

A new method to identify earthquake swarms applied to seismicity near the San Jacinto Fault, California

Qiong Zhang^{1,2} and Peter M. Shearer¹

¹*Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA, USA.*

E-mail: qiongzhang87@gmail.com

²*Seismological Laboratory, California Institute of Technology, Pasadena, CA, USA*

Accepted 2016 February 17. Received 2016 February 5; in original form 2015 August 31

SUMMARY

Understanding earthquake clustering in space and time is important but also challenging because of complexities in earthquake patterns and the large and diverse nature of earthquake catalogues. Swarms are of particular interest because they likely result from physical changes in the crust, such as slow slip or fluid flow. Both swarms and clusters resulting from aftershock sequences can span a wide range of spatial and temporal scales. Here we test and implement a new method to identify seismicity clusters of varying sizes and discriminate them from randomly occurring background seismicity. Our method searches for the closest neighbouring earthquakes in space and time and compares the number of neighbours to the background events in larger space/time windows. Applying our method to California's San Jacinto Fault Zone (SJFZ), we find a total of 89 swarm-like groups. These groups range in size from 0.14 to 7.23 km and last from 15 min to 22 d. The most striking spatial pattern is the larger fraction of swarms at the northern and southern ends of the SJFZ than its central segment, which may be related to more normal-faulting events at the two ends. In order to explore possible driving mechanisms, we study the spatial migration of events in swarms containing at least 20 events by fitting with both linear and diffusion migration models. Our results suggest that SJFZ swarms are better explained by fluid flow because their estimated linear migration velocities are far smaller than those of typical creep events while large values of best-fitting hydraulic diffusivity are found.

Key words: Earthquake interaction, forecasting, and prediction; Seismicity and tectonics; Statistical seismology.

1 INTRODUCTION

Clusters are common across earthquake catalogues in the form of both main-shock–aftershock (M–A) sequences and swarms. The aftershock sequences triggered by main shocks have been well studied and can be simulated by statistical modelling of event-to-event triggering, such as the Epidemic-Type Aftershock Sequence (ETAS) model, which bears on the empirical Omori's law and Gutenberg–Richter (G–R) magnitude–frequency relationship (e.g. Ogata 1988; Helmstetter & Sornette 2002). In contrast, swarms are typically but somewhat vaguely defined as clusters without obvious main shocks, or more generally as clustering that does not appear to result primarily from earthquake-to-earthquake triggering. Characterizing swarms is important for understanding their underlying physical driving mechanisms and the physical properties of areas prone to swarms. Swarm occurrence is often associated with external aseismic processes, such as fluid migration and aseismic slip (e.g. Hill 1977; Vidale & Shearer 2006; Lohman & McGuiire 2007). Other studies have suggested that areas of high heat flow, which results in

low viscosity and small accumulated elastic strain, tend to generate more swarms than M–A sequences (Ben-Zion & Lyakhovskiy 2006; Yang & Ben-Zion 2009).

To fully characterize earthquake swarms, it is desirable to compile as complete a swarm catalogue as possible. Previous swarm studies mainly have applied two kinds of approaches: (1) Select clustered events with higher densities in fixed space and time windows than background events, and then separate swarms from M–A sequences (Vidale & Shearer 2006); and (2) Fit earthquake catalogues with ETAS parameters, including background seismicity rate, aftershock productivity and branching extent, and identify sequences with anomalous parameters or an excess of events compared to the ETAS model predictions (e.g. Hainzl & Ogata 2005; Enescu *et al.* 2009; Llenos *et al.* 2009; Roland & McGuiire 2009). These sequences are likely to be swarm-like in character because most aftershock sequences will be predicted by the ETAS model. Here we largely adopt the first approach, but implement a variable-window-size method to naturally and automatically accommodate a wide range of cluster sizes.

We choose the San Jacinto Fault Zone (SJFZ) as a proof-of-concept test site because it may host a number of swarms due to possible deep creep (Wdowinski 2009) and the presence of geothermal fluids, especially near its southern end. As one of the most seismically active regions in southern California, the SJFZ has abundant microseismicity and five $M \geq 5$ earthquakes have occurred there since 1981. Due to the difficulty of detection at depth, clear observations of slow slip within the SJFZ have not been established. However, slow slip has been hypothesized for some time because of observations near Anza, where strainmeter data show anomalies possibly related to slow slip (Agnew & Wyatt 2005) and substantial afterslip at depth following some earthquakes can be inferred from strain and aftershock data (Inbal *et al.* 2013). Furthermore, the moderate earthquakes to the south of Anza are observed to have relatively large aftershock productivity (Felzer & Kilb 2009). High-frequency tremor, which is often observed to accompany slow slip, has also been found near Anza (Gomberg *et al.* 2008; Wang *et al.* 2013). In addition, the southern SJFZ includes patches near the Salton Sea and Salton Sea geothermal field, where numerous swarms have been observed (Lohman & McGuire 2007; Chen & Shearer 2011). The objective of this paper is to characterize the features of the SJFZ swarms identified by our new approach, including their spatial and temporal distributions and migration behaviour.

2 DATA AND METHOD

Because swarms are often very compact (i.e. less than 1 km across), it is desirable to use an earthquake catalogue that has been relocated using high-resolution methods. Currently the most accurate and complete earthquake catalogue in southern California is the relocated HYS catalogue (Hauksson *et al.* 2012), which results from differential traveltimes relocation from waveform cross-correlation and absolute traveltimes relocation using an improved 3-D velocity structure. From the HYS catalogue, we choose events between 1981 and 2014 within 20 km away from the strike of the SJFZ. The magnitude completeness is estimated to be 2.0 for this region according to G–R magnitude–frequency analysis. However, the majority of the earthquakes in the catalogue are smaller than M 2.0 and many swarms are defined largely by these smaller events. Thus, for the purpose of detecting the maximum number of swarms we analyse

77 377 SJFZ earthquakes down to M 1.0, recognizing that the catalogue is not complete to this level. Because it is improbable that the detection threshold varies significantly over short time and length scales, it is unlikely that this results in any false detection of clusters.

2.1 Identification of clusters

We describe clusters using the nearest-neighbour approach from Zaliapin & Ben-Zion (2013a), which provides a general way to separate clusters from background events. The distance η between two events in space and time is defined as

$$\eta = dt \times dr^d$$

where dt is the time separation between the two events, dr is the 3-D space separation and d is the fractal dimension. We experimented with different values of d and found that the cluster identifications did not change very much, thus we use $d = 1.6$ following the study of (Zaliapin & Ben-Zion 2013a). Note that we modify the definition of η in Zaliapin & Ben-Zion (2013a) by dropping the magnitude dependence because we seek to identify all clusters, not just aftershock sequences explained by earthquake-to-earthquake triggering models; thus we do not search for the most likely triggering events among events before the target event (Zaliapin & Ben-Zion 2013a) but instead search for the closest events to each target event (our study). Note also that the specific units used to measure dt and dr are unimportant because we consider only relative values of η . For a given target event, the event in the catalogue with the smallest η value is termed its nearest neighbour.

Our algorithm aims to detect event clustering in space and time by comparing local event density with the surrounding distribution of ‘background’ events. We sequentially treat each event in the catalogue as a target event and find its n nearest neighbours (for n from 3 to 200) from subsequent events. For each value of n , we save details of the space/time window, that is, the maximum temporal and spatial distance (t_{\max} and r_{\max}) from the target event. This defines a reference space/time window, which is shown in green in Fig. 1. Similar to the idea of STA/LTA (short-time-average through long-time-average) triggering algorithms (e.g. Vanderkulk *et al.* 1965), we also define a larger ‘background’ window that scales with t_{\max} and r_{\max} . This window includes: (1) events with distance $dr \leq 3r_{\max}$

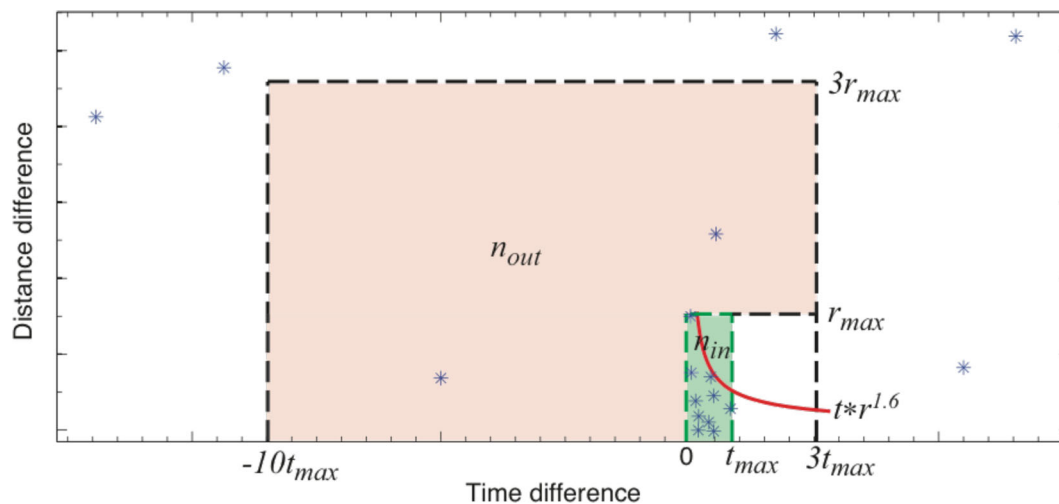


Figure 1. Illustration of our cluster detection method. A cluster is identified when the number of neighbours ($n_{\text{in}} = 10$) within a window of r_{\max} km and t_{\max} days (green) is significant compared to the number of background events ($n_{\text{out}} = 2$) in a larger window (pink). The maximum neighbouring distance from the target event is shown by the red curve.

and time $|dt| \leq 10t_{\max}$ prior to the target event, and (2) events within $r_{\max} \leq dr \leq 3r_{\max}$, and an interval of $dt \leq 3t_{\max}$ following the target event. This background window is shown in pink in Fig. 1. Our strategy for identifying space/time clusters is to find seismicity groups with a relatively large number of events in the reference window compared to the background window, which both precedes and surrounds the reference window. The reference window dimensions of $10t_{\max}$ and $3r_{\max}$ are somewhat arbitrary and were chosen following limited experimentation with other values; we find that they work quite well in discriminating seismicity clusters from more random seismicity for our data set. We also tested background spatial windows of $2r_{\max}$ and $6r_{\max}$, and found that they yielded numbers of clusters are only about 10–15 per cent different from our preferred results. Note that to avoid including possible late aftershock events, we do not include events within $dr \leq r_{\max}$ and $t_{\max} \leq dt \leq 3t_{\max}$ as part of the background window. This is consistent with the fact that aftershock sequences usually decay more slowly in time (p , Omori's law parameter, typically $p = 1$) than in space ($q = 1.5$ to 2.5; e.g. Shearer 2012). Comparing the number of events falling in the windows, we quantify the clustering strength as

$$Q = \frac{n_{\text{in}}}{n_{\text{out}} + 1}$$

where n_{in} is the number of neighbours in the reference window and n_{out} is the number of events falling in the background window. For each target event, the largest value of Q (Q_{\max}) from the 198 values of n is chosen and the corresponding n_{in} neighbours are saved as daughter events in the cluster group.

Our next step is to remove duplicate clusters. In the case that a target event is included as a daughter event in other cluster groups with larger values of Q_{\max} , we remove the target event and its group. In addition, when one daughter event has multiple parent events with different Q_{\max} values, we keep the parent event with the largest value of Q_{\max} . From our catalogue of 77 377 events this results in a

total of 24 689 possible clusters. However, most of these have low Q_{\max} values or very small numbers of events. To focus on the most clearly defined clusters, we analyse only groups with $Q_{\max} \geq 2$ and containing at least 10 events. Using these criteria, a total of 179 clusters are found in our study area. Three examples of identified clusters are shown in Fig. 2.

2.2 Separation of swarms and M–A sequences

From the 179 clusters, we then attempt to separate the swarm-like and M–A sequences. The clearest difference between these groups is seen in their magnitude–time distribution. Swarm-like sequences do not usually start with their largest event, while in M–A sequences the main shocks tend to occur near the beginning of the cluster. To quantify this magnitude behaviour, we adopt two measures: the timing of the largest event t_m normalized by the mean value, and the skewness value μ . For a given sequence, the skewness of the moment release with time can be calculated as (Roland & McGuire 2009; Chen & Shearer 2011)

$$\mu = \frac{\sum_1^N (t_i - t^*)^3 M_i}{\sigma^3},$$

where the numerator is the third central moment of moment release, t^* is the centroid occurrence time, the mean value of t_i weighted by moment release (M_i) and σ is the standard deviation of central moment.

For the 179 clusters, the median value of t_m is 0.77 and the median μ is 1.77. Note that because a large aftershock sequence may fall into several smaller clusters, to avoid the artefact that parts of aftershock sequences are treated as swarms, we include events within the reference window and its following surrounding window to calculate the values of t_m and μ for each cluster. The t_m parameter is a robust measure of the timing of the largest event, but is

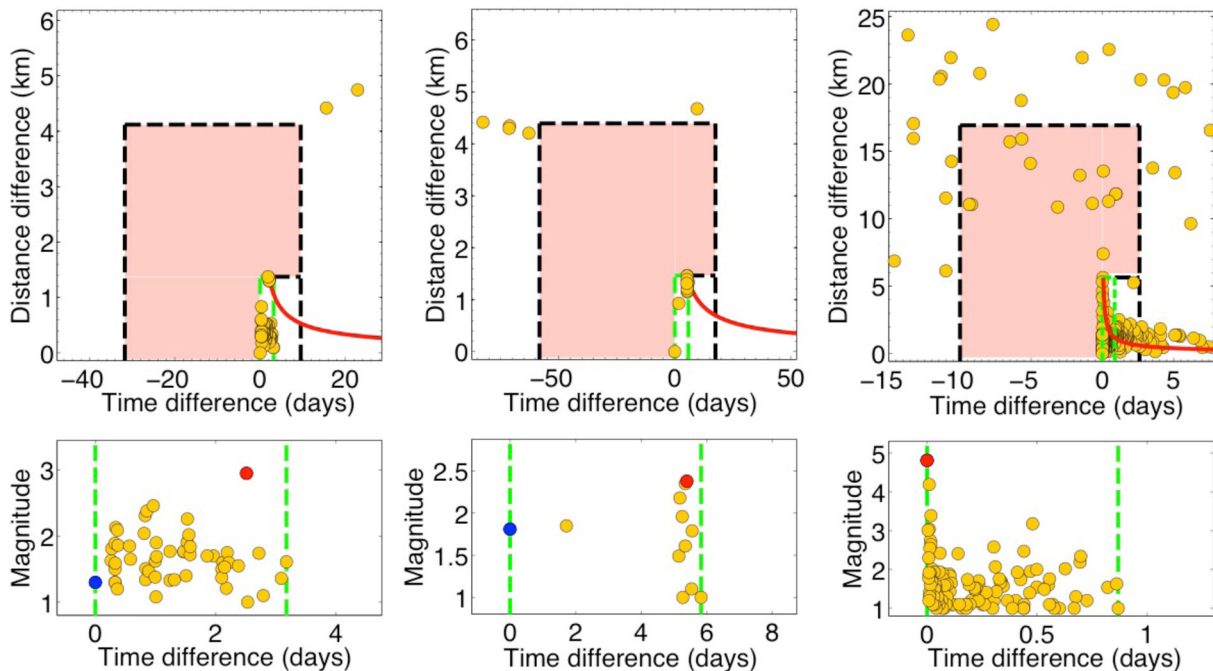


Figure 2. Three examples of identified seismicity clusters including, from left to right, one large swarm, one small swarm and one main-shock–aftershock (M–A) sequence. The spatiotemporal windows for each group are shown in the top panels. Symbols are the same as used in Fig. 1. Below are the distributions of magnitudes versus time. Note that the largest events (red dots) occur later in swarm-like sequences and earlier in the M–A sequence. The target events, also the first events in each sequence, are denoted by blue dots.

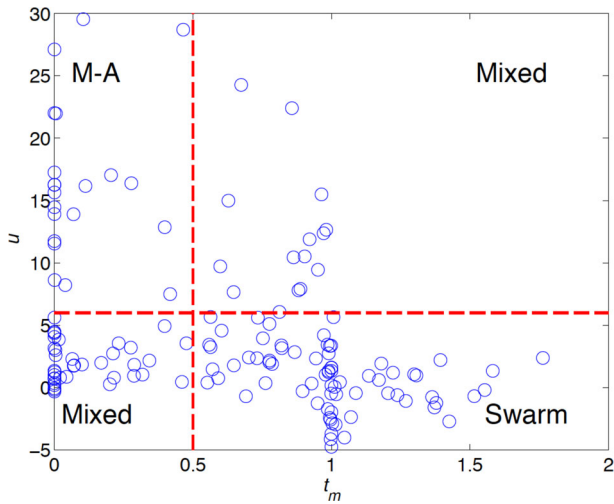
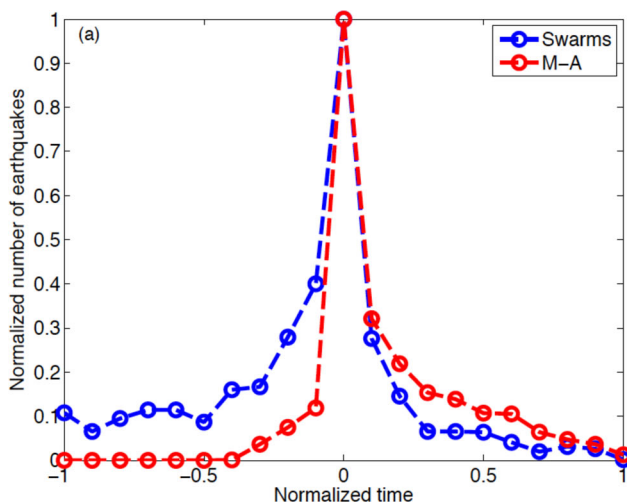


Figure 3. Skewness μ and t_m for 179 seismicity clusters. We have divided these clusters into swarm, M–A and mixed type categories using empirical thresholds.

not a good measure of the overall moment distribution, which may result in misidentification of foreshock–main–shock sequences as swarms. While skewness measures the overall asymmetry of the moment distribution, it is sensitive to the duration of sequence, that is, a long duration can bias μ to a larger value, and may involve misidentification of swarms as M–A sequences. Therefore, we categorize the clusters into swarms and M–A sequences using both measures.

As the largest event tends to occur later in swarm-like sequences and earlier in M–A sequences, we expect that swarm-like sequences have relatively large t_m and small or negative μ , while M–A sequences have small t_m and large positive μ . A sequence with $t_m > 0.5$ indicates events occur in the later part of the sequence, and the empirical threshold to differentiate swarms and M–A for μ is around 5 (e.g. Roland & McGuire 2009). Fig. 3 shows that 71.5 per cent of the clusters fall into the two categories using both measures. Namely, a total of 89 clusters with $t_m > 0.5$ and $\mu < 6$ are identified as swarms, and 29 clusters with $t_m \leq 0.5$ and $\mu \geq 6$ as M–A sequences. We assign the rest of the clusters to a mixed type category. These 61



clusters cannot be definitively identified as M–A or swarms but contain a diversity of behaviour. For example, it can include sequences with the largest event occurring at the beginning, and multiple other events nearly as large occurring throughout the sequence, which leads to a small t_m and a small μ . While some of these clusters are more ‘swarm-like’ than others, we do not attempt to divide them any further and focus only on the other two categories, where more definitive identifications are possible.

Fig. 4 illustrates the differences between M–A sequences and swarms, showing their temporal distribution relative to the largest event, and the relationship between their skewness values and a measure of their spatial migration with time. Earthquakes in many swarms are observed to exhibit spatial migration (e.g. Hainzl 2004; Lohman & McGuire 2007; Chen & Shearer 2011). As a rough estimate of whether the sequences migrate in space, we calculate the spatial separation between the centroids of the first half and the second half of the sequences, and normalize it by the radius of the cluster, defined as the median distance from the centre. For each type, we shift the event occurrence time relative to the largest event in each sequence and stack the number of events from all sequences in time windows normalized by the largest time (Fig. 4a). For the M–A-like sequences, most events occur following the largest event with power law decay. In contrast, the swarm-like clusters are much more time symmetric with respect to their largest event, as expected since the largest earthquake usually occurs relatively late in swarm sequences. Generally, swarm-like sequences have larger separation distances and smaller skewness values, and M–A-like sequences have smaller separation distances and higher skewness values (Fig. 4b). However, the variety of separation distances suggests that not all swarms migrate and that they may migrate with different velocities. We will quantify the migration behaviour more directly in Section 3.3.

3 RESULTS

3.1 Map views

We map the spatial distribution of the 2035 swarm-like events between 1981 and 2014 in Fig. 5, in which large and small clusters

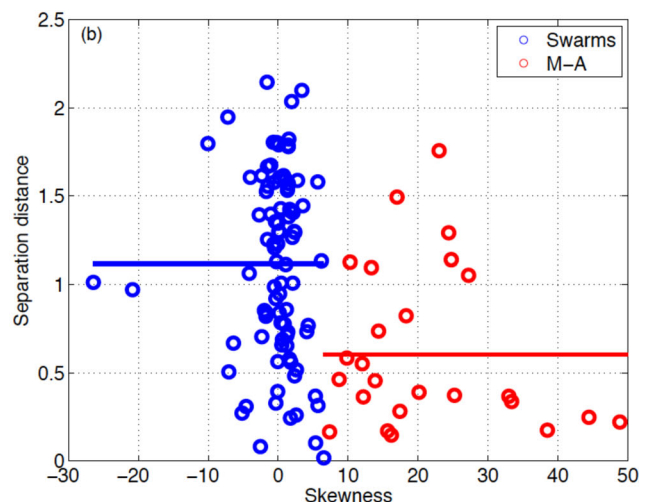


Figure 4. Difference of main–shock–aftershock (M–A) sequences and swarms in terms of temporal distribution, skewness values and separation distance. (a) Stacked number of events versus normalized time relative to the largest event. Note that a large number of events occur prior to the largest event for the swarm-like groups while the M–A-like sequences are dominated by events following the largest event. (b) Relationship between skewness values and separation distance. Each circle denotes one cluster group. The mean values of separation distance for each type are given by lines in different colours.

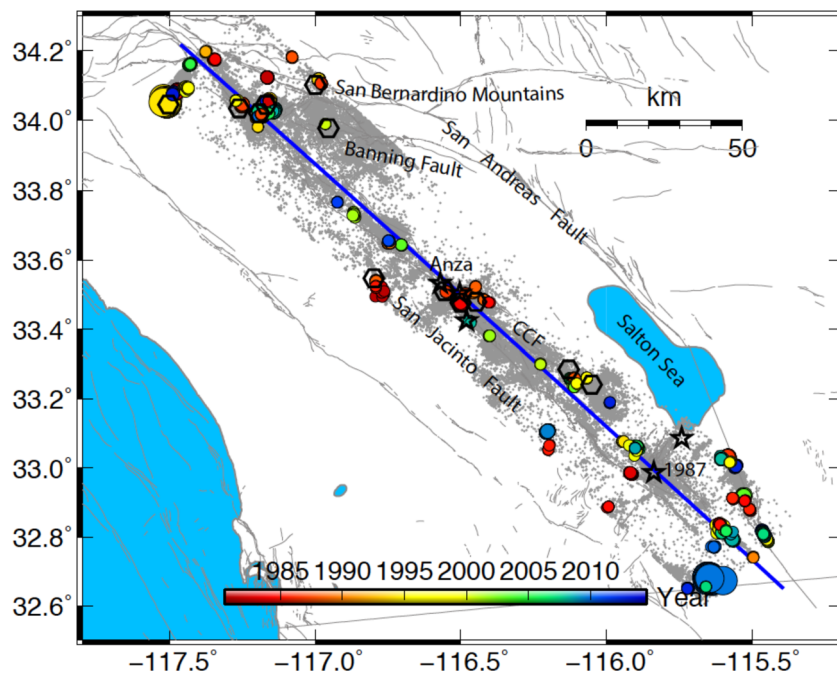


Figure 5. Map view of the detected swarm (solid circle) events in the SJFZ. The circles are colour-coded by occurrence time and scaled with the total number of events in each group. Hexagons represent the swarms identified from Vidale & Shearer (2006). Stars mark five $M \geq 5$ earthquakes including two Superstition Hills in 1987 and three occurring south of Anza. Blue line shows the strike of the SJFZ, along which cross-sections are shown in Figs 5 and 6. Grey dots denote the 77 377 catalogue events between 1981 and 2014. CCF: Coyote Creek Fault. Note that the events shown on the map are only SJFZ events, and the events in surrounding faults are masked.

are defined based on the number of events in each group. These swarms span our study region but do not appear to be distributed uniformly within the seismicity. The northern and southern ends of the SJFZ have more frequent and larger swarms, such as the area in the San Bernardino Mountains near the intersection of the San Andreas Fault and the SJFZ, and the area close to the Salton Sea. Regions of the SJFZ with a high background seismicity rate do not always produce swarms, such as the areas close to the Banning Fault and the Coyote Creek Fault (CCF) and the region near south of the 1987 $M 6.6$ Superstition Hills.

In the same study area, Vidale & Shearer (2006) identified 12 swarms, each of which contains at least 40 events occurring within 4 weeks and a 2 km radius. The swarms identified from our method range from 0.14 to 7.23 km in radius and last from 15 min to 22 d. Because swarm sequences can contain gaps in space or time, the parent and daughter events for a certain sequence identified using two different methods might be different. We map the locations of the 12 swarms from Vidale & Shearer (2006) and find that they coincide with swarms in our analysis (Fig. 5). In general, our results add additional swarms compared to Vidale & Shearer (2006) and span a wider range of sizes.

3.2 Cross-sections

To examine the spatial distribution of swarms in more detail, we plot a depth cross-section along the strike of the SJFZ (Fig. 6a). For comparison, we also plot the cross-section of our detected 29 M–A sequences (Fig. 6b). While the occurrence of both types of clusters is not homogeneous, swarms are distributed differently from M–A events. The two ends of the fault are dominated by numerous large swarms, while sparser and smaller swarms occur in the central segment, which we define as from 50 km to 150 km along the

strike from the northern end point (Fig. 6a). M–A events are mainly concentrated in two regions where three $M \geq 5$ earthquakes occurred near Anza and further south where the two Superstition Hills earthquakes occurred. The other regions with a moderate number of M–A events (about 10–30) do not correlate with the occurrence of swarms. Fig. 6(c) plots histograms of catalogue and swarm events along the fault strike. The distribution of swarms does not always follow the background events. The background events are distributed relatively evenly along strike with two peaks in the northern and central segments, whereas the highest swarm rate occurs at the southern end and the central segment hosts the lowest swarm rate. Fig. 6 also shows the maximum depth of seismicity becomes shallower from north to south, especially at the Anza gap and the CCF, which could be caused by regional heat flow changes (Doser & Kanamori 1986). Fig. 6(d) shows that the depths of catalogue events are broadly distributed and extend to about 17 km depth, while the swarms tend to occur at shallower depths with a peak around 8 km and few swarms below 10 km.

We now consider the temporal distribution of the SJFZ swarms and explore their relationship with large regional earthquakes. The non-swarm part of the catalogue appears consistent with a roughly constant background rate (i.e. Poissonian), punctuated by clusters caused by aftershock sequences. However, SJFZ swarms appear to exhibit temporal variations (see Fig. 7). For example, during the period between 1991 and 1998, fewer swarms are found for the whole SJFZ. Because the magnitude of completeness for the entire catalogue is about 2.0, we tested this result by identifying swarms from $M \geq 2$ events and the results still show a lower rate of swarms in this period, which is particularly pronounced for the central segment where a number of swarms occurred in the 1980 s, followed by quiescence in the 1990 s, and then a few swarms in the 2000 s. However, statistical tests show that this apparent rate change could be due to random chance, that is, we cannot definitively exclude

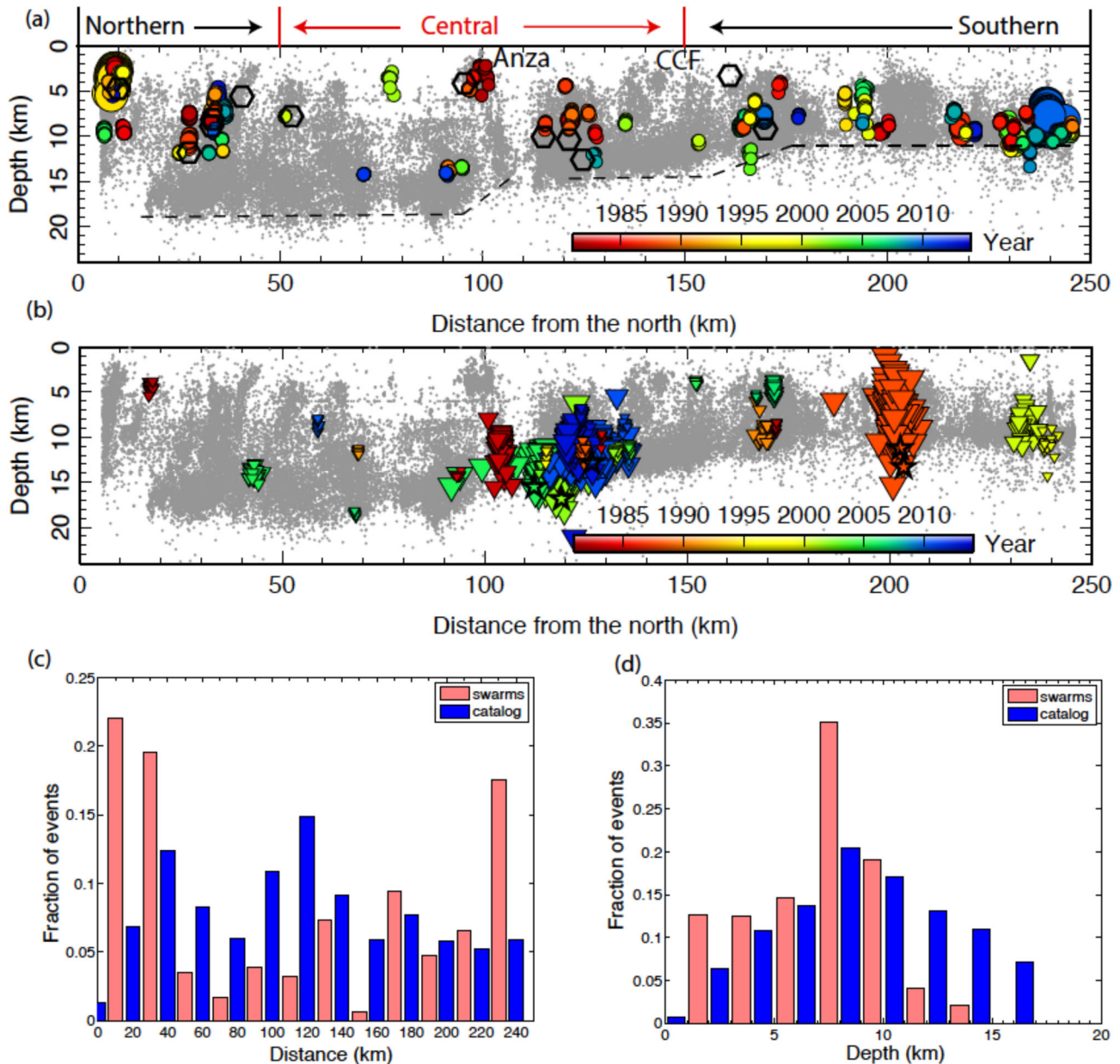


Figure 6. (a) Cross-section of swarm-like events versus depth. Symbols are the same as in Fig. 4. Dashed line marks the maximum seismogenic depth. Note that the central segment has fewer swarm events than the two ends. (b) Cross-section of M–A events versus depth. Triangles denote the M–A events. Histograms of catalogue events and swarm events along strike (c) and in depth (d) show differences in spatial distributions.

the possibility that swarm occurrence times are Poissonian. Fig. 7 also shows that the number of catalogue events jumps notably with the five moderate main shocks and increases slightly following four large regional earthquakes. In contrast, swarm occurrence does not appear to be affected by the four $M \geq 6.5$ regional earthquakes. For example, abundant swarms activate at the southern end of the SJFZ after 2009, with locations close to the US–Mexico boundary. However, their relationship to the 2010 $M 7.2$ Baja California earthquake is unclear because the swarms begin several months before the earthquake.

3.3 Swarm geometry and migration

To explore the possible causes of the SJFZ swarms and their relation to local geology, we estimate the geometry of swarm sequences

based on their 3-D event locations. Following the method described in Vidale & Shearer (2006), we compute the three eigenvalues ($\lambda_1, \lambda_2, \lambda_3$) and corresponding eigenvectors from autocorrelation of hypocentral coordinates. Swarms fall on a plane if $\lambda_1, \lambda_2 \gg \lambda_3$, thus planarity is defined by $1 - \lambda_3/\lambda_2$. A majority of the 89 swarms have planarity more than 0.8 and the highest number of swarms falls in the bin with planarity near 1, as shown in Fig. 8. This strong planarity suggests that these swarms are related to pre-existing faults. The strike and dip of the best-fitting seismicity plane can be calculated from the eigenvectors. Fig. 8 shows that most swarms dip steeply with angles greater than 60° . This suggests that faulting type may be an important factor for the occurrence of swarms because normal faults and strike-slip faults tend to have steeper dips. A more detailed analysis of the relation between swarms and faulting type is described in Section 4.

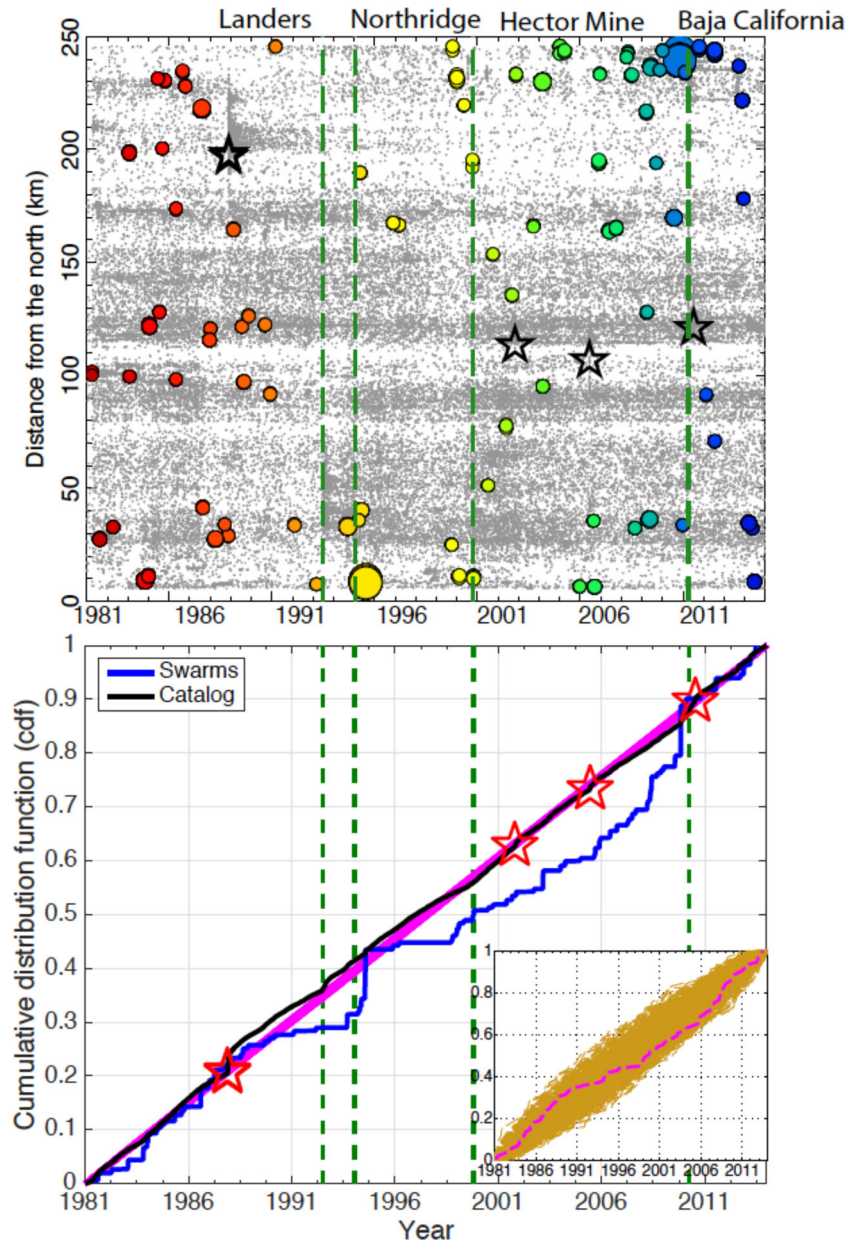


Figure 7. Spatiotemporal distribution of swarm events compared to the catalogue events in the SJFZ. The top panel maps the time and location of events and shares the same symbols as in Fig. 5. Besides the five $M \geq 5$ earthquakes in the study area (two Superstition Hills events occurred on the same day in 1987), four $M \geq 6.5$ regional earthquakes are marked by the dashed lines. The panel below shows the cumulative distributions versus time. Pink lines represent 1000 random catalogues with the same number of catalogue events. Inset shows the distributions of swarm occurrence (pink line) and 1000 random catalogues (yellow lines) with the same number of swarm occurrences (89).

Since the separation distance we calculated earlier indicates that many swarms migrate with time, we further quantify the migration velocity from the event times and locations. Following the approach in Chen & Shearer (2011) we model two types of behaviour, that is, linear and diffusion migration. To obtain reliable results we model only the 44 swarms containing at least 20 events. We fit event time and locations in each sequence with these two types of models using an inversion algorithm (Chen & Shearer 2011), in which a weight function is applied to penalize early arrivals more than late arrivals. In this way, the migration front of seismicity initiation is modelled with events falling in the upper triangle window in the time–distance plot (Fig. 9). For linear migration, we find the best-fitting parameters of migration speed (v) and migration direction

represented by a strike and dip. The diffusion model is based on the equation $r = \sqrt{4\pi Dt}$, where D is the hydraulic diffusivity, also called the diffusion coefficient, and r is the 3-D distance between one event and the initial migrating event (Shapiro *et al.* 1997; Chen & Shearer 2011). We assume the migration starts from the first event in each sequence, and then grid-search the optimal diffusion coefficients for all sequences.

To differentiate which migration model is more favourable for each sequence, we compare the rms (root mean square) of residuals in the two models. Since both models fit the migration front, which means to fit the furthest migrating distance at each discretized time, the rms could be large in both models due to scattered data. In this case, we visually check which type can better model the migrating

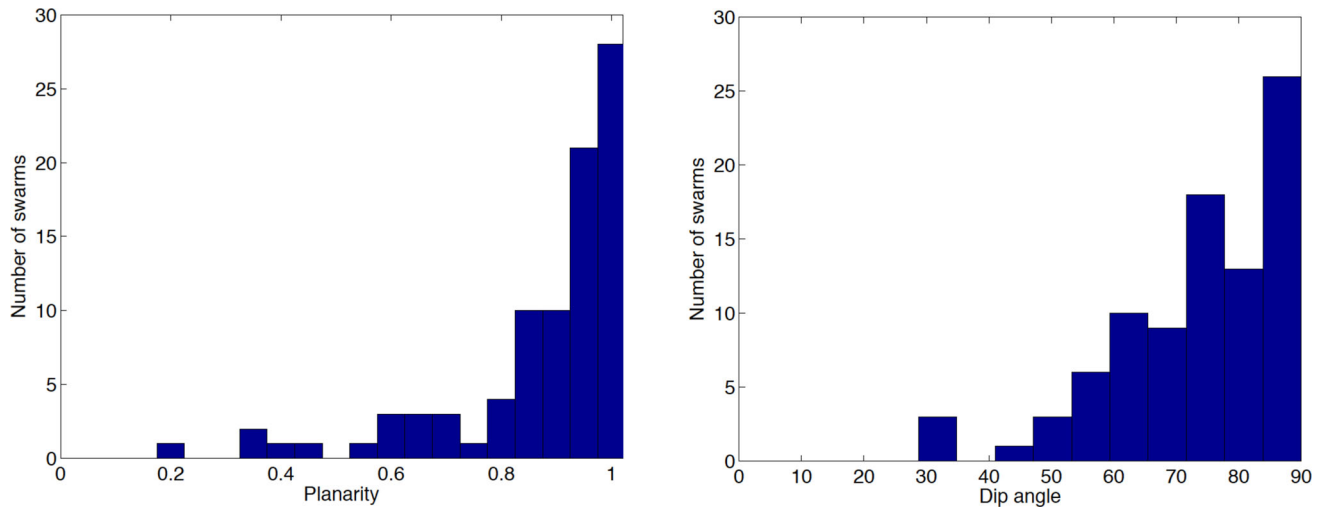


Figure 8. Histograms of planarity and dip angles for the 89 swarm groups. Most of the swarm groups have large planarity near 1 and steep dips of more than 60°. See the text for more details.

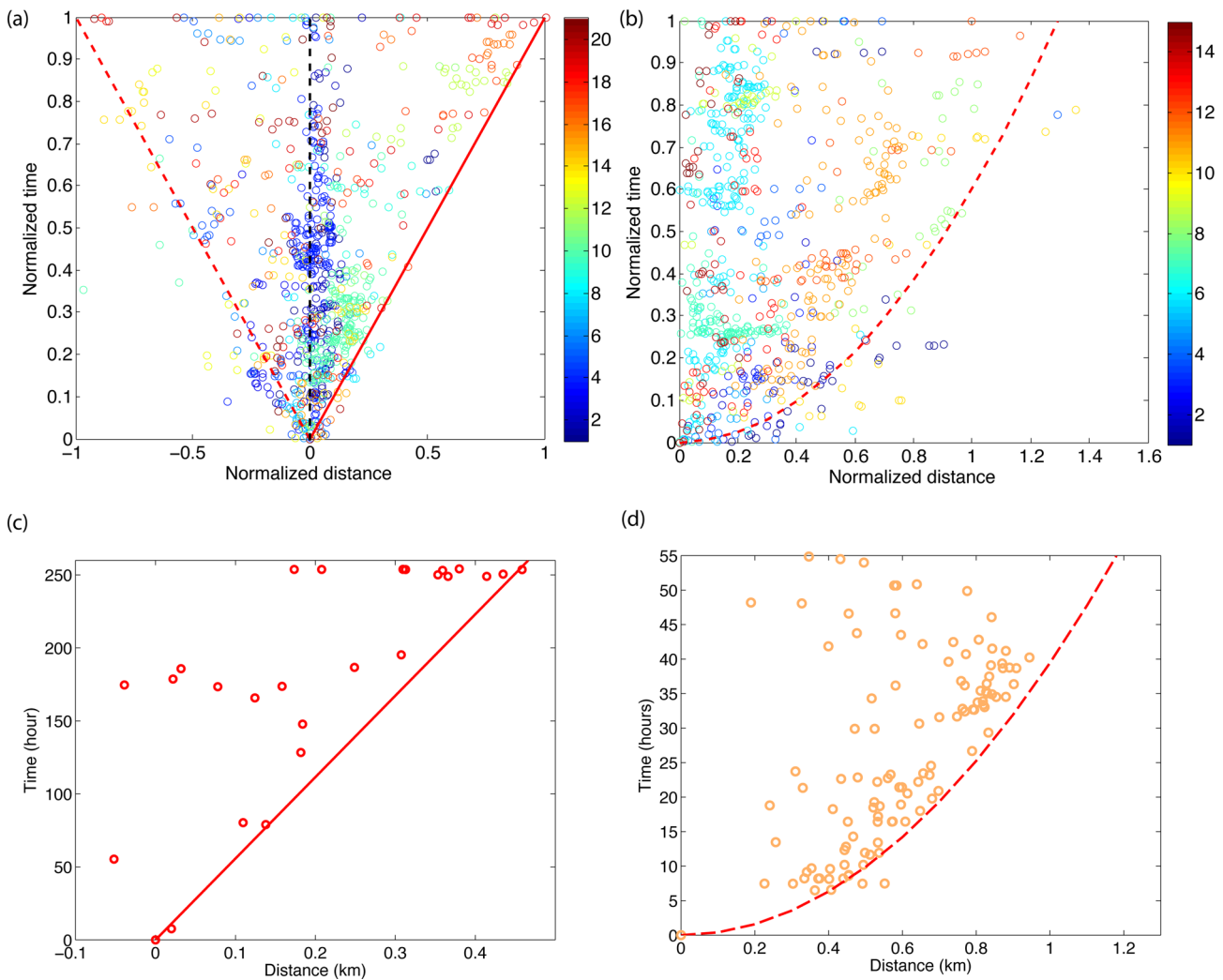


Figure 9. Time–distance plots obtained by fitting with the two different migration models. (a) Stacked version of 21 swarms fit with linear migration. Each sequence is represented by coloured dots. The distance is first normalized by the linear migration velocity and then normalized by the maximum time. The red line represents the expected migration front and the dashed line shows the opposite migration direction. (b) Stacked version of 15 swarm-like groups fitted to diffusion migration. The distance is normalized by $\sqrt{4\pi D}$ and then normalized by the square root of the maximum time. The dashed line shows the migration front with the median diffusivity of 0.6 m s^{-2} for these swarms. (c) An individual sequence with the corresponding colour in panel (a) in the real time–distance window. (d) An individual sequence with the corresponding colour in panel (b) in the real time–distance window and the best fit of diffusivity is 0.8 m s^{-2} .

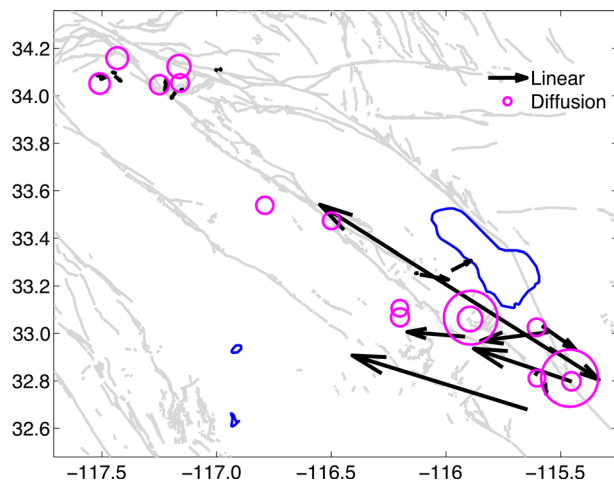


Figure 10. Map view of the best-fitting migration velocities for 36 swarms. The velocities are mapped at the locations of the first event in each group. Black arrows represent the speed and direction of the optimal linear migration. Pink circles scale with the estimated hydraulic diffusivity.

line or curve in the time–distance plot. The sequence is not categorized if neither type can fit well with a small rms. This results in 21 sequences fit with linear migration, and 15 sequences with diffusion migration. Fig. 9 shows the stacked sequences with distance normalized by its migrating vector (v or D). In Fig. 9(a), some events migrate in opposite directions, which indicates that they migrate bilaterally rather than laterally and we assume that the two directions are symmetrical. We also show two individual examples for each migrating type in which migration fronts are fit reasonably well although the data show considerable scatter (Figs 9c and d).

We map the estimated migration velocity vectors for the 36 swarms to explore whether swarm migration varies spatially and with possible tectonic implications. Fig. 10 shows the linear migration speed and strike, and the diffusion coefficient at the location of the initial event in each sequence. The linear migration velocities range from 0.0014 to 0.18 km hr⁻¹ and the diffusion coefficient varies from 0.01 to 8.81 m² s⁻¹. The detailed migration velocities for each swarm can be found in Supporting Information Tables S1 and S2. In general, the southern segment shows faster migration velocities than the northern segment for both types of migration. The region near the Salton Sea hosts the largest migration velocities (v and D), consistent with previous studies in the Salton Trough and Salton Sea geothermal field (e.g. Lohman & McGuire 2007; Chen & Shearer 2011). These migrating swarms are likely driven by fluid and/or aseismic slip.

4 DISCUSSION

Our approach assigns 4.6 per cent of catalogue events within the SJFZ to be within swarms or swarm-like clusters. Because we exclude clusters with fewer than 10 events, the true fraction of swarm events may be even higher. However, before accepting that all of our identified swarms are real, it is important to test whether many of them could be due to random chance, that is, the accidental clustering of seismicity in catalogues characterized by a Poissonian background rate and realistic aftershock triggering behaviour. To perform this check, we generate synthetic ETAS earthquake catalogues, in which we model earthquake occurrence as a point process and use the modified Omori's law and G–R relation to produce earthquake sequences (Ogata 1988, 1998). We adopt the

values suggested from previous studies in southern California for the input parameters including the Omori's decay parameter c and p ($c = 0.001$ d, $p = 1$), b value ($b = 1$) and an α value related to the main-shock magnitude dependence of triggering ($\alpha = 1$; Felzer *et al.* 2004; Helmstetter *et al.* 2005; Shearer 2012). Because the aftershock productivity K can span a wide range, we conduct a series of simulations by experimenting with K ranging from 0.01 to 0.08 with the upper limit set just low enough to avoid a runaway explosion of seismicity. The simulated catalogues contain a comparable number of the real catalogue events ($\sim 77\,000$). We then apply our swarm detection method to the simulated catalogues and find the number of clusters ranges from 41 to 239 with increasing K . However, among these clusters few swarms are found, and even the catalogue with the highest value of $K = 0.08$ produces only 17 swarms, which is far less than we find in the SJFZ catalogue. This shows that the ETAS modelling approach cannot easily explain our observed number of swarm-type sequences, which was also noted in Zaliapin & Ben-Zion (2013b). In addition we note that such a high value of K is unrealistic because it produces aftershock statistics that violate Bath's law (Shearer 2012).

We observe that swarms in the SJFZ are unevenly distributed, that is, the northern and southern ends have more frequent and larger swarms than the central segment. As suggested by the dip angles, most of these swarms likely occur in normal and strike-slip faults. To further explore the relationship between swarm occurrence and stress regime, in Fig. 11 we map the stress in a 5 km grid across our study area, as inverted from focal mechanisms (Yang & Hauksson 2013). The cells containing swarms are shown in colour. Although the overall tectonic regime is dominated by strike-slip faulting, normal faulting is also found in the areas near the northern and southern ends of the SJFZ where more swarms occur. Indeed, we find that a higher fraction of the normal faulting earthquakes (8.6 per cent) occur as swarms than for the other faulting types (1.8–3.9 per cent). Among the three types of faulting, the minimum shear stress to rupture for normal faulting is the smallest (Sibson 1974), thus normal faults may be more sensitive to small stress changes caused by creep events. Alternatively fluid migration episodes may be more common in extensional environments. However, some of these larger swarms also occur within strike-slip zones, so the presence of normal faulting is not the only controlling factor.

Occurrence of swarms is also suggested to correlate with high heat flow (Enescu *et al.* 2009; Zaliapin & Ben-Zion 2013b) and the high surface heat flow (>100 mW m⁻²) at the southern end of the SJFZ may help explain the abundant swarms there. In our cross-section (Fig. 6a), the region at distances between 20 and 100 km and depths below 12 km is devoid of swarms and has the deepest focal depths, where only two spots are available with heat flow data from the U.S. Geological Survey. Although there is not enough sampled heat flow data, the deep seismogenic depth and relatively low heat flow (around 50 mW m⁻²) at the two locations suggest this crustal block could have lower heat flow. Still, due to its poor spatial sampling, the surface heat-flow data cannot account for the local variations in swarm density. For example, swarms are absent near the CCF, while the nearest area with available heat flow data is 20 km south and shows large values (>100 mW m⁻²). Since fluid flow is one form of heat transfer, high heat flow is generally accompanied by hydrothermal circulation, which could be another key factor to determine the occurrence of swarms.

The leading candidates for the physical driving mechanisms of swarms are fluid flow and slow slip. In some cases, slow slip events are also associated with fluid flow (Schwartz & Rokosky 2007). Slow slip occurs episodically with a long duration of days to months

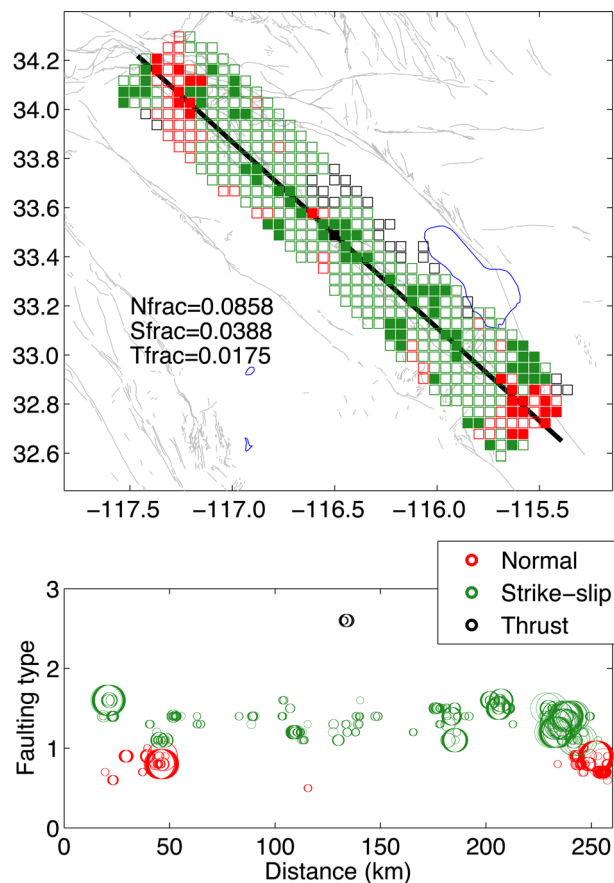


Figure 11. Map view of stress regimes and areas with swarm events. Coloured squares denote the stress regimes for each 5 km grid and are filled when swarm events occur within the grid. The faulting type is defined by the Anderson fault parameter (Simpson 1997), which ranges 0.0–1.0 for normal faulting, 1.0–2.0 for strike-slip faulting and 2.0–3.0 for reverse faulting. Note that more normal faulting events (red) are found at the two ends of the SJFZ. The fraction of events occurring as swarms for the three faulting types is also given. The panel below shows the faulting types for the areas with swarm events along the strike. Circles scale with the total number of events in each group.

and a wide source area (Obara *et al.* 2004). Even the localized events are distributed over a spatial extent of several kilometres such as in the Salton Trough (Lohman & McGuire 2007). In contrast, fluid flow is distributed more locally and usually migrates slower than slow-slip events. According to our results, most of swarms in the SJFZ are found at 5 to 8 km depth, but several swarms northwest of Anza are found at ~ 13 km depth, which is consistent with the depth where tremors are observed (Wang *et al.* 2013). However, it is still difficult to directly associate our results with slow slip. In our analysis, the median estimated linear migration velocity is 0.0072 km hr^{-1} , slower than typical creep or slow slip events (Linde *et al.* 1996; Lohman & McGuire 2007). On the other hand, most of the radii of our identified swarms are less than 2 km, which is much smaller than the spatial extent of known slow slip events and thus could be more related to the fluid flow. Our estimated fluid diffusion coefficients fall in the range of the expected crustal diffusivity between 0.01 and 10 m^2 s^{-1} (Scholz 2002), so it is possible that fluid migration plays a role in controlling the occurrence of swarms, especially since the largest estimated hydraulic diffusivity occurs near the Salton Sea geothermal field, which hosts a large number of swarms. Of course,

with a variety of spatial and temporal scales, our identified swarms may be driven by different mechanisms.

5 CONCLUSIONS

We have developed a new method to search for swarms by comparing the number of neighbours to the number of background events in scalable spatiotemporal windows. Applying the method in the SJFZ, we find swarms at a wide range of spatiotemporal scales. The SJFZ contains strong spatial and temporal variations of swarm events: the two ends of the SJFZ are dominated by more and larger swarms and the years between 1991 and 1998 host fewer swarms than the decades before and after. In general more swarms are found in normal faulting regions. While most of the estimated swarm linear migration speeds are lower than those of typical creep or slow slip events, large values of estimated hydraulic diffusivity are found near the southern end of the SJFZ (Salton Sea), which indicates that these swarms could be more correlated with the local fluid flow.

ACKNOWLEDGEMENTS

We thank the editor, two anonymous reviewers, and John Vidale for their constructive comments. The paper benefited from discussion with Xiaowei Chen and Egill Hauksson. Funding for this work was provided by the USGS-NEHRP program and the Southern California Earthquake Center.

REFERENCES

- Agnew, D. & Wyatt, F., 2005. Possible triggered aseismic slip on the San Jacinto Fault, in *SCEC Annual Meeting*, Palm Springs, CA.
- Ben-Zion, Y. & Lyakhovskiy, V., 2006. Analysis of aftershocks in a lithospheric model with seismogenic zone governed by damage rheology, *Geophys. J. Int.*, **165**, 197–210.
- Chen, X. & Shearer, P.M., 2011. Comprehensive analysis of earthquake source spectra and swarms in the Salton Trough, California, *J. geophys. Res.*, **116**, B09309, doi:10.1029/2011JB008263.
- Doser, D.I. & Kanamori, H., 1986. Depth of seismicity in the Imperial Valley region (1977–1983) and its relationship to heat flow, crustal structure and the October 15, 1979, earthquake, *J. geophys. Res.*, **91**, 675–688.
- Enescu, B., Hainzl, S. & Ben-Zion, Y., 2009. Correlations of seismicity patterns in Southern California with surface heat flow data, *Bull. seism. Soc. Am.*, **99**, 3114–3123.
- Felzer, K.R. & Kilb, D., 2009. A case study of two $M \sim 5$ mainshocks in Anza, California: Is the footprint of an aftershock sequence larger than we think?, *Bull. seism. Soc. Am.*, **99**, 2721–2735.
- Felzer, K.R., Abercrombie, R.E. & Ekström, G., 2004. A common origin for aftershocks, foreshocks, and multiplets, *Bull. seism. Soc. Am.*, **94**, 88–98.
- Gomberg, J., Rubinstein, J.L., Peng, Z., Creager, K.C., Vidale, J.E. & Bodin, P., 2008. Widespread triggering of nonvolcanic tremor in California, *Science*, **319**, 173–173.
- Hainzl, S., 2004. Seismicity patterns of earthquake swarms due to fluid intrusion and stress triggering, *Geophys. J. Int.*, **159**, 1090–1096.
- Hainzl, S. & Ogata, Y., 2005. Detecting fluid signals in seismicity data through statistical earthquake modeling, *J. geophys. Res.*, **110**, doi:10.1029/2004JB003247.
- Hauksson, E., Yang, W. & Shearer, P.M., 2012. Waveform relocated earthquake catalog for southern California (1981 to June 2011), *Bull. seism. Soc. Am.*, **102**, 2239–2244.
- Helmstetter, A. & Sornette, D., 2002. Diffusion of epicenters of earthquake aftershocks, Omori's law, and generalized continuous-time random walk models, *Phys. Rev. E*, **66**, 061104, doi:10.1103/PhysRevE.66.061104.

- Helmstetter, A., Kagan, Y.Y. & Jackson, D.D., 2005. Importance of small earthquakes for stress transfers and earthquake triggering, *J. geophys. Res.*, **110**, doi:10.1029/2004JB003286.
- Hill, D.P., 1977. A model for earthquake swarms, *J. geophys. Res.*, **82**, 1347–1352.
- Inbal, A., Ampuero, J.P. & Avouac, J., 2013. Afterslip and aftershocks triggered by moderate events on the San Jacinto Fault, *AGU Annual, Fall Meeting*, abstract #T51D-2499.
- Linde, A.T., Gladwint, M.T., Johnston, M.J.S., Gwyther, R.L. & Bilham, R.G., 1996. A slow earthquake sequence on the San Andreas fault, *Nature*, **383**, 65–68.
- Llenos, A.L., McGuire, J.J. & Ogata, Y., 2009. Modeling seismic swarms triggered by aseismic transients, *Earth planet. Sci. Lett.*, **281**, 59–69.
- Lohman, R.B. & McGuire, J.J., 2007. Earthquake swarms driven by aseismic creep in the Salton Trough, California, *J. geophys. Res.*, **112**, doi:10.1029/2006JB004596.
- Obara, K., Hirose, H., Yamamizu, F. & Kasahara, K., 2004. Episodic slow slip events accompanied by non-volcanic tremors in southwest Japan subduction zone, *Geophys. Res. Lett.*, **31**, doi:10.1029/2004GL020848.
- Ogata, Y., 1988. Statistical models for earthquake occurrences and residual analysis for point processes, *J. Am. Stat. Assoc.*, **83**, 9–27.
- Ogata, Y., 1998. Space-time point-process models for earthquake occurrences, *Ann. Inst. Stat. Math.*, **50**, 379–402.
- Roland, E. & McGuire, J.J., 2009. Earthquake swarms on transform faults, *Geophys. J. Int.*, **178**, 1677–1690.
- Simpson, R., 1997. Quantifying Anderson's fault types, *J. geophys. Res.*, **102**, 17 909–17 919.
- Scholz, C.H., 2002. *The Mechanics of Earthquakes and Faulting*, Cambridge Univ. Press.
- Schwartz, S.Y. & Rokosky, J.M., 2007. Slow slip events and seismic tremor at circum-Pacific subduction zones, *Rev. Geophys.*, **45**, doi:10.1029/2006RG000208.
- Shapiro, S.A., Huenges, E. & Borm, G., 1997. Estimating the crust permeability from fluid-injection-induced seismic emission at the KTB site, *Geophys. J. Int.*, **131**, F15–F18.
- Shearer, P.M., 2012. Space-time clustering of seismicity in California and the distance dependence of earthquake triggering, *J. geophys. Res.*, **117**, doi:10.1029/2012JB009471.
- Sibson, R.H., 1974. Frictional constraints on thrust, wrench and normal faults, *Nature*, **249**, 542–544.
- Vanderkulk, W., Rosen & Lorenz, S., 1965. Large aperture seismic array signal processing study, *IBM Final Report*, ARPA Contract Number SD-296.
- Vidale, J.E. & Shearer, P.M., 2006. 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.*, **111**, B05312, doi:10.1029/2005JB004034.
- Wang, T.-H., Cochran, E.S., Agnew, D. & Oglesby, D.D., 2013. Infrequent triggering of tremor along the San Jacinto fault near Anza, California, *Bull. seism. Soc. Am.*, **103**, 2482–2497.
- Wdowinski, S., 2009. Deep creep as a cause for the excess seismicity along the San Jacinto fault, *Nat. Geosci.*, **2**, 882–885.
- Yang, W. & Ben-Zion, Y., 2009. Observational analysis of correlations between aftershock productivities and regional conditions in the context of a damage rheology model, *Geophys. J. Int.*, **177**, 481–490.
- Yang, W. & Hauksson, E., 2013. The tectonic crustal stress field and style of faulting along the Pacific North America Plate boundary in Southern California, *Geophys. J. Int.*, **194**, 100–117.
- Zaliapin, I. & Ben-Zion, Y., 2013a. Earthquake clusters in southern California I: identification and stability, *J. geophys. Res.*, **118**, 2847–2864.
- Zaliapin, I. & Ben-Zion, Y., 2013b. Earthquake clusters in southern California II: classification and relation to physical properties of the crust, *J. geophys. Res.*, **118**, 2865–2877.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this paper:

Table S1. Swarms fitting to linear migration model.

Table S2. Swarms fitting to fluid diffusion migration model.

(<http://gji.oxfordjournals.org/lookup/suppl/doi:10.1093/gji/ggw073/-/DC1>).

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.