Characterization of Induced micro-seismicity associating with one hydraulic fracturing experiment near the San Andreas fault, Central California

Report for SCEC Award #15137
Submitted March 20, 2016

Investigators: Chen Ji (UCSB).

I. Project Overview

A. Abstract

In the box below, describe the project objectives, methodology, and results obtained and their significance. If this work is a continuation of a multi-year SCEC-funded project, please include major research findings for all previous years in the abstract. (Maximum 250 words.)

We apply match filter (e.g., Shelly et al., 2007) and match&locate (M&L, Zhang and Wen, 2015) to study the micro-seismicity induced by a hydraulic fracture experiment in a central California site only 8 km away from the San Andreas fault. Two modifications have been made to improve the detection capability of M&L algorithm and reduce its computational cost. Using a 16-geophones array that located less than 400 m away from the treatment well, the new M&L algorithm detects 12,944 earthquakes within a 7.5 hour injection period, 12 times more than what were registered in the industry catalog. The events span a magnitude range from -4.17 to -0.77 with a magnitude completeness M_c of -3.0, in contrast with M_c -2.2 of the industry catalog. The new catalog reveals a magnitude-frequency distribution during the injection period that clearly deviates from the Gutenberg-Richter relation.

B. SCEC Annual Science Highlights

Each year, the Science Planning Committee reviews and summarizes SCEC research accomplishments, and presents the results to the SCEC community and funding agencies. Rank (in order of preference) the sections in which you would like your project results to appear. Choose up to 3 working groups from below and re-order them according to your preference ranking.

Seismology
Computational Science
Central California Seismic Project (CCSP)

C. Exemplary Figure

Select one figure from your project report that best exemplifies the significance of the results. The figure may be used in the SCEC Annual Science Highlights and chosen for the cover of the Annual Meeting Proceedings Volume. In the box below, enter the figure number from the project report, figure caption and figure credits.

Figure 2. Magnitude-frequency distribution of earthquakes occurring during a 7.5 hour period of hydraulic fracture injection. GR denotes Gutenberg and Richter relation. Caputo and Utsu denote two non-GR models that match this distribution well.

D. SCEC Science Priorities

In the box below, please list (in rank order) the SCEC priorities this project has achieved. See *https://www.scec.org/research/priorities* for list of SCEC research priorities. *For example: 6a, 6b, 6c*

2f, 2a

E. Intellectual Merit

How does the project contribute to the overall intellectual merit of SCEC? *For example: How does the research contribute to advancing knowledge and understanding in the field and, more specifically, SCEC research objectives? To what extent has the activity developed creative and original concepts?*

The work further improves a newly developed earthquake detection and relocation algorithm. The result reveals that during the injection period, the magnitude-frequency distribution of induced seismicity deviates from Gutenberg-Richter statistics.

F. Broader Impacts

How does the project contribute to the broader impacts of SCEC as a whole? *For example: How well has the activity promoted or supported teaching, training, and learning at your institution or across SCEC? If your project included a SCEC intern, what was his/her contribution? How has your project broadened the participation of underrepresented groups? To what extent has the project enhanced the infrastructure for research and education (e.g., facilities, instrumentation, networks, and partnerships)? What are some possible benefits of the activity to society?*

The software package can be used to study the earthquakes in other region. The research improves our understanding to the seismicity induced by hydraulic fracture experiments, an area with extensive public attention. The fund is used to support the research of graduate student Trevor Smith.

G. Project Publications

All publications and presentations of the work funded must be entered in the SCEC Publications database. Log in at *http://www.scec.org/user/login* and select the Publications button to enter the SCEC Pubications System. Please either (a) update a publication record you previously submitted or (b) add new publication record(s) as needed. If you have any problems, please email *web@scec.org* for assistance.

II. Technical Report

A. Introduction

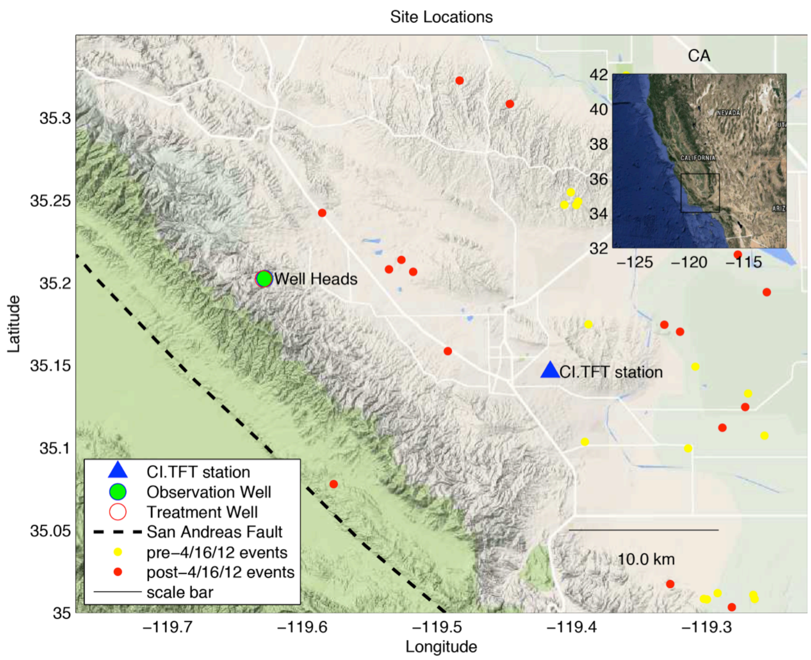

Figure 1. Location of treatment well (red circle) and observational well. The long dashed line denotes San Andreas fault. The red and yellow dots denote the earthquakes in SCEC data center during a 4 yr period since 2010. The minimum magnitude is 0.8. As it can be seen before and after this experiment, no M >0.8 earthquake happened in the vicinity of treatment well.

With an agreement between UCSB and oil company Venoco incorporation, my group is allowed to access its micro-seismic observations of one seven-stage hydraulic experiment for education and research purpose. As shown in Figure 1, the experiment was conducted on a Central California site only 8 km away from the San Andreas fault. The closest CISN station of is TFT, which is 17 km away. The hydraulic fracturing was stimulated in seven stages between approximate elevations of -6250 ft and -7540 ft. The induced micro-seismicity was observed using a sub-vertical borehole array composed of 16 3-component geophones. The treatment and observation wells are separated by about 1200 ft horizontally. The spatial interval between two nearby geophones is 50.8 ft. The total length of this array is then about 760 ft. The entire dataset includes 57 hours of records, which is sampled 4000 times per second. We received $30k from SCEC to study this project, which is used to support Master student Trevor Smith. Our research last year focused on improving the available catalog using waveform cross-correlation technique.

B. Method: Match filter and match & locate algorithms

Following the work of Shelly et al. [2007], recently match filter technique is widely used to improve the detection of microseismic events in noisy environment. It employs some events with known locations as templates and detects small events through stacking cross-correlograms between waveforms of the template events and potential small event signals in the continuous waveforms over multiple stations and components. As we show next, when correlation between template records and target records are high, it can be detected using the observations with signal-noise-ratio (SNR) much less than 1.

Let's use $T(t)$ to denote the template event and $O(t)$ for the continuous waveform. The signal we intend to detect is $O_s(t)$, and the background noise is $O_n(t)$. The SNR of data can be defined as $[\frac{\int O_s{}^2(t)dt}{\int O_n{}^2(t)dt}]^{1/2}$. The correlation between target signal and template (CCS) as well as that between the observed noise and template (CCN) can be represented as

$$CCS(\tau) = \frac{\int O_s(t-\tau)T(t)dt}{[\int O_s{}^2(t-\tau)dt \int T^2(t)dt]^{1/2}} \,\&\, CCN(\tau) = \frac{\int O_n(t-\tau)T(t)dt}{[\int O_n{}^2(t-\tau)dt \int T^2(t)dt]^{1/2}} \quad (1)$$

In the case the signal is very noisy or target event is small, SNR<<1 and $O(t)\sim O_n(t)$, the correlation between original continuous signal $O(t)$ and template records (hereafter referred as CC) can be approximated as

$$CC(\tau) = \frac{\int O(t-\tau)T(t)dt}{[\int O^2(t-\tau)dt \int T^2(t)dt]^{1/2}} \sim SNR \cdot CCS(\tau) + CCN(\tau) \quad (2)$$

Note that the contribution of cross-correlation between target signal and template ($CCS(\tau)$) to the overall correlation $CC(\tau)$ is weighted by SNR. If $CCN(\tau)$ is uncorrelated within the network, we can reduce it by stacking.

$$StackCC \sim \frac{1}{N}\sum SNR^i \cdot CCS^i + \frac{1}{N}\sum CCN^i \quad (3)$$

In our practice, we notice that the standard deviation of $\frac{1}{N}\sum CCN^i$ decreases with the number of station N as $1/\sqrt{N}$. It is about 0.02 using our 16-station array. Following Shelly et al., [2007], an event is detected when $StackCC$ is larger than 8 times of median absolute mean (MAD). This is equivalent to 5.33 times of standard deviation of $\frac{1}{N}\sum CCN^i$. Or on average $SNR \cdot CCS$ need be larger than 0.11. If CCS is 0.5, the corresponding event can be detected from the observations even with an average SNR of 0.22! It also can be seen that further improving the correlation between template signals and target signals, smaller event (with lower SNR) can be picked up from noise data.

The conventional match filter algorithm [e.g., *Huang and Beroza*, 2015; *Meng et al.*, 2013; *Shelly et al.*, 2007] ignores the location difference between template event and target event. Match&Locate algorithm [M&L, *Zhang and Wen*, 2015] is similar to match filter method. But rather than simply stacking the crosscorrelograms at multiple stations and components, in M&L algorithm the stacking is performed after making relative travel-time corrections based on the relative locations of the template event and potential small event scanning through a 3-D region around the template. Zhang and Wen [2015] showed that using this approach, $StackCC$ could be improved by nearly a factor of two if the relative location between template and target events is not negligible. M&L algorithm then can detect more events than match filter method. However, its computational cost is proportional to the number of spatial grid points and therefore is much slower than the match filter method. The memory requirement of the original version of M&L algorithm is proportional to the product of number of spatial grid points and length of time window [*Zhang and Wen*, 2015]. A large computer cluster is then required to conduct this study.

We have made two modifications to the M&L algorithm, aiming to improve its detection capability and reduce its computational requirement. First, in our approach the crosscorrelograms are interpolated before stacking to avoid the error caused by coarse sampling interval. Previously, this is achieved by resampling the template and continuous waveforms before the relocation analysis. The computational cost increases a factor of M^2 and the memory required increases by a factor of M, M is the ratio of new and old sampling rates. The new approach does not require more memory and the computational cost just increases slightly. Second, in the original version of the M&L algorithm, a temporal search is first conducted at each spatial grid point, following by a spatial search over all grid points. We found that if switching the sequence of temporal search and spatial search, the memory requirement can be independent of the number of spatial grid points. It reduces the memory requirement by two or three orders of magnitude. We then can use a workstation rather than a large cluster to conduct M&L analysis. Our numerical test also suggests that this modification slightly improves the search speed (~20%).

C. Analysis

Table 1. Comparison of the numbers of events registered in three catalogs

Catalogs for stages 5-7	Industry	Match Filter (e.g., Shelly et al., 2007)			Match&Locate (Zhang and Wen, 2015)		
		MAD>16	MAD>12	MAD>8	MAD>16	MAD>12	MAD>8
# of Events	1089	1474	1943	2529	1800	3071	12943

*Select 220 events in the industry catalog for this study.

Because of the efforts in method developments and the availability of computational resource, the systematic analysis to the observations is still undergoing. Here we only report our preliminary analysis to the 7.5-hour continuous waveforms recorded during the injection periods of stages 5-7, which occurred in less than 400 m away from the center of the geophone array. Both match filter and M&L algorithms have been used. The data has been bandpass filtered from 30 Hz to 1000 Hz to improve the SNR and we assume that there is only one event within any 1-s window as *a priori* condition [*Zhang and Wen*, 2015]. In the catalog provided by Venoco, 1089 events were registered during this period. We selected 220 of them as templates based on SNR. The M&L analysis took 15 days using an ERI cluster.

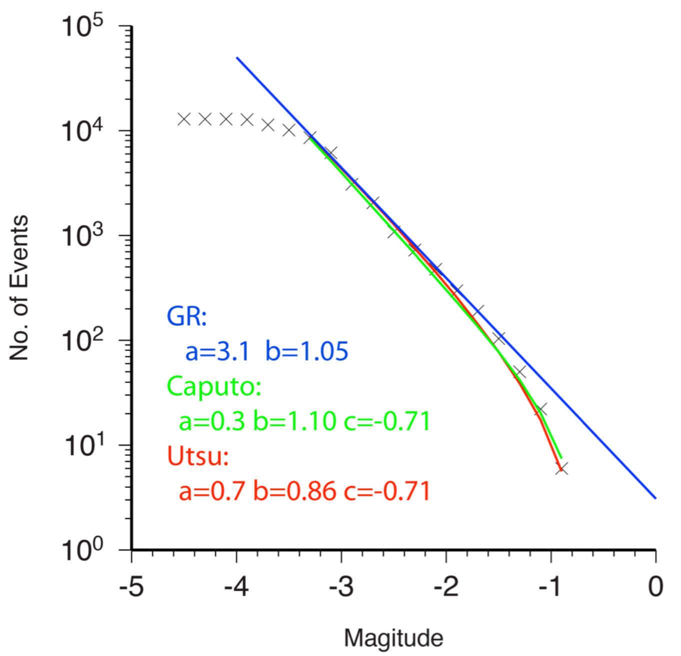

Figure 2. Magnitude-frequency distribution of earthquakes occurring during the period of hydraulic fracture injection. GR denotes Gutenberg and Richter relation. Caputo and Utsu denote two non-GR models [*Utsu*, 1999] that match this distribution better.

We summarize the result in Table 1. As it can be seen that the catalog based on match filter method has 2529 detected events (~2.5 times of these in industry catalog) and that based on M&L method has 12943 events (>12 times) within this 7.5-hour period. Increasing detection threshold shall reduce the false detections but will also significantly reduce the number of detected events. However, even using a very conservative threshold (stacking crosscorrelogram is larger than 16 times of MAD), M&L method still detects 70% more events (Table 1). Once an event is detected, its magnitude is computed based on the median value of the peak amplitude ratio between the detected event and the template event among all channels [*Meng et al.*, 2013; *Zhang and Wen*, 2015].

To explore the quality of match filter and M&L results, we check the 869 events that were included in the industry catalog but were not selected as templates. We find that 91% of them are detected by both match filter and M&L analyses. The 9% events that M&L failed to detect all occurred within 1 sec of detected events. As we only allow one event to be detected in any 1 s time window, they were ruled out. Hence, the catalog might still miss ~9% of events.

D. Magnitude-frequency relationship during injection period

We divide the M&L catalog into many 0.2 magnitude unit bins. The numbers of events within these bins are shown in Figure 2. Using the maximum curvature method [*Wiemer and Wyss*, 2000], we find the magnitude of completeness (M_c) of this catalog is -3.0. The same analysis to the industry catalog yields an estimate of -2.2. The largest earthquake detected has a magnitude of -0.71, suggesting a range of 2.3 magnitude units for the subsequent analysis of magnitude-frequency relationship. The number of events within this magnitude range is 6175. Matching the data within the magnitude range from -3.0 to -1.6 with the GR relation yields a b-value estimate of 1.05 (blue line in Figure 2), close to the b-value estimate for relatively large natural earthquakes in this region [*Mori and Abercrombie*, 1997]. However, it can be seen that the number of events with magnitudes larger than -1.6 is much less than what is predicted using the GR relation. As shown in Figure 2, the distribution can be better modeled with non-GR relations such as Caputo and Utsu equations [*Utsu*, 1999], though the b value estimated using Utsu relation is significantly smaller (0.86, Figure 2). Hence, our result reveals that the earthquake magnitude distribution during the injection period does not follow GR relation.

While it is not clear yet whether this is held for all hydraulic fracture experiments, it is consistent with what Huang and Beroza [2015] found during the study of Guy-Greenbrier earthquake sequence. Currently, we are analyzing the 9-hour post-injection records and intend to explore whether its magnitude-frequency distribution return to GR relation. The result is planned to report during SSA.

E. References

Huang, Y. H., and G. C. Beroza (2015), Temporal variation in the magnitude-frequency distribution during the Guy-Greenbrier earthquake sequence, *Geophys Res Lett*, *42*(16), 6639-6646.

Meng, X. F., Z. G. Peng, and J. L. Hardebeck (2013), Seismicity around Parkfield correlates with static shear stress changes following the 2003 M(w)6.5 San Simeon earthquake, *J Geophys Res-Sol Ea*, *118*(7), 3576-3591.

Mori, J., and R. E. Abercrombie (1997), Depth dependence of earthquake frequency-magnitude distributions in California: Implications for rupture initiation, *J Geophys Res-Sol Ea*, *102*(B7), 15081-15090.

Shelly, D. R., G. C. Beroza, and S. Ide (2007), Non-volcanic tremor and low-frequency earthquake swarms, *Nature*, *446*(7133), 305-307.

Utsu, T. (1999), Representation and analysis of the earthquake size distribution: A historical review and some new approaches, *Pure Appl Geophys*, *155*(2-4), 509-535.
Wiemer, S., and M. Wyss (2000), Minimum magnitude of completeness in earthquake catalogs: Examples from Alaska, the western United States, and Japan, *B Seismol Soc Am*, *90*(4), 859-869.
Zhang, M., and L. X. Wen (2015), An effective method for small event detection: match and locate (M&L), *Geophys J Int*, *200*(3), 1523-1537.