Characterizing Seismic Swarm Morphology

Sutton Chiorini

11/24/15

Advisor: Prof. Vedran Lekic

GEOL393

Abstract

Seismic swarms are characterized by an anomalously large number of earthquakes happening in a relatively small area, typically ranging from a few to several kilometers, over a short period of time, typically ranging from days to weeks. However, how and why swarms occur is poorly understood, which poses an interesting set of questions within the greater body of geologic research. In this study, I propose that previous methods of identifying seismic swarms from larger bodies of earthquake catalogs are not effective in characterizing the full range of possible swarm behaviors. Furthermore, I propose that if a full seismic swarm catalog were to be compiled and analyzed, the sequences will demonstrate a much more even distribution in space and that a larger fraction of swarms will be shown to migrate in both time and space. Finally, I propose that over time, a smaller fraction of swarms will exhibit an exponential decay in event occurrence over time.

I will use a set of analyses implemented in MATLAB to, first, analyze two different methods for isolating seismic swarms from larger earthquake catalogs, proposed by Vidale and Shearer (2006) and Zaliapin et al. (2008), respectively, before proceeding to demonstrate that while both methods do identify seismic swarms, the former method does not identify all the seismic swarms identified by the latter, and the seismic swarms it does identify contain less events than the seismic swarms identified by the former. I will then discuss why this might be the case by analyzing the spatial distribution of events within seismic swarms. I will finally conclude by analyzing the spatial distribution, migration characteristics, and decay behavior of all the seismic swarms that the Vidale and Shearer method missed, in order to more accurately characterize the full range of observed swarm behavior.

1) Introduction

Seismic swarms, sequences of earthquakes occurring in a relatively small area over a relatively short period of time, are defined in a variety of ways in the seismological literature. Corral (2003) defines swarms as seismic activity that is not associated with a main event, and therefore do not follow Omori's law, in which the frequency of aftershocks decreases approximately with the inverse of time after the main shock. Vidale and Shearer (2006; henceforth VS2006) define swarms to include many earthquakes striking in a limited space over a limited time window. Using the VS2006 approach to analyze swarms in Southern California, Chen and Shearer (2011) conclude that seismic swarms are thought to be mainly triggered by physical processes. Although Zaliapin et al. (2008; henceforth Z2008) does not propose a specific definition for seismic swarms, they do detail that clustered events they have identified are much closer to each other in time and space than would be expected for a uniform Poisson distribution.

Seismic swarms are generally composed of low-magnitude events, the largest of which generally does not exceed M_w 5.0, and on average events run between M_w 0.0 and 2.0. Because of their small size, swarm earthquakes do not pose an immediate threat to surrounding communities and infrastructure upon initiation of the sequence. Nevertheless, swarms have been shown to be associated with magma or fluid activity (Ruppert et al. 2011), as well as stress loading in fault areas and slow aseismic slip events (Vidale and Shearer, 2006). Therefore, swarms are related to the processes that result in volcanic eruptions and stress accumulation that can lead to large, destructive earthquakes. Mechanisms for where and why swarms occur are not well understood, and an improved understanding of seismic swarms can elucidate the relationship between swarms and volcanic eruptions and large earthquakes, or point to yet other triggering mechanisms for seismic swarms that have not been explored yet.

Seismic swarms have also been demonstrated to have a variety of applications to other geological questions. Savage et al. (2015) showed that seismic swarms can be used as a predictor of volcanic eruptions, providing an additional method for hazard mitigation in volcanic areas. Umeda et al. (2015) demonstrated that seismic swarms can precede a megathrust event, in their specific case the 2011 Tohoku earthquake. As this event was extremely damaging both in terms of lives and infrastructure, this could offer a new way to predict and avoid another such disaster. Finally, Shapiro and Dinske (2008) demonstrated seismic triggering due to fluid injection exhibits swarm-like behavior, which could help in investigations regarding seismic activity associated with hydraulic fracturing. All of these methods demonstrate that better understanding how and why seismic swarms could provide a fruitful source of information for other geological investigations.

Earthquake catalogs are large and contain hundreds of thousands of events for a seismically active region like Northern California over the course of three decades. Therefore, identifying what pattern is a swarm and what is simply background seismicity is a complicated task with no straightforward methodology. Geophysicists have identified a variety of methods by which to identify seismic swarms from large catalogs of earthquakes. Previous research has identified seismic swarms specific to volcanic regions (Fischer et al. 2003; Ruppert et al. 2011), and presented methods for identifying seismic swarms (Vidale and Shearer 2006; Zaliapin et al. 2008; Chen and Shearer 2011) in non-volcanic settings. I will base my research on the methods proposed by Vidale and Shearer (2006) and Zaliapin et al. (2008). The first research method, VS2006, uses arbitrary-chosen parameters defining spatial, temporal, and earthquake count thresholds to identify 71 seismic bursts in Southern California. The second method, Z2008, uses a statistical methodology to analyze spatio-temporal distributions of earthquakes and identify anomalously clustered sets within the population of all earthquakes.

Vidale and Shearer (2006)

Vidale and Shearer picked seismic bursts using 166,525 events from the waveform relocated catalog SHLK_1.01 in Southern California based on an arbitrary-chosen set of criteria, which consists of the following: 1) An initial event must be followed by at least 39 events within a radius of 2 km in 28 days; 2) There must be 3 or fewer events in the prior 28 days within the same 2 km radius; and, 3) No more than 20% more events can occur between 2 and 4 km from the initiating event in the same 28 days afterwards.

Based on these threshold parameters, Vidale and Shearer identified 71 seismic clusters. Fourteen of these bursts resembled main shock/aftershock sequences, eighteen exhibited “swarm-like” sequences or behaviors lacking a clear main shock, and 39 were considered to be “average” sequences, falling between the mainshock/aftershock and swarm sequence behavior. Vidale and Shearer then analyzed the eighteen swarms identified using their particular choice of thresholding parameters (Figure 1), and concluded that seismic swarms tended to exhibit an interval of steady seismicity rate, and that the largest event in the swarm tended to strike later in the sequence. They also found a weak correlation between the number of events in each burst and the magnitude of the largest event in each burst, and that shallow sequences were most likely to be swarm-like. Finally, they found that swarms were most likely to occur in extensional (normal faulting) settings. They concluded that seismic swarms were most likely driven by pore fluid pressure fluctuations and that they are most likely a general feature of tectonic faulting, rather than specific to volcanic or geothermal regions.

Despite its interesting findings and influence on the swarm community (the study has been cited 99 times according to Google Scholar), Vidale and Shearer's analysis of seismic swarms is not without limitations. Their use of an arbitrarily-chosen set of parameters to define what is and is not a swarm makes it difficult to quantify how many swarms might have been missed by their analysis. Therefore, whether the conclusions of their analysis can be generalized to all swarms remains an open question. In addition, their sample set of seismic swarms, limited to eighteen bursts, is a fairly small sample size, making it difficult to generalize the characteristics they identify to all swarms globally, or even throughout California, since the tectonics of Southern California differ substantially from those in Northern California. Further research on swarm sequences carried out using an objective set of parameters to define a swarm, and done outside the Southern California setting, would help answer these questions.

Zaliapin et al. (2008)

Zaliapin et al. proposed using a statistical methodology for analyzing the clustering of seismicity in the time-space-energy domain. They established the existence of two statistically distinct populations of earthquakes: clustered and non-clustered. Clustered earthquakes can be considered to belong to a swarm population, and the non-clustered events to a Poisson or non-swarm population. This method was developed based on the analysis of Baiesi and Paczuski (2004), and is built upon the parameter η_{ij}, or pairwise earthquake distance in space, time, and energy; the smallest η_{ij} across all *j* is associated with a particular earthquake *i*. The catalog used in their research was produced by the Advanced National Seismic System.

Zaliapin et al. first identified the location, time, and magnitude of each event in their catalog, and then used those parameters to calculate an intercurrence time (T) between pairs of events, denoted by subscript *i* and *j*, as well as a spatial distance (R) between the two events. They then normalized R_{ij} and T_{ij} by the magnitude of earthquake *i*. By multiplying these two parameters, they produced a nearest neighbor distanceη_{ij}. A scatter plot of values of R_{ij} and T_{ij} corresponding to the minimum $\eta_{ij}$for each earthquake can then be used to identify two statistically distinct earthquake populations (Figure 2a): A population of earthquakes that follow a Poisson distribution (background seismicity) found at larger distances from the origin and tracing out a field with linear and negatively sloped distribution in the R-T space; A population of clustered seismicity (including swarms) at smaller distances from the origin. Histograms of the values of η (Figure 2b) follow a bimodal distribution between the clustered (smaller average η) and non-clustered (larger average η) populations of earthquakes (Figure 2). Based on this analysis, Zaliapin et al. concluded that seismic swarms could be identified using their clustering in the spatial-temporal-energy domain.

As it is based on statistical analysis of the clustering of seismicity in the energy-normalized spatio-temporal domain, the Zaliapin et al. method dispenses with the need for using arbitrary threshold parameters to identify swarms. Nevertheless, the method does have some limitations. Rather than identifying specific bursts of seismicity as Vidale and Shearer did, Zaliapin et al. identifies all swarm events without identifying the swarm they belong to. In other words, while all the swarm events in a catalog can be identified using this approach, there is no way to tell which event belongs to which swarm, or if events belong to multiple swarms. Therefore, events could be misidentified with an unrelated earthquake swarm if they happen to coincide in time and space with that swarm.

2) Hypothesis and Proposed Work

Despite the fact that a number of methods have been established to identify seismic swarms from earthquake catalogs, there is no clear procedure for determining whether all the seismic swarms in an earthquake catalog have been identified, or if all events being identified as swarm events are unique to a swarm. As a result of these difficulties, to date, there are no comprehensive compilations of swarm earthquakes, even though complete catalogs of earthquakes have existed for decades.

I propose to compile a comprehensive catalog of swarm earthquakes – as opposed to the incomplete swarm compilations that are typically analyzed – and analyze the spatial and temporal distributions of the swarm events. I hypothesize that these distributions will not fully match the conclusions advanced by previous investigators. More specifically, I propose a two-part hypothesis: 1) that the swarms I identify will not exhibit an exponential decay curve in terms of a normalized distribution over time, and that they will exhibit a much more even distribution in space than was found by Vidale and Shearer (2006); and, 2) that a larger fraction of swarms will be shown to migrate than found by Chen and Shearer (2011). I will use the hypocenter double difference catalog compiled by Waldhauser (2013) as my full earthquake catalog in Northern California, and apply methods proposed by Vidale and Shearer (2006) and Zaliapin et al. (2008) in order to compile and analyze the results, morphology, number of events and locations of my seismic swarm catalog.

I expect that the significance of my work will add to the understanding of seismic swarm mechanisms, regarding both where and why they occur. In addition, my work will build upon previous research done by Vidale and Shearer (2006) and Zaliapin et al. (2008) by expanding their categorization of seismic bursts and swarm events in the form of a complete swarm catalog. Finally, my research will add to the existing body of information regarding the temporal and spatial distribution of swarms as they migrate in space and time, which in turn will help categorize their structure and possibly elucidate their relationship to other geologic processes.

3) Methods

HypoDD Algorithm and Resulting Catalog

Waldhauser and Schaff (2008) developed a catalog of 513,474 events in Northern California spanning 27 years, from the digital seismic waveform archives of Northern California. The catalog is produced using both waveform cross correlation and double difference methods, in which pairs of events with correlated waveforms are then inverted for the precise relative locations of events using a hypocenter double difference *(hypoDD)* algorithm. This method relates the observed and the predicted travel-time differences for pairs of earthquakes observed at common stations to their hypocenters in order to link them through a chain of nearest neighbors, resulting in a high-resolution relative hypocenter locations over a large area (Figure 3).

For identification and analysis of seismic swarms, I modified and wrote analysis codes in MATLAB based on the work of Vidale and Shearer, and Zaliapin et al. For the former, I used a program written by Jeff Gay that applies the five parameters identified by V2006 in order to isolate all events that fit within those specified values. For the latter, I wrote a program based on Z2008 that calculates an intercurrence time, distance between all events, normalizes both parameters based on the magnitude distribution of the events, and then multiplies them together to yield a nearest neighbor distance.

Schuster Test

In order to confirm whether all swarms were being identified in the catalog by both methods of analysis, I applied a Schuster test to the catalog after removing swarms isolated by each method. A Schuster test, first created by Arthur Schuster in 1897, computes the probability that the timing of events in a catalog varies according to a sine-wave function of period T. The probability that the distribution of event times arises from a uniform seismicity rate is referred to as the Schuster p-value. The lower this p-value, the higher the probability that the distribution of the timings of events stacked over the period T is non-uniform, which is usually interpreted as the probability of a periodicity at period T (Ader and Avouac 2013). Therefore, low Schuster p-values indicate non-uniformities in the catalog, which, in our case, point to swarm events still remaining that have not been identified by the method involved. Once the Schuster p-values have been generated, they can be plotted (Figure 4) to demonstrate whether the observed periodicities at periods T exceed expected values at 99% confidence. Larger numbers of significant periodicities can be interpreted to indicate a greater number of swarm events still remaining in the catalog.

Evaluating Uncertainties in Earthquake Location

The earthquake events in the catalog I will be working with have been relocated, as previously stated, using the hypoDD, or hypocenter double difference algorithm. This highly precise algorithm results in high-resolution relative hypocenter locations over a large area for each earthquake event, reducing uncertainty by over an order of magnitude compared to catalog locations (Walhauser and Ellsworth, 2000). Although error estimated assigned to relocated hypocenters need to be reviewed, especially when station distribution is sparse or if azimuthal coverage of available phases is not optimal (Waldhauser 2001), Waldhauser and Ellsworth (2000) reviewed a number of error estimates using a battery of tests with the hypoDD algorithm. They were able to conclude that the relocation method is able to image very fine-scale structure of seismicity along fault zones. In addition, they were able to safely conclude that the algorithm allowed for the consistent relocation of seismicity with high resolution along entire fault systems, therefore corroborating its efficiency and efficacy. Using these conclusions, the vertical relative location error in kilometers at the 95% confidence level is reported as part of Waldhauser’s earthquake catalog of Northern California, so those location errors will serve as my uncertainties for the vertical locations of events.

Budgetary and Work Plan Considerations

The time schedule for the research I propose is somewhat difficult to estimate, as the identification of each swarm can take a variable amount of time. However, currently I aim to have the full catalog of swarms with dates, locations, depths, magnitudes and error values recorded by the end of December, with at least half of the swarms that comprise the catalog to be analyzed for spatial and temporal migration patterns, number of events, magnitudes and depths. The materials I will be using are as follows:

1) Waldhauser's 2013 hypoDD earthquake catalog of Northern California
2) MATLAB R2011b to analyze earthquake events, programs, etc.

At this time, I do not anticipate to incur any costs in pursuance of this research, since the MATLAB license and the earthquake catalog are provided to me free of charge. However, additional costs might result from having to purchase software from MATLAB or from another company in order to continue an aspect of my research in more detail, such as being able to better characterize seismic swarm structure.

4) Results

Clustering Analysis Plots: Zaliapin vs. Vidale and Shearer

I first modified or wrote the MATLAB programs for carrying out earthquake catalog analysis outlined in both of the aforementioned methods. To validate my implementation of the Z2008 method, I plotted the log(T)-log(R) scatter plots for nearest neighbor earthquakes in the Northern California catalog. Figure 5 shows that the clustering in time and space (Panel a) and bimodal distribution of nearest neighbor distances (Panel b) previously noted by Zaliapin and collaborators is, indeed, a characteristic of the Waldhauser's Northern California catalog. Therefore, my implementation of the Z2008 method was indeed capable of identifying clustered distributions of quakes. In order to be able to compare these results to the seismic swarms identified by the VS2006 method, I separated this bimodal distribution into two groups – swarm and non-swarm events – with a cutoff between the clustered and non-clustered distributions at $\eta = -6$.

Next, I applied the VS2006 approach, using the same threshold values as used in their study, in order to identify swarms in Northern California. Then, I applied the Zaliapin clustering analysis to the earthquakes identified as swarms with the VS2006 approach, in order to see where they fell on the log(T)-log(R) and nearest neighbor distance plots (Figure 6). Despite the fact that the clustered events distribution is similar for both methods, and the swarm events fall exactly where we expect them to in the log(T)-log(R) diagram, it is evident from the histograms in Figure 6b that the VS2006 method only identifies a very small subset of all swarm events.

To analyze the spatial distribution of swarm seismicity identified using the two methods, I plotted them on a map using a program that reads topography and bathymetry data from the Sandwell Database (Sandwell et al. 2009) and plots a topography section (Figure 7) using a set of specified coordinates to produce the desired figure. To confirm that more events were being identified by the former than the latter, I used a program that reads bathymetry data from the Sandwell Database and plots a topography section (Figure 7) using a set of specified coordinates to produce the desired figure. Running this program, I plotted a topography map of Northern California using the latitude limits of 35 to 42 N and longitude limits of -117 to -127 W. I then plotted the seismic swarm events identified by the Z2008 and VS2006 methods in three sets of plots. The swarm events identified by the Z2008 method totaled 21,248 out of a possible 63,705 for a magnitude of completeness $m_c > 2.0$, or about 33% of the catalog. The swarm events identified by VS2006 totaled 4,684 out of 63,705, or about 7.3% of the catalog.

Noting this difference, I then moved on to analyze whether all of the swarm events were being identified by the respective programs using a Schuster Test. The results of this analysis are plotted in Figure 8. I find that a strongly statistically significant periodicity is present in the Schuster spectrum for the catalog with VS2006-identified events removed; this indicates that not all of the seismic swarms present in the catalog are being identified by the approach. In contrast, removing the swarms identified with the Z2008 approach results in no significant periodicities being present in the Schuster spectrum. I take this as an indication that the Z2008 method identifies all the swarms.

Having confirmed that the choice of thresholds used by VS2006 did not identifying all the seismic swarms present in Waldhauser's Northern California catalog (from here out referred to as NCA), I proceeded to analyze the effects of varying each of the five threshold values. The threshold number of days was varied between 10 and 48, minimum number of earthquakes between 21 and 59, the maximum number of earthquakes preceding a swarm sequence between 0 and 17, and the minimum and maximum distances between 1 and 15 and 0 and 18, respectively. These variations on the threshold values resulted in a total of 5 different swarm catalogs based on the VS2006 approach, which I then compared against the swarm events identified by the Z2008 analysis. I noted how adjusting each threshold parameter affected the number of swarm events identified.

In order to compare the adjusted V&S thresholds to the clustering analysis, I processed the swarm events identified by the adjusted V&S thresholds to find the nearest neighbor distance given by the clustering analysis, and then plotted all five of the nearest neighbor distance distributions given by the adjusted thresholds against the clustering analysis nearest neighbor distance distribution (Figure 8). I used the first 50,000 earthquakes to examine these distributions due to time constraints associated with running the analyses on the entire catalog for each threshold adjustment.

For the time threshold, an increase in the number of days resulted in the distribution becoming much more spread out in space without much increase in the number of events identified. This was the same for the N2 parameter, or the threshold that describes the number of earthquakes (at least 39) after the initial event in the swarm sequence). Increasing the N1 threshold, or the number of earthquakes preceding the initiating event (3 or fewer), resulted in a slight increase in the number of events identified. The minimum and maximum distance thresholds, however, had the largest effect on the number of events identified. As these distances were increased, there was a dramatic increase in the number of events identified and an increase in similarity to the swarm distribution inferred by the Z2008 analysis.

Having completed this analysis, I proceeded to plot the distribution of swarm and non-swarm events in time, depth, and magnitude in order to investigate how the events identified by both methods compared based on these parameters (Figure 9, 10, 11).

In all three of the different plotted distributions, it was evident that while both methods were identifying the same sets of swarms, the VS2006 approach was not identifying all the events constituent of those swarms, while the Z2008 method seemed to be. To investigate why, I decided to manually identify discrepancies between the two plots and identify which swarms the Z2008 approach had identified but that VS2006 had missed.

Preliminary Results of Swarms Identified

The first seismic swarm I have identified from the Z2008 results (Figure 12) consists of 20 events spanning from June 1st, 1986 to June 14th, 1986 (14 days) in Northern California, near San Mateo. The swarm occurs over a total distance of 4 km, with 6 events occurring within the first 2 km and 14 events occurring between 2-4 km. The morphology of the swarm is linear in map view, but does not appear to have much temporal migration (Figure 13). Based on the spatial linearity that the swarm exhibits in conjunction with the rest of the earthquake events in the full catalog, it is reasonable to conclude that the swarm occurs along a fault structure, specifically the San Andreas fault zone, peninsula section (earthquake.usgs.gov).

I determined two reasons why this swarm was missed by the threshold-based VS2006 analysis. For one thing, 20 events is less than the minimum number of earthquakes needed to identify a swarm. Next, while the maximum distance that the swarm migrates in space is just over the 4 km limit, more than 20% of the events in the swarm occur after the 2 km cutoff between 2-4 km. However, as the swarm does appear to be occurring along a fault, this preliminary result might suggest that seismic swarms are more common on faults than previously thought. More research will need to be done to investigate this conclusion.

5) Discussion and Conclusions

It is evident from my analysis and results thus far that while both methods are effective in identifying seismic swarms, the Z2008 clustering analysis method is more effective in identifying not only the sequences which swarm events belong to, but a higher percentage of the events than the arbitrary thresholds proposed by VS2006. However, the VS2006 approach is more effective in only identifying seismic swarm sequences without mistakenly involving other, unrelated. This is in contrast to the Z2008 method, which identifies more earthquakes related to one another in space and time without regard to which swarm those events belong. It is interesting to note that while the two methods identified very different numbers of total swarm events, neither program seemed to be affected by the background seismicity present in the Wauldhauser catalog. In addition, it is also interesting to note that from the time distribution, the VS2006 approach identified no events that occurred in 1985, and more analysis will have to be done regarding why this might be the case.

Further examination of why the method proposed by Vidale and Shearer is not as effective as the method proposed by Zaliapin et al. suggests that the arbitrary parameters set by Vidale and Shearer are too stringent to be able to capture all seismic bursts. Although preliminary analysis of the seismic swarms sets missed by the V&S methods, as well as nearest neighbor distance distribution comparisons would indicate that the minimum and maximum distance thresholds exert the most influence on which swarms are identified, more analysis needs to be done to confirm this conclusion.

For my next steps, I aim to analyze more seismic swarms that the Vidale and Shearer parameters failed to identify fully, or even at all. I will then analyze the spatial and temporal relations of the events constituent of those swarms to the parameters identified by Vidale and Shearer in order to confirm which parameter exerts the most influence over which swarms are identified. In addition, I also aim to identify what thresholds for V&S will most closely match the clustering analysis in terms of distribution and number of events identified. I will therefore re-run the analyses of Vidale and Shearer on the seismic swarms I have obtained in Northern California from my initial analysis using VS2006 and VS2008 to see whether they lead me to different conclusions than those found by Vidale and Shearer in Southern California. If analysis of the catalogs results in the same conclusions, then while the Vidale and Shearer catalog might not be complete, it is not biased. However, if the Z2008 catalog differs from that of the VS2006 catalog, then it is reasonable to conclude that Vidale and Shearer's conclusions regarding swarms are not as supported as might have once been thought. I will therefore run through my time analysis distribution, (Figure 9c) to pull all event distributions that resemble seismic swarms and assemble them into a full catalog of seismic swarms. I will then run these swarms back through both of the programs to see how analysis of these events by both methods differs. I will ideally have this accomplished by December-January.

6) Summary

In summary, I have used two different methods to identify seismic swarms from Vidale and Shearer and Zaliapin et al., respectively and used them on an earthquake catalog identified by Waldhauser in Northern California in order to identify sequences of swarm events in the aim of seeing which method more effectively identified swarm events. I have tentatively concluded from analysis of timing and frequency of events that the Zaliapin et al. method more effectively identifies swarm events in comparison to Vidale and Shearer, and have now moved on to analysis of specific swarm sequences that were identified by the clustering analysis, but were completely missed by Vidale and Shearer in order to better understand how the threshold parameters affect what events are identified or missed in a swarm, as well as to better understand seismic swarm morphology.

Bibliography

[1] Ader, T. J., and J.-P. Avouac. Detecting periodicities and declustering in earthquake catalogs using the Schuster spectrum, application to Himalayan seismicity, Earth Planet. Sci. Lett., 377–378, 97–105, 2013.

[2] Baiesi, Marco, and Maya Paczuski. "Scale-free Networks of Earthquakes and Aftershocks." Physical Review E Phys. Rev. E 69.6, 2004.

[3] Chen, X., and P. M. Shearer. "Comprehensive Analysis of Earthquake Source Spectra and Swarms in the Salton Trough, California." J. Geophys. Res. Journal of Geophysical Research 116.B9, 2011.

[4] Chouliaras, G., G. Drakatos, K. Pavlou, and K. Makropoulos. "Stress Distribution and Seismicity Patterns of the 2011 Seismic Swarm in the Messinia Basin, (South-Western Peloponnesus), Greece." Nat. Hazards Earth Syst. Sci. Natural Hazards and Earth System Science 13.1 (2013): 45-51.

[5] Corral, Álvaro. "Local Distributions and Rate Fluctuations in a Unified Scaling Law for Earthquakes." Physical Review E Phys. Rev. E 68.3, 2003. [1] Holtkamp, S.g., and M.r. Brudzinski. "Earthquake Swarms in Circum-Pacific Subduction Zones." Earth and Planetary Science Letters 305.1-2 (2011): 215-25.

[6] Fischer, Tomáš, and Josef Horálek. "Space-time Distribution of Earthquake Swarms in the Principal Focal Zone of the NW Bohemia/Vogtland Seismoactive Region: Period 1985–2001." Journal of Geodynamics January-March 35.1-2 (2003): 125-44.

[7] Ruppert, N. A., S. Prejean, and R. A. Hansen. Seismic swarm associated with the 2008 eruption of Kasatochi Volcano, Alaska: Earthquake locations and source parameters, J. Geophys. Res., 116, 2011.

[8] Sandwell, David T., and Walter H. F. Smith. "Global Marine Gravity from Retracked Geosat and ERS-1 Altimetry: Ridge Segmentation versus Spreading Rate." *J. Geophys. Res. Journal of Geophysical Research* 114.B1 (2009).

[9] Savage, M. K., V. Ferrazzini, A. Peltier, E. Rivemale, J. Mayor, A. Schmid, F. Brenguier, F. Massin, J.-L. Got, J. Battaglia, A. Dimuro, T. Staudacher, D. Rivet, B. Taisne, and A. Shelley. "Seismic Anisotropy and Its Precursory Change before Eruptions at Piton De La Fournaise Volcano, La Réunion." *J. Geophys. Res. Solid Earth Journal of Geophysical Research: Solid Earth* 120.5 (2015): 3430-458.

[10] Shapiro, S.a., and C. Dinske. "Fluid-induced Seismicity: Pressure Diffusion and Hydraulic Fracturing." *Geophysical Prospecting* 57.2 (2009): 301-10.

[11] "Schuster Spectrum – Searching for Periodicities in the timing of Earthquakes." California Institute of Technology Tectonics Observatory, 5 Oct. 2012.

[12] Umeda, Koji, Koichi Asamori, Ayumu Makuuchi, Kazuo Kobori, and Yuki Hama. "Triggering of Earthquake Swarms following the 2011 Tohoku Megathrust Earthquake." *J. Geophys. Res. Solid Earth Journal of Geophysical Research: Solid Earth* 120.4 (2015): 2279-291.

[13] U.S. Geological Survey and California Geological Survey, 2006, Quaternary fault and fold database for the United States, accessed Nov 12, 2015, from USGS web site: http//earthquake.usgs.gov/hazards/qfaults/.

[14] Vidale, John E., and Peter M. Shearer. "A Survey of 71 Earthquake Bursts across Southern California: Exploring the Role of Pore Fluid Pressure Fluctuations and Aseismic Slip as Drivers." *J. Geophys. Res. Journal of Geophysical Research* 111.B5 (2006).

[15] Waldhauser, F., HypoDD: A computer program to compute double-difference earthquake locations, USGS Open File Rep., 01-113, 2001.

[16] Waldhauser F. and W.L. Ellsworth, A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, Bull. Seism. Soc. Am., 90, 1353-1368, 2000.

[17] Waldhauser, Felix, and David P. Schaff. "Large-scale Relocation of Two Decades of Northern California Seismicity Using Cross-correlation and Double-difference Methods." *J. Geophys. Res. Journal of Geophysical Research* 113.B8 (2008).

[18] Zaliapin, Ilya, and Yehuda Ben-Zion. "Asymmetric Distribution of Aftershocks on Large Faults in California." Geophysical Journal International 185.3 (2011): 1288-304.

[19] Zaliapin, I., and Y. Ben-Zion. Earthquake clusters in southern California I: Identification and stability, J. Geophys. Res. Solid Earth, 118, 2847–2864, 2013.

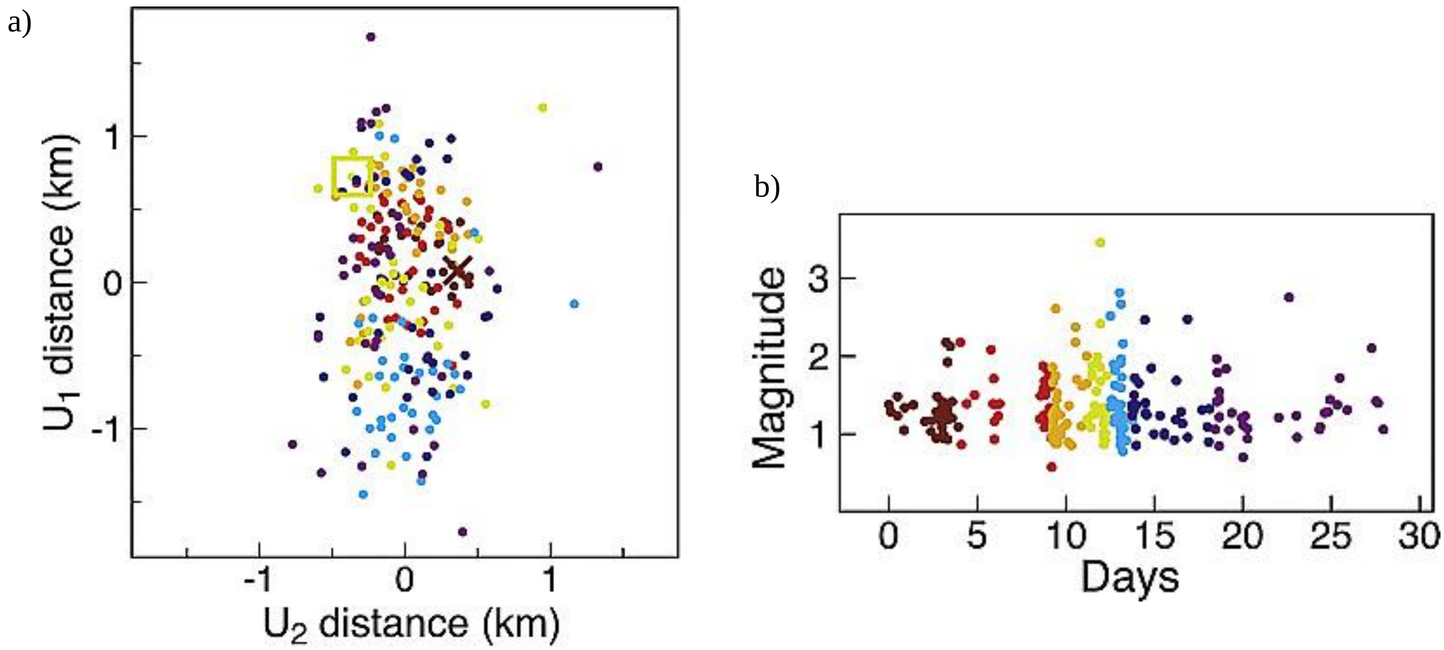


Fig. 1. a) Figure from Vidale and Shearer (2006), illustrating a swarm-like cluster of 230 earthquakes. Colors denote divisions of days. Note the linearity of the morphology of the swarm, which is characteristic of swarm sequences. b) Figure from Vidale and Shearer (2006) demonstrating the distribution of swarm event magnitudes over days. Note that the largest event does not occur at the beginning, but, rather in the middle of the sequence, which is characteristic of swarms.

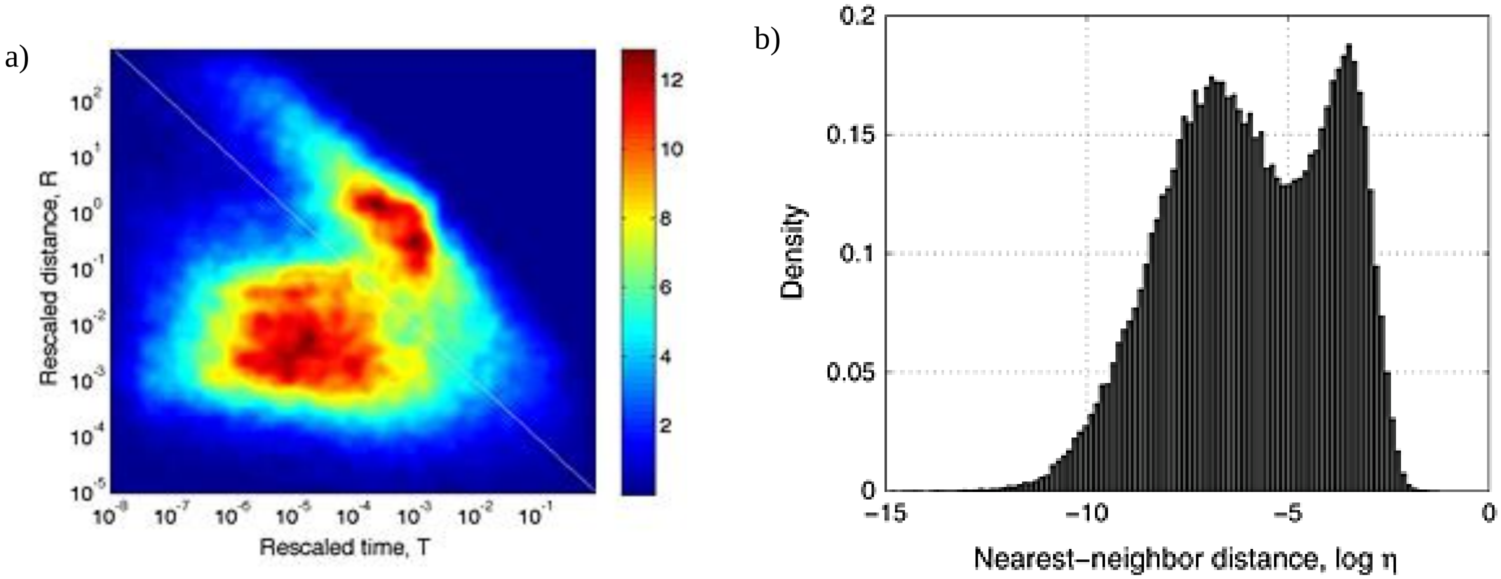


Fig. 2. a) Figure from Zaliapin and Ben-Zion (2013) illustrating the magnitude normalized nearest-neighbor time-distance plot showing a statistically distinct bimodal distribution of the clustered and non-clustered events in space and time. b) Figure from Zaliapin and Ben-Zion (2013) illustrating the bimodal distribution of η that results from the nearest neighbor distance distribution. Clustered (including swarm) events are to the left of -5, non-clustered events to the right.

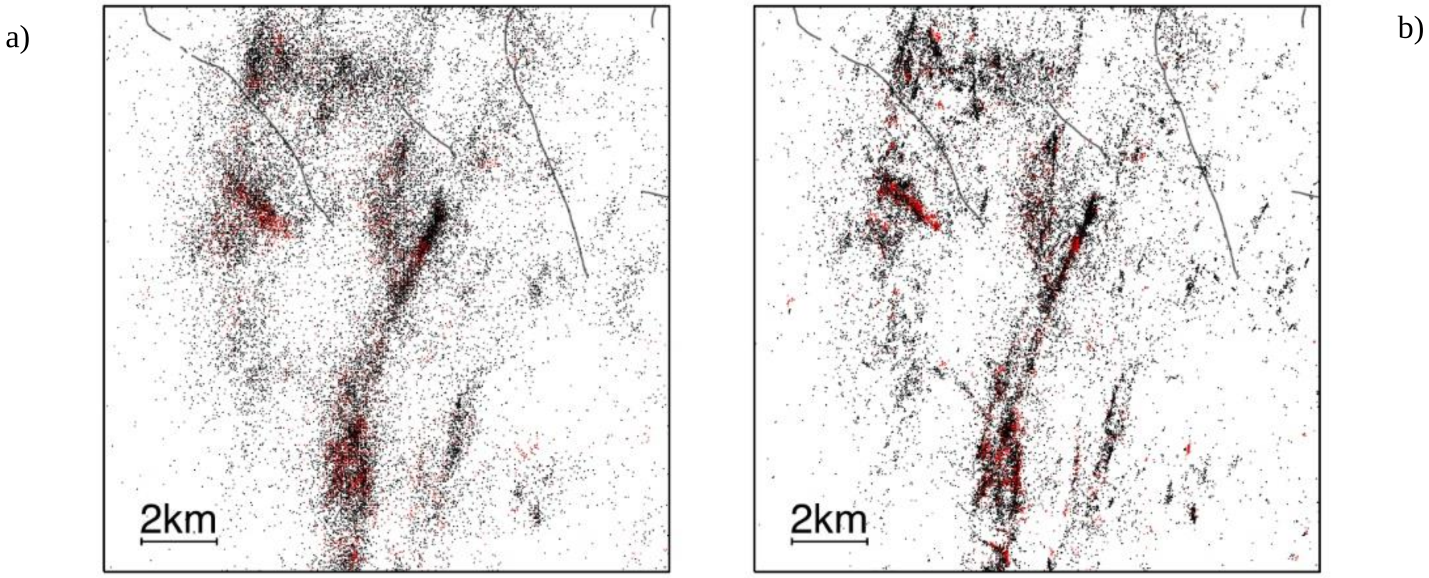


Fig. 3. a) Figure from Waldhauser (2008) demonstrating earthquake locations before application of hypoDD algorithm correction. b) Earthquake locations after application of hypoDD algorithm correction. Note how much more clearly earthquake locations are plotted in space, demonstrating the precision of the correction algorithm.

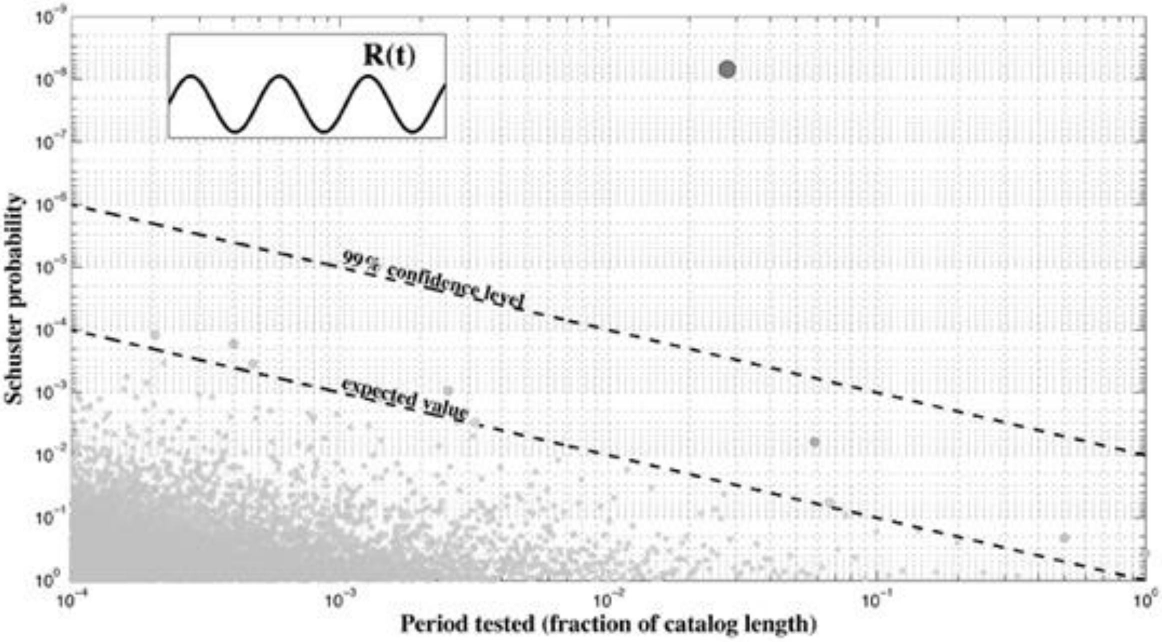
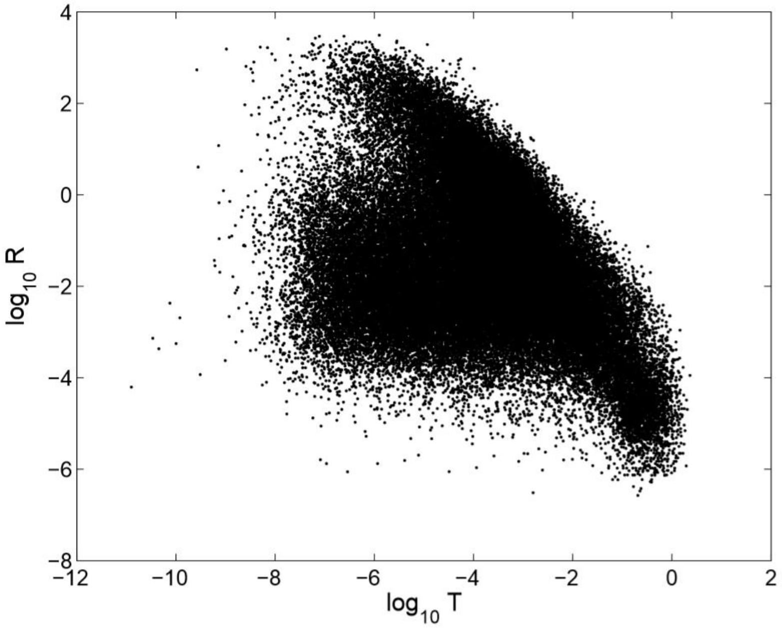


Fig. 4. Schuster plot courtesy of California Institute of Technology Tectonics Observatory (tectonics.caltech.edu). Note the one periodicity above the 99% confidence level, indicating that events are distributed non-uniformly in time with a certain periodicity. In our analyses, such an anomalously high probability could potentially point to the presence of swarm events in the catalog analyzed.

a)



b)

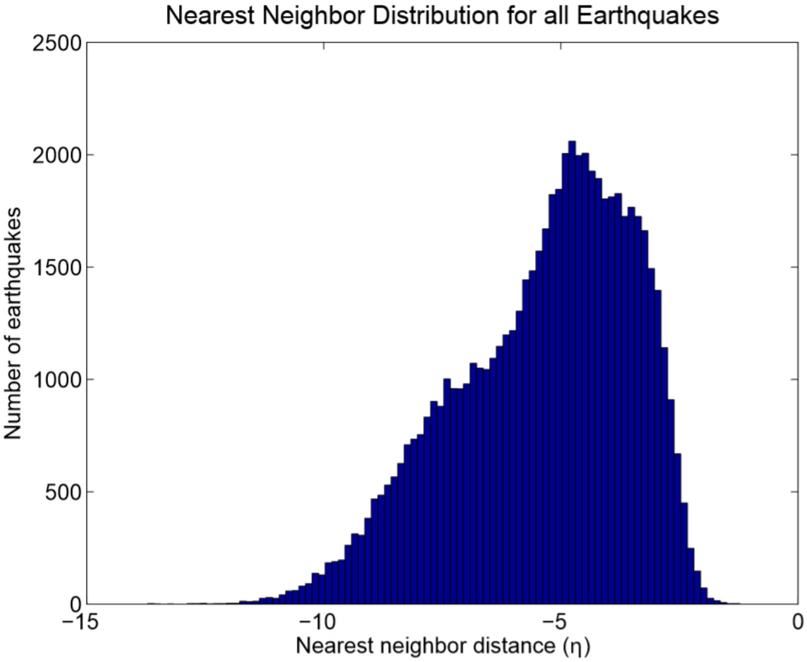


Fig. 5. a) log(T)-log(R) plot. Note the line running linearly to divide the two distributions. b) Nearest neighbor distance distribution. Note the bimodal distribution cutoff at -6 dividing the clustered events from the non-clustered events.

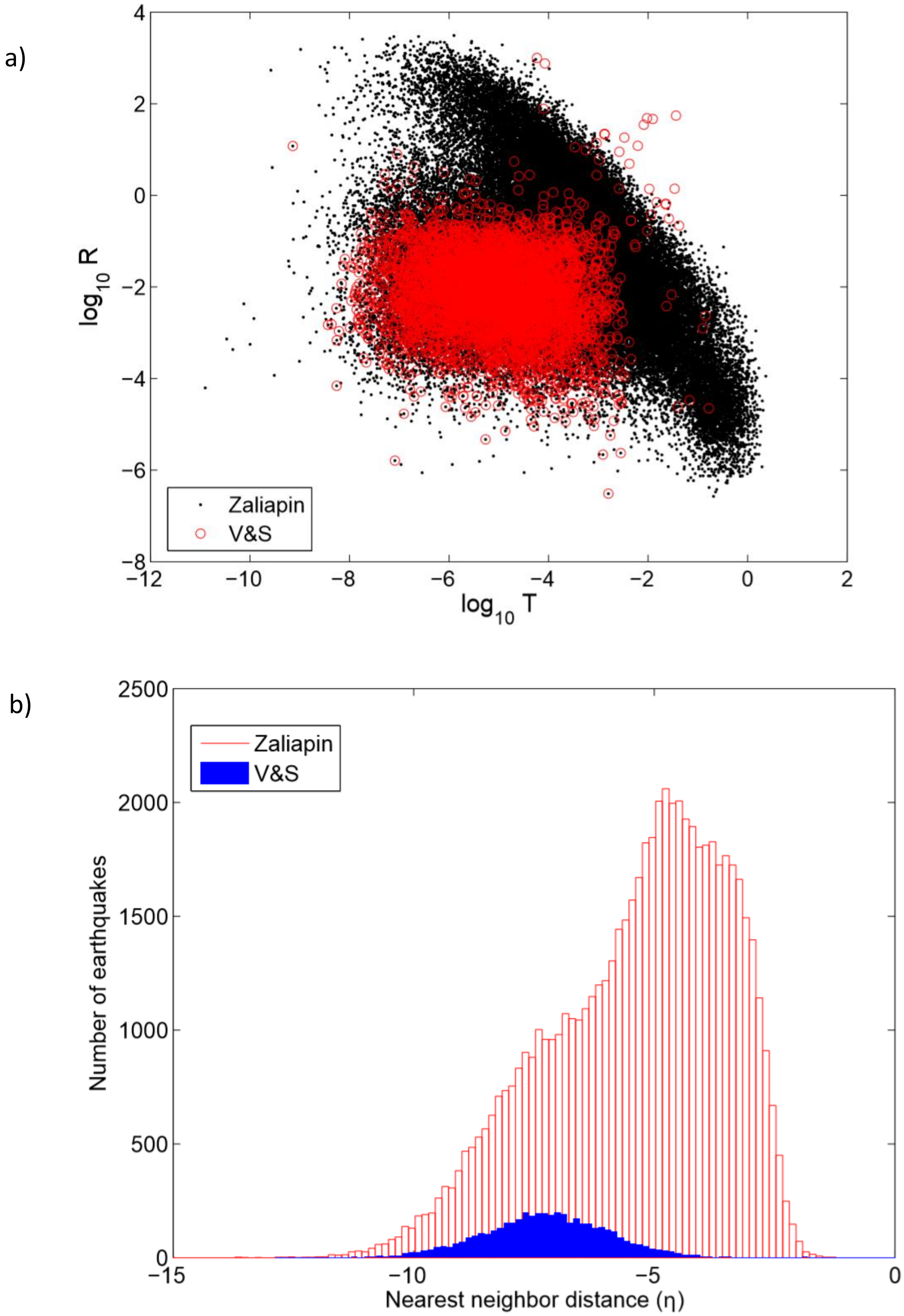
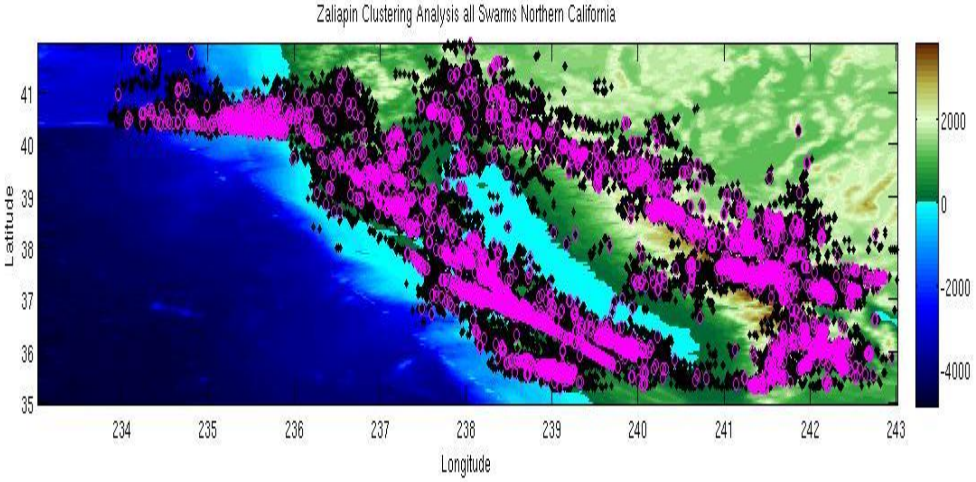
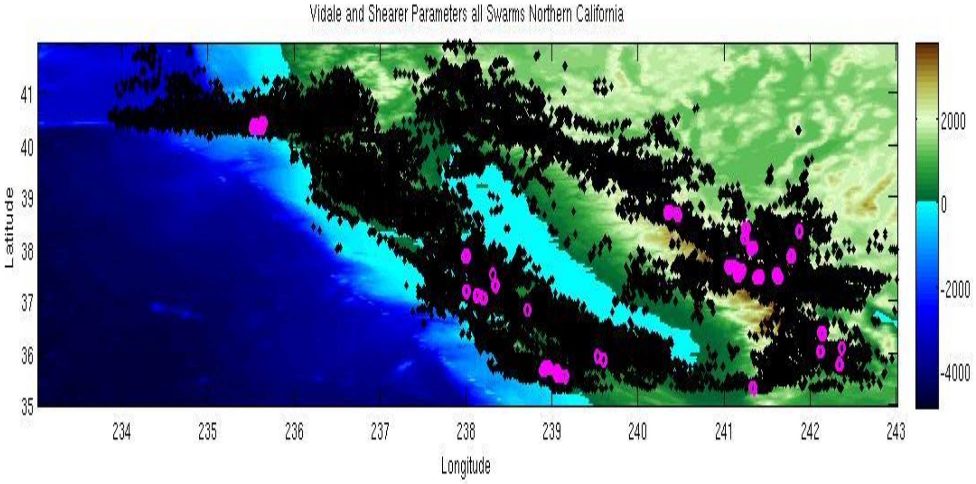


Fig. 6. a) log(T)-log(R) plot following the Z2008 analysis for all earthquakes (black) and earthquakes identified as swarm events by the VS2006 (red) method. Note that the swarm events identified using the VS2006 method fall in the region of the plot expected for clustered seismicity. b) Nearest neighbor distance distributions for all events (red) and events identified as swarms by the VS2006 (blue) method. Note that the VS2006 swarm events are much more normally distributed and fewer in number than the clustered part of the distribution.

a)


Zaliapin Clustering Analysis all Swarms Northern California
Latitude
Longitude
35
36
37
38
39
40
41
234
235
236
237
238
239
240
241
242
243
2000
0
-2000
-4000

b)


Vidale and Shearer Parameters all Swarms Northern California
Latitude
Longitude
35
36
37
38
39
40
41
234
235
236
237
238
239
240
241
242
243
2000
0
-2000
-4000

c)

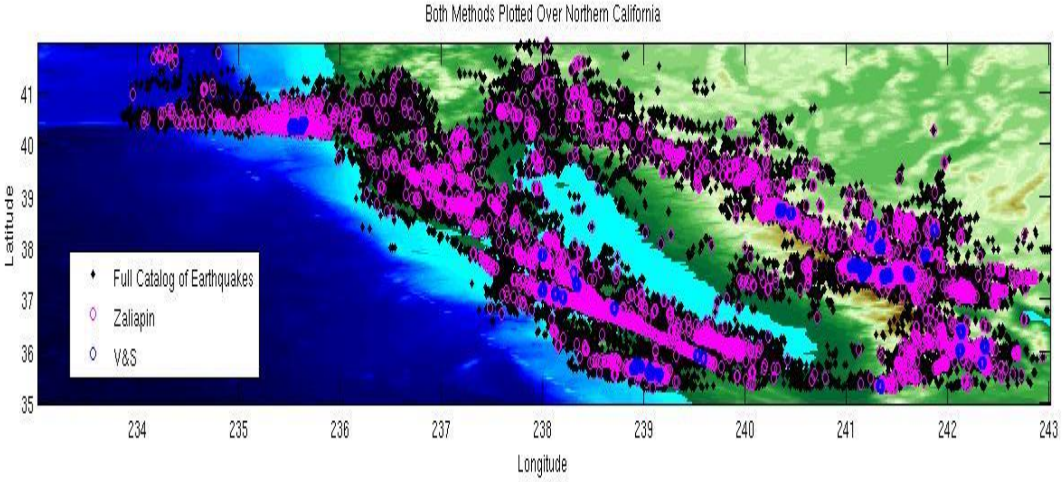

Both Methods Plotted Over Northern California
Latitude
Longitude
Full Catalog of Earthquakes
Zaliapin
V&S
35
36
37
38
39
40
41
234
235
236
237
238
239
240
241
242
243

Fig. 7. a) Full map of swarm distributions over Northern California. In black are all the events contained in Waldhauser's 2013 catalog, in pink are the swarm events identified by Zaliapin. b) Full map of swarm distributions over Northern California. In black are all the events contained in Waldhauser's 2013 catalog, in pink are the swarm events identified by Vidale and Shearer. c) Full map of swarm distributions over Northern California. In black are all the events contained in Waldhauser's 2013 catalog, in pink are the swarm events identified by Zaliapin, and in blue are the events identified by V&S.

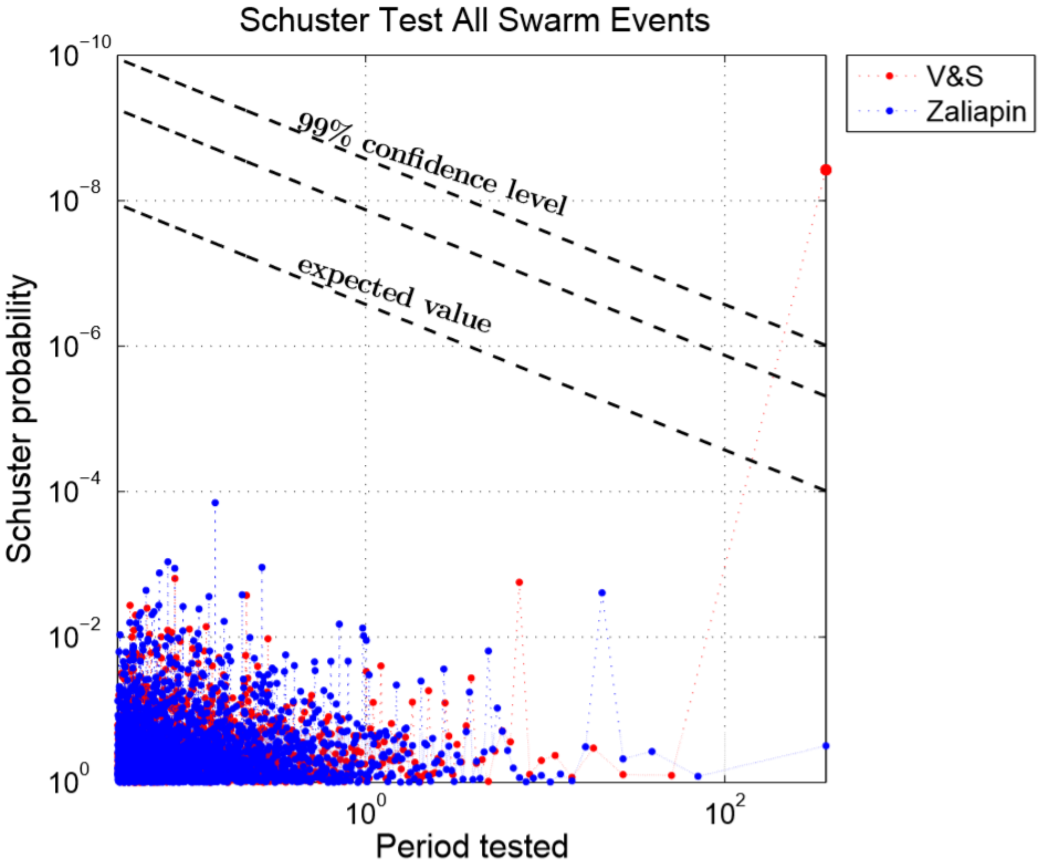
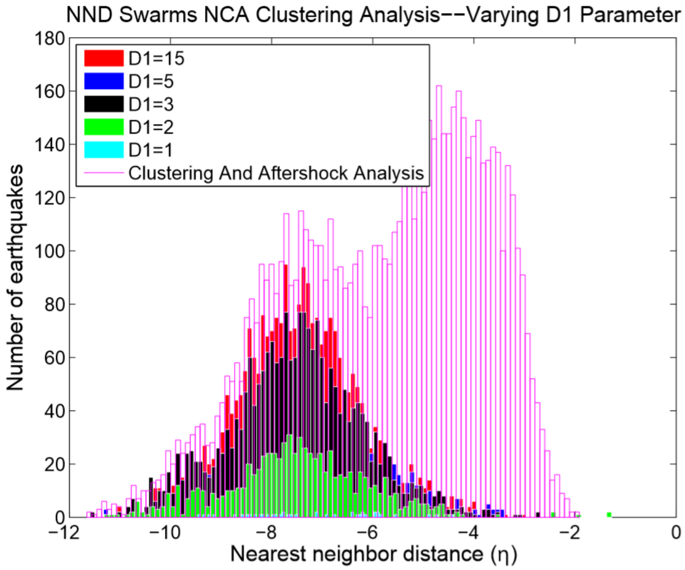


Fig. 8. Schuster spectrum plotted from testing the periodicities present in the data once all swarm events identified by both methods are removed. VS2006 is in red and Z2008 is in blue. Note that the one periodicity in red above the 99% confidence level line demonstrates that there are periodicities present in the data, indicating that VS2006 does not identify all swarm events present Waldhauser's catalog.

a)



b)

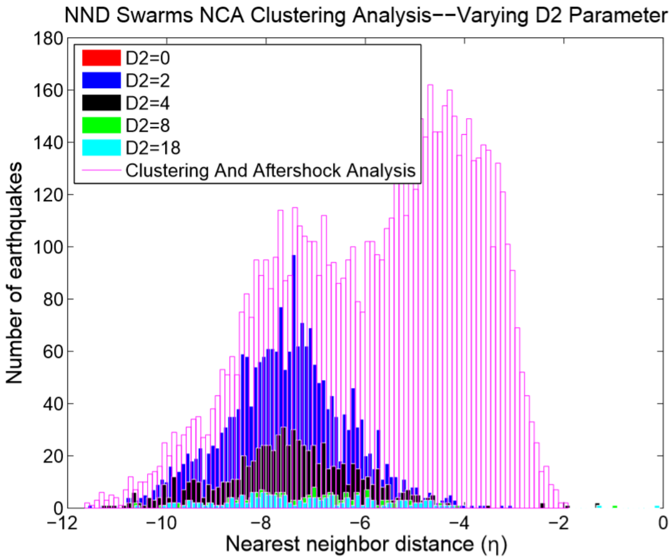


Fig. 9. a) Nearest neighbor distance distribution for threshold variation of the minimum distance parameter. Note that as the minimum distance threshold is increased, the number of events identified increases and begins to match the swarms identified by the clustering analysis. The minimum distance varied is D1=1 due to the fact that the minimum distance cannot be 0. b) Nearest neighbor distance distribution for the threshold variation of the maximum distance parameter. Note that as the maximum distance threshold is increased, the number of events identified increases and begins to match the swarms identified by the clustering analysis. Only these two figures are shown because they show the greatest results for the variation of threshold parameters, and only the first 50,000 earthquakes from the Waldhauser catalog were used to make these figures due to time constraints.

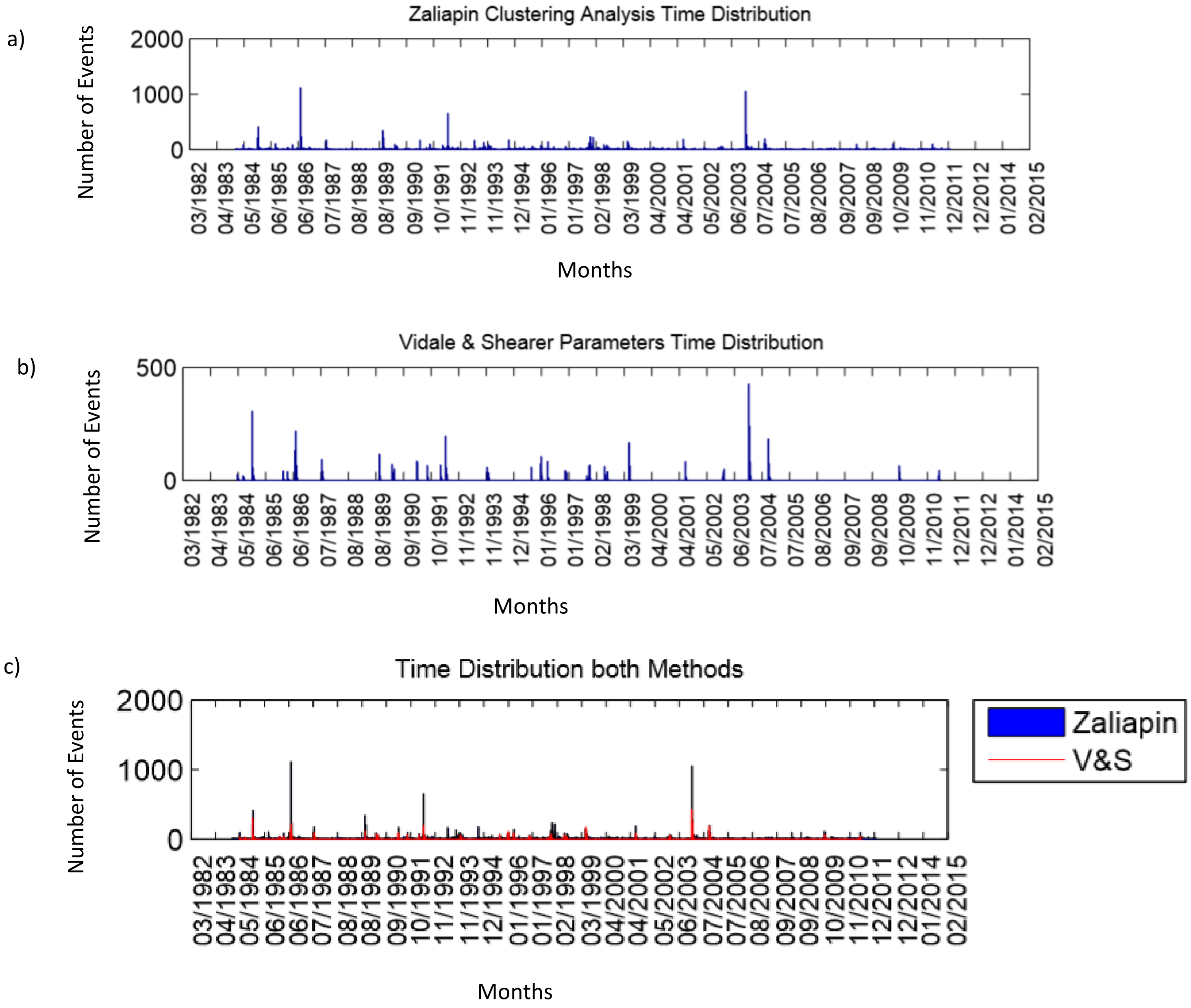


Fig 10. a) Event timing distribution for the Zaliapin method. b) Event timing distribution for the V&S method. c) Event timing for both methods, Zaliapin in blue and V&S in red. Note that the peaks, which correspond to swarm sequences, in the V&S distribution do not reach that of the Zaliapin distribution.

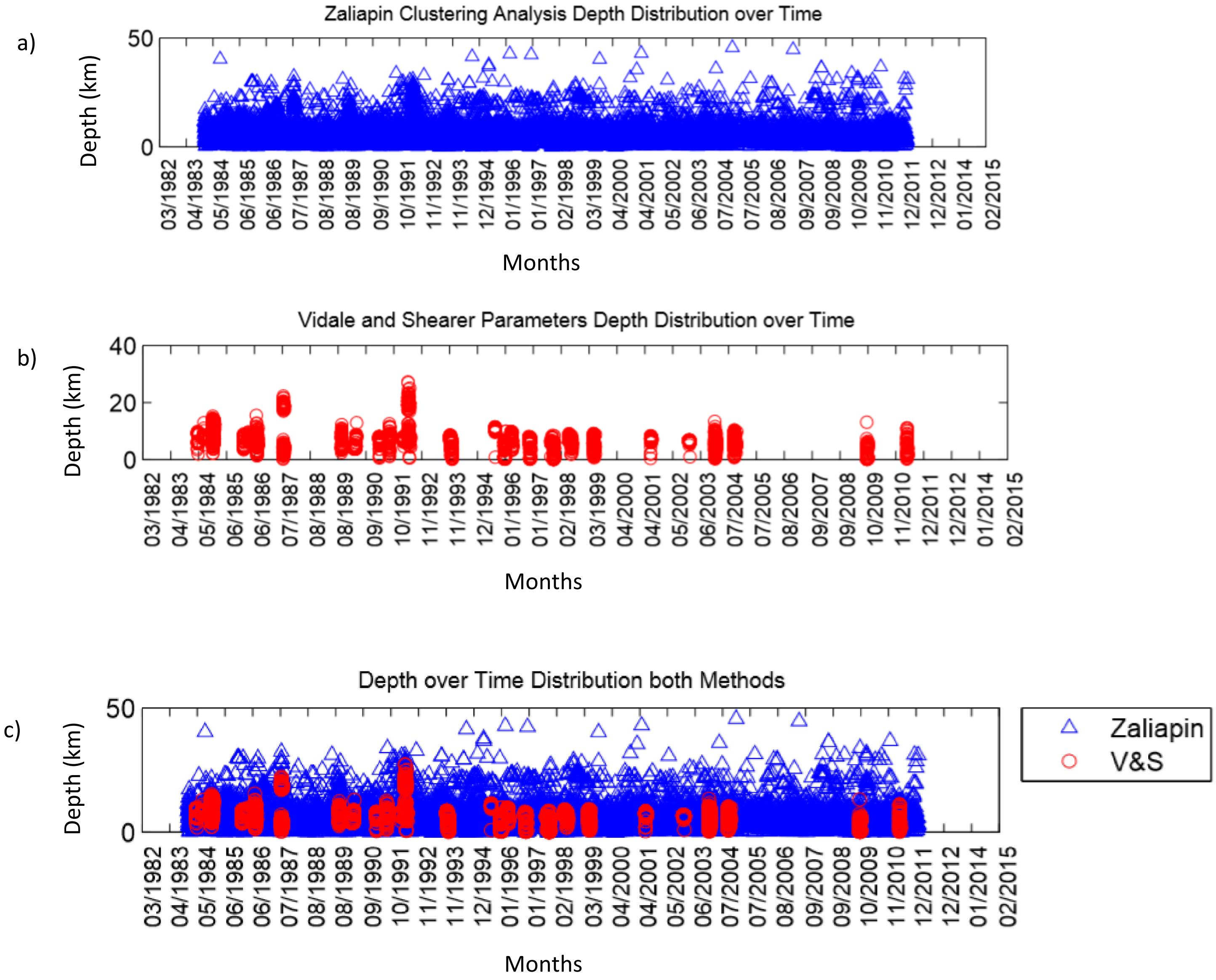


Fig. 11. a) Depth over time distribution for the Zaliapin method analysis. b) Depth over time distribution for the V&S parameters analysis. c) Depth over time distribution for both methods, Zaliapin in blue and V&S in red. Note that the V&S parameters identify the same swarms, but not the same number of events as the Zaliapin method.

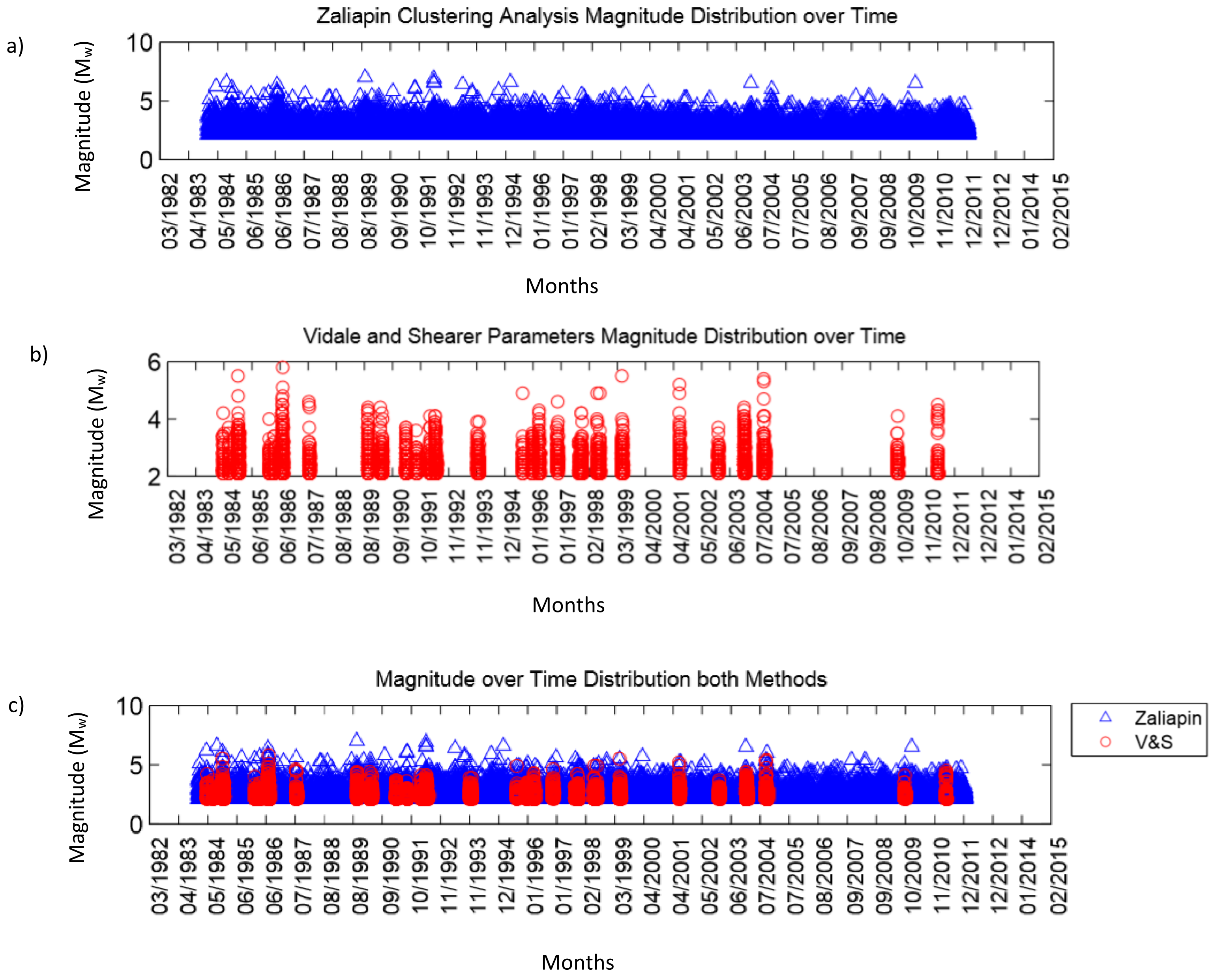
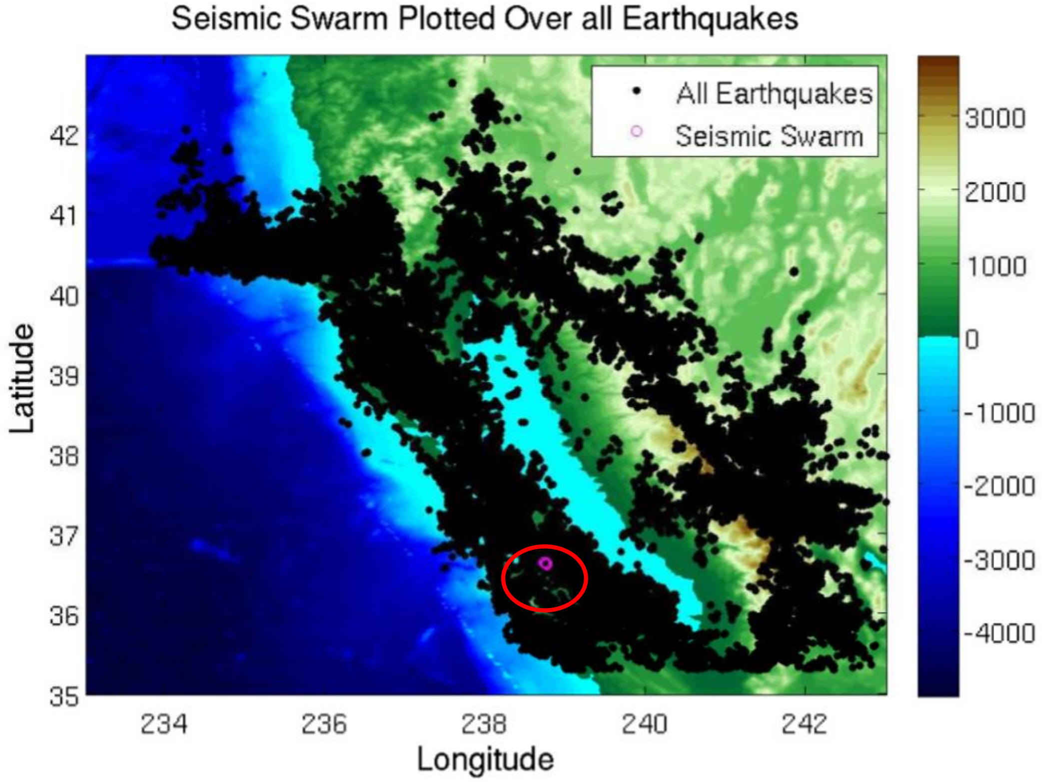


Fig 12. a) Magnitude over time distribution for the Zaliapin method analysis. b) Magnitude over time distribution for the V&S parameter analysis. c) Magnitude over time distribution for both methods, Zaliapin in blue and V&S in red. Note that while V&S identifies the same swarm sequences, it doesn't identify the same number of events constituent of each swarm that Zaliapin does.

a)


Seismic Swarm Plotted Over all Earthquakes
All Earthquakes
Seismic Swarm
Latitude
42
41
40
39
38
37
36
35
234
236
238
240
242
Longitude
3000
2000
1000
0
-1000
-2000
-3000
-4000

b)

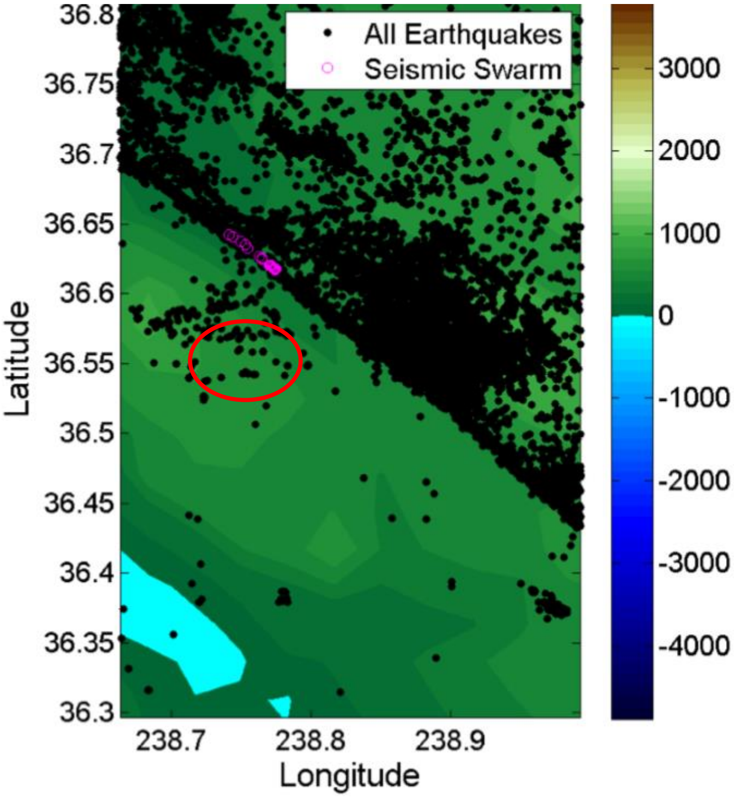

All Earthquakes
Seismic Swarm
Latitude
36.8
36.75
36.7
36.65
36.6
36.55
36.5
36.45
36.4
36.35
36.3
238.7
238.8
238.9
Longitude
3000
2000
1000
0
-1000
-2000
-3000
-4000

Fig 13. Topography map of Northern California. In black are all earthquakes from the Waldhauser catalog, in pink is the seismic swarm. b) Topography map of Northern California, zoomed to show the linear spatial migration of the seismic swarm near San Mateo, along the San Andreas fault. The swarm is circled on both maps in red.

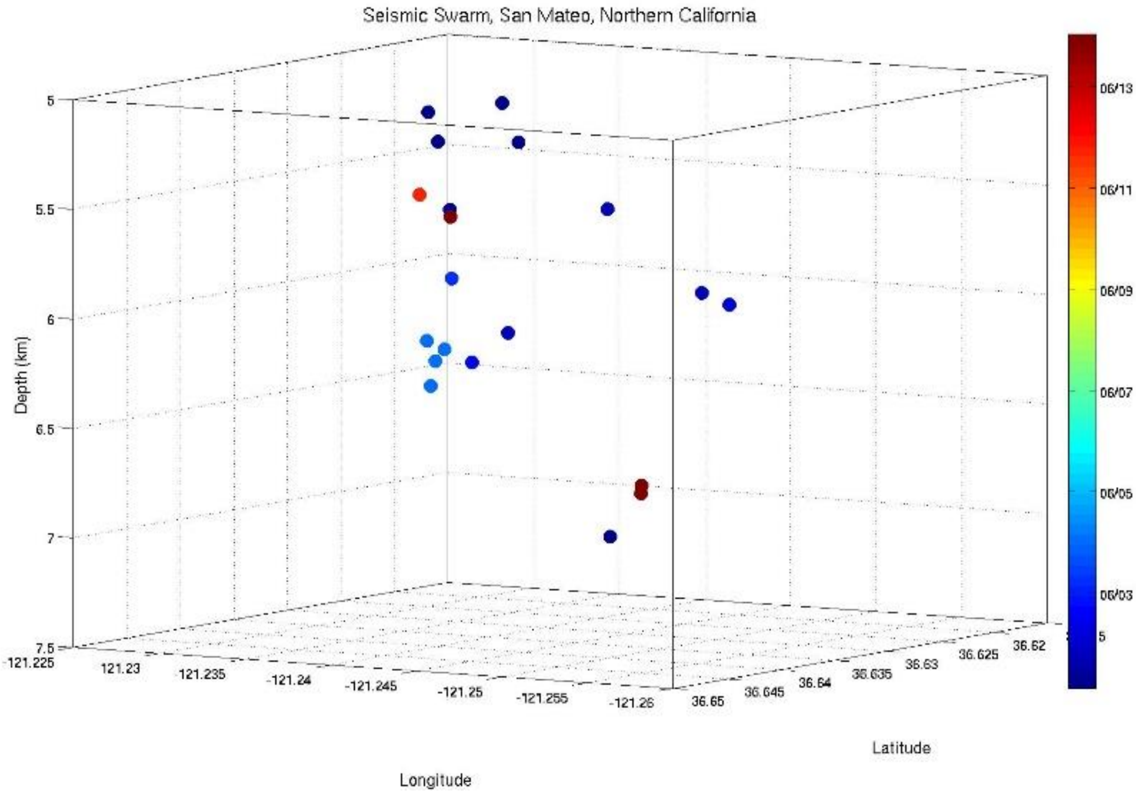


Fig 14. Seismic swarm near San Mateo plotted over longitude, latitude and depth. Colorbar demonstrates range of date. Temporal migration demonstrates loop, rather than linear migration, with last event originating near the first.