Pseudo-prospective evaluation of the Foreshock Traffic Light System in Ridgecrest and implications for aftershock hazard assessment

Laura Gulia[1*^], Stefan Wiemer[1] and Gianfranco Vannucci[2]

[1] Swiss Seismological Service, ETH Zurich, Switzerland.

[2] Istituto Nazionale di Geofisica e Vulcanologia, Bologna, Italy

*corresponding author: Laura Gulia (lgulia@ethz.ch, laura.gulia@unibo.it),

^ now at the University of Bologna, Department of Physics and Astronomy, Viale Berti Pichat 8, Bologna.

Abstract

The Mw7.1 Ridgecrest earthquake sequence in California in July 2019 offered an opportunity to evaluate in near real-time the temporal and spatial variations in the average earthquake size distribution (the b-value) and the performance of the newly introduced Foreshock Traffic Light System (FTLS). In normally decaying aftershock sequences, the b-value of the aftershocks was in past studies found, on average, to be 10%-30% higher than the background b-value. A drop of 10% or more in 'aftershock' b-values was postulated to indicate that the region is still highly stressed and that a subsequent larger event is likely. In this Ridgecrest case study, after analysing the magnitude of completeness of the sequences, we find that the quality of the monitoring network is excellent, which allows us to determine reliable b-values over a large range of magnitudes within hours of the two mainshocks. We then find that in the hours after the first Mw6.4 Ridgecrest event, the b-value drops by 23% on average, compared to the background value, triggering a red foreshock traffic light. Spatially mapping the changes in b, we identify an area to the north of the rupture plane as the most likely location of a subsequent event. After the second, magnitude-7.1 mainshock, which did occur in that location as anticipated, the b-value increased by 26% over the background value, triggering a green traffic light. Finally, comparing the 2019 sequence with the Mw5.8 sequence in 1995, where no mainshock followed, we find a b-value increase of 29% after the mainshock. Our results suggest that the real-time monitoring of b-values is feasible in California and may add important information for aftershock hazard assessment.

Introduction

It is well known and almost universally observed that the stress changes caused by a major earthquake strongly affect seismic activity in the vicinity, and the rate of earthquakes increases near the mainshock rupture by several orders of magnitude (Okada, 1992; Stein, 1999; Ebel et al., 2000). In most sequences, on average, this aftershock activity then decays exponentially back to the previous background rate (e.g. Reasenberg and Jones, 1990), a process first described by Omori in 1895 and nowadays often described with reference to the concept of Epidemic Type Aftershock Sequences (ETAS) (Ogata, 1988). This systematic aftershock behaviour can be satisfactorily explained and well modelled using models combining Coulomb stress changes and rate-and-state friction (Dietrich et al., 2000; Toda and Stein, 2003). It also constitutes the baseline of probabilistic assessments of aftershock probabilities (e.g. Reasenberg and

Jones, 1994; Marzocchi et al, 2017; Omi et al., 2019). Today, the term Operational Earthquake Forecasting (OEF) is often used when referring to aftershock forecasting in near real- time (Zechar et al., 2016; Jordan et al., 2014).

Far less well established and not currently used in OEF is the fact that the stress redistribution caused by a mainshock also systematically influences relative earthquake size distribution, the b-value of the Gutenberg and Richter relationship (Gutenberg and Richter, 1944; Ishimoto and Iida, 1939). Laboratory measurements taken since the 1960s have established that b-values are sensitive to stress (Scholz, 1968; Goebel et al., 2013; Amitrano, 2003) and this inverse dependency of b-value and the applied stress is fully consistent with a number of observed b-value variations with depth, faulting style and the loading state of faults (e.g. Petruccelli, 2019 a, b; Staudenmeier et al., 2019; Scholz, 2015; Tormann et al., 2015; Gulia and Wiemer, 2010; Narteau et al., 2009). Mainshock stress changes are therefore expected to systemically change b-values, as suggested by a number of case studies (Wiemer and Katsumata, 1999; Wyss and Wiemer, 2000; Enescu and Ito, 2002). Just recently, Gulia et al. (2018) confirmed this hypothesis in a systematic study. To establish generic b-value behaviours in aftershock sequences, they applied a stacking approach to 31 high-quality aftershock sequences from California, Japan, Italy and Alaska and demonstrated that the b-values of those sequences generically increase by 20% after the mainshock. The higher b-value results suggest a far lower probability of a subsequent large event. Gulia et al. (2018) also presented a model based on Coulomb stress changes that explains the observations and the observed dependencies on distance, magnitude and faulting style.

Based on these findings, Gulia and Wiemer (2019) postulated the hypothesis that sequences in which the b-value of the aftershock decreased by 10% or more instead of increasing as expected would indicate that a bigger event was not yet to occur. The authors then extended their b-value analysis by successfully testing this hypothesis on three sequences where a secondary larger mainshock occurred, and proposed a Foreshock Traffic Light System (FTLS) which, taking b-value evolution over time as an indicator of the average stress condition of faults in a region, defines three alert (or concern) levels that can be used to determine in near real-time whether an ongoing sequence is likely. The lowest, 'green' alert is triggered by a normally decaying aftershock sequence (b-value increases by 10% or more). The highest, 'red' alert indicates a precursory sequence that is more likely to be followed by a larger event (b-value decreases by 10% or more). Sequences falling between these extremes trigger 'orange' alerts. Gulia and Wiemer (2019) tested the FTLS on 58 sequences and found it to be more than 95% accurate. Differential b-value maps are proposed as an additional step to estimate the likely location of subsequent larger events. The FTLS is thus proposed as a tool for real-time discrimination between foreshocks and aftershocks, but the authors also point out that additional, ideally fully prospective tests, are needed before FTLS can be used in Operational Earthquake Forecasting (OEF) systems.

Key to the robustness of b-value based forecast is a correct assessment of the completeness of reporting, Mc, for this variable fluctuates dramatically during aftershock sequences (Hainzl, 2016; Helmstetter et al., 2006; Woessner and Wiemer, 2005). In the past, if often took weeks or even years to post-process the rich catalogues of aftershock sequences to make them fully useful for statistical seismology. Consequently, another objective of our study is to investigate the reliability of assessed statistical parameters of aftershock sequences in the light of improved modern-day network-processing capabilities and automation. A further, related objective is to analyse whether high-precision and more complete datasets based on cross-correlation, provided by Shelly (2020), can improve the reliability and lower the latency of aftershock forecasting. We also investigate another potential limitation of near-real time application, the availability of reliable focal mechanism data.

In many ways, the Ridgecrest sequence is an ideal case study for investigating the effects of mainshock on the size distribution of aftershocks, and our study is the first prospective evaluation of the FTLS as a purely data-driven decision support system. Finally, we discuss the implications of our analysis for aftershock hazard assessment.

The 2019 Mw7.1 Ridgecrest sequence

On the morning of 4 July (at 17:33 UTC time) a Mw6.4 earthquake hit eastern California in the Mojave Desert (Ross et al., 2019), injuring about 20 people and damaging numerous buildings in the Ridgecrest area (earthquake.usgs.gov). Over the past 40 years, this part of southern California has experienced several moderate earthquakes, the largest being a Mw5.8 event on 20 September 1995, about 13 km away from the Mw6.4 event.

The earthquakes following the Mw6.4 quake outline two lineaments: one SW-NE and the other NW-SE, on an unmapped fault, exhibiting a distinctive 'T' pattern created by the simultaneous activation of two or more faults (Ross et al., 2019; Hobbs, 2019). During the hours after the mainshock, United States Geological Survey (USGS) seismologists estimated in near-real tine probabilities of aftershocks and subsequent mainshocks , using in essence the Reasenberg and Jones (1990) approach (https://earthquake.usgs.gov/earthquakes/eventpage/ci38457511/oaf/commentary). Immediately after the mainshocks, this model estimated the weekly probability of one quake being followed by a second mainshock of equal or larger magnitude at about 9%

(Michael et al., 2020; Hardebeck et al., 2019). This figure was higher than the default value of 5% obtained when using the standard Reasenberg and Jones (1990) parameter, because of the higher than average aftershock productivity in the region (Hardebeck et al., 2019). Just one day later, a Mw7.1 earthquake struck (at 8:20 p.m. local time on 6 July, or 03:20 UTC) at a distance of about 7 km.

The aforementioned probabilities of a subsequent larger earthquake occurring, as is common in California, were also cited in public. For example, after the second event Dr Lucy Jones tweeted: "*So the M6.4 was a foreshock. This was a M7.1 on the same fault as has been producing the Searles Valley sequence. This is part of the same sequence.*" This was followed by: "*You know we say 1 in 20 chance that an earthquake will be followed by something bigger? This is that 1 in 20 time.*" And then: *Yes, we estimate that there's about a 1 in 10 chance that Searles Valley will see another M7. That is a 9 in 10 chance that tonight's M7.1 was the largest*".

Here, we monitor fluctuating b-values and apply the FTLS in a near real-time application, comparing the FTLS forecast with currently used aftershock probabilities for California. We then compare the FTLS's performance with preliminary, revised and high-resolution datasets. A key aim in our research was to evaluate the feasibility of using b-value fluctuations for real-time hazard assessment.

Data and method

To compute a reliable and detailed b-value time series, we used a window approach, moving a window with a fixed sample size, event by event. In order to provide a prospective evaluation, the method strictly adheres to the approach used by Gulia and

Wiemer (2019; in review before the Ridgecrest mainshocks). The codes used can be downloaded from https://doi.org/10.3929/ethz-b-000357449. Here is a brief description of the approach and sequence-specific aspects. Using the quick focal mechanism of the Mw6.4 (GCMT, Dziewonski et al., 1981; Ekström et al., 2012) and the Wells and Coppersmith relationships (Wells and Coppersmith, 1994) corresponding to the tectonic style of the event -strike slip-, we built two possible fault planes, with a 1-km spaced grid. To decide quickly and automatically which was the most likely fault plane, we selected all events recorded in the sequences within the first hour and within a radius of 3 km from each grid point of the fault plane (from now on, the box), then selected the plane where most of the aftershocks occurred. While more sophisticated rupture planes using multiple fault segments, among other things, are often available for larger events within days, we opted to apply a simple, quick and robust approach that will facilitate independent testing as well as real-time application. We divided the dataset into two parts: a *pre*- and *post*-initiating event catalogue. The start time of the pre-catalogue depended on the quality and completeness of the local network: for the Californian seismicity we downloaded from the ANSS Comprehensive Earthquake Catalog (ComCat) via the FDSN web service (https://earthquake.usgs.gov/fdsnws/event/1/); we started the analysis of the background seismicity from 1981, when the network was greatly improved. The data were first downloaded on July 14th 2019, and then updated week by week.

The computation of b-values critically depends on correct estimates of the magnitude of completeness (Mc) (e.g. Mignan & Woessner, 2012). A specific Mc was assessed for each window (250-event-long) after a pre-cutting level, established using the Maximum Curvature method with a correction factor of 0.2 (Wiemer and Wyss, 2000). A b-value was then calculated for each window, applying the maximum likelihood method (Aki, 1965). We then defined a pre-event reference b-value, which was the median of all the single estimates preceding the Mw6.4.

For the post-event catalogue processing, we had to consider the temporal changes of the magnitude of completeness following a big event (Helmstetter, et al., 2006; Tormann et al., 2013), which can easily mask or bias the space-time b-value fluctuations. During the first hours after a large event, Mc typically changes by two orders of magnitude, resulting in a somewhat heterogeneous dataset. Changes in completeness are network-specific, but also depend on mainshock magnitude (Helmstetter et al., 2006). Our analysis of Ridgecrest's completeness (Figure 1) was fully consistent with previous experience, since Mc increased much more and over a longer time span after the Mw7.1 than after the Mw6.4 event. Specifically, after the Mw6.4 Mc increased from the background value (Mc=1.2) to about 1.8, before dropping back to a near-to-background value within 12 hours. After the Mw7.1 event, it increased to between 3.3 and 3.5, then recovered within three days to near-to-background values.

While we subsequently estimated Mc in each sample before computing a- and b-values, a common observation is that during periods of very strong gradients the Mc estimate is not conservative enough (e.g. Woessner and Wiemer, 2005), which potentially biases the analysis towards lower b-values. Based on our Mc analysis (see Figure 1), typically in keeping with such an analysis (e.g. Gulia et al., 2018), we therefore excluded from the dataset those events recorded during the initial, most heterogeneous period after the Mw6.4 and Mw7.1 events and introduced a minimum cut-off magnitude. In the aftermath of the Mw6.4, we excluded events occurring during the first 12 hours and pre-cut the dataset at M1.7. For the Mw7.1, we removed events occurring during the first 48 hours and pre-cut at M1.2 (see the shaded areas in Figure 1). This 'no-alert-time' is of course one of the limitations affecting the method's practical application: for the shorter this no-alert-time is, the more use FTLS decision support can be for practical mitigating actions. We subsequently tested the choice of these expert-selected parameters for sensitivity and confirmed that they did not critically influence our results. Subsequently we also used an alternative, revised and higher-resolution dataset (Shelly, 2020) to challenge and refine our analysis. Computing the percentage difference compared to the reference b-value was the final step. The values thus obtained allowed us to define the level of alert. If the percentage difference of the post-Mw6.4 event was plus or minus 10%, the alert was designated green or red, otherwise it was classified as orange.

Figure 3 schematically illustrates schematically the process of constructing b-value time-series and FTLS values for the Ridgecrest earthquake sequence. This figure contains the b-value difference in percentage respect to the reference value to allow comparison between the two fault planes. After the occurrence of the first event with M greater or equal than 6, we calculate the b-value time-series on its box, as explained in the previous lines, till the occurrence of a bigger event (Step 1 in Figure 3). Once a larger event occurs, we automatically refocus the analysis of b-value changes and FTLS on this new event, using the same procedures: We re-select the fault plane with the highest number of early aftershocks, re-select a new dataset and finally re-run the code that estimates the background b-value (Note: from 1981 to the M64, only, excluding the aftershocks and mainshock of the first sequence) and aftershock b-values (Step 2 in Figure 3). We normalize always the b-values relative to the background value, allowing for comparisons between different sequences in one timeline (Figure 3B). Refocusing on the new. Larger fault area is sensible, since the stress changes introduced by this event (larger and more recent) will dominate the changes in seismicity, this is now also the area of highest concern for larger events and the area the most seismicity for analysis. Note that in essence all steps are automated and follow the procedure by Gulia and Wiemer (2019), the only 'free' parameter is the starting date of the background b-value analysis (here: 1981).

Mapping of b-values to provide additional information on spatial changes was performed on a regular 1x1-km grid, selecting the closest 200 events within a maximum radius of 10 km. For the time series, we used the Maximum Curvature method (Wiemer and Wyss, 2000) for Mc, after pre-cutting the dataset at the same Mc level already adopted for the *pre* and *post* time period. We plotted the percentage difference of the post Mw≥6 events with respect to the b-value map obtained for the background (i.e. the time span from 1981 up to the last event preceding the Mw6.4).

The sub-catalogs generated for each fault plane and for the three different catalogs are provided as text files in the Supplement.

Results

Automatic fault selection

The Mw6.4 earthquake on 4 July in Ridgecrest ruptured two conjugate strike-slip faults, which intersected to form an 'T' shape. It took days before geodetic, seismic and relocated seismicity data provided an overall view of this complex sequence (Ross et al., 2019; Hobbs, 2019). By kinematically inverting for subevents using seismograms from the dense regional seismic network and global seismic stations, Ross et al. (2019)

identified three simultaneous subevents and hypothesised that the rupture had been a cascading phenomenon, rather than a single continuous process. The three identified subevents coincided with at least three faults: the 6-km-long northwest-trending fault that slipped first; then the rupture propagated over a short southwest-trending fault with only about 5 km of surface break, and finally the jump to a larger southwest-trending fault roughly 15 km long (Ross et al., 2019).

The FTLS method was developed to be applied in near real-time, when little other information apart from data from the focal mechanism and the automatically derived network catalogue is both known and publicly accessible. The seismic source used in our analysis is thus represented by a single plane. Following the method described by Gulia and Wiemer (2019), once the GCMT provided the focal mechanism, the algorithm built the two fault planes, centred on the local hypocentre catalogue (see https://www.fdsn.org/networks/). Between the two fault planes, the one with the largest number of early aftershocks within a 3 km radius was selected as the likely fault plane. For Mw6.4, this purely statistical method chose the northwest-trending fault plane (Figure 2) that represented the initial rupture, in the process described by Ross et al. (2019), and is the one aligned with the eventual Mw7.1 hypocentre. The background or reference b-value for this box containing 1275 events above M1 since 1981 is b = 0.97 (blue symbols in Figure 4).

Seismicity preceding Mw7.1

Figure 4A shows the b-value time-series. All b-values after the Mw6.4 event are substantially lower than the background b-value. A comparison of the Frequency Magnitude Distributions (FMDs) of events occurring between 4 and 6 July is in Figure

4B. During the time interval between the two big events, the b-value decreases from 0.97 to 0.75, a decrease by 23%, resulting in a red FTSL status (Figure 4B). We also calculated the respective daily probability (Pr) commonly derived by extrapolating the observed frequency-magnitude distribution to an Mw6.4 event or larger earthquake (Figure 5C). These probabilities reached a peak value of 66% on 5 July a value about one order of magnitude larger than the aforementioned ones derived by the USGS (https://earthquake.usgs.gov/data/oaf/overview.php, 5% using default values, 9% using sequence specific values according to Michael et al., (2020) and Hardebeck et al., (2019)).

Next, we mapped the spatial distribution of the differential b-value (i.e. the background b-value map subtracted from the current episode map) to infer information on the likely nucleation region of a subsequent mainshock (Figure 6A). The expectation described by Gulia and Wiemer (2019) is that a subsequent mainshock would nucleate near the strongest b-value decrease, in our conceptual model represented by high stress asperities. In the case of the Ridgecrest sequence, this low b-value patch locates to the NW of the Mw6.4 epicentre and corresponds closely to the location of the subsequent Mw7.1 quake on 6 July (marked in Figure 6B).

Seismicity following the Mw7.1

We then analysed b-value evolution over time in the Mw7.1 source volume, constructed following the same procedure as described above for the Mw6.4 event. We also determine a new background b-value of 0.87 for this much larger source volume, compared to the volume of Mw6.4. The b-value time series, plotted in Figure 3 and starting two days after the Mw7.1 earthquake, indicated a general increase from the normalised background value of more than 10%, reaching a peak of 26% within the first week (Figure 3C). This qualified it for green FTLS status and suggested that the chance of a subsequent even larger event was lower than average. Figure 4C shows the FMD's of the background (b=0.87) compared to the aftershocks (b = 1.1). We again calculated the probability of a subsequent event of equal or larger magnitude at 0.4% per day two days after the event and falling to 0.004% per day in subsequent weeks. These values were one order of magnitude lower than the USGS aftershock probabilities communicated during the sequence. The differential b-value map for events occurring in the first week with respect to their background (Figure 6B) indicated a general rise in b-values throughout the region.

Revised and high-resolution datasets

While this manuscript was under review, revised GCMT and ComCat catalogues (downloaded on 21 January 2020) became available, so we repeated our analysis, to compare it with the FTLS's performance using near real-time data. The revised GCMT focal mechanisms, available online since 8 November 2019, are very similar to their quick equivalents (Table 1), both in orientation and dip. We then re-computed the fault planes centred on the hypocentres of the two mainshocks (Mw6.4 and Mw7.1) for the revised ComCat catalogue as well as for the high-resolution catalogue compiled by Shelly (2020).

Minor displacement (by approx. 0.2 km) of the epicentre of the 4 July mainshock in the revised ComCat catalogue makes the revised boxes imperceptibly different with respect to their quick counterparts (Table 2). The overall completeness of the catalogues remains largely unchanged. Consequently, the result showed the same almost imperceptible difference, with the overall b-value during the time interval between the two biggest events rising from 0.75 to 0.76, and the red alert from -23% to -22%. After the mainshock, we obtained the same b-values and the same green alert (+26%).

In addition, Shelly (2020) published a revised, higher-resolution catalogue containing 34,000 events during the period 4-16 July for the Ridgecrest sequence, allowing us for the first time to evaluate the b-value evolution and FTLS performance with a partially independently calculated and presumably higher-quality dataset. This earthquake catalogue is based on cross-correlation analysis of continuous wave-forms and according to Shelly (2020) substantially more complete in magnitude, more consistent through time and more precise in hypocentres. Shelly (2020) points out that cross-correlation is not well suited for relocating M>5 earthquakes, especially the two events with the highest magnitudes, because its wave forms are too dissimilar to those of smaller events. Indeed, in this dataset, the two epicentres roughly correspond to the location provided by USGS, albeit having different depths, with the Mw6.4 deeper (from 10.5 to 15 km) and the Mw7.1 shallower (from 8 to 3 km). For this reason, we use the same source volumes determined for the previous analysis (i.e. revised GCMT moved to the ComCat hypocentre).

This catalogue contains only 38 events preceding the Mw6.4 quake, not enough to establish a reference b-value for the FTLS, so we used the revised ComCat catalogue to estimate that value for the boxes of the Mw6.4 and Mw7.1 mainshocks. As shown in Shelly (2020), the cross-correlation analysis substantially lowers these events' overall magnitude of completeness, a finding supported by our Mc(t) analysis (Figure 8). The Shelly catalogue reaches an Mc of about 0.7, roughly half degree of magnitude lower than the standard ComCat catalogue. However, the increase in Mc immediately after the mainshock is almost as high (rising to roughly Mc = 3.0-3.5), but completeness recovers faster and more systematically. Completeness for M1.5 is reached 24 hours earlier than using standard datasets (Figure 7). This improvement is extremely important for our approach, but also for other real-time methods used to assess time-dependent earthquake probabilities.

Using the Shelly catalogue, we repeated the b-value analysis using the same time-windows but lower completeness and found almost identical results (-21% after the Mw6.4 and +29% after the Mw7.1), confirming that the results based on near-real time data are in line with the more homogeneous, higher-quality catalogue. To exploit the possible improvements of higher quality data for aftershock hazard assessment, we then moved the start of our analysis closer to the mainshock origin time, thus shortening our no-alert-time. After the Mw6.4 earthquake, we were able to cut this no-alert-time from 12 hours to just one, and after the Mw7.1 from 48 hours to 24 hours (using Mc pre-cuts of 1.5 in both cases). The time series of b-values is shown in Figure 8. The overall trend, the b-values themselves and FTLS status all remain unchanged. However, it is worth noting that we can establish a low b-value after the M6.4 with just one hour of no-alert-time when high-quality data is available.

Sensitivity analysis

Our method contains essentially three free parameters that we determined based on data analysis and expert choices: 1) the magnitude of completeness, 2) the no-alert-time, and the 3) the sample size analysed. The first two we have determined based on the completeness analysis (Figures 1 and 6), the last is a commonly used value in studies. We introduce a novel sensitivity analysis to evaluate the impact of the changes on the result our study. We scan systematically the parameter space of the pre-cut Mc and no-alert-time parameters. The results shown in Figure 9 for the revised ComCat and the Shelly catalogue are fully consistent with the previous interpretations: For all choices of Mc and no-alter times, there is a string decrease in b-value (red colours and red FTLS status) subsequent to the M6.4. Following the M7.1, the picture is somewhat different: for value at or below the estimated completeness (black dashed line in Figure 9), there is decrease in b-value – an expected bias due to incompleteness. Above Mc, however, green colours indicate an increase in b-value and green FTLS status.

Seismic sequence in 1995

In 1995, an Mw5.8 earthquake occurred in the same region, a few kilometres away from the Mw7.1 (Figure 2A). That event was not followed by a larger one. For comparison, we also applied the FTLS approach to this sequence, too. Figure 10 shows the FMDs and time series relative to the 1995 sequence, indicating a roughly 30% increase in the b-value, resulting in a correct green traffic light classification. This result suggests that the FTLS approach can also be extended to events of smaller magnitude than the currently used Mw≥6.0 reference.

Discussion and conclusion

Our analysis shows that the Ridgecrest earthquake sequence not only impacted the seismic activity rate, increasing the productivity of earthquakes near the fault by between 3 and 5 orders of magnitude; it also changed the relative size distributions, the b-values, in both space and time. This should come as no surprise, since the size distribution is known to be sensitive to the applied shear stress on faults (e.g. Goebel et al., 2013), and also to depend on location (e.g. Tormann et al., 2015). Thus, b-values are linked to the seismotectonic context and evolution of events, but they also constitute an important factor influencing the probability of a subsequent larger event. The FTLS concept introduced by Gulia and Wiemer (2019) exploits the systematic differences in b-values observed between the majority of aftershock sequences that will normally decay over time and the small percentage of sequences that are followed by an even larger event. The FTLS method and codes were developed in the first half of 2019, but only published in October 2019 (Gulia and Wiemer, 2019). The Ridgecrest sequence, representing one of the best-monitored large mainshock-aftershock sequences, presented us with an ideal opportunity to test the FTLS hypothesis and developed software. The analysis presented is here is not yet a truly prospective, real-time application, because we were (and still are not) set up computationally and, to a certain extent, methodologically to conduct such an urgently needed but challenging test. However, it is meaningful in a pseudo-prospective sense, an analysis that reproduces real-time condition. Our pseudo-prospective study is, however, more rigours and we would argue more meaningful than typical such studies, because the method and codes used to conduct the automatic analysis have been published before and were here used unchanged from the version of the method submitted for publication. In other words, they could not have been optimised to provide the best outcome for our hypothesis.

The results obtained and presented in this paper support the FTLS hypothesis: seismicity following the Mw6.4 event showed a substantially lower b-value (a drop of 23%, Figure 4), resulting in its correct red traffic light designation. The b-value also rose by 26% after the Mw7.1 quake, resulting in a correct green classification. This adds one correct positive and one correct negative to the confusion matrix analysis presented in Gulia and Wiemer (2019), increasing the accuracy assessment to above 96%. A correct green traffic light was also attributed after the 1995 Mw5.8 earthquake. Since the FTLS hypothesis is proposed and evaluated for events with a magnitude of 6.0 and above, the Mw5.8 results are not factored into the (retrospective) error matrix score.

The FTLS hypothesis itself needs to be further tested, and the error matrix approach carried out on future sequences in a fully prospective, independently conducted way. Such tests are now planned as part of the Collaboratory for the Study of Earthquake Predictability (CSEP, Schorlemmer et al., 2018), financed by the European RISE project (www.rise-eu.org). In addition, the observed changes in b-values can and should also be directly converted into time-dependent earthquake probabilities, as shown in Tormann et al. (2016) and Gulia et al. (2016) for example. These probabilities are also reported for the Ridgecrest sequences (Figure 5), which are very consistent with the FTLS results and will be tested in comparison to other models, such as the Reasenberg-Jones or ETAS models. Note that the FTLS green alert may turn out to be the most important one in terms of its practical implications, for the vast majority (80%) of all sequences will fall into this category, and knowing that a larger event is unlikely will be extremely valuable information. Indeed, after the M7.1, we estimate about a factor 10 lower probability for a subsequent larger one that the standard USGS model.

Naturally, in principle it would be great to extend the FTLS model to smaller mainshocks, because more data could be used to test the hypothesis. However, the data would have to be of very high quality and their inclusion would probably increase the uncertainty of the analysis. The smaller size of the fault planes involved in such events (an M5.5 source, for example, would be about 6 km long) would make it more challenging to identify the active fault. Because smaller mainshocks will generally result in fewer aftershocks, the spatiotemporal resolution of b-values is reduced and the useful magnitude range between the largest events and Mc decreases, making it more difficult to establish reliable b-values. Probably scaling works in such a way that we would have to select events even closer to the mainshock fault only, which in turn makes pinpointing the location even more challenging. Also, sample sizes may be too small for robust analyses. Similarly, the relevant background (i.e. the reference level) would be even more local and thus harder to determine. In addition, the Coulomb stress and failure modelling in Gulia et al. (2018) suggests that the amplitude of the b-value increase is magnitude-dependent, so it is unclear whether b-value transients are scale-invariant. So, it needs to be explored whether the evaluation of the FTLS hypothesis can be extended to smaller events, but this will necessitate a very thorough analysis of any uncertainties and their influence on the stability of the analysis. An analysis of that kind is beyond the scope of this Ridgecrest case study.

The spatial patterns of changes in b-values have been proposed as additional information on the future location of subsequent larger events, and here too the Ridgecrest case study is well in line with this loosely formulated and as yet not formally tested hypothesis: the Mw7.1 event occurred near the area of the steepest b-value decrease (Figure 6). More research and testing in needed to integrate this spatial information into aftershock forecasting in an automate way, for now we consider the information contained in b-value or earthquake probability maps additional information for experts to be considered.

Establishing with confidence a b-value time series critically hinges on the quality of the seismic network, and judging from our analysis the southern California network performed extremely well (Figure 1) in near real-time (much of our analysis was in fact conducted within days of the Mw6.4 event). The magnitude of completeness rapidly decreased (Figure 1) and the frequency magnitude distribution (Figures 4 and 9) is among the best we have ever analysed, closely following a linear Gutenberg-Richter distribution and leading within hours to reliable observations of b-value changes. Based on our experience, the differential b-value maps computed (Figure 6) are also very reliable. Progress made in station coverage and automated network processing approaches are clearly delivering very rapidly high-quality data that are useful for scientific analysis and risk assessment. Further improvements using advanced automated post-processing methods may be feasible and desirable to decrease no-alert-time. Our test using the higher-resolution catalogue provided by Shelly (2020) supports this (Figures 8 and 9). The catalogue confirms every aspect of the results obtained using ComCat real-time data, so we consider the likelihood of data imperfection influencing our analysis to be very low. Equally importantly, the Shelly catalogue allows us to reduce no-alert-time to just one hour. Since the approach implemented by Shelly in principle reveals the real-time capabilities of seismic networks in the not-too-distant future, we suggest that it may be possible to produce an FTLS assessment within just one or a few hours. We also suggest that the sensitivity analysis to Mc and no-alter-time we introduce in Figure 9 is a powerful tool to quickly evaluate the robustness of an FTLS results. This may be also in real-time a graphical representation a seismologist wants to consult in a crisis to ensure the results are not critically dependent on the choice of parameters,

The FTLS hypothesis is quite new, and while the successful Ridgecrest case provides additional support for it, in our view it is too early to use it routinely for making decisions about civil protection or public communications. More extensive sensitivity and robustness studies are needed, the hypothesis should be independently evaluated by other research teams and the hypothesis needs to be formally tested. There are plans for this, but it will take time. At the same time, numerical modelling may allow the formulation of a better physical understanding and maybe enhanced forecasting abilities. These efforts will take time, but given the potential implications and greater understanding, we consider them highly worthwhile.

Data and resource

The ComCat catalogue by USGS was downloaded from the website https://earthquake.usgs.gov/fdsnws/event/1/catalogs and ZMAP (Reyes and Wiemer, 2019).

Acknowledgment

Figures were made using Generic Mapping Tools (www.soest.hawaii.edu/gmt; Wessel and Smith, 1998). The authors thank the Editors, Allison Bent and Anastasia Pratt, Andy Michael and one anonymous reviewer for helping in improving and clarifying the manuscript. This study was supported by the Real-time Earthquake Risk Reduction for a Resilient Europe (RISE) project, funded by the European Union's Horizon 2020 research and innovation programme under grant agreement no. 821115.

References

Aki, K. (1965). Maximum likelihood estimate of b in the formula log n = a _ bM and its confidence limits, *Bull. Earthquake Res. Inst. Univ. Tokyo*, 43, 237–239.

Amitrano, D. (2003). Brittle-ductile transition and associated seismicity: Experimental and numerical studies and relationship with the b-value. *J. Geophys. Res.*, 108(B1), 2044. https://doi.org/10.1029/2001JB000680

Dieterich, J. H., Cayol, V., & Okubo, P. G. (2000). The use of earthquake rate changes as a stress meter at Kilauea volcano. Nature, 408(6811), 457–460. https://doi.org/10.1038/35044054

Dziewonski, A. M., Chou, T. A., and Woodhouse, J. H. (1981). Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *J. Geophys. Res*., 86, 2825–2852. https://doi.org/10.1029/JB086iB04p02825

Ebel, J. E., Bonjer, K.-P., & Oncescu, M. C. (2000). Paleoseismicity: Seismicity evidence for past large earthquakes. *Seismol. Res. Lett.*, 71(2), 283–294. https://doi.org/10.1785/gssrl.71.2.283

Ekström, G., Nettles, M., and Dziewoński, A. M. (2012). The global CMT project 2004–2010: Centroid-moment tensors for 13,017 earthquakes. *Phys Earth Planet In*, 200–201, 1–9.

Goebel, T. H. W., Schorlemmer, D., Becker, T. W., Dresen, G., and Sammis, C. G. (2013). Acoustic emissions document stress changes over many seismic cycles in stick-slip experiments. *Geophys Res Lett*, 40, 2049–2054. https://doi.org/10.1002/grl.50507

Gulia, L., Rinaldi, A. P., Tormann, T., Vannucci, G., Enescu, B., and Wiemer, S. (2018). The effect of a mainshock on the size distribution of the aftershocks. *Geophys Res Lett*, 45. https://doi.org/10.1029/2018GL080619

Gulia L. and Wiemer S., (2019). Real-time discrimination of earthquake foreshocks and aftershocks. *Nature,* 574, 193-199.

Gulia, L., and Wiemer, S. (2010). The influence of tectonic regimes on the earthquake size distribution: A case study for Italy. *Geophys Res Lett*, 37, L10305. https://doi.org/10.1029/2010GL043066

Gulia, L., Tormann, T., Wiemer, S., Herrmann, M., & Seif, S. (2016). Short-term probabilistic earthquake risk assessment considering time dependent b values. *Geophys Res Lett*, 43, 1100–1108. https://doi.org/10.1002/2015GL066686

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

Hainzl, S. (2016). Rate-Dependent Incompleteness of Earthquake Catalogs, Seismol. Res. Lett., 87(2A), 337–344.

Hardebeck J.L, Llenos A.L., Michael A.J., Page M.T. and van der Elst N. (2019). Updated California Aftershock Parameters, *Seismol. Res. Lett., 90(1), 262-270.*

Helmstetter, A., Kagan, Y. Y. and Jackson, D. D. (2006). Comparison of short-term and time-independent earthquake forecast models for Southern California. *Bull. Seismol. Soc. Am*., 96(1), 90–106. https://doi.org/10.1785/0120050067

Hobbs, T.E. (2019), Can science tell when a large earthquake will be followed by an even larger one? Temblor, http://doi.org/10.32858/temblor.056

Ishimoto, M., and Iida, K., (1939). Observations sur les séismes enregistrés par le microséismographe construit dernièrement, *Bull. Earthqu. Res. Int.*, 17, 443-478.

Jordan, T. H., Marzocchi, W., Michael, A. J. and Gerstenberger, M. C. (2014). Operational earthquake forecasting can enhance earthquake preparedness, Seismol. Res. Lett., 85(5), 955–959.

Marzocchi W., Taroni M., Falcone G. (2017), Earthquake forecasting during the complex Amatrice-Norcia seismic sequence, Science advances 3 (9), e1701239

Michael, A. J., McBride, S. K. , Hardebeck, J. L. , Barall, M. , Martinez, E. , Page, M. T. , van der Elst, N. ,. Field, E. H , Milner, K. R. and Wein, A. M. (2020). Statistical Seismology and Communication of the USGS Operational Aftershock Forecasts for the 30 November 2018 Mw 7.1 Anchorage, Alaska, Earthquake, *Seismol. Res. Lett*., 91 (1): 153–173

Mignan, A., and Woessner, J. (2012). Estimating the magnitude of completeness for earthquake catalogs, Community Online Resource for Statistical Seismicity Analysis. https://doi.org/10.5078/corssa-00180805

Narteau, C., Byrdina, S., Shebalin, P. and Schorlemmer D., (2009). Common dependence on stress for the two fundamental laws of statistical seismology, Nature, 462, 642–645, doi:10.1038/nature08553.

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. Journal of the American Statistical Association, 83(401), 9–27. https://doi.org/10.1080/01621459.1988.10478560

Okada, Y. (1992). Internal deformation due to shear and tensile faults in a half-space, *Bull. Seismol. Soc. Am*., 82, 1018–1040.

T Omi, Y Ogata, K Shiomi, B Enescu, K Sawazaki, K Aihara (2019). Implementation of a Real-Time System for Automatic Aftershock Forecasting in Japan, *Seismol. Res. Lett.,* 90 (1), 242-250

Omori, F. (1895). On the aftershocks of earthquakes. Journal of the College of Science, Imperial University of Tokyo, 7, 111–200.

Page, M. T., N. van der Elst, J. Hardebeck, K. Felzer, and A. J. Michael (2016). Three ingredients for improved global aftershock forecasts: Tectonic region, time-dependent catalog incompleteness, and intersequence variability, Bull. Seismol. Soc. Am. 106, 2290–2301, doi: 10.1785/0120160073.

Petruccelli A. Schorlemmer D., Tormann T., Rinaldi A.P., Wiemer S., Gasperini P. and G. Vannucci (2019a). The influence of faulting style on the size-distribution of global earthquakes. *Earth Plan. Sci. Lett.,* 527, doi: 10.1016/j.epsl.2019.115791

Petruccelli A., Gasperini P., Tormann T., Schorlemmer D., Rinaldi A.P., Vannucci G. and S. Wiemer (2019b). Simultaneous dependence of the earthquake-size distribution on faulting style and depth. *Geophys. Res. Lett*., doi: 10.1029/2019GL083997

Reasenberg, P. A. and Jones, L. M., (1990). California aftershock hazard forecast. *Science* 247, 345–346.

Reyes, C. G. and Wiemer, S., (2019). ZMAP7, a refreshed software package to analyze seismicity. *Geophysical Research Abstracts*, Vol. 21, EGU2019-13153, 2019

Ross E. Z., Idini B., Jia Z., Stephenson O.L., Zhong M., Wang X., Zhan Z., Simons M., Fielding E.J., Yun S., Hauksson E., Moore A.W., Liu Z., Jung J, (2019). Hierarchical interlocked orthogonal faulting in the 2019 Ridgecrest earthquake sequence, Science, 366, p.346-351. DOI: 10.1126/science.aaz0109

Scholz, C. H. (1968). The frequency-magnitude relation of microfracturing in rock and its relation to earthquakes. *Bull. Seism. Soc. Am.*, 58, 399–415.

Scholz CH (2015). On the stress dependence of the earthquake b value, Geophys Res Lett 42:1399–1402. doi:10.1002/2014GL062863

Schorlemmer, D., Werner, M., Marzocchi, W., Jordan, T., Ogata, Y., Jackson, D., Mak, S., Rhoades, D., Gerstenberger, M., Hirata, N., Liukis, M., Maechling, P., Strader, A., Taroni, M., Wiemer, S., Zechar, J., Zhuang, J. (2018). The Collaboratory for the Study of Earthquake Predictability: Achievements and Priorities. *Seismol Res Lett*, 89. doi: 89. 10.1785/0220180053.

Shelly, D. R. (2020). A High-Resolution Seismic Catalog for the Initial 2019 Ridgecrest Earthquake Sequence: Foreshocks, Aftershocks, and Faulting Complexity. *Seismol. Res. Lett.,* in press., 1–8, doi: 10.1785/0220190309.

Shi, Y. and Bolt, B. (1982). The standard error of the magnitude-frequency b value. *Bull. Seism. Soc. Am.*, 72, 1677–1687.

Staudenmaier N., Tormann T., Edwards B., Mignan A. and Wiemer S., (2019). The frequency-size scaling of non-volcanic tremors beneath the San Andreas Fault at Parkfield: Possible implications for seismic energy release. *Earth Plan. Sci. Lett.* 516:77-107. DOI: 10.1016/j.epsl.2019.04.006

Stein, R. S. (1999). The role of stress transfer in earthquake occurrence. *Nature*, 402(6762), 605–609. https://doi.org/10.1038/45144

Tormann, T., Wiemer, S., Metzger, S., Michael, A., and Hardebeck, J. L. (2013). Size distribution of Parkfield's microearthquakes reflects changes in surface creep rate. *Geophys J Int*, 193(3), 1474–1478. https://doi.org/10.1093/gji/ggt093

Tormann, T., Enescu, B., Woessner, J. and Wiemer, S., (2015). Randomness of megathrust earthquakes implied by rapid stress recovery after the Japan earthquake. *Nat. Geosci.* 8, 152–158.

Tormann, T., Wiemer S., Enescu B. and Woessner, J. (2016), Normalized rupture potential for small and large earthquake along the Pacific Plate off Japan. Geophys. Res. Lett., 43, 7468-7477.

Wells, D. L. and Coppersmith, K. J., (1994). New Empirical Relationships among Magnitude, Rupture Length, Rupture Width, Rupture Area, and Surface Displacement. *Bull. Seismol. Soc. Am.* 84, 974–1002.

Wessel, P. and Smith W.H.F. (1995). New version of the generic mapping tools released, *Eos. Trans*., 76, 329.

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

Wiemer, S., (2001). A software package to analyze seismicity: ZMAP. *Seismol Res Lett*, 72(3), pp.373-382. doi: https://doi.org/10.1785/gssrl.72.3.373

• **Email address of each author**

L. Gulia: Laura Gulia (lgulia@ethz.ch, laura.gulia@unibo.it); University of Bologna, Department of Physics and Astronomy, Viale Berti Pichat 8, Bologna.

S. Wiemer: Stefan Wiemer (stefan.wiemer@sed.ethz.ch); Swiss Seismological Service, ETH, NO H61, Sonneggstrasse 5, CH-8092 Zurich

G. Vannucci: Gianfranco Vannucci (gianfranco.vannucci@ingv.it); Istituto Nazionale di Geofisica e Vulcanologia, via Donato Creti 12, 40128 Bologna

Tables

Table 1: Nodal planes (np1 and 2) of the quick and revised GCMT catalogue for the two events on 4 July 2019, 17:33 UTC (Day 04) and 6 July 2019, 03:19 UTC (Day 06)

GCMT	Day	Strike np1	Dip np1	Rake np1	Strike np1	Dip np1	Rake np1	Length (km)	Width (km)
Quick	04	228	81	0	318	90	-171	27.28	9.65
	06	322	78	-177	231	87	-12	61.9	13.79
Revised	04	227	86	3	137	87	176	26.84	9.58
	06	321	81	180	51	90	9	61.3	13.73

Table 2: Vertices of the fault planes (FP1 and FP2) corresponding to the nodal planes in Table 1. See Table 1 for details of the symbols.

GCMT	Day	Lon FP1 (deg.)	Lat FP1 (deg.)	Depth FP1 (km)	Lon FP2 (deg.)	Lat FP 2 (deg.)	Depth FP2 (km)
Quick	04	-117.3882	35.7822	5.9452	-117.4049	35.614	5.8858
		-117.6127	35.6181	5.9452	-117.6071	35.7963	5.8858
		-117.6238	35.6281	15.4748	-117.6071	35.7963	15.5342
		-117.3993	35.7923	15.4748	-117.4049	35.614	15.5342

	06	-117.4007	35.5422	1.2576	-117.3302	35.9421	1.1164
		-117.823	35.9809	1.2576	-117.8634	35.5918	1.1164
		-117.798	35.9968	14.7424	-117.8684	35.5969	14.8836
		-117.3756	35.5581	14.7424	-117.3353	35.9472	14.8836
Revised	04	-117.3926	35.7854	5.7213	-117.6032	35.7951	5.7162
		-117.61	35.6208	5.7213	-117.4004	35.6186	5.7162
		-117.6151	35.6252	15.2787	-117.4045	35.6155	15.2838
		-117.3977	35.7898	15.2787	-117.6072	35.7921	15.2838
	06	-117.3948	35.5492	1.2208	-117.8633	35.596	1.1363
		-117.8224	35.9776	1.2208	-117.3353	35.943	1.1363
		-117.8039	35.9898	14.7792	-117.3353	35.943	14.8637
		-117.3763	35.5614	14.7792	-117.8633	35.596	14.8637
Shelly, 2020	04	-117.3874	35.7885	10.2753	-117.598	35.7982	10.2702
		-117.6048	35.6238	10.2753	-117.3952	35.6216	10.2702
		-117.6098	35.6282	19.8327	-117.3993	35.6185	19.8378
		-117.3924	35.7929	19.8327	-117.602	35.7951	19.8378
	06	-117.3896	35.5515	-3.5382	-117.8582	35.5984	-3.6227
		-117.8172	35.9799	-3.5382	-117.3302	35.9453	-3.6227
		-117.7987	35.9921	10.0202	-117.3302	35.9453	10.1047
		-117.3711	35.5637	10.0202	-117.8582	35.5984	10.1047

Figure captions

Figure 1 –A): time/magnitude plot for the events following the Mw 6.4 on 4 July. Shaded areas indicate times when the dataset was least complete. B): time series of the magnitude of completeness (red lines) estimated using the maximum curvature method for samples containing 300 events, moved through the data in overlapping windows. Grey lines represent uncertainty estimates obtained by bootstrapping.

Figure 2– A) Seismicity map with the events (white stars) on 4 July - Mw 6.4 (M64), 6 July - Mw 7.1 (M71) and subsequent events in black and red respectively. The two green fault planes indicate the Mw 6.4 GCMT focal mechanism, with strike and dip directions. B) 3-D view of Figure 2a, from a 200° azimuth and 40° elevation.

Figure 3 – A) Schematic representation of the single time-series obtained on the M64 and M7.1 fault planes and B) the summary one with the 2 fault planes in the near-real-time analysis of the Ridgecrest earthquake sequence.

Figure 4 – Performance of the FTLS in near real-time. A) b-value time series for the Mw 7.1 sequence superimposed on the FTLS assessment (Wiemer and Gulia, 2019); the blue dashed line is the reference b-value; the black dashed vertical lines indicate Mw 6.4 and Mw 7.1 respectively. The black rectangle zooms in on the time series in the interval between the two M>6 events. All the estimates are below the reference value. Grey indicates uncertainty (one standard deviation by Shi and Bolt, 1982). B) Frequency-magnitude distributions for the source of the Mw 6.4 event for two time periods: background in blue, time between the two Mw>6 events in red. C) Frequency-magnitude distributions for the source of the Mw 7.1 event for two time periods: background in blue, maximum b-value reached in the first week of aftershocks.

Figure 5 A-F: Daily time series on the fault planes of the two major events. A-C) Fault plane of the Mw 6.4 event: b-value (A), daily a-value (B) and daily probability of a Mw 6.4+ (C). D-F)

Fault plane of the Mw 7.1 event: b-value (D), daily a-value (E) and daily probability of a Mw 6.4+ (F). The blue dashed lines represent the mean value of all the background estimates.

Figure 6 – Mapped b-values with the difference in percentage with respect to the background for two different periods: A) between Mw 6.4 and Mw 7.1; B) the first week after Mw 7.1. The original maps were produced by ZMAP (Wiemer, 2000; Reyes and Wiemer, 2019) and post-processed in the Matlab using GMT, generic mapping tools (http://gmt.soest.hawaii.edu)..

Figure 7 – Time series of the magnitude of completeness (red lines) in the catalogue by Shelly (2020) estimated using the maximum curvature method for samples containing 300 events, moved through the data in overlapping windows. The grey lines represent uncertainty estimates obtained by bootstrapping.

Figure 8 – Performance of the FTLS with the high-resolution catalogue by Shelly (2020): A) b-value time series for the Mw 7.1 sequence superimposed on the FTLS assessment (Wiemer and Gulia, 2019); the blue dashed line is the reference b-value; the black dashed vertical lines indicate Mw 6.4 and Mw7.1 respectively. Grey indicates uncertainty (one standard deviation by Shi and Bolt, 1982). B) Frequency-magnitude distributions for the source of the Mw 6.4 event for two time periods: background in blue, time between the two Mw>6 in red. c) Frequency-magnitude distributions for the source of the Mw 7.1 event for two time periods: background in blue, maximum b-value reached in the first week of aftershocks.

Figure 9 – A-B) Sensitivity analysis on no-alert-time and completeness. Color-coded is the b-value difference in percentage with respect to the reference b-value as a function of magnitude cut-off and time after the Mw 6.4 (left) and Mw 7.1 (right). We always analyzed the first 300 events above this magnitude and after this time. Black dashed line: A) the estimated magnitude of completeness for the ComCat catalog reported in Figure 1; gray dashed line: the same with

the 0.2 correction factor, as adopted in our modeling; B) the estimated magnitude of completeness for the high-resolution catalogue by Shelly (2020) reported in Figure 7; gray dashed line: the same with the 0.2 correction factor, as adopted in our modeling.

Figure 10 – A) Frequency-magnitude distributions for the Mw 5.8 sequence in 1995: in blue, the background b-value (1981-1995) and in green the highest b-value reached by the aftershocks during the first week after Mw 5.8; B) b-value time series for the same sequence. The blue dashed line represents the reference b-value (see Data and method).