Systematic detection of clustered seismicity beneath the Southwestern Alps

Eric Beaucé, William B. Frank, Anne Paul, Michel Campillo, Robert D. van Der Hilst

Systematic Detection of Clustered Seismicity Beneath the Southwestern Alps

Eric Beaucé[1], William B. Frank[2], Anne Paul[3], Michel Campillo[3,1], Robert D. van der Hilst[1]

[1]Massachusetts Institute of Technology, USA
[2]University of Southern California, USA
[3]Institut des Sciences de la Terre, France

Key Points:

- Automated earthquake detection combining beamforming, event classification and template matching applied to the Southwestern Alps.
- Systematic analysis of earthquake clustering by estimating the fractal dimension of time series of earthquake occurrence.
- Burst-like seismicity in fractured medium in the Briançonnais and Dora Maira vs fluid-driven swarm-like seismicity in the Ubaye valley.

Corresponding author: Eric Beaucé, ebeauce@mit.edu

Abstract

We present a new automated earthquake detection and location method based on beamforming (or back projection) and template matching, and apply it to study the seismicity of the Southwestern Alps. We use beamforming with prior knowledge of the 3D variations of seismic velocities as a first detection run to search for earthquakes that are used as templates in a subsequent matched-filter search. Template matching allows us to detect low signal to noise ratio events, and thus to obtain a high spatiotemporal resolution of the seismicity in the Southwestern Alps. We describe how we address the problem of false positives in energy-based earthquake detection with supervised machine learning, and how to best leverage template matching to iteratively refine the templates and the detection. We detected 18,754 earthquakes over one year (our catalog is available online), and observed temporal clustering of the earthquake occurrence in several regions. This statistical study of the collective behavior of earthquakes provides insights into the mechanisms of earthquake occurrence. Based on our observations, we infer the mechanisms responsible for the seismic activity in three regions of interest: the Ubaye valley, the Briançonnais and the Dora Maira massif. Our conclusions point to the importance of fault interactions to explain the earthquake occurrence in the Briançonnais and the Dora Maira massif, whereas fluids seem to be the major driving mechanism in the Ubaye valley.

1 Introduction

Earthquake catalogs are the cornerstone of many studies in seismology, such as characterizing the seismic source (*e.g.* Abercrombie, 1995; Ide et al., 2003), estimating the amount of stress released at plate margins and understanding the role of repeating seismicity in this releasing process (*e.g.* Nadeau et al., 1995; Wech & Creager, 2011; Shelly et al., 2011; Frank et al., 2014), constructing reference earth models (*e.g.* Dziewonski & Anderson, 1981; B. Kennett & Engdahl, 1991; B. L. Kennett et al., 1995), seismic tomography (*e.g.* Dziewonski & Woodhouse, 1987; Van der Hilst et al., 1997; Li et al., 2008), seismic hazard estimation (*e.g.* on California Earthquake Probabilities, 1995), or modeling of the earthquake cycle (model calibration, *e.g.* Richards-Dinger & Dieterich, 2012). The first generation of regional and global catalogs were based on phase arrival picks on analog records (*e.g.* Engdahl et al., 1998). With the advent of digital recording, energy-based detection methods such as the short-term/long-term average (STA/LTA, Allen, 1982) method became popular.

The transition to digital recording and storage, the implementation of protocols for data curation and sharing, the increasing availability of data from networks and arrays, and the recognition of different types of earthquake signals motivated the development of more sophisticated earthquake detection and location algorithms, based, for instance, on array processing (*e.g.* Meng & Ben-Zion, 2017), or learning methods, such as neural networks (*e.g.* Perol et al., 2018). Automated data processing is not only essential for extracting signal from large, and rapidly increasing, data volumes, it also leads to uniform catalog quality.

Analysis of the seismic wavefield recorded at multiple sensors leverages the coherency of the signal across the station array to detect seismic phases which human eyes would have failed to identify. Network-based detection has led to the identification of phenomena such as low frequency earthquakes (*e.g.* Shelly et al., 2007; Brown et al., 2008; Frank et al., 2014) and non-volcanic tremor (*e.g.* Obara, 2002; Rogers & Dragert, 2003).

We develop an earthquake detection method that combines array processing, or, more precisely, a beamformed network response (Frank & Shapiro, 2014) and template matching (Gibbons & Ringdal, 2006; Shelly et al., 2007; Frank & Shapiro, 2014; Ross et al., 2019). Template matching is known to be efficient at detecting low signal-to-noise ratio (SNR) signals (*i.e.* with SNR < 1), and the required prior knowledge of the target seismicity is obtained from the beamformed network response.

We applied this new detection algorithm to one year of seismic data from 87 stations located in the Southwestern Alps, between August 2012 and August 2013, including 55 stations from the temporary network CIFALPS (*cf.* Zhao et al., 2016, and see more information in Data and Resources). Although the Western Alps have been studied for a long time, the mechanisms driving the seismicity are still not well understood (*cf.* Nocquet, 2012, and references therein), and a more complete earthquake catalog will make possible new studies to investigate the tectonic processes that cause them. The Alps were formed following the closure of the Alpine Tethys ocean, due to converging motion between Europe and Africa. The mountain range is located at the border between the Eurasian plate and the Adriatic plate (*cf.* Figure 1). In the Western Alps, Chopin (1984) gave the first petrological evidence for continental subduction, which was later confirmed by several geophysical studies (*e.g.* Nicolas et al., 1990; Zhao et al., 2015). It is unclear, however, whether subduction is still taking place. Even though geodetic data show that the Adriatic plate is rotating counterclockwise with respect to stable Europe (*e.g.* Serpelloni et al., 2007), there is no observation of shortening in the Western Alps and part of the seismic activity is observed to occur under an extensional regime (*c.f.* analysis of earthquake focal mechanisms, Delacou et al., 2004). Various studies (*e.g.* Delacou et al., 2004; Nocquet et al., 2016; Walpersdorf et al., 2018) show that the earthquake activity in the Southwestern Alps is likely to be due to a complex combination of plate tectonic forces and other forces such as buoyancy forces or post glacial rebound. A more detailed characterization of seismic activity, which is indicative of active deformation, will help address these issues.

Figure 1. Interpretative cross-section of the Western Alps. Following the closure of the Alpine Tethys ocean, the collision of the European and Adriatic margins formed the Alps and the subduction complex illustrated here. A clear understanding of what is driving the deformation and the seismic activity in these complex geological units is still lacking. Abbreviations: FPF – Frontal Penninic Fault, Srp – serpentinized, RMF – Rivoli-Marene deep fault. We show the locations of the CIFALPS stations on the topographic profile of the cross-section. The onset shows the location of the transect in the Western Alps, Europe. Figure modified from Zhao et al. (2015) and Solarino et al. (2018).

We first describe the earthquake detection method, and then present the earthquake catalog we thus obtained in the Southwestern Alps. We gain new insights into the seismicity of the study region by investigating the collective behavior of earthquakes, made possible by the large number of detected events. We then discuss the importance of earthquake interaction in the observed behavior of clustered seismicity.

2 Earthquake Detection Method

Detecting low SNR seismic signals by means of template matching requires knowledge of the type of signal to search for in the data. This can be obtained from an existing earthquake catalog or from a preliminary detection run. Since the former is not publicly available for our study area, we produced a preliminary catalog using the energy-based detection method from Frank and Shapiro (2014), which is described in the following. The events thus found were then used as template events in a subsequent matched-filter search.

2.1 Data Pre-processing

We used seismic data recorded between August 2012 and August 2013 at 87 seismic stations in the Southwestern Alps. The network includes 55 broadband sensors from

the temporary CIFALPS array (China-Italy-France Alps survey, Zhao et al., 2016, sampling at 100 Hz), and 32 broadband sensors from French and Italian networks (sampling at 100 Hz or 125 Hz, see Data and Resources). The data are downsampled to 50 Hz and filtered in the band 1-12Hz, which we found was a good compromise between targeting the frequency band of interest for observing local earthquakes and removing undesired signal.

2.2 Energy-based Detection (Composite Network Response)

The beamformed network response method due to Frank and Shapiro (2014) seeks to determine the origin, in time and space, of the seismic energy recorded at an array. This approach leverages the coherency of seismic energy across a receiver array for automatic event detection. Using wave speeds according to a 3-D reference model (Potin, 2016), the apparent travel times measured in the seismograms are then associated with a source location.

As a toy example, let us consider the earthquake whose location is indicated by a yellow star in Figure 2, and whose waveforms are recorded at multiple stations at the surface. Because spatial coherency of the seismic waveforms is not ensured (*e.g.* due to crustal heterogeneities or focal mechanism), we prefer to work with the envelopes of the waveforms. The envelope is the amplitude of the analytical representation of a time series, it is calculated after the preprocessing described in Section 2.1 and the processing of the data is illustrated in Figure S1. We first discretize the volume beneath the study region into a grid of points, each of which representing a possible location of the seismic source (*cf.* Figure 2A). Each of these hypothetical sources is associated with a collection of P- and S-wave travel times to each of the stations. For a sufficiently accurate velocity model, the travel times from the potential source closest to the real source will provide the best alignment with the envelopes of the seismic data (*cf.* Figure 2B). We define the stack of the shifted envelopes as the network response:

$$\mathrm{NR}_k(t) = \sum_{s,c} f\left(u_{s,c}(t + \tau^k_{s,c})\right). \tag{1}$$

In Equation 1, k identifies a potential source and s, c are the station and the component indexes, respectively. We use the S-wave travel times on the horizontal components and the P-wave travel times on the vertical component; $\tau^k_{s,c}$ is the travel time from potential source k to station s on component c. u is the data and f is some transformation of the seismic waveforms. In our case f relates to the function "envelope" (see Supplementary Material Figure S1). The source k^* that yields the largest network response is found by a grid search and represents a proxy of the real source location. Locating earthquakes through such a grid search, that is, shifting and stacking seismic energy, is also known as back projection or migration (*e.g.* Ishii et al., 2005; Walker et al., 2005; Honda & Aoi, 2009), but the objective here is detection.

For earthquake detection purposes, the quantity of interest is the largest network response of the grid at each time step. We define the composite network response (CNR) as:

$$\mathrm{CNR}(t) = \max_k \{\mathrm{NR}_k(t)\} = \mathrm{NR}_{k^*}. \tag{2}$$

The process of searching for NR_{k^*}, continuously in time, is illustrated in Figure S3.

Figure 2C shows an example of CNR from real data. We postprocess the CNR by removing the baseline – a curve connecting the local minima – to set the noise level to zero (which explains the negative values in CNR). The peaks of CNR that exceed a user-defined threshold are detections of events, and the source locations are given by the corresponding k^*. We use the following time-dependent threshold:

$$\mathrm{threshold}(t) = \mathrm{median}\left(\mathrm{CNR}\right)(t) + 10 \times \mathrm{MAD}\left(\mathrm{CNR}\right)(t), \tag{3}$$

Figure 2. Top left panel (A): Spatial discretization of the volume beneath the study region. Using a velocity model, each point of the grid is associated with a collection of source-receiver travel times. The grid points are called potential seismic sources. As an example, let us consider an earthquake with location shown by the yellow star, and recorded at multiple stations. **Right panel (B):** The envelopes of the earthquake waveforms are shifted using the travel times of a potential seismic source close to the real location (*yellow star*). The shifted envelopes are then stacked to calculate the network response (*green waveform*, *cf.* Equation 1). The resulting network response is intrinsically related to the potential seismic source from which the travel times were calculated: different potential seismic sources give different network responses. **Bottom panel (C):** Composite network response (*cf.* Equation 2) calculated over one day. We subtract a curve connecting the local minima of the CNR to set its baseline to zero. To adapt to variations in the level of noise, we use a time-dependent threshold: the value "median $+10 \times$ MAD" is evaluated every 30 minutes and a linear interpolation makes the threshold varying continuously within each 30-minute bin. Using small bin sizes enables the threshold to adapt to locally noisy episodes, but at the risk of discarding actual events: a 30-minute bin size is a good compromise between the two. We perform the peak selection on a smoothed CNR and impose a minimum peak distance, which explains why some of the values above threshold are not selected.

where MAD stands for median absolute deviation. We evaluate median $(\mathrm{CNR}) + 10 \times$ MAD (CNR) in 30-minute bins and make a continuously varying threshold by linearly interpolating the values obtained every 30 minutes.

Each detection yields a so-called template event (located at k^*), and the template for that event is then built by extracting waveforms using the detection time, travel times from k^* to each of the stations considered in the template (in our case, the 20 stations that are closest to k^*), and a window length (we choose 8 seconds). For our application in Section 3, we considered potential sources 1 km apart on a regular 3D cartesian grid (to 80 km depth) beneath a geographic area from 5.5°-9.0°E in longitude and 43.5°-46.0°N in latitude. This 1 km spacing is a good compromise between computation time, array sizes and detection performances.

2.3 Classification of Seismic Signals

Before using a template in a matched-filter search it is important to verify that the signal is due to an earthquake, because the CNR can be influenced by non-earthquake signals, such as proximal noise sources, electronic noise, and by issues in the preprocessing. For this purpose, we conduct a signal classification step prior to template matching.

For automated analysis and signal classification we use supervised machine learning: to discriminate earthquakes from non-earthquakes, an algorithm is trained on a relatively small set of examples classified by a human expert. Our algorithm computes a linear combination of the signal features to generate a scalar that is fed into the logistic function (bounds the output between 0 and 1), which gives the probability of being an earthquake. Therefore, our algorithm is a binary logistic classifier. More information on the structure of the classifier is provided in Figure 3. For each three-component record extracted from the 20 stations, we calculate five features:

1. the amplitude maximum,
2. the first three statistical moments of the distribution of the peaks of the waveform autocorrelation function: variance, skewness, and kurtosis,

Figure 3. Left panel (A): We randomly sample detections from the database of candidate template events and identify each channel as *earthquake* or *non-earthquake*. We attribute the label *earthquake* to the detections with more than nine channels identified as earthquakes (*non-earthquake* otherwise). This arbitrary choice can be tuned in order to select more or less low SNR earthquakes in the template database. **Right panel (B):** Structure of our binary logistic classifier. The signal features are first preprocessed by standardizing them (*i.e.* removing the mean and setting the standard deviation to one) and bounding them between -1 and 1 through the use of hyperbolic tangent. A linear combination of the preprocessed signal features generates a scalar, which is fed into the logistic function (also called sigmoid function). The resulting output is bounded between 0 and 1, and is interpreted as the probability of being an earthquake. An output greater than 0.5 means the detection is more likely to be an earthquake than a non-earthquake. This algorithm was built using the Python library Keras (Chollet et al., 2015).

3. the maximum of the moving kurtosis along the extracted time series,

for a total of 300 features per event detection. The amplitude maxima help identify strong signals, the maximum of the moving kurtosis is sensitive to seismic phase arrivals, and the statistical moments of the autocorrelation function discriminate spikes (with large kurtosis) from impulsive earthquake waveforms. These features are not dependent on the relative phase of the signals, which renders them insensitive to small source mislocation.

For our application, Section 3.1, we manually labeled the waveforms of 500 detections as earthquakes or noise (any non-earthquake signals). We note that labeling the waveforms currently prevents the full automation of the method, but it has to be done only once. In the training dataset, a 60 channel (20 stations $\times$ 3 components) template event is labeled as an earthquake if more than nine channels were individually identified as earthquake waveforms by eyes. This somewhat arbitrary criterion is used to reject the low SNR earthquakes that would not be interesting for use as template events, or which are not identified as earthquakes with high confidence. For training the algorithm, we split the dataset into two independent sub datasets: the training dataset (75% of the detections) and the validation dataset (25% of the detections). Each of these datasets were then augmented by a factor 100 by shuffling the channels in the templates (the classification output must not depend on the order in which the input features are given). While optimizing the classifier with gradient descent on the training dataset, we evaluated the error on the validation dataset and stopped optimizing when this error began to increase. This method, which is known as early stopping (*e.g.* Yao et al., 2007), implicitly regularizes the classifier by providing a criterion for stopping the training when further updating the parameters would only overfit the data. On average, for several randomly selected training and validation datasets, we had a training accuracy of 0.92 and a validation accuracy of 0.90. Eventually, the classification process outputs a database of template events to be used in template matching.

2.4 Template Matching

In seismology, we often approximate the Earth as a linear filter and write an earthquake seismogram as the convolution of a source term with a propagation term and an instrument term:

$$u(r,t) = \underbrace{S(t)M(r;\xi)}_{\text{source}} * \underbrace{G(r,t;\xi)}_{\text{propagation}} * \underbrace{I(t)}_{\text{instrument}} . \tag{4}$$

In Equation 4, the source term is the product of the source time function S and the focal mechanism M that describes effects due to preferred directions in the rupture process (*e.g.* rupture on a fault plane). The propagation term G, the Green's function, describes how the earth responds to an impulsive source for a given travel path. We include site effects in the Green's function. I represents how the recording device distorts actual ground motion. The receiver location is r and the source location is ξ. Equation 4 shows that colocated earthquakes produce similar waveforms because of similar Green's functions. Moreover, similarity is high when the source functions have the same shape (similar focal mechanisms and magnitudes). Template matching leverages this expected similarity to detect new events.

Figure 4. Left panel (A): The waveforms of a template event (*red waveforms*), on 12 stations and each of the 3 components, match well the data (*blue waveforms*): a new earthquake is detected. The correlation coefficient (CC) is given on each channel. **Right panel (B):** Comparison of the template waveform on one channel (*red waveform*) with the waveforms of a few detected events (*blue waveforms*).

Template matching consists of scanning continuous recordings in search for matches between data and the waveforms that constitute a template. This method has proven to be efficient at detecting events with low SNR (SNR < 1, e.g. Gibbons & Ringdal, 2006; Shelly et al., 2007; Frank et al., 2014; Ross et al., 2019). Formally, scanning the data means calculating the correlation coefficient between the template waveforms and the data, continuously in time. We use the following definition of the average correlation coefficient:

$$\mathrm{CC}(t) = \sum_{s,c} w_{s,c} \sum_{n=1}^{N} \frac{T_{s,c}(t_n) u_{s,c}(t + t_n + \tau_{s,c})}{\sqrt{\sum_{n=1}^{N} T_{s,c}^2(t_n) \sum_{n=1}^{N} u_{s,c}^2(t + t_n + \tau_{s,c})}}. \quad (5)$$

In Equation 5, N is the length of the template waveform, n is a temporal index, and $w_{s,c}$ is the weight attributed to station s and component c. If all weights are equal, with $w_{s,c} = 1/N_s N_c$ (with N_s, N_c being the number of stations and components), then it is equivalent to calculating the arithmetic mean. For station s and component c, $T_{s,c}$ is the waveform template, $u_{s,c}$ the continuous data, and $\tau_{s,c}$ the moveout (or time shift) in $u_{s,c}$. The time t is the detection time, meaning that the template window starts at time $\tau_{s,c}$ after the detection time. The template windows start four seconds before the S wave on the horizontal components and one second before the P wave on the vertical component. We note that Equation 5 assumes the mean of $T_{s,c}$ and $u_{s,c}$ within each sliding window of length N is zero. We have shown in previous work that this assumption is correct when the data are filtered such that the lower non-zero period in the data is shorter than the window length (*cf.* Data and Resources and Beaucé et al., 2017). In the application presented in Section 3, template matching was done with a detection threshold of eight times the daily root mean square (RMS) of the correlation coefficient time series. This detection threshold is more conservative than the commonly used threshold of 8×MAD (*e.g.* Shelly et al., 2007; Brown et al., 2008; Baratin et al., 2018, $8 \times \mathrm{RMS} \approx 12 \times \mathrm{MAD}$).

Evaluating the correlation coefficient over long periods of time, and for many templates, requires high performance computing to do it within a reasonable amount of time. We use the software Fast Matched Filter (Beaucé et al., 2017), which is particularly quick when run on graphics processing units (GPUs). The scanning process is illustrated in Figure 4. In the application to data from the Southwestern Alps we use just over 1,400 templates, a template duration of 8 s (with 50 samples per second), and one year of continuous data from 87 3-component stations, and we evaluated CC(t) every sample. Eight seconds is a good compromise between extracting a representative chunk of the target waveform, and a reasonable computation time. Running our codes simultaneously on 12 nodes equipped with one Tesla K20m GPU each took 12 h. As expected, reading operations (I/O) of data and templates is the most time consuming task.

2.5 Second Generation Templates

As illustrated in Figure 4, a matched-filter search provides us with many repetitions of the same target waveform. By stacking the waveforms of the detected events we can enhance the SNR in the template waveform, which decreases the unwanted correlation component of the CC between data and noise in the template, thus improving the quality of the detection, and allows the template events to be located better.

Non-linear stacking, like the Nth-root stack or the phase-weighted stack, greatly improves the SNR with respect to the linear stack, but also distorts the target waveform because of their non-linear nature. Even if it does not enhance SNR as much as non-linear stacking, we prefer the Singular Value Decomposition-based Wiener Filter (SVDWF) because it does not distort the waveform. SVDWF is based on the association of spectral filtering (keeping a limited number of singular vectors from the singular value decomposition) and Wiener filtering, and was initially developed for processing noise correlation functions (Moreau et al., 2017). For each station and each component, the matrix of detected events is first denoised using SVDWF, and a new template waveform is then obtained by stacking the denoised waveforms. Figure S4 illustrates the performance of these different stacking strategies.

Detection and location involve finding the optimal network response for a given f in Equation 1. For detection purposes, we prefer using the envelope for f, but for location purposes, we choose f to be the kurtosis-based transform presented in Figure 5A (from Baillard et al., 2014). This transform makes the signal more sensitive to seismic phase arrivals and, thus, biases the CNR towards finding the travel times that align well the seismic phase arrivals. Performing this relocation process on the second generation template waveforms reduces the spatial spread of the potential sources that yield a large CNR (*cf.* Figure 5, more details in Appendix A).

The second generation templates are used in a subsequent matched-filter search to detect more events. This process – new template generation and matched-filter search – can be iterated several times until the earthquake catalog does not show notable updates between two iterations. During successive iterations, we optimize the template database by regrouping template events with same location and similar waveforms (template events with locations closer than 20 km and with average waveform correlation coefficient greater than 0.8) to avoid redundant matched-filter searches.

Figure 5. Relocation of the second generation templates. **Top panel (A):** The denoised and stacked waveforms obtained from the SVDWF (*blue waveforms*) are transformed following Baillard et al. (2014) to get a signal that is sensitive to phase arrivals (*orange waveforms*). The arrival times predicted by the new location are shown by black and red bars for the P- and S-wave, respectively. **Bottom left panel (B):** The composite network response (*blue curve*) is calculated using the orange signal shown in **A**. The neighborhood of the maximum of the CNR is analyzed to build a weighting function (*red curve*, *cf.* Appendix A for details). This weighting function is used to calculate a weighted average of the distance to the best potential seismic source (*cf.* Equation A3 in Appendix A), *i.e.* the potential source associated with the highest CNR. We define this weighted average as the uncertainty on the location. **Bottom right panel (C):** Each sample of the CNR shown in **B** is associated with a potential source in the grid; the color codes for the value of the CNR and the transparent points are those for which the weighting function is zero. In this example, the location uncertainty is 3.05 km.

3 Seismicity of the Southwestern Alps

We applied the earthquake detection method presented in Section 2 – that is, the combination of the Composite Network Response (CNR), signal classification, and template matching (with SVDWF) – to the preprocessed seismic data described in Section 2.1.

3.1 Catalog

Calculating the CNR as described in Section 2.2 yielded a total of 50,262 detections (candidate template events). After applying the classifier described in Section 2.3, we were left with 1,725 template events. We further reduced this number to 1,406 by regrouping redundant template events (*cf.* Section 2.5); Figure 6 shows their locations. The matched-filter search yielded 18,754 non-redundant detections, with redundancy defined as events with similar waveforms (average CC > 0.8), detected within a time interval of three seconds and from template earthquakes located within 20 km from each other. This arbitrary choice may remove actual earthquakes from the catalog and leave some double counted events but produces a reasonable number of detected events. Our earthquake catalog is available online (see Data and Resources).

Figure 6. Locations of the 1,406 template events. Template events relocated with an uncertainty $\Delta r < 15$ km are shown with filled dots, and template events for which we did not find a reliable location are shown with open diamonds; the color scale codes for the depth of the events. Black inverted triangles are the seismic stations used in this study. We note that the uncertainty estimation described in Section 2.5 does not always perform well for deep events, which do not only feature simple P- and S-wave arrivals as assumed in the calculation of the network response. Therefore, a few events with $\Delta r < 15$ km still show odd locations (*e.g.* deep events located out of the group of deep earthquakes around Torino). The purple star indicates the epicenter of a M_L3.9 earthquake that occurred in early October 2012, and which is important for the discussion in Section 4. The onset shows the position of the Western Alps in Europe. The *black dashed line* corresponds to the axis along which the stations from the CIFALPS network are deployed; this axis is used to project the locations of the template events for 2D cross sections.

To evaluate how well our detection method performs, we compared our catalog to the SISmalp catalog of Potin (2016). The number of events detected and located by our algorithm is more than an order of magnitude larger than the approximately 1,200 included in the SISmalp catalog for our study region; more details on the comparison with this catalog are given in Figures S5 and S6. The events that we seem to have missed all have magnitude less than one and most less than 0.4 (*cf.* Figure S7), which might explain inconsistencies in reported location or non-detection. We note here that other catalogs are also publicly available for this region, such as the Réseau National de Surveillance Sismique catalog with 383 events, and the Istituto Nazionale di Geofisica e Vulcanologia catalog with 743 events.

The temporal distribution of the 18,754 events is shown in Figure 7A with the daily seismic rate. We also report the magnitude of the events for earthquakes with M > 1 and located with high confidence ($\Delta r < 5$ km). These local magnitudes are based on waveform amplitude ratios, they were estimated following the procedure described in Appendix B. Amplitude ratios of events with M < 1 are contaminated by noise and therefore the resulting magnitude estimates are not meaningful. $M = 1$ is also where we observe the Gutenberg-Richter relation to break down (see Figure S8). The daily seismic rate shows continuous seismic activity in the Southwestern Alps, and reveals the existence of episodes of strong, burst-like seismicity (*e.g.* October 2012 and January 2013). Figure 7B shows the earthquake temporal distribution on recurrence time versus detec-

tion time graphs for three templates in distinct geographical regions: the Ubaye valley, the Briançonnais and the Dora Maira massif (*cf.* locations in Figure 1). The recurrence time is the time interval between two colocated earthquakes, and thus is defined template-wise. These three templates offer a representative view of the diversity of seismic behaviors observed in our study region. The Ubaye valley hosts continuous seismic activity without clear sequences of foreshocks-mainshock-aftershocks, but the seismicity of the Briançonnais and the Dora Maira massif are dominated by burst-like episodes. These episodes are characterized by recurrence times spanning many orders of magnitudes, which is the signature of temporal clustering. Seismicity in the Ubaye valley also differs from the burst-like seismicity observed in the Briançonnais and the Dora Maira massif by the smaller magnitude range it spans (*cf.* Figure 7B).

Figure 7. Left panel (A): Daily seismic rate (left axis, *blue continuous curve*) and daily magnitude distribution (right axis, *red dots*). Details on the local magnitude scale are given in Appendix B. **Right panels (B):** Recurrence time vs detection time for three templates located in three distinct geographic regions. The Briançonnais and the Dora Maira massif are dominated by episodes of burst-like seismicity, and the Ubaye valley hosts continuous seismic activity that does not feature clear foreshocks-mainshock-aftershocks sequences. Local magnitudes are coded in color: we observe a smaller magnitude range in the Ubaye valley than for the earthquake sequences in the Briançonnais and in the Dora Maira massif.

3.2 Temporal Clustering of the Seismicity

Unlike Poisson seismicity, clustered earthquake sequences have earthquake occurrence that is not random in time: instead, time clustered seismicity suggests that past events influence the occurrence of future ones. We emphasize that an earthquake sequence with high seismic rate does not have to be clustered in time, but can be Poissonian (*e.g.* Frank et al., 2018). Temporal clustering is often observed for sequences of foreshocks-mainshock-aftershocks (*e.g.* Utsu, 1961; Knopoff, 1964; Gardner & Knopoff, 1974; Zaliapin & Ben-Zion, 2013a) and is thought to be the signature of stress redistribution on neighboring faults taking place during the seismic rupture (*e.g.* Burridge & Knopoff, 1967; Dieterich, 1992; Stein, 1999). More generally, temporal clustering can be explained by various mechanisms implying interactions between earthquakes (*e.g.* Frank et al., 2016). The observation of temporal clustering thus provides a window into the mechanisms of earthquake occurrence.

Quantifying the degree of temporal clustering requires characterization of the time series of earthquake occurrence. While accurate knowledge of the earthquake locations and magnitudes allows sophisticated characterization of clustering in the time-space-energy domain (*e.g.* Zaliapin et al., 2008; Zaliapin & Ben-Zion, 2013b), restricting the analysis to the time-space domain is an appropriate choice for the Southwestern Alps since earthquake magnitudes are small. To describe seismic activity, we introduce the event count $e(t)$ (*cf.* Figure 8A), that is, the number of events in narrow time windows (bins). We characterize clustering by means of the autocorrelation and spectrum of $e(t)$ (Figure 8B and C). By definition, temporal clustering implies temporal correlation of the earthquake occurrence at non-zero correlation time in the autocorrelation function. We observe that clustered earthquake sequences exhibit power-law dependence of $e(t)$ on frequency ($\tilde{e}(f) \propto f^{-\beta}$, similar to Frank et al., 2016). The strength of temporal clustering is quantified by β, referred to as clustering coefficient, which can be estimated from the slope of the spectrum in log-log space (Figure 8C). A strongly clustered earthquake sequence has a large β whereas an earthquake sequence close to a Poisson sequence has a small β, and $\beta = 0$ indicates a purely random sequence (flat spectrum).

Processes exhibiting a power-law spectrum are scale-invariant processes, within a certain range of scales limited by natural bounds. For instance, we expect the power-law $\tilde{e}(f) \propto f^{-\beta}$ to hold between the period of activation of the fault/seismic source (smallest frequency) and the smallest time interval we can resolve between two earthquakes (highest frequency). A powerful analysis tool for scale-invariant time series comes from the theory of fractal clustering (*e.g.* Turcotte, 1997; Lowen & Teich, 2005). Fractal analysis, which has been applied to earthquake occurrence in various studies (*e.g.* Smalley Jr et al., 1987; Lee & Schwarcz, 1995), consists of counting earthquakes in time intervals of variable width. In the case of fractal clustering, the fraction of occupied intervals x has a power-law dependence on the size of the intervals τ, *i.e.* $x \propto \tau^{1-D}$. The fractal dimension D is zero for a Poisson distributed earthquake occurrence, and is typically larger than 0.2 for clustered seismicity (*cf.* Figure 8D). We used correlation time, clustering coefficient β and fractal dimension D to characterize the temporal clustering in our study region. We found that the clustering coefficient was well appropriate for studying clustering over short times, whereas the fractal dimension gave the most contrasted results for studying the long-term clustering (see Supplementary Material Figure S9 and Figure S10). We present our observations of temporal clustering in Figure 9.

Figure 8. Quantification of temporal clustering. **Top left panel (A):** Event count number $e(t)$ for earthquakes detected with two different templates. The event count number is calculated by dividing the time axis into 5-minute bins, and counting the number of events within each bin. **Top right panel (B):** Autocorrelation function of the event count number. We define the correlation time τ as the time interval over which the autocorrelation function is greater than the threshold plotted with the dashed black line (arbitrarily set to 0.12). **Bottom left panel (C):** Power spectral density of the event count number. The spectrum of the event count number has a power-law dependence on the frequency when temporal clustering occurs. We define the power-law exponent β as the clustering coefficient. **Bottom right panel (D):** Fractal analysis of the earthquake sequences. Within a limited range of size of time intervals, the fraction of occupied intervals follows a power-law, whose exponent is related to the fractal dimension of the earthquake occurrence.

Comparison between Figure 9A and Figure 9B shows that there is no trivial correlation between the number of earthquakes per template (*i.e.* number of earthquakes in some volume around the template location) and temporal clustering. We distinguish three geographic regions of high seismic activity: from west to east, the Ubaye valley, the Briançonnais and the Dora Maira massif. The largest temporal clustering is observed beneath the western part of the Dora Maira massif (*cf.* the geological cross-section in Figure 1). The fractal dimension of the event count reveals large temporal clustering also in the southwestern part of the Briançonnais (fractal dimension $D \gtrsim 0.2$). Although we detected a large number of earthquakes beneath the Ubaye valley, we do not observe significant temporal clustering. The seismic activity in the Ubaye valley features a mixture of continuous unclustered seismicity punctuated by episodes of strong, clustered seismicity (see Supplementary Material Figure S9). The Ubaye valley is known to host a seismic swarm (*e.g.* Jenatton et al., 2007; Daniel et al., 2011; Leclère et al., 2012, 2013) that was reactivated in February 2012 by a M3.9 earthquake (Thouvenot et al., 2016). In the following discussion, we refer to swarms as episodes of high seismic activity without substantial temporal clustering (as in, for example, Zaliapin & Ben-Zion, 2013a).

4 Discussion

Frank et al. (2016) present a model where a group of stationary Poisson point processes can lead to a clustered event occurrence if there is interaction between the point

processes. They show that without interaction a coherent acceleration of the Poisson event rates cannot reproduce the clustered distribution as in Figure 8. Poisson point processes describe earthquake occurrences on faults experiencing constant tectonic loading. Therefore, temporal clustering is the signature of earthquake interaction rather than an increase of the external forcing of the faults (*e.g.* because of aseismic slip occurring in the vicinity). We note that elastic interactions are commonly invoked to explain time clustered events (*e.g.* Knopoff, 1964; Dieterich, 1994; Stein, 1999). Thus, assuming there exists a constant loading acting on the faults, we can expect systems with many interacting elements – dense fault networks or single faults with many asperities – to be able to produce strong short-term clustering whereas clustering in sparser networks takes place on longer time scales. With our observations we cannot differentiate between multiple faults or single faults with multiple asperities at the sub-template scale (*i.e.* for events detected with the same template).

Figure 9. Cross-section along the CIFALPS axis showing 976 templates that were well relocated ($\Delta r < 15$ km). **Top panel (A):** Number of detected earthquakes per template. **Bottom panel (B):** Sources with fractal dimension $D > 0.2$, *i.e.* sources exhibiting temporal clustering. The fractal dimension was calculated by taking the event count $e(t)$ of each template plus all the templates within a 10-km radius, over the whole study period. Even though intense seismic activity is located in the Ubaye valley, this seismicity is not associated with significant temporal clustering, showing that there is no systematic relation between temporal clustering and number of events per unit volume. The purple star indicates the location of the M_L3.9 earthquake that we mention in the discussion (Section 4). The red structures are reported from the geological cross-section in Figure 1.

Both regions where we observe significant temporal clustering, the Briançonnais and the Dora Maira massif, seem to share a common mechanism for clustering. Solarino et al. (2018) observed high V_p/V_s ratios (low V_s) in the uppermost part of the Briançonnais, where we observe high temporal clustering. They suggested that low shear wave velocities V_s could be explained by the widespread fault network observed in the Briançonnais (*e.g.* Tricart et al., 2004). In the Dora Maira massif, all the templates detecting seismic activity with fractal dimension $D > 0.3$ (see Figure 9B) are located around the M_L3.9 earthquake that occurred on October 3rd, 2012 (*cf.* location in Figures 6 and 9, *cf.* our catalog for the local magnitude). This highly clustered seismicity took place over about four days (see Figure 7A), and can be seen as a sequence of foreshocks-mainshock-aftershocks. The locations shown in Figure 9 are substantially spreaded, which suggests that seismicity is occurring on multiple faults. Given the limited temporal extent of the episode, we expect fault interactions to be a major contribution to temporal clustering in this area. Moreover, it is known that the Dora Maira massif is made of ultra-high pressure metamorphic rocks, *i.e.* of European crust subducted to 90 km depth and later exhumed (Chopin, 1984), it is very likely to be fractured. Thus, along with geological evidence, our observations of temporal clustering support the idea of fault interactions in dense fault networks as a driving mechanism for clustering in the Briançonnais and the Dora Maira massif.

Despite the high density of seismic sources beneath the Ubaye valley, temporal clustering is limited (only a few templates with $D \gtrsim 0.2$), which is an expected feature for seismic swarms. Thus, our measurements of temporal clustering suggest that the driving mechanism for seismicity in the Ubaye swarm differs from the one in the Briançonnais and the Dora Maira massif. Multiple studies (*e.g.* Daniel et al., 2011; Leclère et al., 2012; De Barros et al., 2019) emphasized the role of fluids in the stressing mechanism driving the seismicity of the Ubaye swarm. Furthermore, Ben-Zion and Lyakhovsky (2006) studied numerically the influence of damage rheology on the production of earthquakes. Their

model shows that cold, brittle media produce burst-like seismicity (high temporal clustering) whereas regions with high fluid activity produce more diffuse, swarm-like seismicity (low temporal clustering). Our observations of high seismic activity with low temporal clustering in the Ubaye swarm thus support the important role of fluid activity in this region. We realize, however, that such swarm-like behavior could also be the signature of aseismic processes (*e.g.* Lohman & McGuire, 2007). Whether aseismic slip is an important factor (Leclère et al., 2013) or not (De Barros et al., 2019) is still an ongoing debate, and our observations are not enough to support one scenario over the other. The clustered seismicity we detected in the Ubaye valley is consistent with the observations in De Barros et al. (2019) of coexisting aftershock sequences and swarm-like seismicity in this area. Studying a longer period of time, including the 2003-2004 and 2012-2015 Ubaye seismicity, could provide information on the stationarity of temporal clustering in the Ubaye valley and the rest of the Southwestern Alps.

5 Conclusion

In this paper we present a new method for automated earthquake detection and location, based on template matching and beamforming (or back projection), and use it for high (spatiotemporal) resolution characterization of seismicity in the Southwestern Alps. We address the problem of false positives in energy-based detection with signal classification based on supervised machine learning (Section 2.3), and we construct low noise templates by combining the singular value decomposition Wiener filter (SVDWF) with subsequent stacking (Section 2.5).

In our application to data from CIFALPS (Zhao et al., 2016), a semi-linear seismic network, and other permanent seismic stations in the Southwestern Alps, we detected in one year over an order of magnitude more events (18,754 vs. approximately 1,200) than an existing catalog based on traditional phase picking. We analyzed the statistical properties of the seismicity, and observed and characterized temporal earthquake clustering. We observed that regions of high seismic activity and high temporal clustering coincided with regions that are highly fractured (Briançonnais) or likely to be fractured (Dora Maira massif). Seismicity in the Dora Maira massif during the study period was dominated by the sequence of foreshocks and aftershocks associated with the 2012-10-03 M_L3.9 earthquake. We also identified one region of high seismic activity and low temporal clustering coinciding with the Ubaye swarm. Our results support interpretations invoking an important role of fluids in swarm seiscimity (Daniel et al., 2011; Leclère et al., 2012, 2013; De Barros et al., 2019).

The efficiency of this method increases when the database of templates gets more complete. Thus, processing longer times is likely to give better results as the opportunities of detecting new template events grow. The systematic application of this method to the Western Alps data, or even to the whole mountain range, will help gathering new observations of the seismicity and understanding the tectonic context of the region. We note that even though we presented an application to a semi-linear seismic network, our method can be applied to any network geometry. If 3D wave speed variations are sufficiently well known on the scale of study, it is possible to perform comprehensive studies of 3D seismicity structures by applying this method with 2D seismic arrays.

Data and Resources

The timings and locations of the 18,754 earthquakes we detected are available at E. Beaucé's personal website `https://ebeauce.github.io/` in the Material section. The reported times are the origin times, so that users can retrieve the P- and S-wave data by adding the origin times and the travel times, also provided in the catalog. Our codes are available at `https://github.com/ebeauce/earthquake_detection_EB_et_al_2019` (last accessed 08/16/2019), and are provided with a real-data example.

We created the map in Figure 6 using the topographic data from the Shuttle Radar Topographic Mission (SRTM) 90m database (`http://www.cgiar-csi.org/data/srtm-90m-digital-elevationdatabase-v4-1`, last accessed May 2019). Our data come from the temporary experiment CIFALPS (Zhao et al. (2016), DOI: `http://dx.doi.org/10.15778/RESIF.YP2012`) and permanent French (FR and RD RESIF (1995)) and Italian (GU University of Genova (1967), IV INGV Seismological Data Centre (2006), MN MedNet Project Partner Institutions (1990) and MT OGS (Istituto Nazionale di Oceanografia e di Geofisica Sperimentale) and University of Trieste (2002)) networks. The RENASS and INGV catalogs we mention in Section 3.1 can be obtained at `https://renass.unistra.fr/recherche` and `http://cnt.rm.ingv.it/en`, respectively.

Our study showing that the simplified definition of the correlation coefficient we use in this work is valid is available at `https://github.com/beridel/fast\_matched\_filter/blob/master/consequences\_nonzero.pdf`.

Acknowledgments

The authors thank Laurent Stehly and the SISmalp team for helping them evaluating the quality of their new earthquake catalog. They also thank Bertrand Potin for providing P- and S-wave travel times calculated from his velocity model. Finally, they thank the members of the CIFALPS team, in particular Coralie Aubert, Stefano Solarino, and Simone Salimbeni for their excellent work in data acquisition in the field. CIFALPS was funded by IGGCAS (Institute of Geology and Geophysics, China Academy of Sciences, Beijing, P.I. Liang Zhao) and Labex OSUG@2020 (France). E. B. was supported by funds associated with RvdHs Schlumberger chair and by the European Research Council (ERC) under the European Unions Horizon 2020 Research and Innovation Program (grant agreement N° 789742335,F-IMAGE).

The reader is referred to the Data and Resources section to find the origin of the data.

Appendix A Template Relocation

The weighting function presented in Figure 5 is defined by:

$$w(t_n) = \begin{cases} A \exp\left(-\dfrac{(\mathrm{CNR}(t_n) - \mathrm{CNR}_{\max})^2}{4\sigma}\right) & \text{if } t_n \in \mathcal{V}(t_{\max}), \\ 0 & \text{otherwise.} \end{cases} \tag{A1}$$

In Equation A1, the neighborhood $\mathcal{V}(t_{\max})$ is defined by:

$$\mathcal{V}(t^*) = \left\{ t^- \leq t_k \leq t^+ \mid [t^-, t^+] \text{ is a convex set, } t_{\max} \in [t^-, t^+], \mathrm{CNR}(t_k) > 0.75 \times \mathrm{CNR}_{\max} \right\}, \tag{A2}$$

and $t_{\max} = \underset{t_n}{\mathrm{argmax}}\,(\mathrm{CNR})$. A is a normalization factor such that $\sum_{n=1}^{N} w(t_n) = 1$, and σ is the standard deviation of the CNR within $\mathcal{V}(t_{\max})$.

Using the locations of the potential sources from the composite network response, we calculate the average distance to the best test source:

$$\Delta r = \sum_{n=1}^{N} w(t_n) |r_n - r_{\mathrm{best}}|. \tag{A3}$$

In Equation A3, N is the temporal length of the stacked waveforms, r_n is the potential source location associated with the CNR at time t_n and r_{best} is the location of

the potential source associated at time $t_{\max}$, *i.e.* the location of the second generation template.

Appendix B Magnitude Estimation

Our local magnitude is calculated from the amplitude ratio of the peak velocities with a reference event. Thus, we first need to estimate the magnitude of at least one event per template to calibrate our local magnitude scale. For each family of earthquakes detected with the same template, we proceed as follows:

1. calculate the S-wave spectrum on every station and component,
2. calculate the noise spectrum in a window taken just before the P-wave arrival,
3. average the spectra over all the stations and components, including only the samples satisfying the SNR criterion (similarly to Uchide & Imanishi, 2016), according to:

$$\bar{S}(f) = \frac{1}{\sum_{s,c} \mathbf{1}_{\mathrm{SNR}>5}\left[S_{s,c}(f)\right]} \sum_{s,c} \alpha_{s,c} S_{s,c}(f) \mathbf{1}_{\mathrm{SNR}>5}\left[S_{s,c}(f)\right]. \tag{B1}$$

   In Equation B1, $\mathbf{1}_{\mathrm{SNR}>5}\left[S(f)\right]$ is the indicator function testing whether $S(f)$ has SNR greater than 5 (equal to 1) or not (equal to 0). The SNR is calculated at every frequency by taking the ratio of the S-wave spectrum to the noise spectrum. $\alpha_{s,c}$ is a corrective factor that we describe further.
4. The average spectra are converted to displacement spectra by using the relationship

$$|u_{\mathrm{velocity}}(f)| = f \times |u_{\mathrm{displacement}}(f)|, \tag{B2}$$

5. the average displacement spectra are fitted with the Boatwright model (Boatwright, 1978):

$$S_{\mathrm{Boatwright}}(f) = \frac{\Omega_0}{\left(1 + \left(\frac{f}{f_c}\right)^4\right)^{1/2}}, \tag{B3}$$

   where Ω_0 is the low-frequency plateau, related to the seismic moment, and f_c is the corner frequency.
   The corrective factors $\alpha_{s,c}$ are defined such that the low-frequency plateau Ω_0 can be identified to the seismic moment M_0. Assuming a double-couple source, a displacement amplitude spectrum can be written as (following Boatwright, 1978):

$$\begin{aligned} |u^S(f)| &= \frac{R^S}{2\rho\beta^3 r} \frac{M_0}{\left(1 + \left(\frac{f}{f_c^S}\right)^4\right)^{1/2}} \exp\left(-\frac{\pi f t^S}{Q^S}\right), \\ \implies M_0 &= \Omega_0 \frac{2\rho\beta^3 r}{R^S} \exp\left(\frac{\pi f t^S}{Q^S}\right), \\ \implies \alpha_{s,c} &= \frac{2\rho\beta^3 r_{s,c}}{R^S} \exp\left(\frac{\pi f t^S_{s,c}}{Q^S_{s,c}}\right). \end{aligned} \tag{B4}$$

In Equation B4, we use typical values for the S-wave velocity β (3000 km/s), the density ρ (2700 kg/m^3) and the average S-wave radiation pattern R^S ($\sqrt{2/5}$ from Aki & Richards, 2002). The seismic moment M_0 gives the magnitude moment M_w through:

$$M_w = \frac{2}{3}\left(\log M_0 - 9.1\right). \tag{B5}$$

The reference events are those for which fitting a Boatwright model to the average spectrum results in a variance reduction greater than 0.95. Figure B1 shows an example of an average spectrum that was fitted correctly, and therefore kept as a reference event. The local magnitude of all the other events are determined by:

$$M_i = M_{\text{ref}} + \underset{s,c}{\text{Median}} \left\{ \log \frac{A^i_{s,c}}{A^{\text{ref}}_{s,c}} \right\}, \tag{B6}$$

or more generally if there are several reference events:

$$M_i = \underset{k}{\text{Median}} \left\{ M_{\text{ref},k} + \underset{s,c}{\text{Median}} \left\{ \log \frac{A^i_{s,c}}{A^{\text{ref},k}_{s,c}} \right\} \right\}. \tag{B7}$$

Figure B1. Magnitude estimation of the reference event. For each template, we use the highest SNR detections to calculate the average S-wave spectrum (Equation B1) and fit it with the Boatwright model (Equation B3). The low-frequency plateau gives us the seismic moment M_0. The average is calculated over all the stations and components that satisfy the SNR criterion. Thus, for each frequency sample the number of channels included in the average may vary, as we can see with the color scale. Since frequency samples with a higher number of channels are more reliable, we give them larger weight in the inversion.

References

Abercrombie, R. E. (1995). Earthquake source scaling relationships from- 1 to 5 ml using seismograms recorded at 2.5-km depth. *Journal of Geophysical Research: Solid Earth*, *100*(B12), 24015–24036.

Aki, K., & Richards, P. G. (2002). *Quantitative seismology.*

Allen, R. (1982). Automatic phase pickers: their present use and future prospects. *Bulletin of the Seismological Society of America*, *72*(6B), S225–S242.

Baillard, C., Crawford, W. C., Ballu, V., Hibert, C., & Mangeney, A. (2014). An automatic kurtosis-based p-and s-phase picker designed for local seismic networks. *Bulletin of the Seismological Society of America*, *104*(1), 394–409.

Baratin, L.-M., Chamberlain, C. J., Townend, J., & Savage, M. K. (2018). Focal mechanisms and inter-event times of low-frequency earthquakes reveal quasi-continuous deformation and triggered slow slip on the deep alpine fault. *Earth and Planetary Science Letters*, *484*, 111–123.

Beaucé, E., Frank, W. B., & Romanenko, A. (2017). Fast Matched Filter (FMF): An efficient seismic matched-filter search for both cpu and gpu architectures. *Seismological Research Letters*, *89*(1), 165–172.

Ben-Zion, Y., & Lyakhovsky, V. (2006). Analysis of aftershocks in a lithospheric model with seismogenic zone governed by damage rheology. *Geophysical Journal International*, *165*(1), 197–210.

Boatwright, J. (1978). Detailed spectral analysis of two small new york state earthquakes. *Bulletin of the Seismological Society of America*, *68*(4), 1117–1131.

Brown, J. R., Beroza, G. C., & Shelly, D. R. (2008). An autocorrelation method to detect low frequency earthquakes within tremor. *Geophysical Research Letters*, *35*(16).

Burridge, R., & Knopoff, L. (1967). Model and theoretical seismicity. *Bulletin of the seismological society of america*, *57*(3), 341–371.

Chollet, F., et al. (2015). *Keras.* `https://keras.io`.

Chopin, C. (1984). Coesite and pure pyrope in high-grade blueschists of the western alps: a first record and some consequences. *Contributions to Mineralogy and Petrology*, *86*(2), 107–118.

Daniel, G., Prono, E., Renard, F., Thouvenot, F., Hainzl, S., Marsan, D., ... others (2011). Changes in effective stress during the 2003–2004 ubaye seismic swarm, france. *Journal of Geophysical Research: Solid Earth*, *116*(B1).

De Barros, L., Baques, M., Godano, M., Helmstetter, A., Deschamps, A., Larroque, C., & Courboulex, F. (2019). Fluid-induced swarms and coseismic stress transfer: a dual process highlighted in the aftershock sequence of the 7 april 2014 earthquake (ml 4.8, ubaye, france). *Journal of Geophysical Research: Solid Earth*.

Delacou, B., Sue, C., Champagnac, J.-D., & Burkhard, M. (2004). Present-day geodynamics in the bend of the western and central alps as constrained by earthquake analysis. *Geophysical Journal International*, *158*(2), 753–774.

Dieterich, J. (1992). Earthquake nucleation on faults with rate-and state-dependent strength. *Tectonophysics*, *211*(1-4), 115–134.

Dieterich, J. (1994). A constitutive law for rate of earthquake production and its application to earthquake clustering. *Journal of Geophysical Research: Solid Earth*, *99*(B2), 2601–2618.

Dziewonski, A. M., & Anderson, D. L. (1981). Preliminary reference earth model. *Physics of the earth and planetary interiors*, *25*(4), 297–356.

Dziewonski, A. M., & Woodhouse, J. H. (1987). Global images of the earth's interior. *Science*, *236*(4797), 37–48.

Engdahl, E. R., van der Hilst, R., & Buland, R. (1998). Global teleseismic earthquake relocation with improved travel times and procedures for depth determination. *Bulletin of the Seismological Society of America*, *88*(3), 722–743.

Frank, W. B., & Shapiro, N. M. (2014). Automatic detection of low-frequency earthquakes (LFEs) based on a beamformed network response. *Geophysical Journal International*, *197*(2), 1215–1223.

Frank, W. B., Shapiro, N. M., & Gusev, A. A. (2018). Progressive reactivation of the volcanic plumbing system beneath tolbachik volcano (kamchatka, russia) revealed by long-period seismicity. *Earth and Planetary Science Letters*, *493*, 47–56.

Frank, W. B., Shapiro, N. M., Husker, A. L., Kostoglodov, V., Gusev, A. A., & Campillo, M. (2016). The evolving interaction of low-frequency earthquakes during transient slip. *Science Advances*, *2*(4). Retrieved from `http://advances.sciencemag.org/content/2/4/e1501616` doi:

10.1126/sciadv.1501616

Frank, W. B., Shapiro, N. M., Husker, A. L., Kostoglodov, V., Romanenko, A., & Campillo, M. (2014). Using systematically characterized low-frequency earthquakes as a fault probe in Guerrero, Mexico. *Journal of Geophysical Research: Solid Earth*, *119*(10), 7686–7700.

Gardner, J., & Knopoff, L. (1974). Is the sequence of earthquakes in southern california, with aftershocks removed, poissonian? *Bulletin of the Seismological Society of America*, *64*(5), 1363–1367.

Gibbons, S. J., & Ringdal, F. (2006). The detection of low magnitude seismic events using array-based waveform correlation. *Geophysical Journal International*, *165*(1), 149–166.

Honda, R., & Aoi, S. (2009). Array back-projection imaging of the 2007 niigataken chuetsu-oki earthquake striking the worlds largest nuclear power plant. *Bulletin of the Seismological Society of America*, *99*(1), 141–147.

Ide, S., Beroza, G. C., Prejean, S. G., & Ellsworth, W. L. (2003). Apparent break in earthquake scaling due to path and site effects on deep borehole recordings. *Journal of Geophysical Research: Solid Earth*, *108*(B5).

INGV Seismological Data Centre. (2006). Rete Sismica Nazionale(RSN), Istituto Nazionale di Geofisica e Vulcanologia (INGV), Italy.
doi: 10.13127/SD/X0FXnH7QfY

Ishii, M., Shearer, P. M., Houston, H., & Vidale, J. E. (2005). Extent, duration and speed of the 2004 sumatra–andaman earthquake imaged by the hi-net array. *Nature*, *435*(7044), 933.

Jenatton, L., Guiguet, R., Thouvenot, F., & Daix, N. (2007). The 16,000-event 2003–2004 earthquake swarm in ubaye (french alps). *Journal of Geophysical Research: Solid Earth*, *112*(B11).

Kennett, B., & Engdahl, E. (1991). Traveltimes for global earthquake location and phase identification. *Geophysical Journal International*, *105*(2), 429–465.

Kennett, B. L., Engdahl, E., & Buland, R. (1995). Constraints on seismic velocities in the earth from traveltimes. *Geophysical Journal International*, *122*(1), 108–124.

Knopoff, L. (1964). The statistics of earthquakes in southern california. *Bulletin of the Seismological Society of America*, *54*(6A), 1871–1873.

Leclère, H., Daniel, G., Fabbri, O., Cappa, F., & Thouvenot, F. (2013). Tracking fluid pressure buildup from focal mechanisms during the 2003–2004 ubaye seismic swarm, france. *Journal of Geophysical Research: Solid Earth*, *118*(8), 4461–4476.

Leclère, H., Fabbri, O., Daniel, G., & Cappa, F. (2012). Reactivation of a strike-slip fault by fluid overpressuring in the southwestern french-italian alps. *Geophysical Journal International*, *189*(1), 29–37.

Lee, H.-K., & Schwarcz, H. P. (1995). Fractal clustering of fault activity in california. *Geology*, *23*(4), 377–380.

Li, C., Van Der Hilst, R. D., Engdahl, E. R., & Burdick, S. (2008). A new global model for p wave speed variations in earth's mantle. *Geochemistry, Geophysics, Geosystems*, *9*(5).

Lohman, R. B., & McGuire, J. J. (2007). Earthquake swarms driven by aseismic creep in the salton trough, california. *Journal of Geophysical Research: Solid Earth*, *112*(B4).

Lowen, S. B., & Teich, M. C. (2005). *Fractal-based point processes* (Vol. 366). John Wiley & Sons.

MedNet Project Partner Institutions. (1990). Mediterranean Very BroadbandSeismographic Network (MedNet). Istituto Nazionale di Geofisica e Vulcanologia (INGV), Italy.
doi: 10.13127/sd/fbbbtdtd6q

Meng, H., & Ben-Zion, Y. (2017). Detection of small earthquakes with dense array data: example from the san jacinto fault zone, southern california. *Geophysical Journal International*, *212*(1), 442–457.

Moreau, L., Stehly, L., Boué, P., Lu, Y., Larose, E., & Campillo, M. (2017). Improving ambient noise correlation functions with an svd-based wiener filter. *Geophysical Journal International*, *211*(1), 418–426.

Nadeau, R. M., Foxall, W., & McEvilly, T. (1995). Clustering and periodic recurrence of microearthquakes on the san andreas fault at parkfield, california. *Science*, *267*(5197), 503–507.

Nicolas, A., Hirn, A., Nicolich, R., & Polino, R. (1990). Lithospheric wedging in the western alps inferred from the ecors-crop traverse. *Geology*, *18*(7), 587–590.

Nocquet, J.-M. (2012). Present-day kinematics of the mediterranean: A comprehensive overview of gps results. *Tectonophysics*, *579*, 220–242.

Nocquet, J.-M., Sue, C., Walpersdorf, A., Tran, T., Lenôtre, N., Vernant, P., ... others (2016). Present-day uplift of the western alps. *Scientific reports*, *6*, 28404.

Obara, K. (2002). Nonvolcanic deep tremor associated with subduction in southwest japan. *Science*, *296*(5573), 1679–1681.

OGS (Istituto Nazionale di Oceanografia e di Geofisica Sperimentale) and University of Trieste. (2002). North East Italy Broadband Network. International Federation of Digital Seismograph Networks, Other/SeismicNetwork. doi: 10.7914/SN/NI

on California Earthquake Probabilities, W. G. (1995). Seismic hazards in southern california: Probable earthquakes, 1994 to 2024. *Bulletin of the Seismological Society of America*, *85*(2), 379–439.

Perol, T., Gharbi, M., & Denolle, M. (2018). Convolutional neural network for earthquake detection and location. *Science Advances*, *4*(2), e1700578.

Potin, B. (2016). *The Western Alps: Tomography, earthquake location and moho topography* (Unpublished doctoral dissertation). Universit Grenoble Alpes.

RESIF. (1995). RESIF-RLBP French Broad-band network, RESIF-RAPstrong motion network and other seismic stations in metropoli-tan France. RESIF - Réseau Sismologique et géodésique Français. doi: 10.15778/resif.fr

Richards-Dinger, K., & Dieterich, J. H. (2012). Rsqsim earthquake simulator. *Seismological Research Letters*, *83*(6), 983–990.

Rogers, G., & Dragert, H. (2003). Episodic tremor and slip on the cascadia subduction zone: The chatter of silent slip. *Science*, *300*(5627), 1942–1943.

Ross, Z. E., Trugman, D. T., Hauksson, E., & Shearer, P. M. (2019). Searching for hidden earthquakes in southern california. *Science*, eaaw6888.

Serpelloni, E., Vannucci, G., Pondrelli, S., Argnani, A., Casula, G., Anzidei, M., ... Gasperini, P. (2007). Kinematics of the western africa-eurasia plate boundary from focal mechanisms and gps data. *Geophysical Journal International*, *169*(3), 1180–1200.

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

Shelly, D. R., Peng, Z., Hill, D. P., & Aiken, C. (2011). Triggered creep as a possible mechanism for delayed dynamic triggering of tremor and earthquakes. *Nature Geoscience*, *4*(6), 384.

Smalley Jr, R., Chatelain, J.-L., Turcotte, D., & Prévot, R. (1987). A fractal approach to the clustering of earthquakes: applications to the seismicity of the new hebrides. *Bulletin of the Seismological Society of America*, *77*(4), 1368–1381.

Solarino, S., Malusà, M. G., Eva, E., Guillot, S., Paul, A., Schwartz, S., ... others (2018). Mantle wedge exhumation beneath the dora-maira (u) hp dome unravelled by local earthquake tomography (western alps). *Lithos*, *296*, 623–636.

Stein, R. S. (1999). The role of stress transfer in earthquake occurrence. *Nature*, *402*(6762), 605.

Thouvenot, F., Jenatton, L., Scafidi, D., Turino, C., Potin, B., & Ferretti, G. (2016). Encore ubaye: earthquake swarms, foreshocks, and aftershocks in the southern french alps. *Bulletin of the Seismological Society of America*, *106*(5), 2244–2257.

Tricart, P., Schwartz, S., Sue, C., & Lardeaux, J. (2004). Differential exhumation in the inner western alpine arc evidenced by late normal faulting (eastern queyras schistes lustrés). *Journal of Structural Geology*, *26*, 1633–1645.

Turcotte, D. L. (1997). *Fractals and chaos in geology and geophysics*. Cambridge university press.

Uchide, T., & Imanishi, K. (2016). Small earthquakes deviate from the omega-square model as revealed by multiple spectral ratio analysis. *Bulletin of the Seismological Society of America*, *106*(3), 1357–1363.

University of Genova. (1967). Regional Seismic Network of North Western Italy. International Federation of Digital Seismograp Networks, Other Seismic Network.
doi: 10.7914/SN/GU

Utsu, T. (1961). A statistical study on the occurrence of aftershocks. *Geophys. Mag.*, *30*, 521–605.

Van der Hilst, R. D., Widiyantoro, S., & Engdahl, E. (1997). Evidence for deep mantle circulation from global tomography. *Nature*, *386*(6625), 578.

Walker, K. T., Ishii, M., & Shearer, P. M. (2005). Rupture details of the 28 march 2005 sumatra mw 8.6 earthquake imaged with teleseismic p waves. *Geophysical Research Letters*, *32*(24).

Walpersdorf, A., Pinget, L., Vernant, P., Sue, C., Deprez, A., & team, R. (2018). Does long-term gps in the western alps finally confirm earthquake mechanisms? *Tectonics*, *37*(10), 3721–3737.

Wech, A. G., & Creager, K. C. (2011). A continuum of stress, strength and slip in the cascadia subduction zone. *Nature Geoscience*, *4*(9), 624.

Yao, Y., Rosasco, L., & Caponnetto, A. (2007). On early stopping in gradient descent learning. *Constructive Approximation*, *26*(2), 289–315.

Zaliapin, I., & Ben-Zion, Y. (2013a). Earthquake clusters in southern california II: Classification and relation to physical properties of the crust. *Journal of Geophysical Research: Solid Earth*, *118*(6), 2865–2877.

Zaliapin, I., & Ben-Zion, Y. (2013b). Earthquake clusters in southern california I: Identification and stability. *Journal of Geophysical Research: Solid Earth*, *118*(6), 2847–2864.

Zaliapin, I., Gabrielov, A., Keilis-Borok, V., & Wong, H. (2008). Clustering analysis of seismicity and aftershock identification. *Physical review letters*, *101*(1), 018501.

Zhao, L., Paul, A., Guillot, S., Solarino, S., Malusà, M. G., Zheng, T., ... others (2015). First seismic evidence for continental subduction beneath the Western Alps. *Geology*, *43*(9), 815–818.

Zhao, L., Paul, A., & Solarino, S. (2016). Seismic network YP: CIFALPS temporary experiment (China-Italy-France Alps seismic transect); RESIFRéseau Sismologique et géodésique Français.
doi: http://dx.doi.org/10.15778/RESIF.YP2012

Figure 1.

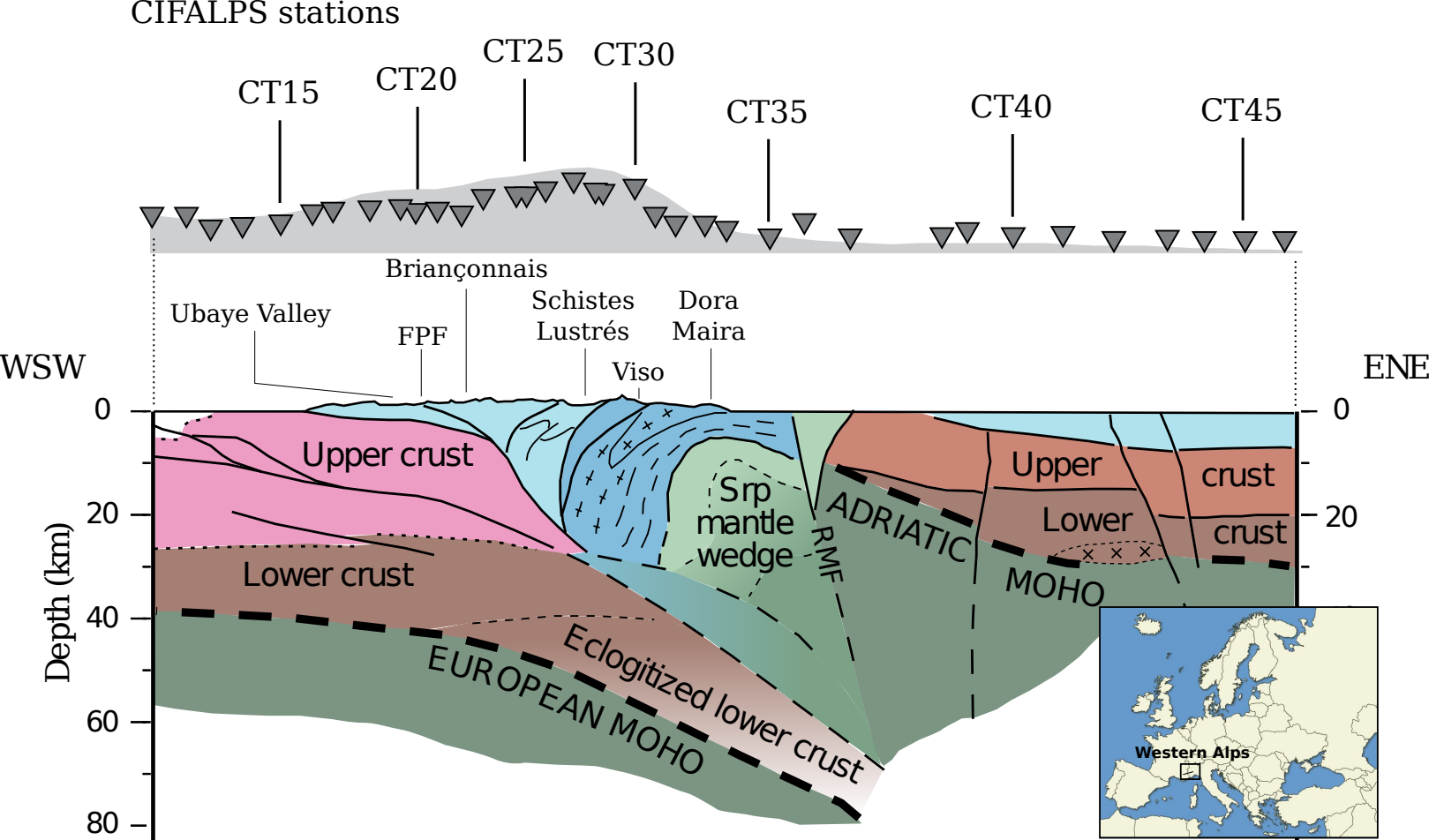

CIFALPS stations
CT15
CT20
CT25
CT30
CT35
CT40
CT45
Briançonnais
Ubaye Valley
FPF
Schistes
Lustrés
Dora
Maira
Viso
WSW
ENE
0
20
40
60
80
Depth (km)
Upper crust
Lower crust
Srp
mantle
wedge
RMF
ADRIATIC
MOHO
Upper
crust
Lower
crust
Eclogitized lower crust
EUROPEAN MOHO
Western Alps

Figure 2.

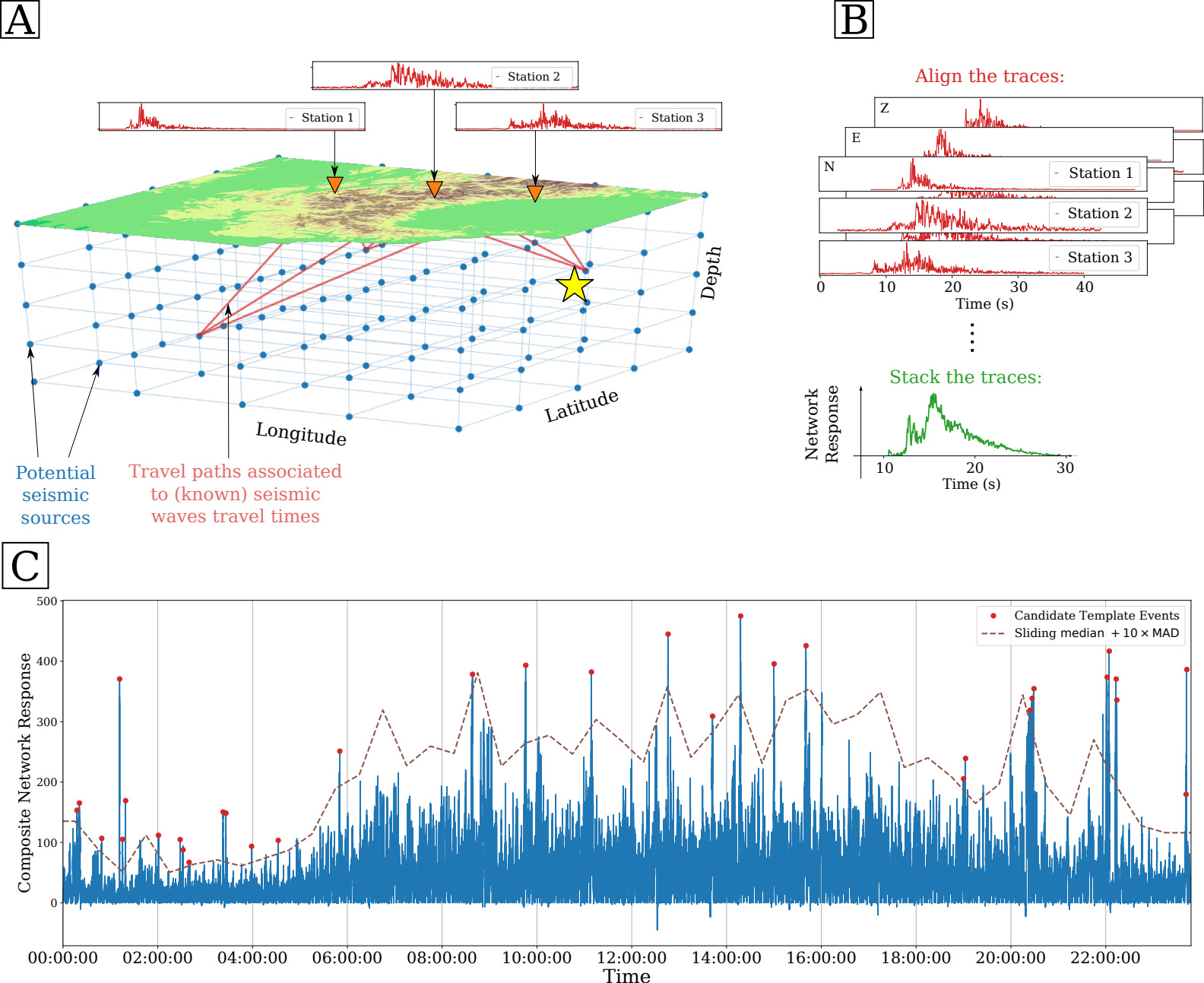

A
Station 1
Station 2
Station 3
Depth
Longitude
Latitude
Potential seismic sources
Travel paths associated to (known) seismic waves travel times
B
Align the traces:
Z
E
N
Station 1
Station 2
Station 3
0
10
20
30
40
Time (s)
Stack the traces:
Network Response
10
20
30
Time (s)
C
Candidate Template Events
Sliding median + 10 × MAD
Composite Network Response
500
400
300
200
100
0
00:00:00
02:00:00
04:00:00
06:00:00
08:00:00
10:00:00
12:00:00
14:00:00
16:00:00
18:00:00
20:00:00
22:00:00
Time

Figure 3.

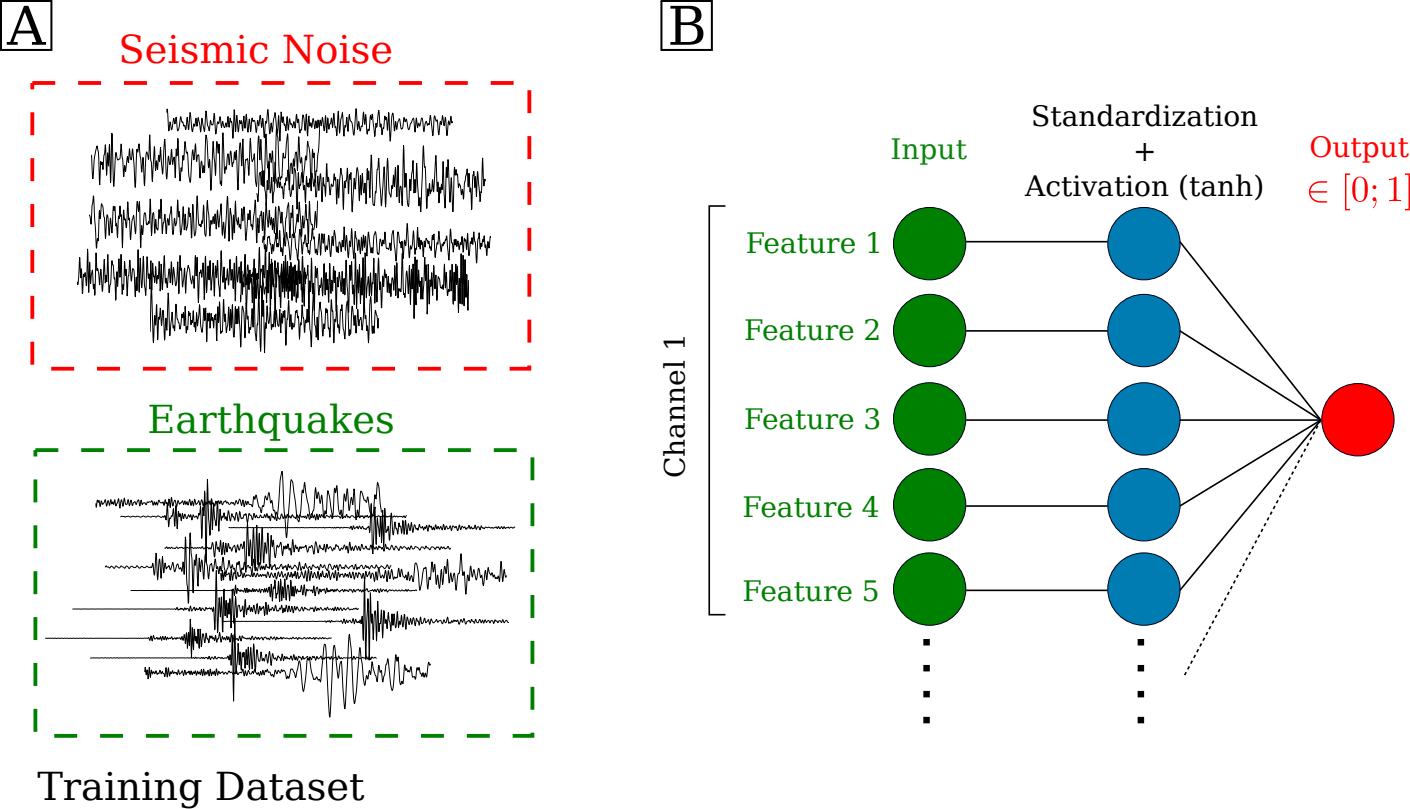

A
Seismic Noise
Earthquakes
Training Dataset
B
Input
Standardization
+
Activation (tanh)
Output
$\in [0;1]$
Channel 1
Feature 1
Feature 2
Feature 3
Feature 4
Feature 5

Figure 4.

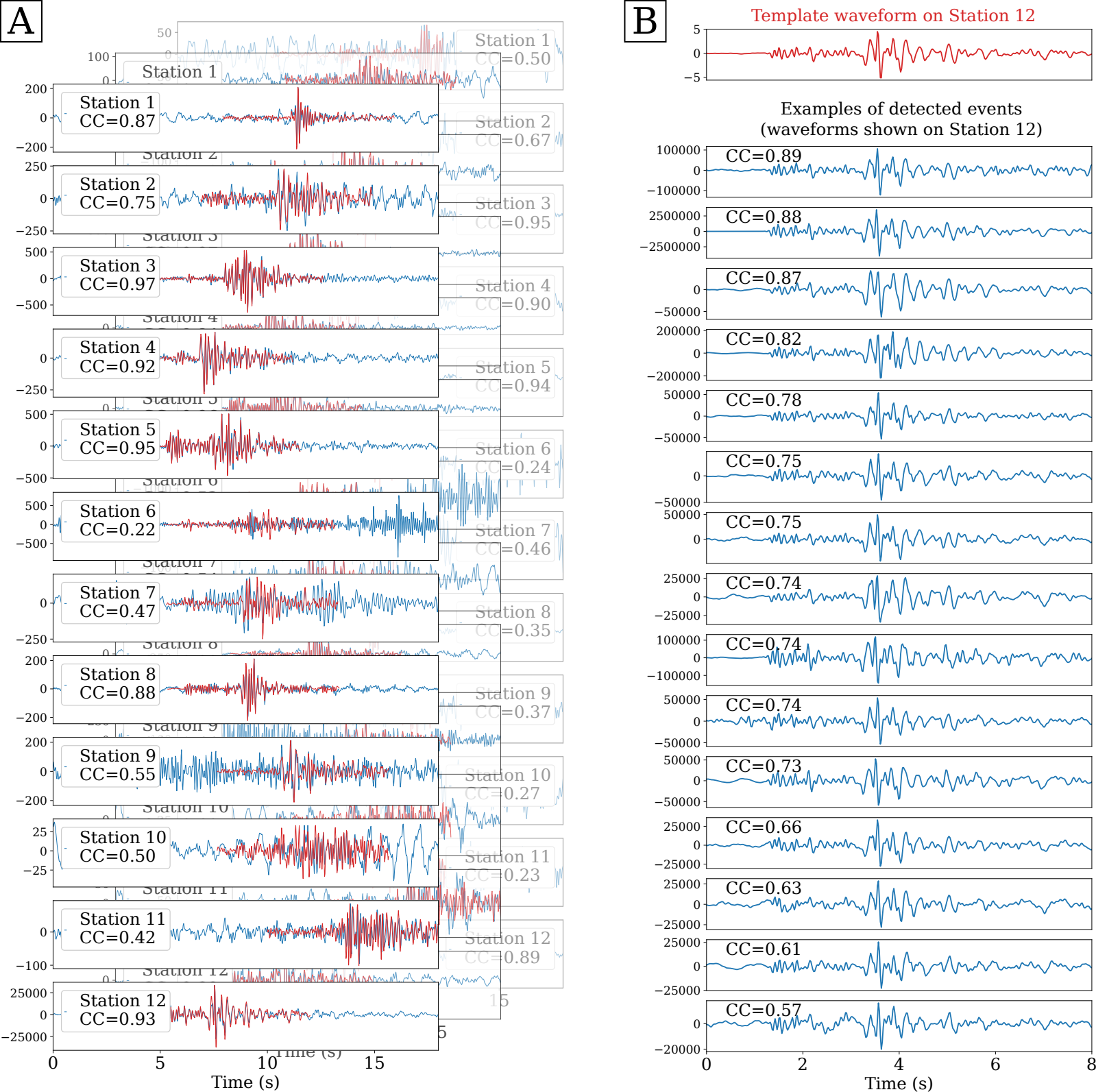

A
Station 1
CC=0.87
Station 2
CC=0.75
Station 3
CC=0.97
Station 4
CC=0.92
Station 5
CC=0.95
Station 6
CC=0.22
Station 7
CC=0.47
Station 8
CC=0.88
Station 9
CC=0.55
Station 10
CC=0.50
Station 11
CC=0.42
Station 12
CC=0.93
Station 1
CC=0.50
Station 2
CC=0.67
Station 3
CC=0.95
Station 4
CC=0.90
Station 5
CC=0.94
Station 6
CC=0.24
Station 7
CC=0.46
Station 8
CC=0.35
Station 9
CC=0.37
Station 10
CC=0.27
Station 11
CC=0.23
Station 12
CC=0.89
Time (s)
B
Template waveform on Station 12
Examples of detected events
(waveforms shown on Station 12)
CC=0.89
CC=0.88
CC=0.87
CC=0.82
CC=0.78
CC=0.75
CC=0.75
CC=0.74
CC=0.74
CC=0.74
CC=0.73
CC=0.66
CC=0.63
CC=0.61
CC=0.57
Time (s)

Figure 5.

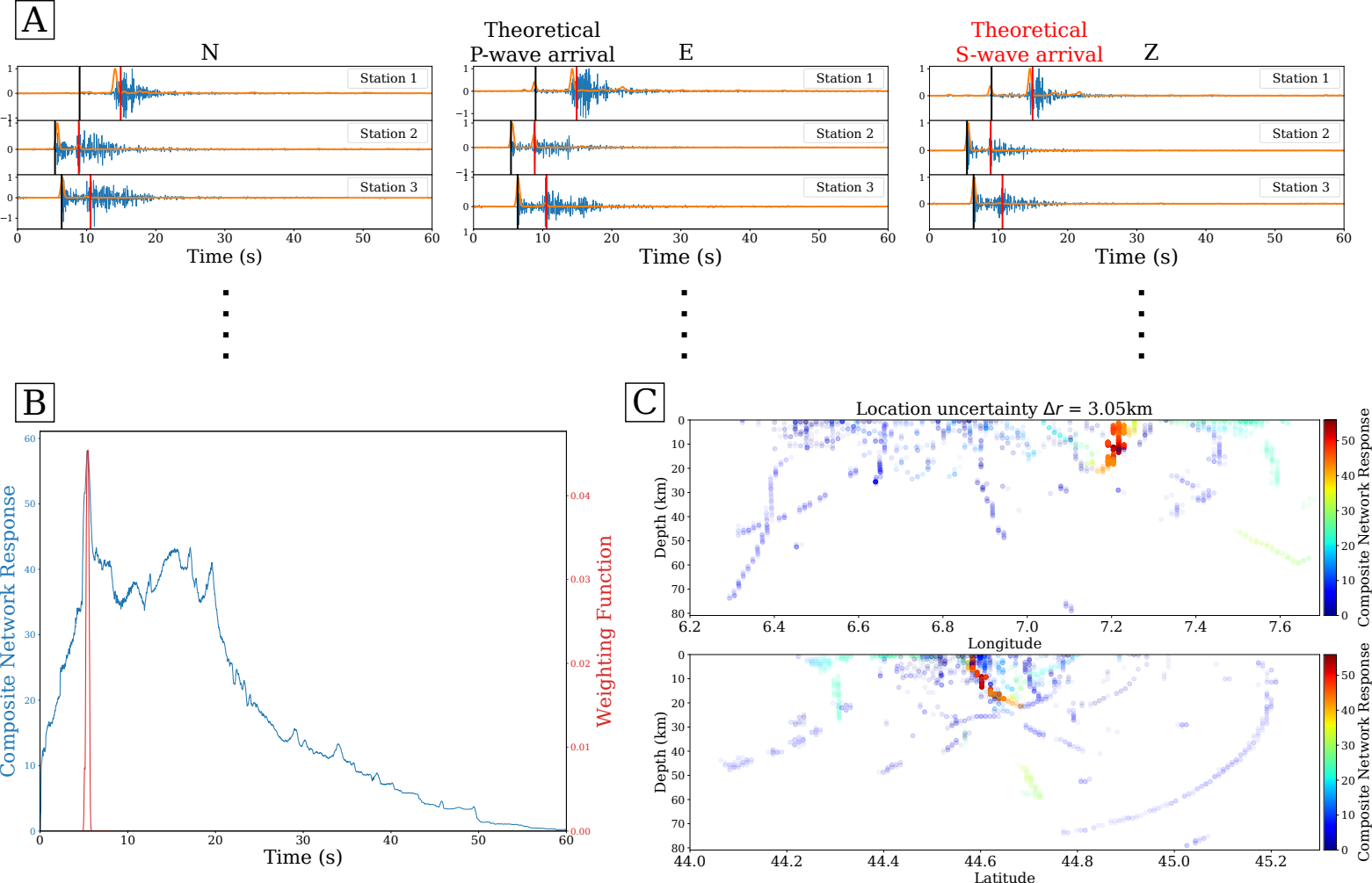

A
N
Theoretical
P-wave arrival
E
Theoretical
S-wave arrival
Z
Station 1
Station 2
Station 3
Time (s)
B
Composite Network Response
Weighting Function
Time (s)
C
Location uncertainty Δr = 3.05km
Depth (km)
Longitude
Latitude
Composite Network Response

Figure 6.

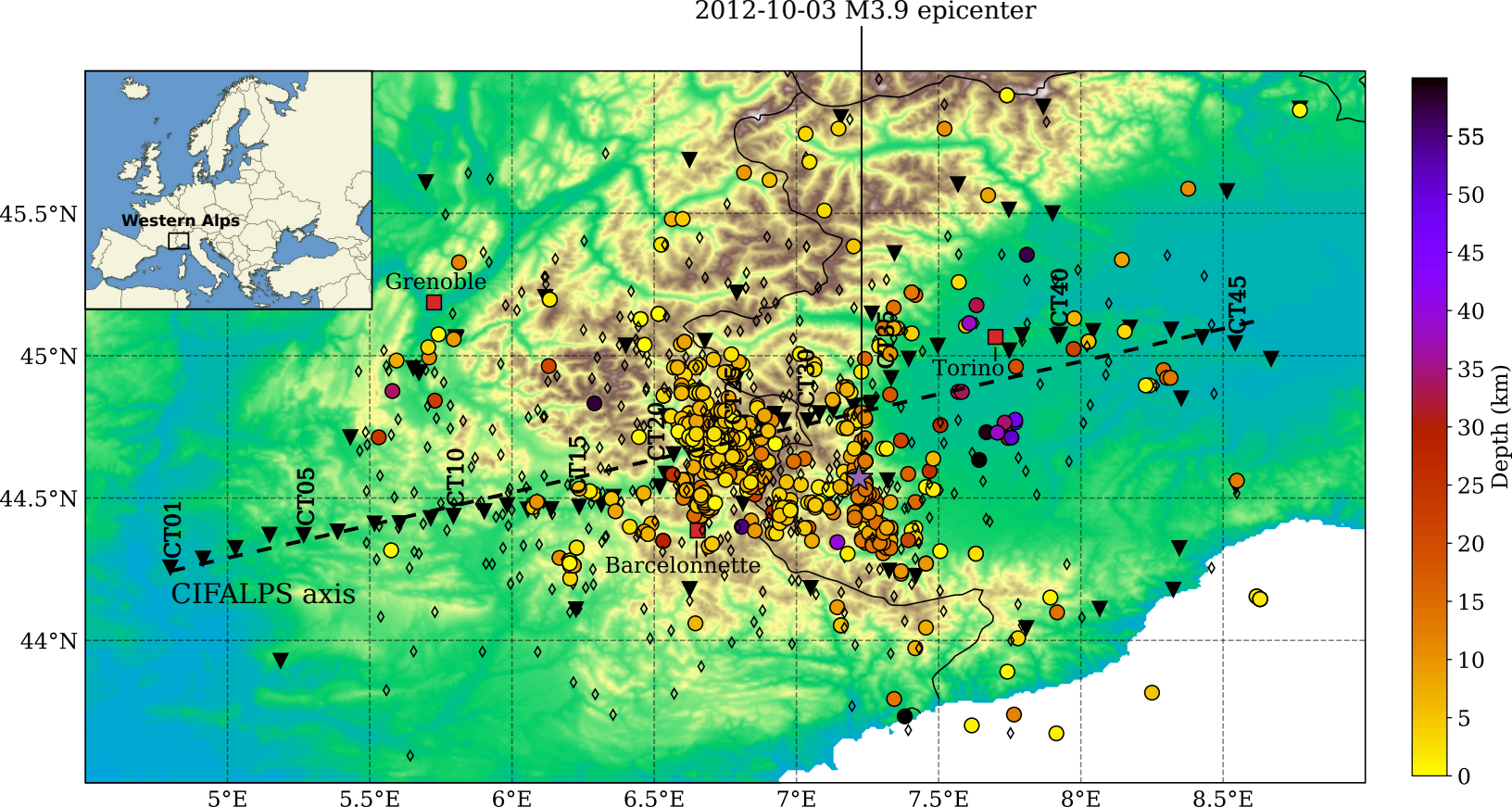

2012-10-03 M3.9 epicenter
Western Alps
Grenoble
Torino
Barcelonnette
CIFALPS axis
CT01
CT05
CT10
CT15
CT20
CT30
CT35
CT40
CT45
45.5°N
45°N
44.5°N
44°N
5°E
5.5°E
6°E
6.5°E
7°E
7.5°E
8°E
8.5°E
Depth (km)
0
5
10
15
20
25
30
35
40
45
50
55

Figure 7.

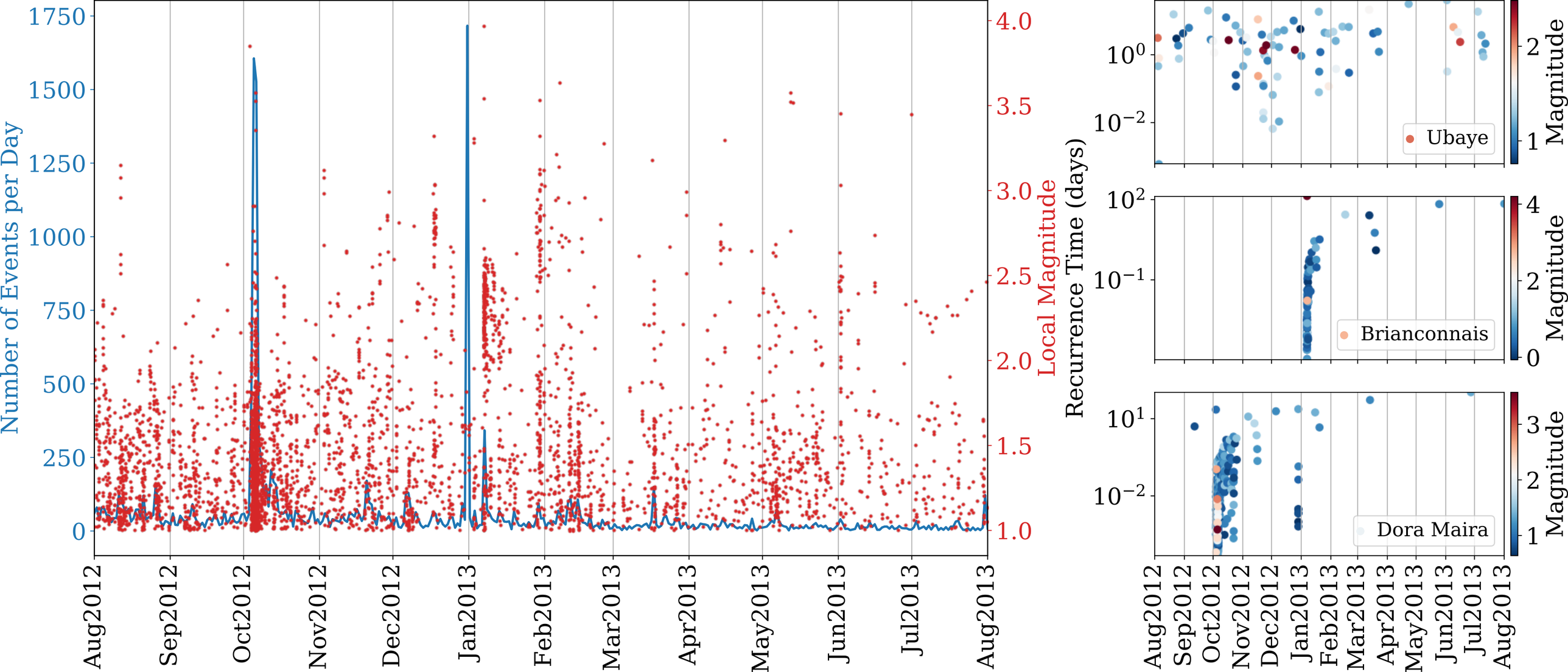

Number of Events per Day
Local Magnitude
Recurrence Time (days)
Magnitude
Ubaye
Brianconnais
Dora Maira
Aug2012
Sep2012
Oct2012
Nov2012
Dec2012
Jan2013
Feb2013
Mar2013
Apr2013
May2013
Jun2013
Jul2013
Aug2013

Figure 8.

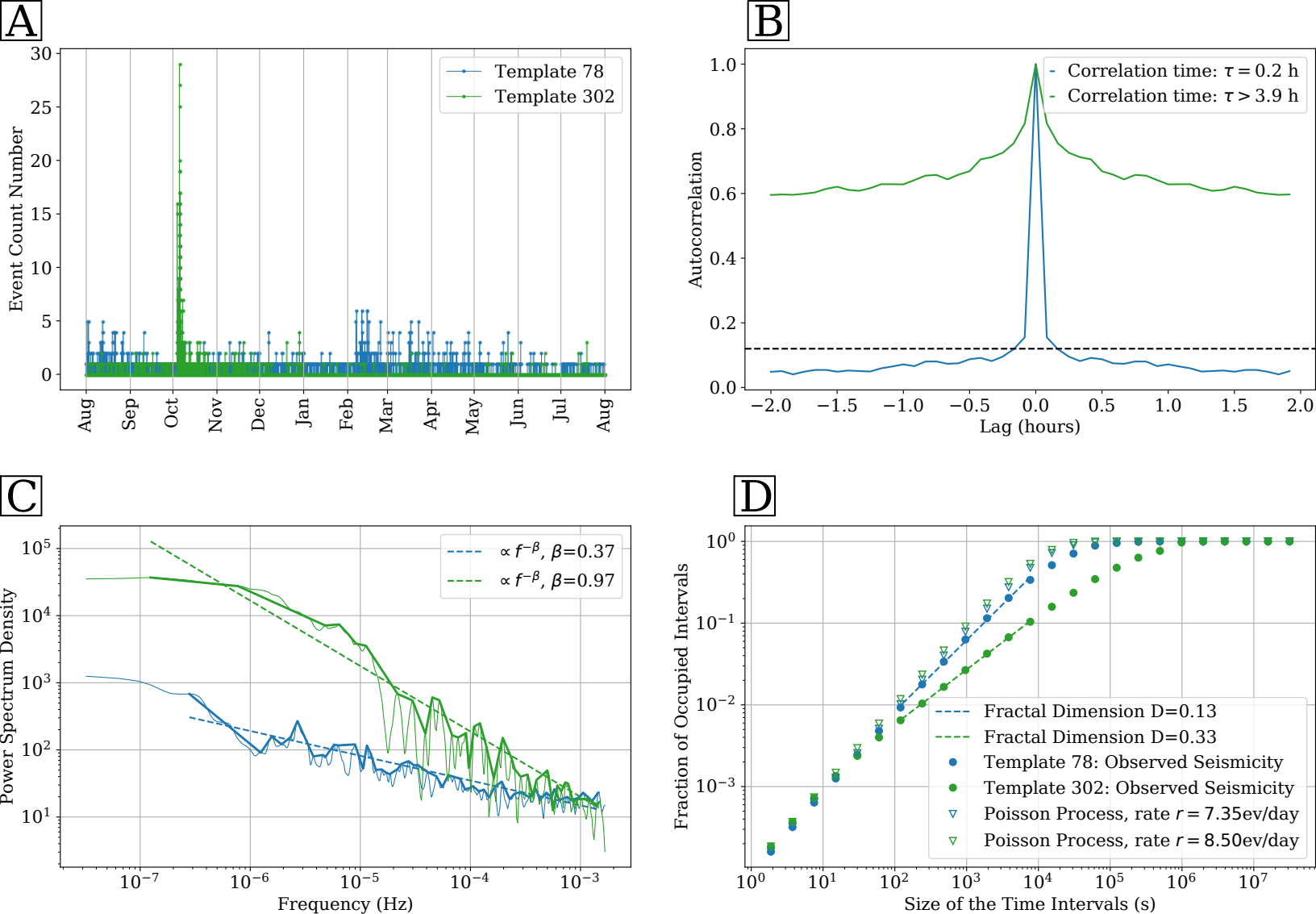

A
Event Count Number
Template 78
Template 302
Aug
Sep
Oct
Nov
Dec
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
B
Autocorrelation
Correlation time: τ = 0.2 h
Correlation time: τ > 3.9 h
Lag (hours)
C
Power Spectrum Density
∝ f^−β, β=0.37
∝ f^−β, β=0.97
Frequency (Hz)
D
Fraction of Occupied Intervals
Fractal Dimension D=0.13
Fractal Dimension D=0.33
Template 78: Observed Seismicity
Template 302: Observed Seismicity
Poisson Process, rate r = 7.35ev/day
Poisson Process, rate r = 8.50ev/day
Size of the Time Intervals (s)

Figure 9.

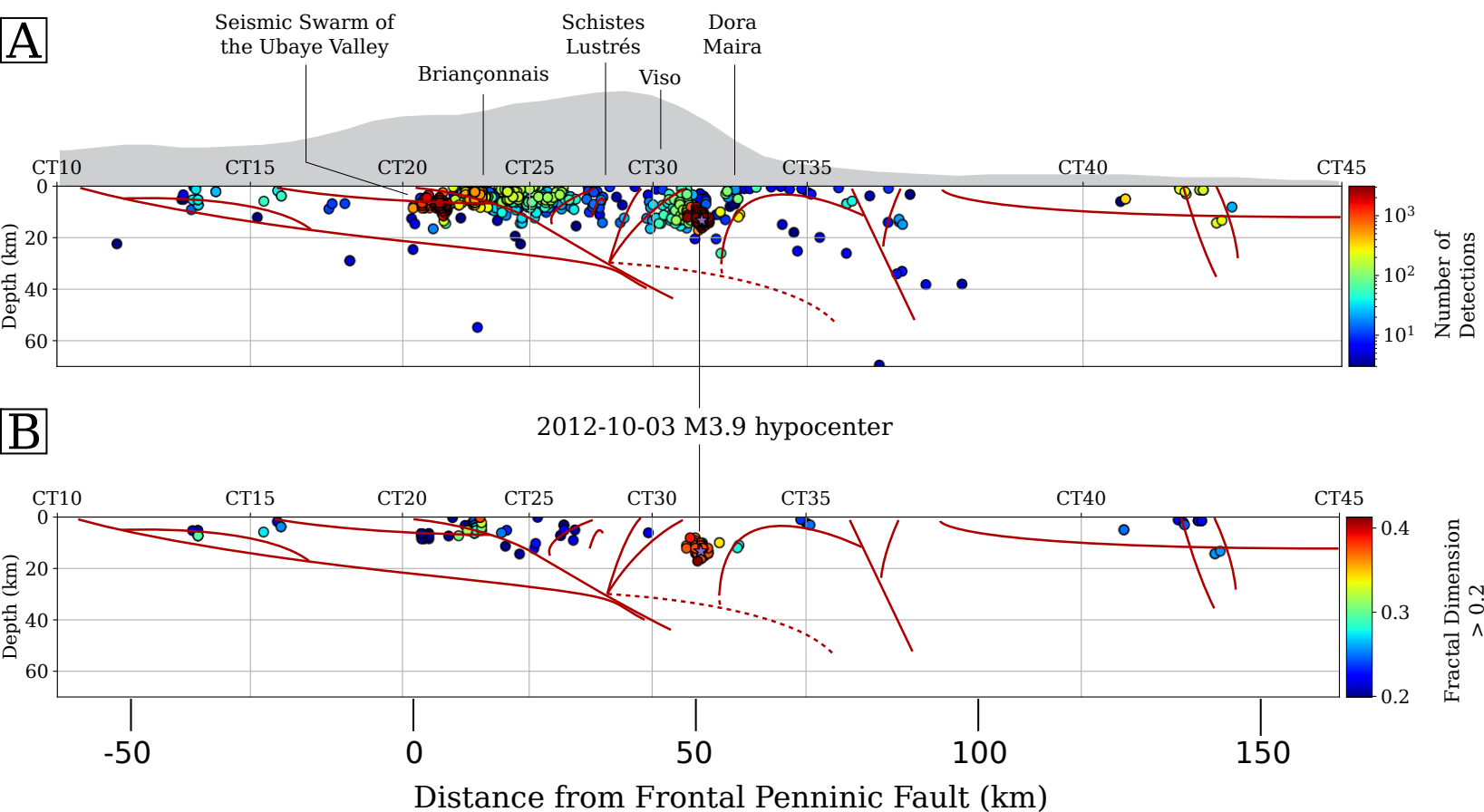
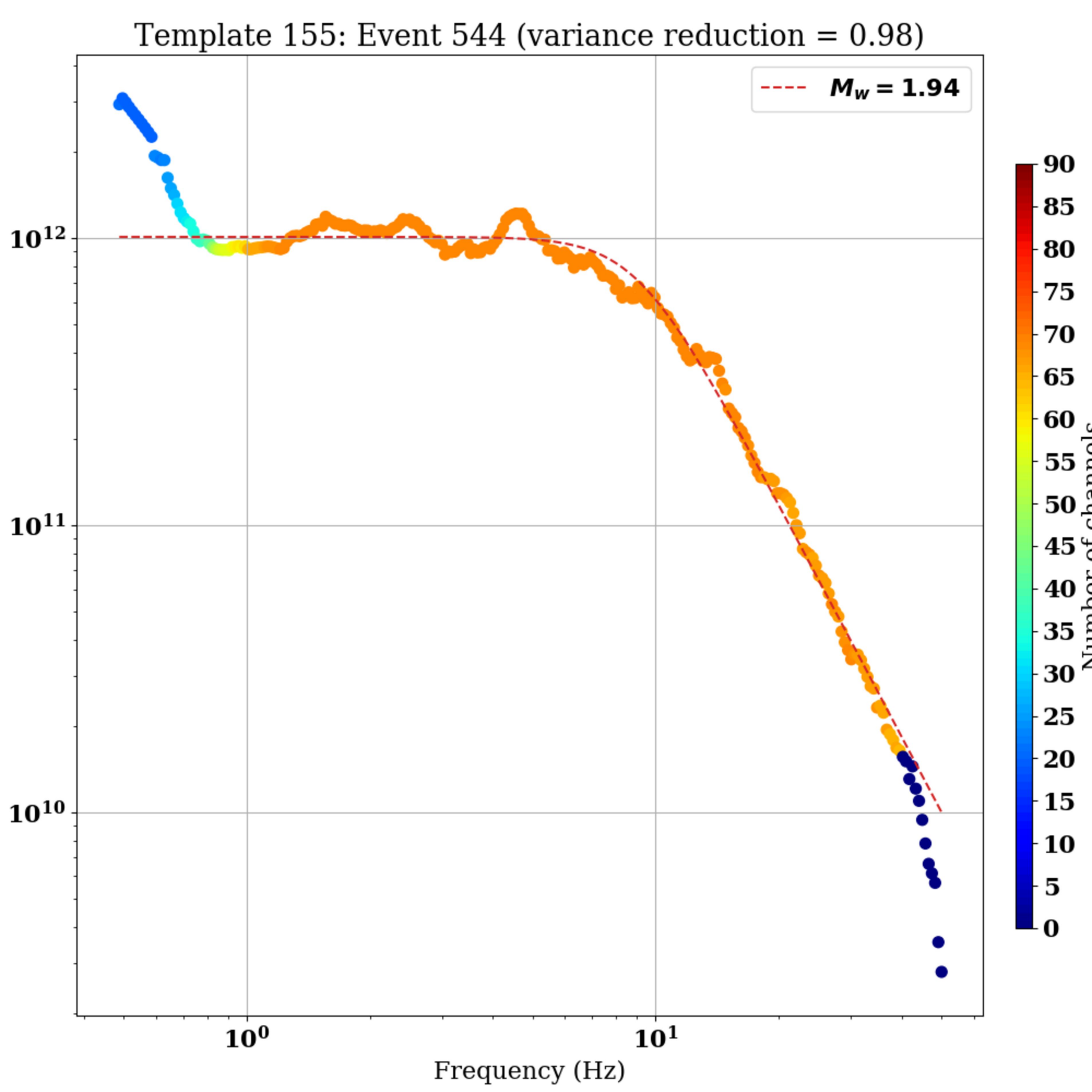

A
Seismic Swarm of the Ubaye Valley
Briançonnais
Schistes Lustrés
Viso
Dora Maira
CT10
CT15
CT20
CT25
CT30
CT35
CT40
CT45
Depth (km)
0
20
40
60
Number of Detections
10^3
10^2
10^1
2012-10-03 M3.9 hypocenter
B
Fractal Dimension > 0.2
0.4
0.3
0.2
-50
0
50
100
150
Distance from Frontal Penninic Fault (km)

Figure B1.


Template 155: Event 544 (variance reduction = 0.98)
$M_w = 1.94$
10^{12}
10^{11}
10^{10}
10^0
10^1
Frequency (Hz)
Number of channels
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0