin{aligned}
 P_{Hist}(T \leq t \leq T + \Delta T | T < t, T_{unknown}) &= \int_0^\infty p(T) P(T + \Delta T < t | T < t) dT \\
 &= \int_0^\infty \frac{1 - F(T)}{\tau} \frac{1 - F(T + \Delta T)}{1 - F(T)} dT \\
 &= \frac{\int_0^\infty (1 - F(T + \Delta T)) dT}{\tau}. \quad (S3)
 \end{aligned}$$

It defines the probability to have no earthquakes in the time period T and $T + \Delta T$, knowing that there has been no earthquake until T but not knowing T . This probability takes into account all possible T given its initial BPT assumptions, avoiding us to sample it.

We now explore the range of possible magnitude and frequency for the largest possible earthquake needed to close the budget.

We use again the MCMC method to sample these probabilities. Each sample is taken from a uniform PDF between 6.4 and 9 for magnitudes, and between 10^{-6} and 10^4 yrs⁻¹ for frequencies. We calculate for each sample k the value proportional to the probability $P_{G,k}$ as described in the method section. For each estimate, we run the sampler for 900000 steps, rejecting the first 500 steps to ensure accurate sampling of the PDF. The space of magnitude-frequency-time of the last event is then binned into cells of size 0.01 for the moment magnitude axis and logarithmically binned into cells of size 0.01 concerning the frequency axis. The number of samples in each cell divided by the total number of samples approximates the probability $P_{G,k}$ as defined by the MCMC method. The results are shown in Figure S1.

The probabilities computed using the BPT model for low magnitude and high frequency events are smaller than Poisson model. As already mentioned in the main text in the 2nd approach section, the marginal cumulative probabilities do not give us any useful information. This is due to the fact that the maximum magnitude and lowest frequency are unbounded a priori. The magnitude for which 95% of the maximum non-cumulative probability is reached can instead give us an indication on the magnitude and frequency needed to close the moment budget (Figure S2). We obtain almost the exact same values as for the Poisson model: a $M7.42$ with a recurrence time of 2000 years is sufficient to close the moment budget at a 95% confidence level. However, the normalized probabilities are lower for the BPT model between $M6.4$ and $M7$ than Poisson's model. Note that the 95% value does not actually represent a probability due to the normalization of the non-cumulative probabilities. Higher magnitudes and recurrence times than those estimations are as plausible but would not be necessary.

The hazard probability estimation, P_{Hazard} , for renewal models is given by Field and Jordan, 2015:

$$P_{Hazard}(T \leq t \leq T + \Delta T | T < t, T_{unknown}) = \int_0^{\infty} p(T) P((T \leq t \leq T + \Delta T | T < t)) dT$$

$$= \frac{\Delta T - \int_0^{\Delta T} F(T) dT}{\tau}, \quad (S4)$$

where $p(T)$ is the probability distribution for the time since the last event, T . It defines the probability to have an earthquake in the time period T and $T + \Delta T$, knowing that there has been no earthquake until T but not knowing T .

Figures S3 shows the hazard estimation for BPT model derived from the probabilities of the magnitude-frequency of the largest earthquake using our second approach described above for a period of time of 30 years. $\bar{P}_{Hazard}$ for the Poisson model is also indicated in the figure.

96 The $\bar{P}_{Hazard}$ of the BPT model is equivalent to the Poisson's model for $M_{Test} \lesssim 4$; they are
97 both almost equal to 1. They then differentiate, the BPT model having higher probabilities
98 for $4 \lesssim M_{Test} \lesssim 6$. The BPT model $\bar{P}_{Hazard}$ joins back to Poisson's model $\bar{P}_{Hazard}$ for larger
99 M_{Test} . This highlights the fact that $\bar{P}_{Hazard}$ for the BPT model is significantly higher for
100 events over M_{Test} that have frequencies around ΔT .

101

References

Field, E. H., and T. H. Jordan (2015), Time-Dependent Renewal-Model Probabilities When Date of Last Earthquake is Unknown, *Bull. Seismol. Soc. Am.*, 105(1), 459–463, doi:10.1785/0120140096.

Matthews, M. V, W. L. Ellsworth, and P. A. Reasenberg (2002), A Brownian Model for Recurrent Earthquakes, *Bull. Seismol. Soc. Am.*, 92(6), 2233–2250.

Topozada, T. R., D. M. Branum, M. S. Reichle, and C. L. Hallstrom (2002), San Andreas Fault Zone , California : $M \geq 5.5$ Earthquake History, *Bull. Seismol. Soc. Am.*, 92(7), 2555–2601.

Table S1: Flow chart of methodology

	<i>Steps</i>	<i>Parameters used</i>
Calculation of P_{Budget}	<ol style="list-style-type: none"> 1. Choose variables (M_{max}, τ_{max}) 2. Assume Bath law: ΔM between the mainshock and the larger aftershock 3. Assume contributions from aftershocks and background seismicity 4. Assume aseismic contribution 5. Calculate $P_{Budget}(M_{max}, \tau_{max})$ 	M_{max}, τ_{max} $B (=0 \text{ or } 1)$ b α $\dot{m}_0$
Calculation of P_{Hist}	<ol style="list-style-type: none"> 1. Choose variables (M_{max}, τ_{max}) 2.a Assume Poisson model 2.b Assume renewal model 3.a Approach1: Assume the largest earthquake follows the GR law determined by known seismicity 3.b Approach2: The largest earthquake does not need to follow the GR law determined by known seismicity. Only the duration (t_{hist}) of the catalogue matters. 4. Calculate P_{Hist} 	M_{max}, τ_{max} τ τ, α_{BPT} a, b t_{hist} Equation 13 and 15
Calculate of P_G		Equation 18

(a)

(b)