

Response to Reviewer 3, “The impact of earthquake cycle variability on neotectonic and paleoseismic slip rate estimates”

Richard Styron

Reviewer comments denoted as [Bx] where x is the comment number. Comments referenced as [Ax] or [Ox] refer to those from Reviewer 1 and 2, respectively.

Review by Anonymous Reviewer #3

[B1] Uncertainty in estimates of earthquake recurrence and fault slip rate are important parameters, pursued by conscientious investigators of seismic hazard. The author perceives a lack of statistical support, and offers in this paper “insights” in variously tones of “friend of the practitioner” and “trust me, I’m the numerate one here”. Neither is convincing.

Someone’s grumpy!

Since this review seems to aim to shame rather than help, the number of changes are few.

[B2a] A couple of omissions in this paper are particularly striking.

[B2b] First, how do we have a paper addressed to “variability” in fault slip rate, addressing particularly the problem of small samples, without mentioning the methods of estimation for censored samples? There is an extensive statistical literature to estimate parameters and uncertainties and numerous recent papers applying it in paleoseismic contexts. This literature provides real quantitative methods to deal with the open intervals, long or short, that affect the geologist’s estimate of fault slip rate and recurrence estimation. These are real equations, with real uncertainties. One would look in vain in this paper for anything of similar substance.

The paper doesn't go into depth on the topics of open intervals and censored data because this paper does not use any samples of earthquake recurrence intervals or per-event displacements in the rate calculations. Though they are present behind the scenes, all of the numerical results in the paper are slip rates derived as the total offset over the total time (equation 1)—no information on how many earthquakes are present in that time window, or what the duration of open intervals may be, is used in the calculations. This is intended to simulate a neotectonic-type slip rate estimation (as written in page 2, lines 24-25 of the original manuscript), where an offset geological marker layer is dated and the offset is measured. I have reinforced this by restating in the methods that neither the number of events or the length of the closed or open intervals are known.

The decision to avoid discussing the topic of including censored data in parameter estimation is also warranted because the most robust method to incorporate open intervals are in maximum-likelihood estimation, which is a complex topic to bring up, as are survival analysis and related methods. Since these topics are complimentary, not central, to this paper I don't see the benefit of adding a dense paragraph (much less any equations) for methods that aren't relevant.

I don't disagree that more substance is offered by methods developed through the decades by many researchers than I can offer in one short paper. Sorry!

Changes: brief reinforcement that this study uses neotectonic (age/offset) slip rate measurements, not those from multiple events as in a paleoseismic context.

[B2c] Second, pages of this paper could be replaced (and improved) by a presentation and discussion of the properties of the standard error. E.g., given an estimate a sample- based estimate of the mean, how far might the population (or true) mean be from the estimate? S.E. is estimated by the sample standard deviation divided by the square root of the number of samples. So, of course, estimates from small samples from a fuzzy log normal converge more slowly than from a well-defined (quasi-periodic) lognormal. Instead of a small equation ($SE=s/\sqrt{n}$), our paper back-calculates the result using 2 million years of samples, and presents the results like a new discovery . And again, with little by way of meaningful uncertainties (e.g., p1, lines 11- 13, 14- 16).

The standard error does share the same basic decrease in dispersion with increasing number of events, but it is symmetric even though the distributions used here are not, and it does not account well for the systematic bias in short time windows that are observed in these results. The misfit between the symmetric and asymmetric approximations becomes particularly bad when estimating the 'epistemic' uncertainty as in Figure 6, which is the uncertainty that one should apply to a measurement to account for earthquake-cycle variability. I have calculated the standard error and added to this graph show in Figure 1 below (not added to manuscript). The results do not support the use of the standard error here. The upper error envelope under-approximates the uncertainty at short time windows (due to the measurement bias), but then becomes acceptable. However, the lower error envelope stays at or below

the 5th percentile, though it should be the 1-standard deviation approximation.

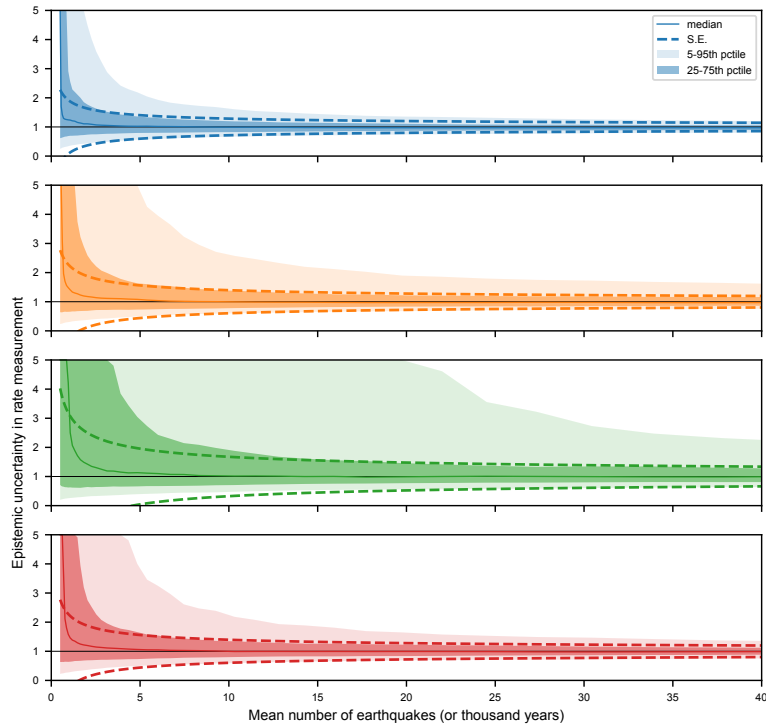


Figure 1: Standard error (S.E.) plotted on top of Figure 6 from the manuscript.

The standard error is also not a great fit for neotectonic slip rate estimations: They are typically derived by dividing net offset over net time n is not known, nor are the durations of the open intervals. It would be different if this paper was at all concerned with the calculation of recurrence intervals, per-event displacements, or other parameters of individual earthquakes where a number of samples is used, but that is a different topic.

Finally, meaningful uncertainties are given in Table 1 for all distributions at a wide range of times. It's not appropriate to reduce all of these to a single equation.

Changes: Short discussion of standard error added to 'slip rate calculations' section.

[B3] 1 L 9,10: We read that the most important parameter is the coefficient of variation. First, this equation is the arithmetic coefficient of variation, and not the CV for a lognormal distribution. The CV of a

lognormal does not depend on the mean.

The reviewer is really confused about the lognormal distribution (see [B6] and [B7] where the reviewer continues to make mistakes and level false accusations at me).

I was quite worried that I made a mistake after reading this, so I re-checked my math, my code and re-read a range of references (e.g., *Statistical Distributions* by Evans, Hastings and Peacock, 1993; *Confidence Intervals for the Coefficient of Variation for the Normal and Lognormal Distributions*, Koopmans, Owen, and Rosenblatt, *Biometrika*, v. 51, no. 1/2, 1964, as well as a lot of internet pages). Nowhere do I find alternative definitions of the CV for a lognormal distribution, except in the Wikipedia page for Coefficient of Variation where two conflicting definitions for the Geometric Coefficient of Variation are given. Both of these involve estimating the CV from samples rather than as an algebraic transformation of the equations for a lognormal distribution. The first definition is mathematically the same as the standard definition; the difference lies in whether the estimation procedure is done on the raw data or log-transformed data. The second is simply the square of the first.

Just to reiterate, the arithmetic coefficient of variation for the lognormal distribution is the standard definition. It is also equivalent to the geometric coefficient of variation for a theoretical longormal distribution (although not necessarily for sample-based estimates).

The reviewer's confusion may stem from the fact that there are two equations for the coefficient of variation that get regularly passed around. However, as I demonstrate in the response to [B7c], they are mathematically equivalent. The simple definition, which I use in the paper, is $CV = \frac{\sigma}{\mu}$, where σ and μ are the standard deviation and mean of the distribution. The other definition, $CV = \sqrt{e^{\phi^2} - 1}$, is given in terms of the standard deviation ϕ of the natural log of the lognormal distribution.

OK, this is pretty funny: All of the confusion that the reviewer has about the lognormal distribution appears to be because of a misreading of the Wikipedia pages for the lognormal distribution and for the coefficient of variation. I mentioned the geometric CV (non-)issue above. The quip that the 'CV of a lognormal does not depend on the mean' is... kind of true? As demonstrated in the response to [B7c], the standard deviation of a lognormal is a *function of the mean* in that the mean is a coefficient in the standard deviation equation if it's expressed in the typical shape and scale parameters. Therefore $CV = \sigma/\mu = \mu A/\mu$ where A is another mathematical expression (see [B7c]). So, yes, technically it is true but only because μ is on the top and the bottom of the fraction and therefore cancels out; furthermore, you can't derive the expression A without knowing both μ and σ . Regardless, the quip gives all outward appearance of being a slight modification of the sentence "Contrary to the arithmetic standard deviation, the arithmetic coefficient of variation is independent of the arithmetic mean" which is below the equation for the CV of a lognormal on the Wikipedia page. It is not a statement that makes any practical sense unless one is beginning with log-transformed data or distributions.

So to conclude, my math, code and analysis are correct in terms of the use of the CV.

Changes: None.

[B4] We could stop here, but a central flaw in the paper is exposed – nothing in this paper addresses how to obtain this most important parameter. If attempted, the essential emptiness of a 2,000,000 year sample would emerge. No real data set in paleoseismology resolves the mean and standard deviation to better than maybe 50%. Typical sites do well to resolve it to a factor of 2. P3, line 22- 23 reflect this reality.

“the essential emptiness”. . . I’m surprised, and perhaps a bit concerned, to find that I have tapped into the existential despair of Anonymous Reviewer #2.

This study investigates how slip rates measured as *offset / time* of an offset marker are influenced by earthquake cycle variation. It is *not* a paper discussing methods or uncertainties involved in the calculation of the mean and standard deviation of earthquake recurrence intervals or displacements from paleoseismology. Indeed, if “There is an extensive statistical literature. . .” (comment [B2b]) on this topic, why should I rehash it?

I fully agree that a typical paleoseismological study cannot resolve the mean and standard deviation with a high degree of accuracy and I never stated otherwise. Partly for this reason, by using CVs of 0.5, 1, and 2, I cover a pretty wide range of values that can give some idea of the expected behavior of the almost any fault system if the basic recurrence distributions can be estimated or assumed.

A discussion of ‘guesstimating’ the CV is actually already present in the manuscript, starting on p.8, l. 31 in the original submission. However, I do not give an in-depth treatment of estimating the CV on a fault–this topic would be suitable in a paper on interpreting field methods or a general overview of statistical paleoseismology, perhaps, but not in a numerical simulation paper that takes values of the CV *a priori*.

Changes: None.

[B5a] It is not obvious that the author has material experience words “aleatoric” and “epistemic”. Line 1, “aleatoric uncertainty” is a contradiction in terms.

It’s really not a contradiction in terms. ‘Aleatory’ does not mean ‘the absence of uncertainty’, which it would have to in order to contradict the term ‘uncertainty’.

In any case, the term ‘aleatoric uncertainty’ is commonly understood and used by statisticians in addition to us proles. See for example O’Hagan, T. (2004). Dicing with the unknown. *Significance*, 1(3), 132-133.

Changes: None.

[B5b] Bird, Zechar and Frankel all know better than to use the method the author alleges in lines 21 and 22 to arrive at epistemic uncertainty in

slip rate. They would more likely consider the allegation a misreading of their work.

This is a very perplexing comment. The line in question states, “The uncertainty in the resulting slip estimate is typically treated as epistemic, and quantified through the propagation of the measurement uncertainties on the offset and time quantities (e.g., Bird, 2007; Zecher and Frankel, 2009).” Both of the papers referenced are focused on estimating fault slip rates (or *long-term offset rates* in Bird’s terminology) that incorporate the measurement uncertainty in both age and offset measurements. Both of their methods are based in a convolution of the probability distributions for age and offset of the marker features (i.e., the data with their uncertainties). Perhaps the reviewer takes ‘the propagation of the measurement uncertainties...’ to mean that it is done through adding them in quadrature to the final rate, but I clearly do not specify the methods used here and the reviewer may want to reign in the assumptions.

Changes: None.

[B5c] More broadly, the lack of care in writing makes one wonder how to understand this paper.

Sorry!

No changes.

[B5d] p2, L5: A perturbation in slip rate would mean it was slipping at rate X, then changes to Y.

Yes. This is both the intended and the obvious interpretation of the sentence. The slip rate at the fault trace varies because of variability in the frequency and magnitude of accumulating offsets.

No changes.

[B5e] p.2, L14-17 have careful paleoseismologists doing reasonable things in one sentence, then imply they would make plainly rookie mistakes in the next.

The paragraph in question states that the open interval since the last earthquake is often taken into consideration, however the variability in closed intervals is not nearly as often taken into consideration—this is why I have written the paper. Should I not state that paleoseismologists and neotectonicists who consider the effects of the terminal open interval are careful? I don’t feel like being less generous. Should I not point out that variability within the closed intervals is important and rarely considered? No, I think that I should. The first reference in this paragraph makes both the ‘reasonable’ decision and then the ‘rookie mistake’. I am not simply making this stuff up.

Changes: None.

[B5f] From here these read like inexperienced generalizations.

Sorry!

No changes. Changes: None.

[B6a] p.4, L15-23: The descriptions of the lognormal variables here give one pause. First, log-normal parameters do not have units.

This comment is completely off the mark. The shape and scale parameters of a lognormal distribution are indeed unitless, as they are the mean and standard deviation of $\ln L$, where L is a lognormal distribution (this transforms the lognormal distribution into a normal distribution).

However, the mean and standard deviation of the actual lognormal distribution L are in the units of the distribution itself. Nowhere in the manuscript did I state that μ and σ , which are clearly defined as the mean and standard deviation of any distribution in question, are the shape and scale parameters of L . They are simply the mean and standard deviation of the distribution.

I have added a footnote clarifying this and giving the derivation of the shape and scale parameters from the mean and standard deviation.

Changes: Footnote with explanation and equations added.

[B7b] Second, the mean recurrence interval is not the location parameter of a log normal. This is just wrong.

Nowhere in the manuscript did I state this. In fact, the word 'location' is not present anywhere in the submitted manuscript (it has been added to an expanded discussion of probability distributions and their parameters in the revised manuscript as described in the response to Comment [O14], though I still do not make the mistake that I have been accused of).

Furthermore, the lognormal distribution doesn't have a location parameter—it is a 2-parameter distribution specified by shape and scale parameters, as stated in P 3, L. 24 of the submitted manuscript.

Changes: None.

[B7c] Third, if one uses the CV equation for the lognormal distribution (e.g., https://en.wikipedia.org/wiki/Log-normal_distribution#Arithmetic_coefficient_of_variation), the CV will not match the COV alleged here. Given that the study depends on these distributions, we can't really use subsequent conclusions.

This is false. The equations are mathematically equivalent.

The mean μ of a lognormal distribution L is often given in terms of the mean m and standard deviation ϕ of the log-transformed distribution $\ln L$, which is a normal distribution. The equation is:

$$\mu = e^{m + \frac{1}{2}\phi^2}$$

and similarly, the standard deviation σ of L may be given in terms of m and ϕ :

$$\sigma = e^{m + \frac{1}{2}\phi^2} \sqrt{e^{\phi^2} - 1},$$

which is clearly equivalent to

$$\sigma = \mu \sqrt{e^{\phi^2} - 1}.$$

Therefore,

$$CV = \frac{\sigma}{\mu} = \frac{\mu \sqrt{e^{\phi^2} - 1}}{\mu} = \sqrt{e^{\phi^2} - 1}$$

which is the equation for the CV of a lognormal distribution given in any source including the Wikipedia page. Please note that in most sources, the symbols ' μ ' and ' σ ' are what I have here defined as ' m ' and ' ϕ ' because I use the symbols μ and σ to represent the mean and standard deviation of L , consistent with the manuscript where the same symbols are used equivalently across distributions.

The methods and results of this study are consistent and this alleged error did not occur.

Changes: None.

[B7] p.6, L20. If the number of samples is really $n = N - t + 1$, the samples are correlated by virtue of the overlap in the windows. No accounting has been made of the correlation structure.

There is certainly a lot of autocorrelation of the samples if one were to look at them sequentially. Nonetheless, even if sequential samples are correlated, given enough samples, the final distribution of the samples won't depend on the autocorrelation in the samples (this principle underlies the family of powerful and commonly-used numerical methods based on Markov Chain Monte Carlo sampling, for instance). This is why I use 2,000,000 years of simulation: it ensures that the sampling is ample enough to not worry that the serially-correlated samples have not sufficiently sampled the distribution. I would have to use a much longer simulation (tens to hundreds of millions of years) otherwise.

Changes: None.

[B8a] p.6, L26: Starts a narrative of the consequences of the standard error, as though the standard error was never invented.

I have added a brief discussion of the standard error, which is an OK approximation of the results beyond 5-10 mean earthquake cycles (once the systematic bias in very short time windows has gone away), but the asymmetry of the results are not well approximated by the standard error. One could hack on it a bit by log-transforming

the data, applying the standard error of the log-transformed distributions, and then anti-log transforming it, but that seems a bit much.

Changes: None.

[B8b] The fuzzy, back-of-the-envelope estimates start to get thick here. Real uncertainty estimates would serve better.

The uncertainty estimates at a range of percentiles are available in Figure 5, and the uncertainties for a slip rate measurement are in Figure 6 and Table 1.

Changes: None.

[B9a] p.7, L1-3: Two observations: First, as written, the practicing geologist is being asked to believe that 60 earthquake cycles have passed with zero displacement. I can guess what was intended, but should not have to.

Actually, the naive reading of the sentence, that 60 *mean* earthquake cycles passed without an earthquake on a fault in the simulation, is exactly what is intended.

No changes.

[B9b] Second, what probability is associated with this 60-cycle thing? I ask because practicing hazard geologists have to make estimates, and give weights to extreme events. What is the probability of 60 cycles, a CV of 2.0, ...? Hard to imagine that the author has thought much about what these results would mean or how to use them if they were true.

I happen to be a practicing hazard geologist who is responsible for the implementation of fault sources for probabilistic seismic hazard analysis, so I have actually considered this. It's very important when deciding whether to include faults that have very clear bedrock geomorphic signatures (say, from rapid deformation in the Miocene) and are favorably oriented for slip in the current stress regime, but lack evidence of Quaternary deformation. Are they to be included in the source models, or not? I tend to add them, with low slip rates and high uncertainties, for just this reason. The analysis in this paper shows that very long recurrence intervals are possible on active faults.

There are two probabilities that are asked here: 1. *The probability of a fault with a CV of 2.0 going 60 mean earthquake cycles without accruing much displacement.*

This is answered very easily using the equation for the lognormal probability distribution with a CV of 2. The cumulative distribution function for this recurrence distribution at 60,000 years (60 mean earthquake cycles) is 0.9994..., which means that there is a probability of about $p = 0.0006$ that any given recurrence interval will exceed 60,000 years, or 60 mean earthquake cycles. But the probability P of a recurrence interval this long being present in a sequence of n earthquakes is $P = 1 - (1 - p)^n$, and for 2000 earthquakes (which is the approximate number in these simulations) the probability of a single recurrence interval exceeding 60 mean earthquake cycles is 0.1 or 10%. It becomes more likely than not that such a

recurrence interval will be present in a sequence at about 12,500 earthquakes. Are you going to see this in a trench? Almost certainly not even if you were to have the appropriate sedimentation to preserve the signal. But that does not mean that it's an impossible or even improbable event in a long time series.

Changes: I have added the following sentence to the paragraph under discussion: "It is highly unlikely that any given recurrence interval will be this long, but given thousands of earthquakes over millions of years, the chance of such an event occurring at least once is far more likely." I don't find it necessary to go through the probability calculations in the manuscript as I did here.

2. *The probability of a fault having a CV of 2.0.*

The best way to calculate this probability is basically to look at the frequency of such a CV occurring in some sample of faults. I don't have a big collection of faults with paleoseismic data to work with, and I'm not going to put one together to satisfy a reviewer, but I do calculate that the schematic, representative model for southeastern Australian intraplate faults by Clark et al., 2017 (cited in the paper) has a CV of 2. Unfortunately, Clark et al. cannot precisely date all of the earthquakes they infer so no real CV can be calculated from their data, but this is *their* best guess as to how the faults behave, based on their evidence for clusters of events with recurrence intervals of estimated mean 8,000 years separated by quiescent intervals of 0.5-2 million years.

Similarly, the USGS Quaternary Faults and Folds dataset has a number of faults that are thought to have last ruptured in the 'Middle and Late Quaternary' (<1.6 Ma) but not in the 'Late Quaternary' (<130,000 years). Disregarding potential misclassifications of these faults, it seems reasonable to assume that a few of them may be experiencing very long recurrence intervals and may have a CV approaching 2. Many of these faults are along strike of fault segments with Holocene ruptures and there is no obvious geological difference between the segments with recent rupture and those without—the Steens fault zone in the lovely Steens rift of southeast Oregon is an example.

Changes: None.