NOTICE CONCERNING COPYRIGHT RESTRICTIONS

Discovering New Events Beyond the Catalog— Application of Matched Field Processing to Salton Sea Geothermal Field Seismicity

Jingbo Wang[1], Dennise C. Templeton[1], David B. Harris[2]

[1]Lawrence Livermore National Laboratory, Livermore CA
[2]Deschutes Signal Processing, LLC, Maupin OR
wang64@llnl.gov

Keywords

Geothermal, Salton Sea, matched field processing, event detection, event location, seismic array, microseismicity

ABSTRACT

We aim to detect and locate more microearthquakes using the empirical matched field processing (MFP) method than can be detected using only conventional earthquake detection techniques. We propose that empirical MFP can complement existing catalogs and techniques. In the Southern California Earthquake Data Center (SCEDC) earthquake catalog, 2972 events were identified in our study area during January 2008 and December 2010. We use this earthquake catalog to identify the best potential empirical master templates. We create 242 master templates with at least four stations with good quality. We test our method on continuous seismic data collected at the Salton Sea Geothermal Field during January 2010. The MFP method successfully identified 1115 events. Therefore, we believe that the empirical MFP method combined with conventional methods significantly improves network detection capabilities.

Introduction

Accurate identification and mapping of large numbers of microearthquakes is one technique that provides diagnostic information when determining the location, orientation and length of underground crack systems for use in reservoir development and management applications. Conventional earthquake location techniques are often employed to locate microearthquakes. These techniques require picking individual seismic phase onsets across a network of sensors and work best on seismic records containing a single well-recorded event with low signal-to-noise ratio. Additionally, fluid injection frequently induces a large number of events with overlapping waveforms, which can complicate the picking of phases or completely obscure the onset of smaller signals.

To aid in the seismic characterization of reservoir fracture networks, we propose to complement traditional earthquake detection and location techniques with the empirical matched field processing (MFP) method. MFP, as applied in seismology, matches the spatial structure of incoming seismicity observed by a network of sensors to master templates keyed to potential event locations.

Empirical MFP develops a catalog of matching templates from a collection of representative microearthquakes that uniformly samples the study volume. The earthquakes for the empirical master templates initially will have to be located using conventional earthquake location techniques and subsequently relocated using advanced processing techniques, however all future seismicity can be mapped using the computationally efficient MFP algorithm. In this paper, we apply this technique to recent seismic swarms that occurred in the Salton Sea Geothermal Field during January 2010.

Geological Background

The Salton Sea Geothermal Field lies on the southeastern shore of the Salton Sea. The Salton Sea is the lowest part of the
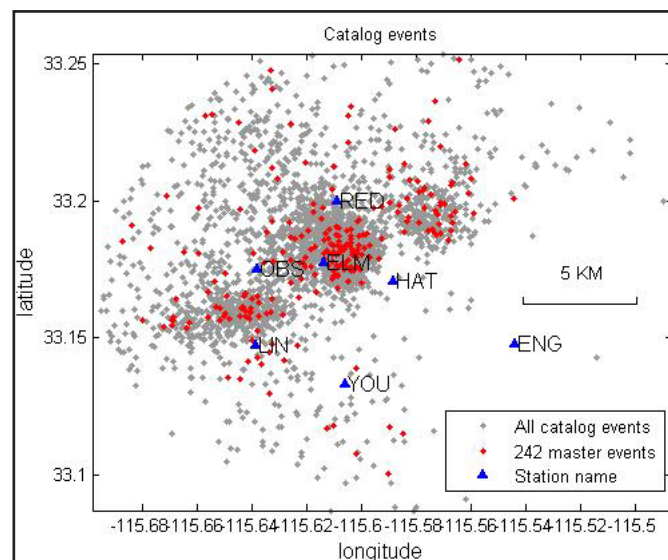
Figure 1. Seven stations (blue triangle with station names) map view in EN network. The grey dots are catalog events from SCEDC during January 2008 and December 2010. The master templates are marked in red.

Salton Trough, a tectonic depression. Near the southern end of the Salton Sea, the San Andreas Fault appears to terminate at a spreading center called the Brawley seismic zone. This zone is the most northerly in a series of spreading centers distributed along the length of the Gulf of California that forms part of the East Pacific Rise. Rifting and intrusions produce high heat flow that metamorphoses the sedimentary rocks to shallow depths (Fuis *et al.*, 1984).

Data

Seven three-component seismic stations are located within the geothermal production field (Figure 1). This array is maintained by Caltech/USGS and continuous data has been archived at the Southern California Earthquake Data Center (SCEDC) since January 2008. Earthquake catalog locations and phase data is also available at the SCEDC.

Methodology

Our MFP technique is an adaptation of a signal-processing technique originally developed to locate continuous underwater acoustic sources. MFP can steer the array explicitly in the frequency domain using the complex phase and amplitude factors obtained by solving the wave equation through a propagation model. However, it is difficult to develop realistic Earth

models to predict the structure of seismic wavefields at frequencies much above a tenth of a Hertz (Harris and Kvaerna, 2010). An alternative to calculating the wavefield structure across an array is to estimate the structure directly from field calibration data, i.e., previous seismic events. We refer to this strategy as empirical MFP. In empirical MPF, the master templates that are created from the seismograms of previously detected micro-earthquakes contain contributions from direct and scattered seismic energy.

An example work flow is described in Figure 2. To determine which events we should choose for our field calibration events, we first obtained the SCEDC catalog. The objective was to choose master template events as the input of the empirical MFP method. We visually inspected each potential master template to make sure
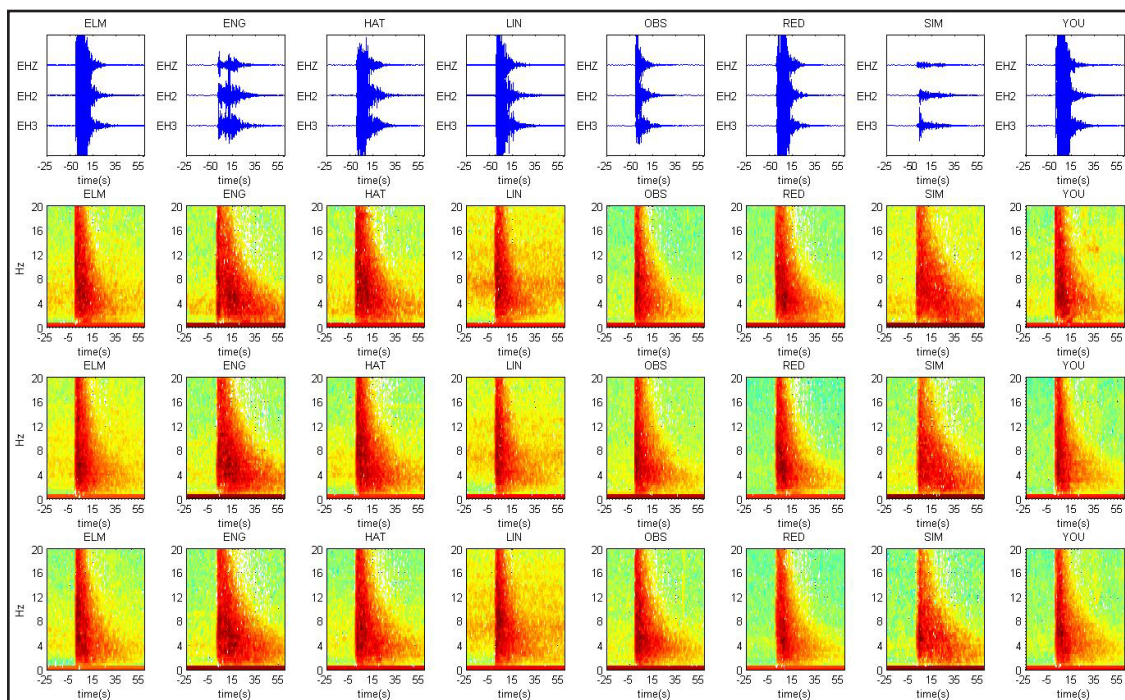
Figure 3. Example of one master event with origin time at January 23, 2008 18:15:57. The first row contains the waveform with the station name on top. The second, third and fourth rows show spectrograms of the EHZ, EH2 and EH3 components between 0 – 20 Hz, respectively.
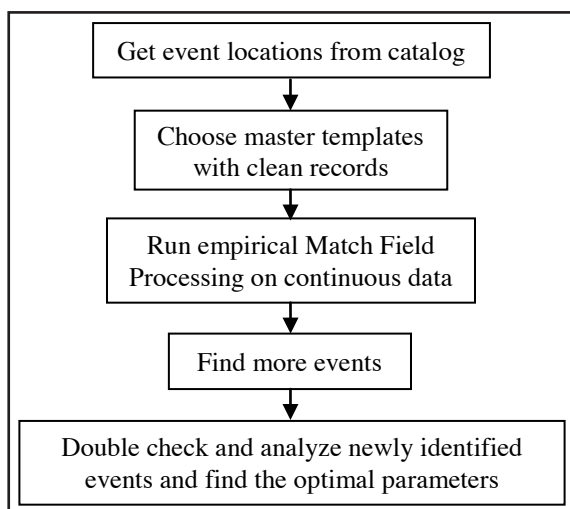
Figure 2. Empirical MFP work flow.

that there were no overlapping events or noise spikes. Then, we ran the empirical MFP code on the continuous seismic data from January 2010 and identified events in the data stream that match the master templates. We investigate different frequency bands: 2-8, 4-10, and 6-12 Hz and different thresholds from 3 to 25. Finally, we double check the new detections by eye and analyze the relationship between newly identified events and master templates.

Master Events Selection

The number of events that have occurred within 10 km of station HAT in three years between January 2008 and December 2010 is 2972. We have looked at the waveforms of all catalog events and chose the best quality data as the master templates. These events have little noise in the frequency range we are investigating and do not have overlapping events within the 70 sec time window that is needed to create the master templates. We choose 242 master templates out of 2972 catalog events. The master templates include at least four stations with good quality recordings.

Figure 3 shows an example of a master event with origin time at January 23, 2008 18:15:57. In general, the background noise is incoherent across the 70 sec time window at stations ENG, HAT, OBS, RED, SIM and YOU. In this particular example station LIN displays higher background noise than other stations over the frequency band in which we are searching for events. In general, LIN is consistently noisy especially when events are weak. Therefore, we decided to remove station LIN due to its

Figure 4. Empirical MFP detection results from four master templates during January 10th – 20th, 2010. The threshold is indicated by the horizontal line in each plot.

Figure 5. HypoDD relocation of the January 2010 seismicity and master templates. The four detectors shown in Figure 4 are marked by red symbols. The background seismicity and other master templates are represented by the black dots.

poor quality of recording. The other stations do not consistently display background noise energy.
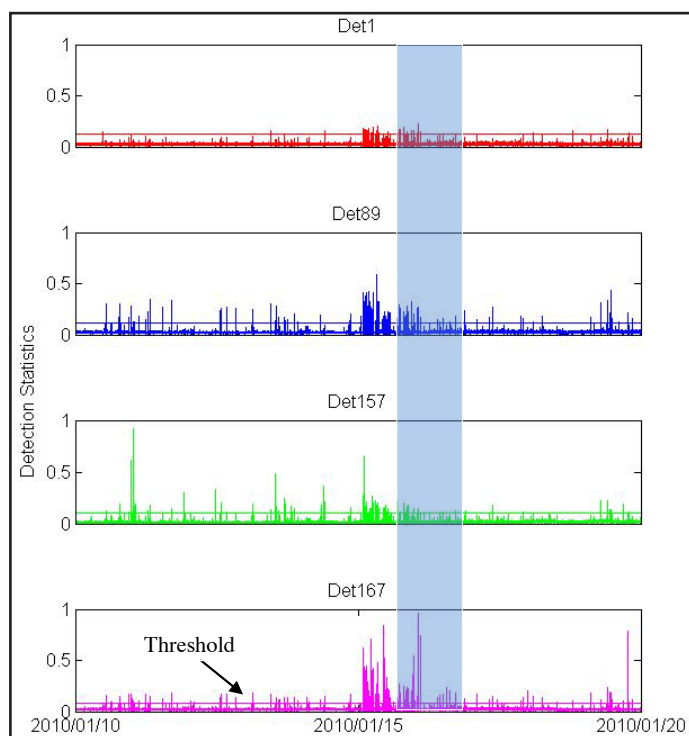
MFP Results

For this study, we focus on continuous data collected in January 2010. The data includes the largest seismic swarm that occurred in the Salton Sea Geothermal Field since continuous seismic data has been archived at the SCEDC. Unfortunately, seismic station SIM was not operational at this time.

The empirical MFP code performs its calculation on the continuous data using a 70-sec window which steps forward 1 second at a time. Figure 4 shows a 10-day example of results for the time period January 10th – 20th, 2010. This segment of data is band-pass filtered between 4 - 10 Hz. The y-value at each time point indicates the normalized detection statistic. A value of 1 would indicate an exact match between the template and the incoming seismicity at that particular time. Threshold levels for each detector are calculated over each 10-day period and are a function of the average detection statistic value (Figure 4). Detection statistics above the threshold are compared to detections at other detectors. If two detectors identify the same event, the detector with the largest detection statistic is then determined to have detected the event. Figure 5 gives the location of the detectors in 3D view.

As illustrated in Figure 4, most detectors show an elevated detection statistic during the January 15 swarm (see the blue shadow zone). These four detectors show representative behavior. Compared to other events, Detector 89 (origin time: April 17, 2009 04:35:34) and 167 (origin time: January 16, 2010 01:26:20) detected events in a relatively similar pattern. This is reasonable because these two events are closely located and so may have similar wavefields along the path (Figure 5). Detector 167 matches the incoming seismicity better than Detector 89 as indicated by the higher detection statistics. Detector 157 (origin time: January 11, 2010 00:25:36) is somewhat more distant from Detectors 89 and 167. This detector is able to detect some other events that are not detected by Detectors 89 and 167. This illustrates the fact that the more spatially evenly sampled the master events are, the higher chance we will have to be able to detect more events in that area. The detection statistics from Detector 1 (origin time: 2008/01/21, 03:29:28) mostly fall below the threshold. This detector is far away from the swarm and not many events occurred close to it during this particular time period (Figure 5).

Threshold and Frequency Band Study

We identified 1115 new events by visually verifying the MFP detections in the 2 – 8 Hz frequency band and by setting the threshold to 3 times the average detection statistic value within the 10-day time frame. We use this set of events as the reference set of real detections. We compare the reference set with the detections identified from the three frequency bands 2 – 8 Hz, 4 – 10 Hz and 6 – 12 Hz, as well as thresholds ranging from 3 to 25. Lowering the threshold value increases the number of real events identified, however the number of false positive detections increases exponentially possibly due to matching of the background noise instead of a real signal. Figure 6 shows the number of real events identified is a function of the threshold level. As we can see, the false alarm rate quickly decreases when the threshold is set higher.
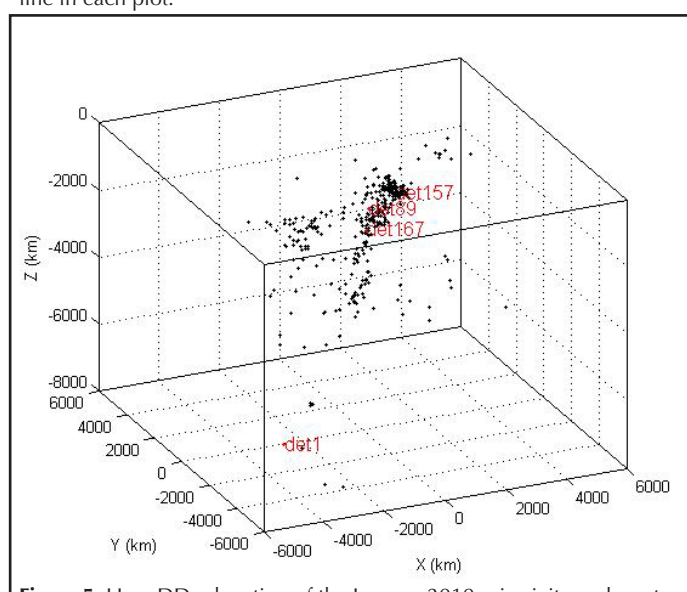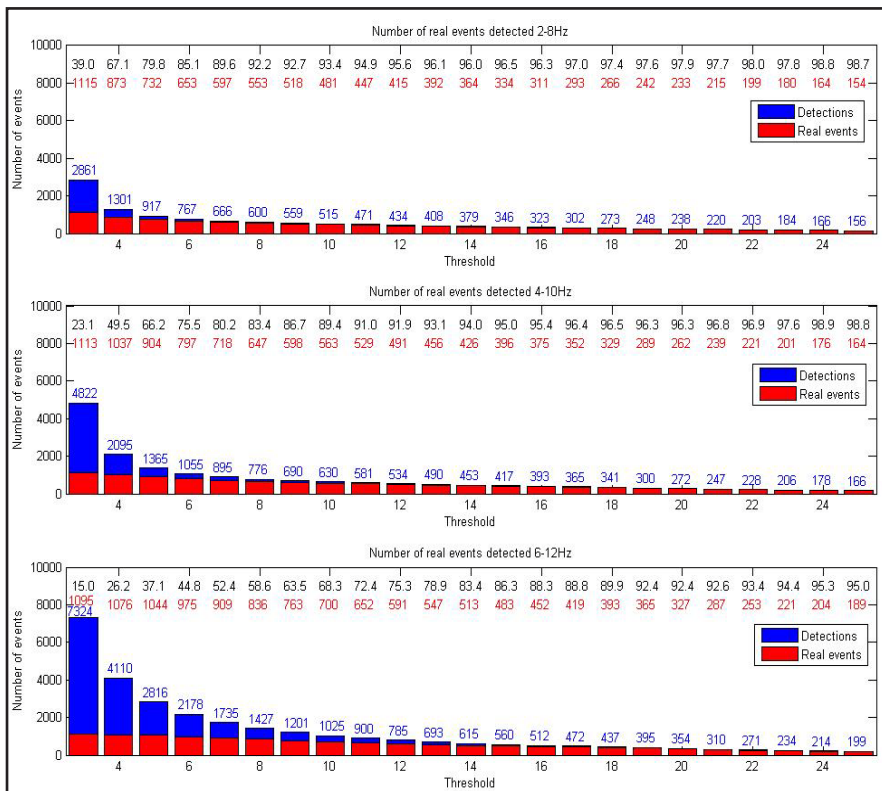
Figure 6. Number of real events identified by the MFP algorithm compared with the total number of MFP detections within different frequency bands and above different thresholds. The number of detections is marked on top of each bar in blue. The number of real events identified is indicated in red. The black number on top is the percentage of real events within the detections. The same scale for the y axis is used for the three frequency bands: 2-8 Hz, 4-10 Hz, and 6-12 Hz, from top to bottom.

Figure 7 shows the reconciliation between the real events and the catalog events. During January 2010, 333 events were identified in the SCEDC catalog. 318 catalog events were detected by the empirical MFP method and 15 catalog events were not detected by the MFP.

Figure 6 and 7 demonstrate that more real events are identified with a higher frequency range at lower thresholds, but the price is the introduction of more false detections. When the threshold is increased, the alarm rate increases correspondingly. For frequency band 2 – 8 Hz, the true alarm rate is approximately 95% when the threshold is 12. For 4 – 10 Hz, the threshold has to be 15 in order to reach 95%, but the number of real events is smaller than that in the 2 – 8 Hz frequency band. For 6 – 12 Hz, the number of real events identified decreases even more quickly as the threshold increases. This behavior could potentially be improved by further removal of noisy stations.

Discussion

Among 1115 real events identified by the MFP with threshold set to 3 using 242 master templates, we plot the number of real events identified by each master template in Figure 8. We divide the master templates into three groups (see Table 1).

Figure 9 shows the magnitude and location of the master templates marked in each group by color. The grey dots are background seismicity in January 2010. We have found that the master templates with low identification ability are mainly in the western half of the study area. The master templates with high identification ability are scattered throughout the study area.
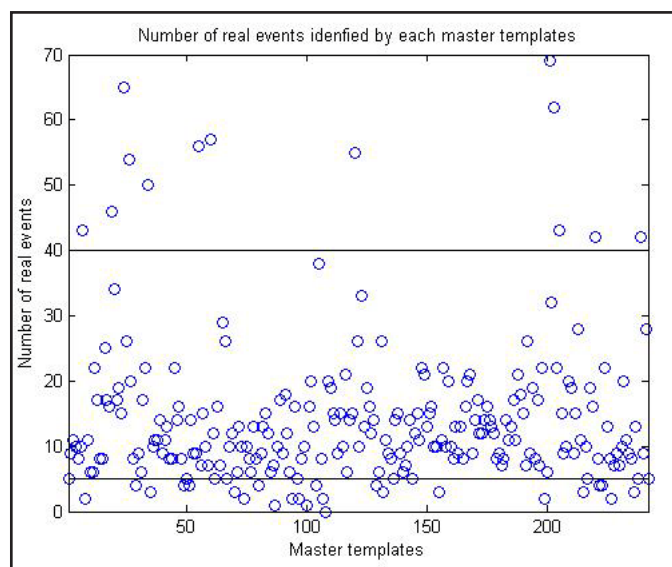
The performance of the master templates is heavily affected by whether the templates have a
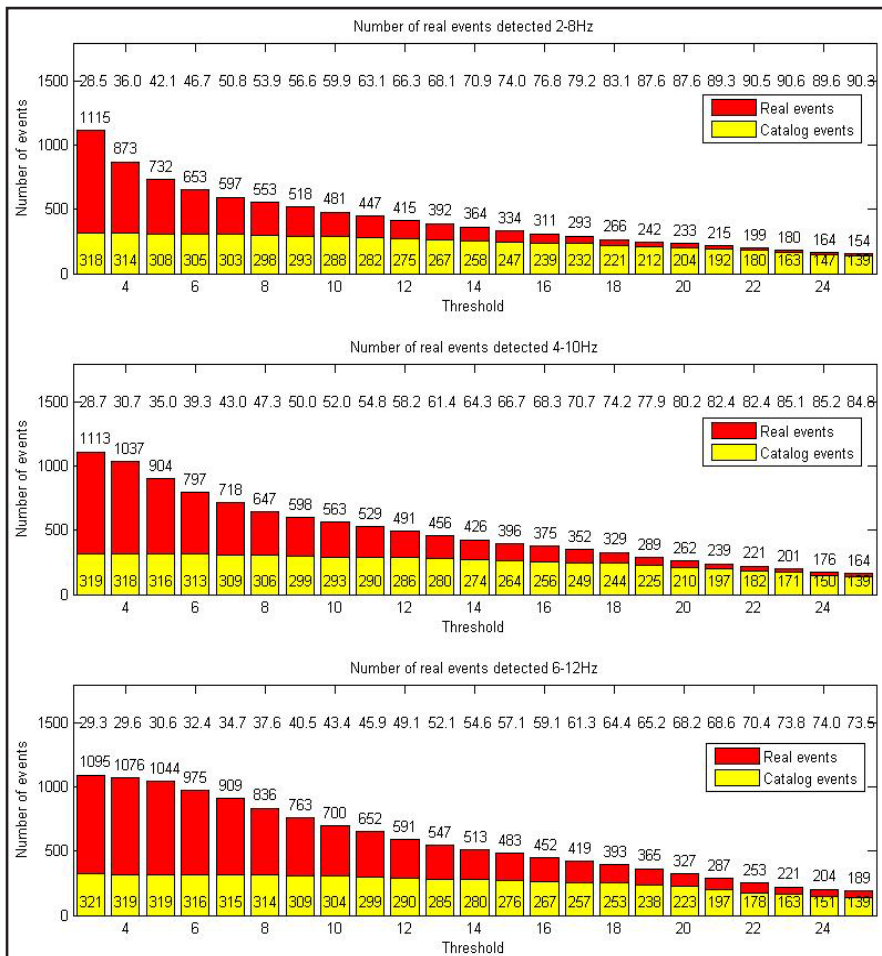
Figure 7. Comparison of the number of real events identified by the MFP and the events included in the official SCEDC catalog for different frequency bands and thresholds. The number of real events is marked on top of each bar. The number of catalog events reconciliated is inside the yellow bars. The number on top is the percentage of the catalog events that were identified in each threshold level. The same scale for y axis is used for the three frequency bands: 2-8 Hz, 4-10 Hz, and 6-12 Hz, from top to bottom

Table 1. Three groups of master templates.

Color	Number of real events identified	Comments
Red group A	Num >=40	High
Blue group B	40>Num>=5	Medium
Green group C	Num<5	Low

Figure 8. Number of real events identified by each master template. Two lines: num=5 and num=40 are the criteria to divide them into three groups by the performance of the master templates.
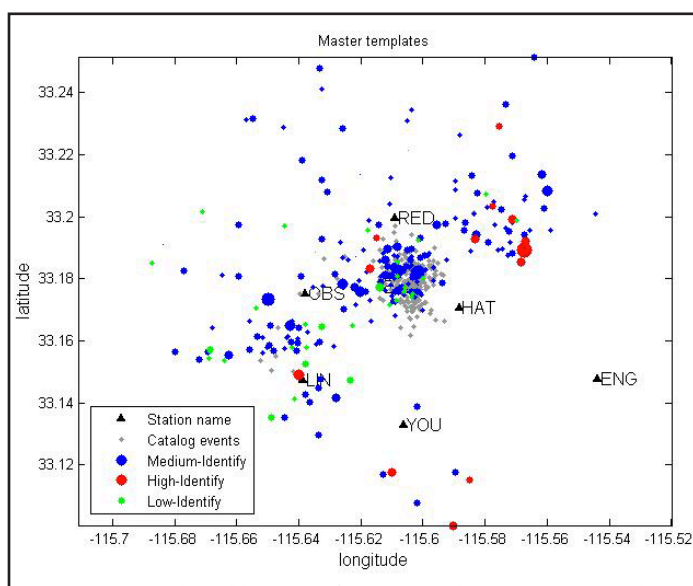
Figure 9. Magnitude and location of master templates in map veiw. The color indicates the group mark. The size of the dot indicates the magnitude of the master templates.

clean signal or are compromised by high background noise. Other factors that affect the detection ability of the master templates will be investigated in future research. For example, the magnitude of the master templates is another potential factor to affect the detection capability. Whether larger master template identify more

events, is still an open question. However, Figure 9 shows that the master templates with various magnitudes are randomly allocated over the whole region. Another factor is the location of the master templates. As we know that most of the events are located around the major swarm, a master template located nearby the swarm is expected to identify more events than the one located farther away from the swarm. But in Figure 9, the master templates which identify the most real events are not located at the center of the main swarm. Performance of the master templates may also be influenced by the source mechanism of the master event. These issues will be investigated in a future study.

Conclusions

MFP with empirically calibrated master templates is able to detect more events than can be detected using conventional techniques. Unlike most array processing methods, empirical MFP does not require a plane wave assumption. Therefore, empirical MFP has more adaptability to varying noisy environments, as long as the master templates adequately cover the area where future events will possibly occur. Our test on the continuous data during January 2010 demonstrates the detection capability using empirical MFP. There are 1115 events detected in total by MFP, while the catalog reports only 333 events. Thus the empirical MFP algorithm significantly improves seismic array detection capability.

Acknowledgements

References

Waldhauser, F. (2001), "hypoDD -- A program to compute double-difference hypocenter locations," *U.S. Geological Survey Open-File Report*, 01-113.

Waldhauser, F., and William L. E. (2000), "A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, California," *Bulletin Seismological Society of America*, 90, 1353-1368.

Hooland, A. A. (2002), Microearthquake Study of the Salton Sea Geothermal Field, California: Evidence of stress triggering. Idaho National Engineering and Environmental Laboratory, Geosciences Department. *Thesis*, The University of Texas at El Paso.

Magistrale, H., S. Day, R. W. Clayton, and R. Graves, (2000), The SCEC southern California reference 3D seismic velocity model Version 2, *Bulletin Seismological Society of America*, v. 90, no. 6B, S65-S76

Kohler, M., H. Magistrale, and R. Clayton, (2003), Mantle heterogeneities and the SCEC three-dimensional seismic velocity model version 3, *Bulletin Seismological Society of America,* 93, 757-774.

Harries, B. D., and T. Kvaerna, (2010), Superresolution with seismic arrays using empirical matched field processing, *Geophysical Journal International*, 182, 1455-1477.

Fuis, G. S., W. D. Mooney, J. H. Healey, G. A. McMechan, and W. J. Lutter, (1984), A seismic reflection survey of the Imperial valley, California. *Journal of Geophysical Research*, 89, 1165-1189.

Allis R., J.N. Moore, J. McCulloch, S. Petty, and T. DeRocher, 2000. "Karaha-Telaga Bodas, Indonesia: A Partially Vapor Dominated Geothermal System." Geothermal Resources Council *Transactions*, v. 24, p. 217-222.