Reply to "Comment on 'Two Foreshock Sequences Post Gulia and Wiemer (2019)' by Kelian Dascher-Cousineau, Thorne Lay, and Emily E. Brodsky" by Laura Gulia and Stefan Wiemer

Kelian Dascher-Cousineau*[1], Thorne Lay[1], and Emily E. Brodsky[1]

Abstract

Gulia and Wiemer (2019; hereafter, GW2019) proposed a near-real-time monitoring system to discriminate between foreshocks and aftershocks. Our analysis (Dascher-Cousineau *et al.*, 2020; hereinater, DC2020) tested the sensitivity of the proposed Foreshock Traffic-Light System output to parameter choices left to expert judgment for the 2019 Ridgecrest M_w 7.1 and 2020 Puerto Rico M_w 6.4 earthquake sequences. In the accompanying comment, Gulia and Wiemer (2021) suggest that at least six different methodological deviations lead to different pseudoprospective warning levels, particularly for the Ridgecrest aftershock sequence which they had separately evaluated. Here, we show that for four of the six claimed deviations, we conformed to the criteria outlined in GW2019. Two true deviations from the defined procedure are clarified and justified here. We conclude as we did originally, by emphasizing the influence of expert judgment on the outcome in the analysis.

Introduction

GW2019 proposed a near-real-time monitoring system to discriminate between foreshocks and aftershocks based on temporal changes in *b*-value relative to background rates, applying it to 29 well-monitored earthquake sequences, two of which included foreshocks. The objectives of DC2020 were to (1) independently replicate the procedure for two new well-recorded sequences with clear foreshocks, and (2) test how sensitive the resulting Foreshock Traffic-Light System (FTLS) output is to several decisions left to expert judgment. We concluded that key decisions affect the FTLS outcome for the 2019 Ridgecrest earthquake sequence and give ambiguous results, whereas Gulia *et al.* (2020; hereafter, GWV2020) found good performance of the FTLS in their application to the 2019 Ridgecrest earthquake sequence.

In their comment on our article, Gulia and Wiemer (2021; hereafter, GW2021) assert that our replication of their analysis deviates in at least six different ways from the proposed method (as given in GW2019). These purported deviations lead to different pseudoprospective warning levels, particularly for the Ridgecrest aftershock sequence.

Here, we show that for four of the six claimed deviations, DC2020 conformed to the criteria outlined in GW2019, and that, in some cases, the criticisms in GW2021 are in direct contradiction with the guidelines in GW2019. There are two true deviations from the defined procedure that we should have better articulated. We explain and discuss the rationale for these deviations. One attempts to reconcile the code distributed by GW2019 with the published documentation in GW2019. The other stems from a decision to encompass a volume that was robust to uncertainty in early hypocenter depths, as would be required for a real-time application. We conclude, as we did originally in DC2020, by emphasizing the influence of expert judgment in the analysis. Considering the importance and public impact of real-time earthquake warnings and the scarce opportunities for validating proposed methodologies, a next generation of the FTLS needs to be robust to differences in expert judgment, propagating model uncertainty into warning levels.

To facilitate comparison with GW2021, we follow the same labeling. We group deviations 1.3 and 3.1 together as they reflect the same issue. Where it is relevant, we directly quote GW2019 to avoid misrepresentation.

Deviation 1.1: DC2020 uses a catalog from the year 2000, while we would advise (and do so in Gulia *et al.*, 2018, 2020) to use data from 1981 (GW2021)

We did use the year 2000 as a starting point for our background catalog for the Ridgecrest earthquake sequence. The start date is

1. Department of Earth and Planetary Sciences, University of California Santa Cruz, Santa Cruz, California, U.S.A., https://orcid.org/0000-0002-3118-8056 (KD-C); https://orcid.org/0000-0003-2360-4213 (TL); https://orcid.org/0000-0002-6855-6860 (EEB)

*Corresponding author: kdascher@ucsc.edu

important for defining the *b*-value of background activity within a target event source volume and is a decision left to expert judgment. Our choice followed the guidelines in GW2019:

> *The start time of the pre-event catalogue depends on the quality and completeness of the local network and sometimes on avoiding overlap with past sequences (in our case, we choose 1 January 2012 for both Japan and Italy; in Italy, to avoid overlap with the L'Aquila aftershocks and in Kumamoto to avoid the influence of the 2011 Tohoku $M_w = 9$ megathrust event). The pre-event period should ideally contain several years of seismicity for a robust estimate. (GW2019)*

Figure 1 shows local seismicity from 1980 up to the Ridgecrest sequence. The 1980–2000 period is seismically active, with apparent triggering from large regional earthquakes including 1992 Landers and, to a lesser extent, 1999 Hector Mine, along with the strong local 1995 Ridgecrest sequence. In the accompanying comment, the authors do not mention the local 1995 Ridgecrest sequence and, despite a notable spike in local seismicity in 1992, they argue that the Landers (and Hector Mine) events are too far away to influence seismicity within the source volumes of the Ridgecrest sequence. It was our judgment that the background should be established after 2000 when seismicity is not perturbed by past local sequences or triggering from large events, as prescribed in GW2019. The background catalog interval of 18 yr more than doubled the intervals used to establish a background for the two foreshock sequences analyzed in GW2019. Our choice for the catalog start date is a decision left to expert judgment by GW2019, not a deviation of the method as asserted. Some discrepancies discussed subsequently are a direct consequence of this choice, which we view as more reasonable than extending the catalog back to 1981.

The sensitivity to the catalog start date is the first point we emphasized in our original discussion and something we explored more fully using start dates ranging from 2000 to 2012 in the Monte Carlo sampling.

Deviation 1.2: DC2020 apply erroneously an additional M_c increment 0.4 rather than 0.2 (GW2021)

DC2020 does indeed apply the magnitude of completeness correction twice. We used the code distributed by GW2019 which necessitated adaptation for individual mainshocks because of hard-wired parameters which we struggled to reconcile with the procedure as described in the Methods section of GW2019. In GW2019's primary code (Run_TLS_Gulia_Wiemer.m), the multistep correction procedure is inconsistent across the background, foreshock, and aftershock intervals. The background events are screened for events below the completeness thresholds of M_{MAXC} which is implemented with a +0.2 increment at line 162 and a −0.2 increment at line 172. The foreshocks are screened with $M_{MAXC} + 0.2$ (lines 168–176). The aftershocks are screened with $M_c = 1.8$ (line 237). After these initial screenings, the data are further screened for completeness with an additional +0.2 correction (function calls on lines 192 or 214 for the background period; lines 226–227 for the foreshock period and lines 244–245 for the aftershock period). Thus, aggregate corrections of 0.2, 0.4, and 0.2 were applied respectively. Faced with a nontransferrable processing scheme that both did not describe how an initial $M_c = 1.8$ was selected for the aftershock interval and differed in our reading from the documentation provided in the published paper, we opted to uniformly use the conservative and consistent correction of 0.4 to the maximum curvature estimates (+0.2 globally, with additional +0.2 in every event window for our primary result) that is consistent with the coded method in GW2019 for the most important period, that is, the foreshocks.

Estimating catalog completeness is susceptible to the peculiarities of local catalogs. We therefore also included in DC2020 a range of corrections ranging from 0.1 to 0.3 in our Monte Carlo simulation with the corrections applied at both screening steps, resulting in aggregate corrections of 0.2–0.6. Using more conservative completeness thresholds trade-off precision (reported with the standard deviation at every timestep), for accuracy.

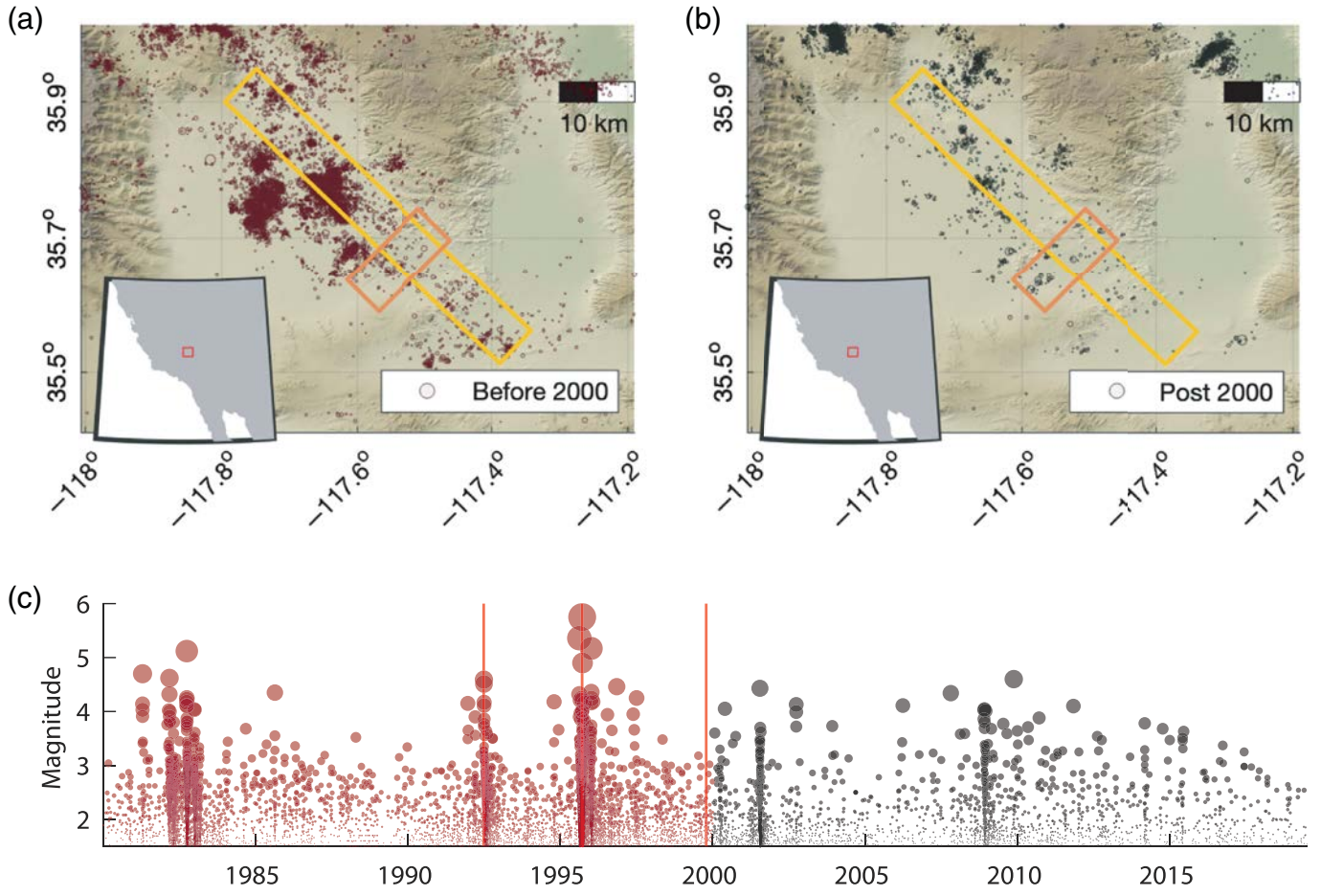
For the sake of completeness, we have now additionally performed the primary calculation of DC2020 with a single +0.2 correction (Fig. 2). Lower relative *b*-values particularly during the aftershock sequence toggle the aftershock period from a yellow (ambiguous) to red warning indicating an impending larger earthquake that has not yet been observed. Because the original code provided with GW2019 hard-wired specific mainshocks into the code, we provide at the end of this article our version that produce the figures in DC2020 as well as Figure 2.

Deviations 1.3 and 3.1: To establish the reference *b*-value, DC2020 sample events in a circular region of about 10 km around the epicenter, while Gulia *et al.* (2020) use events in a box within 3 km of the rupture plane (GW2021)

The background seismicity in the Ridgecrest source volumes from 2000 to 2018 is very sparse (Fig. 1b), presenting difficulty estimating a stable *b*-value. For sparse background seismicity, GW2019 indicate the following procedure:

> *(b) If fewer than N_{pre} [250] events are available, we use the N_{pre} [250] events that are nearest to the epicentre and then compute a single regional background b value as reference, following the computational approach defined in (a). This procedure was used for the $M_w = 6.5$ Kumamoto event (Fig. 2a), sampling a distance of up to 17 km from the epicenter. (GW2019)*

N_{pre} (250 events) refers to the minimum number of background events remaining after the first completeness



correction in the target source volume (a box within 3 km of the estimated rupture surface inferred from scaling relationships for the event magnitude). Because we used a later catalog start date (2000 instead of 1981, as discussed earlier) in a region with low-true background activity, we have far fewer background events in the two choices of foreshock and the mainshock source volumes for the Ridgecrest sequence (Fig. 1). This situation triggers the above-stated processing step in the original code. Thus, instead of using the deficient number of events in the box within 3 km from the target faults in Ridgecrest, we use the nearest 250 events to the epicenter in the regional catalog spanning the catalog start time to the occurrence of the M_w 6.4 foreshock, as specified by GW2019. Low-background activity levels trigger this condition for both the foreshock ("deviation 1.3") and mainshock ("deviation 3.1") for the Ridgecrest and Puerto Rico events.

Figure 1. Local seismicity along with approximate source volumes used for the 2018 Ridgecrest foreshock (orange box) and the mainshock (yellow box) for the period (a) prior to 2000 and (b) after 2000. The true source volume is 3D with lateral dimensions scaling with magnitude and orientation prescribed by the Global Centroid Moment Tensor solution. Inset Maps indicate the corresponding map extent in (a–b). (c) Time sequence of regional seismicity in the Ridgecrest area from 1980 to 2019 prior to the Ridgecrest sequence. The area features triggered seismicity in 1992 following the Landers earthquake and intense activity in 1995 associated with the regional M_w 5.7 earthquake along the 2018 Ridgecrest mainshock rupture.

Deviation 2.1: DC2020 selected aftershock of the first hour, rather than the first 24 hr to define the active fault. They thus selected the alternate fault plane for estimating the *b*-values of the aftershocks following the first mainshock (GW2021)

We note the guidance from GW2019 on this point:

> *Typically, one to several hours of aftershocks are sufficient to select the right plane, and rapid source-inversion approaches can also deliver a finite fault model within 1–2 h.*

GW2019 offer no further guidelines. DC2020 indeed selected the southwest–northeast fault plane based on the first hour of aftershocks for the main solution. No point in the original paper specifies that the first 24 hr should be used to define the fault. We note that early finite-fault models also assumed rupture on the southwest–northeast plane because the early aftershock distribution has greater numbers and

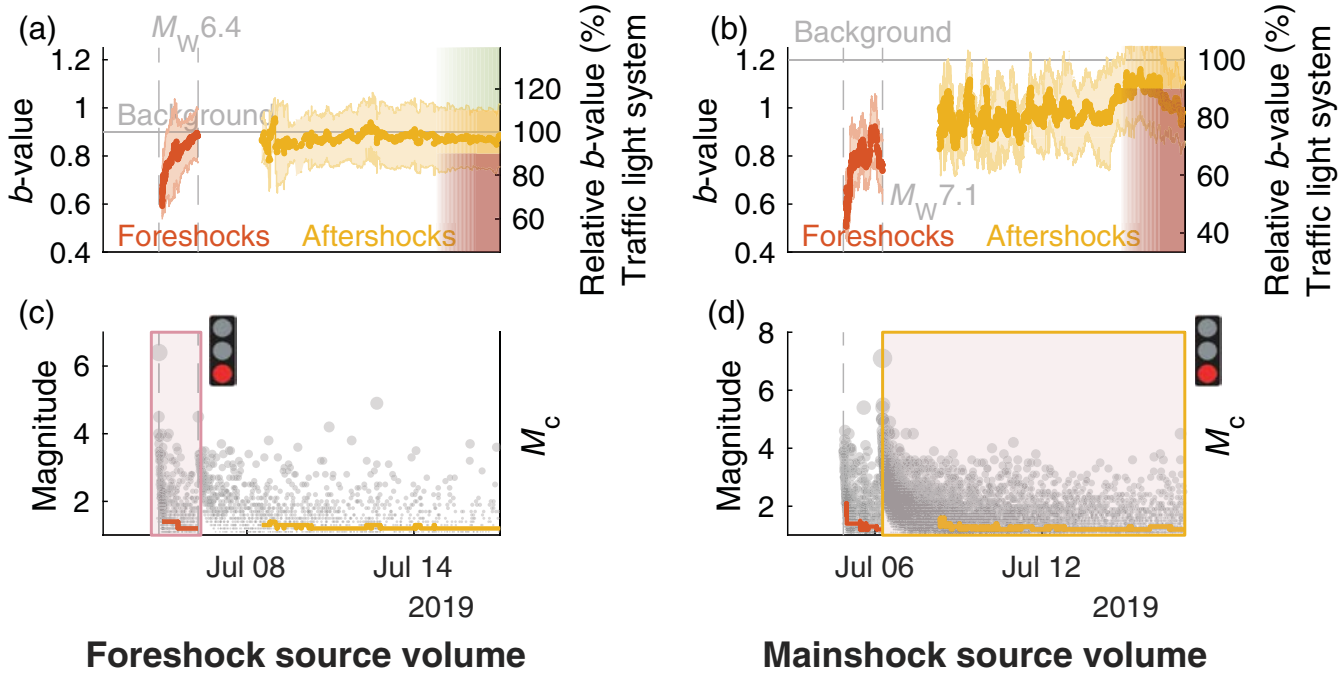


Figure 2. Primary result for Ridgecrest with a single completeness correction. (a,b) Time series of *b*-value estimates during the sequence with 1σ error bars for the corresponding source volumes indicated below (c,d) and shown in Figure 1b. Dashed lines indicate the timing of the 4 July 2019 M_w 6.4 foreshock and the 5 July 2019 M_w 7.1 mainshock. The traffic-light criteria relative to the background level are indicated on the right. (c,d) Time series of event magnitudes during the sequence in the corresponding volumes (Fig. 1b). Colored curves indicated the time-varying catalog completeness M_c during the intervals of the foreshock and aftershock sequences used for *b*-value computation. Note that with a single correction the traffic light warning for both the foreshock (red box in c) and aftershock (yellow box in d) periods are red, indicating that there should be an impending larger earthquake that has not yet occurred.

length on that plane than on the orthogonal plane of the "L-shaped" rupture. The 2018 Ridgecrest foreshock is truly ambiguous in terms of defining the "right" source volume in the procedure of GW2019, given that both planes ruptured simultaneously (Liu *et al.*, 2019). There is no guidance in GW2019 for a case involving simultaneous rupture of two orthogonal planes. We therefore included either east–west or north–south choices of source volume in the Monte Carlo suite of calculations. The deficiency of background events from 2000 to 2018 triggers the "nearest 250" event option for estimating a reference *b*-value for either choice of the foreshock target source volume, and with the epicenter being near the intersection of the two planes, this gives the same reference *b*-value estimate for both cases.

Deviations 2.2: DC2020 do not limit the analysis to events with 3 km depth below and above the fault plane, but extend the sampling down to 20 km.

We thank Gulia and Wiemer for raising the issue of the depth-extent of the source volume. This is a challenge for near-real-time applications due to time-varying estimates of hypocentral position. In DC2020, we established source–volume boxes centered upon the hypocentral locations and rotated according to the strike and dip of the respective foreshock and mainshock faulting mechanisms. These source volumes extend 3 km on either side of the rupture plane, as prescribed in GW2019. Estimated horizontal dimensions of the source–volume boxes were defined by the Wells and Coppersmith (1994) relationship for strike-slip faulting subsurface rupture length (RLD in their table 2A). This gave lengths of 25 km for the foreshock and 68 km for the mainshock. Wells and Coppersmith (1994) relationships also estimate down-dip subsurface width (RW) for strike-slip faulting, which gives 9.3 km (5.25–16.44 km range) for the foreshock and 14.4 km (7.9–26 km range) for the mainshock. One can then center the box on the hypocenter (as prescribed by GS2019). We deviated from this depth windowing procedure by including source volumes extending across the ~15 km seismogenic layer depth; this is a true deviation from the defined procedure that we should have articulated, with a justification as given subsequently.

Early catalog locations gave hypocentral depths of 10.7 km for the foreshock (available by 9:21 a.m. on 15 July 2019) and 17 km for the mainshock (available on 6 July 2019), but the latter was quickly revised to 8 km. Our postprocessing of the sequences used the catalog of Shelly (2020) which gives hypocentral depth estimates of 15 km and 3 km for the

foreshock and mainshock, respectively. This substantial variability in hypocentral depths combined with a choice of centering the source–volume box on the uncertain hypocentral depth significantly impacts the placement of the source volume used for fault-plane preference and *b*-value time-series calculations. The vast majority of seismicity in this region is from 1 to 15 km depth. Immediate aftershock hypocenters of both events spanned this full depth range. Using the final catalog depth of 8 km for the mainshock, with central value for RW gives a depth range of 8 ± 7.2 km, which spans the main seismogenic zone. By allowing a greater depth extent, only a few additional events are included in our analysis of the mainshock sequence. There would be a significant truncation of activity if we centered the source volume on hypocenters at either 17 or 3 km for the same RW value. For the foreshock hypocenter at 10.7 km, using the central value for RW gives a depth range of 10.7 ± 4.7 km or 10.7 ± 8.2 km for the maximum range of RW. Using the Shelly catalog depth of 15 km, centering the source volume leads to up-dip truncation of the seismicity. Our choice of extending the source volume across the seismogenic zone has a greater impact on the foreshock sequence than on the mainshock sequence. Still, with early seismicity extending across the depth range, and considering that Wells and Coppersmith (1994) themselves calibrate rupture dimensions on early aftershocks, it is not clear that our decision is any less or more correct regarding the down-dip width of the source volume.

Uncertainty in actual hypocentral depth and ambiguity as to whether the hypocenter is at the top, bottom, or center of the actual rupture area, warranted extending the source volumes throughout the depth range of the activity. In Southern California, where most seismicity is less than 15 km depth, this would be our recommended way of dealing with the uncertainty in hypocentral depth estimates and down-dip width of the source volume for large events. For other regions, this may not be appropriate.

Puerto Rico Case Study

As the accompanying comment notes, DC2020 explicitly stated that the analysis of the Puerto Rico Sequence did not strictly comply with the criteria of GW2019, in part due to the low magnitude of the largest foreshocks and in part due to paucity of events within the putative M_w 5.0 foreshock source volume, which was widened from 3 to 10 km from the fault to capture sufficient activity. In the same way that GW2019 included the Tohoku sequence without giving weight to the findings for it, we included application to a second clear foreshock sequence despite the catalog limitations. The near-coast sequence is unusually well monitored for an offshore event (in contrast to the Tohoku case); and its aftershock sequence was vigorous and ongoing during our submission process, making it a fully prospective test (albeit of a slightly adapted hypothesis).

Conclusion

Readers that have studied GW2019, DC2020, GWV2020, and the comment and reply may find this discussion somewhat unsatisfactory. The prognostic value of the traffic-light system and the physical implications thereof remain ambiguous. GW2019, GWV2020, and to some degree DC2020, all suggest that there may indeed be a precursory signal in the *b*-value time series of foreshocks. Following our best judgment, DC2020 found that the Ridgecrest foreshock sequence did have reduced *b*-value along the east–west plane that dipped below background level during the foreshock sequence (red warning). This is consistent with GWV2020 who used an earlier catalog start time and the north–south plane for estimating the background *b*-value. The magnitude of this dip and its recovery with time, however, are found to be sensitive to seemingly innocuous decisions associated with expert judgement that determine whether options in the FTLS method of GW2019 are exercised. For the Ridgecrest mainshock, the main difference in DC2020 and GWV2020 is that the latter paper estimates a lower background *b*-value apparently because their time window includes a prior earthquake sequence and triggered activity. This enhances the *b*-value increase of the aftershock sequence relative to background, giving a green light in GWV2020, whereas the higher background *b*-value estimate in DC2020 yields a yellow light, or ambiguous definition of aftershock activity.

The goal of DC2020 was not to stunt progress with this approach but rather to contribute to its advancement. Independent research groups replicating scientific methods are a desirable feature that we hope to continue fostering in our community. The outcome of this exercise illustrates the concerns we raised in our original submission. The reliance on expert judgement suggests that the method, as stated, lacks the specificity to robustly monitor the underlying physical process accurately. The method, as defined in GW2019, needs the following improvements: (1) explicit and physically motivated criteria for determining background *b*-values free of contamination by prior sequences and remotely triggered activity, particularly for low-background level regions, (2) a more robust method for the determination of catalog completeness for *b*-value estimation, and (3) a provision for multifault ruptures and uncertain mainshock rupture source dimensions. Foregoing empirical calibration in favor of early aftershock distributions to directly estimate the source volume may help address this last point. Improving robustness of the *b*-value estimation with respect to time-varying catalog completeness may greatly help reduce short-term fluctuations and increase usable data during an ongoing sequence. Other concerns arise with respect to using sufficient numbers of observations for any *b*-value estimate to be meaningful. Addressing these issues will ensure that prospective tests are comparable across multiple research groups without ambiguity.

Data and Resources

The data and MATLAB scripts used in this reply and to reproduce the analysis are available at https://github.com/keliankaz/DC2020_reply (last accessed February 2021). The MATLAB is available at www.mathworks.com/products/Matlab (last accessed February 2020).

Declaration of Competing Interests

The authors acknowledge there are no conflicts of interest recorded.

Acknowledgments

The authors would like to thank the members of the University of California Santa Cruz (UCSC) seismology laboratory for providing thoughtful insight and lively debate on the topic. This work was funded by National Science Foundation (NSF)-EAR Grant 1761987 (E. B.) and NSF-EAR Grant 1802364 (T. L.). To the best of their knowledge, no author has any conflict of interest publishing this research.

References

Dascher-Cousineau, K., T. Lay, and E. E. Brodsky (2020). Two foreshock sequences post Gulia and Wiemer (2019), *Seismol. Soc. Am.* **91,** no. 5, 2843–2850.

Gulia, L., and S. Wiemer (2019). Real-time discrimination of earthquake foreshocks and aftershocks, *Nature* **574,** no. 7777, 193–199.

Gulia, L., and S. Wiemer (2021). Comment on "Two foreshock sequences post Gulia and Wiemer (2019)", *Seismol. Res. Lett.* doi: 10.1785/0220200428.

Gulia, L., A. P. Rinaldi, T. Tormann, G. Vannucci, B. Enescu, and S. Wiemer (2018). The effect of a mainshock on the size distribution of the aftershocks, *Geophys. Res. Lett.* **45,** no. 24, 13–277.

Gulia, L., S. Wiemer, and G. Vannucci (2020). Pseudoprospective evaluation of the foreshock traffic-light system in Ridgecrest and implications for aftershock hazard assessment, *Seismol. Soc. Am.* **91,** no. 5, 2828–2842.

Liu, C., T. Lay, E. E. Brodsky, K. Dascher-Cousineau, and X. Xiong (2019). Coseismic rupture process of the large 2019 Ridgecrest earthquakes from joint inversion of geodetic and seismological observations, *Geophys. Res. Lett.* **46,** no. 21, 11,820–11,829.

Shelly, D. R. (2020). A high-resolution seismic catalog for the initial 2019 Ridgecrest earthquake sequence: Foreshocks, aftershocks, and faulting complexity, *Seismol. Res. Lett.* **91,** no. 4, 1971–1978.

Wells, D. L., and K. J. Coppersmith (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement, *Bull. Seismol. Soc. Am.* **84,** no. 4, 974–1002.

Manuscript received 1 March 2021
Published online 28 July 2021