Tectonophysics 804 (2021) 228756

Contents lists available at ScienceDirect

Tectonophysics

journal homepage: www.elsevier.com/locate/tecto

Three-dimensional paganica fault morphology obtained from hypocenter clustering (L'Aquila 2009 seismic sequence, Central Italy)

Brennan Brunsvik [a,b,*], Gabriele Morra [a,c], Gabriele Cambiotti [d], Lauro Chiaraluce [e], Raffaele Di Stefano [e], Pasquale De Gori [e], David A. Yuen [f,g]

[a] School of Geosciences, University of Louisiana at Lafayette, LA 70504, USA
[b] Department of Earth Science, University of California Santa Barbara, CA 93106, USA
[c] Department of Physics, University of Louisiana at Lafayette, LA 70504, USA
[d] Department of Earth Sciences, University of Milan, Via Cicognara 7, 20133 Milano, Italy
[e] Istituto Nazionale di Geofisica e Vulcanologia, Via di Vigna Murata, 605 00143 Rome, Italy
[f] Dept. of Earth Sciences and Minnesota Supercomputing Institute, University of Minnesota, Minneapolis, MN 55455, USA
[g] Dept. of Applied Physics and Applied Mathematics, Columbia University, New York, NY 10026, USA

ARTICLE INFO

Keywords:
Fault morphology
Clustering
L'Aquila earthquake sequence

ABSTRACT

In seismic modelling, fault planes are normally assumed to be flat due to the lack of data which can constrain fault morphology. However, incorporating 3D fault morphology is important for modelling several phenomena, for example calculating mainshock induced stress changes. We utilize a data-analytical method to unveil the 3D rupture morphology of faults using unsupervised clustering techniques applied to earthquake hypocenters in seismic sequences. We apply this method to the 2009 L'Aquila seismic sequence which involved a M_W 6.1 mainshock on April 6th. We use a dataset of about 50,000 relocated events, mostly microearthquakes, reaching magnitude of completeness equal to 0.7. Clustering distinguishes the earthquakes as occurring in three main clusters along with other minor fault segments. We then represent the morphology of the main Paganica fault system (responsible for the largest mainshock) using splines. This method shows promise as a step toward robustly and quickly obtaining 3D rupture morphologies where earthquake sequences have been monitored. The 3D model is presented interactively online, and the processing is presented in an interactive Jupyter Notebook (https://bit.ly/2MnCFdj).

1. Introduction

The 3D morphology of fault planes is generally unknown at depth due to the lack of information which can constrain them. We present an approach where we use machine learning to find the 3D morphology of faults based on the spatial distribution of aftershocks in seismic sequences. We apply this method to model the Paganica fault system which was the source of the April 6th, 2009 M_W 6.1 mainshock of the thoroughly monitored L'Aquila seismic sequence in the central Apennines, Italy (Scognamiglio et al., 2010).

This approach to automating the inversion of 3D fault morphology can be useful in many ways. Inaccuracies in fault morphology can interfere with earthquake source inversions, resulting in less realistic slip distributions (Ragon et al., 2018) and affecting coulomb stress calculations and aftershock forecasts (Kaven and Pollard, 2013; Mildon et al., 2016; Verdecchia et al., 2018; Mancini et al., 2019). In addition, the morphology of faults influences dynamic stress (e.g. Pelties et al., 2012; Galvez et al., 2014; Zhang et al., 2014), earthquake magnitude and location (Barka and Kadinsky-Cade, 1988; Zielke et al., 2017), tsunami generation (Moore et al., 2007), and other phenomena. Along megathrusts (e.g. the M_W 7.8 2015 Gorkha earthquake, Hubbard et al., 2016) the determination of the 3D morphology of slabs and faults has been taken into account (Qiu et al., 2016; Landry and Barbot, 2018). This can provide information on locked zones (Avouac, 2008; Avouac et al., 2015), which are crucial to understand as they can promote intense seismicity. For the major subduction zones of the Earth, it has been shown that curvature influences earthquake rupture area and magnitude (Bletery et al., 2016). Ultimately, progress toward automatic and objective fault reconstruction is an important task in seismology.

A variety of methods are used to estimate fault morphology. Active

* Corresponding author at: 40 S. Patterson Ave 204, Santa Barbara, CA, USA.
E-mail address: brennanbrunsvik@ucsb.edu (B. Brunsvik).

https://doi.org/10.1016/j.tecto.2021.228756
Received 29 April 2020; Received in revised form 25 January 2021; Accepted 27 January 2021
Available online 4 February 2021

source seismology is used to build high resolution fault models (e.g. Lohr et al., 2008; Røe et al., 2014). This is the case, among others, of the low angle Altotiberina normal fault in the northern Apennines (Italy) for which 3D morphology was obtained by merging seismic imaging, borehole analysis, geologic maps, and field surveys. Unfortunately, the costs of active source seismology and drilling make this a difficult option. Fault geometry can also be illuminated from passive source methods. For example, modelling of seismic waves trapped in low-velocity damage zones can provide information on fault geometry (Peng et al., 2003; Ben-Zion et al., 2003; Calderoni et al., 2012).

The distribution of foreshocks and aftershocks can also provide information on fault morphology. Analysts can attribute earthquakes to their source faults by visual inspection of earthquake locations and surfaces can be fit to the aftershocks to model these faults (e.g. Carena et al., 2002). The combination of hypocenter interpretation along with various other data types (e.g. fault traces, borehole data, geodetic data, geologic cross-sections) has been used to construct 3D geologic models of faults associated with the L'Aquila sequence (Lavecchia et al., 2012; Lavecchia et al., 2017; Castaldo et al., 2018) and the 2016 Amatrice/Visso/Norcia sequence (Lavecchia et al., 2016; Tung and Masterlark, 2018). While these methods are successful, they involve qualitative and subjective elements which influence the results based on the analysts' background. Moreover, it is not quickly reproducible, particularly in quasi real-time as a seismic sequence evolves.

On the argument of automation and fast, reproducible data analysis, machine learning is becoming increasingly useful in Earth science (Kong et al., 2019; Bergen et al., 2019). Some examples include determining earthquake magnitudes (Mousavi and Beroza, 2019), analyzing lab quakes (Rouet-Leduc et al., 2017), and picking phase arrivals (Zhu and Beroza, 2019). Machine learning includes supervised learning which requires pre-labeled data for training, and unsupervised learning where no pre-labeled data is required for training. Clustering is one application of unsupervised machine learning. If data points are distributed in distinct groups, clustering finds these groups and labels each datum according to which group it belongs to (e.g. Omran et al., 2007; Joshi and Kaur, 2013).

There have been several applications of clustering in Earth Science. For example, it has been applied to earthquake relocation (Trugman and Shearer, 2017), investigation of earthquake processes and precursors (Dzwinel et al., 2003; Dzwinel et al., 2005; Yuen et al., 2009), and distinguishing dominant faulting patterns in focal mechanism data-sets (Rietbrock et al., 1996; Custódio et al., 2016). Hypocenter locations have been clustered to establish seismic zones for probabilistic seismic hazard analysis (Weatherill and Burton, 2009; Ansari et al., 2015; Novianti et al., 2017; Hall et al., 2018; Scitovski, 2018).

Clustering has also been used to investigate fault geometry based on hypocenter information (Ouillon et al., 2008; Ouillon and Sornette, 2011; Wang et al., 2013). In these cases, clusters were distinguished by fitting them with a series of planes. The RANdom SAmple Consensus (RANSAC) algorithm, which fits a model to the dominant trend in data while attempting to reject outlier data (Fischler and Bolles, 1981), has also been used to find faults in hypocenter data and to model them with 3D surfaces (Kaven and Pollard, 2013; Skoumal et al., 2019). Ultimately, clustering has the capability to automatically identify groups of hypocenters common to fault zones to aid in inverting fault morphology in a quick and objective fashion.

Here, we present an approach based on the unsupervised machine learning technique clustering which can automate fault identification using hypocenter locations. We apply our clustering-based method to the L'Aquila seismic sequence in the central Apennines (Italy) which was particularly well monitored with over 50,000 high-resolution relocated aftershocks identified throughout 2009 (Valoroso et al., 2013). This large amount of data is well suited for data-driven techniques. Further, the fault morphologies associated with this sequence are relatively simple, making this a good region to test fault morphology inversion algorithms.

1.1. 2009 L'Aquila seismic sequence

The faults associated with the L'Aquila seismic sequence have been thoroughly studied (Figs. 1 and 2; e.g. Lavecchia et al., 2012; Valoroso et al., 2013). The L'Aquila sequence activated a dominantly south-west dipping normal fault system about 50 km long. Aftershocks are broadly associated with the ~N140° striking Paganica fault system, the ~N150° striking Campotosto (Gorzano) fault, and a cluster near the Cittareale village. The mainshock occurred in the Paganica fault system, which is composed of several enechelon faults extending approximately from the SE end of the San Demetrio fault system through the Mt. Stabiata Fault. The Cittareale cluster is composed dominantly of several NW-SE striking and NE dipping faults, as well as a SW dipping fault likely associated with the Montereale fault system.

Chiaraluce (2012) provided a review of the 2009 L'Aquila sequence (Fig. 1), which began in approximately January 2009. The foreshocks included a M_W 4.0 earthquake on March 30th before the April 6th M_W 6.1 NW-SE striking and SW dipping mainshock (Chiaraluce et al., 2011). Southeast of the mainshock, on April 7th, a M_W 5.4 aftershock occurred on a ~NE dipping fault with a non dip-slip component (rake of ~58°: see Fig. 1 for focal mechanisms of large events) (Scognamiglio et al., 2010; Valoroso et al., 2013). Campotosto fault seismicity included a M_W 5.0 event on April 6th and M_W 5.0 and 5.2 events on April 9th, which were dominantly dip slip, SE to SSE striking, and SW to WSW dipping (Scognamiglio et al., 2010). These events occurred on the shallower, steeply dipping portion of the listric Campotosto fault which appears kinked at its base where seismicity abruptly changes to low dip (Chiaraluce, 2012). A M_W 4.4 event occurred later on June 22nd, 2009 beneath this kink, on the gently dipping portion of the Campotosto fault with nodal planes showing dip of 88° or 14° (Valoroso et al., 2013). On June 25th, a M_W 3.5 (Scognamiglio et al., 2010) earthquake occurred within the Cittareale cluster. This seismic cluster is approximately the northward extent of the L'Aquila sequence.

Varying terminology has been used to describe the fault system that we focus on. Many refer to the fault as the "L'Aquila fault system" (e.g. Chiaraluce et al., 2011), and many others use variations on the "Mt. Stabiata – Paganica – San Demetrio fault system" (e.g. Boncio et al., 2010; Lavecchia et al., 2012). The mainshock occurred on the L'Aquila normal fault system, most precisely on the portion corresponding in the field to the Paganica fault (Chiaraluce, 2012). Essentially the Mt. Stabiata, Mt. Castellano, Colle Enzano, Paganica, and San Demetrio faults can be viewed as part of the same semi-continuous fault system (Fig. 1; Civico et al., 2017). However, it must be recognized that the relationship between outcropping faults and hypocenters can be complex (e.g. Tondi, 2000), and our goal is to more generally model the morphology of the fault zone where it was either coseismically activated by the mainshock or is continuous with the coseismically activated portion of the fault zone. For this reason, and to emphasize the relation of the modelled fault to the mainshock, we call the system the "Paganica fault system" to emphasize the importance of the mainshock.

2. Methods

The aim of the proposed method is to use aftershock locations to invert for fault morphology in an objective, automated way. For this method, it is first necessary to group aftershocks according to the faults they occurred on. Then the 3D morphology of a specific group of aftershocks can be inverted using splines (Virtanen et al., 2020). The reader is referred to the Supplementary Jupyter Notebook which presents the methods and results interactively (https://bit.ly/2MnCFdj).

The first step is to apply clustering to hypocenters. Then, hypocenter clusters can be associated with source faults (similar to Ouillon et al., 2008). Note that because this method relies on the occurrence of earthquakes, faults which are not activated or are only partially activated will not be fully accounted for. To increase the quality of the results, we run two iterations of clustering: one with spectral clustering

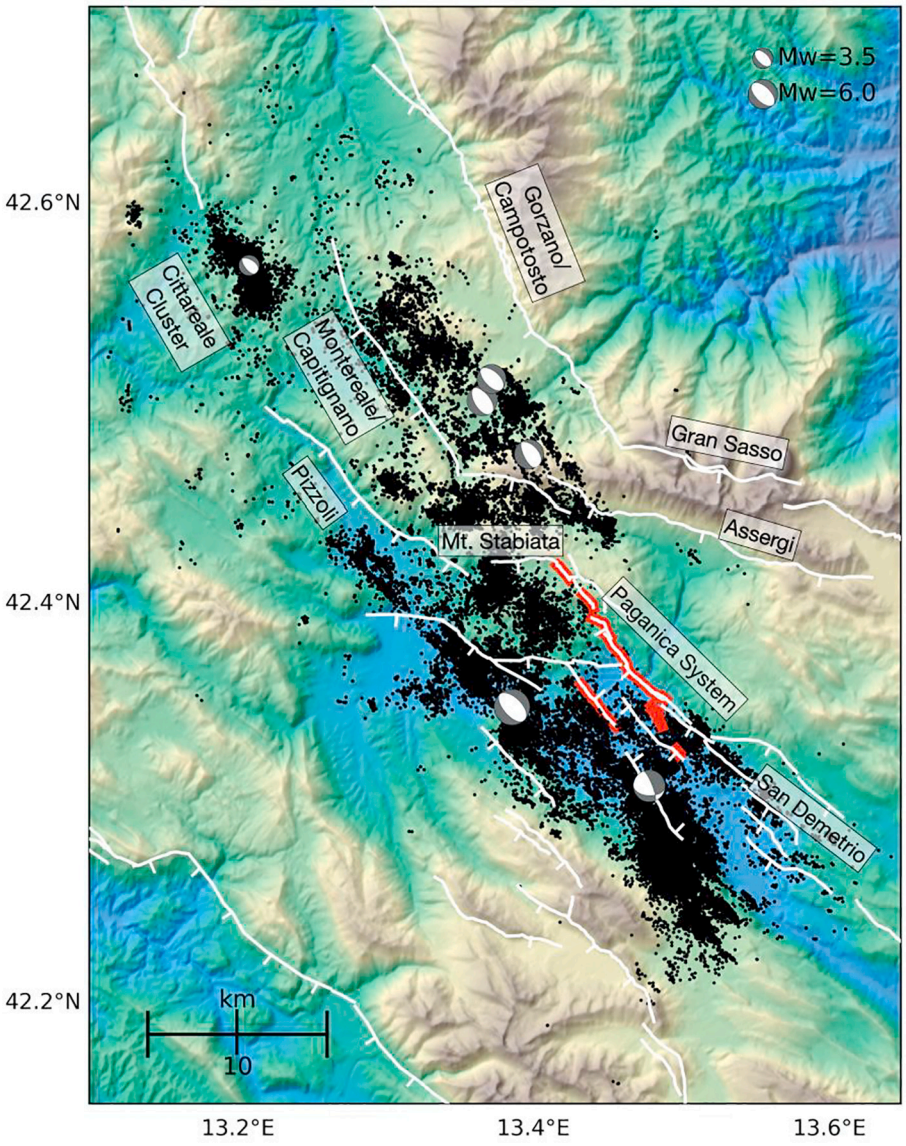

Fig. 1. Map view of seismicity of the L'Aquila seismic sequence. Hypocenters are from Valoroso et al. (2013). Fault traces and surface ruptures are modified from Boncio et al. (2010) and Valoroso et al. (2013). Fault and fault system names discussed in the text are labeled (names do not refer to city locations). Normal fault dip directions are indicated by barbs. Focal mechanisms are plotted for events with $M_W \geq 5$ in addition to the M_W 3.5 event in the Cittareale area (Scognamiglio et al., 2010).

(Von Luxburg, 2007), and another with Density-Based Spatial Clustering of Applications with Noise (DBSCAN) (Ester et al., 1996). Finally, after clusters of hypocenters are categorized according to their source faults, we focus on specifically the Paganica fault system for which we interpolate a continuous 3D surface.

We use the earthquake catalog of Valoroso et al. (2013), which is publicly available (Valoroso et al., 2020). The catalog was generated by applying STA/LTA event detection, applying an automated phase arrival picking procedure, removing duplicate events, cross-correlating records of event pairs, and finally applying the double-difference relocation algorithm *hypoDD* (Waldhauser and Ellsworth, 2000). Based on 95% confidence intervals obtained from a bootstrap analysis, the authors estimated mean location errors of 0.087, 0.178, and 0.039 km in the vertical and two horizontal directions, respectively.

2.1. Application of clustering to identify faults

The machine learning clustering approach attempts to find relationships in data based on their similarity. In our case, where the locations of different earthquakes are sufficiently close, clustering can recognize that these earthquakes occurred on the same fault. If the earthquakes are sufficiently distant from each other, clustering can recognize that the earthquakes occurred on different faults. Several challenges arise using this approach. Where faults intersect, hypocenters of different faults can be close to each other and clustering algorithms can be blind to this. Further, aftershock location inaccuracy is inevitably present due to imperfect knowledge of crustal 3D velocities and elastic properties. Thus, hypocenters tend to erroneously appear between multiple faults, and to which fault these hypocenters belong to is unclear. To mitigate this, clustering is done in two iterations; first with a spectral method, and second with DBSCAN.

During the first clustering iteration, we used the spectral clustering algorithm of Scikit-Learn (Pedregosa et al., 2011). Von Luxburg (2007) provided an overview of spectral clustering. Spectral clustering is appropriate because it works well for non-convex clusters (a cluster is non-convex when a line drawn between two points can escape the cluster, such as hypocenters belonging to a curved fault). We first compute the similarity matrix between points which defines clustering weights according to the distance between hypocenters: $\mathbf{s_{i,j}} = \exp(-\|\mathbf{x_i} - \mathbf{x_j}\|^2/2\sigma^2)$, where $\mathbf{x_i}$ and $\mathbf{x_j}$ are the positions of the ith and jth hypocenters, and σ is the width of the gaussian kernel. Then, the normalized Laplacian matrix is constructed: $\mathbf{L} = \mathbf{D} - \mathbf{W}$. $\boldsymbol{W}$ is the weighted adjacency matrix of each point. We use $\boldsymbol{W} = \mathbf{s}$. The degree matrix $\boldsymbol{D}$ is a diagonal matrix containing the degrees of the weight vector $d_i = \Sigma_{j=1}^{n} w_i$,

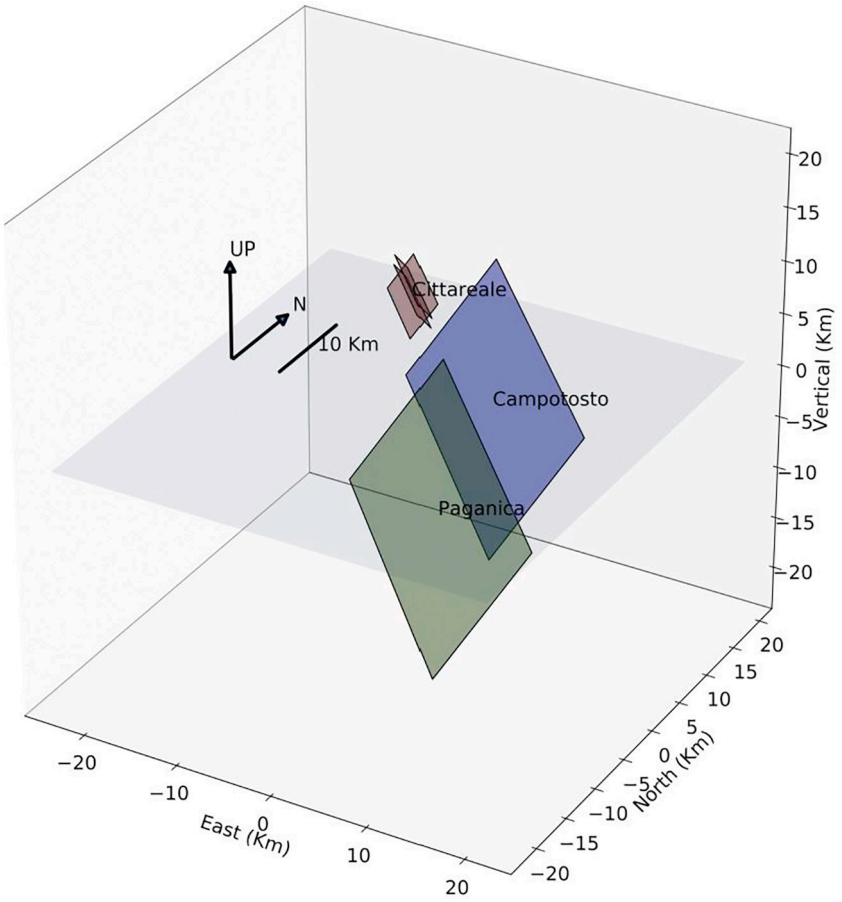

Fig. 2. Simplified schematic of the dominant hypocenter clusters involved in the L'Aquila sequence. Seismic activity occurred dominantly in a cluster about the Paganica fault system (including from the Mt. San Franco fault through the San Demetrio faults), the Campotosto cluster, and the Cittareale cluster (consisting of conjugate faults in part corresponding to the Montereale fault system).

$_j$. Then, the first k eigenvectors $\boldsymbol{u_1}$, …, $\boldsymbol{u_k}$ of $\boldsymbol{L}$ are computed to construct $\boldsymbol{U} \in R^{n\times k}$. Finally, in k-dimensional space, the n points within $\boldsymbol{U}$ are assigned labels using k-means.

For the spectral clustering application, we used only the 542 earthquakes with $M_L \geq 2.3$ (Table 1). This helps to remove small aftershocks that blur the distinction between different faults. It also makes the memory demanding spectral clustering algorithm more manageable. Spectral clustering, like many other clustering algorithms, requires prior knowledge of the number of clusters. We choose $k = 3$, corresponding to the Paganica, Campotosto, and Cittareale clusters (Table 1). Note that this choice of $k = 3$ only applies to the spectral clustering step: the number of clusters found during the later application of DBSCAN will be independent. Finally, we choose $\sigma = 2000$ m (Table 1). However, for 1250 m $< \sigma <$ 4300 m, the results vary only negligibly. This spectral clustering step separated the hypocenters into three dominant clusters (Fig. S1).

Before applying the second clustering iteration, we modify the hypocenter data to enhance fault like trends. Using the 396 hypocenters which were preliminarily identiffed as corresponding to the Paganica fault system using spectral clustering, we solve for a least squares distance best fft plane. The positions of all 51,339 hypocenters, which includes those outside of the Paganica cluster, were then stretched by a factor $f = 5$ in the direction normal to the best fit plane (Table 1). f is chosen to enhance planar trends in the hypocenters, particularly those matching the Paganica fault's strike and dip. This preliminary operation only minimally influences hypocenters close to the fault surface identified by the best fit plane, while deviations of earthquakes far from this surface are amplified. With these Paganica fault-like planar trends more pronounced and bias from deviant earthquakes reduced, clustering algorithms are much more successful at both recognizing the full extent of faults and keeping neighboring faults separate. This step also reduces the problem that some groups of hypocenters may lie along the same fault, but there are so few hypocenters between these groups that clustering algorithms normally find them as separate clusters. This spatial modification also allows for the use of less computationally expensive clustering methods.

Table 1
Parameters and properties for the two clustering methods. M_{min} is the minimum magnitude of events used for that step; we also report the number of remaining earthquakes. We report the number of clusters. This was specified as a parameter for the Spectral clustering step only, while the number of clusters reported for DBSCAN was a result of the method. Eps and MinPts are specified parameters for the DBSCAN step only, and σ is specified for Spectral clustering only. f is the stretching factor which was applied to hypocenters prior to applying DBSCAN. dx1 and dx2 describe the horizontal and vertical spacing of nodes for the spline inversion.

Method	Spectral	DBSCAN	Splines
M_{min}	2.3	None	
Number of earthquakes	542	51,339	
Number of clusters	3	9	
Eps		2 km	
σ	2 km		
MinPts		80	
f		5	
dx1			5668 m
dx2			2544 m

Once DBSCAN is applied during a later step, the results for the Paganica fault system can be closely matched without applying f given careful tuning of other clustering parameters (Fig. S2). However, without applying f, we observe that clustering will be less successful when considered over the entire seismic sequence. For example, for $f = 1$, we find no combination of DBSCAN parameters which both work well for the Paganica fault while simultaneously distinguishing the Campotosto fault from the Mt. San Franco fault. For $f > 5$, we do not observe further improvement to clustering.

With preparatory data operations completed, the final clusters can be generated using DBSCAN (Ester et al., 1996; Pedregosa et al., 2011; Schubert et al., 2017). Intuitively, DBSCAN finds dense clusters of points which are separated by regions of low point density. The two important parameters are MinPts and Eps which essentially control the number of points in clusters and the width of clusters, respectively. DBSCAN relies on the distinction of core points, border points, and noise. A point p is considered a core point if there are greater than MinPts points q with dist (p,q) less than Eps. Such points q are considered direct density-reachable

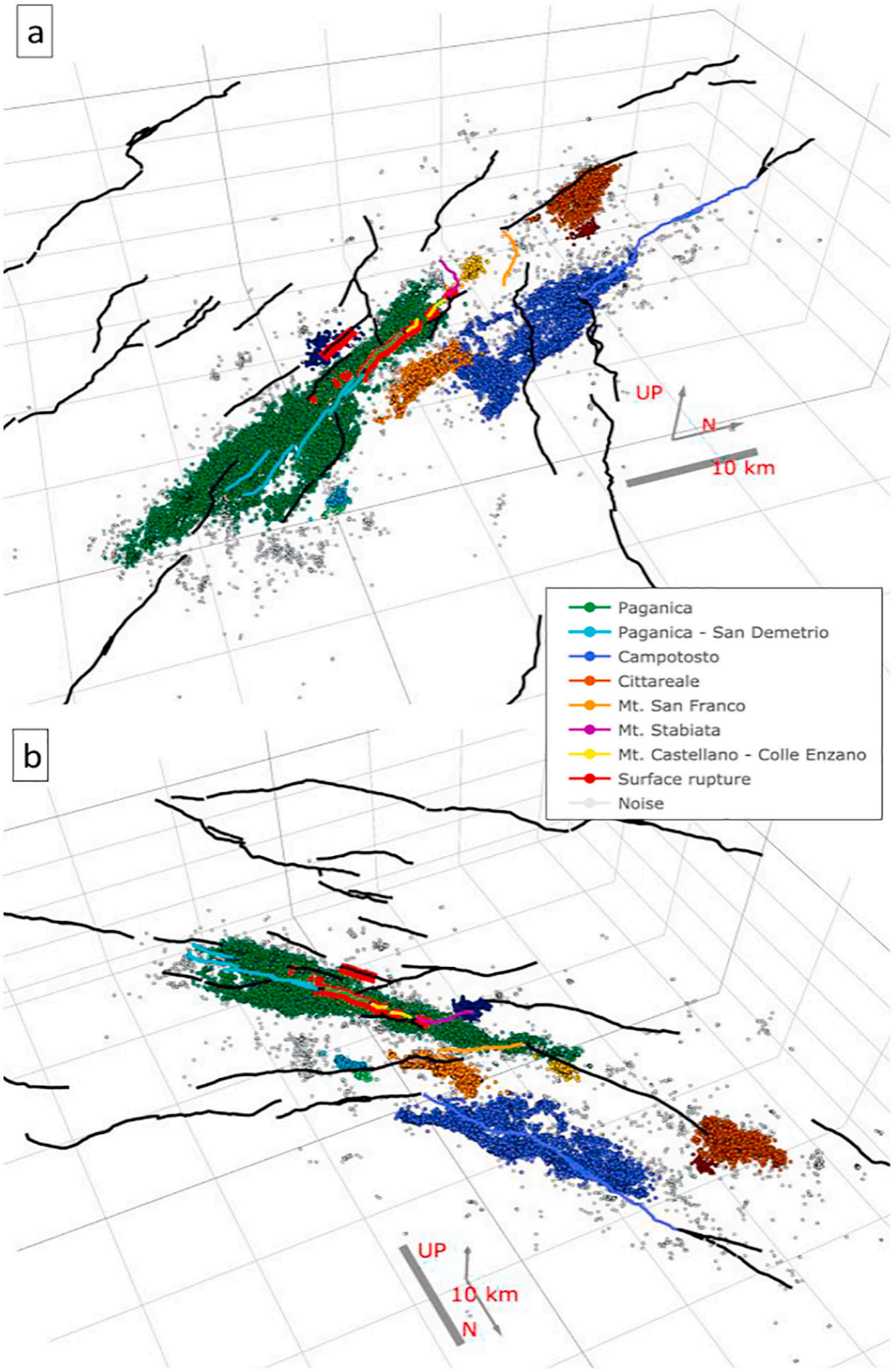

Fig. 3. Also available interactively online (see supplementary material or at https://bit.ly/3k7ykqg). Looking northwest (a) and southwest (b). Final results from clustering hypocenters from the L'Aquila seismic sequence. Spectral clustering was first used to find three primary clusters. DBSCAN was then used on all hypocenter location data. The gray dots were considered by DBSCAN to be noise. Three hypocenter clusters are colored according to potentially affiliated geological fault traces: the Paganica fault, Campotosto fault, and Mt. San Franco fault. We plot the individual Paganica fault trace as green, though our Paganica cluster represents Paganica fault system including faults between the Mt. Stabiata and San Demetrio faults. The Cittareale cluster, containing several faults, is colored brown. Other clusters were found for which we do not consider their relationship to known fault outcroppings, and these are colored randomly. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

from p and are thus established as part of the same cluster. A point q that does not qualify as a core point, but with dist(p,q) < Eps, is considered a border point in the same cluster as the core point p. A point p_n is considered density reachable from each core point p_j in the chain p_1,p_2... p_n if each core point in the chain is direct density reachable to the core point before it; that is, if p_{i+1} is direct density-reachable from p_i for $1 < i < n - 1$. Note that p_n itself can be either a core point or border point. This allows for the growth of clusters by density reachability where all density reachable points from the core point p are part of the same cluster. A point q with dist(p_j,q) > Eps to all core points p_j is considered noise. Thus, a cluster is composed of core points p and their density reachable border points q. We empirically found that Eps = 2 km and MinPts = 80 (Table 1) are good choices for the Paganica fault system. However, based on a grid search of Eps and MinPts, the DBSCAN clustering results for this fault are only negligibly sensitive to Eps and MinPts (Fig. S3).

The result of DBSCAN is that hypocenters are grouped and established as clusters (Fig. 3). Then, we identify the hypocenter clusters according to which previously geologically identified faults they most closely correspond to (Fig. 3). We discuss the relationship of geologically identified faults and hypocenter clusters in the results section.

2.2. Inversion of Paganica fault system morphology model

Hypocenters within individual clusters can next be spatially interpolated to find a best fit 3D fault surface. Although we identified several hypocenter clusters, this work focuses on modelling the morphology only of the Paganica fault system. Assuming the aftershocks in the Paganica hypocenter cluster are distributed in a damage zone and are approximately symmetric about the Paganica fault system surface (Valoroso et al., 2013), a least squares best-fit surface will represent the average rupture location and thus closely resemble the fault surface.

We interpolated the hypocenters of the Paganica cluster using splines. We choose splines because high-order polynomials tend to suffer from unrealistic oscillations between interpolated data whereas splines can converge to a solution using stable, low-order polynomials (e.g. Fornberg and Zuev, 2007). The spline algorithm constructs a piece-wise polynomial which is continuous and smooth. To apply splines, we choose a coordinate system that is defined by the best fit plane to the Paganica hypocenter cluster where x_1' faces along strike to the southeast, x_2' faces down dip along the fault plane, and x_3' faces the hanging wall, normal to the fault plane (Fig. 4). We use 4th order bivariate B-splines (Prautzsch et al., 2002; Virtanen et al., 2020) defined over the region $0 \leq x_1 \leq 15.2$ km and $0 \leq x_2 \leq 34.0$ km which contains all identified Paganica hypocenters. We distribute throughout the plane a

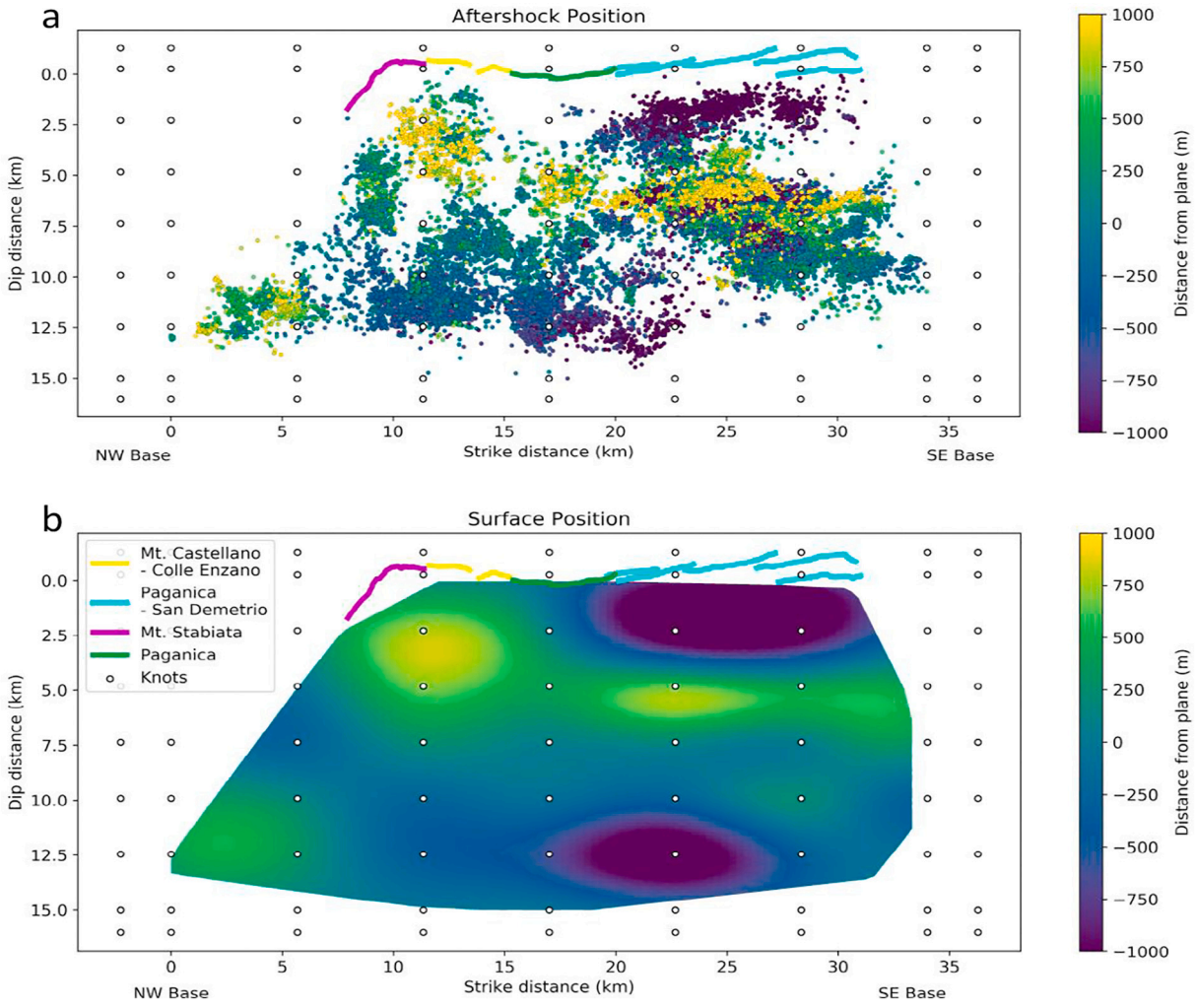

Fig. 4. (a) 2D representation of the aftershocks that were part of the Paganica fault cluster, and (b) the surface that interpolates them. Position is shown on the best-fitting plane to the hypocenters where strike distance corresponds to x_1'. The y axis corresponds to distance along dip, not depth, and corresponds to x_2'. Colour represents distance in meters from the best fitting plane on which spline knots were distributed (distance in x'$_3$). Selected fault traces are shown as their projection onto the same plane. The viewing angle is such that the viewer looks northeast and down.

grid of nodes with spacing of $dx_1 = 5668$ m, and $dx_2 = 2544$ m (Fig. 4; Table 1). In order to have 4th order splines defined through the fault model, exterior nodes must also be used; these are distributed at distances of ½ dx_1 and ½ dx_2 outside of our fault model (Fig. 4). However, the style of exterior nodes implementation is not important for our results. Finally, the spline coefficients are inverted using least squares to minimize the distance between the hypocenters and the spline surface. The resulting splines indicate the best fit fault surface model of the Paganica hypocenter cluster (Fig. 4).

3. Results

The final clusters, the results of DBSCAN, can be seen in Fig. 3. Nine clusters were recognized; among these, the Paganica, Campotosto, and Cittareale clusters are most clearly represented. Of the 51,339 events, 29,367 were clustered as part of the Paganica fault system, 9588 on the Campotosto fault, and 5384 in the Cittareale cluster. We also find a cluster between the Paganica and Campotosto faults which appears to correspond to the secondary Mt. San Franco fault and Capitignano fault (part of the Montereale fault system) according to the interpretation of Lavecchia et al. (2012).

The general correspondence between these clusters and the potentially associated geologically mapped fault traces can also be seen in Fig. 3. Although the Paganica fault is our primary focus, and clustering parameters were chosen for optimal performance on the Paganica cluster, we describe other clusters briefly.

Although the hypocenters identified as the Campotosto fault maintain a depth of at least 4.5 km, they show a close association to the Gorzano fault trace (e.g. Lavecchia et al., 2012; Castaldo et al., 2018). First, the dip of the hypocenters shows relatively good alignment with the fault trace, as noted by Chiaraluce et al. (2011). Further, the apparent strike of the hypocenters in the Campotosto cluster varies along strike, matching well with the fault trace (Fig. 3b), supporting a connection between them. Similarly, Castaldo et al. (2018) modelled a fault connecting these hypocenters and the fault trace. Clustering was also successful at identifying both the shallow and steeply dipping portions of the listric Campotosto fault (Valoroso et al., 2013) as the same feature. The correspondence of the Mt. San Franco fault and the mapped surface trace is less clear. For example, the strike of the hypocenter distribution is ~N140°, while the strike of the overall surface trace varies from ~N140° to ~N100°. The Cittareale cluster accurately contains the June 25th M_w 3.5 event. This cluster contains a group of faults, and is actually dominated by NE dipping, antithetic faults. The SW fault appears to align with the Montereale Fault system NW of the Mt. San Franco fault.

The separation between major faults and more scattered and independent aftershocks (referred to as noise and rejected by the DBSCAN algorithm) appears successful. However, sets of faults antithetic to and

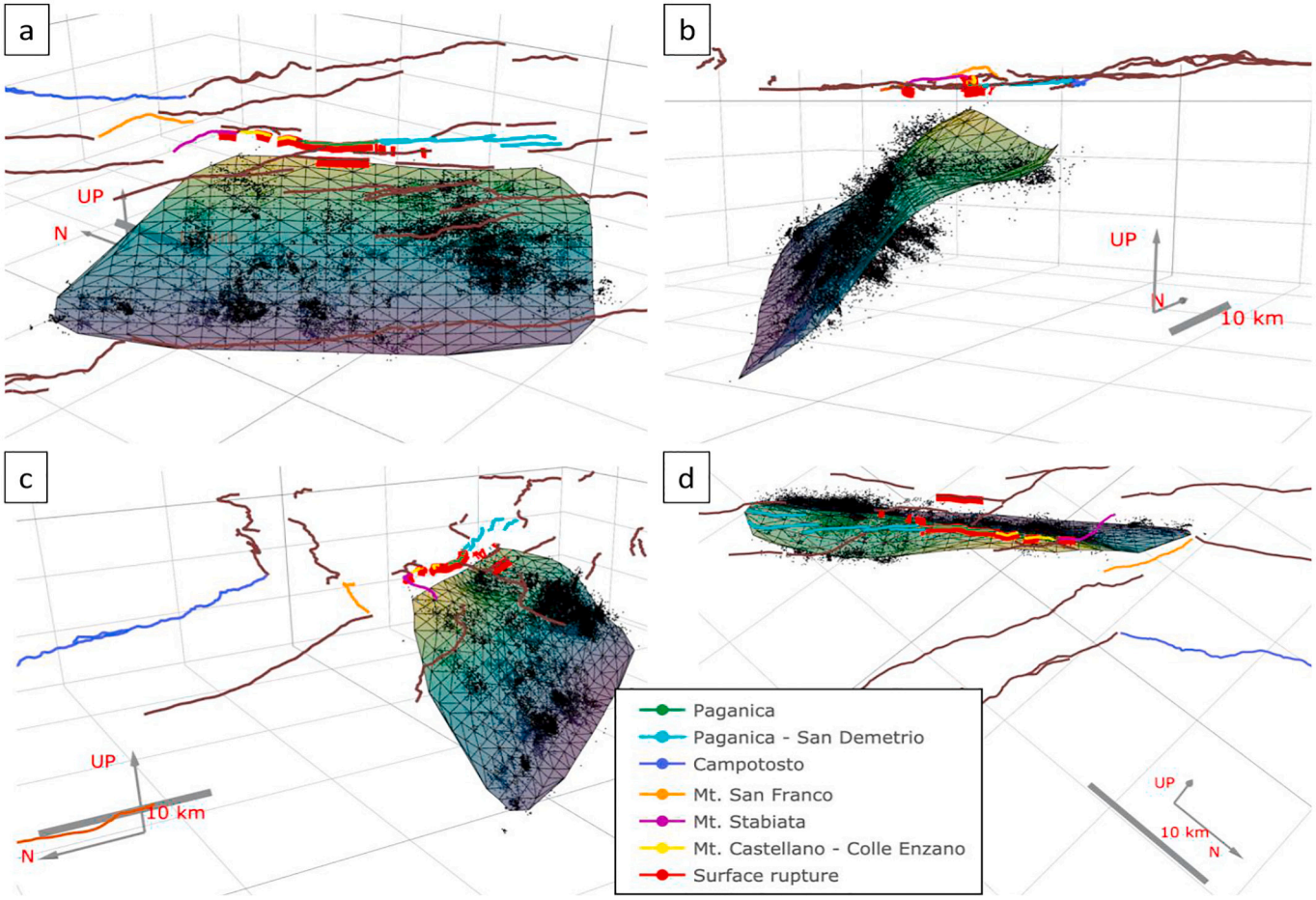

Fig. 5. Also available interactively (see supplementary material or online at https://bit.ly/3k7ykqg). Interpolated fault surface compared to locations of hypocenters in our Paganica fault system model. We plot the individual Paganica fault trace as green, though our Paganica fault model represents the whole Paganica fault system including faults between the Mt. Stabiata and San Demetrio faults. The perspectives are such that (a) views northeast, (b) views northwest, (c) views southeast, and (d) views southwest. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

intersecting the Paganica cluster contained very dense aftershocks and these were not distinguishable from the Paganica fault through clustering with our chosen parameters. In particular, there are three faults at the south-east portion of and antithetic to the main Paganica fault model which were too closely connected to the Paganica fault to be distinguished (Fig. 5). As reported in Ouillon et al. (2008), allowing the number of clusters to increase separates hypocenters which belong to the same fault, while having less clusters tends to lump hypocenters of different faults together. In general, the choice of clustering parameters involves a trade-off between capturing the full seismogenic portion of a fault within a single cluster versus separating out the minor fault splays (Fig. S3). Another example of this trade-off is illuminated where the fault which hosted the M_W 5.4 aftershock was considered mostly noise due its low aftershock density, despite hosting a large earthquake. Using focal mechanisms with manual hypocenter interpolation, Lavecchia et al. (2017) instead modelled this and nearby aftershocks as occurring on an east dipping fault.

The 3D Paganica fault surface that best fits the clustered hypocenters can be seen in Figs. 4 and 5. The modelled fault surface has a length of ~20 km at the shallow edge of the fault. This is longer than what is often used for modelling the mainshock (often between 12 km and 17 km: see reviews of Chiaraluce, 2012; Vannoli et al., 2012), and longer than geologically measured deformation (e.g. Boncio et al., 2010; Vannoli et al., 2012). However, this length is consistent with evidence suggesting that the Paganica-San Demetrio fault system essentially acts as a SW dipping normal fault system extending approximately 19 km (Calderoni et al., 2012; Civico et al., 2015; Civico et al., 2017). The northwest tip of our fault model extends beneath mapped surface ruptures (Boncio et al., 2010) of the fault system including at the Mt. Castellano-Colle Enzano faults and the similarly striking portion of the Mt. Stabiata fault. The NW shallow edge of the fault model is where the Mt. Stabiata fault transitions to *E*-W striking. This contrasts slightly with the models of Lavecchia et al. (2012, 2017), where the E-W striking portion of the Mt. Stabiata fault is also combined into their Paganica fault system models. The fault surface extends SE to the termination of the San Demetrio fault. This same SE fault extent agrees with other Paganica fault system models (e.g. Lavecchia et al., 2012; Civico et al., 2017). Ultimately, our fault model extent is coherent with the radiography of the seismogenic Paganica fault and with mapped faults as discussed in Chiaraluce (2012) and Valoroso et al. (2013) based on the same dataset. However, a notable difference is that our fault model extends ~34.0 km along strike at its base. This is because a group of aftershocks at the NW termination of our fault model lies in alignment with but coherently with the more traditionally established Paganica fault system hypocenters (Figs. 3 and 5).

The fault surface extends to about 10 km depth which is the maximum depth of the hypocenters (Valoroso et al., 2013). The width of the fault (along dip) is 15.4 km, which is similar to values often used to model the mainshock (about 11 km - 17 km, see Chiaraluce, 2012). These are also similar dimensions to those of the 3D model of Lavecchia et al., (2012): 15 km along dip and maximum depth of 10.5 km.

Visually speaking, the surface is smooth and does not deviate greatly from a plane. It shows curvature similar to other 3D fault models generated with similar methods in other locations (Carena and Suppe, 2002; Kaven and Pollard, 2013) and for the Paganica fault system (Lavecchia et al., 2012; Lavecchia et al., 2017; Castaldo et al., 2018). The basic plane from which the surface is built (where $x_3' = 0$), which indicates the average geometry of our Paganica fault model, has a strike and dip of N143° and 42.6°, respectively. This is in close agreement with (Lavecchia et al., 2012), whose 3D reconstruction suggested a strike of ~N140° and dip of ~45°. A compilation of 9 previous L'Aquila mainshock models shows values of strike between N133° and N144° with an average of N137°, and values of dip between 47° and 56° with an average of 52° (Chiaraluce, 2012). The focal mechanism of the mainshock shows a strike and dip of N135° and 55°. Measurements of surface ruptures strike dominantly N130°–N140° (Boncio et al., 2010; EMERGEO Working Group, 2010).

Approximately southeast of documented Paganica fault surface ruptures, the model wavers as it interacts with the more complex distribution of hypocenters. Here, moving upward, the dip abruptly flattens at about 3 km in depth (Fig. 5b). This flattened portion of the fault model is likely responsible for the lower-than-average dip of our Paganica fault system model. The shallow, low dip portion of the fault model is likely not truly associated with the Paganica fault system, but is possibly associated with the Gran Sasso fault to the east (Chiaraluce et al., 2011). The apparent continuity of these low dipping hypocenters with those of the Paganica fault system is the reason clustering identifies them as one fault surface. The incorporation of fault traces during clustering could help to produce fault models which are more consistent between hypocenter depths and fault traces along the entire length of faults, although automation of such a procedure is outside the scope of this paper.

We note specifically two locations on the Paganica fault system model where issues in the interpolation might have arose; (i) near the south-east edge of the fault where there are conjugate faults intersecting the Paganica fault (Fig. 5b), and (ii) at the north-west edge of the fault where there were few aftershocks for interpolation (Fig. 4, 5a).

Regarding problem (i), because the clustering did not distinguish between some conjugate planes that were closely connected to the Paganica fault system, the surface interpolation is slightly shifted by these conjugate fault hypocenters (Fig. 5b). Because the hypocenters were widely distributed near the fault in the direction normal to the fault (the aftershocks deviate from the 3D surface by ~600 m on average), small variation in the fault surface does not conflict with the hypocenter distribution. It is thus likely not a problem that the surface is shifted slightly in this location. Ultimately, manual fault model construction models did not suffer from this problem (Lavecchia et al., 2012; Lavecchia et al., 2017; Castaldo et al., 2018).

Regarding problem (ii), although the splines were inverted with regularly distributed knots, the hypocenters which were used to find the knot parameters were distributed unevenly. Particularly at the upper NW corner of the fault, interpolation was done where few hypocenters were available. The spline parameters in this location are thus underconstrained. To address this problem, we applied a convex hull to the hypocenters. We only report the portion of our fault model that is within this polygon.

To evaluate the accuracy of our 3D fault model, we compared its orientation to the orientation of all 3422 available earthquake focal mechanisms (Chiaraluce et al., 2021; Fig. 6). This comparison is quantified as the angle between the normal vectors at the fault surface model and the normal vectors of the aftershock focal mechanisms. From the focal mechanisms, both the fault planes and auxiliary planes are used. We compare this angular mismatch result for our 3D fault model with a planar model of Cirella et al. (2012). For both our 3D fault and for the planar model, there is alignment at two primary peaks of 20° and 90° which correspond to the fault and auxiliary planes of the focal mechanisms, respectively (Fig. 6b). Because earthquakes with greater magnitude tend to involve larger rupture area and greater slip, we test the fault surface to focal mechanism alignment weighted according to the magnitude. While a weight of 32^M would be more consistent with the moment and energy released from an earthquake, the results would be completely dominated by the largest earthquake. We thus choose 10^M as a weight to be consistent with the M_L, M_b, and M_S scales. The weighted results are also shown in Fig. 6b, where instead of plotting a histogram of each angular mismatch, we plot a normalized histogram of each angular mismatch weighted by 10^M. Focal mechanism to fault plane orientation agreement improves with this weighting scheme. Note that we do not include the focal mechanism of the M_W 6.1 mainshock in the weighted analysis because its high magnitude and weight would make it dominate the orientation results, and the alignment of our fault model created from aftershocks compared to the mainshock orientation is only of secondary importance.

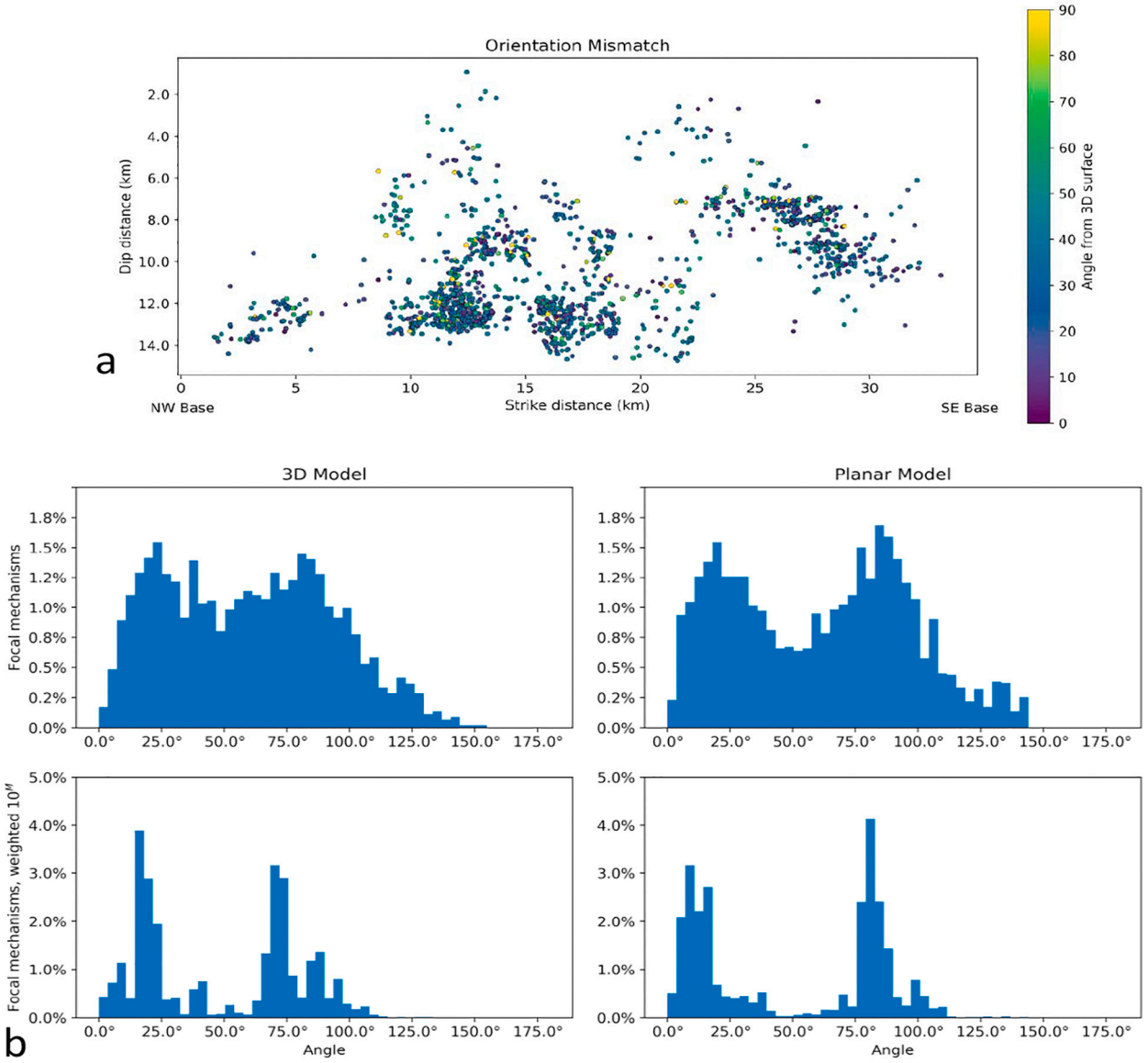

Fig. 6. Angular mismatch between the normal vectors of the interpolated fault surface and the normal vectors of the available earthquake focal mechanisms. (a) Angular mismatch compared to location of hypocenters in the cluster that corresponds to the Paganica fault. Earthquake locations are plotted according to their distance along strike and along dip on the fault surface. Only the best matching angle for each aftershock is used for (a). (b) Histograms showing the number of focal planes with any given angular mismatch (top) weighted by 10^M (bottom). We show both our 3D fault model (left panels) and the planar model used in Cirella et al. (2012) (right panels). Both focal planes of each focal mechanism were included in (b). This results in alignment peaks near 90° that correspond with the auxiliary planes of focal mechanisms. For the fault planes of the focal mechanisms, alignment peaks are closer to 20°. Orientation match is improved when using the magnitude as a weight.

4. Discussion and conclusions

There is a need to develop and improve algorithms which can find the 3D morphology of faults. The primary purpose here is to explore a new algorithm to do so which is fast and objective and to investigate the utility of machine learning for this purpose. We used successive iterations of clustering to construct a quantitative model of the fault segments associated with the L'Aquila seismic sequence. The approach, based on a combination of spectral and DBSCAN clustering, provides several advantages. The method can be applied in only a few seconds on a standard laptop, saving time in the case of analyzing a seismic sequence in real time, while it is also objective in its determination of faults. In this way, it provides advantages over the high-resolution options of manual fault interpolation. It is advantageous over planar clustering approaches in that it estimates the continuous first order geometry of a fault zone. Curved faults, for instance the listric Campotosto fault, can thus be represented with one cluster (Fig. 3). Using planar clustering, a curved fault would instead be represented with several disconnected planes (e.g. Wang et al., 2013) which complicates modelling the rupture of a mainshock along its entire extent. Our approach is further advantageous over algorithms which cluster based on the distribution of points around each cluster's average point, such as k-means, which would find diffuse clouds of seismicity that are suitable for methods which are not dependent on fault morphology (Weatherill and Burton, 2009).

4.1. Parameter choice

Although the algorithm is objective, the parameter choices (Table 1) require care as these can have important influences on the results. The

reader can experiment with different parameter choices in the Supplementary Jupyter Notebook.

For our application of DBSCAN, the choice of Eps, MinPts, and f requires human supervision. By decreasing f, DBSCAN incorporates hypocenters which are more widely distributed about faults and it tends to group multiple faults together. However, by changing MinPts and Eps to reflect different values of f, similar results to our final clusters can be ultimately obtained. We showed an example of our Paganica cluster and fault using $f = 5$ (Figs. 3–5) and $f = 1$ (Fig. S2). Fig. S3 shows the influence of Eps and MinPts on the Paganica hypocenter cluster. Increasing MinPts increases the tendency to reject potential clusters as noise, resulting in fewer and smaller clusters, and specifically reducing the number of hypocenters in the Paganica cluster. Increasing Eps increases the tendency for hypocenters to be clustered together, potentially merging faults together, and increasing the number of hypocenters in the Paganica hypocenter cluster. We performed a grid test to quantify the influence of Eps and MinPts (Fig. S3). Variations of the parameters about our chosen values did not impart important changes to the Paganica cluster which consistently had approximately the same 29,367 hypocenters. Because of this, the parameters do not necessarily need fine tuning as a wide variety of parameters produce similar results.

The choice of spectral clustering parameters is slightly less crucial here, as it is primarily used for preliminary data operations which are designed to make DBSCAN clustering more successful. The spectral clustering step requires prior knowledge of the number of major hypocenter clusters. We chose three based on visual inspection and trial and error; increasing the number of clusters broke the Paganica fault up along strike which could also be acceptable based on several other Paganica fault system models having smaller distance along strike (e.g. Vannoli et al., 2012). σ was tuned manually to maintain maximal separation between the three dominant clusters (Fig. S1). However, for 1250 m $< \sigma <$ 4300 m, the results varied only negligibly.

4.2. Implications for fault structure

We identified key hypocenter clusters in part according to which geologically identified fault segment they most closely correspond to. However, there are important differences between the hypocenter clusters and faults. Clustering will be blind to portions of faults which were never activated. Thus, the true maximum extent of a fault can be greater than indicated by the hypocenter clusters. Similarly, for separate faults which are coherent with each other in space and shape, clustering may indicate these as being the same fault. This is why our fault model is longer at its base than most other fault models, at ~34.0 km. Clustering does not discriminate in accordance with fault identification made geologically at fine scales, but rather identifies coherent hypocenters. While this may not be desirable for recognizing geological structures on fine scales, it can be more appropriate for modelling earthquakes where rupture is not confined to a single fault plane (e.g. Tondi, 2000; Civico et al., 2017).

We tested the quality of the interpolated Paganica fault surface through several approaches. One approach was to find how well the interpolated fault surface matches the orientation of the earthquakes that it interpolates (Fig. 6). When comparing the match of aftershock orientations to our 3D model and a 2D model (Cirella et al., 2012), there does not appear to be an improvement moving from the planar model to our 3D surface. This is probably due to the fact that the L'Aquila mainshock ruptured on a relatively planar fault segment. Faults with more curvature, such as the Campotosto fault, will likely show more certain improvement by incorporating 3D morphology.

While ideally the aftershock orientations should have matched the 3D fault surface model well, when comparing the mismatch with each aftershock weighted equally, the aftershock orientations only mildly agree with either the planar model or 3D model (Fig. 6). However, when the magnitude of aftershocks (10^M) is used as a weight, the match between aftershocks and the 3D Paganica fault model increases. This suggests that larger earthquakes are more closely aligned with their fault surface and/or the focal mechanisms of smaller events are more poorly constrained. Another plausible reason could be that some aftershocks may have been associated with conjugate faults instead of the main fault surface (Fig. 5b). There may be more unnoticed, smaller conjugate faults as well.

However, the misalignment between the fault model and focal mechanisms may have implications for the structure of the fault. Fault structures can become very complex with damage zones involving various scales and patterns (Peacock and Sanderson, 1991; Valoroso et al., 2014), and aftershock clusters often occur at the most complex regions of a fault surface where damage zones are the most extreme (Sibson, 1989; Kim and Sanderson, 2008; Michele et al., 2020). Not only damage zones, but also the complexity of fault connectivity provides a means for aftershocks to rupture in directions and locations not in agreement with the main fault surface. One model for fault growth requires several smaller faults to link together, for instance through the breaching of relay ramps, in order to form larger faults (Childs et al., 2009). This results in complex rupture geometries. Relay ramps allow for a fault to be connected by many discontinuous faults at small and large scales. Relays have indeed been observed on surface expressions of the Paganica fault (Roberts et al., 2010). Faults can also be connected by extensional steps and contractional steps (Kim et al., 2004). The tips of a fault can be particularly complex as well. Where fault displacement dies out at the tips of faults, displacement tends to end in splays (Kim et al., 2004). Splaying has been observed in the Paganica fault geologically (Cinti et al., 2011) and seismically (Chiaraluce et al., 2011). Given these issues, and that we are modelling a fault zone as a single 3D surface, we expect discontinuous sections of the fault zone to contain aftershocks which align poorly to the main fault surface.

4.3. Future directions

This approach could aid in seismic hazard assessment. Most models of stress transfer that address aftershocks in the L'Aquila seismic sequence rely on flat fault models (e.g. Serpelloni et al., 2012). However, the 3D morphology of the mainshock as well as any recipient faults will influence this stress and earthquake interaction, influencing the development of earthquake sequences (Mildon et al., 2016; Kaven and Pollard, 2013). Our 3D fault model can be used as the basis for a 3D mainshock geometry, and the slip distribution can be inverted using geodetic or waveform data (e.g. Kaven and Pollard, 2013; Castaldo et al., 2018). This approach may illuminate important foreshock-mainshock or mainshock-aftershock interactions that are otherwise not clear.

This method can be applied to other seismic sequences. The L'Aquila sequence itself is just one part of a larger extensional fault system that has experienced several recent seismic sequences extending over more than 150 km. This includes the 1979 Umbria sequence, the 1997 Colfiorito sequence, and the 2016 Amatrice/Visso/Norcia sequence. Other locations where seismicity is thoroughly monitored, for instance throughout California, could be good sites to apply this technique. Further, applying this method to faults in developing seismic sequences could help illuminate locked zones where a region of a fault experiences seismic quiescence but is surrounded by an active portion of the fault (Avouac et al., 2015).

Funding

This work was partly funded by the US Department of Energy Grant DE-SC0019759 (D.A.Y.) and by the National Science Foundation (NSF) Grant EAR-1918126 (D.A.Y.).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence

the work reported in this paper.

Acknowledgements

We thank Raphael Gottardi, Carl Richter, and Martin Mai for helpful comments and advice on writing. We thank Antonella Cirella, Olaf Zielke, and Mehdi Nikkhoo for giving guidance on completing the code. We thank two anonymous reviewers for helping to improve the manuscript. 3D figures were made and provided interactively using Plotly (https://plotly.com/).

The Python code for this work is written in Jupyter Notebooks (Kluyver et al., 2016; Morra, 2018) where the interactive code is embedded within the manuscript. This is available in GitHub (https://github.com/brennanbrunsvik/Fault-morphology-clustering) where it can be ran through the cloud service Binder. Online interactive 3D figures are available as well for the clustering and fault model as supplemental html files and online (https://bit.ly/2MnCFdj).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tecto.2021.228756.

References

Ansari, A., Firuzi, E., Etemadsaeed, L., 2015. Delineation of seismic sources in probabilistic seismic-hazard analysis using fuzzy cluster analysis and monte carlo simulation. Bull. Seismol. Soc. Am. 105, 2174–2191.
Avouac, J.P., 2008. Dynamic processes in extensional and compressional settings-mountain building: from earthquakes to geological deformation. Treatise Geophys. 6, 377–439.
Avouac, J.P., Meng, L., Wei, S., Wang, T., Ampuero, J.P., 2015. Lower edge of locked Main Himalayan Thrust unzipped by the 2015 Gorkha earthquake. Nat. Geosci. 8.
Barka, A.A., Kadinsky-Cade, K., 1988. Strike-slip fault geometry in Turkey and its influence on earthquake activity. Tectonics. 7, 663–684.
Ben-Zion, Y., Peng, Z., Okaya, D., Seeber, L., Armbruster, J.G., Ozer, N., Aktar, M., 2003. A shallow fault-zone structure illuminated by trapped waves in the Karadere–Duzce branch of the North Anatolian Fault, western Turkey. Geophys. J. Int. 152, 699–717.
Bergen, K.J., Johnson, P.A., Maarten, V., Beroza, G.C., 2019. Machine learning for data-driven discovery in solid Earth geoscience. Science. 363, eaau0323.
Bletery, Q., Thomas, A.M., Rempel, A.W., Karlstrom, L., Sladen, A., De Barros, L., 2016. Mega-earthquakes rupture flat megathrusts. Science. 354, 1027–1031.
Boncio, P., Pizzi, A., Brozzetti, F., Pomposo, G., Lavecchia, G., Di Naccio, D., Ferrarini, F., 2010. Coseismic ground deformation of the 6 April 2009 L'Aquila earthquake (central Italy, Mw6. 3). Geophys. Res. Lett. 37.
Calderoni, G., Di Giovambattista, R., Vannoli, P., Pucillo, S., Rovelli, A., 2012. Fault-trapped waves depict continuity of the fault system responsible for the 6 April 2009 Mw 6.3 L'Aquila earthquake, Central Italy. Earth Planet. Sci. Lett. 323, 1–8.
Carena, S., Suppe, J., 2002. Three-dimensional imaging of active structures using earthquake aftershocks: the Northridge thrust, California. J. Struct. Geol. 24, 887–904.
Carena, S., Suppe, J., Kao, H., 2002. Active detachment of Taiwan illuminated by small earthquakes and its control of first-order topography. Geology. 30, 935–938.
Castaldo, R., De Nardis, R., DeNovellis, V., Ferrarini, F., Lanari, R., Lavecchia, G., Tizzani, P., 2018. Coseismic Stress and Strain Field Changes Investigation Through 3-D Finite Element Modeling of DInSAR and GPS Measurements and Geological/Seismological Data: The L'Aquila (Italy) 2009 Earthquake Case Study. J. Geophys. Res. Solid Earth 123, 4193–4222.
Chiaraluce, L., 2012. Unravelling the complexity of Apenninic extensional fault systems: a review of the 2009 L'Aquila earthquake (Central Apennines, Italy). J. Struct. Geol. 42, 2–18.
Chiaraluce, L., Valoroso, L., Piccinini, D., Di Stefano, R., De Gori, P., 2011. The anatomy of the 2009 L'Aquila normal fault system (central Italy) imaged by high resolution foreshock and aftershock locations. J. Geophys. Res. Solid Earth 116.
Childs, C., Manzocchi, T., Walsh, J.J., Bonson, C.G., Nicol, A., Schöpfer, M.P., 2009. A geometric model of fault zone and fault rock thickness variations. J. Struct. Geol. 31, 117–127.
Cinti, F.R., Pantosti, D., De Martini, P.M., Pucci, S., Civico, R., Pierdominici, S., Patera, A., 2011. Evidence for surface faulting events along the Paganica fault prior to the 6 April 2009 L'Aquila earthquake (central Italy). J. Geophys. Res. Solid Earth 116.
Cirella, A., Piatanesi, A., Tinti, E., Chini, M., Cocco, M., 2012. Complexity of the rupture process during the 2009 L'Aquila, Italy, earthquake. Geophys. J. Int. 190, 607–621.
Civico, R., Pucci, S., De Martini, P.M., Pantosti, D., 2015. Morphotectonic analysis of the long-term surface expression of the 2009 L'Aquila earthquake fault (Central Italy) using airborne LiDAR data. Tectonophysics. 644, 108–121.
Civico, R., Sapia, V., Di Giulio, G., Villani, F., Pucci, S., Baccheschi, P., Pantosti, D., 2017. Geometry and evolution of a fault-controlled Quaternary basin by means of TDEM and single-station ambient vibration surveys: The example of the 2009 L'Aquila earthquake area, central Italy. J. Geophys. Res. Solid Earth 122, 2236–2259.
Chiaraluce, L., Gori, P., Stefano, R., Valoroso, L., 2021. L'Aquila 2009 seismic sequence: integrated dataset of automatic first motion polarities focal mechanisms and RMT with HypoDD high quality relative earthquake locations (Version 1.0). Zenodo. https://doi.org/10.5281/zenodo.4024065.
Custódio, S., Lima, V., Vales, D., Cesca, S., Carrilho, F., 2016. Imaging active faulting in a region of distributed deformation from the joint clustering of focal mechanisms and hypocentres: Application to the Azores–western Mediterranean region. Tectonophysics. 676, 70–89.
Dzwinel, W., Yuen, D., Kaneko, Y., Boryczko, K., Ben-Zion, Y., 2003. Multi-resolution clustering analysis and 3-D visualization of multitudinous synthetic earthquakes. Vis. Geosci. 8, 1–32.
Dzwinel, W., Yuen, D.A., Boryczko, K., Ben-Zion, Y., Yoshioka, S., Ito, T., 2005. Cluster analysis, data-mining, multi-dimensional visualization of earthquakes over space, time and feature space. Nonlinear Process. Geophys. 12, 117–128.
EMERGEO Working Group, 2010. Evidence for surface rupture associated with the Mw 6.3 L'Aquila earthquake sequence of April 2009 (Central Italy). Terra Nova 22, 43–51.
Ester, M., Kriegel, H.P., Sander, J., Xu, X., 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. Kdd. 96, 226–231.
Fischler, M.A., Bolles, R.C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM 24, 381–395.
Fornberg, B., Zuev, J., 2007. The Runge phenomenon and spatially variable shape parameters in RBF interpolation. Comput. Math. Appl. 54, 379–398.
Galvez, P., Ampuero, J.P., Dalguer, L.A., Somala, S.N., Nissen-Meyer, T., 2014. Dynamic earthquake rupture modelled with an unstructured 3-D spectral element method applied to the 2011 M 9 Tohoku earthquake. Geophys. J. Int. 198, 1222–1240.
Hall, T.R., Nixon, C.W., Keir, D., Burton, P.W., Ayele, A., 2018. Earthquake Clustering and Energy Release of the African–Arabian Rift System. Bull. Seismol. Soc. Am. 108, 155–162.
Hubbard, J., Almeida, R., Foster, A., Sapkota, S.N., Bürgi, P., Tapponnier, P., 2016. Structural segmentation controlled the 2015 Mw 7.8 Gorkha earthquake rupture in Nepal. Geology. 44, 639–642.
Joshi, A., Kaur, R., 2013. A review: Comparative study of various clustering techniques in data mining. Int. J. Adv. Res. Comp. Sci. Software Eng. 3.
Kaven, J.O., Pollard, D.D., 2013. Geometry of crustal faults: Identification from seismicity and implications for slip and stress transfer models. J. Geophys. Res. Solid Earth 118, 5058–5070.
Kim, Y.S., Sanderson, D.J., 2008. Earthquake and fault propagation, displacement and damage zones. Struct. Geol. 1, 99–117.
Kim, Y.S., Peacock, D.C., Sanderson, D.J., 2004. Fault damage zones. J. Struct. Geol. 26, 503–517.
Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B.E., Bussonnier, M., Frederic, J., Ivanov, P., 2016. Jupyter Notebooks-a publishing format for reproducible computational workflows. ELPUB. 87-90.
Kong, Q., Trugman, D.T., Ross, Z.E., Bianco, M.J., Meade, B.J., Gerstoft, P., 2019. Machine Learning in Seismology: Turning Data into Insights. Seismol. Res. Lett. 90, 3–14.
Landry, W., Barbot, S., 2018. Fast, accurate solutions for curvilinear earthquake faults and anelastic strain. In: arXiv preprint. arXiv:1802.08931.
Lavecchia, G., Ferrarini, F., Brozzetti, F., De Nardis, R., Boncio, P., Chiaraluce, L., 2012. From surface geology to aftershock analysis: Constraints on the geometry of the L'Aquila 2009 seismogenic fault system. Ital. J. Geosci. 131, 330–347.
Lavecchia, G., Castaldo, R., De Nardis, R., De Novellis, V., Ferrarini, F., Pepe, S., Boncio, P., 2016. Ground deformation and source geometry of the 24 August 2016 Amatrice earthquake (Central Italy) investigated through analytical and numerical modeling of DInSAR measurements and structural-geological data. Geophysical Research Letters. 43, 12–389.
Lavecchia, G., Adinolfi, G.M., de Nardis, R., Ferrarini, F., Cirillo, D., Brozzetti, F., Zollo, A., 2017. Multidisciplinary inferences on a newly recognized active east-dipping extensional system in Central Italy. Terra Nova. 29, 77–89.
Lohr, T., Krawczyk, C.M., Oncken, O., Tanner, D.C., 2008. Evolution of a fault surface from 3D attribute analysis and displacement measurements. J. Struct. Geol. 30, 690–700.
Mancini, S., Segou, M., Werner, M.J., Cattania, C., 2019. Improving physics-based aftershock forecasts during the 2016–2017 Central Italy Earthquake Cascade. J. Geophys. Res. Solid Earth 124, 8626–8643.
Michele, M., Chiaraluce, L., Di Stefano, R., Waldhauser, F., 2020. Fine-scale structure of the 2016–2017 Central Italy seismic sequence from data recorded at the Italian National Network. J. Geophys. Res. Solid Earth 125.
Mildon, Z.K., Toda, S., Faure Walker, J.P., Roberts, G.P., 2016. Evaluating models of Coulomb stress transfer: is variable fault geometry important? Geophys. Res. Lett. 43, 12–407.
Moore, G.F., Bangs, N.L., Taira, A., Kuramoto, S., Pangborn, E., Tobin, H.J., 2007. Three-dimensional splay fault geometry and implications for tsunami generation. Science. 318, 1128–1131.
Morra, G., 2018. Pythonic Geodynamics (Lecture Notes in Earth System Sciences).
Mousavi, S.M., Beroza, G.C., 2019. A machine-learning approach for earthquake magnitude estimation. Geophys. Res. Lett. 47.
Novianti, P., Setyorini, D., Rafflesia, U., 2017. K-Means cluster analysis in earthquake epicenter clustering. Int. J. Adv. Intell. Inform. 3, 81–89.
Omran, M.G., Engelbrecht, A.P., Salman, A., 2007. An overview of clustering methods. Intell. Data Anal. 11, 583–605.

Ouillon, G., Sornette, D., 2011. Segmentation of fault networks determined from spatial clustering of earthquakes. J. Geophys. Res. Solid Earth 116.
Ouillon, G., Ducorbier, C., Sornette, D., 2008. Automatic reconstruction of fault networks from seismicity catalogs: Three-dimensional optimal anisotropic dynamic clustering. J. Geophys. Res. Solid Earth 113.
Peacock, D.C.P., Sanderson, D.J., 1991. Displacements, segment linkage and relay ramps in normal fault zones. J. Struct. Geol. 13, 721–733.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Vanderplas, J., 2011. Scikit-learn: Machine learning in Python. Journal of machine learning research. 12, 2825–2830.
Pelties, C., Puente, J., Ampuero, J.P., Brietzke, G.B., Käser, M., 2012. Three-dimensional dynamic rupture simulation with a high-order discontinuous Galerkin method on unstructured tetrahedral meshes. J. Geophys. Res. Solid Earth 117.
Peng, Z., Ben-Zion, Y., Michael, A.J., Zhu, L., 2003. Quantitative analysis of seismic fault zone waves in the rupture zone of the 1992 Landers, California, earthquake: evidence for a shallow trapping structure. Geophys. J. Int. 155, 1021–1041.
Prautzsch, H., Boehm, W., Paluszny, M., 2002. Bézier and B-spline techniques (Springer Science & Business Media).
Qiu, Q., Hill, E.M., Barbot, S., Hubbard, J., Feng, W., Lindsey, E.O., Tapponnier, P., 2016. The mechanism of partial rupture of a locked megathrust: The role of fault morphology. Geology. 44, 875–878.
Ragon, T., Sladen, A., Simons, M., 2018. Accounting for uncertain fault geometry in earthquake source inversions–I: theory and simplified application. Geophys. J. Int. 214, 1174–1190.
Rietbrock, A., Tiberi, C., Scherbaum, F., Lyon-Caen, H., 1996. Seismic slip on a low angle normal fault in the Gulf of Corinth: evidence from high-resolution cluster analysis of microearthquakes. Geophys. Res. Lett. 23, 1817–1820.
Roberts, G.P., Raithatha, B., Sileo, G., Pizzi, A., Pucci, S., Walker, J.F., Guerrieri, L., 2010. Shallow subsurface structure of the 2009 April 6 M w 6.3 L'Aquila earthquake surface rupture at Paganica, investigated with ground-penetrating radar. Geophysical Journal International. 183, 774–790.
Røe, P., Georgsen, F., Abrahamsen, P., 2014. An uncertainty model for fault shape and location. Math. Geosci. 46, 957–969.
Rouet-Leduc, B., Hulbert, C., Lubbers, N., Barros, K., Humphreys, C.J., Johnson, P.A., 2017. Machine learning predicts laboratory earthquakes. Geophys. Res. Lett. 44, 9276–9282.
Schubert, E., Sander, J., Ester, M., Kriegel, H.P., Xu, X., 2017. DBSCAN revisited, revisited: why and how you should (still) use DBSCAN. ACM Trans. Database Syst. (TODS). 42, 1–21.
Scitovski, S., 2018. A density-based clustering algorithm for earthquake zoning. Comput. Geosci. 110, 90–95.
Scognamiglio, L., Tinti, E., Michelini, A., Dreger, D.S., Cirella, A., Cocco, M., Piatanesi, A., 2010. Fast determination of moment tensors and rupture history: What has been learned from the 6 April 2009 L'Aquila earthquake sequence. Seismological Research Letters. 81, 892–906.
Serpelloni, E., Anderlini, L., Belardinelli, M.E., 2012. Fault geometry, coseismic-slip distribution and Coulomb stress change associated with the 2009 April 6, M W 6.3, L′ Aquila earthquake from inversion of GPS displacements. Geophys. J. Int. 188, 473–489.
Sibson, R.H., 1989. Earthquake faulting as a structural process. J. Struct. Geol. 11, 1–14.
Skoumal, R.J., Kaven, J.O., Walter, J.I., 2019. Characterizing seismogenic fault structures in Oklahoma using a relocated template-matched catalog. Seismol. Res. Lett. 90, 1535–1543.
Tondi, E., 2000. Geological analysis and seismic hazard in the central Apennines (Italy). J. Geodyn. 29, 517–533.
Trugman, D.T., Shearer, P.M., 2017. GrowClust: a hierarchical clustering algorithm for relative earthquake relocation, with application to the Spanish Springs and Sheldon, Nevada, earthquake sequences. Seismol. Res. Lett. 88, 379–391.
Tung, S., Masterlark, T., 2018. Resolving Source Geometry of the 24 August 2016 Amatrice, Central Italy, earthquake from InSAR Data and 3D Finite-Element ModelingResolving Source Geometry of the 2016 Amatrice Earthquake from InSAR Data and 3D FEMs. Bull. Seismol. Soc. Am. 108, 553–572.
Valoroso, L., Chiaraluce, L., Piccinini, D., Di Stefano, R., Schaff, D., Waldhauser, F., 2013. Radiography of a normal fault system by 64,000 high-precision earthquake locations: the 2009 L'Aquila (Central Italy) case study. J. Geophys. Res. Solid Earth 118, 1156–1176.
Valoroso, L., Chiaraluce, L., Collettini, C., 2014. Earthquakes and fault zone structure. Geology. 42, 343–346.
Valoroso, L., Chiaraluce, L., Piccinini, D., Stefano, R., Schaff, D., Waldhauser, F., 2020. The 2009 Mw 6.1 L'Aquila normal fault system imaged by 64,051 high-precision foreshock and aftershock locations (Version 1.0). Zenodo. https://doi.org/10.5281/zenodo.4036248.
Vannoli, P., Burrato, P., Fracassi, U., Valensise, G., 2012. A fresh look at the seismotectonics of the Abruzzi (Central Apennines) following the 6 April 2009 L'Aquila earthquake (Mw 6.3). Ital. J. Geosci. 131, 309–329.
Verdecchia, A., Pace, B., Visini, F., Scotti, O., Peruzza, L., Benedetti, L., 2018. The role of viscoelastic stress transfer in long-term earthquake cascades: Insights after the Central Italy 2016–2017 seismic sequence. Tectonics. 37, 3411–3428.
Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., van Mulbregt, P., 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature methods. 17, 261–272.
Von Luxburg, U., 2007. A tutorial on spectral clustering. Stat. Comput. 17, 395–416.
Waldhauser, F., Ellsworth, W.L., 2000. A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, California. Bull. Seismol. Soc. Am. 90, 1353–1368.
Wang, Y., Ouillon, G., Woessner, J., Sornette, D., Husen, S., 2013. Automatic reconstruction of fault networks from seismicity catalogs including location uncertainty. J. Geophys. Res. Solid Earth 118, 5956–5975.
Weatherill, G., Burton, P.W., 2009. Delineation of shallow seismic source zones using K-means cluster analysis, with application to the Aegean region. Geophys. J. Int. 176, 565–588.
Yuen, D.A., Dzwinel, W., Ben-Zion, Y., Kadlec, B., 2009. Visualization of earthquake clusters over multidimensional space. Encyclopedia Comple. Syst. Sci. 2347–2371.
Zhang, Z., Zhang, W., Chen, X., 2014. Three-dimensional curved grid finite-difference modelling for non-planar rupture dynamics. Geophys. J. Int. 199, 860–879.
Zhu, W., Beroza, G.C., 2019. PhaseNet: a deep-neural-network-based seismic arrival-time picking method. Geophys. J. Int. 216, 261–273.
Zielke, O., Galis, M., Mai, P.M., 2017. Fault roughness and strength heterogeneity control earthquake size and stress drop. Geophys. Res. Lett. 44, 777–783.