Analyzing the Collective Behavior of Earthquakes to Understand Fault Mechanisms Better

by

Eric Beaucé

Submitted to the Department of Earth, Atmospheric and Planetary Sciences
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy of Geophysics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2021

© Massachusetts Institute of Technology 2021. All rights reserved.
Analyzing the Collective Behavior of Earthquakes to Understand Fault Mechanisms Better

by

Eric Beaucé

Submitted to the Department of Earth, Atmospheric and Planetary Sciences on August 30, 2021, in partial fulfillment of the requirements for the degree of Doctor of Philosophy of Geophysics

Abstract

The Gutenberg-Richter law tells us that there is a tenfold increase in the number of earthquakes of magnitude $m > M$ when $M$ decreases by one unit. Thus, the vast majority of earthquakes occur at magnitudes so small that the vibrations they cause can barely be recorded at the surface of Earth. Given that earthquakes are the symptoms of motion on faults, observing small earthquakes brings valuable information on fault mechanisms. In this thesis, not only do I focus on studying small-to-moderate size earthquakes ($M < 4$), but I study properties that emerge when many of these earthquakes interact. Many of my conclusions are drawn from observations of earthquake temporal clustering.

I present the automatic earthquake detection and location method that I developed for collecting the time and space coordinates of as many earthquakes as possible, and base all subsequent analyses on these. My investigations covered two study regions: the Southwestern Alps, and the western section of the North Anatolian Fault that last broke in August 1999. In both studies, I demonstrate how different fault systems produce seismicity with different temporal clustering properties. Observations of temporal clustering describe seismicity patterns between two end-members: the swarm-like seismicity with little inter-event triggering, and the cascade-like seismicity with strong earthquake interaction.

Temporal clustering and the analysis of earthquake source characteristics in the Southwestern Alps helped explain differences in fault mechanisms in the two most active areas of the study region. My results also point towards non self-similar earthquakes. Along the North Anatolian Fault, in addition to temporal clustering, I analyzed the earthquake focal mechanisms, used them to infer the state of stress in the fault zone, and thus provided a comprehensive description of the study region. A major conclusion of this study is that strongly time clustered seismicity developed in normal fault systems several years after the 1999 Izmit earthquake, and may indicate the inter-play between seismic and aseismic slip on these faults.
Thesis Supervisor: Robert D. van der Hilst
Title: Schlumberger Professor of Earth and Planetary Sciences

Thesis Supervisor: Michel Campillo
Title: Professor
Acknowledgments

This thesis is the outcome of five years of my life at MIT. Many people helped me go all the way to the end of these five years, which – as many would tell you – was quite uncertain in the beginning. In fact, my plan was initially to shorten my time in the US by leaving MIT after two years with a Master’s degree, and move back to France. And this almost happened: I got the degree but, in the end, decided to stay for the whole PhD program. My Master’s committee members certainly played a major role in making this happens when they enthusiastically suggested I stay, and I can now safely thank them for that: Matěc Peč, Brad Hager, Sai Ravela, and, of course, Michel and Rob. Rob found the right words to motivate me (in particular at the celebratory barbecue he organized at his place, after a couple refreshments). Behind this life choice were also an assortment of very special and quirky people at EAPS at the time: Aurélien, Eva, Harry, Shujuan, our friends from the second floor Aarti and Katya, and my roommates Loïc and Blake.

My research has always been deeply affected by exchanges with others, and I cherished my interactions with the postdocs there in my early PhD times, Aurélien, Piero and William. William holds a special place in my curriculum since he initiated my first scientific publication. My research group, Eva, Harry, Shujuan, Hongjian, and later Jing, always had valuable critical comments, even on topics sometimes remote to their own research interests. My two supervisors, Michel and Rob, have always provided me with the ideas and guidance that kept me doing great research all along. Despite their ever heavily occupied schedules, they always found time for my problems and always showed excitement about our research, which is the best motivation a student can get. In five years, I also shared many great non-research related moments with them. Finally, I want to acknowledge all the evenings I spent with Piero and Léonard, during my visits in Grenoble, that have considerably shaped my scientific thinking.
I thank all the members of my thesis committee: Leigh Royden, Rachel Abercrombie, Brad Hager, Michel, and Rob, for their time and precious comments, in particular towards the end of my PhD. I thank all my EAPS and ERL friends for making this experience more than work, with a special acknowledgement to the COVID team: Patrick, Aatri, Katya, Hervé, and Norah. I thank Loïc for the most entertaining year of my life, Blake for sharing my excitement about espressos and cheese, and Katya for our passionate arguments about earthquakes (and so many other topics) and her support. The list is long, and cannot be fully enumerated here, but I, of course, think about all the people in Lyon (la team), Grenoble, and my family in Normandy, who always supported my choices even when it involved being away for an extended period of time.
Contents

1 Introduction 45
  1.1 The Earthquake Phenomenon 45
    1.1.1 Birth of the Modern Understanding of Earthquakes 45
    1.1.2 Earthquakes as Sliding Instabilities 47
    1.1.3 Scaling Laws in Earthquakes 52
  1.2 The Earthquake Detection Problem 56
    1.2.1 Motivations 56
    1.2.2 A Brief Tour from Beamforming to Artificial Intelligence 58
  1.3 Collective Properties of Earthquakes 60
    1.3.1 The Gutenberg-Richter Law 60
    1.3.2 Cascades, Interactions, and Temporal Clustering 62
  1.4 Goals and Structure of the Thesis 67

2 Systematic Detection of Clustered Seismicity Beneath the Southwestern Alps 77
  2.1 Introduction 78
  2.2 Earthquake Detection Method 81
    2.2.1 Data Pre-processing 81
    2.2.2 Energy-based Detection (Composite Network Response) 81
    2.2.3 Classification of Seismic Signals 84
    2.2.4 Template Matching 87
    2.2.5 Second Generation Templates 90
  2.3 Seismicity of the Southwestern Alps 91
3 The Rupture Complexity of Small Earthquakes in the Southwestern Alps

3.1 Introduction .................................................. 124
3.2 Method ..................................................... 126

3.2.1 Data ................................................ 126
3.2.2 Spectral Ratios ........................................ 126
3.2.3 Theoretical Model and Inversion .............. 131

3.3 Results .................................................. 134

3.3.1 Scaling Law ........................................ 134
3.3.2 Departure from Self-Similarity ................. 138

3.4 Discussion .................................................. 141

3.4.1 Unconventional Rupture Models .............. 141
3.4.2 Fluids vs. Asperity: Ubaye vs. Dora Maira .... 142
3.4.3 Template Matching and Unconventional Earthquakes 145

3.5 Conclusions .................................................. 146

3.6 Supplementary Material .................................. 148

3.6.1 Model Comparison ................................ 148
3.6.2 Supplementary Results ......................... 148

4 Seismotectonic Study of the North Anatolian Fault Zone Thirteen Years After the 1999 M7.4 Izmit Earthquake

4.1 Introduction .................................................. 160
4.2 Earthquake Detection and Location .............. 163
B  An Iterative Linear Method for Estimating the Stress Tensor from Earthquake Focal Mechanism Data: Method and Examples

B.1  Introduction ................................................. 258
B.2  Methodology ................................................. 259
   B.2.1  Previous Work ........................................... 259
   B.2.2  This Study .............................................. 262
B.3  Synthetic Experiments ........................................ 266
B.4  Real Data Applications ....................................... 270
   B.4.1  Case Study 1: Central Crete ............................ 270
   B.4.2  Case Study 2: Western North Anatolian Fault ........ 271
B.5  Discussion on Shear Stress Magnitudes ...................... 274
B.6  Summary and Concluding Remarks ............................ 275
B.A  Stress Tensor Sign Convention ............................... 277
   B.A.1  Compression Positive ................................. 279
   B.A.2  Tension Positive ................................. 279
   B.A.3  Implications for the Stress Tensor Inversion ....... 280
List of Figures

1.1.1 A mass sits on an inclined table and is subject to several forces: its weight $W$ causes the normal force $F_n$ due to the reaction of the table. The weight has a component that is tangential to the table $F_t$, and this force is resisted by friction $F_f$ between the mass and the table. . 48

1.1.2 Spring-slider model. A: A block, the slider, lies on a surface and is coupled to a loading point moving at constant velocity $V_L$ through a spring with stiffness $k$. Thus, the slider is pulled by an elastic force $k\delta$ where $\delta$ is the elongation of the spring. This loading force is resisted by the friction $F_f$. B: Numeric simulation of the spring-slider model with rate-and-state friction and aging law (cf. Equation (1.2)). The slider is stationary most of the time (inter-seismic phase, see zero slip rate, in solid red, and flat displacement curve, in solid blue), but catches up with the motion of the loading point (dashed blue line) in periodic, fast slip events. The stick-slip behavior ensures that the spring does not keep accumulating energy indefinitely. C: Focus on the fast slip event (the rupture) when the inertial term of the equations of motion cannot be neglected (left hand side of Equation (1.1)). All axes are non-dimensional quantities. Courtesy of Ekaterina Bolotskaya. . . . . 49
1.1.3 Stability domains predicted by rate-and-state friction (Equation (1.2)).
The middle panel shows a measure of stability, $\zeta$, as a function of depth. These stability domains are sketched for a subduction interface (left) and a crustal fault (right). The rightmost panel shows the depth distribution of earthquakes for a section of the San Andreas fault near Parkfield. Figure from Scholz (1998).

1.1.4 Fault area $A$ against seismic moment $M_0$. The slope in the log-log domain is $2/3$, indicating that $A \propto M_0^{2/3}$. Lines of equal stress drops for circular cracks are shown. Figure from (Kanamori and Anderson, 1975).

1.1.5 Amplitude spectrum corner frequency $f_c$ against seismic moment $M_0$ (the corresponding moment magnitude $M_w$ is given at the top). The red dashed lines shown the constant stress drop scaling. Figure from Allmann and Shearer (2009).

1.2.1 A: Sketch of a plane wave reaching the surface. B: Normal sum vs. beamformed sum, i.e. shift and sum. Figure from Rost and Thomas (2002).

1.3.1 Frequency-magnitude relationship, referred to as the Gutenberg-Richter law (cf. Equation (1.7)). The slope in the log-log domain indicates a b-value $b = 1$. Figure 23 from Kanamori and Brodsky (2004), see their caption for description of the data.

1.3.2 Coulomb stress changes from the M7.6 1872 Owens valley are still consistent with the occurrence of moderate-to-large $M \geq 4$ earthquakes more than a century later. Areas of Coulomb stress increase are shown in red (promote failure) and areas of coulomb stress decrease are shown in blue (inhibit failure). Figure from Stein et al. (2019).
1.3.3 Quantification of the strength of earthquake clustering. An example of a weakly (blue) and strongly (green) time clustered sequences are given. **A:** Event count number, *i.e.* the number of events per unit time. **B:** Autocorrelation of the event count number. The weakly clustered sequence shows the characteristic delta autocorrelation of random sequences, whereas the strongly clustered sequence keeps a high correlation over a long time scale. **C:** Measure of the power law exponent of the event count spectrum. Clustered sequences exhibit large power law exponents. **D:** Fraction $x$ of time bins with size $\tau$ occupied by at least one earthquake: $x \propto \tau^{1-F}$. Clustered sequences have large fractal dimensions $F$. Figure from (Beaucé et al., 2019).

2.1.1 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).
2.2.1 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 2.1). 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 2.2) 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 × 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.
2.2.2 **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).

2.2.3 **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*).
2.2.4 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. 2.A for details). This weighting function is used to calculate a weighted average of the distance to the best potential seismic source (*cf. Equation 2.8 in 2.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.
2.3.1 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 2.2.5 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 $M_L 3.9$ earthquake that occurred in early October 2012, and which is important for the discussion in Section 2.4. 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.

2.3.2 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 2.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.
2.3.3 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 $\tau$ 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 $\beta$ 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.

2.4.1 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 2.4). The red structures are reported from the geological cross-section in Figure 2.1.1.
2.B.1 Magnitude estimation of the reference event. For each template, we use the highest SNR detections to calculate the average S-wave spectrum (Equation 2.9) and fit it with the Boatwright model (Equation 2.11). 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.

2.S.1 Left panel (A): Earthquake recorded at multiple stations. The waveforms are filtered in the band 1-12Hz and downsampled from 100 or 125 Hz to 50 Hz. Right panel (B): The envelopes of the seismic data are calculated and standardized: the daily median is removed, and the resulting signal is divided by its daily MAD (Median Absolute Deviation). Eventually, we cut out the 95th percentile of the signals by saturating the standardized envelopes \( u(t) \) with hyperbolic tangent: 
\[
\tilde{u}(t) = p_{95} \times \tanh \left( \frac{u(t)}{p_{95}} \right),
\]
where \( p_{95} \) is the 95th percentile of \( u(t) \). This processing ensures the stations to have equal noise level before stacking (cf. Figure 2.S.2), and decreases the effect of undesired spurious signals in the data. The three superimposed layers show the three components: north, east and vertical.

2.S.2 Statistics of the envelope data for a given day. Different whisker boxes are for different stations, with each component (north, east, vertical) in different subplots. Legend of the whisker boxes: orange line: median, lower side of the box: Q1, upper side of the box: Q3, lower whisker: Q1 - 1.5(Q3-Q1), upper whisker: Q3 + 1.5(Q3-Q1). Top panel: Raw envelopes. Bottom panel: MAD-normalized envelopes: 
\[
\tilde{u}(t) = \frac{u(t) - \text{Median}(u(t))}{\text{MAD}(u(t))}.
\]
After normalization, the stations exhibit similar distributions.
2.S.3 Composite network response (CNR, cf. Equation 2.2 in main material). **Left panel (A):** Illustration of the maxima searching operation achieved to calculate the composite network response from all the network responses of the grid. **Right panel (B):** Histogram of the CNR samples presented in (C). The threshold median $+10 \times$ MAD is given for information.

2.S.4 Second generation templates: increasing the SNR of the template waveforms. **Top panel:** Template matching provides us with many noisy repetitions of the same waveform. Different stacking methods can extract the coherent information from this collection of noisy records. **Bottom panel:** Stacking methods such as the $N^{th}$-root stack (*red waveform*) or the phase-weighted stack (*yellow waveform*) greatly improve the SNR with respect to the linear stack, but also distort waveforms because of non-linear operations. Our preferred method is the SVDWF (*green waveform*), which only performs linear operations. It exhibits a better SNR than the linear stack, and preserves the shape of the target waveform.

2.S.5 Comparison of the existing catalog with our catalog inside the dashed box; this region is where the geometry of the network allows best performances. Two events match if their origin times and locations are less than $\Delta T$ and $\Delta r$, respectively. $\Delta T$ and $\Delta r$ are two arbitrarily chosen thresholds. The unmatched events are shown with filled dots and the color codes their depth. For information, our template locations are shown with open diamonds. We missed 142 out of the 825 events (17%) documented in the catalog. However, we detected 16,430 new events, *i.e.* we detected a total of 17,113 earthquakes (21 times more detections with respect to the existing catalog). We show the same comparison for different $\Delta T$ and $\Delta r$ in Figure 2.S.6. Some of the unmatched events presented here are likely to be associated with inconsistencies in reported locations.
2.S.6 We perform the same comparison as in Figure 2.S.5 for different thresholds in order to investigate the effect of such arbitrary criteria. As the number of unmatched events decreases when we relax the criteria on origin time and location, it suggests that some of the unmatched events are only due to inconsistencies in location between the catalogs.

2.S.7 Distribution of magnitude of the 142 undetected events within the restricted area shown in Figure 2.S.5. Since the locations of the undetected events are close to the locations of our template events, we likely missed them because of their low magnitudes rather than because of the configuration of the station network.

2.S.8 Frequency magnitude distribution of the final catalog, only using templates with $\Delta r < 15$ km. The maximum likelihood estimate (red curve) is made on the range [1, $+\infty$). We observe the Gutenberg-Richter relation to break down at $M_L \approx 1$ (black dashed line).

2.S.9 On both panels, the y-axis reports the position of the templates projected along the axis defined by the linear network CIFALPS. **Top panel:** Number of detected earthquakes per 10-day sliding window for each template. **Bottom panel:** Clustering coefficient (cf. definition in main material Figure 2.3.3) per 10-day sliding window for each template. Comparison of the two panels shows that strong seismic activity is often associated with high temporal clustering, but also that this is not always true. We observe continuous seismic activity beneath the Ubaye valley (with projected location around CT20), but only few episodes of high temporal clustering. This is consistent with previous observations pointing at a mixture of swarm-like seismicity and foreshocks-mainshock-aftershocks sequences in the Ubaye valley Daniel et al. (2011); Leclère et al. (2012, 2013); De Barros et al. (2019).
2.3.1 Cross-section along the CIFALPS axis showing 976 templates that were well relocated ($\Delta r < 15$ km). The color codes for different attributes in each cross-section, from top to bottom: number of detections, correlation time, fractal dimension and clustering coefficient of the one-year earthquake sequence of each template. Visual inspection seems to reveal that the fractal dimension offers a more contrasted image of temporal clustering than the clustering coefficient.  

3.2.1 Centroid locations (filled dots) of the 81 groups of similar template earthquakes. Color shows the centroid depth. Each group detected at least 10 events, and these similar earthquakes are used for computing spectral ratios (Equations (3.2) and (3.3)). Black inverted triangles are the 82 seismic stations used in this study. The seismicity in the Ubaye valley and in the Dora Maira massif is discussed at length in this manuscript.  

3.2.2 Waveform alignment and spectral ratio of a pair of events on a single station. **A-C:** Three component P-wave waveforms of event 1 (black) and event 2 (orange). The first P-wave arrivals have a correlation coefficient (CC) of 0.75 after shifting event 2 backward by one sample. **D-F:** Three component S-wave waveforms of event 1 (black) and event 2 (orange). The first S-wave arrivals have a correlation coefficient of 0.87 after shifting event 2 forward by 6 samples. Note: the CC was computed on a narrow window around the first arrivals, and not over the whole time window shown here. **G-I:** P- (dashed blue) and S-wave (solid orange) spectral ratios (Equation (3.3)) of the three component waveforms shown above.
3.2.3 Multi-station average spectral ratio for a given pair of events. **A, B:** P-wave and S-wave average spectral ratio, respectively. Lines colored with a shade of blue are single-channel spectral ratios. The color shows the correlation coefficient between the first arrivals of the P or S wave (see Figure 3.2.2). The orange line is the median stack (see Equation (3.4) and text) and the shaded grey area shows the uncertainty on the stack. Only single-channel spectral ratios with CC greater than 0.5 are included in the average, in addition to the signal-to-noise ratio (SNR) criterion (see panels below). **C, D:** P-wave and S-wave single-channel SNR. The SNR is computed with respect to a noise window taken just before the P-wave arrival. Solid lines are SNRs of the first event of the pair, dotted lines are SNRs of the second event. Color is the same as for panels A and B. The dashed black line is the threshold on SNR above which a frequency bin is included in the average. Both SNRs of event 1 and 2 are required to be above threshold.

3.2.4 Multi-station average spectral ratios between the largest magnitude event and all others (thick colored lines) and model predicted spectral ratios (thin colored lines) for **A:** the P wave and **B:** the S wave. Curves are colored according to the catalog magnitude of the second event. The corner frequencies of the two events $f_{c,1}$, $f_{c,2}$ of each pair are shown with black circled dots. Note that this group of earthquakes was taken from the Dora Maira massif (see location on Figure 3.2.1), and that the model does not seem to explain well the low-frequency part of the ratios ("bump" in the 3-8 Hz band).

3.3.1 Inverted seismic moments and corner frequencies for a given group of similar earthquakes (the same as in Figure 3.2.4), on the P (blue dots) and the S (orange dots) waves. The power-law exponent $\alpha$ of the $f_c^{-\alpha} M_0$ scaling relation (Equation (3.12)) is measured by a robust linear regression in the log-log domain (11-norm minimization). Uncertainties are estimated with bootstrap resampling.
3.3.2 Corner frequency vs seismic moment scaling and other parameters. **A:** Centroid depth. **B:** Power-law exponent $\alpha$ of Equation (3.12). The color scale is centered around the self-similarity power-law exponent: $\alpha_{ss} = 1/3$ (white). All exponents that were not different from $\alpha_{ss}$ within the 1-$\sigma$ uncertainty were set to $\alpha_{ss}$ on this map. **C:** Number of events that were successfully inverted in each group (P- and S-wave inversion). **D:** Fractal dimension of the time series of event count per unit time (see Beaucé et al., 2019). Large values mean strong temporal clustering. Diamonds are results from groups with less than 10 successfully inverted events (see panel C). The rest of the groups are shown with dots.

3.3.3 All corner frequencies and seismic moments that were successfully inverted. **A:** P-wave inversion. **B:** S-wave inversion. Fitting a power-law to the data $f_c \propto M_0^{-\alpha}$ yields low exponents $\alpha$: 0.08 and 0.14 for the P and S waves, respectively. These data could also be explained by the canonical scaling law, i.e. $\alpha_{ss} = 1/3$, and a spread of stress drop values (see 0.1, 1, 10, 100 and 1000 MPa lines, computed with the model from Madariaga, 1976). Grey diamonds are results from groups with less than 10 successfully inverted events (see Figure 3.3.2C).
3.3.4 Unusual corner frequency - seismic moment scaling relationship in the Dora Maira massif. Multi-station average displacement amplitude spectra of A: P waves and B: S waves corrected for geometrical spreading and attenuation (assuming a $Q \propto f^{-0.5}$ model, e.g. Campillo et al., 1985). We use the magnitudes and corner frequencies inferred with the spectral ratio inversion to further average amplitude spectra and corner frequencies in magnitude bins (black diamonds are the average corner frequencies). Dashed black lines show the self-similar ($\propto f^{-3}$) and the inverted ($\propto f^{-6.67}$, i.e. $\alpha = 0.14$) scaling. C: P-wave and D: S-wave amplitude spectra are unsuccessfully collapsed using the self-similar scaling. E: P-wave and F: S-wave amplitude spectra are satisfactorily collapsed using the inverted scaling.

3.3.5 Unusual corner frequency - seismic moment scaling relationship in the Ubaye valley. Multi-station average displacement amplitude spectra of A: P waves and B: S waves corrected for geometrical spreading and attenuation (assuming a $Q \propto f^{-0.5}$ model, e.g. Campillo et al., 1985). We use the magnitudes and corner frequencies inferred with the spectral ratio inversion to further average amplitude spectra and corner frequencies in magnitude bins (black diamonds are the average corner frequencies). Dashed black lines show the self-similar ($\propto f^{-3}$) and the inverted ($\propto f^{-11.34}$, i.e. $\alpha = 0.09$) scaling. C: P-wave and D: S-wave amplitude spectra are unsuccessfully collapsed using the self-similar scaling. E: P-wave and F: S-wave amplitude spectra are satisfactorily collapsed using the inverted scaling.

3.4.1 Corner frequencies and seismic moments estimated in A: the Ubaye valley and B: the Dora Maira massif, for the P (blue symbols) and S (orange symbols) waves. Solid lines indicate different levels of constant stress drop scaling (computed following the model of Madariaga, 1976). The inverted scaling laws are shown with the dashed lines.
3.S.1 Multi-station average spectral ratios between the largest magnitude event and all others (thick colored lines) and model predicted spectral ratios (thin solid line: l2 loss, thin dashed line: soft l1 loss). **Left column (A, C, E):** P wave. **Right column (B, D, F):** S wave. **Top row (A, B):** Brune model ($\gamma = 1$ in Equation (3.6) in main text). **Middle row (C, D):** Boatwright model ($\gamma = 2$ in Equation (3.6) in main text). **Bottom row (E, F):** Custom model ($\gamma$ is a free parameter). The Boatwright and custom models with the l2 loss do not converge to a satisfactory solution on the largest magnitude event.

3.S.2 Residuals between log observations and log models for P (blue distributions) and S (orange distributions) waves. **First row (A, B, C):** L2 loss with the Brune, Boatwright and custom model. **Second row (D, E, F):** Soft l1 loss with the Brune, Boatwright and custom model. The Boatwright model with the soft l1 loss does marginally best.

3.S.3 Inverted corner frequencies $f_c$ and seismic moments $M_0$ for all events. The exponent of the scaling law $f_c \propto M_0^{-\alpha}$ was estimated with a linear regression in the log-log domain using the least absolute value criterion (dashed lines). Uncertainties were estimated by bootstrapping the data set and repeating the regression. **First row (A, B, C):** L2 loss with the Brune, Boatwright and custom model. **Second row (D, E, F):** Soft l1 loss with the Brune, Boatwright and custom model. Preferred inversions show consistent results between the P and S waves. Boatwright and soft l1 loss also produce the smallest residuals (*cf.* Figure 3.S.2), therefore we choose this model in the study.
3.S.4 P-S averaged power-law exponent $\alpha$ (cf. Equation (3.12)) vs. **A**: Centroid depth, **B**: Root mean square (RMS) residual, **C**: Total number of inverted events (P- and S-wave inversions), and **D**: Fractal dimension of the event count per unit time (see main text). Red diamonds are groups for which less than 10 events were successfully inverted, therefore for which the power-law exponent is not very reliable. Blue dots are the rest of the groups.
4.1.1 **A:** Large scale view of the North Anatolian Fault Zone. Abbreviations: NAFZ - North Anatolian Fault Zone, EAFZ - East Anatolian Fault Zone. The red arrows indicate the direction of coseismic motion. Our study region is located at the western end of the North Anatolian Fault (NAF). **B:** Magnified view of the fault zone in our study region. Larger font names are the main geologic units: Istanbul Zone, Armutlu Block, Almacik Mountains and Sakarya Terrane. The smaller font, italic names are segments and faults of the NAF: the Izmit-Sapanca segment, the Sapanca lake step-over, the Sapanca-Akyazi segment (which together constitute the Northern strand), the Karadere segment and the Southern strand (names following Barka et al., 2002). The Sapanca-Akyazi segment is made of the Sakarya fault and the Akyazi fault. The flat area around the Akyazi fault is referred to as the Akyazi plain. Both Lake Sapanca and the Akyazi plain are pull-apart basins. The large red star indicates the epicenter of the $M_w7.4$ Izmit earthquake, and the small purple star indicates the epicenter of the $M_w7.2$ Düzce earthquake. **C:** The seismic stations used in this study are from the temporary experiment DANA (70 stations, red triangles; DANA, 2012) and the permanent network (9 stations, black triangles; Kandilli Observatory And Earthquake Research Institute, Boğaziçi University, 1971). Each column of the DANA array is indexed by a letter and each row is indexed by a number (DA01, DA02, ..., DB01, ...).

4.2.1 Summary flowchart of the earthquake detection and location method. For clarity, only a subset of stations is shown in the above panels, but all the analysis is carried on the 79 stations together. Template matching is performed on the 10 stations closest to the source and the detection threshold is set to $8 \times$ RMS of the correlation coefficients in a 30-minute sliding window. See Data and Resources for code availability.
4.3.1 Map view of the locations of the template earthquakes detected and used in this study. Only templates with maximum horizontal uncertainty less than 15 km and depth less than 20 km are shown (total of 3,320 templates). Filled dots are for natural earthquakes (1,471 templates), and squares are for mining-related events (1,849 templates; see text for details about identifying templates as mining templates). **A:** Event depths. **B:** Cumulative number of event detections per template. Most of the earthquake signals found in the dataset actually originate from outside the North Anatolian Fault Zone. **C:** Maximum vertical uncertainty derived from the uncertainty ellipsoids. **D:** Maximum horizontal uncertainty derived from the uncertainty ellipsoids. As expected, template earthquakes that are located further from the stations have larger location uncertainties.

4.3.2 Spatio-temporal distribution of the earthquake activity in the study region. The longitude of each event is shown against its origin time, and the color codes the latitude. **A:** We detected 31,356 events with the 3,320 template earthquakes presented in Figure 4.3.1 from 2012-05-04 to 2013-09-20. **B:** The templates due to natural seismicity detected 16,708 earthquakes. The seismic activity taking place on the NAF (latitudes 40.40°N-40.80°N) is eclipsed by the numerous earthquakes occurring elsewhere.
4.3.3 Earthquakes in the North Anatolian Fault Zone. **A:** Locations of the template earthquakes with color coded depths. We define nine subregions along the different segments of the fault. Only in this figure the Sapanca-Akyazi region is subdivided into a fault parallel and a fault perpendicular sections. The dashed lines inside each colored box define either fault parallel or fault perpendicular cross-sections (see bottom panels, **C**). The color shading of each box is only to help distinguish between them. **B:** Earthquake hypocenters successfully relocated with the double-difference method and color coded by depth. Events for which relocation was not successful were attributed the template location. **C:** Depth cross-sections of the different areas introduced above. The earthquake locations contained in the boxes are projected onto the boxes’ central axis (thin black dashed lines). The bottom x-axes are distances along the cross-section axes in kilometers, and the top x-axes are the geographic coordinates relevant to each cross-section (either longitude or latitude).

4.3.4 Earthquake activity seen on recurrence time vs detection time graphs for different subsets of the earthquake catalog (refer to Figure 4.3.3 for the name of the areas). The recurrence time is the time between two consecutive events detected by a same template. Note that the y-axis is in log scale and that some seismic episodes span many orders of magnitude of recurrence time. These episodes are characteristic of burst-like, or cascade activity (see text). The color codes the local magnitude, and inverted grey triangles are events for which no reliable estimates were obtained.
4.3.5 **A:** Map view of template earthquakes with color-coded fractal dimension (cf. Equation (4.4)) showing the strength of temporal clustering. **B:** Map view of template earthquakes with color coded cumulative number of detections. In both top panels, the shaded areas refer to the regions introduced in Figure 4.3.3. **C:** Template earthquakes with color coded fractal dimension on fault parallel and fault perpendicular cross-sections (same color scale as panel A). Hypocenters are projected along the dashed axes shown on the map view. High fractal dimensions mean strongly time clustered activity (i.e. past events strongly influence the timings of future events).

4.3.6 Clustering vs. depth vs. event density. Inside each region, templates are binned per distance from the bottom of the seismogenic zone and the fractal dimension is averaged among the 10% largest values (one dot per bin). The location of the bottom of the seismogenic zone is approximated by the depth of the locally deepest template. Dots are colored according to the average inter-event distance in the cloud of earthquakes detected by the selected template; this is a proxy for event density. Darker colors mean higher event density. Strongest clustering tends to occur at the bottom of the seismogenic zone, i.e. at the transition zone between stable and unstable sliding.
4.4.1 Estimation of focal mechanisms. **A:** We find the synthetic waveforms that fit best the first P- and S-wave arrivals on all seismic stations and components. Black waveforms are data, blue and red dashed waveforms are synthetic P and S waves, respectively (Cotton and Coutant, 1997, computed with Axitra,). The waveforms are bandpass filtered in 2-6 Hz. Fine waveform realignment is performed by cross-correlating the synthetics and data to account for errors in location. **B:** To estimate uncertainties on focal mechanisms, we generate 100 focal mechanism solutions by randomly excluding 10 stations each time (leave-10-out jackknife resampling). The colored focal mechanism on the left is the solution from the minimum of the averaged 100 objective functions, and on the right is the solution from averaging the 100 (normalized) moment tensors. We chose to use the latter as our best solution. **C:** Location of the template earthquake taken as an example. The solution shows a right-lateral strike-slip fault striking to the northeast.

4.4.2 Focal mechanisms of all 436 template earthquakes located in the vicinity of the North Anatolian Fault, following the method described in Section 4.4.1. The beachballs are lower hemisphere P-wave radiation patterns and their size is scaled according to their uncertainty (smaller beachballs have larger errors). **A:** Strike-slip faulting earthquakes. **B:** Reverse faulting earthquakes. **C:** Normal faulting earthquakes. **D:** Distribution of faulting regimes with depth. Each event of the catalog is attributed the focal mechanism of the template to which it correlates best. **E:** Focal mechanisms located in the so-called Kaverina diagram (Kaverina et al., 1996), which we use to categorize the faulting regimes. The colors of the symbols match the other panels. This panel was created with a plotting routine from the focal mechanism analysis software FMC (Álvarez-Gómez, 2019).
4.4.3 Inverted local stress tensors linked on the map to their corresponding subregion. The stereographic projections show the orientations of the three principal stresses \( \sigma_1 \), \( \sigma_2 \) and \( \sigma_3 \), from most compressive to least compressive. The grey symbols are the 2000 solutions obtained by repeating the inversion on randomly sampled data sets from the 100 focal mechanism solutions available for each template earthquake. See Table 4.2 for numerical values of azimuth and plunge. We use these empirical distributions to estimate the 80\% and 90\% confidence intervals, shown here with the darker and lighter contour lines. The histograms show the distributions of shape ratios \( R \), see Equation (4.6)\) and the vertical red bars indicate the shape ratios obtained by inverting the set of best focal mechanisms. In terms of deviatoric stress, \( R < 0.5 \) indicates that \( \sigma_2 \) is compressional, whereas \( R > 0.5 \) indicates that \( \sigma_2 \) is extensional.

4.5.1 Comparison of the two step-overs bounding the Sapanca-Akyazi segment on both sides. **A:** Eastern side of Lake Sapanca. **B:** Beneath the Akyazi plain, between Sapanca-Akyazi and Karadere. **Top panels:** Lower-hemisphere focal mechanisms of template earthquakes colored according to their depth. Arrows indicate the predominant different slip directions among the focal mechanisms. The side panels show the distribution of faulting styles with depth. Each detected event is attributed the focal mechanism of the template earthquake with which it correlates best. **Bottom panels:** Frequency-magnitude distribution.
4.S.1 **Top left panel:** Mining-related seismicity is characterized by predominantly diurnal seismicity, whereas we expect no preferred time for natural seismicity. In fact, natural seismicity shows slightly more events at night because noise is generally lower, and earthquake detection is easier. **Bottom left panel:** Mining-related seismicity also often shows no earthquakes on Sundays. **Right panels:** The waveforms produced by these mining-induced earthquakes have all characteristics of natural earthquakes, with clear P and S waves.

4.S.2 Average S-wave spectrum fitted with the Brune model (red curve). This is a weighted average of all single-channel S-wave spectra (thin grey spectra, Equation (4.7)). The weight of each frequency bin of each channel is proportional to the excess signal-to-noise ratio (SNR) defined as \( w(f) = \text{SNR}(f) - \text{SNR}_t(f) \), where \( \text{SNR}_t(f) \) is the minimum SNR value that the frequency bin \( f \) must exceed in order to contribute to the average. Every frequency bin of the average spectrum also has a weight that is equal to the sum of the single-channel weights. Note that because we correct the single-channel spectra for geometric spreading and attenuation, the low-frequency plateau shown here gives directly the seismic moment \( M_0 \).
4.S.3 Seismicity of the 2012-07-07 $M_L$4.1 Serdivan earthquake sequence. **Top left panel:** Cumulative number of earthquakes and local magnitudes $M_L$. The mainshock was followed by about 30 events in the next four hours, but only we recorded only 10 events in the next 26 days. **Top right panel:** Epicenters colored by time relative to the mainshock. Epicenters’ alignment and the largest events’ focal mechanisms suggest the existence of two conjugate faults and a network of secondary faults. **Bottom left panels:** Spatio-temporal evolution of the earthquake sequence. Successfully relocated hypocenters do not show any migration pattern consistent with fluid diffusion with diffusivity $D \approx 0.2 \text{–} 0.3 \text{m}^2/\text{s}$ typically observed for swarm seismicity (Shapiro et al., 2002) nor fast linear migration (> 30 km/day), suggesting that the earlier part of the sequence was controlled by static and dynamic stress changes. However, we can visually identify a southeastward migration of the seismicity in the later days of the sequence (top right panel).
4.S.4 Quantification of the strength of temporal clustering. Top left panel: Earthquake occurrence time series, alias the event count, i.e. number of events per unit time. Top right panel: Autocorrelation of the event count. A random time series would show a perfect Dirac function. Bottom left panel: Spectrum of the event count. The spectrum of a time clustered event count follows a power law $\propto f^{-\beta}$ and $\beta$ quantifies the strength of temporal clustering. Bottom right panel: Fractal analysis of the event count. We measure the number of time bins occupied by at least one event when dividing the time axis into smaller and smaller time bins. The slope of the curve gives the fractal dimension, which also quantifies the strength of temporal clustering. We fit the curve between 100 s (to avoid dealing with short time scales that can be corrupted by events counted twice, even though we normally take care of that) and the inverse of the average event rate (to avoid fitting the time scales where fractality trivially breaks down).

4.S.5 Earthquake clustering along the North Anatolian Fault Zone. Top panel: Fractal dimension (as introduced in Figure 4.S.4). The eastern Marmara Sea and Lake Sapanca show the strongest clustering along the NAF. Bottom panel: Cumulative number of detections per template.

4.S.6 Objective function used in the focal mechanism estimate. Top row: Root Mean Square (RMS) waveform amplitude difference. Middle column: Negative waveform correlation (minimizing the negative correlation means maximizing the correlation). Bottom row: The objective function is made of the sum of these two terms. Note that the RMS and the correlation are already scaled so that they can be summed without having one term dominate the sum. The three columns show orthogonal 2D slices containing the best solution.
4.S.7 Faulting regime vs depth in each region. Each event is attributed the focal mechanism of the template earthquake with which it correlates best. If the corresponding template earthquake has no focal mechanism estimate, the event is not included in these histograms.

4.S.8 Regional stress state: All focal mechanism data, both from the northern and the southern strands, were inverted. **Left panel:** Pressure (P, red inverted triangles) and tension (T, blue dots) axes of all focal mechanisms. The study region taken as a whole exhibits a wide variety of P/T axes, which violates the assumption of uniform stress state. **Middle panel:** Directions of the principal stresses on an equal area stereographic projection. The principal stresses are ordered from most compressive ($\sigma_1$) to least compressive ($\sigma_3$). The numbers in the legend are the azimuth (angle from north) and the plunge (angle from horizontal) of the axes. The contours are the 90% and the 95% confidence regions. **Right panel:** Shape ratio, defined as $R = \frac{\sigma_1 - \sigma_2}{\sigma_1 - \sigma_3}$. It measures the relative magnitude of the principal stresses. In both the middle and right panels, the distribution of possible solutions was obtained by randomly sampling the set of possible focal mechanisms generated for each template earthquake (see main text).
5.2.1 Decomposition of GNSS displacement time series based on local seismicity. The GNSS station is IZMT, Turkey, located north of the NAF. The original displacement time series (grey inverted triangles) is decomposed into the low seismicity displacement time series (black squares) and the high seismicity displacement time series (orange dots). Decomposition is performed on differentiated time series (speed), and decomposed time series are then integrated (back to displacement). A: East component. B: North component. C: Vertical upward component. The long term speed is first estimated on the initial time series with a linear regression (grey solid line), and the low seismicity and high seismicity speeds (black and orange solid lines, respectively) are then determined such that the modelled displacements sum to the initially modelled displacement. That is, the black and orange solid lines put together are forced to match the grey line when inverting for the high and low seismicity speeds. The vertical component is usually harder to interpret because subject to environmental noise such as rainfall.

5.2.2 The sand pile model traditionally used to describe a self-organized critical system. a: Critical state of a sand pile. b: A large avalanche occurred on the right side of the pile, thus the system is locally no longer at the critical state on this side. The pile has to be rebuilt before developing large avalanches on this side. c: A small avalanche occurred on the right side, making the right side more unstable and future avalanches more likely. d: As the right side of the pile approaches the critical state, moderate-size avalanches occur. Figure from Sykes et al. (1999).
5.2.3 Sites of a 2-D grid are occupied with probability $p$. 

**A:** $p = 0.420$. 

**B:** $p = p_c = 0.593$ (i.e. at the critical state, see Stauffer and Aharony, 2018). A cluster that links the two sides of the grid appear at the critical state (i.e. there is a path for percolation from one side to the other). 

**C:** $p = 0.850$. The size of the clusters is given by the color scale. 

**D-F:** At $p = p_c$, clusters exhibit scale invariance. When zooming in the grid, there is no length scale that gives a sense of scale (the size of a pixel becomes actually visible in F).

---

**A.2.1 Workflow of the CPU implementation of our program Fast Matched-Filter (FMF).** A single large I/O operation is achieved at the beginning by reading the templates’ waveforms and the continuous data. Beside, the sums of the squared templates and the cumulative sum of the squared data are computed before entering the loops to avoid redundant operations. After that, the iterative computation of the CC starts and is parallelized with OpenMP: different sliding windows (chunks) of the data are assigned to different threads. The two dashed boxes indicate which part is executed by the wrapper, and which one by the C code.

**243**

**A.3.1 Workflow of the GPU implementation of FMF.** The continuous seismic data and the templates are first read from the disk by the CPUs and transferred to the GPUs. The GPUs take advantage of the collective behavior of the threads and their quick access to shared memory to efficiently parallelize the computation. One GPU sets up several blocks ($\sim 10$) whom each creates many threads (512) that computes the CC on all the components of a given template/station and at different times ($CC_{s,c}(t_n)$). The average correlation coefficients are computed through a weighted average, and are eventually transferred back to the CPUs. The two dashed boxes indicate which part is executed by the wrapper, and which one by the CUDA C code.

**246**

---

39
A.3.2 Organization of the grid of thread blocks set up by the GPU to compute the correlation between one template and the data. The block size (the number of threads per block) is 512 and the grid size (the total number of blocks) is $N_B$. The grid size depends on the length of the data array and on the temporal step; $N_B$ is just enough so that all the correlation coefficients are computed. Note that this scheme holds for a single template, but the same scheme runs in parallel when several GPUs are available.

A.4.1 Comparison of the run times of matched-filter searches with different codes. Matched-filtering is achieved between one-day long synthetic seismograms on 5 stations, 3 components and a set of 8-second long templates whose size varies from 1 to 10. The sampling rate is 10 Hz and the temporal step used in the CC computation is 1 sample. Note the log scale on the y-axis.

A.4.2 Comparison of the run times of several matched-filter searches using the CPU and the GPU implementations of our FMF and EQcorrscan. Matched-filtering is achieved between a set of one-day long synthetic seismograms on 12 stations, 3 components and a set of 8-second long templates whose size varies from 1 to 50. The sampling rate is 50 Hz and the temporal step takes the values 1, 5 and 10. Note the log scale on the y-axis.

A.5.1 Matched-filter search of one template in the Western Alps, France over a single day. **Top left panel:** Map of the region showing the template’s location (star) and the neighboring stations (inverted triangles). **Bottom left panel:** Correlation coefficients (Equation A.2) over the whole day. The threshold of $10 \times \text{MAD}\{\text{CC}(t)\}$ is plotted with the dashed line. A total of 83 earthquakes was detected with constraining each event to be separated by at least 3 seconds. **Right panel:** Template’s channel CT23.HHZ (top record) and 10 examples of detections on this same channel.
B.2.1 Definition of the instability parameter (Equation (B.13), following Vavryčuk et al., 2013) in the Mohr space with the negative compression convention. The red straight lines are the failure lines whose slopes are controlled by the friction $\mu$. The most unstable fault has coordinates $(\sigma_c, \tau_c)$ in the Mohr space. The $\sigma_i$'s are the principal stresses ordered from most compressive to least compressive.

B.3.1 Synthetic experiment 1. The true stress tensor (leftmost column, middle row, large black symbols) promotes right-lateral strike-slip faulting on east-west oriented vertical faults. The shape ratio is 0.50. The fault orientations are randomly chosen from a range of parameters that is physically sensible given the stress state (see text), and the rakes are chosen such that slip is along the maximum shear stress direction. 

- **a, e, i:** Data set with only the 100 true fault planes.
- **b, f, j:** Noise free data set with both the fault planes and their auxiliary planes.
- **c, g, k:** Data set with noisy fault planes with strikes/dips/rakes perturbed by random values in $[-3^\circ; +3^\circ]$, and their auxiliary planes.
- **d, h, i:** Data set with noisy fault planes with strikes/dips/rakes perturbed by random values in $[-10^\circ; +10^\circ]$, and their auxiliary planes.

**d:** Fault planes (black lines) and auxiliary planes (grey lines). 

**e, f, g, h:** Lower hemisphere, equal area stereographic projections of the principal stress axes and their 95% confidence intervals (CI) estimated from 1000 bootstrap resamplings: solid lines = $\sigma_1$ CI, dashed lines = $\sigma_2$ CI, dot-dashed lines = $\sigma_3$ CI. The legend shows the inverted shape ratios $R$, and the mean angle $|\Delta \theta|$ between the predicted shear directions on the true fault planes and the true slip directions. Circles, squares and triangles are the most compressive ($\sigma_1$), intermediate ($\sigma_2$) and least compressive ($\sigma_3$) stresses, respectively. 

**i, j, k, l:** The distributions of shape ratios from the 1000 bootstrap resamplings. The vertical black line indices the true shape ratio. The proposed method is labeled "Iterative failure criterion" and is shaded for clarity.
B.3.2 Predicted shear stress magnitudes. a, c: Synthetic experiment 1. b, d: Synthetic experiment 2. The linear inversion does not predict the relative shear stress magnitudes correctly, but our iterative procedure retrieves the true values. 

B.3.3 Synthetic experiment 2. The true stress tensor (e) promotes right-lateral oblique strike-slip faulting with a normal faulting component on east-west oriented faults. The shape ratio is 0.70. Same legend as Figure B.3.1.

B.4.1 Inversion of the central Crete data set from Angelier (1979). a: Fault plane orientations. b: Principal stress directions of the inverted stress tensor with the linear method due to Michael (1984) (blue symbols) and the iterative method introduced in Section B.2.2. In the original publication, Angelier (1979) find similar stress directions and a shape ratio of \( R = 0.93 \). The 95% confidence intervals (CI) derived from 1000 bootstrapped data sets are shown: solid lines = \( \sigma_1 \) CI, dashed lines = \( \sigma_2 \) CI, dot-dashed lines = \( \sigma_3 \) CI. c: Empirical shape ratio distributions estimated from the bootstrap resampling. d: Distribution of shear stress magnitudes resolved on the fault planes.

B.4.2 Inversion of the North Anatolian Fault data set from Poyraz et al. (2015). a, b, c, d: Focal mechanisms from the DANA data set (Dense Array for North Anatolia DANA, 2012). e, f, g, h: Focal mechanisms from the 1999-08-17 \( M_w7.6 \) Izmit earthquake and some of its aftershocks. Columns are the same as in Figure B.4.1. The measure of misfit \( |\Delta \theta| \) and the shear stress magnitudes (right panel) were computed on the fault planes selected by the instability parameter (Equations (B.13) and (B.14)).

B.A. East-west right-lateral strike-slip fault under northwest-southeast maximum compression and southwest-northeast minimum compression. The intermediate stress is vertical.
List of Tables

1.1 Summary of scaling relations between source parameters. N.B.: These relations only hold for earthquakes that do not saturate the seismogenic width. ................................................. 54

4.1 1D velocity model due to Karabulut et al. (2011) used in this study. . 166

4.2 Compilation of results. †: Method from Gephart and Forsyth (1984). ‡: Method from Michael (1984, 1987). *: Method from Beaucé et al. (2021a). Note that for axes with plunges close to 0°, azimuth differences near 180° mean the axes are nearly colinear. . . . . . . . . . . . . . . 186
Chapter 1

Introduction

1.1 The Earthquake Phenomenon

1.1.1 Birth of the Modern Understanding of Earthquakes

For most of Human’s history, earthquakes were considered of supernatural origin or, if otherwise such as in ancient Greece, explanations were based on little scientific grounding. The first occurrences of measuring earthquakes were reported in China about 2000 years ago with the invention of the seismoscope by the scientist Zhang Heng. The device measured the incident direction of passing waves produced by a distant earthquake, and was used to determine approximately the location of the earthquake in order to send aid to affected regions in a timely manner. Ironically, building an efficient replica of Zhang Heng’s device in the modern era appeared to be challenging.

It was only in 1884 that a complete description of what is now called the earthquake cycle was made by Gilbert (1884). Gilbert sought to explain the formation of mountains and observed, in the Great Basin, California, the presence of "fractures" on the side of the mountains. He concluded that slow accumulation of strain over long time scales was responsible for the uprising of the mountains and that strain was relieved by sudden slip along these fractures, that he also called fault scarps.
He attributed earthquakes to these sudden slip phenomena, and set the basis for all modern descriptions of earthquake rupture by describing the cycle of accumulation and release as a stick-slip phenomenon due to friction on the fault.

Even though Gilbert’s theory did not offer a quantitative description, it introduced for the first time the concepts of a cycle with a long and a short time scales, earthquakes caused by sudden slip on a fault, and friction as the balancing force. Twenty-two years later, the 18 April 1906 San Francisco M7.9 earthquake struck and devastated the city. Triangulation surveys performed before and after the earthquake delivered, for the first time, a description of the deformation field caused at the surface by an earthquake. A few years after the 1906 event, based on these new observations of surface deformation, Reid published his elastic rebound theory, in which he described the accumulation of strain energy localized along faults and its sudden release in earthquakes (Reid, 1910).

The elastic rebound theory was still to be completed with a mathematical description of the wavefield radiated by the sudden slip on a fault, or an equivalent set of forces. The notion of double-couple first emerged among Japanese seismologists (Nakano, 1923, 1930; Honda, 1957) and became accepted worldwide as a successful model to describe the forces that act during faulting and generate the observed waves. Representation theorems were derived to provide mathematical formulas to compute the wavefield given a displacement discontinuity across a fault, or equivalently, a set of double-couples (e.g. Burridge and Knopoff, 1964; Aki and Richards, 1980). The understanding of the pattern of radiated waves from a given type of faulting gave, in turn, ways of interpreting seismograms to infer fault characteristics (cf. "beachballs", Stauder, 1962; Sykes, 1967).

The theoretical developments of the first half of the 20th century were limited by the lack of data to validate or disprove theories. Earthquake science, and seismology in general, entered its modern era when the first global networks of seismographs
were deployed in the early 1960s, notably with the Worldwide Standardized Seismic Network managed by the US Geological Survey. Since then, observations have become always more abundant and led to new discoveries about the complexity of the earthquake phenomenon (\textit{e.g.} non-volcanic tremors, \textit{Obara}, 2002).

### 1.1.2 Earthquakes as Sliding Instabilities

The previous section has remotely touched the topic of \textit{kinematic} descriptions of faulting by mentioning representation theorems, namely, descriptions that assume the slip distribution on the fault and compute the radiated wavefield. I will now present a \textit{dynamic} description of the full earthquake cycle, that is, a description given by solving the equations of motion (Newton’s second law) on the fault.

In his seminal paper (Gilbert, 1884), Gilbert described earthquakes as the fast slipping phase of a stick-slip phenomenon due to the competing tectonic and frictional forces on the fault. Stick-slip occurs in systems that alternate between a stable state (stick) and an unstable state (slip). Friction related phenomena have been known to exhibit such a stick-slip behavior. Early work by Amontons in 1699 (itself based on unpublished work by Leonardo da Vinci) and Coulomb in 1785 showed that:

(i) the frictional force \( F_f \) resisting the motion of a mass is proportional to the normal force \( F_n \) exerted on this mass (see Figure 1.1.1): \( F_f = \mu F_n \), where the coefficient of proportionality \( \mu \) is called the coefficient of friction.

(ii) If this same mass is stationary and pulled by a tangential force \( F_t \), this mass will remain stationary as long as \( F_t \leq \mu_s F_n \) and friction balances this force: \( F_f = F_t \). \( \mu_s \) is called the coefficient of static friction.

(iii) Once the mass has initiated motion because of \( F_f > \mu_s F_n \), the coefficient of friction suddenly drops to \( \mu_d \), the coefficient of dynamic friction.

If the coefficient of dynamic friction \( \mu_d \) is smaller than the coefficient of static
Figure 1.1.1: A mass sits on an inclined table and is subject to several forces: its weight $W$ causes the normal force $F_n$ due to the reaction of the table. The weight has a component that is tangential to the table $F_t$, and this force is resisted by friction $F_f$ between the mass and the table.

friction $\mu_s$, then the system is unstable: the force that was resisting right before the onset of motion, $F_f = \mu_s F_n$, suddenly drops to a lower value, $F_f = \mu_d F_n$, and the motion accelerates even more. If one were to incline the table shown in Figure 1.1.1, the normal force $F_n$ would gradually get smaller while the tangential force $F_t$ would be growing until overcoming $\mu_s F_n$. The mass would then slide all the way down the table. Once the motion stops, for example because the force that was driving the motion drops, the coefficient of friction starts increasing. This healing process was observed by Coulomb. Everyone has probably experienced the difference between static and dynamic friction, for example when moving a heavy piece of furniture on the floor: the initial effort to initiate motion is the most demanding because one has to overcome this "starting friction".

Thus, earthquakes can be described as frictional instabilities on faults. The simplest model of the earthquake cycle is the spring-slider model (cf. Figure 1.1.2A). A block rests on a surface and is coupled to a loading point through a spring with stiffness $k$. The loading point moves with the constant velocity $V_L$ and loads the block with the elastic force $k\delta$ due to the elongation $\delta$ of the spring. This elastic force is resisted by friction $F_f$ between the block and the surface, and therefore the equation
of motion is:

\[
\frac{d^2u}{dt^2} = k\delta - F_f; \quad \delta = V_L t - u. \tag{1.1}
\]

In Equation (1.1), \(\delta\) is the elongation of the spring and \(u\) is the position, or displacement, of the block.

![Diagram of the spring-slider model](image)

Figure 1.1.2: Spring-slider model. **A:** A block, the slider, lies on a surface and is coupled to a loading point moving at constant velocity \(V_L\) through a spring with stiffness \(k\). Thus, the slider is pulled by an elastic force \(k\delta\) where \(\delta\) is the elongation of the spring. This loading force is resisted by the friction \(F_f\). **B:** Numeric simulation of the spring-slider model with rate-and-state friction and aging law (cf. Equation (1.2)). The slider is stationary most of the time (inter-seismic phase, see zero slip rate, in solid red, and flat displacement curve, in solid blue), but catches up with the motion of the loading point (dashed blue line) in periodic, fast slip events. The stick-slip behavior ensures that the spring does not keep accumulating energy indefinitely. **C:** Focus on the fast slip event (the rupture) when the inertial term of the equations of motion cannot be neglected (left hand side of Equation (1.1)). All axes are non-dimensional quantities. Courtesy of Ekaterina Bolotskaya.

The block stays still until the elastic force overcomes the starting friction. If the dynamic friction is lower than the static friction, the onset of motion also marks the sudden acceleration of the block: this is the analog of an earthquake. Sliding stops
when the spring has shortened enough and then healing starts. This simple model
describes a stick-slip phenomenon due to the interplay between an elastic force and
friction, and has therefore been used to provide a simple mathematical description
of the earthquake cycle. The model also provides the means of investigating how the
specific form of the friction force $F_f$ influences the cycle.

An underdeveloped characteristic of friction in Coulomb’s work was the transition
from $\mu_s$ to $\mu_d$ as the mass slides. This transition controls the dynamics of the slip
event, and therefore the friction law describing this transition is essential. Various
friction laws have been proposed and all of them rely on at least one characteristic
length scale $D_c$ over which the transition occurs, possibly representing the smoothing
of the contact area between the block and the surface as sliding occurs (e.g. Dieterich,
1978; Ruina, 1983). The simplest law describes a linear drop from $\mu_s$ to $\mu_d$ as a func-
tion of slip $D$ and over the distance $D_c$. This law is called slip weakening.

More complex friction laws were proposed to describe laboratory experiments
more satisfactorily than slip weakening. Today, the most commonly used law is the
rate-and-state dependent friction, whose single state variable $\theta$ form is (Dieterich,
1992; Ruina, 1983):

$$
\mu(V, \theta) = \mu_0 + a \ln \left( \frac{V}{V^*} \right) + b \ln \left( \frac{\theta V^*}{D_c} \right).
$$

(1.2)

In Equation (1.2), $V$ is the slip rate, $\theta$ is the state variable, $a$ and $b$ are material-
dependent parameters that controls the dependence on slip rate and state variable,
as well as the stability of the system. $\mu_0$ is the reference coefficient of friction for
$V = V^*$ and $\theta = D_c/V^*$. Equation (1.2) needs to be completed with an equation that
describes the evolution of the state variable (aging law or slip law).

The success of rate-and-state friction lies in its ability to reproduce many exper-
imentally observed features such as slip instability, and healing, slip rate dependent
steady states (Dieterich, 2007). Tuning the values of $a$ and $b$, rate-and-state friction can also describe stable sliding ($\mu_d > \mu_s$, slip strengthening). Thus, rate-and-state friction elegantly explains the depth distribution of earthquakes by showing that seismogenic depths coincide with the depth range where values of $a$ and $b$ are such that frictional instability is possible (see Figure 1.1.3 and Scholz, 1998).

Figure 1.1.3: Stability domains predicted by rate-and-state friction (Equation (1.2)). The middle panel shows a measure of stability, $\zeta$, as a function of depth. These stability domains are sketched for a subduction interface (left) and a crustal fault (right). The rightmost panel shows the depth distribution of earthquakes for a section of the San Andreas fault near Parkfield. Figure from Scholz (1998).

Despite the success of rate-and-state friction in modeling friction on faults and laboratory samples, the parameters suffer from a lack of interpretability. Just as entropy was unexplained before the era of statistical mechanics, rate-and-state friction cannot be currently derived from fundamental physics. Furthermore, there is little certainty about how $a$, $b$, and $D_c$ scale from the laboratory to Earth. These drawbacks actually apply to other friction laws. In fact, it seems that any law that describes an unstable slip event would provide a good description of the dynamic rupture, given an appropriate set of parameters. Work remains to be done to understand the forces at play in earthquakes. I purposely entitled this section "Earthquake as Sliding Instabilities" to emphasize that the essence of the current description of earthquake
dynamics is not friction but instability.

The spring-slider model is useful to capture first order features of the earthquake cycle and dynamics, and multiple interacting sliders can even be considered to reproduce statistical features of seismicity (Burridge and Knopoff, 1967, and see Section 1.3). However, a 0-D description (a point) of an earthquake is obviously an oversimplification that fails to describe important aspects of the earthquake dynamics, such as the fault geometry and the influence of radiated waves.

1.1.3 Scaling Laws in Earthquakes

The spring-slider model taught us that sliding events occur to restore the elongation of the spring $\delta$ to zero. Equivalently, these sliding events restore the loading force exerted on the slider to zero. The drop in loading force $\Delta T$ during an event is set by the elongation of the spring just before the onset of sliding $\delta_0$: $\Delta T = k\delta_0$. It appears that $\delta_0$ is also the slip $D$ of the sliding event. If $A$ is the contact area between the slider and the support, then the static stress drop $\Delta \sigma$ of a sliding event is $\Delta \sigma = kD/A$. Thus, we have shown that slip is proportional to stress drop, and it can also be shown that the maximum slip rate and acceleration are proportional to stress drop too (Nur, 1978; Scholz, 2019, Chapter 2). Stress drop thus appears to be a fundamental parameter to describe earthquake processes.

Unlike the spring-slider model, real earthquakes extend over 2-D surfaces and another parameter that describes the size of an earthquake is useful: the seismic moment $M_0$. The seismic moment $M_0$ and the stress drop $\Delta \sigma$ are defined by:

$$M_0 = GA\bar{D},$$  \hspace{1cm} (1.3)

$$\Delta \sigma = CG\frac{\bar{D}}{L}.$$  \hspace{1cm} (1.4)
In Equations (1.3) and (1.4), $G$ is the shear modulus (also called rigidity $\mu$, in Pa), $A$ is the fault area ($m^2$), $L$ is the linear dimension of the fault (m), $\bar{D}$ is the average slip on the fault (m), and $C$ is a non-dimensional shape factor ($\sim 1$). Using Equation (1.4) to express $G\bar{D}$ as a function of $\Delta \sigma$, we can write explicitly the relation between seismic moment and stress drop:

$$M_0 = \frac{1}{C} \Delta \sigma AL.$$  \hspace{1cm} (1.5)

Estimates of seismic moment and fault area have shown a remarkable linear relation between $\log A$ and $\log M_0$ indicating $A \propto M_0^{2/3}$ or, equivalently, $M_0 \propto A^{3/2}$ (Kanamori and Anderson, 1975). This scaling is interpreted as the scale invariance of stress drop $\Delta \sigma$, and $L \propto A^{1/2}$. Equivalently, $M_0 \propto L^3$. The latter implies that the shape ratio of faults is constant, which seems true for most but the largest earthquakes (Denolle and Shearer, 2016).

Figure 1.1.4: Fault area $A$ against seismic moment $M_0$. The slope in the log-log domain is 2/3, indicating that $A \propto M_0^{2/3}$. Lines of equal stress drops for circular cracks are shown. Figure from (Kanamori and Anderson, 1975).

The scale invariance of stress drop is a central result in earthquake seismology and suggests the similarity between small and large earthquakes, or self-similarity. The
concept of self-similar earthquakes was first introduced by Aki (Aki, 1967) to develop a model of earthquake amplitude spectra defined by a single parameter: the linear fault dimension $L$. Self-similarity also implies a constant rupture velocity $v_r$ among earthquakes. In turn, these predict that the rupture duration $\tau$ is proportional to the linear dimension $L$: $\tau \propto L$. Aki showed that his model of amplitude spectrum could be used to estimate the rupture duration from the spectrum’s corner frequency, $f_c \propto 1/\tau$, and produced observations of $\tau \propto L$. Additional scaling observations confirmed that the approximation of constant rupture velocity was reasonable (e.g. $M_0 \propto \tau^3$, Kanamori and Brodsky, 2004).

Following the pioneering work of Aki (1967), amplitude spectrum corner frequencies are now commonly used to estimate the source dimension. Results over a wide range of magnitudes ($M_w$0-9) indicate that the corner frequency scales as $f_c \propto M_0^{-1/3}$ (see Figure 1.1.5, Allmann and Shearer, 2009). Scatter in the data is explained by differences in stress drop, all events showing stress drops between 0.1 MPa and 100 MPa. This is a strong evidence for the scale invariance of stress drops. The scaling laws mentioned in this section are summarized in Table 1.1.

<table>
<thead>
<tr>
<th>Source Parameters</th>
<th>Reference</th>
<th>Scaling Relation</th>
</tr>
</thead>
<tbody>
<tr>
<td>seismic moment $M_0$, fault area $A$</td>
<td>Kanamori and Anderson (1975)</td>
<td>$M_0 \propto A^{3/2}$</td>
</tr>
<tr>
<td>seismic moment $M_0$, source duration $\tau$</td>
<td>Kanamori and Brodsky (2004)</td>
<td>$M_0 \propto \tau^4$</td>
</tr>
<tr>
<td>seismic moment $M_0$, linear dimension $L$</td>
<td>Denolle and Shearer (2016)</td>
<td>$M_0 \propto L^3$</td>
</tr>
<tr>
<td>corner frequency $f_c$, linear dimension $L$</td>
<td>Aki (1967)</td>
<td>$f_c \propto 1/L$</td>
</tr>
<tr>
<td>seismic moment $M_0$, corner frequency $f_c$</td>
<td>Allmann and Shearer (2009)</td>
<td>$M_0 \propto f_c^{-3}$</td>
</tr>
</tbody>
</table>

Table 1.1: Summary of scaling relations between source parameters. N.B.: These relations only hold for earthquakes that do not saturate the seismogenic width.

It is known and well understood that large earthquakes are not similar to small ones in the sense that all scaling laws assuming $A \propto L^2$ break down when earthquakes saturate the seismogenic width. In this case, an earthquake can only grow laterally and the fault area becomes (asymptotically) proportional to its lateral length $A \propto L$. The scaling laws in Table 1.1 need to be changed accordingly (Denolle and Shearer,
Figure 1.1.5: Amplitude spectrum corner frequency $f_c$ against seismic moment $M_0$ (the corresponding moment magnitude $M_w$ is given at the top). The red dashed lines shown the constant stress drop scaling. Figure from Allmann and Shearer (2009).

Despite the dependence of shape ratio on magnitude for large earthquakes, which implies the break of self-similarity, the question remains whether small and large earthquakes are dynamically similar. A parameter that I have not discussed so far, the ratio of energy radiated in seismic waves $E_R$ to the seismic moment $M_0$, called scaled energy $\tilde{e} = E_R/M_0$, informs about the dynamics of earthquakes. Mitigated results sometimes show scale invariant $\tilde{e}$ (Baltay et al., 2011; Denolle and Shearer, 2016), but others suggest a systematic dependence on magnitude (Abercrombie, 1995). Accurate measures of $E_R$ for small earthquakes is difficult due to noise. The scaling of $\tilde{e}$ is key for earthquake forecasting and early warning as a scale independent $\tilde{e}$ would mean that the earthquake initiation process is scale invariant and that nothing can tell if an earthquake is going to grow into a large rupture when observing the first P-wave arrival at a seismometer.

Even though clear scaling relationships between source parameters are observed
(see Figures 1.1.4, 1.1.5), details seem to be obscured by their simplicity. Breakdown of self-similarity for smallest size earthquakes have long been sought as it is thought that there exists a lower limit to earthquake sizes (the nucleation length), which would break scale invariance (Cattania, 2020). Proving departures from these laws is difficult as observational limits are significant. Estimating a correct corner frequency from amplitude spectra strongly relies on the correction of path effects that corrupt the source information in recordings, and on instruments with sufficiently large frequency bandwidths (Abercrombie, 1995; Ide and Beroza, 2001; Abercrombie, 2015, 2021). However, several studies have produced results that show (maybe apparent) deviations from self-similarity (Lin et al., 2016; Farge et al., 2020, Chapter 3 of this thesis).

1.2 The Earthquake Detection Problem

1.2.1 Motivations

Seismology is largely an observational science. The introduction to the earthquake phenomenon given in the previous section showed that earthquake seismology would not exist without data. Nowadays, field observations are completed by the abundant recordings of the ground motion on seismometers. In order to use the recordings of earthquake signals to study the seismic source or Earth’s interior, one evidently needs to find these recordings that are associated with a particular event. The earthquake detection problem consists of determining when a seismometer, or an array of seismometers, is recording an earthquake signal, as opposed to ambient seismic noise travelling in the earth. With earthquake detection methods, earthquake signals are identified in recordings and their origin times are stored in a catalog. Note that the detection problem is often solved together with the location problem so that both time and space information is available for subsequent studies.

The most trivial detection method is certainly that of looking with the naked
eye through continuous recordings of the ground motion for signals with earthquake characteristics: impulsive arrival, multiple phases (e.g. the P- and then the S-wave arrivals), coda wave, etc. While still routinely used in observatories because it ensures making (hopefully) good quality catalogs with picks of P and S arrival times, this method is limited by human capacities. An analyst is likely to miss events when working for a long time, can make errors when picking the P and S arrival times, and typically fails at identifying earthquakes when the signals are weak and hidden in the ambient noise. Most importantly, the capacity of an analyst to combine information from multiple seismograms at once is limited. The notion of signal visibility, or signal-to-noise ratio (SNR), is fundamental in the earthquake detection problem. Furthermore, since all humans work slightly differently, the catalogs based on visual picks are noisy and uncertainty in individual picks is hard to know.

The deployment of large scale arrays of seismometers in the 1960s, such as the Large Aperture Seismic Array (LASA) in Montana and the Norwegian Seismic Array (NORSAR) in Norway, fed the rapid development of array detection methods. Early detection methods relied on heavy mathematics from the statistical theory of signal detection (Vanderkulk et al., 1965; Capon, 1970). The base of array methods is to sum together potentially noisy but coherent signals from multiple seismometers. The noise, if uncorrelated between seismometers, will interfere destructively when summing. Thus, array methods provide the means for detecting weak earthquake signals with low SNR but recorded at multiple locations. Such array methods were motivated not only by the will to improve our understanding of earthquakes and Earth’s interior, but also by the necessity of monitoring nuclear tests during the Cold War. Therefore, earthquake detection methods were also a widely documented topic in the Soviet scientific literature, and later English translations made it available to a wider audience (Kushnir et al., 1992).
1.2.2 A Brief Tour from Beamforming to Artificial Intelligence

To sum coherently earthquake signals observed at different seismometers with different time delays, one needs to time-shift the recordings according to the relative time delays prior to summation. If a plane wave with known incidence is recorded by multiple seismometers at the surface (cf. Figure 1.2.1A), then the relative time delays at which the wave is recorded at each seismometer can be computed given some knowledge of the wave velocity. These time delays are used to shift and sum the recordings $u_i(t)$ collected by the seismometers in order to form a beam $b(t)$:

$$b(t) = \sum_i u_i(t - \tau_i). \quad (1.6)$$

In Equation (1.6), summation is over the seismograms indexed by $i$, and $\tau_i$ is the relative time delay computed at seismometer $i$. The beam offers obvious advantages over a naive summation of all seismograms, and may make earthquake signals hidden in noise become visible (see Figure 1.2.1B, and Rost and Thomas, 2002). Of course, the incidence of the target plane wave is in general not known, and one may build many beams over a grid of possible incidences, and retain only the beam that gives the largest sum. This method of shifting and summing seismograms by assuming different possible plane wave incidences is called beamforming.

The beamforming method described by Equation (1.6) is naive in that it assumes that, once correctly shifted in time, seismograms will sum constructively. Due to the radiation pattern of earthquakes, the sum might in fact be zero because of opposite wave polarities. Seismologists have introduced the use of characteristic functions in place of the raw seismograms to better serve the detection purposes. Such a characteristic function could simply be the absolute value, in order to avoid the destructive summation of opposite polarity waves. A popular characteristic function measures the ratio between the short term average to the long term average seismic energy (STA/LTA Freiberger, 1963; Allen, 1978). The STA/LTA method enhances the onset of impulsive signals and adapts to changes in the background noise (e.g. night vs
day time). Various characteristic functions have been used (signal envelope, higher order statistical moments, etc). Moreover, the assumption of plane waves can be relaxed and travel times can be computed in a more realistic Earth. This generalized beamforming has been successfully applied in different geological contexts to detect various kinds of seismic events (Frank et al., 2014; Poiata et al., 2016, and Chapter 2 of this thesis).

The use of characteristic functions partly aims at compensating for incoherency emerging from directivity and propagation effects. In fact, these effects are responsible for most of an earthquake signal’s complexity. They can be fully accounted for by using a given earthquake as a template pattern to look for all events with similar signals and time delays. This similarity based method, called template matching (Gibbons and Ringdal, 2006; Ross et al., 2019, Chapters 2, 4 and Appendix A of this thesis), is highly effective at detecting earthquakes originating from similar location and faulting (i.e. same directivity and propagation effects). The main drawback of the method is that it relies on some prior knowledge of the target seismicity, and does not generalize to earthquakes that are not located in the proximity of the template.
earthquakes. Template matching plays a key role in this thesis, and is therefore discussed at length in the following chapters.

Recently, the field of earthquake detection has undergone profound changes with the rise of statistical learning methods (Bergen et al., 2019). The large amount of information stored in earthquake catalogs made by the past generations of seismologists with classic techniques, such as those mentioned above, provides all the data necessary to train complex models to detect earthquakes (but also phase picking and association, location and others Zhu and Beroza, 2019; Mousavi et al., 2020; Majstorović et al., 2021). Deep learning models, sometimes referred to as artificial intelligence, have proven particularly efficient at this task and recent models have shown good abilities at generalizing to multiple data sets (for example, see the use of PhaseNet in Chapter 4 of this thesis Zhu and Beroza, 2019). These models will certainly become the standard usage in the next years.

1.3 Collective Properties of Earthquakes

Even though the scaling laws described earlier in Section 1.1.3 are visible when studying groups of earthquakes, they still only relate properties of single earthquakes. Here, I give an introduction to important collective properties of earthquakes, i.e. properties that emerge when several earthquakes interact. Observation and discussion of earthquake interactions is a central topic of this thesis.

1.3.1 The Gutenberg-Richter Law

A well-known scaling law that was left aside in Section 1.1.3 is the frequency-magnitude relation, better known as the Gutenberg-Richter law (Gutenberg and Richter, 1941). This law describes the expected number of earthquakes $N(M)$ with magnitude $m$
greater than a certain value $M$ within a given period of observation (see Figure 1.3.1):

$$\log N(M) = a - b \log M.$$  \hspace{1cm} (1.7)

In Equation (1.7), $a$ and $b$ are constants, and are not related to the $a$ and $b$ parameters mentioned in the rate-and-state friction law (Section 1.1.2). $a$ is the total number of earthquakes within the space-time region of observation, and $b$, originally called the b-value, characterizes the relative frequency between small and large magnitude earthquakes.

![Figure 1.3.1: Frequency-magnitude relationship, referred to as the Gutenberg-Richter law (cf. Equation (1.7)). The slope in the log-log domain indicates a b-value $b = 1$. Figure 23 from Kanamori and Brodsky (2004), see their caption for description of the data.](image)

Worldwide observations of the Gutenberg-Richter law show b-values $b = 1$ (cf. Figure 1.3.1, Kanamori and Brodsky, 2004). This first tells us that there is a ten-fold increase in $N(M)$ every time $M$ decreases by one unit. For example, every year we can expect to observe 10 times more $M \geq 7$ than $M \geq 8$. Deeper implications come when this law is interpreted. Such a power law frequency-size distribution is intimately related to the concepts of criticality and fractals (Turcotte, 1989). Aki (1981) proposed an explanation of the Gutenberg-Richter law with a fractal distribution of fault sizes, assuming the scaling laws of self-similarity (see Table 1.1). The cascading of events among these faults of all scales produce the observed frequency-magnitude
distribution. Given that critical phenomena are characterized by local interactions that are able to build up into long range interactions, other models have tried, and succeeded, to reproduce the Gutenberg-Richter law with many simple elements interacting with their neighbors (Burridge and Knopoff, 1967; Bak and Tang, 1989). Thus, the Gutenberg-Richter law is understood as a collective phenomenon that would not exist without interactions.

1.3.2 Cascades, Interactions, and Temporal Clustering

Cascades of earthquakes have long been observed as sequences of aftershocks following sizable earthquakes. The first mathematical description of the enhanced seismic activity following a large earthquake appeared in the late 19th century with the Omori law \( n(t) = K(t + c)^{-1} \) (Omori, 1894), which was later generalized by Utsu as (Utsu, 1961, 2002):

\[
n(t) = \frac{K}{(t + c)^p}.
\]  

In Equation (1.8), \( t \) is the time measured from the reference event called the main-shock, \( n(t) \) is the number of earthquakes per unit time, and \( K, c \) are constants. The power law exponent \( p \) describes how fast the seismicity rate drops after the main-shock, and is usually around 1 (as stated in the original law).

Cascades of earthquakes such as those observed with aftershock sequences are now understood as resulting from interactions between faults or subfaults (Burridge and Knopoff, 1967). A simple model that helps understand triggering through interaction is the Coulomb stress model, which is closely related to the frictional instability discussed in Section 1.1.2. We saw that the onset of motion of a block occurred when the tangential force exceeded a threshold set by the coefficient of static friction \( \mu_s \), which constitutes the Coulomb failure criterion. Thus, the proximity of a fault to the
Coulomb failure criterion can be quantified by the Coulomb stress $C$:

$$C = \tau - \mu_s \sigma_n,$$  (1.9)

where $\tau$ is the shear stress resolved on the direction of slip on the fault, and $\sigma_n$ is the normal stress on the fault (including effects such as pore-fluid pressure). If $C > 0$, then the failure criterion is satisfied and the fault slips. The change in Coulomb stress $\Delta C$ produced by an outside perturbation, such as a nearby earthquake, quantifies whether this perturbation took the fault closer or further to failure (King and Cocco, 2001):

$$\Delta C = \Delta \tau - \mu_s \Delta \sigma_n.$$  (1.10)

The $\Delta$ symbol indicates the difference between after and before the stress perturbation. If the Coulomb stress change is larger than zero, $\Delta C > 0$, then the perturbation promoted failure on the fault, or otherwise inhibited the failure if $\Delta C < 0$. This method has successfully explained the patterns of moderate-to-large earthquakes in terms of stress perturbations due to past earthquakes (King and Cocco, 2001; Stein et al., 1997, and see Figure 1.3.2), and is useful to understand earthquake triggering in terms of stress transfer. However, the assumptions it relies on may limit its applicability in practice. When computing the Coulomb stress on a fault, one needs to know its orientation in order to correctly decompose stress into its normal and tangential components. Faults may have complex geometries, and the Coulomb stress change is often simply computed on "optimally oriented" faults. Furthermore, the tangential component is taken along an a priori slip direction (e.g. same direction as the regional sense of motion), but faults, particularly at small scales, may slip in different directions depending on how they are stressed.

The static stress changes involved in the Coulomb failure criterion are not the only way earthquakes can interact. It has been observed that dynamic stress changes caused by elastic waves can trigger earthquakes (Fan and Shearer, 2016), and that the disturbance of a stably sliding fault by a nearby earthquake could, in turn, trigger...
other earthquakes located in the vicinity (Dublanchet et al., 2013; Cattania, 2019). Any process caused by an earthquake that perturbs the stress field plays a role in earthquake interaction. All these mechanisms have in common that they alter the timings of fault ruptures, and, consequently, the patterns of earthquake occurrence seen in earthquake catalogs.

Triggering of earthquakes by other earthquakes has been observed in earthquake catalogs as deviations from a purely random earthquake occurrence typically described by a Poisson law (Gardner and Knopoff, 1974). The cascading of events due to earthquake interactions constitutes the core property of time clustered earthquake occurrence. In this thesis, temporal clustering refers exclusively to these cascade-like sequences. The most represented manifestation of temporal clustering is certainly the production of aftershocks after a large earthquake. It is now worth mentioning
that, given mechanisms such as static stress transfers, we now know that any earthquake produces its own sequence of aftershocks, although aftershock productivity is higher for larger earthquakes (Marsan and Lengline, 2008). Such time clustered sequences exhibit time scale invariance as the distribution of inter-event times follow a power law (up to a certain time scale when scale invariance breaks). In the case of aftershock sequences, we actually expect such a power law distribution from the generalized Omori law (Equation (1.8)). Because of this time scale invariance, time clustered sequences obey fractal patterns (Smalley Jr et al., 1987, and this thesis), which sometimes is interpreted as cascades in a critically stressed Earth (Main, 1995).

Several methods exist to characterize the clustering of earthquakes. Historically, the focus has been on analyzing the distribution of inter-event times with a single scalar: the ratio between the standard deviation and the average of this distribution, named the coefficient of variation $C_v$ (Kagan and Jackson, 1991):

$$C_v = \begin{cases} 
0 & \text{periodic seismicity} \\
1 & \text{poissonian seismicity} \\
> 1 & \text{clustered seismicity}. 
\end{cases} \quad (1.11)$$

Equation (1.11) can be interpreted as follows: for periodic seismicity, $C_v = 0$, the probability of an earthquake is larger after a period of quiescence, while for clustered seismicity, $C_v > 1$, this probability is lower than would be for poissonian seismicity. Other methods lie on the time scale invariance of clustered seismicity and look for strong power law or fractal characteristics in earthquake occurrence. The spectrum of the event number per unit time of a clustered sequence shows a power law dependence on frequency, $\propto f^{-\beta}$, and the power law exponent $\beta$ measures the strength of clustering (cf. Figure 1.3.3, Frank et al., 2016). The fractality of the event number per unit time becomes apparent when observing the power law between the fraction $x$ of time bins with duration $\tau$ that are occupied by at least one event and the bin
duration $\tau$ \citep[cf. Figure 1.3.3,][]{SmalleyJr1987}:

$$x \propto \tau^{1-F}. \quad (1.12)$$

In Equation (1.12), the power law exponent $1 - F$ involves the fractal dimension $F$ of the time series \citep{LowenTeich2005}. The fractal dimension $F$ of the event count per unit time is extensively used to characterize temporal clustering in this thesis. Note that the fractal dimension is called $D$ in the following chapters, which should not be confused with slip on the fault $D$ mentioned in Sections 1.1.2 and 1.1.3.

Figure 1.3.3: Quantification of the strength of earthquake clustering. An example of a weakly (blue) and strongly (green) time clustered sequences are given. \textbf{A:} Event count number, \textit{i.e.} the number of events per unit time. \textbf{B:} Autocorrelation of the event count number. The weakly clustered sequence shows the characteristic delta autocorrelation of random sequences, whereas the strongly clustered sequence keeps a high correlation over a long time scale. \textbf{C:} Measure of the power law exponent of the event count spectrum. Clustered sequences exhibit large power law exponents. \textbf{D:} Fraction $x$ of time bins with size $\tau$ occupied by at least one earthquake: $x \propto \tau^{1-F}$. Clustered sequences have large fractal dimensions $F$. Figure from \citep{Beaucé2019}.
Work has been done on understanding the topology of earthquake clusters either in the time-space domain (Frohlich and Davis, 1990), or in the energy-time-space domain (Zaliapin et al., 2008). These studies involve defining distances in a certain domain and analyzing the distances between nearest neighbors. Earthquake sequences attributed to different types of seismicity (e.g. swarm-like vs. cascade-like) and geological contexts were identified based on their topological differences (Zaliapin and Ben-Zion, 2013; Martínez-Garzón et al., 2019). While powerful descriptors of clustering processes, the products of such methods are also more difficult to analyze than a simple scalar such as $C_v$ or $D$, which we preferred in this thesis.

1.4 Goals and Structure of the Thesis

This chapter has introduced useful concepts for understanding the motivations and conclusions of my thesis. As the thesis’ title suggests, my generic goals were to:

- produce new observations of seismicity with automated earthquake detection and location methods,
- systematically characterize faults with the time clustering property of their seismicity,
- advance the understanding of how clustering relates to fault mechanisms, in particular with aseismic slip.

My investigations targeted two regions: the Alps, in Europe, and the western section of the North Anatolian Fault, in Turkey. Because of the history of the North Anatolian Fault, my work was deeply related to understanding how the last major rupture in 1999 had perturbed the fault system.

Chapter 2 "Systematic Detection of Clustered Seismicity Beneath the Southwestern Alps" presents a detailed study of earthquake detection and location in the Southwestern Alps, describing in detail the method, and shows the first observations of tem-
poral clustering of the thesis. Time clustering properties, or its absence, are discussed along other geological evidence to relate them to fault properties and environmental factors.

Chapter 3 "The Rupture Complexity of Small Earthquakes in the Southwestern Alps" pushes further the analysis of the Southwestern Alps earthquake catalog (Beaucé et al., 2019) by analyzing earthquake spectra via the spectral ratio method to characterize earthquake sources. Spectra’s corner frequencies and the seismic moments are estimated, and variations in the \( f_c \propto M_0^{-\alpha} \) scaling relationship are investigated. Evidence of apparent departures from self-similarity is shown and discussed in terms of source physics, and related to the observations of temporal clustering made in Chapter 2.

Chapter 4 "Seismotectonic Study of the North Anatolian Fault Zone Thirteen Years After the 1999 M7.4 Izmit Earthquake" presents an extensive study of the seismicity in the western section of the North Anatolian Fault, where the Izmit earthquake propagated (Beaucé et al., 2021b). The earthquake detection and location method from Chapter 2 was improved and applied to dense array data. In addition to the space-time catalog and quantification of temporal clustering, this study produces a large data set of focal mechanisms and use them in an inversion scheme to estimate the stress tensor in several areas. The inversion method is the topic of Appendix B (Beaucé et al., 2021a). Results are discussed altogether to draw conclusions about the state of the fault with respect to before and shortly after the Izmit earthquake. A major finding of this study is that highly clustered seismicity tends to occur at the bottom of the seismogenic zone and is likely to be associated with normal faults slipping aseismically.

Chapter 5 "Conclusions and Perspectives" summarizes the results of this thesis and the contributions to advancing the interpretation of earthquake clustering. Quantification of clustering is a valuable seismic observable that we believe will help
characterize the state of the crust in future studies. Finally, perspectives are given on extending the systematic observation of clustered seismicity to longer times and interesting fault systems, and on the inclusion of geodetic data into seismological studies (preliminary results are presented).

Appendix A "Fast Matched Filter (FMF): An Efficient Seismic Matched-Filter Search for Both CPU and GPU Architectures" describes the high performance computing software for template matching that I co-wrote with William B. Frank (Beaucé et al., 2018). The success of the algorithm lies in its GPU implementation that allows speeds far greater than any CPU architecture. FMF has already emerged as a popular template matching software in the community.

Appendix B "An Iterative Linear Method for Estimating the Stress Tensor from Earthquake Focal Mechanism Data: Method and Examples" describes the new methodology we have developed to invert focal mechanisms for stress tensor (Beaucé et al., 2021a). This work accompanies the results presented in Chapter 4. The method takes a novel approach to combine the efficiency of the classic linear inversion and an iterative scheme that relaxes a constraining assumption. We have released a Python package that fully implements the proposed method.

During my time as a student at MIT, I have put particular efforts into developing open source numerical tools that are available to the community, and I highly encourage the interested reader to check out my Github account at https://github.com/ebeauce and ask questions if needed.
Bibliography


J. Dieterich. Applications of rate-and state-dependent friction to models of fault slip and earthquake occurrence. 2007.


Chapter 2

Systematic Detection of Clustered Seismicity Beneath the Southwestern Alps

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.

2.1 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 and Creager, 2011; Shelly et al., 2011; Frank et al., 2014), constructing reference earth models (e.g. Dziewonski and Anderson, 1981; Kennett and Engdahl, 1991; Kennett et al., 1995), seismic tomography (e.g. Dziewonski and 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 and 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 and 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...
We develop an earthquake detection method that combines array processing, or, more precisely, a beamformed network response (Frank and Shapiro, 2014) and template matching (Gibbons and Ringdal, 2006; Shelly et al., 2007; Frank and 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 (Zhao et al., 2016, cf.>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 (Nocquet, 2012, cf.>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 2.1.1). 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 2.1.1: 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.
2.2 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.

2.2.1 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 (Zhao et al., 2016, China-Italy-France Alps survey; 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.

2.2.2 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 2.2.1, 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 2.2.1 and the processing of the data is illustrated in Figure 2.S.1. 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 2.2.1A). 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 2.2.1B). We define the stack of the shifted envelopes as the network response:

$$NR_k(t) = \sum_{s,c} f(u_{s,c}(t + \tau_{s,c}^k)) .$$  \hspace{1cm} (2.1)

In Equation 2.1, $k$ 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; $\tau_{s,c}^k$ 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 2.S.1). 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 and 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
Figure 2.2.1: **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 2.1). 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 2.2) 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 × 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.
(CNR) as:
\[ \text{CNR}(t) = \max_k \{ \text{NR}_k(t) \} = \text{NR}_{k^*}. \] (2.2)

The process of searching for \( \text{NR}_{k^*} \), continuously in time, is illustrated in Figure 2.2.3.

Figure 2.2.1C 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:

\[ \text{threshold}(t) = \text{median} \left( \text{CNR}(t) \right) + 10 \times \text{MAD} \left( \text{CNR}(t) \right), \] (2.3)

where MAD stands for median absolute deviation. We evaluate \( \text{median} \left( \text{CNR} \right) + 10 \times \text{MAD} \left( \text{CNR} \right) \) 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 2.3, 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.

### 2.2.3 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
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 2.2.2. For each three-component record extracted from the 20 stations, we calculate five features:

- the amplitude maximum,

- the first three statistical moments of the distribution of the peaks of the waveform autocorrelation function: variance, skewness, and kurtosis,

- 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 2.3.1, 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 × 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 gradi-
ent 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.

2.2.4 Template Matching

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

\[ u(r, t) = S(t)M(r; \xi) * G(r, t; \xi) * I(t) . \]  

(2.4)

In Equation 2.4, 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 loca-
tion is \( \xi \). Equation 2.4 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.

Template matching consists of scanning continuous recordings in search for matches
between data and the waveforms that constitute a template. This method has proven
Figure 2.2.3: **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*).

... to be efficient at detecting events with low SNR (SNR < 1, e.g. Gibbons and 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:

\[
CC(t) = \sum_{s,c} w_{s,c} \frac{1}{N} \sum_{n=1}^{N} \frac{\sum_{n=1}^{N} T_{s,c}(t_n) u_{s,c}(t + t_n + \tau_{s,c})}{\sqrt{\sum_{n=1}^{N} T_{s,c}^2(t_n) \sum_{n=1}^{N} u_{s,c}^2(t + t_n + \tau_{s,c})}}. \tag{2.5}
\]

In Equation 2.5, \(N\) 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 2.5 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 2.3, 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 \(8 \times \text{MAD}\) (Shelly et al., 2007; Brown et al., 2008; Baratin et al., 2018, e.g. \(8 \times \text{RMS} \approx 12 \times \text{MAD}\)).

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 2.2.3. 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.

2.2.5 Second Generation Templates

As illustrated in Figure 2.2.3, 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 2.8.4 illustrates the performance of these different stacking strategies.

Detection and location involve finding the optimal network response for a given \( f \) in Equation 2.1. 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 2.2.4A (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 2.2.4, more details in 2.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 simi-
lar 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.

2.3 Seismicity of the Southwestern Alps

We applied the earthquake detection method presented in Section 2.2 – 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 2.2.1.

2.3.1 Catalog

Calculating the CNR as described in Section 2.2.2 yielded a total of 50,262 detections
(candidate template events). After applying the classifier described in Section 2.2.3,
we were left with 1,725 template events. We further reduced this number to 1,406
by regrouping redundant template events (cf. Section 2.2.5); Figure 2.3.1 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

91
Figure 2.2.4: 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. 2.A for details). This weighting function is used to calculate a weighted average of the distance to the best potential seismic source (*cf. Equation 2.8 in 2.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.

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).

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
Figure 2.3.1: 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 2.2.5 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 $M_L 3.9$ earthquake that occurred in early October 2012, and which is important for the discussion in Section 2.4. 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.

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 2.5 and 2.6. The events that we seem to have missed all have magnitude less than one and most less than 0.4 (cf. Figure 2.7), 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 2.3.2A 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 2.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 2.S.8). 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 2.3.2B shows the earthquake temporal distribution on recurrence time versus detection time graphs for three templates in distinct geographical regions: the Ubaye valley, the Briançonnais and the Dora Maira massif (cf. locations in Figure 2.1.1). 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 2.3.2B).

2.3.2 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
Figure 2.3.2: **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 2.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.

for sequences of foreshocks-mainshock-aftershocks (*e.g.* Utsu, 1961; Knopoff, 1964; Gardner and Knopoff, 1974; Zaliapin and Ben-Zion, 2013b) and is thought to be the signature of stress redistribution on neighboring faults taking place during the seismic rupture (*e.g.* Burridge and 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 and Ben-Zion, 2013a), 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 2.3.3A), 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 2.3.3B 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 ($\bar{e}(f) \propto f^{-\beta}$, similar to Frank et al., 2016). The strength of temporal clustering is quantified by $\beta$, referred to as clustering coefficient, which can be estimated from the slope of the spectrum in log-log space (Figure 2.3.3C). A strongly clustered earthquake sequence has a large $\beta$ whereas an earthquake sequence close to a Poisson sequence has a small $\beta$, 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 $\bar{e}(f) \propto f^{-\beta}$ 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 and Teich, 2005). Fractal analysis, which has been applied to earthquake occurrence in various studies (e.g. Smalley Jr et al., 1987; Lee and 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 $\tau$, 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 2.3.3D). We used correlation time, clustering coefficient $\beta$ 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 2.S.9 and Figure 2.S.10).
present our observations of temporal clustering in Figure 2.4.1.

Figure 2.3.3: 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 $\tau$ 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 $\beta$ 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 2.4.1A and Figure 2.4.1B 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 2.1.1). 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 2.S.9). The Ubaye valley is known to host a seismic swarm (e.g. Jeannot 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 and Ben-Zion, 2013b).

2.4 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 2.3.3. 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 obser-
vations 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 2.4.1: Cross-section along the CIFALPS axis showing 976 templates that were well relocated (∆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 $M_L 3.9$ earthquake that we mention in the discussion (Section 2.4). The red structures are reported from the geological cross-section in Figure 2.1.1.

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 2.4.1B)
are located around the $M_L 3.9$ earthquake that occurred on October 3rd, 2012 (cf. location in Figures 2.3.1 and 2.4.1, cf. our catalog for the local magnitude). This highly clustered seismicity took place over about four days (see Figure 2.3.2A), and can be seen as a sequence of foreshocks-mainshock-aftershocks. The locations shown in Figure 2.4.1 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, \textit{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 (\textit{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 (\textit{e.g.} Lohman and 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.

2.5 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 2.2.3), and we construct low noise templates by combining the singular value decomposition Wiener filter (SVDWF) with subsequent stacking (Section 2.2.5).

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_L$3.9 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 seismic-
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 2.3.1 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 2.3.1 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.

Acknowledgements

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 RvdH’s Schlumberger chair and by the European Research Council (ERC) under the European Union’s Horizon 2020 Research and Innovation Program (grant agreement No 789742335, F–IMAGE).

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

2.A Template Relocation

The weighting function presented in Figure 2.2.4 is defined by:
\[ w(t_n) = \begin{cases} 
A \exp \left( -\frac{(\text{CNR}(t_n) - \text{CNR}_{\text{max}})^2}{4\sigma} \right) & \text{if } t_n \in \mathcal{V}(t_{\text{max}}), \\
0 & \text{otherwise.} 
\end{cases} \quad (2.6) \]

In Equation 2.6, the neighborhood \( \mathcal{V}(t_{\text{max}}) \) is defined by:

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

and \( t_{\text{max}} = \arg\max_{t_n} \text{CNR} \). \( A \) is a normalization factor such that \( \sum_{n=1}^{N} w(t_n) = 1 \), and \( \sigma \) is the standard deviation of the CNR within \( \mathcal{V}(t_{\text{max}}) \).

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

\[ \Delta r = \sum_{n=1}^{N} w(t_n) |r_n - r_{\text{best}}|. \quad (2.8) \]

In Equation 2.8, \( N \) is the temporal length of the stacked waveforms, \( r_n \) is the potential source location associated with the CNR at time \( t_n \) and \( r_{\text{best}} \) is the location of the potential source associated at time \( t_{\text{max}} \), i.e. the location of the second generation template.

### 2.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:

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

$$\bar{S}(f) = \frac{1}{\sum_{s,c} 1_{\text{SNR}>5} [S_{s,c}(f)]} \sum_{s,c} \alpha_{s,c} S_{s,c}(f) 1_{\text{SNR}>5} [S_{s,c}(f)].$$  \hspace{1cm} (2.9)

In Equation 2.9, $1_{\text{SNR}>5} [S(f)]$ 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.

- The average spectra are converted to displacement spectra by using the relationship

$$|u_{\text{velocity}}(f)| = f \times |u_{\text{displacement}}(f)|,$$  \hspace{1cm} (2.10)

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

$$S_{\text{Boatwright}}(f) = \frac{\Omega_0}{\left(1 + \left(\frac{f}{f_c}\right)^4\right)^{1/2}},$$  \hspace{1cm} (2.11)

where $\Omega_0$ 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 $\Omega_0$ 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):
\[ |u^S(f)| = \frac{R^S}{2\rho \beta^3 r} \frac{M_0}{\left(1 + \left(\frac{f}{f_0}\right)^4\right)^{1/2}} \exp \left( -\frac{\pi ft^S}{Q^S} \right), \]

\[ \implies M_0 = \Omega_0 \frac{2\rho \beta^3 r}{R^S} \exp \left( \frac{\pi ft^S}{Q^S} \right), \tag{2.12} \]

\[ \implies \alpha_{s,c} = \frac{2\rho \beta^3 r_{s,c}}{R^S} \exp \left( \frac{\pi ft^S_{s,c}}{Q^S_{s,c}} \right). \]

In Equation 2.12, we use typical values for the S-wave velocity \( \beta \) (3000 km/s), the density \( \rho \) (2700 kg/m\(^3\)) and the average S-wave radiation pattern \( R^S \) (\( \sqrt{2/5} \) from Aki and Richards, 2002). The seismic moment \( M_0 \) gives the magnitude moment \( M_w \) through:

\[ M_w = \frac{2}{3} (\log M_0 - 9.1). \tag{2.13} \]

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 2.B.1 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:

\[ M_i = M_{ref} + \text{Median}_{s,c} \left\{ \log \frac{A_{s,c}^i}{A_{s,c}^{ref}} \right\}, \tag{2.14} \]

or more generally if there are several reference events:

\[ M_i = \text{Median}_k \left\{ M_{ref,k} + \text{Median}_{s,c} \left\{ \log \frac{A_{s,c}^i}{A_{s,c}^{ref,k}} \right\} \right\}. \tag{2.15} \]

**2.S Supplementary Material**
Figure 2.B.1: Magnitude estimation of the reference event. For each template, we use the highest SNR detections to calculate the average S-wave spectrum (Equation 2.9) and fit it with the Boatwright model (Equation 2.11). 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.
Figure 2.S.1: **Left panel (A):** Earthquake recorded at multiple stations. The waveforms are filtered in the band 1-12Hz and downsampled from 100 or 125 Hz to 50 Hz. **Right panel (B):** The envelopes of the seismic data are calculated and standardized: the daily median is removed, and the resulting signal is divided by its daily MAD (Median Absolute Deviation). Eventually, we cut out the 95th percentile of the signals by saturating the standardized envelopes $u(t)$ with hyperbolic tangent: $	ilde{u}(t) = p_{95} \times \tanh \left( \frac{u(t)}{p_{95}} \right)$, where $p_{95}$ is the 95th percentile of $u(t)$. This processing ensures the stations to have equal noise level before stacking (**cf.** Figure 2.S.2), and decreases the effect of undesired spurious signals in the data. The three superimposed layers show the three components: north, east and vertical.
Figure 2.S.2: Statistics of the envelope data for a given day. Different whisker boxes are for different stations, with each component (north, east, vertical) in different subplots. **Legend of the whisker boxes:** orange line: median, lower side of the box: Q1, upper side of the box: Q3, lower whisker: Q1 - 1.5(Q3-Q1), upper whisker: Q3 + 1.5(Q3-Q1). **Top panel:** Raw envelopes. **Bottom panel:** MAD-normalized envelopes: $\tilde{u}(t) = \frac{u(t) - \text{Median}(u(t))}{\text{MAD}(u(t))}$. After normalization, the stations exhibit similar distributions.
Figure 2.S.3: Composite network response (CNR, cf. Equation 2.2 in main material).

**Left panel (A):** Illustration of the maxima searching operation achieved to calculate the composite network response from all the network responses of the grid. **Right panel (B):** Histogram of the CNR samples presented in (C). The threshold median + 10 × MAD is given for information.
Figure 2.S.4: Second generation templates: increasing the SNR of the template waveforms. **Top panel:** Template matching provides us with many noisy repetitions of the same waveform. Different stacking methods can extract the coherent information from this collection of noisy records. **Bottom panel:** Stacking methods such as the $N^{th}$-root stack (*red waveform*) or the phase-weighted stack (*yellow waveform*) greatly improve the SNR with respect to the linear stack, but also distort waveforms because of non-linear operations. Our preferred method is the SVDWF (*green waveform*), which only performs linear operations. It exhibits a better SNR than the linear stack, and preserves the shape of the target waveform.
Figure 2.S.5: Comparison of the existing catalog with our catalog inside the dashed box; this region is where the geometry of the network allows best performances. Two events match if their origin times and locations are less than $\Delta T$ and $\Delta r$, respectively. $\Delta T$ and $\Delta r$ are two arbitrarily chosen thresholds. The unmatched events are shown with filled dots and the color codes their depth. For information, our template locations are shown with open diamonds. We missed 142 out of the 825 events (17%) documented in the catalog. However, we detected 16,430 new events, i.e., we detected a total of 17,113 earthquakes (21 times more detections with respect to the existing catalog). We show the same comparison for different $\Delta T$ and $\Delta r$ in Figure 2.S.6. Some of the unmatched events presented here are likely to be associated with inconsistencies in reported locations.
Figure 2.S.6: We perform the same comparison as in Figure 2.S.5 for different thresholds in order to investigate the effect of such arbitrary criteria. As the number of unmatched events decreases when we relax the criteria on origin time and location, it suggests that some of the unmatched events are only due to inconsistencies in location between the catalogs.
Figure 2.S.7: Distribution of magnitude of the 142 undetected events within the restricted area shown in Figure 2.S.5. Since the locations of the undetected events are close to the locations of our template events, we likely missed them because of their low magnitudes rather than because of the configuration of the station network.

Figure 2.S.8: Frequency magnitude distribution of the final catalog, only using templates with $\Delta r < 15$ km. The maximum likelihood estimate (red curve) is made on the range $[1, +\infty)$. We observe the Gutenberg-Richter relation to break down at $M_L \approx 1$ (black dashed line).
Figure 2.S.9: On both panels, the y-axis reports the position of the templates projected along the axis defined by the linear network CIFALPS. **Top panel:** Number of detected earthquakes per 10-day sliding window for each template. **Bottom panel:** Clustering coefficient (cf. definition in main material Figure 2.3.3) per 10-day sliding window for each template. Comparison of the two panels shows that strong seismic activity is often associated with high temporal clustering, but also that this is not always true. We observe continuous seismic activity beneath the Ubaye valley (with projected location around CT20), but only few episodes of high temporal clustering. This is consistent with previous observations pointing at a mixture of swarm-like seismicity and foreshocks-mainshock-aftershocks sequences in the Ubaye valley Daniel et al. (2011); Leclère et al. (2012, 2013); De Barros et al. (2019).
Figure 2.S.10: Cross-section along the CIFALPS axis showing 976 templates that were well relocated ($\Delta r < 15$ km). The color codes for different attributes in each cross-section, from top to bottom: number of detections, correlation time, fractal dimension and clustering coefficient of the one-year earthquake sequence of each template. Visual inspection seems to reveal that the fractal dimension offers a more contrasted image of temporal clustering than the clustering coefficient.
Bibliography


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


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


Chapter 3

The Rupture Complexity of Small Earthquakes in the Southwestern Alps

Abstract

We estimate the corner frequencies and seismic moments of small-to-moderate size earthquakes ($M < 4$) in the Southwestern Alps on multi-station average spectral ratios of similar earthquakes. This method allows to non-parametrically remove propagation and site effects corrupting the source signal in seismograms. The source parameters are obtained by fitting the Boatwright model to all spectral ratios together. Our results reveal apparent departures from self-similarity, in particular in the Dora Maira massif and in the Ubaye valley. Our previous study described a fluid-driven swarm-like seismicity in the Ubaye valley, and a cascade-like seismicity driven by earthquake interactions in the Dora Maira massif. We explain the apparent departure from self-similarity as pore-fluid controlled stress drops (and, therefore, seismic moments) in the Ubaye valley, whereas observations suggest asperity-like ruptures in the Dora Maira massif. Although not describing a systematic breakdown of self-similarity for small earthquakes, this study emphasizes the existence of rupture complexity in small
earthquakes that is poorly described by crack-like models (e.g. Brune or Boatwright model), and which may pass unseen when studying large ensembles of earthquakes from different geological settings.

3.1 Introduction

Observations of scaling laws between earthquake source parameters have helped constrain the physics of earthquakes. The concept of self-similarity, first introduced by Aki (1967), found strong support in these scaling laws, which imply that parameters such as the static stress drop $\Delta \sigma$ or the rupture velocity $v_r$ are independent of the size of the earthquake (e.g. Kanamori and Anderson, 1975; Ide and Beroza, 2001; Allmann and Shearer, 2009). However, the universality of self-similarity have been challenged either because models predict it to break down when earthquake sizes approach the nucleation length (e.g. Cattania, 2020) or because seismic ruptures of different natures could produce different scaling laws (e.g. Boatwright, 1988; Lin et al., 2016; Farge et al., 2020), whose details may be lost when analyzing global data sets.

Characterizing earthquake sources from surface (or borehole) measurements of the ground motion is not an easy task due to the poor spatial sampling and the large source-receiver distance usually involved in seismic data sets. These seismic data are nonetheless our best source of information about earthquakes that are too small to produce a significant permanent signal at the surface that could be measured by geodetic instruments (e.g. GPS/GNSS, InSAR, see studies such as Reilinger et al., 2000; Cakir et al., 2003). Earthquake waveforms in the far field carry information about source characteristics such as its duration and the size of the event (e.g. Aki, 1967; Brune, 1970). After propagating over long distances inside Earth and being recorded at the surface, earthquake waveforms not only are the signature of the source, but also of – potentially strong – propagation and local site amplification effects. On the one hand, propagation reduces the wave amplitude as seismic energy becomes more spread (geometrical spreading, e.g. Aki and Richards, 2002), and filters out
some of the original signal frequency content because of attenuation (the earth acts as a low-pass filter). On the other hand, when a seismic station is located in a sedimentary basin with low impedance, seismic waveforms can be strongly amplified \textit{(e.g. Viens et al., 2018)} and high frequencies strongly attenuated \textit{(e.g. Abercrombie, 1997)}.

The propagation and local site effects need to be accounted for if one wants to study earthquake source properties. Ideally, one needs to know the full response – \textit{i.e.} the Green function – of the earth to retrieve the source signal. Parametric methods for correcting the medium response typically use the approximation of homogeneous medium for geometric spreading, and a model of Earth’s properties for attenuation (attenuation models also derived from seismic data, \textit{e.g.} Campillo et al., 1985). A powerful non-parametric method is the use of an empirical Green function: an earthquake that is small enough to be seen as a point source – in time and space – in the utilized waveform frequency band constitutes a good approximation for the earth’s Green function between the source and the receiver \textit{(Hartzell, 1978; Hough, 1997)}). A generalization of this method is the use of spectral ratios of co-located earthquakes with similar focal mechanisms \textit{(i.e. same direction of slip on similarly oriented faults; Berckhemer, 1962; Aki, 1967; Ide et al., 2003; Abercrombie and Rice, 2005)}. Dividing the amplitude spectra of such earthquakes produces spectral ratios that are free of the medium response, which was common to both earthquakes, and therefore that are well suited for studying source properties.

This study analyzes the source properties of small earthquakes (\(M<4\)) in the Southwestern Alps, based on the high resolution earthquake catalog from Beaucé \textit{et al.} (2019). We identify groups of co-located and similar earthquakes, and study their source characteristics – seismic moment and corner frequency – with the spectral ratio method. We focus the analysis on the scaling laws that relate these source parameters and discuss the implications of our results for rupture physics and self-similarity in small earthquakes.
3.2 Method

3.2.1 Data

The earthquakes studied here were detected and located with an automatic method (see Beaucé et al., 2019) applied to one year of continuous recordings from 82 broadband seismic stations. Among these, 55 stations were part of the temporary linear array CIFALPS (China-Italy-France Alps survey, Zhao et al., 2016), and 32 stations were part of French and Italian permanent networks (see Data and Resources). Sampling rates were 100 Hz for most stations and 125 Hz for some of them.

3.2.2 Spectral Ratios

When approximating the earth as a linear filter, a seismogram \( u \) can be written as the convolution of a source term \( S \), a path term \( G \) (the Green function) and an instrument term \( I \):

\[
u(t; r) = (S(\cdot; \xi) * G(\cdot; r, \xi) * I)(t).
\]

In Equation (3.1), \( r \) is the location of the seismic station, \( \xi \) is the source location, and * is the temporal convolution. \( G \) include both propagation and local site effects. Because convolution is a multiplication in the Fourier domain, the ratio of the Fourier transforms of two event waveforms 1 and 2, recorded at the same station, is:

\[
\hat{u}_1(f; r) \hat{u}_2(f; r) = \frac{\hat{S}_1(f; \xi_1) \hat{G}(f; r, \xi_1) \hat{I}(f)}{\hat{S}_2(f; \xi_2) \hat{G}(f; r, \xi_2) \hat{I}(f)}.
\]

Equation (3.2) is called the spectral ratio between event 1 and 2 (at a given station). Therefore, one can see that if these two events are colocated, \( \xi_1 = \xi_2 = \xi \), the spectral ratio simplifies to:

\[
\frac{\hat{u}_1(f; r)}{\hat{u}_2(f; r)} = \frac{\hat{S}_1(f)}{\hat{S}_2(f)}.
\]

Only source terms are left in Equation (3.3).

Our earthquake catalog (Beaucé et al., 2019) was produced with the template
matching detection method (Gibbons and Ringdal, 2006; Shelly et al., 2007; Beaucé et al., 2018), which detects co-located earthquakes with similar waveforms. Thus, the catalog provides groups of co-located earthquakes with similar focal mechanisms that can be analyzed by the spectral ratio method described above. To build families of co-located and similar earthquakes, we first group together templates that have an average waveform correlation coefficient (CC) larger than 0.50 (see the centroid locations of these groups in Figure 3.2.1). We only keep groups that detected at least 10 events, which leaves 81 groups. We select, at most, 100 earthquakes that were detected by each of these groups. The selected earthquakes are the ones with larger magnitudes and higher similarity with the templates. The spectral ratio analysis involves all pairs of events within each family and therefore demands computation time and resources that grow quadratically with the number of events: 100 events is a good compromise between the latter and the number of source parameters we can estimate.

We compute the spectral ratio (Equations (3.2) and (3.3)) between all pairs of
events, and on each channel of the 40 stations closest to the centroid location (cf. Figure 3.2.2). Spectral ratios are computed on time series of 256 samples or 512 samples, for P and S waves, respectively, starting 0.2 s before the predicted first P- or S-wave arrival. Given that most sampling rates are 100 Hz, the time series duration are 2.56 s and 5.12 s for the P and S waves, respectively. We sometimes have to shorten the P-wave time window such that it does not overlap with the S-wave. Event pair windows are systematically realigned by finding the lags that maximize the correlation coefficient between the first arrivals (we discuss further how CCs are used in the processing). Spectral ratios are computed with the multiple-taper spectral correlation method proposed in Park and Levin (2000) and the Python library mtspec (Krischer, 2016) that wraps the multitaper spectrum estimation library from Prieto et al. (2009) (see Data and Resources). The spectral ratios between the P- and S-wave first arrivals and a noise window taken just before the P wave are also computed in order to obtain the signal-to-noise ratio (SNR) in the frequency domain.

Single-channel spectral ratios are subject to rupture directivity effects that can bias the estimate of source parameters (e.g. Madariaga, 1976), and also to noise. In order to mitigate these problems, we create multi-station average spectral ratios $\hat{R}$ by median stacking the single-channel observations $R_{ch}$ in the log domain (Prieto et al., 2004; Kaneko and Shearer, 2015; Uchide and Imanishi, 2016), and under several constraints:

$$\log \hat{R}(f) = \text{Median} \left\{ \log R_{ch}(f) \right\}_{ch \in C} ; \quad C = \left\{ ch \in S \mid \text{SNR}_{ch}^{1,2}(f) \geq \text{SNR}_{t}(f) \text{ and } \text{CC}_{ch} \geq 0.50 \right\}. \quad (3.4)$$

$C$ is the set of channels $ch$ that satisfy the constraints: the SNR of both events of the pair ($\text{SNR}_{ch}^{1,2}(f)$) have to be larger than the threshold ($\text{SNR}_{t}(f)$, cf. black dashed line in Figure 3.2.3), and the inter-event first arrival CC has to be larger than 0.5. The SNR threshold is 3 and frequencies higher than 45 Hz are rejected because they are affected by the instrument limits. These two criteria ensure that the average spectral ratio does not include noise, and only include similar waveform pairs. Note
Figure 3.2.2: Waveform alignment and spectral ratio of a pair of events on a single station. **A-C:** Three component P-wave waveforms of event 1 (black) and event 2 (orange). The first P-wave arrivals have a correlation coefficient (CC) of 0.75 after shifting event 2 backward by one sample. **D-F:** Three component S-wave waveforms of event 1 (black) and event 2 (orange). The first S-wave arrivals have a correlation coefficient of 0.87 after shifting event 2 forward by 6 samples. Note: the CC was computed on a narrow window around the first arrivals, and not over the whole time window shown here. **G-I:** P- (dashed blue) and S-wave (solid orange) spectral ratios (Equation (3.3)) of the three component waveforms shown above.

That Equation (3.4) is equally applied to P and S waves.

The set of channels $\mathcal{C}$ that satisfy the constraints at each frequency $f$ vary depending on the noise spectra. Therefore, the number of contributing channels at a given frequency in the median stack can be small. To avoid stacking an excessively low number of channels, we discard frequency bins with less than 12 contributing channels (see Section 3.2.3 for details on discarding). This criterion also helps ensure the reliability of the estimate of uncertainty in the stack, which we obtain by measuring the deviation of single-channel observations about the stack at each frequency. We use the median absolute deviation (MAD) scaled by the cardinality of $\mathcal{C}$:

$$
\epsilon(f) = \frac{\text{MAD} \left( \{ \log R_{ch}(f) \}_{ch \in \mathcal{C}} \right)}{|\mathcal{C}|}; \quad \text{MAD}(x) = \text{Median} \left( |x - \text{Median}(x)| \right), \quad (3.5)
$$
Figure 3.2.3: Multi-station average spectral ratio for a given pair of events. A, B: P-wave and S-wave average spectral ratio, respectively. Lines colored with a shade of blue are single-channel spectral ratios. The color shows the correlation coefficient between the first arrivals of the P or S wave (see Figure 3.2.2). The orange line is the median stack (see Equation (3.4) and text) and the shaded grey area shows the uncertainty on the stack. Only single-channel spectral ratios with CC greater than 0.5 are included in the average, in addition to the signal-to-noise ratio (SNR) criterion (see panels below). C, D: P-wave and S-wave single-channel SNR. The SNR is computed with respect to a noise window taken just before the P-wave arrival. Solid lines are SNRs of the first event of the pair, dotted lines are SNRs of the second event. Color is the same as for panels A and B. The dashed black line is the threshold on SNR above which a frequency bin is included in the average. Both SNRs of event 1 and 2 are required to be above threshold.

where $|C|$ is the number of elements in $C$ and $\epsilon(f)$ is the error of the stacked log spectrum. Since the stacked spectrum behaves approximately log-normally, $\epsilon(f)$ can be used to determine an uncertainty interval symmetric about the stack (cf. Figure 3.2.3A and B). Because of the division by $|C|$, $\epsilon(f)$ is only proportional to the actual error. We do not seek to obtain absolute errors as we use $\epsilon(f)$ for relative weighting of the frequency bins in the inversion (see Section 3.2.3).
3.2.3 Theoretical Model and Inversion

Simple crack-like models relate source displacement spectra to the seismic moment $M_0$ and the corner frequency $f_c$. The general expression for these is:

$$\tilde{u}(f) = \frac{\Omega_0}{(1 + (f/f_c)^\gamma)^{1/\gamma}}.$$  \hfill (3.6)

In Equation (3.6), $\Omega_0$ is the low-frequency plateau that is proportional to the seismic moment $M_0$, $f_c$ is the corner frequency that is related to the rupture duration (and the source dimension, for simple sources), $n$ is the high-frequency falloff rate, and $\gamma$ controls the sharpness of the corner of the spectrum. Well-known cases of Equation (3.6) are the Brune model with $\gamma = 1$ (Brune, 1970) and the Boatwright model with $\gamma = 2$ (Boatwright, 1978). The falloff rate $n$ is generally found to be around 2 both in data (e.g. Prieto et al., 2004) and models (e.g. Haskell, 1964; Aki, 1967; Brune, 1970). In this study, we fix the falloff rate to 2 to reduce the number of model parameters.

Using Equation (3.6) with $n = 2$, we obtain the expression for a spectral ratio between events 1 and 2:

$$R_{12}(f) = \frac{\tilde{u}_1(f)}{\tilde{u}_2(f)} = \frac{\Omega_{0,1}}{\Omega_{0,2}} \left(\frac{1 + (f/f_{c,2})^{2\gamma}}{1 + (f/f_{c,1})^{2\gamma}}\right)^{1/\gamma} = \frac{M_{0,1}}{M_{0,2}} \left(\frac{1 + (f/f_{c,2})^{2\gamma}}{1 + (f/f_{c,1})^{2\gamma}}\right)^{1/\gamma}. \hfill (3.7)

In Equation (3.7), the second equality comes from $\Omega_0 \propto M_0$. Equation (3.6) shows that one can only constrain the ratio of the two seismic moments rather than their absolute values with a spectral ratio. When computing the spectral ratios of all event pairs among $N$ events, we obtain $N^2$ spectral ratios and $N(N - 1)/2$ of them are relevant ($R_{12} = R_{21}^{-1}$, and the $R_{ij}$'s are trivial). Thus, we can write the system of
\[ N(N - 1)/2 \text{ non-linear equations:} \]

\[
\begin{align*}
\hat{R}_{12}(f) &= R_{12}(f) \\
\vdots & \\
\hat{R}_{ij}(f) &= R_{ij}(f), \ j > i \\
\vdots & \\
\hat{R}_{(N-1)N}(f) &= R_{(N-1)N}(f).
\end{align*}
\]

The system of equations (3.8) involves \(2N\) independent variables: the \(N\) corner frequencies \(f_{c,i}\), the \(\gamma\) parameter, and \(N - 1\) seismic moment ratios \(M_{0,i}/M_{0,i_{\text{ref}}}\). To retrieve absolute seismic moments, one needs to know a priori the value of one seismic moment \(M_{0,i_{\text{ref}}}\). In this study, we know the seismic moment of at least one event per family (see Beaucé et al., 2019) and we add one equation to the system to force the inverted \(M_{0,i_{\text{ref}}}\) to be equal to its prior value. Other studies force the average of the inverted seismic moments to be equal to the average of the prior seismic moments (only works if all seismic moment estimates are available a priori, e.g. Ide et al., 2003; Imanishi and Ellsworth, 2006).

In practice, \(N(N - 1)/2\) spectral ratios can rapidly lead to intractable computation during the inversion of the system of equations (3.8). We therefore remove some of the bad quality event pairs, but taking care of not creating any isolated subsets of events (e.g. pair \(i-j\) is featured, but none of the other pairs involve \(i\) or \(j\)). We solve the system of non-linear equations (3.8) by minimizing the \(p\)-norm of the weighted residuals between the observed log spectral ratios \(\hat{R}_{ij}(f)\) and the modelled log spectral
The loss function written in Equation (3.9) is a non-linear function of the model parameters. The residuals (term in parenthesis) are weighted by $w_{ij}(f)$ that is inversely proportional to the stack uncertainty $\epsilon_{ij}(f)$ (see Equation (3.5)). We find the optimal parameters that minimize $\mathcal{L}$ by iteratively linearizing the problem (e.g. Tarantola and Valette, 1982), and the uncertainties on the inverted parameters are estimated with the Hessian of the loss function. We selected a set of max 100 events, and minimized both the l2-norm (least squares criterion, $p=2$) and the l1-norm (least absolute values criterion, $p=1$) of the weighted residuals, and tested the Brune ($\gamma = 1$), Boatwright ($\gamma = 2$) and custom models ($\gamma$ is a free parameter). Using this linearized approach is subject to getting stuck in a local minimum of the loss function (cf. Figure 3.S.1). The l1-norm seemed to give better convergence overall, and both the Boatwright and Brune models gave similar performances (cf. Figure 3.S.2). The custom model suffered from convergence issues. Since the Boatwright model with the least absolute values criterion gave slightly smaller residuals than the Brune model (cf. Figure 3.S.2), all subsequent analysis is done with the Boatwright model and l1-norm minimization. Examples of fitted models are shown in Figure 3.2.4. We acknowledge that further comparison between the Brune and Boatwright models should be done to see how the model influences the final results.
Figure 3.2.4: Multi-station average spectral ratios between the largest magnitude event and all others (thick colored lines) and model predicted spectral ratios (thin colored lines) for A: the P wave and B: the S wave. Curves are colored according to the catalog magnitude of the second event. The corner frequencies of the two events $f_{c,1}$, $f_{c,2}$ of each pair are shown with black circled dots. Note that this group of earthquakes was taken from the Dora Maira massif (see location on Figure 3.2.1), and that the model does not seem to explain well the low-frequency part of the ratios ("bump" in the 3-8 Hz band).

### 3.3 Results

#### 3.3.1 Scaling Law

Two fundamental, and observable, parameters to describe earthquakes are the seismic moment $M_0$ (Equation (3.10)) and the static stress drop $\Delta \sigma$ (Equation (3.11)):

\[
M_0 = GA\bar{D},
\]  

\[
\Delta \sigma = CG\frac{\bar{D}}{L}.
\]

In Equations (3.10) and (3.11), $G$ is the shear modulus (also called rigidity $\mu$, in Pa), $A$ is the fault area ($m^2$), $L$ is the linear dimension of the fault (m), $\bar{D}$ is the average slip on the fault (m), and $C$ is a non-dimensional shape factor ($\sim 1$). Empirical scaling
laws between source parameters have shown strong evidence for the self-similarity of earthquakes (Kanamori and Anderson, 1975; Kanamori and Brodsky, 2004; Allmann and Shearer, 2009), originally postulated by Aki (1967). The static stress drop $\Delta \sigma$ and the rupture velocity $v_r$ appeared to be independent of the size of the earthquake. Studies have shown evidence for self-similarity in earthquakes down to $M - 1$ (Abercrombie, 1995).

The corner frequency of omega-square models (e.g. Aki, 1967; Brune, 1970) is interpreted as a measure of the rupture duration $\tau$: $f_c \propto 1/\tau$. Self-similarity predicts the $f_c \propto M_0^{-1/3}$ scaling between corner frequency and seismic moment. In this study, we investigate potential deviations from this canonical scaling by searching which power-law exponent $\alpha$ best describes the data:

$$f_c \propto M_0^{-\alpha}.$$  \hspace{1cm} (3.12)

We will refer to the power-law exponent of the self-similar scaling as $\alpha_{ss} = 1/3$.

![Figure 3.3.1: Inverted seismic moments and corner frequencies for a given group of similar earthquakes (the same as in Figure 3.2.4), on the P (blue dots) and the S (orange dots) waves. The power-law exponent $\alpha$ of the $f_c-M_0$ scaling relation (Equation (3.12)) is measured by a robust linear regression in the log-log domain (l1-norm minimization). Uncertainties are estimated with bootstrap resampling.](image)

135
We use the results of the inversion (see Section 3.2.3) to estimate the power-law exponent $\alpha$ of Equation (3.12) with a robust linear regression in the log-log domain (l1-norm minimization, cf. Figure 3.3.1). Bootstrapping (Efron and Tibshirani, 1986) is used to further reduce the influence of outliers and to estimate uncertainties $\Delta \alpha$ on the power-law exponent. We combine the power-law exponents estimated from the P- and S-wave inversions into a single number by taking their weighted average, as well as for their uncertainties. Weights are the inverse of the uncertainties $\Delta \alpha$. Groups that produced an average uncertainty larger than 0.20 are discarded. After applying all the quality criteria (SNR and CC, bad quality event pairs and $\Delta \alpha < 0.2$), we were left with 46 template groups. Our results are shown in Figure 3.3.2, and groups with less than 10 successfully inverted events are marked differently. These groups seem to produce correct estimates of corner frequencies and seismic moments (see Figure 3.3.3), but the low number of data points makes the power-law exponent regression little reliable.

We compare the power-law exponents (Figure 3.3.2B) with other parameters: the centroid depth (Figure 3.3.2A), the total number of inverted events (Figure 3.3.2C) and the fractal dimension of the time series of event count per unit time (Figure 3.3.2D). The latter quantifies the strength of temporal clustering (explanations in Beaucé et al., 2019), which is usually interpreted as symptomatic of earthquake interactions (e.g. Marsan and Lengline, 2008; Dublanchet et al., 2013). Large deviations from the canonical scaling $f_c \propto M_0^{-1/3}$ are observed, particularly with exponent 2-3 times lower than $\alpha_{ss}$ (blue dots in Figure 3.3.2B). We observe that the groups with strongest temporal clustering ($D > 0.35$, Figure 3.3.2D), located at $\sim 12$-15 km depth at the western edge of the Po plain, correlate well with low $\alpha$ ($\approx 0.14$). These are also well resolved due to the large number of successfully inverted events (Figure 3.3.2C). However, groups with many inverted events but low clustering, in the Ubaye valley, also show a weak scaling ($\alpha \approx 0.09$), while being located in the middle of many events exhibiting self-similarity. All groups with many successfully inverted events show weak scaling $\alpha < \alpha_{ss}$, but these also have larger depths and/or fractal
Figure 3.3.2: Corner frequency vs seismic moment scaling and other parameters. 

A: Centroid depth. B: Power-law exponent $\alpha$ of Equation (3.12). The color scale is centered around the self-similarity power-law exponent: $\alpha_{ss} = 1/3$ (white). All exponents that were not different from $\alpha_{ss}$ within the 1-$\sigma$ uncertainty were set to $\alpha_{ss}$ on this map. C: Number of events that were successfully inverted in each group (P- and S-wave inversion). D: Fractal dimension of the time series of event count per unit time (see Beaucé et al., 2019). Large values mean strong temporal clustering. Diamonds are results from groups with less than 10 successfully inverted events (see panel C). The rest of the groups are shown with dots.

Measuring meaningful scaling relationships require observing source characteristics over a wide range of magnitudes to capture the general trend among often apparently noisy data. When fitting a power-law to the whole inverted data set (cf. Figure 3.3.3), we still find a different scaling than self-similarity, although scatter is large. Beside errors in source parameter estimates, significant scatter in the $f_c$-$M_0$ space is due to variations in stress drops. We used Equations (3.13) (Madariaga, 1976) and (3.14) (Eshelby, 1957) to compute the corner frequency given the seismic moment and stress drop:

\[
f_c^{PS} = k^{PS} \beta \frac{M_0}{L}, \tag{3.13}
\]

\[
\Delta \sigma = \frac{7}{16} \frac{M_0}{L^3}, \tag{3.14}
\]
Figure 3.3.3: All corner frequencies and seismic moments that were successfully inverted. **A:** P-wave inversion. **B:** S-wave inversion. Fitting a power-law to the data $f_c \propto M_0^{-\alpha}$ yields low exponents $\alpha$: 0.08 and 0.14 for the P and S waves, respectively. These data could also be explained by the canonical scaling law, i.e. $\alpha_{ss} = 1/3$, and a spread of stress drop values (see 0.1, 1, 10, 100 and 1000 MPa lines, computed with the model from Madariaga, 1976). Grey diamonds are results from groups with less than 10 successfully inverted events (see Figure 3.3.2C).

with $k^P = 0.32$ and $k^S = 0.21$ (i.e. considering $v_r = 0.9\beta$, Madariaga, 1976), and $\beta = 3000$ m$^2$/s the shear wave speed. Note that this stress drop computation assumes a circular fault geometry and self-similarity as the rupture velocity $v_r$ is fixed for all values of $M_0$. The results show that part of the data could follow the $f_c \propto M_0^{-1/3}$ scaling with various stress drops, but some data fall out of the range of reasonable stress drops ($\Delta\sigma > 100$ MPa, Figure 3.3.3).

### 3.3.2 Departure from Self-Similarity

Self-similarity has been observed over a wide range of magnitudes (Kanamori and Anderson, 1975; Abercrombie, 1995), and apparent departures from this scaling for small earthquakes have been attributed to observational limits (e.g. attenuation or narrow bandwidth recording Abercrombie, 1995; Ide et al., 2003; Abercrombie, 2015). Given the limited bandwidth we utilized in this study, up to 50 Hz only due to instru-
ment limitations, we interpret our results with caution. To investigate further the apparent breakdown of self-similarity seen in this study, we choose one group from each of the two clusters with high number of inverted events and low power-law exponents (Ubaye valley and Dora Maira massif), and visualize the earthquake spectra (see Figures 3.3.4 and 3.3.5).

Figure 3.3.4: Unusual corner frequency - seismic moment scaling relationship in the Dora Maira massif. Multi-station average displacement amplitude spectra of A: P waves and B: S waves corrected for geometrical spreading and attenuation (assuming a $Q \propto f^{-0.5}$ model, e.g. Campillo et al., 1985). We use the magnitudes and corner frequencies inferred with the spectral ratio inversion to further average amplitude spectra and corner frequencies in magnitude bins (black diamonds are the average corner frequencies). Dashed black lines show the self-similar ($\propto f^{-3}$) and the inverted ($\propto f^{-6.67}$, i.e. $\alpha = 0.14$) scaling. C: P-wave and D: S-wave amplitude spectra are unsuccessfully collapsed using the self-similar scaling. E: P-wave and F: S-wave amplitude spectra are satisfactorily collapsed using the inverted scaling.

We compute the multi-station average amplitude spectra with the same median stacking and SNR criterion as described for spectral ratios (see Equation (3.4)). Single-channel amplitude spectra are corrected for geometrical spreading and atten-
Figure 3.3.5: Unusual corner frequency - seismic moment scaling relationship in the Ubaye valley. Multi-station average displacement amplitude spectra of A: P waves and B: S waves corrected for geometrical spreading and attenuation (assuming a $Q \propto f^{-0.5}$ model, e.g. Campillo et al., 1985). We use the magnitudes and corner frequencies inferred with the spectral ratio inversion to further average amplitude spectra and corner frequencies in magnitude bins (black diamonds are the average corner frequencies). Dashed black lines show the self-similar ($\propto f^{-3}$) and the inverted ($\propto f^{-11.34}$, i.e. $\alpha = 0.09$) scaling. C: P-wave and D: S-wave amplitude spectra are unsuccessfully collapsed using the self-similar scaling. E: P-wave and F: S-wave amplitude spectra are satisfactorily collapsed using the inverted scaling.

Therefore, given the scaling $f_c \propto M_0^{-\alpha}$, amplitude spectra of various magnitude events are satisfactorily collapsed using the inverted scaling.
should all fall along the $\propto f^{-1/\alpha}$ curve. Thus, one can test the validity of the scaling relationship by collapsing the amplitude spectra using the $M_0 \propto f_c^{-1/\alpha}$ relationship. We transform each average spectra $|\tilde{u}_i(f)|$ according to:

$$|\tilde{u}_i(f)| \longrightarrow \frac{\bar{M}_0}{M_{0,i}} |\tilde{u}_i \left( (\bar{M}_0/M_{0,i})^{-\alpha} f \right)|.$$

(3.16)

where $M_{0,i}$ is the seismic moment of event $i$ and $M_0$ is the average seismic moment of the family of events (the value of $\bar{M}_0$ does not matter for collapsing the spectra). We test the canonical scaling $\alpha_{ss} = 1/3$ and the inverted scaling for collapsing the amplitude spectra. Results are shown in Figures 3.3.4C-F and 3.3.5C-F. The self-similar scaling does poorly at collapsing the spectra, while the inverted scaling succeeds.

### 3.4 Discussion

#### 3.4.1 Unconventional Rupture Models

In Section 3.3, we showed that many groups of earthquakes follow $f_cM_0$ scaling relationships that deviate substantially from the widely observed self-similar scaling $f_c \propto M_0^{-1/3}$. The low sampling rate (100 Hz) may make us discard event pairs with corner frequencies higher than the Nyquist frequency (Ide and Beroza, 2001), or may lead to underestimating the corner frequencies (Abercrombie, 2015). However, our study still identifies many relatively low corner frequencies ($f_c < 20$ Hz) that are resolvable by our method. The entire set of inverted source parameters suggest that the observed deviation could only be reconciled with self-similarity under extreme stress drop variations from 0.1 MPa to 1000 MPa, with largest stress drops for $M>4$ earthquakes (see Figure 3.3.3). Such large stress drops, in particular for moderate size earthquakes, are unlikely to exist.

We checked the validity of our results on two groups of earthquakes, located in the Dora Maira massif and the Ubaye valley (cf. Figures 3.3.4 and 3.3.5). The results con-
firm that these earthquakes follow an unusual scaling where the corner frequency varies more weakly with the seismic moment than would be expected from self-similarity. Thus, these observations suggest that particular source properties are the cause of the observed low power-law exponent \( \alpha < \alpha_{ss} \). Multiple studies have reported a weak magnitude dependence of corner frequency upon magnitude (e.g. Nadeau and Johnson, 1998; Bostock et al., 2015; Farge et al., 2020). Such a weak dependence is sometimes explained by variable slip and stress drop on fixed size sources (Lengliné et al., 2014; Lin et al., 2016), by an event duration fixed by an external process such as transient pore pressure changes (e.g. Lin et al., 2016; Shapiro et al., 2018; Farge et al., 2021), or by a rupture scenario that differs from the crack model of Brune-like models (Brune, 1970).

An example of the latter are asperity models that consider a strong patch embedded in a weaker fault where slip propagates following the breaking of the asperity, thus constituting an heterogeneous event (McGarr, 1981; Boatwright, 1988). These models feature two characteristic time scales (therefore, two corner frequencies): the rupture duration of the asperity, and the rupture duration of the surrounding area. Consequently, asperity earthquake spectra may show a wide range of seismic moments, which is mostly controlled by the slip on the surrounding area, with the same corner frequency that reflects the size of the asperity. Various studies have built models of closely located and strongly interacting asperities embedded in a weak, creeping fault (Sammis and Rice, 2001; Johnson and Nadeau, 2002) to solve the apparent problem of very high stress drops (\( \sim \) GPa) in repeating earthquakes (e.g. Nadeau and Johnson, 1998).

3.4.2 Fluids vs. Asperity: Ubaye vs. Dora Maira

In our previous work (Beaucé et al., 2019), we showed that the seismic activity in the Ubaye valley and the Dora Maira massif behaved differently: the Ubaye valley hosts a continuous high seismic activity with low temporal clustering, whereas the
seismicity in the Dora Maira massif occurs in strongly time clustered bursts of events. We quantified the strength of temporal clustering by computing the fractal dimension of the time series of event count per unit time (see Figure 3.3.2D). The swarm-like seismicity in the Ubaye valley has been attributed to fluid-related processes (Daniel et al., 2011; Leclère et al., 2013; De Barros et al., 2019), and we have attributed the strong burst-like seismicity in the Dora Maira massif to a densely fractured medium (Beaucé et al., 2019). In the light of the new $f_c \propto M_0^{-\alpha}$ scaling observations, and the models of unconventional rupture mentioned in Section 3.4.1, we suggest that the peculiar scalings observed in these two regions have different physical origins.

Figure 3.4.1: Corner frequencies and seismic moments estimated in A: the Ubaye valley and B: the Dora Maira massif, for the P (blue symbols) and S (orange symbols) waves. Solid lines indicate different levels of constant stress drop scaling (computed following the model of Madariaga, 1976). The inverted scaling laws are shown with the dashed lines.

The corner frequencies of $M_w \leq 2$ earthquakes in the Ubaye valley seems to be almost independent of the seismic moment, and many events occurred at low stress drops ($\Delta\sigma < 1$ MPa, cf. Figure 3.4.1A). We interpret this weak dependence of $f_c$ on $M_0$ ($\alpha = 0.09$) as a decrease of stress drop with decreasing seismic moment due to transient pore-fluid changes. An increase of pore-fluid pressure results in a decrease of the effective normal stress acting on the fault, and therefore in the earthquake stress drops (shear stress is proportional to normal stress). For a fixed source dimension, a lower stress drop also means a lower seismic moment (see Equation (3.11)). Thus, pore-fluid-controlled changes in effective normal stress acting on the population of
seismogenic fault patches in the Ubaye valley may be responsible for low stress drop, low seismic moment events within a narrow range of corner frequencies, which are only determined by the fixed dimensions of the seismogenic sources.

Even though we also found a weak dependence of corner frequency on seismic moment in the Dora Maira seismicity ($\alpha = 0.14$), observations now suggest anomalously high stress drops for moderate-size earthquakes ($M_w > 3$ and $\Delta \sigma > 100$ MPa, cf. Figure 3.4.1B). This, along our previous observations of temporal clustering, discard the hypothesis of fluid-driven seismicity. We propose that these earthquakes result from asperity-like ruptures, for which earthquake spectra are expected to be complex because of the multiple length scales involved in the rupture (Boatwright, 1988). A close look at spectral ratios in Figure 3.2.4 reveals a systematic discrepancy between the Boatwright model and the observations at low frequencies. Instead of a clear plateau at low frequencies, we observe two bumps and a trough, and our inversion seems to pick the second bump as the corner frequency. We suggest that the second bump is related to the smallest length scale of the rupture, the size of the strong asperity, while the first bump is controlled by the largest length scale, the size of the weak area. It has also been suggested that areas with strong asperities embedded in a weak fault could produce composite events where not only the weaker area controls the seismic moment, but also the number of asperities that break together due to elastic interactions (Johnson and Nadeau, 2002). Thus, it is the attempt to explaining these spectral ratios with a simple crack model that leads to the incorrect conclusion of high stress drops ($\Delta \sigma > 100$ MPa), because the inverted corner frequencies are not representative of the total rupture areas.

Both the hypotheses of pore-fluid controlled stress drops in the Ubaye valley and interacting asperities embedded in a weak fault in the Dora Maira massif are consistent with our observations of temporal clustering (cf. Figure 3.3.2D). Fluid-driven seismicity tends to show weak temporal clustering, as opposed to an asperity-asperity interaction driven seismicity that implies strong temporal clustering. Furthermore,
in Chapter 4 of this thesis, we show that earthquakes located closer to the bottom of the seismogenic zone were more prone to be part of strongly time clustered sequences. We interpret it as resulting from enhanced interactions due to creep mediated stress transfers that occur when seismogenic patches are embedded in a weak, creeping fault (Dublanchet et al., 2013; Cattania, 2019). Given the depth range of the events in the Dora Maira massif (12-15 km), creep mediated stress transfers are a plausible explanation for strong temporal clustering, which also fits into the asperity model discussed here. However, the weak temporal clustering observed in the Ubaye valley does not support the existence of aseismic slip taking place around the seismogenic sources, as could be caused by low effective normal stresses (Lengliné et al., 2014). If aseismic slip occurs in the Ubaye valley, then it must be spatially separated from the seismic sources so that it does not enhance interactions (Lohman and McGuire, 2007).

3.4.3 Template Matching and Unconventional Earthquakes

Many studies that found unusual $f_c$-$M_0$ scaling relationships used earthquake catalogs made with the template matching detection method (e.g. Bostock et al., 2015; Farge et al., 2020, and this study). This raises the question of whether the detection method biases catalogs towards these scalings or not. Events that show almost no $f_c$-$M_0$ scaling exhibit amplitude spectra that differ only by a proportionality factor (see similar spectra in Figures 3.3.5 and 3.3.4). Therefore, in an ideal noise-free setting these events of different magnitudes still produce highly correlated waveforms. Thus, template matching is expected to perform particularly well on this type of earthquakes. We therefore suggest that template matching catalogs could be biased towards weak $f_c$-$M_0$ scaling events. If the detected events are, after all, self-similar, our results would still suggest that the detected events span a very wide range of stress drops, if interpreted by standard models (see Figures 3.3.3 and 3.4.1).

Another source of bias towards weak $f_c$-$M_0$ scaling could be introduced by the correlation criterion on event pairs used to ensure that events are sufficiently well
co-located to cancel out path effects at high frequencies in spectral ratios. Instead, event pair correlation coefficients should be carefully computed using frequencies below both events’ corner frequencies, which may not be practically possible due to SNR limitations.

3.5 Conclusions

We applied the spectral ratio method to small-to-moderate earthquakes (M < 4) in the Southwestern Alps. This method allows correcting earthquake spectra for propagation and instrument effects in a non-parametric way. We used the dense linear array CIFALPS, and additional permanent stations, in order to average the spectral ratios over multiple azimuths, and thus limit directivity effects.

We made groups of up to 100 similar earthquakes, inside which we performed an inversion to fit a Boatwright model with falloff rate $n = 2$ to all event pairs at once. The results revealed apparently significant deviations from self-similarity. We focused our analysis on two groups, located in the Ubaye valley and in the Dora Maira massif, that produced large amounts of successfully inverted source parameters. These groups show a remarkably weak scaling between corner frequency and seismic moment ($f_c \propto M_0^{-\alpha}$, $\alpha < 1/3$), which is well evidenced when collapsing the earthquake spectra onto one curve using the inverted scaling (see Figures 3.3.4 and 3.3.5).

Our previous study (Beaucé et al., 2019) showed that the swarm-like seismicity of the Ubaye valley indicates weak earthquake interactions and suggests that some external forcing is at play to drive the seismicity, which we attributed to fluid related processes (in agreement with other studies, Daniel et al., 2011; Leclère et al., 2013; De Barros et al., 2019). The same work also described the cascade-like seismicity of the Dora Maira massif as strongly driven by interactions between earthquakes, possibly indicating a dense population of interacting seismogenic sources.
We interpreted the $\alpha = 0.09$ power-law exponent of the Ubaye Valley seismicity and low stress drops as reflecting a decrease in stress drop $\Delta \sigma$ with decreasing seismic moment $M_0$ due to pore-fluid pressure changes (see Section 3.4.2). These transient changes decrease the effective normal stress on faults, and may occur due to fluid migration. The drop in effective normal stress would trigger subsequent rupturing of the same seismogenic fault patches with low stress drops, and thus explain the quasi-constant corner frequency of low magnitude events. The weak temporal clustering observed in the Ubaye valley does not support the scenario of seismogenic patches embedded in a creeping fault, as could be expected from a drop in normal stress (Lengliné et al., 2014).

We attributed the $\alpha = 0.14$ power-law exponent of the Dora Maira seismicity and apparently high stress drop earthquakes to the misleading interpretations of inadequate crack-like models, such as Brune’s (see Section 3.4.2). Instead, we proposed that an asperity-like rupture scenario (Boatwright, 1988) was a more likely explanation to the population of events observed in the Dora Maira massif. Asperity-like earthquakes involve heterogeneous fault properties, and possibly composite events, that imply a degree of rupture complexity that is often overlooked in the description of small earthquakes. Furthermore, such asperity models describe strong seismogenic patches embedded in a weak creeping fault, which constitutes favorable conditions for the observed strong temporal clustering (Dublanchet et al., 2013; Cattania, 2019; Beaucé et al., 2021b).

Finally, we stress that this work did not evidence a systematic breakdown in self-similarity due to the small size of the ruptures. Instead, we observed departures from the self-similar scaling as a consequence of specific physical processes. We suggest that these apparent departures from self-similarity would not be seen if we were to analyze a larger data set with a wider range of geological settings. Therefore, if the self-similar scaling laws derived from crack-like models do seem to govern earthquake
characteristics at first order, their universality may only be an approximation that makes us lose information on the actual physics of earthquakes.

Data and Resources

We created the maps in Figures 3.2.1 and 3.3.2 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 July 2021). 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.

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union’s Horizon H2020 research and innovation program (grant agreement No 742335). E.B. was also supported by funds associated with Robert D. van der Hilst’s Schlumberger chair.

3.S Supplementary Material

3.S.1 Model Comparison

3.S.2 Supplementary Results
Figure 3.S.1: Multi-station average spectral ratios between the largest magnitude event and all others (thick colored lines) and model predicted spectral ratios (thin solid line: l2 loss, thin dashed line: soft l1 loss). Left column (A, C, E): P wave. Right column (B, D, F): S wave. Top row (A, B): Brune model ($\gamma = 1$ in Equation (3.6) in main text). Middle row (C, D): Boatwright model ($\gamma = 2$ in Equation (3.6) in main text). Bottom row (E, F): Custom model ($\gamma$ is a free parameter). The Boatwright and custom models with the l2 loss do not converge to a satisfactory solution on the largest magnitude event.
Figure 3.S.2: Residuals between log observations and log models for P (blue distributions) and S (orange distributions) waves. **First row (A, B, C):** L2 loss with the Brune, Boatwright and custom model. **Second row (D, E, F):** Soft l1 loss with the Brune, Boatwright and custom model. The Boatwright model with the soft l1 loss does marginally best.
Figure 3.S.3: Inverted corner frequencies $f_c$ and seismic moments $M_0$ for all events. The exponent of the scaling law $f_c \propto M_0^{-\alpha}$ was estimated with a linear regression in the log-log domain using the least absolute value criterion (dashed lines). Uncertainties were estimated by bootstrapping the data set and repeating the regression. **First row (A, B, C):** L2 loss with the Brune, Boatwright and custom model. **Second row (D, E, F):** Soft l1 loss with the Brune, Boatwright and custom model. Preferred inversions show consistent results between the P and S waves. Boatwright and soft l1 loss also produce the smallest residuals (cf. Figure 3.S.2), therefore we choose this model in the study.
Figure 3.S.4: P-S averaged power-law exponent $\alpha$ (cf. Equation (3.12)) vs. A: Centroid depth, B: Root mean square (RMS) residual, C: Total number of inverted events (P- and S-wave inversions), and D: Fractal dimension of the event count per unit time (see main text). Red diamonds are groups for which less than 10 events were successfully inverted, therefore for which the power-law exponent is not very reliable. Blue dots are the rest of the groups.
Bibliography


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


L. Krischer. mtspec python wrappers 0.3. 2. *Dataset, doi*, 10, 2016.


Chapter 4

Seismotectonic Study of the North Anatolian Fault Zone Thirteen Years After the 1999 M7.4 Izmit Earthquake

Abstract

The 17 August 1999 $M_{\text{w}}7.4$ Izmit earthquake ruptured the western section of the North Anatolian Fault and strongly disturbed the properties of the fault zone. The stress changes induced by the coseismic and postseismic phenomena affected the spatial distribution and the productivity of earthquakes, as well as the local stress state and the slip rates observed at the surface. Using data recorded by the dense seismic array DANA thirteen years after the Izmit earthquake, in 2012-2013, we built a new earthquake catalog with an automated detection and location method. We combined backprojection, template matching and automatic phase picking in a single workflow. We utilize the abundance of detected earthquakes to characterize the state of the fault zone with first order observations such as the spatial distribution of events, and higher order observations such as temporal clustering and focal mechanisms. We use the focal mechanisms to determine the regional and the local stress state, compare it with the pre- and post-Izmit state, and our results suggest that the weakening of the fault observed just after the Izmit earthquake had terminated by the time of the DANA experiment. However, this study also brings new seismological evidence

that postseismic phenomena are still disturbing the fault system. A careful analysis and interpretation of earthquake temporal clustering suggests the existence of seismic sources on predominantly creeping normal faults at the interface between unstable and stable sliding beneath Lake Sapanca, which could help explain the north-south surface displacement observed in the region.

4.1 Introduction

The North Anatolian Fault (NAF) is a 1,500 km long strike-slip fault that marks the boundary between the Anatolian plate in the south and the Eurasian plate in the north (cf. Figure 4.1.1A). The fault slips, overall, in a right-lateral manner to accommodate the westward motion of Anatolia with respect to Eurasia due to the subduction along the Hellenic trench and the Cyprus trench in the southwest and the south, and the collision with Arabia in the southeast (Le Pichon and Angelier, 1979; McClusky et al., 2000; Reilinger et al., 2006). Near the Gulf of Izmit, in western Turkey, the NAF splits into a Northern strand and a Southern strand. These two strands bound the Almacik mountains and the Armutlu block, and separate the Istanbul Zone in the north from the Sakarya Terrane in the south, the remains of the passive margin of the Intra-Pontide Ocean (see Figure 4.1.1B).

The 17 August 1999 M7.4 Izmit earthquake and the 12 November Düzce M7.2 earthquake are the most recent (as of 2021) events of a series of westward migrating M>7 earthquakes that ruptured almost entirely the North Anatolian Fault (e.g. Toksöz et al., 1979; Stein et al., 1997). The Izmit earthquake nucleated near the Izmit Bay, propagated bilaterally and broke a 150 km-long, almost vertical section of the fault made of four segments along the Northern strand (Toksoz et al., 1999; Barka et al., 2002). To the east, the rupture propagated at super-shear speeds (Bouchon et al., 2001, 2011) and broke the Izmit-Sapanca, the Sapanca-Akyazi and the Karadere segments (cf. names on Figure 4.1.1B). To the west, the rupture propagated along the Gölcük segment and stopped on the Yalova segment (Langridge et al., 2002), increasing the risk of major failure further west beneath the Marmara Sea (Parsons et al.,
2000). The Düzce earthquake nucleated further east in an area of increased Coulomb stress due to the Izmit rupture (Parsons et al., 2000; Utkucu et al., 2003). The Izmit earthquake significantly perturbed the local seismicity patterns and stress state (e.g. Bohnhoff et al., 2006; Pinar et al., 2010; Ickrath et al., 2015), but also the behavior of the fault. Various studies based on GPS and interferometric synthetic aperture radar (InSAR) documented the postseismic slip following the two earthquakes (Reilinger et al., 2000; Ergintav et al., 2009) and showed that shallow sections of the NAF that ruptured in 1999 started to creep aseismically (e.g. Çakir et al., 2012; Hussain et al., 2016; Aslan et al., 2019).

Thirteen years after the 17 August 1999 Izmit earthquake, the dense seismic array DANA (Dense Array for North Anatolia DANA, 2012, see Figure 4.1.1C) was deployed around the Izmit rupture, and operated from early May 2012 to late September 2013. These data enabled multiple studies that improved our understanding of the complex structures and seismicity patterns in the region (e.g. Poyraz et al., 2015; Kahraman et al., 2015; Papaleo et al., 2018; Taylor et al., 2019). Despite these recent advancements, new geophysical evidence is necessary for refining our view of the complex structures of the North Anatolian Fault and understanding their role in the behavior of the fault. This work brings new seismic observations by automatically analyzing the DANA data set to detect and locate earthquakes, as well as to estimate earthquake focal mechanisms. In Section 4.2, we present the earthquake detection and location method. Section 4.3 describes the seismicity observed in our new catalog and presents an analysis of earthquake temporal clustering. Section 4.4 presents the focal mechanisms and the stress state we derive from them. Finally, in Section 4.5 we discuss the implications of these results on the state of the North Anatolian Fault zone in 2012-2013: we analyze the different behaviors of the fault segments and how the NAF has evolved since the Izmit earthquake, and we propose that extensional normal faults may be important in locally accommodating displacement between the different blocks via mixed seismic and aseismic processes.
Figure 4.1.1: **A:** Large scale view of the North Anatolian Fault Zone. Abbreviations: NAFZ - North Anatolian Fault Zone, EAFZ - East Anatolian Fault Zone. The red arrows indicate the direction of coseismic motion. Our study region is located at the western end of the North Anatolian Fault (NAF). **B:** Magnified view of the fault zone in our study region. Larger font names are the main geologic units: Istanbul Zone, Armutlu Block, Almacik Mountains and Sakarya Terrane. The smaller font, italic names are segments and faults of the NAF: the Izmit-Sapanca segment, the Sapanca lake step-over, the Sapanca-Akyazi segment (which together constitute the Northern strand), the Karadere segment and the Southern strand (names following Barka et al., 2002). The Sapanca-Akyazi segment is made of the Sakarya fault and the Akyazi fault. The flat area around the Akyazi fault is referred to as the Akyazi plain. Both Lake Sapanca and the Akyazi plain are pull-apart basins. The large red star indicates the epicenter of the \( M_w 7.4 \) Izmit earthquake, and the small purple star indicates the epicenter of the \( M_w 7.2 \) Düzce earthquake. **C:** The seismic stations used in this study are from the temporary experiment DANA (70 stations, red triangles; DANA, 2012) and the permanent network (9 stations, black triangles; Kandilli Observatory And Earthquake Research Institute, Boğaziçi University, 1971). Each column of the DANA array is indexed by a letter and each row is indexed by a number (DA01, DA02, ..., DB01, ...).
4.2 Earthquake Detection and Location

4.2.1 Data

The continuous seismic data were recorded by broadband stations from the temporary array DANA (70 stations) and the permanent network KOERI (9 stations, see the locations in Figure 4.1.1, and the Data and Resources section). The time period covered by this study is set by the duration of the DANA experiment: 2012-05-04 to 2013-09-20. Sampling rates are 50 Hz for most stations and 100 Hz otherwise. We bandpass filtered the data between 2 Hz and 12 Hz, allowing us to downsample the time series to 25 Hz to make the computation less intensive and to eliminate low frequency noise. Useful microearthquake signal can also be found between 1 Hz and 2 Hz, but that frequency band is strongly contaminated by anthropogenic noise on some stations.

4.2.2 Method

We analyze the 2012-05-04/2013-09-20 time period with a fully automatized earthquake detection and location method. The workflow, summarized in Figure 4.2.1, consists of three stages:

1. Backprojection: Earthquake detection via systematic backprojection of array seismic energy.

2. Relocation: Automatic picking of the P- and S-wave first arrivals of the previously detected events with the phase picker PhaseNet (Zhu and Beroza, 2019), and relocation of these events by transferring these picks to the NonLinLoc software (Lomax et al., 2000, 2009).

3. Template matching: The successfully relocated earthquakes are used as template earthquakes in a matched-filter search to detect new, smaller earthquakes using the Fast Matched Filter software (Beaucé et al., 2018).
The detection method is discussed at length in Beaucé et al. (2019), although the relocation methodology has been fully automatized here with the addition of PhaseNet and NonLinLoc. We summarize the methodology in this section.

First, we systematically backproject the seismic energy recorded by the array of seismic stations onto a 3D grid that samples locations in the volume beneath the study region. Backprojection is now a widely used earthquake detection and source imaging method \(\textit{e.g.}\) Ishii et al., 2005; Walker et al., 2005; Honda and Aoi, 2009; Frank and Shapiro, 2014). We compute the composite network response (CNR, see Equation (4.1)):

\[
\text{CNR}(t) = \max_k \{ \text{NR}_k(t) \}; \quad \text{NR}_k(t) = \sum_{s,c} \text{env} \left( u_{s,c}(t + \tau_{s,c}^{(k)}) \right).
\]  

(4.1)

In this equation, \(t\) is the calendar time and \(\text{NR}_k(t)\) is the network response for source location indexed by \(k\) at time \(t\). \(\text{NR}_k(t)\) is the stack of the envelopes (the modulus of the analytical signal) of the seismograms \(u_{s,c}\) shifted in time by the moveout \(\tau_{s,c}^{(k)}\) on station \(s\) and component \(c\). The moveouts were computed using the ray-tracing software Pykonal (White et al., 2020) in the 1D velocity model due to Karabulut et al. (2011) shown in Table 4.1. The use of a 1D velocity model in this region can introduce significant errors in the earthquake locations because of the strong lateral velocity variations, in particular across the two strands of the NAF \(\textit{e.g.}\) Karahan et al., 2001; Kahraman et al., 2015; Papaleo et al., 2018). The model shown in Table 4.1 offered both a visually satisfying agreement between earthquake epicenters and fault surface traces, and consistency with a previous study of the same data set (Poyraz et al., 2015).

In the second step, all the events detected from the composite network response are processed with the deep neural network PhaseNet (Zhu and Beroza, 2019) to automatically pick the P- and S-wave first arrivals. These picks are then used by the location software NonLinLoc (Lomax et al., 2000, 2009) to get both the earthquake
The envelopes of the 3-component seismograms of the entire station network are systematically back-projected on a 3D grid of theoretical seismic sources to find when and where energy stacks coherently (composite network response).

The peaks of the composite network response are potential detections of earthquakes.

We automatically pick the first P- and S-wave arrivals with the phase picker PhaseNet (Zhu and Beroza, 2019), and use these picks in the relocation software NonLinLoc (Lomax et al., 2000, 2009).

The events that were successfully relocated form the final database of template earthquakes. Transient noise signals and poor quality earthquakes are discarded during that stage.

Relocated template earthquakes are used in a matched-filter search to detect new, smaller events (FMF routine, Beaucé et al., 2018). Newly detected events are stacked to enhance the SNR of the initial template.

The updated templates can be re-processed to improve their locations and re-run a matched-filter search.

Events are relocated with the double-difference relocation software GrowClust (Trugman and Shearer, 2017).

Figure 4.2.1: Summary flowchart of the earthquake detection and location method. For clarity, only a subset of stations is shown in the above panels, but all the analysis is carried on the 79 stations together. Template matching is performed on the 10 stations closest to the source and the detection threshold is set to $8 \times \text{RMS}$ of the correlation coefficients in a 30-minute sliding window. See Data and Resources for code availability.
<table>
<thead>
<tr>
<th>Depth (top of the layer, km)</th>
<th>$v_P$ (km/s)</th>
<th>$v_S$ (km/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>2.900</td>
<td>1.670</td>
</tr>
<tr>
<td>0</td>
<td>3.000</td>
<td>1.900</td>
</tr>
<tr>
<td>1</td>
<td>5.600</td>
<td>3.150</td>
</tr>
<tr>
<td>2</td>
<td>5.700</td>
<td>3.210</td>
</tr>
<tr>
<td>3</td>
<td>5.800</td>
<td>3.260</td>
</tr>
<tr>
<td>4</td>
<td>5.900</td>
<td>3.410</td>
</tr>
<tr>
<td>5</td>
<td>5.950</td>
<td>3.420</td>
</tr>
<tr>
<td>6</td>
<td>6.050</td>
<td>3.440</td>
</tr>
<tr>
<td>8</td>
<td>6.100</td>
<td>3.480</td>
</tr>
<tr>
<td>10</td>
<td>6.150</td>
<td>3.560</td>
</tr>
<tr>
<td>12</td>
<td>6.200</td>
<td>3.590</td>
</tr>
<tr>
<td>14</td>
<td>6.250</td>
<td>3.610</td>
</tr>
<tr>
<td>15</td>
<td>6.300</td>
<td>3.630</td>
</tr>
<tr>
<td>20</td>
<td>6.400</td>
<td>3.660</td>
</tr>
<tr>
<td>22</td>
<td>6.500</td>
<td>3.780</td>
</tr>
<tr>
<td>25</td>
<td>6.700</td>
<td>3.850</td>
</tr>
<tr>
<td>32</td>
<td>8.000</td>
<td>4.650</td>
</tr>
<tr>
<td>77</td>
<td>8.045</td>
<td>4.650</td>
</tr>
</tbody>
</table>

Table 4.1: 1D velocity model due to Karabulut et al. (2011) used in this study.

locations and the location uncertainties. We only keep events for which at least four P-wave picks and four S-wave picks, and a total of 15 picks were successfully determined by PhaseNet. Some of these events are still left with low quality picks (e.g. seismograms corrupted by other unrelated events) that do not enable NonLinLoc to find a solution to the location problem; these events are discarded. More information about the input parameters given to PhaseNet and NonLinLoc can be found in Supplementary Material.

The events that were successfully relocated are very likely to be actual earthquakes. In the final step, we use them in a matched-filter search to detect new, smaller magnitude earthquakes. Template matching is a powerful method for detecting low signal-to-noise ratio (SNR) events given prior knowledge of the target seismicity (e.g. Gibbons and Ringdal, 2006; Shelly et al., 2007; Ross et al., 2019). It consists of searching for all earthquakes with similar waveforms and moveouts to a known earthquake. These earthquakes share a similar location and focal mecha-
nism. Template matching is implemented by computing the correlation coefficient (CC, Equation (4.2)) between the template waveforms $T_{s,c}$ and the seismograms $u_{s,c}$ shifted by the template moveout $\tau_{s,c}$:

$$CC(t) = \sum_{s,c} w_{s,c} \frac{\sum_{n=1}^{N} T_{s,c}(t_n) u_{s,c}(t + t_n + \tau_{s,c})}{\sqrt{\sum_{n=1}^{N} T_{s,c}^2(t_n) \sum_{n=1}^{N} u_{s,c}^2(t + t_n + \tau_{s,c})}}.$$  \(4.2\)

where $w_{s,c}$ is the weight attributed to station $s$, component $c$, and $N$ is the length of the template waveforms. If all weights are equal and $\sum_{s,c} w_{s,c} = 1$, then Equation (4.2) is the arithmetic mean of the correlation coefficients. We run the matched-filter search on multiple nodes of a super-computer equipped with Graphic Processing Units (GPUs) using the template matching software Fast Matched Filter (Beaucé et al., 2018). We use a template length of 8 seconds and a detection threshold of 8 times the root mean square (RMS) of the correlation coefficients (CC) time series in a 30-minute sliding window ($8 \times$ RMS $\{CC(t)\}$).

The matched-filter search provides many repetitions of the template waveforms. The redundancy of information can easily be leveraged to improve the SNR of the original template waveforms by stacking the newly detected events’ waveforms. We use the Singular Value Decomposition and Wiener Filtering method (Moreau et al., 2017) to efficiently retrieve the coherent signal from a collection of many noisy realizations of the same pattern. These second generation template earthquakes with higher SNR are in turn used for repeating the relocation procedure and the matched-filter search. This detection/stacking/relocation workflow is commonly iterated several times in template matching studies. Iterating template matching and stacking makes neighboring template earthquakes converge to the same template since most of the events they detect are the same. We therefore lose small spatial scale details in the template earthquake database when iterating. In order to trade-off the SNR improvement with the loss of small scale information, we iterate only once in this study.

The final template matching catalog is refined by relocating precisely earthquakes
that are close to the station array with the double-difference method (e.g. Poupinet et al., 1984; Waldhauser and Ellsworth, 2000). P- and S-wave differential arrival times are computed by finding the lag times that maximize the inter-event correlation coefficients. These times are processed by the relocation software GrowClust (Trugman and Shearer, 2017, additional information on parameters are given in Supplementary Material).

4.3 Spatio-Temporal Distribution of the Seismicity

4.3.1 Regional Seismicity

Following the method described in Section 4.2.2, we built a database of 3,790 templates and with them detected 35,503 events. We note that the numbers of events reported here are after post-processing the template matching catalog. Neighboring templates often detect the same events, therefore we keep a single event out of all detections meeting the follow three criteria: i) occurring within three seconds of each other, ii) from templates whose uncertainty ellipsoids are separated by less than 5 km, and iii) with average waveform similarity greater than 0.33. Figure 4.3.1 shows the locations of the 3,320 template earthquakes that are shallower than 20 km and have horizontal uncertainties less than 15 km, as well as the cumulative detection count per template over the whole study period. The majority of earthquakes occurred outside the station array and not in the North Anatolian Fault zone (see Figure 4.3.1, bottom left panel). We purposely present an earthquake catalog for this region that extends far beyond the NAF itself to provide a comprehensive description of the earthquake signals found in the data set. We found that most of the 1,982 events detected with templates deeper than 20 km originated far outside the study region, in particular in the Hellenic and Cyprus subduction zones in the southwest and south of the study region, respectively. Therefore, we discarded these deeper templates for any further analysis. We present the spatio-temporal distribution of the seismicity
Figure 4.3.1: Map view of the locations of the template earthquakes detected and used in this study. Only templates with maximum horizontal uncertainty less than 15 km and depth less than 20 km are shown (total of 3,320 templates). Filled dots are for natural earthquakes (1,471 templates), and squares are for mining-related events (1,849 templates; see text for details about identifying templates as mining templates). A: Event depths. B: Cumulative number of event detections per template. Most of the earthquake signals found in the dataset actually originate from outside the North Anatolian Fault Zone. C: Maximum vertical uncertainty derived from the uncertainty ellipsoids. D: Maximum horizontal uncertainty derived from the uncertainty ellipsoids. As expected, template earthquakes that are located further from the stations have larger location uncertainties.

in Figure 4.3.2. The seismic activity of the region is stronger at the beginning of the study period, when we observe two sequences of slowly decaying activity due to aftershocks south of 39°N and around 40°N. The southernmost earthquake sequence is part of the aftershock activity of the M5.1 2012-05-03 39.18°N/29.10°E/5.4 km earthquake (right before the deployment of DANA), but the cause of the 40°N sequence is unclear.

Template matching lends itself particularly well to identifying sources of mining-
Figure 4.3.2: Spatio-temporal distribution of the earthquake activity in the study region. The longitude of each event is shown against its origin time, and the color codes the latitude. A: We detected 31,356 events with the 3,320 template earthquakes presented in Figure 4.3.1 from 2012-05-04 to 2013-09-20. B: The templates due to natural seismicity detected 16,708 earthquakes. The seismic activity taking place on the NAF (latitudes 40.40°N-40.80°N) is eclipsed by the numerous earthquakes occurring elsewhere.

related earthquakes, and we distinguish these from natural earthquakes in the following. We found that about half of the detected seismicity is due to mining activity. We identify the template earthquakes associated to mining-induced seismicity by analyzing the distribution of detection times within the day. Templates that detect more than 80% of events between 6am and 6pm are categorized as mining-related templates (see Figure 4.S.1 in supplementary), since we do not expect the natural seismicity to occur within preferred times. After removing the mining-related templates, we are left with 16,708 events (cf. Figure 4.3.2, bottom panel). The locations identified as mining-induced activity are indicated by squares in Figure 4.3.1. Both the locations
and the proportion of activity (~50% of the total activity) are consistent with the hand-picked earthquake catalog from Poyraz et al. (2015).

Our earthquake catalog and detection/location codes are available online (see Section Data and Resources, and see Supplementary Material for additional information about the catalog file structure). This analysis of the regional seismicity shows that most of the detected seismic activity occurs outside the North Anatolian Fault Zone (>50 km away from either strand), which may be a feature of this section of the NAF being early in the earthquake cycle (Ben-Zion and Zaliapin, 2020). The seismic activity of our primary interest – the activity in the fault zone – can be studied by focusing on the template earthquakes located in the vicinity of the NAF.

4.3.2 Seismicity in the North Anatolian Fault Zone

We narrow our analysis to the vicinity of the North Anatolian Fault Zone. Figure 4.3.3 shows the locations of the template earthquakes in the fault zone, as well as the relocated earthquakes. The median horizontal and vertical errors on relative locations are 73 m and 91 m, respectively, meaning that they can reliably be interpreted. The seismic activity is not hosted on simple linear faults, instead these observations emphasize the existence of a complex network of faults. In Figure 4.3.3, we introduce nine sub-regions that we will keep referring to in this manuscript. These are organized into four along-strike sections: Izmit-Sapanca, Sapanca-Akyazi, Karadere, and the Southern strand (as a whole), and five fault-perpendicular sections: Lake Sapanca west and east, Akyazi, and the Southern strand west and east. Only in Figure 4.3.3 we also show a fault perpendicular cross-section in the Sapanca-Akyazi region. The Northern strand features several areas of stronger activity: the two ends of Lake Sapanca and around the Akyazi fault, where the deepest events are observed (down to 20 km). A large number of earthquakes occurred north of the Sapanca-Akyazi segment and are part of the 2012-07-07 $M_L$4.1 Serdivan earthquake sequence. The seismicity along the Southern strand is more diffuse and seems to extend further
to the south than to the north of the fault. We observe an area of lesser activity in the middle of the Southern strand. Overall, the relocated hypocenters reveal a spatial distribution dominated by patches of earthquakes, which is partly due to the detection method. Indeed, template matching tends to detect groups of colocated earthquakes, whereas small events located in between template earthquakes may remain undetected.

The fault parallel and fault perpendicular cross-sections in Figure 4.3.3C show the events’ depth distribution. In the north, the seismicity on the Sapanca-Akyazi and the Karadere segments is enhanced in the lower half of the seismogenic zone, with most earthquakes located between 7 km and 15 km and even deeper than 15 km around the Akyazi fault. Earthquakes tend to occur at smaller depths in the west, along the Izmit-Sapanca segment. In addition to this east-west variation, we observe a sharp contrast in earthquake depths across the NAF at the western end of Lake Sapanca with shallow seismicity (∼5 km) south of the fault, probably taking place on shallow structures in the Armutlu bloc, and deeper seismicity (∼14 km) in the northern bloc. Interestingly, Hussain et al. (2016) pointed that a shallow slip deficit existed in the upper crust just west of Lake Sapanca (the Izmit coseismic slip was larger beneath 6 km than above). The observed shallow seismicity might therefore be fed by residual high stresses. In the south, the seismicity along the Southern strand shows a consistent depth range of 5-10 km both in the west and in the east. The shallower events observed in the east (<5 km depth) are most likely associated to secondary faults. The map views and cross-sections in Figure 4.3.3 point towards a narrower deformation zone in the north where seismicity is mostly distributed within 5-10 km of the main fault trace, whereas we observe a wider deformation zone along the Southern strand with seismicity distributed within 15-20 km of the fault trace. An important and unresolved question is whether the NAF is vertical or slightly north dipping along the Izmit-Sapanca and Sapanca-Akyazi segments (Cakir et al., 2003; Ergintav et al., 2009; Kahraman et al., 2015, and see Section 4.5.2). The microseismicity detected in this study does not necessarily illuminate the main fault
plane, but instead it highlights the deformation zone associated with the NAF. The fault perpendicular cross-section of the Sapaca-Akyazi segment could reveal a north dipping deformation zone ($\sim 60^\circ$, see Figure 4.3.3C), as hypothesized by Kahraman et al. (2015), however one could also argue that most earthquakes simply lie around 10 km depth. Similar arguments hold for the cross-section across the Akyazi fault: a small number of earthquakes could indicate a slightly north dipping deformation zone ($\sim 85^\circ$), whereas most earthquakes reveal nearly horizontal alignments. Along the Southern strand, particularly on the eastern section, one could either identify slightly south dipping structures ($\sim 85^\circ$) or strongly north dipping structures ($\sim 70^\circ$).

The temporal patterns of earthquake occurrence provide information about fault properties such as their structural complexity (e.g. fault roughness, fracture density, stress heterogeneity) or how the seismicity is driven locally (e.g. Ben-Zion and Lyakhovsky, 2006; Vidale and Shearer, 2006; Lohman and McGuire, 2007; Dublanchet et al., 2013). In the following, we analyze the temporal organization of seismicity and quantify to which extent earthquakes are self-triggered through cascade mechanisms or forced by an external process (e.g. large-scale tectonic deformation, fluid migration). Figure 4.3.4 shows the temporal distribution of earthquakes in the recurrence time vs. origin time space. The recurrence time is the time interval between two consecutive co-located earthquakes, and the origin time is the time when the earthquake happens. In practice, recurrence times are computed as the time intervals between consecutive events detected by a same template. The local magnitudes shown on this figure were computed from log ratios of peak amplitudes (see Figure 4.S.2 and Section 4.S.1 for more information). The Sapanca-Akyazi segment and its vicinity is the most active region with the largest magnitude events observed during the study period. Among the nine $M_L \gtrsim 3$ natural earthquakes we detected, three occurred near each other, close to the city of Serdivan on a set of conjugate faults (cf. Figure 4.S.3). These include the largest event of the study: the 2012-07-07 $M_L 4.1$ Serdivan earthquake (30.404°E/40.763°N/11.3 km). This relatively large earthquake produced a surprisingly low number of aftershocks (see Figure 4.S.3). The area around

173
Figure 4.3.3: Earthquakes in the North Anatolian Fault Zone. **A:** Locations of the template earthquakes with color coded depths. We define nine subregions along the different segments of the fault. Only in this figure the Sapanca-Akyazi region is subdivided into a fault parallel and a fault perpendicular sections. The dashed lines inside each colored box define either fault parallel or fault perpendicular cross-sections (see bottom panels, **C**). The color shading of each box is only to help distinguish between them. **B:** Earthquake hypocenters successfully relocated with the double-difference method and color coded by depth. Events for which relocation was not successful were attributed the template location. **C:** Depth cross-sections of the different areas introduced above. The earthquake locations contained in the boxes are projected onto the boxes’ central axis (thin black dashed lines). The bottom x-axes are distances along the cross-section axes in kilometers, and the top x-axes are the geographic coordinates relevant to each cross-section (either longitude or latitude).

the Akyazi fault also produced four $M_L > 3$ earthquakes.

Vertical stripes on Figure 4.3.4 indicate bursts of seismicity when recurrence times
span many orders of magnitude and indicate earthquakes occurring at all time scales, from every few days/hours to every few seconds. We strictly refer to burst-like episodes as such sequences with large range of recurrence times. These bursts are usually associated with sequences of foreshocks-mainshock-aftershocks, although in general earthquake sequences can have no clear mainshock (an event of magnitude larger than all other events of the sequence) and still exhibit a strong burst-like behavior. The seismicity at the eastern end of Lake Sapanca is almost exclusively organized into such sequences of burst-like seismicity, whereas the Southern strand hosts much less of these burst-like episodes. Rather than simply reflecting a transient increase in seismicity rate, these bursts are characteristic of time clustered earthquake sequences.

In this study, we define temporal clustering as the property of earthquake sequences where past events influence the timings of future events, as opposed to random sequences where earthquake occurrence is well described by a Poisson point process (e.g. Gardner and Knopoff, 1974; Marsan and Lengline, 2008). Mathematically, temporal clustering does not automatically imply an enhanced seismic activity, and inversely an enhanced activity does not imply temporal clustering. We refer to earthquake occurrence as the time series of event count per unit time (Equation (4.3)):

$$ec(t_n) = \text{Number of events } \in [t_n; t_n + \Delta t], \quad (4.3)$$

where $\Delta t$ is a user-defined time bin duration. The wide range of recurrence times spanned by these burst-like episodes is evidence for time scale invariance. Therefore, the event count of a time clustered sequence exhibits a power law behavior in the spectral domain, and a fractal behavior in the time domain. We measure the fractal dimension of an earthquake occurrence time series by subsequently dividing the time axis into smaller and smaller time bins (varying size $\tau$), and counting the fraction of bins $x$ that are occupied by at least one earthquake (Smalley Jr et al., 1987; Lowen
Figure 4.3.4: Earthquake activity seen on recurrence time vs detection time graphs for different subsets of the earthquake catalog (refer to Figure 4.3.3 for the name of the areas). The recurrence time is the time between two consecutive events detected by a same template. Note that the y-axis is in log scale and that some seismic episodes span many orders of magnitude of recurrence time. These episodes are characteristic of burst-like, or cascade activity (see text). The color codes the local magnitude, and inverted grey triangles are events for which no reliable estimates were obtained.

and Teich, 2005). For a certain range of time bins $\tau$, we observe:

$$ x \propto \tau^{1-D}. $$

In Equation (4.4), $D$ is the fractal dimension of the time series. For a random time series, $D = 0$ and the fraction of occupied bins is inversely proportional to the bin size (a Poisson point process has a fractal dimension $D = 0$, cf. Figure 4.3.4). Large fractal dimensions ($D > 0.2$) characterize cascade-like activity where past events strongly influence the timings of future events. Fractal analysis has been used in mul-

176
multiple studies to characterize earthquake clustering (Smalley Jr et al., 1987; Lee and Schwarcz, 1995; Beaucé et al., 2019). Note that periodic seismicity does not follow a fractal behavior and fitting such a power-law to the data is not achievable.

Figure 4.3.5: A: Map view of template earthquakes with color-coded fractal dimension (cf. Equation (4.4)) showing the strength of temporal clustering. B: Map view of template earthquakes with color coded cumulative number of detections. In both top panels, the shaded areas refer to the regions introduced in Figure 4.3.3. C: Template earthquakes with color coded fractal dimension on fault parallel and fault perpendicular cross-sections (same color scale as panel A). Hypocenters are projected along the dashed axes shown on the map view. High fractal dimensions mean strongly time clustered activity (i.e. past events strongly influence the timings of future events).

We estimate the fractal dimension of the earthquake occurrence of each template’s family of events (see Figure 4.3.5). The most strongly time clustered activity is located in the step-over area between the Izmit-Sapanca and the Sapanca-Akyazi segments,
namely beneath Lake Sapanca. More precisely, strong clustering takes place along the eastern shore of Lake Sapanca, beneath the so-called Rangefront trace. This is evidence that, there, the patterns of earthquake occurrence are shaped by a local feedback mechanism between fault elements (i.e. fault patches interact) rather than by a remote external forcing such as the large scale tectonic forcing. Fault patches (including seismic and aseismic sources) interact through their stress fields, either on short spatial scales (distances $\sim$ fault length) with static stress changes (e.g. King and Cocco, 2001) or dynamic stress changes due to passing waves (e.g. Fan and Shearer, 2016), or on large spatial scales (regional or teleseismic distances) only with dynamic stress changes (Prejean et al., 2004, static stress changes rapidly decrease with distance, e.g.). A complex fault geometry, either with multiple faults or a single rough fault, offers favorable conditions for interactions between earthquake sources on the same fault or on several nearby faults. On short spatial scales, interaction can also take place between the stress fields of a seismic source and a creeping fault (e.g. Dublanchet et al., 2013; Cattania, 2019). Seismogenic asperities embedded in a stably creeping fault tend to produce time clustered earthquake sequences due to interactions enhanced by creep mediated stress transfers (Dublanchet et al., 2013). These earthquake-earthquake or earthquake-creep interactions result in a clock advance or delay in the cycle of the earthquake sources (e.g. Harris et al., 1995; Gomberg et al., 1998). We note that when seismic and aseismic sources are not juxtaposed, slow slip rather promotes swarm-like (weak interactions/low fractal dimension) activity of the seismic sources by increasing the stressing rate on these (e.g. Lohman and McGuire, 2007).

Beneath Lake Sapanca, the high density of faults that is known to exist in this step-over area is one element that creates favorable conditions for earthquake temporal clustering. However, the step-over beneath the Akyazi plain also presents a complex fault network but no significant clustering. We discuss in details the possible mechanisms for strong clustering beneath Lake Sapanca and nonexistent clustering in the Akyazi step-over in Section 4.5.3. Several studies have pointed at the existence
of shallow creep on the Izmit-Sapanca and the Sapanca-Akyazi segments (e.g. Çakir et al., 2012; Hussain et al., 2016; Aslan et al., 2019), but we do not observe temporal clustering nor strong swarm-like activity among shallow templates. A few isolated locations also show strong clustering, and these locations are always at the bottom of the seismogenic zone (10 to 15 km depth, depending on the areas), notably around the Izmit earthquake hypocenter. Figure 4.3.6 shows the general trend for clustering to be stronger at the bottom of the seismogenic zone. It also shows that strong clustering almost always occurs in regions of high event density, but event density alone does not seem to entirely control clustering. Note that event density is not a direct measure of asperity density, since each asperity can produce many earthquakes, and therefore interpreting from data the relationship between clustering and asperity density is limited. Thus, this suggests that clustering is favored by the frictional properties at the transition zone between unstable and stable sliding (cf. discussion in Section 4.5).

Figure 4.3.6: Clustering vs. depth vs. event density. Inside each region, templates are binned per distance from the bottom of the seismogenic zone and the fractal dimension is averaged among the 10% largest values (one dot per bin). The location of the bottom of the seismogenic zone is approximated by the depth of the locally deepest template. Dots are colored according to the average inter-event distance in the cloud of earthquakes detected by the selected template; this is a proxy for event density. Darker colors mean higher event density. Strongest clustering tends to occur at the bottom of the seismogenic zone, i.e. at the transition zone between stable and unstable sliding.

The group of templates near Serdivan (longitude 30.40°E, ~ 10 km depth on the
Sapanca-Akyazi cross-section in Figure 4.3.5) all present moderate fractal dimensions no larger than 0.15. Given that the Serdivan sequence hosted the largest magnitude events recorded during the DANA experiment in the vicinity of the NAF, we would have expected stronger time clustering. However, as mentioned earlier, the aftershock activity after the 2012-07-07 $M_L 4.1$ Serdivan earthquake was extremely short: we detected earthquakes for the following five hours only (7am-12pm, noon). Our analysis did not reveal any signs of fluid migration, which could have explained the observed moderate clustering (see Figure 4.8.3). The Serdivan sequence was not the primary focus of this study, and therefore we did not push the analysis further. However, arguments about clustering detailed in Section 4.5 could apply to this system of faults.

### 4.4 Stress State of the North Anatolian Fault Zone

Previous studies based on long-term and large magnitude seismicity (e.g. Kiratzi, 2002; Pınar et al., 2010) have shown that the maximum compressive stress in the region is horizontal and oriented NW-SE, while the least compressive stress is horizontal and oriented NE-SW (see Table 4.2). Therefore, the predominant faulting regime in large earthquakes is right-lateral strike-slip on E-W oriented faults. However, when characterizing the stress state at smaller scales (e.g. single fault segment), using low magnitude ($M < 4$) earthquake data, the local stress tensors appear to be significantly different from the regional stress tensor (e.g. Pınar et al., 2010; Ickrath et al., 2015).

In this study, we do not focus on the largest magnitude earthquakes, but instead use the detected microearthquakes to characterize stress at the single fault segment scale.

#### 4.4.1 Focal Mechanisms

In this section we present the methodology we followed to estimate the focal mechanisms of all template earthquakes located in the region of interest (regions defined in Figure 4.3.3). Working with template earthquakes instead of individual earthquakes
helps extract information from small magnitude seismicity because of templates’ enhanced SNR, however at the cost of losing fine spatial scale resolution since template earthquakes are centroids of clouds of earthquakes. In that sense, our focal mechanisms are composite focal mechanisms.

For each template, we synthesize waveforms from the template location to every seismic station in the 1D velocity model given in Table 4.1 for an array of possible source mechanisms. The synthetic waveforms are computed with Axitra (Cotton and Coutant, 1997), which is based on the discrete wavenumber method (Bouchon, 1981). Assuming double-couple sources, they are described by the set of three parameters: strike (azimuth of fault), dip (angle from horizontal) and rake (angle between slip direction and horizontal)

We fit the displacement waveforms in the 2-6 Hz frequency band, thus cutting off the low frequency, low SNR content of the data and the high frequencies for which one needs an accurate description of both the source time function and velocity model to fit the complex waveforms. We found that synthesizing the source time function with a Ricker wavelet of width $\sim 0.25\,\text{s}$ fit the first arrivals well. Varying the width of the Ricker wavelet around this value does not significantly influence the results due to the narrow frequency band we utilize. The objective function that we seek to minimize over a 3D grid of strikes/dips/rakes is a combination of the RMS difference and correlation of observed and synthesized waveforms:

$$
\Psi_s(\phi_i, \delta_j, \lambda_k) = A \sqrt{\sum_{p_{hi}=P,S} \sum_{n=1}^{N} \left( u_{\text{pred}, s}^{ph}(t_n; \phi_i, \delta_j, \lambda_k) - u_{\text{obs}, s}^{ph}(t_n) \right)^2} 
- B \sum_{p_{hi}=P,S} \sum_{n=1}^{N} u_{\text{pred}, s}^{ph}(t_n; \phi_i, \delta_j, \lambda_k) u_{\text{obs}, s}^{ph}(t_n),$$

$$
\Psi(\phi_i, \delta_j, \lambda_k) = \sum_{s \in S} w_s \Psi_s(\phi_i, \delta_j, \lambda_k). \tag{4.5}
$$

In Equation (4.5), $\Psi_s(\phi_i, \delta_j, \lambda_k)$ is the objective function on station $s$ for the strike, dip
and rake of grid point \((i, j, k)\). \(u_{\text{pred}}^{ph}(t_n)\) and \(u_{\text{obs}}^{ph}(t_n)\) are the synthetic and observed waveforms of phase \(ph\) (\(P\) or \(S\)), respectively, at time \(t_n\). The two constants \(A\) and \(B\) are designed to make sure the two terms of the objective function are of comparable scales; they are the standard deviations of the RMS and correlation terms over the whole grid, respectively. The addition of the correlation term aims to force the solution to produce correctly polarized waveforms, even for low amplitude arrivals (see the different terms in Figure 4.S.6). All stations from the station set \(S\) are included in a weighted sum, where each weight \(w_s\) is proportional to the SNR and inversely proportional to the source-receiver distance. The focal mechanism is described by the set of best strike, dip and rake \((\phi^*, \delta^*, \lambda^*)\) that minimizes \(\Psi\) (see the waveform fitting in Figure 4.4.1). We estimate the uncertainties by generating an ensemble of 100 solutions for each template earthquake (see Figure 4.4.1). The set of stations \(S\) is perturbed by randomly deleting 10 stations (delete-10 jackknife resampling) and the reduced station set is used to minimize the objective function. The final focal mechanism solution is obtained by averaging the 100 moment tensors (normalized to \(M_w=1\)) obtained from the 100 focal mechanisms.

Figure 4.4.2 shows the 436 focal mechanisms that we estimated for earthquakes located in the vicinity of the NAF. The predominant faulting regimes are right-lateral strike-slip faulting and E-W extending normal faulting: 842 events and 814 events were attributed a strike-slip and normal faulting mechanism, respectively, against 447 events for reverse faulting. Normal faulting is particularly present along the Sapanca-Akyazi segment and at its edges: on the eastern side of Lake Sapanca and beneath the Akyazi plain. Normal faulting is also the most represented faulting regime on the Southern strand, but with a N-S tension axis (see Figure 4.S.7 for distribution of faulting regimes along different segments). Reverse faulting is less common and is often mixed with an oblique faulting component, and does not occur in any clear spatial pattern at the exception of near Lake Sapanca. Indeed, the western edge of Lake Sapanca exhibits a large amount of reverse faulting, at the same location where we observe numerous shallow earthquakes (see Section 4.3.2). The Kaverina diagram
Figure 4.4.1: Estimation of focal mechanisms. **A:** We find the synthetic waveforms that fit best the first P- and S-wave arrivals on all seismic stations and components. Black waveforms are data, blue and red dashed waveforms are synthetic P and S waves, respectively (Cotton and Coutant, 1997, computed with Axitra). The waveforms are bandpass filtered in 2-6 Hz. Fine waveform realignment is performed by cross-correlating the synthetics and data to account for errors in location. **B:** To estimate uncertainties on focal mechanisms, we generate 100 focal mechanism solutions by randomly excluding 10 stations each time (leave-10-out jackknife resampling). The colored focal mechanism on the left is the solution from the minimum of the averaged 100 objective functions, and on the right is the solution from averaging the 100 (normalized) moment tensors. We chose to use the latter as our best solution. **C:** Location of the template earthquake taken as an example. The solution shows a right-lateral strike-slip fault striking to the northeast.

shown in Figure 4.4.2 demonstrates the proximity of many reverse faulting events to the strike-slip space, meaning that pure reverse faulting is quite rare. By not restricting the analysis to the largest magnitude events, this study characterizes the complex network of faults that is expected to exist in shear zones such as the NAF (Dresen, 1991, cf. Riedel shears, *e.g.*), and where the stress state is likely to vary substantially over short spatial scales.

### 4.4.2 Stress Tensor

Our stress tensor inversion uses the iterative linear method developed by Beaucé et al. (2021a), which is based on Michael (1984). This approach relaxes the assumption of constant shear stress magnitudes on all faults. It still relies on the other two assump-
Figure 4.4.2: Focal mechanisms of all 436 template earthquakes located in the vicinity of the North Anatolian Fault, following the method described in Section 4.4.1. The beachballs are lower hemisphere P-wave radiation patterns and their size is scaled according to their uncertainty (smaller beachballs have larger errors). **A:** Strike-slip faulting earthquakes. **B:** Reverse faulting earthquakes. **C:** Normal faulting earthquakes. **D:** Distribution of faulting regimes with depth. Each event of the catalog is attributed the focal mechanism of the template to which it correlates best. **E:** Focal mechanisms located in the so-called Kaverina diagram (Kaverina et al., 1996), which we use to categorize the faulting regimes. The colors of the symbols match the other panels. This panel was created with a plotting routine from the focal mechanism analysis software FMC (Álvarez-Gómez, 2019).

From the focal mechanisms we infer the principal stress directions (the eigenvectors of the stress tensor, see Figure 4.4.3) and a scalar parameter measuring the relative...
magnitudes of the principal stresses called the shape ratio $R$:

$$R = \frac{\sigma_1 - \sigma_2}{\sigma_1 - \sigma_3}. \quad (4.6)$$

In Equation (4.6), the $\sigma_i$ are the principal stresses (the eigenvalues of the stress tensor) ordered from most compressive ($\sigma_1$) to least compressive ($\sigma_3$). We propagated the errors in focal mechanisms into the stress inversion by randomly sampling the set of 100 solutions available for each focal mechanism and thus producing 2000 resampled data sets. Inverting our entire data set yields a regional stress tensor with NE-SW horizontal least compressive axis, NW-SE almost horizontal most compressive axis and a shape ratio $R \approx 0.5$ (see Table 4.2 and Figure 4.8.8). This is consistent with previous studies (Kiratzi, 2002; Pınar et al., 2010; Poyraz et al., 2015). To describe local variations in the stress state, we perform the stress tensor inversion in the nine subsections defined above; results are summarized in Table 4.2. Shape ratios and their uncertainties are shown in Figure 4.4.3.

Along the Northern strand, our analysis reveals that the stress state around the Izmit-Sapanca segment shows NW-SE horizontal compression and NE-SW horizontal extension, thus favoring right-lateral strike-slip faulting on roughly E-W striking faults. These orientations of principal stresses are consistent with the regional stress tensor. However, the shape ratio $R = 0.71$ indicates a mixed reverse faulting component (transpressional regime), which is consistent with the numerous observations of reverse faulting at the western side of Lake Sapanca. When only inverting data from western Lake Sapanca, the stress state is even more favorable to reverse faulting, as the extensional axis is plunging by about $40^\circ$ and the shape ratio is $R = 0.80$. The stress state along the Sapanca-Akyazi segment shows rotated maximum and intermediate compression axes about the minimum compression axis. Although the plane containing $\sigma_1$ and $\sigma_2$ has the same orientation as for the Izmit-Sapanca segment, the maximum compression axis is now almost vertical, meaning that local forces are dominated by extension and thus favor normal faulting. The shape ratio of $R = 0.40$ and the plunge of the principal stresses indicate a mixed strike-slip component. Inverting
Table 4.2: Compilation of results. †: Method from Gephart and Forsyth (1984). §: Method from Michael (1984, 1987). *: Method from Beaucé et al. (2021a). Note that for axes with plunges close to 0°, azimuth differences near 180° mean the axes are nearly colinear.

<table>
<thead>
<tr>
<th>Source</th>
<th>Time Covered</th>
<th>Area Covered</th>
<th>$\sigma_1$</th>
<th>$\sigma_2$</th>
<th>$\sigma_3$</th>
<th>$R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kiratzi (2002)†</td>
<td>1943-1999</td>
<td>NAF</td>
<td>126°/3°</td>
<td>269°/86°</td>
<td>35°/3°</td>
<td>0.5</td>
</tr>
<tr>
<td>Pınar et al. (2010)‡</td>
<td>1943-1999</td>
<td>NAF</td>
<td>308°/3°</td>
<td>91°/86°</td>
<td>218°/2°</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Izmit</td>
<td>284°/15°</td>
<td>154°/55°</td>
<td>25°/24°</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sapanca</td>
<td>5°/65°</td>
<td>286°/5°</td>
<td>198°/25°</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Akyazi</td>
<td>348°/52°</td>
<td>222°/25°</td>
<td>118°/27°</td>
<td>0.83</td>
</tr>
<tr>
<td>Bohnhoff et al. (2006)†</td>
<td>1999-08-17/</td>
<td>NAF</td>
<td>66°/87°</td>
<td>327°/0°</td>
<td>237°/3°</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Izmit-Sapanca</td>
<td>117°/15°</td>
<td>347°/67°</td>
<td>212°/17°</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Akyazi</td>
<td>190°/74°</td>
<td>3°/15°</td>
<td>94°/1°</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Karadere</td>
<td>339°/6°</td>
<td>49°/74°</td>
<td>251°/15°</td>
<td>0.37</td>
</tr>
<tr>
<td>Poyraz et al. (2015)†</td>
<td>1943-2003</td>
<td>NAF</td>
<td>110°/0°</td>
<td>201°/58°</td>
<td>20°/32°</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>2011-2013</td>
<td></td>
<td>103°/27°</td>
<td>256°/61°</td>
<td>7°/11°</td>
<td>0.45</td>
</tr>
<tr>
<td>This study*</td>
<td>2012-05-04/</td>
<td>NAF</td>
<td>305°/25°</td>
<td>127°/65°</td>
<td>35°/1°</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>2013-09-30</td>
<td></td>
<td>309°/10°</td>
<td>182°/73°</td>
<td>42°/13°</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lake Sap. W.</td>
<td>316°/4°</td>
<td>217°/66°</td>
<td>48°/24°</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lake Sap. E.</td>
<td>302°/55°</td>
<td>145°/33°</td>
<td>48°11°</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sap.-Akyazi</td>
<td>295°/61°</td>
<td>148°/25°</td>
<td>51°/14°</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Akyazi</td>
<td>339°/36°</td>
<td>162°/54°</td>
<td>70°/1°</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Karadere</td>
<td>314°/20°</td>
<td>220°/12°</td>
<td>100°/67°</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S. strand</td>
<td>96°/21°</td>
<td>288°/68°</td>
<td>188°/4°</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S. strand W.</td>
<td>90°/47°</td>
<td>262°/43°</td>
<td>356°/4°</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S. strand E.</td>
<td>299°/34°</td>
<td>72°/45°</td>
<td>190°/26°</td>
<td>0.57</td>
</tr>
</tbody>
</table>

only data from eastern Lake Sapanca confirms the predominance of normal faulting at the junction between the Sapanca-Akyazi segment and the step-over area. At the eastern end of the Sapanca-Akyazi segment, beneath the Akyazi plain, the horizontal stresses are rotated clockwise by 40° and the stress state promotes a mixture of strike-slip and normal faulting (transstensional regime, $R = 0.32$). Further east, along the Karadere segment, we found a stress tensor with maximum compression similar to that of the regional state, but the stress tensor favors mixed strike-slip and reverse faulting (transpressional regime, $R = 0.76$). The Karadere segment is the furthest from the station array (cf. Figure 4.1.1) and the source-receiver geometry is far from ideal for robustly constraining the focal mechanisms. Therefore, further interpretation of the Karadere results is not warranted.
Figure 4.4.3: Inverted local stress tensors linked on the map to their corresponding subregion. The stereographic projections show the orientations of the three principal stresses $\sigma_1$, $\sigma_2$ and $\sigma_3$, from most compressive to least compressive. The grey symbols are the 2000 solutions obtained by repeating the inversion on randomly sampled data sets from the 100 focal mechanism solutions available for each template earthquake. See Table 4.2 for numerical values of azimuth and plunge. We use these empirical distributions to estimate the 80% and 90% confidence intervals, shown here with the darker and lighter contour lines. The histograms show the distributions of shape ratios ($R$, see Equation (4.6)) and the vertical red bars indicate the shape ratios obtained by inverting the set of best focal mechanisms. In terms of deviatoric stress, $R < 0.5$ indicates that $\sigma_2$ is compressional, whereas $R > 0.5$ indicates that $\sigma_2$ is extensional.

The stress tensor has a significantly different orientation along the Southern strand. When taken as a whole, the inversion of the Southern strand data reveals an E-W horizontal maximum compression, and the shape ratio $R = 0.30$ indicates
that strike-slip co-occurs with normal faulting (transtensional regime). Right-lateral strike-slip faulting is favored on faults striking roughly NE-SW (which is the global orientation of the Southern strand). Dividing the Southern strand into a western part and an eastern part suggests that the plane containing $\sigma_1$ and $\sigma_2$ rotates clockwise by a few degrees from west to east. The uncertainties show nevertheless that these differences are hardly meaningful and the stress state might be homogeneous along the Southern strand.

4.5 Discussion

4.5.1 Summary of the Izmit Earthquake Sequence

On 17 August 1999, the M7.4 Izmit earthquake nucleated at the eastern end of the Izmit Bay and ruptured four fault segments, including the three segments of the Northern strand studied here (Barka, 1999). The Izmit earthquake produced 1-3 m of coseismic slip on the Izmit-Sapanca segment, around 1 m on the Karadere segment, and up to 5 m on the Sapanca-Akyazi segment (Barka et al., 2002; Langridge et al., 2002). However, almost no coseismic slip was observed at the surface near Lake Sapanca and on the Akyazi fault. The Izmit-Sapanca segment recorded 1.5 m of slip where it enters the western side of Lake Sapanca, and on its eastern side the coseismic slip was about zero but peaked around 5 m just 2 km further east (Langridge et al., 2002). The coseismic slip on the Akyazi fault was no more than 1 m, and even zero at its eastern end, in the Akyazi gap, where the rupture jumped on the Karadere segment. Strong aftershock activity, mostly made of normal faulting events, was observed in the vicinity of the Akyazi fault but little activity, in comparison, was recorded beneath Lake Sapanca (e.g. Ozalaybey et al., 2002; Bohnhoff et al., 2006). Fault slip models (e.g. Delouis et al., 2002; Cakir et al., 2003) suggest that significant coseismic slip occurred at depth beneath Lake Sapanca, which could explain the absence of aftershocks.
4.5.2 Evolution of the Seismicity and Stress State

The middle sections of the Izmit-Sapanca and Sapanca-Akyazi segments were particularly active seismically before the Izmit earthquake, and some clusters of earthquakes were observed beneath Lake Sapanca (Ickrath et al., 2015). The Izmit earthquake is known to have nucleated near a swarm of seismicity that had been active for decades prior to the M7.4 event (Crampin et al., 1985; Lovell et al., 1987). In between the Izmit earthquake and the Düzce event, the seismic activity was strongest on the western side of the Izmit-Sapanca segment, around the Izmit hypocenter, and on the eastern side of the Sapanca-Akyazi segment, in particular beneath the Akyazi gap where little coseismic slip was observed (Ozalaybey et al., 2002; Bohnhoff et al., 2006, 2008).

Our earthquake catalog shows that 13 years after the Izmit earthquake, the strongest seismic activity took place beneath Lake Sapanca, at its eastern end (see Figure 4.3.5). Other areas of high activity are the western end of Lake Sapanca, the eastern section of the Southern strand, and beneath the Akyazi plain. As the spatial distribution of seismicity differs from the pre-Izmit distribution, these areas of enhanced seismicity must be symptomatic of perturbed stress either from Izmit’s coseismic or postseismic slip. These areas are locations of geometric irregularities (step-overs or bifurcations) where we expect coseismic slip of large earthquakes to concentrate stress because of slip barriers. Around the time the seismic data were collected (2012-2013), postseismic slip was still known to occur at significant rates: Ergintav et al. (2009) estimated that GPS stations would return to within 1mm/yr of their pre-Izmit rate in 20-100 years after the mainshock (i.e. by ∼ 2020-2100). In contrast, the middle sections of the fault segments along both strands show low seismic activity, meaning that stresses are still too low to drive a sustained low magnitude seismicity. We note that the distribution of seismicity along the Southern strand cannot be explained by the patterns of Coulomb stress change usually computed for the Izmit earthquake (Parsons et al., 2000; Utkucu et al., 2003). However, the patterns of Coulomb stress change resolved on the secondary faults might be significantly
different from these large scale models, and consistent with the observed seismicity.

The regional stress state consists of a horizontal NW-SE most compressive axis \( \sigma_1 \) and a horizontal NE-SW least compressive axis \( \sigma_3 \) \( \text{(e.g.} \text{Kiratzi, 2002; Pınar et al., 2010, and this study, Figure 4.5.8)} \). This is consistent with the notion that the NAF mostly accommodates the westward motion (relative to stable Eurasia) of the Anatolian plate due to the pushing Arabian plate \( \text{(McKenzie, 1978; Le Pichon and Angelier, 1979)} \). Even before the Izmit earthquake, it was known that the stress state locally differed significantly from the regional stress. The Izmit earthquake nucleated in an area that was characterized by swarm seismicity made of normal faulting events \( \text{(e.g.} \text{Lovell et al., 1987)} \) driven by the N-S extensional forces due to the rollback of the Hellenic trench in the south of Anatolia \( \text{(McKenzie, 1978; Le Pichon and Angelier, 1979)} \). Just prior to the Izmit mainshock, the region between Lake Sapanca and the Karadere segment was mostly characterized by normal faulting events accommodating ENE-WSW extension \( \text{(Ickrath et al., 2015)} \). Following the Izmit earthquake, several studies indicated that the stress state was strongly perturbed \( \text{(Bohnhoff et al., 2006; Pınar et al., 2010; Ickrath et al., 2015)} \). Despite variations in their results, they all point at weakening of the fault induced by coseismic and postseismic slip. The seismicity in the Sapanca-Akyazi region became purely E-W extensional normal faulting \( \text{(i.e.} \sigma_3 \text{parallel to NAF and vertical} \sigma_1 \text{)} \) and a significant normal faulting component appeared along the Izmit-Sapanca segment with a fault-perpendicular \( \sigma_3 \). Note that fault-parallel or fault-perpendicular \( \sigma_1 \) or \( \sigma_3 \) implies low shear stresses, therefore a weak fault (low friction coefficient). \text{Pınar et al. (2010)} \text{proposed that postseismic slip could explain this apparent fault weakening.}

The results of our study (Figure 4.4.3) indicate that the local stress states along the Northern strand all show NW-SE maximum compression (with varying plunge) and SW-NE minimum compression. We conclude that the mechanism(s) that made the NAF weak had terminated by the time of the DANA experiment, which was probably a consequence of both the coseismic slip and the rapid afterslip that oc-
curred in the following months (Ergintav et al., 2009). This may also be the sign of fault weakening and healing following a large earthquake, as described by models such as rate-and-state friction (e.g., Dieterich, 2007). However, the local stress states derived in this study are not identical to pre-Izmit local states (Ickrath et al., 2015). In particular, while the Sapanca-Akyazi region was dominated by almost purely E-W extension prior and just after the Izmit earthquake, we observe a SW-NE extension in the stress tensors (Figure 4.4.3), and in the slip directions given by the normal faulting focal mechanisms (see Figure 4.4.2 and Figure 4.5.1). The seismicity around Akyazi even shows N-S extension.

This N-S component in the surface motion has been identified in InSAR data (Cakir et al., 2003) and GPS site velocities (Ergintav et al., 2009), but has been neglected in subsequent InSAR studies as it is not required to understand the large scale surface displacement field (Hussain et al., 2016; Aslan et al., 2019). Nevertheless, our study highlights the importance of normal faulting and SW-NE extension in the Sapanca-Akyazi area. Given that kinematic and dynamic models of postseismic slip (e.g., Ergintav et al., 2009; Hearn et al., 2009) cannot reproduce the observed N-S motion without invoking a north dipping NAF with a slightly plunging rake (Cakir et al., 2003), or aseismic slip taking place on normal faults (Ergintav et al., 2009), we suggest that the observed seismicity and stress state is related to this missing N-S component in the models.

The microseismicity detected in this study (cf. Figures 4.3.3 and 4.3.5) does not show evidence for a north dipping NAF along the Northern strand. Part of this lack of evidence might be due to the seismicity taking place on secondary faults, which do not inform about the shape of the NAF itself. Our observations of earthquake temporal clustering can help speculate about the role of aseismic slip on normal faults. According to Ergintav et al. (2009), N-S extension is strongest around Lake Sapanca and in the eastern Marmara Sea, and accelerated following the Izmit earthquake. Even though the latter region is not included in our detailed analysis, we detected strongly
clustered seismicity in both regions (see large map of clustering in Figure 4.5.1, clustering beneath Lake Sapanca is well described in Figure 4.3.5). In the following we investigate the differences between Lake Sapanca and Akyazi, two step-over regions, to understand better why Lake Sapanca features such strong time clustering and shed light onto the aseismic processes at play.

4.5.3 Lake Sapanca vs. Akyazi: Two Step-Overs, Two Behaviors

Lake Sapanca and the Akyazi plain are two step-over areas that accommodate motion between misaligned fault segments. Even though the gap between the Akyazi fault and the Karadere segment is a restraining step-over (jump to the left on a right-lateral structure), the Akyazi plain on a larger scale appears to be a pull-apart basin (Bohnhoff et al., 2006), similarly to Lake Sapanca. The whole area between the Izmit rupture trace and the Mudurnu fault is actually made of releasing structures (Pınar et al., 2010). Thus, both regions host extensional normal faulting (see Figure 4.4.2 and 4.5.1) and vertical offsets were observed during the Izmit earthquake (Langridge et al., 2002).

Figure 4.5.1 suggests that the group of earthquakes east of Lake Sapanca accommodate slip in the NE-SW direction, whereas the Akyazi plain shows more complex patterns of mixed normal and strike-slip faulting that accommodate motion along the Akyazi fault, but also NW-SE motion. Thus, the Sapanca-Akyazi segment is bounded in the west by a step-over (Lake Sapanca) made of multiple faults with similar orientations and slip directions, and in the east by a step-over (Akyazi plain) made of multiple faults with heterogeneous orientations and slip directions, spread over a wider depth interval.

We recall that strong temporal clustering is observed east of Lake Sapanca whereas
weak temporal clustering is observed in the Akyazi region (see Figure 4.3.5). Furthermore, the frequency-magnitude distribution (Gutenberg-Richter law) of the seismicity beneath Lake Sapanca is characterized by a higher b-value ($b=1.14$) than that of the seismicity beneath the Akyazi plain ($b=0.69$). This means that the fault system around Akyazi tends to produce a larger fraction of "large" magnitude earthquakes. Interpretation of b-values usually implies that regions of high b-values experience low shear stress, and conversely, (e.g. Wiemer and Katsumata, 1999; Scholz, 2015) on the basis of laboratory experiments and theoretical models (e.g. Scholz, 1968; Wyss et al., 1973). Although it has been proposed that large b-values could be due to strong heterogeneities (Mogi, 1962, i.e. highly fractured medium,), the dependence on stress seems greater (Scholz, 1968, 2015). Various studies have shown that creeping faults produce high b-value seismicity (e.g. Amelung and King, 1997; Wiemer and Wyss, 1997). In agreement with high b-value/low stress drops, numerical models have suggested that asperities embedded in a weak creeping region are prone to producing time clustered seismicity (Dublanchet et al., 2013) with low stress drops due to the fault weakness (Tse and Rice, 1986).
We therefore propose that the differences between the seismicity of Lake Sapanca and that of Akyazi are either due to geometrical or rheological differences. 1) Geometry: The multitude of faults with coherent orientations and slip directions beneath the eastern end of Lake Sapanca favor strong interactions and therefore time clustered, cascade-like seismicity. Conversely, in the Akyazi region, faults with incoherent orientations and slip directions do not promote system-scale interactions (a vertical strike-slip fault barely changes the shear stress on a neighboring almost horizontal normal fault). The high stress environment suggests that residual stresses from the Izmit earthquake are the main driving mechanism of seismic activity around the Akyazi fault. 2) Rheology: The focal mechanisms beneath Lake Sapanca shown in Figure 4.5.1 indicate that NW-SE striking, northward dipping normal faults are at play. As suggested by Ergintav et al. (2009), this normal faulting structure could be experiencing aseismic slip in between seismogenic patches, thus providing another interaction mechanism to explain time clustered sequences (Dublanchet et al., 2013; Cattania, 2019). The high b-value is also consistent with aseismic slip (Amelung and King, 1997), and strongly clustered seismicity embedded in a creeping fault (Dublanchet et al., 2013). The global observation that clustering tends to occur at the bottom boundary of the seismogenic zone (see Figure 4.3.6) suggests a link with aseismic slip. The hypothesized aseismic slip would probably continue downdip of the seismogenic zone, and one would expect this downdip aseismic to modulate the seismic activity. Extending the time span of the earthquake catalog could help study the relationship between the seismicity beneath Lake Sapanca and aseismic slip.

These observations alone are not direct evidence for aseismic slip on normal faults, but they are consistent with it. Therefore, we suggest that this hypothesis should be further tested, for example by modelling the surface displacement caused by slip on normal faults, to clarify the implications of our observations on the dynamics of the fault system, and maybe explain the N-S extensional motion observed in geodetic data (Cakir et al., 2003; Ergintav et al., 2009). Because both the eastern Marmara
Sea and Lake Sapanca were identified as areas of anomalous N-S extension (Ergintav et al., 2009), and because the eastern Marmara Sea is the second region to feature strong temporal clustering beside Lake Sapanca (see Figure 4.S.5), we believe that slow slip on normal faults could be occurring in both regions.

4.6 Conclusions

We processed 1.5 years of continuous data collected during the DANA experiment (May 2012 - September 2013) with an automated earthquake detection and location method (Beaucé et al., 2019, see Section 4.2) and produced an earthquake catalog with 35,503 events (cf. Section 4.3). Due to the heavy mining industry in the study region, we found that about half of the detected events were induced or triggered by mining activity. Most of the seismicity took place outside of the North Anatolian Fault Zone itself, but our analysis brings nonetheless a great amount of information about the fault processes.

We analyzed our earthquake catalog to characterize the state of the North Anatolian Fault thirteen years after the 17 August 1999 $M_w$7.4 Izmit earthquake, in 2012-2013 when the data were collected. The detailed spatial distribution of earthquakes obtained with the double-difference relocation method (Trugman and Shearer, 2017, implemented with GrowClust) does not provide evidence for a dipping NAF neither along the Northern strand nor the Southern strand (cf. Figure 4.3.3). However, the small magnitude seismicity studied here mostly illuminates secondary structures rather than the main fault. Our analysis put a particular emphasis on quantifying the strength of temporal clustering in earthquake sequences, which is a manifestation of fault interactions (see Section 4.3.2). Observations of temporal clustering were complemented by composite focal mechanisms computed on the template earthquakes. These were in turn used to derive the regional and local stress state (Section 4.4).
Our findings show that the North Anatolian Fault Zone had (at least partially) recovered from the loss of frictional strength caused by the Izmit earthquake (Pınar et al., 2010; Ickrath et al., 2015) at the time of the DANA experiment, but also suggest that the fault system was not back to its preseismic state, probably due to postseismic phenomena (Ergintav et al., 2009; Hearn et al., 2009). We observe that the spatial earthquake distribution and the local stress state differed from both the pre- and early post-Izmit state (Ozalaybey et al., 2002; Bohnhoff et al., 2006), most strikingly beneath Lake Sapanca.

In an attempt to deciphering the postseismic processes at play, we analyzed the time clustering properties of the seismicity. The strongest temporal clustering was observed in the extensional normal faulting structures beneath Lake Sapanca, confirming the peculiarity of this area. Previous observations and models of earthquake clustering (Marsan and Lengline, 2008; Dublanchet et al., 2013; Cattania, 2019), the general trend for strong clustering to occur at the bottom of the seismogenic zone (Figure 4.3.6), and the comparison between Lake Sapanca and the Akyazi area (Figure 4.5.1) suggest that the observed time clustering is not only due to the density of seismic sources, but also to creep mediated stress transfers promoted by the frictional properties of the faults embedding these seismic sources. Therefore, as hypothesized previously as a potential explanation for the anomalous extension across Lake Sapanca and the eastern Marmara Sea observed in the Izmit postseismic period (Ergintav et al., 2009), we propose that aseismic slip may be occurring on normal faults beneath Lake Sapanca and, based on our extended study of temporal clustering (Figure 4.S.5), in the eastern Marmara Sea (Section 4.5). Extending the analysis of the seismicity in these two areas to longer times and analyzing it jointly with geodetic data could help elucidate the relation between extension and aseismic slip.
Data and Resources

The earthquake catalog is available for public use at E. Beaucé’s personal website https://ebeauce.github.io/material/. Our earthquake detection and location codes will be available as a Python package at https://github.com/ebeauce/Seismic_BPMF.

The topographic data used for the maps were taken from the Shuttle Radar Topographic Mission (SRTM) 90-m database (https://cgiarcsi.community/data/srtm-90m-digital-elevation-database-v4-1/). The seismic data were recorded by the temporary array DANA (DANA, 2012, DOI: https://doi.org/10.7914/SN/YH_2012) and by the permanent KOERI stations (Kandilli Observatory And Earthquake Research Institute, Boğaziçi University, 1971, DOI: https://doi.org/10.7914/SN/K0).

Figure 4.4.2E was made with the focal mechanism analysis software FMC https://github.com/ElsevierSoftwareX/SOFTX_2018_227 (last accessed June 2021).

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union’s Horizon H2020 research and innovation program (grant agreement No 742335). E.B. was also supported by funds associated with Robert D. van der Hilst’s Schlumberger chair.
4.S  Supplementary Material

4.S.1  Earthquake Catalog

Automated Phase Picking with PhaseNet

The threshold P- and S-wave probabilities to trigger a P- or S-wave pick with PhaseNet (Zhu and Beroza, 2019) is 0.6.

Absolute Earthquake Location with NonLinLoc

NonLinLoc (Lomax et al., 2000, 2009, NLLoc) offers different loss functions to minimize to find the best earthquake location given a set of P- and S-wave arrival times. Beside the classic L2 norm of the residuals, NLLoc can maximize the equal differential time (abbreviated EDT in the software) likelihood function, which is robust to outliers. Since outliers often arise in a fully automated method, the choice of the EDT likelihood function is key for producing correct earthquake locations.

The maximum of the EDT likelihood function is searched with the oct-tree importance sampling algorithm, which combines sampling with grid-search to speed up the grid-search method and use a smart grid that is finer in regions of higher likelihood. Our initial grid has 10 cells in longitude and latitude, and 6 cells in depth. We draw 5000 samples inside each cell and use the station density when deciding which grid cells to further subdivide. The initial grid has 1 km spaced points in the horizontal directions and 0.5 km in the vertical direction.

Double-difference Relative Relocation with GrowClust

GrowClust (Trugman and Shearer, 2017) is an earthquake relative relocation software based on the double-difference method. We compute the inter-event differential times on each station and component by cross-correlating the P-wave and S-wave first arrivals and search for the lag times that maximize the correlation coefficient ($CC$). P- and S-wave windows are 2 s long and start 0.4 s before the P and S wave, respectively,
the sampling rate is 50 Hz, and waveforms are filtered between 2 Hz and 12 Hz.

All differential time observations with $CC > 0.60$ (rmincut = 0.60 in the control file), and an event pair is kept only if the average CC is greater than 0.33 (rpsavgmin = 0.33 in the control file) and at least 5 differential time observations have $CC > 0.50$ (rmin = 0.50 and ngoodmin = 5).

**Earthquake Catalog File**

The earthquake catalog is a csv file with one row per event. The columns of the file are:

- origin_times: Origin times of the events.
- latitudes: Latitudes of the events, in decimal degrees.
- longitudes: Longitudes of the events, in decimal degrees.
- depths: Depths of the events, in km.
- max_hor_uncertainty: Maximum location uncertainty in the horizontal direction, in km.
- max_ver_uncertainty: Maximum location uncertainty in the vertical direction, in km.
- location_quality: 2 - good, 1 - intermediate, 0 - bad (do not trust it).
- magnitudes: Local magnitudes of the events. -10 if no estimate is available.
- fractal_dimensions: Fractal dimension of the earthquake occurrence time series of the template the event was detected with.
- tids: Template ID of the template that detected the event.
- mining_activity: True if the event was detected with a mining related template, False otherwise.
Identifying Mining Templates

See Figure 4.S.1.

Figure 4.S.1: **Top left panel:** Mining-related seismicity is characterized by predominantly diurnal seismicity, whereas we expect no preferred time for natural seismicity. In fact, natural seismicity shows slightly more events at night because noise is generally lower, and earthquake detection is easier. **Bottom left panel:** Mining-related seismicity also often shows no earthquakes on Sundays. **Right panels:** The waveforms produced by these mining-induced earthquakes have all characteristics of natural earthquakes, with clear P and S waves.

Magnitude Estimation

Local magnitudes were computed from the amplitude ratios of peak velocities. This requires estimating the magnitude of at least one event per template to calibrate our local magnitude scale. Within each family of earthquakes detected by a same template, we compute the S-wave spectra with the multi-taper method (Prieto et al., 2009). The SNR is computed in the spectral domain as the ratio of the S-wave spectrum to the spectrum of a noise window taken before the P wave. The SNR is used to compute the multi-channel weighted average of the S-wave spectra (cf.
Equation (4.7) and Figure 4.S.2).

\[
\bar{v}(f) = \frac{1}{W(f)} \sum_{s,c} w_{s,c} \alpha_{s,c} v_{s,c}(f), \quad W(f) = \sum_{s,c} w_{s,c}(f). \tag{4.7}
\]

In Equation (4.7), \( v_{s,c}(f) \) is the velocity spectrum of station \( s \), component \( c \) at frequency \( f \), \( w_{s,c} \) is the corresponding weight (see caption of Figure 4.S.2) and \( \alpha_{s,c} \) is the factor that corrects for geometric spreading and attenuation (see Equation (4.9)). The average spectra are converted to displacement spectra \( u(f) \) and fitted with the Brune model (Brune, 1970, Equation (4.8)),

\[
|u_{\text{Brune}}(f)| = \frac{\Omega_0}{\left(1 + \frac{f}{f_c}\right)^2}, \tag{4.8}
\]

where \( \Omega_0 \) is the low-frequency plateau, which is proportional to the seismic moment \( M_0 \), and \( f_c \) is the corner frequency. The successfully fitted spectra produce a seismic moment estimate using Equation (4.9) (Richards, 1971).

\[
|u^S(f)| = \frac{R^S}{2\rho\beta^3 r} \frac{M_0}{\Omega_0} \exp\left( -\frac{\pi f t_{s,c}^S}{Q_S(f)} \right), \tag{4.9}
\]

\[
\Rightarrow M_0 = \frac{\Omega_0 2\rho\beta^3 r}{R^S} \exp\left( \frac{\pi f t_{s,c}^S}{Q_S(f)} \right),
\]

\[
\Rightarrow \alpha_{s,c} = \frac{2\rho\beta^3 r_{s,c}}{R^S} \exp\left( \frac{\pi f t_{s,c}^S}{Q_S(f)} \right).\]

In Equation (4.9), we use typical values for the S-wave velocity \( \beta \) (3000 km/s), the density of crustal rocks \( \rho \) (2700 kg/m\(^3\)) and the average S-wave radiation pattern \( R^S \) (\( \sqrt{2/5} \) from Aki and Richards, 2002). The source-receiver distance \( r_{s,c} \) and the S-wave travel time \( t_{s,c}^S \) are computed from the source location and velocity model. Finally, a frequency dependent quality factor is obtained from Izgi et al. (2020). The magnitude moment \( M_w \) is obtained with:

\[
M_w = \frac{2}{3} \left( \log M_0 - 9.1 \right). \tag{4.10}
\]
Figure 4.S.2: Average S-wave spectrum fitted with the Brune model (red curve). This is a weighted average of all single-channel S-wave spectra (thin grey spectra, Equation (4.7)). The weight of each frequency bin of each channel is proportional to the excess signal-to-noise ratio (SNR) defined as \( w(f) = \text{SNR}(f) - \text{SNR}_t(f) \), where \( \text{SNR}_t(f) \) is the minimum SNR value that the frequency bin \( f \) must exceed in order to contribute to the average. Every frequency bin of the average spectrum also has a weight that is equal to the sum of the single-channel weights. Note that because we correct the single-channel spectra for geometric spreading and attenuation, the low-frequency plateau shown here gives directly the seismic moment \( M_0 \).

Once moment magnitude estimates \( M_{ref} \) are available for at least one event in a template family, we estimate a local magnitude \( M_{L,i} \) for all other events \( i \) based on log amplitude ratios:

\[
M_{L,i} = M_{ref} + \text{Median}_{s,c} \left\{ \log \frac{A_{i,s,c}}{A_{ref,s,c}} \right\},
\]

or more generally if there are several reference events:

\[
M_{L,i} = \text{Median}_k \left\{ M_{ref,k} + \text{Median}_{s,c} \left\{ \log \frac{A_{i,s,c}}{A_{ref,k,s,c}} \right\} \right\}.
\]

4.S.2 The Serdivan Earthquake Sequence

See Figure 4.S.3.
Figure 4.S.3: Seismicity of the 2012-07-07 $M_L 4.1$ Serdivan earthquake sequence. **Top left panel:** Cumulative number of earthquakes and local magnitudes $M_L$. The mainshock was followed by about 30 events in the next four hours, but only we recorded only 10 events in the next 26 days. **Top right panel:** Epicenters colored by time relative to the mainshock. Epicenters’ alignment and the largest events’ focal mechanisms suggest the existence of two conjugate faults and a network of secondary faults. **Bottom left panels:** Spatio-temporal evolution of the earthquake sequence. Successfully relocated hypocenters do not show any migration pattern consistent with fluid diffusion with diffusivity $D \approx 0.2 - 0.3 m^2/s$ typically observed for swarm seismicity (Shapiro et al., 2002) nor fast linear migration ($> 30$ km/day), suggesting that the earlier part of the sequence was controlled by static and dynamic stress changes. However, we can visually identify a southeastward migration of the seismicity in the later days of the sequence (top right panel).

### 4.S.3 Temporal Clustering


### 4.S.4 Stress State of the North Anatolian Fault Zone

**Focal Mechanisms**

Figure 4.S.4: Quantification of the strength of temporal clustering. **Top left panel:** Earthquake occurrence time series, alias the event count, *i.e.* number of events per unit time. **Top right panel:** Autocorrelation of the event count. A random time series would show a perfect Dirac function. **Bottom left panel:** Spectrum of the event count. The spectrum of a time clustered event count follows a power law $\propto f^{-\beta}$ and $\beta$ quantifies the strength of temporal clustering. **Bottom right panel:** Fractal analysis of the event count. We measure the number of time bins occupied by at least one event when dividing the time axis into smaller and smaller time bins. The slope of the curve gives the fractal dimension, which also quantifies the strength of temporal clustering. We fit the curve between 100 s (to avoid dealing with short time scales that can be corrupted by events counted twice, even though we normally take care of that) and the inverse of the average event rate (to avoid fitting the time scales where fractality trivially breaks down).

**Regional Stress Tensor**

See Figure 4.S.8.
Figure 4.S.5: Earthquake clustering along the North Anatolian Fault Zone. **Top panel:** Fractal dimension (as introduced in Figure 4.S.4). The eastern Marmara Sea and Lake Sapanca show the strongest clustering along the NAF. **Bottom panel:** Cumulative number of detections per template.
Figure 4.S.6: Objective function used in the focal mechanism estimate. **Top row:** Root Mean Square (RMS) waveform amplitude difference. **Middle column:** Negative waveform correlation (minimizing the negative correlation means maximizing the correlation). **Bottom row:** The objective function is made of the sum of these two terms. Note that the RMS and the correlation are already scaled so that they can be summed without having one term dominate the sum. The three columns show orthogonal 2D slices containing the best solution.
Figure 4.S.7: Faulting regime vs depth in each region. Each event is attributed the focal mechanism of the template earthquake with which it correlates best. If the corresponding template earthquake has no focal mechanism estimate, the event is not included in these histograms.
Figure 4.S.8: Regional stress state: All focal mechanism data, both from the northern and the southern strands, were inverted. **Left panel:** Pressure (P, red inverted triangles) and tension (T, blue dots) axes of all focal mechanisms. The study region taken as a whole exhibits a wide variety of P/T axes, which violates the assumption of uniform stress state. **Middle panel:** Directions of the principal stresses on an equal area stereographic projection. The principal stresses are ordered from most compressive ($\sigma_1$) to least compressive ($\sigma_3$). The numbers in the legend are the azimuth (angle from north) and the plunge (angle from horizontal) of the axes. The contours are the 90% and the 95% confidence regions. **Right panel:** Shape ratio, defined as $R = \frac{\sigma_1 - \sigma_2}{\sigma_1 - \sigma_3}$. It measures the relative magnitude of the principal stresses. In both the middle and right panels, the distribution of possible solutions was obtained by randomly sampling the set of possible focal mechanisms generated for each template earthquake (see main text).
Bibliography


J. Dieterich. Applications of rate-and state-dependent friction to models of fault slip and earthquake occurrence. 2007.


Chapter 5

Summary and Concluding Remarks

5.1 Summary

This thesis presents a complete workflow of observational seismology concerning earthquakes and seismicity: we started from the design of automatic methods for earthquake detection and location, then we demonstrated the efficiency and generalizability of these methods by applying them successfully to two data sets (the Southwestern Alps and the North Anatolian Fault), and finally we showed various types of analyses, focusing on studying ensembles of earthquakes rather than isolated events, that pushed further the borders of our understanding of the study regions.

Chapter 2 introduced the earthquake detection and location method (later improved in Chapter 4) that constituted the basis for all analyses of the thesis (Beaucé et al. (2019), and see Appendix A for details on the template matching algorithm Beaucé et al. (2018)). The method was applied on the dense linear array CIFALPS, in the Southwestern Alps, and revealed a continuous activity of $M < 3$ earthquakes (about 20 times more detections than previous catalogs). This chapter also introduced a property of earthquakes that the conclusions of every chapter focused on: the temporal clustering of earthquakes. On the basis that clustering emerges when earthquakes interact (see introduction in Chapter 1), we identified different fault behaviors in the Alps: cascade-like seismicity in densely fractured medium or rough
faults (dense population of asperities), and swarm-like seismicity driven by fluid related stressing mechanisms (e.g. transient changes in pore-fluid pressure due to fluid migration).

Chapter 3 investigated the source properties of the small-to-moderate size earthquakes that we previously detected in the Alps. This study took advantage of the organization of the earthquake catalog into families of similar events, which results from the template matching detection method, to estimate source parameters on ratios of earthquake spectra. We saw that this method is a powerful approach to removing propagation effects in earthquake signals without the necessity of a model. The analysis of source parameters revealed surprising departures from self-similarity in the two most active areas of the Southwestern Alps, where we previously identified strikingly different seismicity patterns (swarm-like vs. cascade-like). We explained the weak dependence of corner frequency upon magnitudes in different ways for these two areas. In agreement with the conclusions of Chapter 2, we proposed that fluids controlled the decrease of stress drop with decreasing magnitude in the swarm-like seismicity (the Ubaye valley), whereas asperity-like ruptures with closely located interacting asperities and possibly creep mediated stress transfers were invoked to explain the weak scaling between corner frequency and seismic moment, and the anomalously high stress drops in the cascade-like seismicity (the Dora Maira massif) as a consequence of the inadequacy of crack-like models.

Chapter 4 reused the method presented in Chapter 2 to analyze a data set from a different region and tectonic context: the North Anatolian Fault (NAF, Beaucé et al., 2021b). We processed data from the dense array DANA, located above the western section of the NAF that broke during the M7.6 1999 Izmit. For this study, we improved the earthquake location method by bringing in deep learning phase picking and double-difference relative relocation. To fully compare the state of the NAF at the time of the DANA experiment (in 2012-2013) and its state right before and after the Izmit earthquake, we estimated the focal mechanisms of earthquakes in the vicin-
ity of the fault and used them to infer the stress state (the stress inversion method is described in Appendix B, Beaucé et al., 2021a). Both the space-time distribution of earthquakes and the stress state on the fault indicated that the NAF was still being affected by postseismic phenomena, notably the existence of residual stresses and aseismic slip triggered by the Izmit earthquake. We completed this analysis with the characterization of temporal clustering, and observed more pronounced temporal clustering at the bottom of the seismogenic zone, namely where frictional properties transition from promoting unstable (seismogenic) sliding to continuous stable sliding. This result suggests that fault systems featuring strong seismogenic asperities embedded in a weak creeping fault promote temporal clustering. Because of the strong temporal clustering observed near Lake Sapanca and in the eastern Marmara Sea, we suggested that these areas are likely to feature a mix of stable and unstable sliding. Furthermore, focal mechanisms revealed that faults near Lake Sapanca act in normal extensional regime (similar analysis not available for the Marmara Sea). The same areas, Lake Sapanca and eastern Marmara Sea, had previously been proposed for hosting slow slip on normal faults, in order to explain the observed north-south extension at the surface in geodetic data. Thus, we brought new, independent observations that support the hypothesis of slow slip on normal faults along the NAF, which should be further investigated with complementary methods (see Perspectives, Section 5.2.1). Beyond the implications for the Alps or the NAF, the results of Chapters 3 and 4 point towards the close relation between temporal clustering and specific frictional properties of faults promoting the co-occurrence of unstable sliding (strong asperities) and stable sliding (weak creeping fault).
5.2 Perspectives

5.2.1 Completing Seismic Recordings with Geodetic Measurements

Most broadband seismometers are only sensitive to frequencies down to 10 or 1 mHz, and thus miss the slow motion of the earth happening over the course of days or months. This thesis has shown the importance of the inter-play between stable and unstable sliding (see Chapters 3 and 4), meaning that signals of tectonic origin must be happening over a broad range of timescales (*cf.* slow "aseismic" slip, Steinbrugge *et al.*, 1960; Ambraseys, 1970), including the long timescales that are not recorded by seismometers. Very low frequencies is the realm of geodesy, which relies on satellite measurements (GPS, GNSS, InSAR, LiDAR, etc) to study Earth’s shape and surface deformation.

Coupling seismic and geodetic data in order to increase the dynamic range of observations has recently become an active field of research (Rivet *et al.*, 2011; Frank and Brodsky, 2019), and has helped identify the ubiquity of slow (aseismic) slip on faults (Jolivet and Frank, 2020). Based on the observation of co-occurring seismic events and slow slip (Rogers and Dragert, 2003), seismic signals have been used to discover weak slow slip signals hidden in GPS/GNSS recordings (Frank, 2016; Rousset *et al.*, 2019). Following previous work (Frank, 2016), we started to investigate how the seismicity on the North Anatolian Fault (see Chapter 4, Beaucé *et al.*, 2021b) used jointly with GNSS data could reveal the slow motion hypothesized from the seismicity. The principle is to decompose a displacement time series obtained from GNSS recordings into a displacement time series occurring during the days of low seismicity, and a displacement time series occurring during the days of high seismicity, the two summing up to the original time series (see Figure 5.2.1).

The successful application of the method to the western section of the North
Figure 5.2.1: Decomposition of GNSS displacement time series based on local seismicity. The GNSS station is IZMT, Turkey, located north of the NAF. The original displacement time series (grey inverted triangles) is decomposed into the low seismicity displacement time series (black squares) and the high seismicity displacement time series (orange dots). Decomposition is performed on differentiated time series (speed), and decomposed time series are then integrated (back to displacement). **A:** East component. **B:** North component. **C:** Vertical upward component. The long term speed is first estimated on the initial time series with a linear regression (grey solid line), and the low seismicity and high seismicity speeds (black and orange solid lines, respectively) are then determined such that the modelled displacements sum to the initially modelled displacement. That is, the black and orange solid lines put together are forced to match the grey line when inverting for the high and low seismicity speeds. The vertical component is usually harder to interpret because subject to environmental noise such as rain falls.

Anatolian Fault has been limited by the very sparse coverage of the study area with permanent GNSS stations (there are only two stations near the fault, and none of them is very close to Lake Sapanca). This method relies on the assumptions that 1) slow slip co-occurs with the detected seismicity, and 2) the deformation produced by the slow slip is detectable by the GNSS stations. The critical parameter of the method is the threshold set on the number of events per unit time (e.g. per day for 1-day GNSS data) to define the periods of high and low seismicity, onto which the
displacement is decomposed (cf. Figure 5.2.1). This parameter can be tuned in a trial-and-error fashion when the strain loading and releasing directions are approximately known for a given fault system (for example, the direction perpendicular to the trench in a subduction zone). Although the NAF exhibits a simple strike-slip right-lateral behavior at large scales, we do not expect the same behavior to apply at the smaller scales investigated in Chapter 4. Instead, I am looking for a potential north-south displacement originating from Lake Sapanca and the eastern Marmara Sea. Therefore, in the case of the NAF tuning the seismicity threshold may lead to strongly misleading results. Overall, the method would benefit from abandoning this binary threshold as slow slip is obviously not a binary process: slow slip may be stronger or weaker on certain days, but is certainly not fully on or fully off.

My preliminary work has explored the limits of the method given the restricted geodetic data at our disposition. I will keep searching for the hidden signal of slow slip by 1) extending the earthquake catalog to longer time periods, 2) developing a threshold free methodology, and 3) using knowledge from the seismotectonic study of Chapter 4 to estimate the expected displacement at the GNSS stations and compare it with the data.

5.2.2 Keeping the Detection Method Up to Date

Extending the earthquake catalog presented in Chapter 4 will be the opportunity to keep my detection methods up to date with the most efficient techniques. In the introduction to this thesis (Chapter 1), I have mentioned the ever growing progress in earthquake detection algorithm driven by the development of machine learning based methods. We saw in Chapter 4 that the addition of the deep learning phase picker PhaseNet (Zhu and Beroza, 2019) greatly improved the location workflow by allowing an automated and accurate picking of P- and S-wave first arrivals. The beamforming method (also referred to as backprojection in this thesis) used to initialize the earthquake catalog before completion with template matching is, for now, based on a simple characteristic function of the seismic waveforms: their envelopes. Beamform-
ing is an efficient and physically justified method for combining these characteristic functions from multiple seismometers in a single trace. In the deep learning community jargon, beamforming is said to perform both detection and phase association. The latter consists of, given P and S picks at several seismometers, inferring how many events are detected and which picks correspond to the same event. Because deep learning detectors are mostly built for single-station data (e.g. Mousavi et al., 2020), detection and phase association are often done separately. Thus, in the next application of my detection method, I will use the output of a deep learning detector as the characteristic function used for beamforming. I expect this modification to drastically reduce the influence of non-earthquake spurious signals in seismograms (e.g. proximal noise source), thus reducing the number of false positive detections, and to improve the capability to detect weak signals.

The improvement of deep learning detectors will soon raise the question: is template matching still necessary? In this thesis, template matching is used as a tool to systematically complete an existing catalog. If this catalog is already complete enough so that the matched-filter search detects a negligible number of new events, then the computational cost will not be justified. Nevertheless, I believe that template matching is here to stay because of its conceptual simplicity. Improvements of the template matching method seem to reside in the choice of data representation, namely whether seismic data are simply expressed in the time domain, or in another domain where earthquakes are better separated from noise. Thus, template matching may benefit from advances due to machine learning, as finding adequate representations of the data is a central task in the learning problem (e.g. Holtzman et al., 2018; Seydoux et al., 2020).

5.2.3 On the Future of Earthquake Clustering Observations

I will close this thesis manuscript with my concluding remarks and thoughts about a major topic of my work: the temporal clustering of earthquakes. In most of this thesis, I analyzed temporal clustering as a stationary property. That is, I considered groups
of nearby earthquakes taken over the whole study period and characterized their tem-
poral clustering with a single scalar: the fractal dimension of their event count per
unit time. By doing so, I lost details about potential temporal changes of the state
of the system. For example, one could think that changes in normal stress (e.g. re-
sulting from nearby slow slip, increase in pore-fluid pressure, etc) would change the
stability of the system, which could result in more or less time clustering. Our study
periods in the Alps (1 year) and on the NAF (1.5 year) were probably short enough
to assume that groups of co-located earthquakes happened in a relatively uniform
fault state. However, developing a framework for investigating temporal variations in
temporal clustering would generalize better to longer study periods, and to systems
where time clustering varies over short timescales. Figure 2.S.9, in the supplementary
information of Chapter 2, shows an attempt to characterize temporal clustering over a
sliding window. As the focus of the study was not on analyzing this time dependence,
this analysis was not pushed further. In the future, I will use this preliminary work
as a starting point for studying how parameters, such as the duration of the sliding
window or the minimum number of earthquakes that a variable-size sliding window
should include, affect the quantification of temporal clustering.

The logical continuation of this work on temporal clustering is to increase the num-
ber of similar studies to improve our understanding of which fault systems produce
time clustered seismicity. Ultimately, we would like to draw reliable conclusions on
the state of the fault system given the observation of temporal clustering. This work
will of course require multiple research groups to systematically include the charac-
terization of temporal clustering in their analyses of seismicity (see interesting studies
in my extended research group: Frank et al., 2016; Reyes et al., 2021; Cabrera et al.,
2021). The accumulation of high quality observations may justify the introduction
of finer description of temporal clustering, for which I believe multifractal analysis
(Mandelbrot, 1974, 1989) would be an adequate extension of the fractal analysis used
in this thesis. Multifractal analysis replaces the characterization of fractal patterns
by a function, instead of a single scalar, and thus may provide a more detailed de-
scription of fractal seismicity.

Finally, I believe that seismologists should try to relate temporal clustering to potential precursory patterns to large earthquakes. We saw in Chapter 1 that the self-similarity of earthquakes and the Gutenberg-Richter law suggest that earthquakes are energy releasing events in an earth at the state of self-organized criticality, and that therefore earthquake prediction is impossible (Bak and Tang, 1989; Geller et al., 1997; Main, 1997). However, the notion that we cannot predict when and where the next large earthquake (e.g. large enough to saturate the seismogenic width in subduction zones: $M_w > 7.5$) will occur seems to only hold if one thinks globally of Earth. On regional scales, the notion of seismic cycle should prevail: after the 2011 $M_w 9.1$ Tohoku earthquake, one would not expect another major rupture in the same area before long because energy is no longer available for large earthquakes. Thus, strain energy will slowly build up until a large region of the crust stores excess energy in sufficient amount to allow a large rupture.

This idea was developed in Sykes et al. (1999), and is summarized in Figure 5.2.2. The sand pile model is used as a typical example of a system in a self-organized critical state: energy is regularly added to the system (falling sand grains) until the system reaches a state, said critical, where energy is irregularly released in avalanches whose sizes follow a power-law (i.e. equivalent to the Gutenberg-Richter law). The critical state is said self-organized when it does not depend upon the initial state of the system. Even though the sand pile as a whole never really departs from the critical state, it can locally strongly depart from the critical state after a large avalanche occurs (Figure 5.2.2b). Similarly to the example I gave about the Tohoku earthquake, the side of the sand pile that slipped in a large avalanche has to be rebuilt before it can produce a large avalanche again. This model also explains that small avalanches make the slope more unstable (Figure 5.2.2c) and therefore favor future avalanches on the same slope. Finally, it describes the occurrence of moderate avalanches when the slope is critical, but these are not large enough to significantly change the slope.
Figure 5.2.2: The sand pile model traditionally used to describe a self-organized critical system. **a:** Critical state of a sand pile. **b:** A large avalanche occurred on the right side of the pile, thus the system is locally no longer at the critical state on this side. The pile has to be rebuilt before developing large avalanches on this side. **c:** A small avalanche occurred on the right side, making the right side more unstable and future avalanches more likely. **d:** As the right side of the pile approaches the critical state, moderate-size avalanches occur. Figure from Sykes et al. (1999).

While the analogy between a sand pile and the crust may not be obvious, this model explains elegantly how 1) a self-organized critical Earth’s crust does not prevent local predictions, and 2) a large event is preceded by a slow build up towards the critical state. In the following, I explain what implications we can expect from the criticality of the crust. Critical phenomena are characterized by long range interactions due to the build up of short range interactions (cf. Ising model or percolation theory). As a consequence of these interactions, physical quantities describing the system become correlated over large scales. Ideal criticality is in fact defined by an infinite correlation length. I use the percolation problem as an illustration of this concept: a 2-D grid contains sites that are occupied with probability $p$, and are un-
occupied otherwise. Neighboring sites form clusters. As the occupancy probability $p$ grows, larger and larger clusters appear, and when $p$ reaches the critical value $p_c$, a cluster extending from one side of the grid to the other appears (see Figure 5.2.3A-C). The size of the largest cluster is of course limited by the size of the grid. In this model, two sites are correlated if they are part of the same cluster, and therefore the correlation length becomes very large at the critical state (is infinity if the grid is infinitely large). Another characteristic of the critical state is the scale invariance (see Figure 5.2.3D-F). The only meaningful length scale in such a system near the critical state is the correlation length, and the system appears to be scale invariant as long as the observation focuses on regions smaller than the correlation length.

Figure 5.2.3: Sites of a 2-D grid are occupied with probability $p$. **A:** $p = 0.420$. **B:** $p = p_c = 0.593$ (*i.e.* at the critical state, see Stauffer and Aharony, 2018). A cluster that links the two sides of the grid appear at the critical state (*i.e.* there is a path for percolation from one side to the other). **C:** $p = 0.850$. The size of the clusters is given by the color scale. **D-F:** At $p = p_c$, clusters exhibit scale invariance. When zooming in the grid, there is no length scale that gives a sense of scale (the size of a pixel becomes actually visible in **F**).
Thus, as the stress on a given region of the crust builds up, this region evolves towards the critical state. As it approaches this state, we expect interactions between faults to occur at longer and longer distances, resulting in some quantities to become correlated over long length/time scales, and we expect scale invariance in some sense. The temporal clustering of earthquakes exhibits both the scale invariance because it is fractal, and long correlation times in the event count per unit time (see Figure 1.3.3B in Chapter 1). Arguments such as these described in Sykes et al. (1999), and the fact that temporal clustering seems to be the manifestation of criticality, make the continuous quantification of temporal clustering a good candidate for revealing precursory signals to large earthquakes. After developing the methodology for time-dependent temporal clustering (discussed above), I plan to analyze the variations in the temporal clustering of the long-term (1-10 years) seismicity preceding large earthquakes. If the ideas described here are correct to some extent, one could expect an increase in the fractal dimension of the event count before the mainshock. The first attempts should focus on good quality earthquake catalogs with a fairly low magnitude of completeness during the years preceding the large earthquake of interest (e.g. the M7.1 2020 Ridgecrest earthquake).

The above speculation on the information contained in the temporal clustering of preseismic seismicity goes beyond the debate on whether foreshocks (that is, earthquakes preceding a large event) can help predict large ruptures or not. While some studies have observed the foreshock activity to accelerate prior to the large event (Bouchon et al., 2013), it has also been argued that foreshocks are regular earthquakes that accidentally trigger a large event (Ellsworth and Bulut, 2018). Foreshocks are thus either a passive marker of earthquake nucleation (preslip model), in which case they can predict the occurrence of a large event, or if they are part of the processes that eventually trigger a large earthquake (cascade model, see summary of the debate in Mignan, 2014; Gomberg, 2018). Regardless of the details of the earthquake nucleation process, temporal clustering may be a tool to assess whether the crust embedding a given fault is near the critical state, in which case a large rupture is possible, or not.
This type of precursory signal may not provide useful data for the short-term prediction (hours to weeks) involved in the discussion on the forecasting value of foreshocks, but may fall in the category of intermediate-term prediction (months to years). Other promising studies have been searching for similar timescale precursory signals, such as seismic source localization prior to large ruptures (Ben-Zion and Zaliapin, 2020; Shi et al., 2021), or changes in the bulk properties of the crust (FaultScan experiment: Brenguier et al., 2019; Sheng et al., 2021).
Bibliography


W. L. Ellsworth and F. Bulut. Nucleation of the 1999 İzmit earthquake by a triggered


W. B. Frank and E. E. Brodsky. Daily measurement of slow slip from low-frequency
earthquakes is consistent with ordinary earthquake scaling. *Science advances*, 5

W. B. Frank, N. M. Shapiro, A. L. Husker, V. Kostoglodov, A. A. Gusev, and
M. Campillo. The evolving interaction of low-frequency earthquakes during tran-


2018.

reveals cyclic changes in seismic source spectra in geysers geothermal field. *Science

R. Jolivet and W. Frank. The transient and intermittent nature of slow slip. *AGU


B. B. Mandelbrot. Intermittent turbulence in self-similar cascades: divergence of high
moments and dimension of the carrier. *Journal of fluid Mechanics*, 62(2):331–358,
1974.

B. B. Mandelbrot. Multifractal measures, especially for the geophysicist. In *Fractals

A. Mignan. The debate on the prognostic value of earthquake foreshocks: A meta-

quake transformer—an attentive deep-learning model for simultaneous earthquake

H. S. S. Reyes, D. Essing, E. Beaucé, and P. Poli. The Imbricated Foreshock and
Aftershock Activities of the Balsorano (Italy) Mw 4.4 Normal Fault Earthquake

D. Rivet, M. Campillo, N. M. Shapiro, V. Cruz-Atienza, M. Radiguet, N. Cotte, and
V. Kostoglodov. Seismic evidence of nonlinear crustal deformation during a large


Appendix A

Fast Matched-Filter (FMF): an Efficient Seismic Matched-Filter Search for both CPU and GPU Architectures

Abstract

Matched-filter searches are an important tool in modern seismology to detect seismic events. The algorithm functions by computing the correlation coefficient between a template earthquake and a sliding window of continuous seismic records. A detection is recorded when the correlation coefficient crosses an established threshold. We present an optimized program, called Fast Matched-Filter (FMF), to efficiently run a network-based matched-filter search with either central processing units (CPUs) or Nvidia graphics processing units (GPUs). Wrappers for both Python and Matlab (CPU only) are provided to easily run FMF on a wide range of computational resources, from multi-core laptops to specialized computing clusters with GPUs. Both implementations leverage the embarrassingly parallel structure of the continuous computation of correlation coefficients in the time domain to achieve rapid performance. The highly parallel architecture of GPUs lends itself perfectly to the matched-filter algorithm, and we achieve the fastest run times with our GPU implementation. FMF allows for seismic network-based matched-filtering between a large set of template waveforms and a large continuous dataset in a reasonable amount of time. Such fast

run times are an important step in expanding the scope of earthquake detection and fostering the reproducibility of such studies.

A.1 Introduction

Many studies in seismology are based on detecting earthquakes and gathering them into a catalog of their locations and timing. Earthquake catalogs provide a window into the solid Earth, including but not limited to, the Earth’s mechanical properties, (e.g. Dziewonski and Anderson, 1981; Van der Hilst et al., 1997), the amount of tectonic stress released at plate margins (e.g. Rogers and Dragert, 2003; Frank, 2016), and the statistical laws that govern earthquake occurrence (e.g. Gutenberg and Richter, 1944; Utsu, 1961; Frank et al., 2016). Systematically generated earthquake catalogs are thus an important resource in seismology. A key to constructing such catalogs is the development of algorithms that can detect earthquakes over a wide range of magnitudes within large volumes of data in a feasible amount of time.

Earthquake catalogs can be made with simple power-based detection algorithms, like the short-term average/long-term average (STA/LTA) detector (Allen, 1982), but such methods often have difficulty discriminating earthquakes from transient noise sources. Earthquake detection is further complicated when the amplitudes of target seismic events are on the same order of magnitude as the ambient seismic noise, measured as the signal-to-noise ratio (SNR). Because the SNR is higher when the magnitude of the event is larger and/or the source-receiver distance is shorter, detection capabilities can vary greatly based on both source position and mechanism, and the geometry of the recording seismic network (e.g. Kwiatek and Ben-Zion, 2016). Through the analysis of multiple seismic stations together, one can lower the SNR threshold for detection by leveraging the coherent information over the seismic network. A method that has proven itself capable of detecting earthquakes with low SNRs on seismic networks is template matched-filtering (e.g. Gibbons and Ringdal, 2006; Shelly et al., 2007; Frank et al., 2014).

Template matched-filtering is motivated by the fact that when two colocated
earthquakes occur with the same focal mechanism, their waveforms are very similar to each other and should have a high correlation coefficient. Matched-filtering algorithms compute the average correlation coefficient (CC) between the waveforms of a reference event, or template, and a sliding window of the continuous seismic records at multiple stations:

\[
CC(t) = \frac{1}{N_s N_c} \sum_{s,c} \frac{\sum_{n=1}^{N} T_{s,c}(t_n) S_{s,c}(t_n + \tau_{s,c})}{\sqrt{\sum_{n=1}^{N} T_{s,c}^2(t_n) \sum_{n=1}^{N} S_{s,c}^2(t_n + \tau_{s,c})}}
\] (A.1)

In Equation A.1, it is important to note that we assume the waveforms are centered around zero, which is the case when the data is highpass filtered to remove periods lower than the template’s duration. s, c and n are respectively the indices for the station, the component and the time sample, and \(N_s\), \(N_c\) and \(N\) are the respective number of stations, components and the template’s length in samples. \(T\), \(S\) are respectively the template and the continuous seismic data, and \(\tau_{s,c}\) is the relative arrival time on station \(s\) and component \(c\). The relative times \(\tau_s\) are also called the network moveout: they describe the delay with which an event is observed on every station relative to the first arrival. The sliding window of continuous data that produces a high CC is considered a newly detected seismic event that is collocated with the template earthquake.

An existing earthquake catalog can be expanded by using its constituent earthquakes as templates in a matched-filter search. Each template will detect similar events over a wide range of SNRs (Warren-Smith et al., 2017). This can, however, be very costly if the catalog contains a large number of events. Making this method computationally efficient is therefore essential to avoid any limitations related to long computation times, such as limiting the number of templates or limiting the duration of the continuous search. One example that demonstrates this is recent work on the postseismic sequence of the 2015 \(M_w 8.3\) Illapel earthquake in Central Chile. Huang et al. (2017) performed a matched-filter search for aftershocks during the month that followed the mainshock. However, a longer detection period of 10 months captured the log-time spatial expansion of aftershock seismicity (Frank et al., 2017), which is
indicative of an afterslip-driven postseismic phase (e.g. Perfettini and Avouac, 2004).

Reproducibility is also crucial in earthquake detection, where studies must often choose some empirical parameters. One common example is choosing a CC threshold to determine when a new event is detected. This threshold represents a trade-off between detecting false positives (detecting non-seismic events that are unwanted) and detecting false negatives (missing target seismic events). Depending on the overarching goal of the detection, one threshold might be favored over another. If the detection algorithm is fast enough, different parameters such as template length, frequency band, or a subset of the seismic network, can be evaluated to select the best ones for the job.

In this context we propose an optimized program, which we call Fast Matched-Filter (FMF), to execute a matched-filter search and compute Equation A.1. We provide two versions implemented for different numerical resources. One version is written in C and can run on compute nodes typically comprised of several central processing units (CPUs), but can also run on any ordinary multi-core laptop or desktop. We refer to this as the CPU implementation. We also provide an Nvidia CUDA C version that runs on computers with Nvidia graphics processing units (GPU). We refer to this version as the GPU implementation. We note that FMF’s GPU implementation does not function with AMD GPUs, as CUDA C is developed for Nvidia GPUs. Both versions are open-source and are hosted on Github: https://github.com/beridel/fast_matched_filter.

A.2 CPU Implementation

The CPU implementation only requires a C compiler (e.g. the GNU compiler collection gcc). The core of the program is written in C, and wrapper functions are available for Python and Matlab to allow for easy use from a higher level programming interface. The continuous computation of the correlation coefficient (Equation A.1) is an embarrassingly parallel algorithm: the computations at different times $t$ and for different channels or templates are completely independent. For that reason, the par-
allelization is only limited by the number of threads we can create, each independently working through Equation A.1 at the same time. We parallelize the program over the temporal loop, meaning that different sliding windows of our data are assigned to different threads.

In this massive computation, the input/output operations (I/O) to access the data are a non-negligible source of slowdown, whose impact depends on the computer’s or computation node’s architecture. Before the CPU can continue computational operations, it must wait during I/O operations for the disk access to finish. We develop the scheme described below and shown in Figure A.2.1 that focuses on an efficient parallelization and only performs a single I/O operation:

- in the Python and Matlab wrapper functions, a single I/O operation is achieved at the beginning of the algorithm,

- noting that the denominator in Equation A.1 involves redundant operations, the Python and Matlab wrapper functions pre-compute the sums of the squared template waveforms and a C function pre-computes the cumulative sum of the squared seismograms within each sliding window,

- the heavy computations are performed in C: for each template the temporal loop is parallelized into several threads using OpenMP (Dagum and Menon, 1998) (OpenMP automatically determines the number of threads that can be created given the available resources).

The arithmetic mean is used to average the correlation coefficients over the stations and components in Equation A.1. When dealing with real data, however, it might be useful to attribute different weights to different stations and components. For example, one could attribute the highest weights to the stations/components that record the highest SNR waveforms. We modify Equation A.1 to take into account a weight per station/component:

\[
CC(t) = \sum_{s,c} w_{s,c} \frac{\sum_{n=1}^{N} T_{s,c}(t_n) S_{s,c}(t_n + \tau_{s,c})}{\sqrt{\sum_{n=1}^{N} T_{s,c}^2(t_n) \sum_{n=1}^{N} S_{s,c}^2(t_n + \tau_{s,c})}}
\]  
(A.2)
where $w_{s,c}$ is the user-designated weight associated with station $s$ and component $c$. If the sum of the weights is normalized to one, then attributing equal weights to every station and component results in computing the arithmetic mean.

The Python and Matlab wrappers ensure the inputs are well formatted for the C functions, and reduce the arguments to:

- templates: array with dimensions\(^2 (N_t, N_s, N_c, N)$ storing the template waveforms, where $N_t$ is the number of templates and the other parameters are the same as in Equation A.1.

- moveouts: array with dimensions\(^2 (N_t, N_s, N_c)$ storing the moveouts in number of samples.

- weights: array with dimensions\(^2 (N_t, N_s, N_c)$ storing the weights attributed to the stations for each template.

- data: array with dimensions\(^2 (N_s, N_c, N_{data})$ storing the seismograms, where $N_{data}$ is the length of a seismogram in number of samples.

- step: number of samples between each computation of the CC.

The data array has to fit in the CPU’s memory. Therefore, to process a long period of continuous data, the user has to call consecutively the wrapper function with successive chunks of data. Typically, continuous seismic data is divided into daily traces and the wrapper function can be used on each of them.

### A.3 GPU Implementation

Computational problems arising from gaming and 3D graphics motivated the development of GPUs: in order to scale and rotate a scene, one has to perform millions of simple matrix operations on the polygons that make up the 3D visualization. GPUs are designed to address such problems with most of their transistors devoted

\(^2\)These dimensions are for C-contiguous arrays. The order is reversed in Matlab, as Matlab uses Fortran-contiguous dimensions.
Figure A.2.1: Workflow of the CPU implementation of our program Fast Matched-Filter (FMF). A single large I/O operation is achieved at the beginning by reading the templates’ waveforms and the continuous data. Beside, the sums of the squared templates and the cumulative sum of the squared data are computed before entering the loops to avoid redundant operations. After that, the iterative computation of the CC starts and is parallelized with OpenMP: different sliding windows (chunks) of the data are assigned to different threads. The two dashed boxes indicate which part is executed by the wrapper, and which one by the C code.

to computation rather than I/O operations, as is the case for CPUs. Thanks to this specialized architecture, GPUs can be much more powerful than CPUs for specific applications (Mu et al., 2017). The embarrassingly parallel nature of matched-filter searches and the fact that our algorithm only performs one I/O operation make this computation well-suited for the use of GPUs. The GPU version of FMF requires a computation node equipped with at least one Nvidia GPU. The code is based on CUDA C (Nickolls et al., 2008) and requires the user to install the CUDA toolkit, which includes the compiler nvcc.
One key aspect of the GPU architecture is the organization of the computation into blocks and threads. Each block is subdivided into threads that each run a single task; in our case, this task is computing the correlation coefficient between a template and one sliding window of continuous data. This structure allows for a much greater degree of parallelization than on CPUs. On a single GPU one typically runs several blocks at the same time, each associated with several hundreds of threads. This amounts to $\sim 1,000 - 10,000$ threads running at the same time, compared with the $\sim 2 - 100$ threads usually running on CPUs. To further optimize the use of GPUs to the matched-filtering algorithm, we leverage the many redundant memory accesses involved in the continuous computation of Equation A.2. During the whole matched-filter search, a data sample is not accessed by a single thread, but by $N$ different threads in total (each thread reads $N$ data samples to compute a single correlation coefficient, $N$ being the template length). These redundant memory accesses can be exploited through the use of shared memory. Each thread of the same block (cf. Figure A.3.1) loads a data or template sample from the GPU’s global memory (its equivalent of RAM) into the block’s shared memory; this sample is then no longer revisited in global memory by other threads. Because shared memory has a throughput $\sim 10 \times$ larger than the throughput of global memory, this optimization represents a significant speedup of the algorithm. Beside, this transfer from global memory to shared memory is such that neighboring threads read data/template samples that are contiguous in the global memory. This organization allows to access the memory in a coalesced manner: neighboring threads read samples at the cost of one memory transaction instead of one transaction each.

The algorithm’s workflow is described below and illustrated in Figure A.3.1:

- in the Python function (no GPU implementation is currently available for Matlab), a single I/O operation is performed at the beginning of the algorithm,

- the wrapper function computes the sums of the squared templates, and formats the inputs for the CUDA C code,
- The communication between the host (the CPUs) and the device (the GPUs) is achieved in CUDA C. The templates and the data (plus the other inputs necessary to compute Equation A.2) are transferred to the device. If multiple GPUs are available, an OpenMP parallelization distributes the templates to different GPUs.

- Many computation blocks, each of them involving many threads, are simultaneously (up to a given number of blocks that is hardware dependent) set up to compute the CCs for different templates/stations/components and at different times. This process is repeated until the entire data has been scanned (cf. Figure A.3.2).

- The CCs are transferred back from the device (GPU) to the host (CPU).

In this algorithm, we skipped the computation of the cumulative sum of the squared data because GPUs typically feature limited amounts of memory (e.g. \( \sim 6\) Gbytes or less) that cannot store the vector of cumulative sums in addition to the vector of continuous data. This does not change, however, the arguments that are passed to the GPU wrapper function, which are the same as the CPU implementation.

### A.4 Speed Comparison

We test FMF running a series of numerical experiments that generate synthetic data and templates, and compute the correlation coefficient between them (i.e. the continuous computation of Equation A.2). We ran our tests on a compute node composed of 2 x 12-core 2.50GHz CPUs (Intel Xeon Processor E5-2680 v3), 256GB RAM and two Nvidia K80 Tesla dual-GPUs, amounting to a total of twenty-four CPU cores and four GPUs.

The first test we present used the following inputs: 5 stations, 3 components, between 1 and 10 templates with a waveform duration of 8 s, one-day long seismograms at a sampling rate of 10 Hz. The correlations are computed with a temporal step
Figure A.3.1: Workflow of the GPU implementation of FMF. The continuous seismic data and the templates are first read from the disk by the CPUs and transferred to the GPUs. The GPUs take advantage of the collective behavior of the threads and their quick access to shared memory to efficiently parallelize the computation. One GPU sets up several blocks (∼10) whom each creates many threads (512) that computes the CC on all the components of a given template/station and at different times (CC_{s,c}(t_{ni})). The average correlation coefficients are computed through a weighted average, and are eventually transferred back to the CPUs. The two dashed boxes indicate which part is executed by the wrapper, and which one by the CUDA C code.

equal to one sample, or 0.1 s. A Matlab code that leverages the same optimizations as CPU implementation (through Matlab’s Parallel Processing Toolbox’s \textit{parfor} loop)
Figure A.3.2: Organization of the grid of thread blocks set up by the GPU to compute the correlation between one template and the data. The block size (the number of threads per block) is 512 and the grid size (the total number of blocks) is $N_B$. The grid size depends on the length of the data array and on the temporal step; $N_B$ is just enough so that all the correlation coefficients are computed. Note that this scheme holds for a single template, but the same scheme runs in parallel when several GPUs are available.

and the EQcorrscan program (Chamberlain, 2017), we used the version 0.2.7 and the back-end FFTW) are compared with the CPU and GPU versions of our program. Figure A.4.1 shows the large difference of performance between the Matlab code and the other programs, emphasizing that Matlab’s high-level language is not suited to high performance computing. To further discriminate the two versions of our program and EQcorrscan, we choose to present another test using more realistic input parameters. Matched-filtering is often performed with more than five stations to take advantage of the seismic network and to obtain more statistically robust correlation coefficients. Sampling rates are also typically higher than 10 Hz.

The second test used the following input parameters: 12 stations, 3 components, between 1 and 50 templates with duration 8 s, one-day long seismograms at a sampling rate of 50 Hz and the cross-correlation is computed with different temporal steps (1, 5 and 10 samples). We ran our CPU version and EQcorrscan on 24 cores and our GPU version on 1 and 4 GPUs; Figure A.4.2 shows all the results. As an example, running 20 templates in a matched-filter search with those parameters and a temporal step of 1 takes:
Figure A.4.1: Comparison of the run times of matched-filter searches with different codes. Matched-filtering is achieved between one-day long synthetic seismograms on 5 stations, 3 components and a set of 8-second long templates whose size varies from 1 to 10. The sampling rate is 10 Hz and the temporal step used in the CC computation is 1 sample. Note the log scale on the y-axis.

- with our CPU code: 55.5 s,
- with EQcorrscan: 15.8 s,
- with our GPU code running on 1 GPU: 11.1 s,
- with our GPU code running on 4 GPUs: 3.5 s.

The GPU code running on a single GPU is respectively 5× and 2× faster than our CPU code and EQcorrscan running on 24 cores. In other words, one GPU is here equivalent to 48 or 120 cores, depending on the program that is running. The speedup associated with the increase of GPUs’ number is almost linear (11.1/3.5=3.2), making the use of many GPUs very attractive.

EQcorrscan, which computes the correlation coefficients in the spectral domain, is ~3.5× faster than the CPU version of our program for this choice of parameters. However this speedup is associated with a very memory-consuming algorithm. Indeed,
the complex numbers required to perform the computation in the spectral domain are 2x larger than double precision float numbers (16 bytes per element) whereas we only use single precision float numbers in our algorithm. Even though our node has enough RAM to run EQcorrscan with 50 templates, we note that memory problems can arise due to insufficient resources on other machines. We also emphasize that using single precision float numbers is part of the optimization on GPUs because Nvidia GPUs are designed to work efficiently on 32-bit numbers. The memory gain and speedup are only at the cost of a small loss of precision (typically of $\sim 10^{-3}$ when large numbers appear in the data). Another advantage of performing matched-filtering in the temporal domain is that we are not constrained to a temporal step of one, whereas it is the only option in the spectral domain because of the intrinsic mathematical formulation of the problem.

Figure A.4.2: Comparison of the run times of several matched-filter searches using the CPU and the GPU implementations of our FMF and EQcorrscan. Matched-filtering is achieved between a set of one-day long synthetic seismograms on 12 stations, 3 components and a set of 8-second long templates whose size varies from 1 to 50. The sampling rate is 50 Hz and the temporal step takes the values 1, 5 and 10. Note the log scale on the y-axis.
Increasing the temporal step between correlation coefficients, or decreasing the number of sliding windows of the continuous dataset, is a great source of speedup, but decreases the time precision of detected events and the overall detection capability. Depending on the frequency content of the events being searched for and the sampling rate, then increasing the temporal step might be a reasonable choice. For the CPU code, the speedup is inversely proportional to the temporal step increase; in Figure A.4.2 the 10-step run takes $10 \times$ shorter than the 1-step run. However, the speedup is not as efficient for the GPU version, due to the fact that a temporal step of one makes an optimal use of the shared memory in our code.

To summarize the results of Figure A.4.2 in another way, if one were to run a matched-filter search with 50 8-second long templates on 12 stations, 3 components, a sampling rate of 50 Hz and one-year-long period of data with a temporal step of one sample, it would take:

- with our GPU code running on 4 GPUs:

\[ 7.95s \times 365 \text{days} = 2901s = 48\text{min}20s \]

- with our CPU code running on 24 cores:

\[ 136.48s \times 365 \text{days} = 49,815s = 13\text{h}50\text{min} \]

\section*{A.5 Application to Real Data}

We present here an application to real data, where we use one template representing near-repeating seismic events from a source in the Western Alps, France to scan one day of data. The data come from the temporary experiment CIFALPS (China-Italy-France Alps Seismic Survey (CIFALPS, 2012-2013; Zhao et al., 2015)) that involved
55 broadband velocimeters and also from 32 permanent stations (from French and Italian networks). Figure A.5.1 summarizes this matched-filter search with the template’s location, the correlation coefficients we computed and some example detections that were found with the template. The correlation coefficient computation is turned into an earthquake detector by establishing a threshold of $10 \times \text{MAD} \{\text{CC}(t)\}$ (MAD stands for Median Absolute Deviation). All time windows with a CC greater than the threshold correspond to detections of earthquakes with waveforms similar to those of the template. We only show a single channel on Figure A.5.1, but the template used in the matched-filter search is composed of waveforms on 12 stations and 3 components. The data was downsampled to 50 Hz and filtered in the band 1.5-25 Hz, thus corresponding to the one-template test case shown in Figure A.4.2.

Figure A.5.1: Matched-filter search of one template in the Western Alps, France over a single day. **Top left panel:** Map of the region showing the template’s location (star) and the neighboring stations (inverted triangles). **Bottom left panel:** Correlation coefficients (Equation A.2) over the whole day. The threshold of $10 \times \text{MAD} \{\text{CC}(t)\}$ is plotted with the dashed line. A total of 83 earthquakes was detected with constraining each event to be separated by at least 3 seconds. **Right panel:** Template’s channel CT23.HHZ (top record) and 10 examples of detections on this same channel.
A.6 Conclusion

We provide an open-source code called FMF to efficiently compute the normalized correlation coefficient between template waveforms and continuous seismic records to achieve earthquake detection. Our open-source code comes with two implementations: a version written in C working with Python or Matlab wrappers that is designed to run on any multi-core computer, and a second version written in CUDA C working with a Python wrapper that is designed to run on computers with an Nvidia GPU.

FMF is faster than parallel codes written in high-level languages like Matlab or Python. Such performance enables large-scale detection via template matched-filtering in a reasonable amount time. For example, running 50 8-second long templates with waveforms on 12 stations, 3 components on a one-year period of data with sampling rate 50 Hz takes 48min with the GPU version running on 4 GPUs and 14h with the CPU version running on 24 cores. The use of GPUs dramatically decreases the computation time, and allows to efficient performances with relatively low cost hardware: our test machine cost about US$15,000 in 2015.

Processing large volumes of data is an important challenge when the number of instruments and data are rapidly increasing. We show here how our program can be used to systematically enhance existing catalogs of seismic events. Finally, we hope that the distribution of this software will help make detection studies more easily reproducible by different groups within the scientific community.

Data and resources

Our program FMF is available at https://github.com/beridel/fast_matched_filter. The topographic data used to generate the map in Figure A.5.1 was extracted from the SRTM 90m database (http://www.cgiar-csi.org/data/srtm-90m-digital-elevation-data).
The data used in the example come from the CIFALPS temporary array (China-Italy-Alps Seismic Survey, doi=\url{http://dx.doi.org/10.15778/RESIF.YP2012}), and some permanents stations belonging to French networks (FR, RD) and Italian networks (GU, IV, MN, MT). EQcorrscan is a Python package for the detection and analysis of repeating and near-repeating seismicity, we used the version 0.2.7 to run our tests (\url{https://doi.org/10.5281/zenodo.893621}).

Acknowledgements

We thank Nikolaï Shapiro for the inspiration behind our adventure into the world of GPUs. W.B.F. was supported by NSF grant EAR-PF 1452375 and E.B. by the Theodore R. Madden Fellowship.
Bibliography


255


Appendix B

An Iterative Linear Method for Estimating the Stress Tensor from Earthquake Focal Mechanism Data: Method and Examples

Abstract

Earthquake focal mechanism data provide information about the stress state at the origin of these earthquakes. The inversion methods that are commonly used to infer the stress tensor from focal mechanisms have varying complexity but always rely on a number of assumptions. We present an iterative method built upon a classic linear stress tensor inversion that allows to relax the assumption on shear stress magnitudes while preserving the computational simplicity of the linear problem. Every iteration of our method computes the least-squares solution of the problem, which makes the method fast enough to estimate the inverted parameter errors with non-parametric resampling methods such as bootstrapping. Following previous studies, this method removes the fault plane ambiguity in focal mechanism data by selecting the set of nodal planes that best satisfies the Mohr-Coulomb failure criterion. We first illustrate the performance of the proposed method on synthetic and real data sets, and then discuss the relationship between the assumption of constant shear stress magnitudes and the presence of non-optimally oriented faults. We provide the Python package ILSI to implement the proposed method.

B.1 Introduction

The sense of motion on faults carries information on the stress state surrounding these faults. Field measurements of fault orientations and slip directions (slickensides) were first used to retrieve the stress tensor using a number of assumptions (e.g. Carey et al., 1974; Angelier, 1979; Angelier et al., 1982). The cornerstone assumption of these methods is that slickensides are oriented along the direction of maximum shear stress resolved on the faults (the so-called Wallace-Bott assumption, Wallace, 1951; Bott, 1959). Stress tensor inversion techniques were extended to the more widely available earthquake focal mechanism data, which describe faulting from seismic observations instead of direct field measurements. In general, the inverse problem is non-linear and solving it requires grid-search or other global optimization methods (Angelier et al., 1982; Gephart and Forsyth, 1984), but with additional assumptions the problem can be linearized (e.g. Michael, 1984). Both non-linear and linear inversion techniques suffer from the ambiguity in earthquake focal mechanism data that provide two possible fault planes per datum, without offering the possibility to identify the actual fault plane (the consequences of choosing the wrong plane are discussed at length in Michael, 1987). While some of the non-linear inversion methods choose the fault planes as the set of planes that minimize their objective function (such as the algorithm implemented in Focal Mechanism Stress Inversion, FMSI, Gephart, 1990), other criteria, physics-based and independent from the objective function, are preferred to avoid data over-fitting. Lund and Slunga (1999) introduced the use of the Mohr-Coulomb failure criterion to select fault planes. In this work, we present an inversion method built upon the linearized problem due to Michael (1984), turning it into an iterative inversion while preserving the computational efficiency of the least-squares solution of a linear problem. We thus relax the assumption made by Michael (1984) that shear stress magnitude is constant across faults or, equivalently, that all faults are optimally oriented within the stress field. We also use the Mohr-Coulomb failure criterion to select fault planes. The proposed method is implemented by our Python package ILSI (Iterative Linear Stress Inversion, see Data and Resources).
In Section B.2, we review the previous work on the topic on stress inversion and introduce our method. Section B.3 demonstrates the advantages of the proposed method with synthetic data sets, and Section B.4 shows its performances on real data sets. Finally, we discuss the implications of the assumption on shear stress magnitudes in Section B.5.

B.2 Methodology

B.2.1 Previous Work

Given a fault plane with orientation described by its unitary normal \( \hat{n} \), the traction on the plane is:

\[
T = \sigma \hat{n},
\]

where \( \sigma \) is the Cauchy stress tensor, and \( T \) is the traction. The normal \( T_n \) and tangential (shear) \( T_t \) components of the traction are:

\[
T_n = (\sigma \hat{n} \cdot \hat{n}) \hat{n} = \sigma_n \hat{n} \quad \text{(B.2)}
\]

\[
T_t = T - T_n = \sigma \hat{n} - (\sigma \hat{n} \cdot \hat{n}) \hat{n} = \tau \hat{t}.
\]

In Equation (B.2), \( \sigma_n \) and \( \tau \) are the magnitudes of the normal and shear tractions, respectively. The direction of shear traction is given by the unitary vector \( \hat{t} \). The stress tensor is often represented by its eigendecomposition: the principal stress directions \( \hat{\sigma}_i \) (eigenvectors) and the principal stresses \( \sigma_i \) (eigenvalues, \( i = 1, 2, 3 \)). A plane that is perpendicular to \( \hat{\sigma}_i \) does not experience any shear, and is under a purely compressional traction of magnitude \( \sigma_i \). By convention, eigenvalues are ordered such that \( \sigma_1 \) is the most compressional stress, \( \sigma_3 \) is the least compressional stress and \( \sigma_2 \) is the intermediate stress. Note that in the earth all stresses are compressional because of lithostatic pressure, and extensional stresses only exist in the sense of deviatoric stresses (the stress minus the lithostatic pressure).
Stress tensor inversion of earthquake focal mechanism data relies on two assumptions:

- the stress tensor is homogeneous in space,
- slip on faults occur in the direction of maximum resolved shear stress (Wallace-Bott assumption, Wallace, 1951; Bott, 1959).

An earthquake focal mechanism is a descriptor of the orientation and slip direction of a fault, based on the radiation pattern of seismic waves. Because of the symmetry of radiation patterns, there exists two fault planes with different slip directions that describe the same focal mechanism. Thus, each focal mechanism datum provides two fault normals \( \hat{n} \) and two slip directions \( \hat{s} \). We describe later a way of solving this ambiguity, but will assume for now that the fault plane and slip direction are known. Based on the Wallace-Bott assumption, one seeks the stress tensor that predicts shear directions \( \hat{t} \) that best match the slip directions \( \hat{s} \). This inverse problem does not provide any information on the absolute stress magnitudes, therefore it can only retrieve the reduced stress tensor, i.e. a normalized deviatoric stress tensor \( \sigma^* \):

\[
\text{Tr}(\sigma^*) = \sum_{k=1}^{3} \sigma^*_{kk} = 0; \quad \sum_{i,j}(\sigma^*_{ij})^2 = 1. \tag{B.3}
\]

For simplicity, hereinafter we keep using \( \sigma \) instead of \( \sigma^* \). One can only obtain four independent parameters from the inverse problem: the three principal stress directions \( \hat{\sigma}_i \), and the shape ratio \( R \),

\[
R = \frac{\sigma_1 - \sigma_2}{\sigma_1 - \sigma_3}. \tag{B.4}
\]

This scalar quantity is a measure of the relative magnitude of the principal stresses. In terms of deviatoric stress, if \( R > 0.5 \), \( \sigma_2 \) is extensional, and conversely if \( R < 0.5 \) then \( \sigma_2 \) is compressional. One can think of \( R \) as describing the position of \( \sigma_2 \) in between \( \sigma_1 \) and \( \sigma_3 \) on the x-axis of a Mohr circle (see Figure B.2.1).
For a population of faults described by their normals $\hat{n}_i$ and their slip directions $\hat{s}_i$, writing Equation (B.2) for each fault and identifying the shear direction $\hat{t}_i$ to the slip direction $\hat{s}_i$ yields the following system of equations for shear tractions:

$$\sigma \hat{n}_i - (\sigma \hat{n}_i \cdot \hat{n}_i) \hat{n}_i = \tau_i \hat{s}_i. \quad (B.5)$$

The inverse problem consists of finding $\sigma$ such that Equation (B.5) is satisfied on each fault $i$. Unfortunately, even though the left-hand side of Equation (B.5) is linear in $\sigma$, the right-hand side is not because of the shear magnitude $\tau_i$. Although one cannot determine the absolute shear magnitudes, the relative magnitudes between faults still matter. Several strategies have be taken to solve the inverse problem: Angelier et al. (1982) solve the non-linear problem iteratively, and the broadly used method due to Gephart and Forsyth (1984) adopts a grid-search over the four independent parameters to minimize their angular misfit (i.e. their objective function does not depend on shear magnitudes). These two methods consider errors not only in the slip directions but also in the orientation of the fault normals. The other widely used method is due to Michael (1984), which stands out by its simplicity, and which many other methods are built upon (for example, Hardebeck and Michael, 2006; Martínez-Garzón et al., 2014). The author made the assumption that shear stress is relatively constant from fault to fault, thus assigning the shear magnitude on the right-hand side of Equation (B.5) a scalar (equal to one). We discuss the implications of the assumption on shear stress magnitudes in Section B.5. The linear problem is:

$$Gm = d \quad (B.6)$$

In Equation (B.6), $d$ is the stack of all slip vectors and $G$ is the stack of all matrices relating the stress tensor elements to the direction of shear stress on the faults. Given that we can only retrieve the deviatoric stress tensor (see Equation (B.3)), the last
diagonal term that does not appear in Equation (B.6) is implicitly defined by \( \sigma_{33} = -\sigma_{11} - \sigma_{22} \). The linear inverse problem defined by Equation (B.6) is usually solved in the least-squares sense, for example with the Tarantola and Valette formula (Tarantola and Valette, 1982):

\[
m = m_{\text{prior}} + (G^T C_D^{-1} G + C_M^{-1})^{-1} G^T C_D^{-1} (d - G m_{\text{prior}}).
\]  
(B.7)

In Equation (B.7), \( m_{\text{prior}} \) is an approximate solution known a priori, \( C_D \) and \( C_M \) are the covariance matrices modeling the prior knowledge on the data and model parameter distributions, respectively. If no prior knowledge on the target solution is available, then \( C_M^{-1} = 0 \) and \( m_{\text{prior}} = 0 \). Note that this linear formulation only considers errors in slip directions and not in fault orientations.

### B.2.2 This Study

We relax the assumption on constant shear stress magnitudes in Michael (1984) by iteratively solving for both the stress tensor elements and the shear stress magnitudes. The algorithm is:

1) Initialize the solution:

\[
m^{(0)} = (G^T C_D^{-1} G + C_M^{-1})^{-1} G^T C_D^{-1} d^{(0)}
\]  
(B.8)

2) Compute shear stress magnitudes at iteration \( t \) and update the set of linear equations (i.e. update the estimate of the right-hand side of Equation (B.5)):

\[
\tau_i^{(t)} = |\sigma_i^{(t)} \hat{n}_i - (\sigma_i^{(t)} \hat{n}_i \cdot \hat{n}_i) \hat{n}_i|; \quad d_i^{(t)} = \tau_i^{(t)} d^{(0)}
\]  
(B.9)

3) Solve the updated linear inverse problem and update the stress tensor elements \( m^{(t)} \) using the previous estimate \( m^{(t-1)} \) as prior knowledge:

\[
m^{(t)} = m^{(t-1)} + (G^T C_D^{-1} G + C_M^{-1})^{-1} G^T C_D^{-1} (d^{(t)} - G m^{(t-1)}).
\]  
(B.10)
4) Repeat 2 and 3 until $|\tau^{(t)} - \tau^{(t-1)}|$ is lower than a user-provided threshold.

The superscript in parenthesis is the iteration index, $d^{(0)}$ are the unitary slip vectors and $C^{-1}_M$ is generally set to zero. $C^{-1}_D$ can be used to give more or less weight to the observations based on their quality, or is equal to identity if all focal mechanisms are equally accurate. In Section B.3, we show how this iterative procedure helps find the exact solution in synthetic examples. In this study, we use the outward footwall normals, and the slip direction of the hanging walls with respect to the footwalls, implicitly setting our stress tensor sign convention to negative compression (see Appendix B.A). The formula relating strike/dip/rake to normal and slip vectors can be found, for example, in Chapter 4.2 of Stein and Wysession (2009).

When dealing with earthquake focal mechanism data sets, one needs to determine the fault planes out of the pairs of nodal planes in order to get an accurate estimate of the stress tensor (Michael, 1987, illustrate how choosing the wrong planes impacts the solution). We use the Mohr-Coulomb failure criterion to assess which planes are more likely to be the fault planes for a given stress tensor (Lund and Slunga, 1999). We recall that this criterion states that a rupture occurs if the shear stress exceeds a critical value given by:

$$\tau_c = C + \mu\sigma_n,$$  \hspace{1cm} (B.11)

where $C$ and $\mu$ are the cohesion and the friction on the fault, respectively. We denote the effective normal stress on the fault by $\sigma_n$, meaning that we include any pore pressure in this term. The closer the shear stress $\tau$ is to the critical value, the more unstable the fault is. Therefore, a measure of fault instability is:

$$\Delta \tau = \tau - \tau_c = \tau - C - \mu\sigma_n = \tau - \mu(\sigma_n + C/\mu) = \tau - \mu\sigma_n^*. \hspace{1cm} (B.12)$$

In Equation (B.12), $\Delta \tau$ is the instability parameter as defined in Lund and Slunga (1999). Note that we included cohesion into the normal stress magnitude ($\sigma_n \rightarrow \sigma_n^*$). Since we do not have access to absolute values of stress, the cohesion is not a
relevant variable in this analysis. This also means that \( \Delta \tau \) as defined here only has a meaning in a relative sense, when comparing different planes. Therefore, following Vavryčuk et al. (2013), Vavryčuk (2014), we express normal stresses with respect to the maximum compression stress \( \sigma_1 \), and normalize the instability parameter by its value at the most unstable plane. The modified instability parameter \( I \) is defined as:

\[
I = \frac{\tau - \mu(\sigma_1 - \sigma)}{\tau_c - \mu(\sigma_1 - \sigma_c)}.
\] (B.13)

This formula was derived assuming negative compression (i.e. \( \sigma_1 < \sigma_2 < \sigma_3 \)) for consistency with the first part of the method. \( \tau_c \) and \( \sigma_c \) are the shear stress and normal stress magnitudes of the most unstable fault. The different terms of Equation (B.13) are defined graphically in Figure B.2.1.

Figure B.2.1: Definition of the instability parameter (Equation (B.13), following Vavryčuk et al., 2013) in the Mohr space with the negative compression convention. The red straight lines are the failure lines whose slopes are controlled by the friction \( \mu \). The most unstable fault has coordinates \((\sigma_c, \tau_c)\) in the Mohr space. The \( \sigma_i \)'s are the principal stresses ordered from most compressive to least compressive.

The shear stress magnitudes are always assumed positive in Equation (B.13), however when choosing the fault plane out of the two nodal planes of each focal mechanism, we need to take into account the direction of the shear stress with respect to the slip direction. Therefore, we multiply the instability parameter \( I \) by the sign
of the dot product between shear stress and slip:

$$\tilde{I} = I \times \text{sign}\left(\hat{s} \cdot \hat{t}\right).$$  \hspace{1cm} (B.14)

Given a stress tensor $\sigma$, the fault plane that is chosen is the one that maximizes $\tilde{I}$ out of the two nodal planes.

Finally, the inversion includes the following steps:

1) Initial guess of $\sigma$ by randomly selecting sets of nodal planes.

2) Choose the fault planes based on $\tilde{I}$.

3) Inner loop: Iteratively run the linear inversion (Equations (B.8)-(B.10)).

4) Repeat 2 and 3 until convergence or a user-defined maximum number of iterations.

Ten iterations are usually sufficient to reach convergence. Even though adding these two layers of iterations to the linear inversion from Michael (1984) makes our algorithm slower, it is still fast enough to be run many times on bootstrapped data sets to infer the parameter confidence intervals (Efron and Tibshirani, 1986). We use this non-parametric method to estimate uncertainties in the applications described in the next sections.

We note that this algorithm may fail to converge in certain situations in which we observe an oscillatory solution. Given the stress tensor $\tau^{(i)}$ at iteration $i$, the selected set of nodal planes $S^{(i)}$ depends on the outcome of the instability criterion $\tilde{I}(\tau^{(i)})$. Inverting this set of nodal planes produces a new stress tensor $\tau^{(i+1)}$ that, in turn, selects a new set of nodal planes $S^{(i+1)}$. The discrete nature of $S$ implies that any change in the selected nodal planes translates into a sharp change in the associated fault normals. Because shear stress is a smooth function of the fault normal (Equation (B.2)), it also implies a sharp change in the inverted stress tensor. De-
spite oscillations of the solution, one can still use this iterative procedure to explore different possible populations of faults, and select in the end the stress tensor that produces the lowest residuals. Ten iterations are again usually sufficient to explore the candidate solutions and select the best one.

### B.3 Synthetic Experiments

We first test our method on synthetic data sets. The first experiment involves a stress tensor with northwest-southeast maximum compression, southwest-northeast least compression and vertical intermediate stress, favoring right-lateral strike-slip faulting on vertical east-west faults, or left-lateral strike-slip on north-south faults. The shape ratio is $R = 0.5$. The 100 fault planes are randomly distributed around azimuth 110° and the dips range from 65° to 90° (see leftmost column in Figure B.3.1). Rake directions are in the direction of the shear tractions determined by the stress tensor in order to be fully consistent with the Wallace-Bott assumption. These strikes/dips/rakes are referred to as the "true fault planes" hereafter. Empirical parameter distributions are estimated by bootstrapping the original data set 1000 times. Confidence intervals are derived from these distributions.

Inverting the true fault planes (that is, without the need of inferring the fault plane) shows that the linear inversion does not retrieve the correct solution when using perfect data. The true solution is also not within the 95% confidence interval. In contrast, our iterative method finds the true stress directions and shape ratio. The inability of the linear method to find the true solution on perfect data is due to the erroneous assumption that all shear stress magnitudes are the same. In this synthetic experiment, there is actually a factor seven between the largest and the smallest shear magnitude values (see Figure B.3.2). Figure B.3.2 shows that the iterative method recovers the exact shear magnitudes while the linear method predicts incorrect values. The low shear stress magnitudes seen in Figure B.3.2 are indicative of non-optimally
Figure B.3.1: Synthetic experiment 1. The true stress tensor (leftmost column, middle row, large black symbols) promotes right-lateral strike-slip faulting on east-west oriented vertical faults. The shape ratio is 0.50. The fault orientations are randomly chosen from a range of parameters that is physically sensible given the stress state (see text), and the rakes are chosen such that slip is along the maximum shear stress direction. a, e, i: Data set with only the 100 true fault planes. b, f, j: Noise free data set with both the fault planes and their auxiliary planes. c, g, k: Data set with noisy fault planes with strikes/dips/rakes perturbed by random values in $[-3\degree; +3\degree]$, and their auxiliary planes. d, h, i: Data set with noisy fault planes with strikes/dips/rakes perturbed by random values in $[-10\degree; +10\degree]$, and their auxiliary planes. a, b, c, d: Fault planes (black lines) and auxiliary planes (grey lines). e, f, g, h: Lower hemisphere, equal area stereographic projections of the principal stress axes and their 95% confidence intervals (CI) estimated from 1000 bootstrap resamplings: solid lines $= \sigma_1$ CI, dashed lines $= \sigma_2$ CI, dot-dashed lines $= \sigma_3$ CI. The legend shows the inverted shape ratios $R$, and the mean angle $|\Delta\theta|$ between the predicted shear directions on the true fault planes and the true slip directions. Circles, squares and triangles are the most compressive ($\sigma_1$), intermediate ($\sigma_2$) and least compressive ($\sigma_3$) stresses, respectively. i, j, k, l: The distributions of shape ratios from the 1000 bootstrap resamplings. The vertical black line indices the true shape ratio. The proposed method is labeled "Iterative failure criterion" and is shaded for clarity.
oriented faults. We discuss in Section B.5 how relaxing the assumption of constant shear stress magnitudes helps deal with these non-optimally oriented faults.

Figure B.3.2: Predicted shear stress magnitudes. a, c: Synthetic experiment 1. b, d: Synthetic experiment 2. The linear inversion does not predict the relative shear stress magnitudes correctly, but our iterative procedure retrieves the true values.

To test the efficacy of the fault plane selection criterion, we augment the data set with the auxiliary planes to synthesize a focal mechanism data set, and we vary the level of noise in the strike/dip/rake values. Even in the noise-free setting, getting the true solution is not trivial because the true fault planes are unknown. We compare four methods: the linear method, the linear method with the failure criterion, the iterative method, and the iterative method with the failure criterion. The latter is the method we introduced in Section B.2.2, and is labeled as "Iterative failure criterion" in the figures. Figure B.3.1 summarizes the results. Both in the noise-free and the low noise case, our method is the only one to retrieve the true solution. All methods fail in the high noise experiment, but even then our method produces a shape ratio
distribution that indicates that the true shape ratio is a likely solution.

Figure B.3.3: Synthetic experiment 2. The true stress tensor (e) promotes right-lateral oblique strike-slip faulting with a normal faulting component on east-west oriented faults. The shape ratio is 0.70. Same legend as Figure B.3.1.

We do the same exercise with a second synthetic data set in which the dominant faulting style is oblique strike-slip with a normal component, and the shape ratio is $R = 0.7$. The 100 fault planes are randomly distributed around azimuth $110^\circ$ and the dips range from $20^\circ$ to $65^\circ$ (see leftmost column in Figure B.3.3). Similar conclusions as for the first experiment can be drawn. The linear method fails to retrieve the true solution when inverting data from the true fault planes, which again is explained by the misfit between the predicted shear stress magnitudes and the true ones (cf. right panels of Figure B.3.2). Our method is also the only one to find the true solution in the noise free and low noise cases. However, the confidence intervals and the shape ratio distribution show that there exists another group of solutions, significantly different from the true solution, that explain this data set well (see the
multiples lobes in the stereonets, and the bimodal distribution of the shape ratios in Figure B.3.3). In the high noise scenario, all methods produce solutions that are off the true solution, but our method’s solution is the closest and, more importantly, it is the only method that captures the true solution inside its 95% confidence interval.

B.4 Real Data Applications

B.4.1 Case Study 1: Central Crete

We now test our method on real data sets. The first application is on the Central Crete data set of field measurements from Angelier (1979), which was also used by Michael (1984) as a test for their linear method. In this data set the true fault planes are known since they were measured in the field (see left panel of Figure B.4.1). The inversion results from the linear and the iterative methods are shown in Figure B.4.1. The linear method finds a solution that is similar to the one shown in Michael (1984), and the iterative method produces a solution that is closer to the one in Angelier (1979). Note that in the original publications, the authors use another definition of the shape ratio: $\Phi = 1 - R$. The measure of misfit $|\Delta \theta|$ shown in Figure B.4.1 is the mean angle between predicted and observed shear traction (using the Wallace-Bott assumption). The values of $|\Delta \theta|$ obtained here are similar to those presented in Angelier (1979) and Michael (1984).

The large azimuthal uncertainty on $\sigma_2$ and $\sigma_3$ in both methods is a physical consequence of the stress state (see Figure B.4.1 middle panel). In this tectonic setting, the two horizontal stresses $\sigma_2$ and $\sigma_3$ have close values ($R \approx 1$, uniaxial deviatoric compression). Therefore, the horizontal component of the shear traction is always much smaller than the vertical component, and the predicted shear directions show little sensitivity to the directions of the horizontal principal stresses. Thus, the azimuths of $\sigma_2$ and $\sigma_3$ are only weakly constrained by this data set. We note that
the confidence intervals estimated in Michael (1984) based on gaussian statistics are smaller than those presented in Figure B.4.1, and failed to capture the physical lack of constraint over the directions $\sigma_2$ and $\sigma_3$ (also discussed in Michael, 1987). This emphasizes the importance of better estimates of confidence intervals, such as via non-parametric methods. With this data set, the stress tensors inverted with the linear and the iterative methods both give shear stress magnitudes that are narrowly distributed around 0.5 (see rightmost panel of Figure B.4.1), which explains why the linear method and the underlying assumption of constant shear stress magnitude do well.

Figure B.4.1: Inversion of the central Crete data set from Angelier (1979). a: Fault plane orientations. b: Principal stress directions of the inverted stress tensor with the linear method due to Michael (1984) (blue symbols) and the iterative method introduced in Section B.2.2. In the original publication, Angelier (1979) find similar stress directions and a shape ratio of $R = 0.93$. The 95% confidence intervals (CI) derived from 1000 bootstrapped data sets are shown: solid lines = $\sigma_1$ CI, dashed lines = $\sigma_2$ CI, dot-dashed lines = $\sigma_3$ CI. c: Empirical shape ratio distributions estimated from the bootstrap resampling. d: Distribution of shear stress magnitudes resolved on the fault planes.

B.4.2 Case Study 2: Western North Anatolian Fault

The second application is on a data set of focal mechanisms from the western North Anatolian Fault Zone, compiled in Poyraz et al. (2015). The authors subdivided the
data set into the 1999 $M_w$7.6 Izmit earthquake and its aftershocks (Izmit data set, 20 focal mechanisms, cf. bottom left panel of Figure B.4.2), and earthquakes recorded a decade later (DANA data set, 41 focal mechanisms, cf. top left panel of Figure B.4.2). They inverted both data sets with the non-linear method FMSI (Gephart, 1990) to estimate the stress tensor in the two time periods. The FMSI solutions from Poyraz et al. (2015) and the solutions obtained with the linear and the proposed iterative methods are presented in Figure B.4.2.

In both cases, the iterative linear and failure criterion method produced solutions that are consistent with the FMSI solutions. On the DANA data set, all three methods agree well on the directions of the principal stresses, however the linear method yields a shape ratio ($R = 0.33$) that is significantly smaller than the value reported in Poyraz et al. (2015) ($R = 0.45$) and the solution of the iterative method ($R = 0.54$). The shape ratio of the regional stress tensor in the western North Anatolian Fault Zone, around the Izmit rupture, is usually estimated to be $R \approx 0.5$ (e.g. Kiratzi, 2002; Pınar et al., 2010). The proposed method gives the lowest measure of angular misfit $|\Delta \theta|$, although this criterion does not allow a fair comparison with the FMSI solution since FMSI is not designed to minimize $|\Delta \theta|$, as it also considers errors in the orientation of the fault plane itself.

Uncertainties and inconsistencies between methods are larger for the Izmit data set, partly due to the low number of events (20 focal mechanisms). The low shape ratios, $R = 0.18$ with the linear method and $R = 0.38$ with the iterative method, indicate that $\sigma_1$ and $\sigma_2$ are close in magnitude. Therefore, similarly to the reason given in Section B.4.1, we observe poorly constrained maximum and intermediate compression axes. The low values of shape ratio inverted for this data set reflect the mixture of strike-slip and normal faulting (transtensional regime) that followed the Izmit earthquake (e.g. Bohnhoff et al., 2006; Pınar et al., 2010). Our method still produced a solution that is consistent with the FMSI solution, and the lowest $|\Delta \theta|$. Here, the FMSI solution displays a surprisingly large $|\Delta \theta|$, which we speculate could
Figure B.4.2: Inversion of the North Anatolian Fault data set from Poyraz et al. (2015). (a), (b), (c), (d): Focal mechanisms from the DANA data set (Dense Array for North Anatolia DANA, 2012). (e), (f), (g), (h): Focal mechanisms from the 1999-08-17 $M_w$ 7.6 Izmit earthquake and some of its aftershocks. Columns are the same as in Figure B.4.1. The measure of misfit $|\Delta \theta|$ and the shear stress magnitudes (right panel) were computed on the fault planes selected by the instability parameter (Equations (B.13) and (B.14)).

be due to discrepancies between the nodal planes selected by their inversion and our failure criterion used to select the planes on which $|\Delta \theta|$ was computed.
B.5 Discussion on Shear Stress Magnitudes

On the North Anatolian data sets, the inverted stress tensors predict shear stress magnitudes that span a large range of values (cf. rightmost column of Figure B.4.2), similarly to what we saw in the synthetic experiments (Figure B.3.2). This shows that data sets where fault orientations are not consistent with constant shear stress do occur in practice, even though it might seem to contradict basic physical arguments at first. Michael (1984) justified the assumption of constant shear stress with the Mohr-Coulomb failure criterion (Equation (B.13), with cohesion $C = 0$): given that normal stress is mostly controlled by the lithostatic stress, it has no reason to vary significantly with the fault orientation. Following this reasoning, the only variable left for explaining a wide range of shear stress magnitudes, as shown in Figures B.3.2 and B.4.2, is the coefficient of friction $\mu$ at the time of failure, which was shown to exhibit little variation in rock samples (0.6-0.85, Byerlee, 1978). However, the previous arguments disregard the importance of pore fluid pressure on controlling the effective normal stress at play in the Mohr-Coulomb failure criterion (Faulkner et al., 2006), and the fact that real faults might have coefficients of friction as low as 0.2 due to various reasons, such as the presence of weak minerals or rock fabric (Collettini et al., 2009).

The wide distribution of shear stress magnitudes (in particular, the low values) is indicative of the presence of non-optimally oriented faults in the inverted average stress tensor. Rupture on these non-optimally oriented faults can be explained by weakness due to low friction coefficient or high pore fluid pressure, or by heterogeneities in the stress field (e.g. interactions between faults). In fact, interactions between applied stress and resulting structures develop local heterogeneities in the stress field, which can lead to the appearance of complex structures (e.g. Riedel shears, Dresen, 1991). In the latter case, fault misorientation is an artifact of the assumption of uniform stress field. Either for real (weak faults) or artificial (invalidity of the uniform stress assumption) reasons, these non-optimally oriented faults are found
in data sets in general. By iteratively inverting for the shear stress magnitudes (see Equations (B.8)-(B.10)), our method learns to identify these non-optimally oriented faults and adjusts the right-hand side of Equation (B.5) accordingly \( d_i^{(t)} = \tau_i^{(t)} d^{(0)} \), cf. Equation (B.9)) in order to achieve low residuals \((d - Gm)^2\) on these faults. Under the assumption of constant shear stress magnitude, the linear method is forced to find a solution that produces neither a good solution on the optimally oriented faults nor on the non-optimally oriented faults (see left column of Figure B.3.1) as a consequence of the least-squares criterion. The identification of non-optimally oriented faults combined with bootstrap resampling is likely to produce more realistic estimates of uncertainties due to violations of the uniform stress assumption than the linear method.

### B.6 Summary and Concluding Remarks

In Section B.2.1, we introduced the stress tensor inversion problem and presented the underlying assumptions and drawbacks. In Section B.2.2, we introduced an iterative inversion method built upon the linear inversion due to Michael (1984), and described how to combine it with a Mohr-Coulomb failure criterion (e.g. Lund and Slunga, 1999; Vavryčuk et al., 2013) to deal with focal mechanism data sets where the true fault planes are generally unknown. We used synthetic examples (Section B.3) to demonstrate the efficacy of our method, its advantages with respect to the linear inversion, and its potential to estimate confidence intervals with the bootstrap resampling method (Efron and Tibshirani, 1986). We then validated the proposed method on real data sets (Section B.4).

The first data set was constituted of field measurements from Central Crete and was inverted in Angelier (1979) and Michael (1984), with which we showed that the proposed method gave consistent results, and more accurate confidence intervals. The second data set was made of focal mechanisms from earthquakes located along the western section of the North Anatolian Fault Zone. We compared our method
against the results presented in Poyraz et al. (2015) that were obtained with the FMSI software that is considered the state-of-the-art stress inversion method (Gephart and Forsyth, 1984; Gephart, 1990). We showed that the solution produced by our method was in good agreement with the FMSI solution, whereas the linear method gave stress tensors with significantly different shape ratios.

Finally, in Section B.5, we discussed the implications of the assumption of constant shear stress magnitudes. We explained that relaxing the assumption on shear stress magnitudes helps our method deal with non-optimally oriented faults that are either symptomatic of stress heterogeneity (i.e., violations of the uniform stress assumption) or of fault weakness. Therefore, we believe that the inversion method proposed here produces accurate solutions, is easy to implement, and is fast enough to allow accurate estimates of uncertainties with non-parametric methods, such the bootstrap resampling method.

The stress inversion method introduced in this article can be implemented using our Python package ILSI (see Data and Resources).

Data and Resources

The first case study uses the Central Crete data set published in Angelier (1979) (their Table 1), and the second case study uses the North Anatolian data set published in Poyraz et al. (2015) (their Table 2).

The Python package ILSI, available at https://github.com/ebeauce/ILSI, implements the method proposed here and provides tutorial scripts to reproduce our figures (version 1.0.0, last accessed July 2021).
Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union’s Horizon H2020 research and innovation program (grant agreement No 742335). E.B. was also supported by funds associated with Robert D. van der Hilst’s Schlumberger chair.

B.A Stress Tensor Sign Convention

Let us consider a stress tensor with northwest/southeast maximum compression and southwest/northeast minimum compression, and intermediate compression in the vertical direction (see Figure B.A.1). The direction of the principal stresses $\hat{\sigma}_i$ ($i=1$ maximum compression, $i=3$ minimum compression) are, in the $(x, y, z)=$(north, west, upward) coordinate system:

$$
\hat{\sigma}_1 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \quad \hat{\sigma}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \hat{\sigma}_3 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}.
$$

(B.15)

The eigendecomposition of the stress tensor $\sigma$ is:

$$
\sigma = V \Sigma V^T,
$$

(B.16)

where $V$ is the matrix of column eigenvectors, and $\Sigma$ is the diagonal matrix of eigenvalues ($\sigma_1$, $\sigma_2$, $\sigma_3$):

$$
V = \begin{pmatrix} -\frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{pmatrix}; \quad \Sigma = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix}.
$$

(B.17)
We parameterize the stress tensor by:

\[ \sigma_1 = \pm 1, \quad \sigma_2 = \pm(1 - 2R), \quad \sigma_3 = \mp 1, \quad R = \frac{\sigma_1 - \sigma_2}{\sigma_1 - \sigma_3}. \]  \hspace{1cm} (B.18)

In Equation (B.18), the upper and lower signs hold for the compression positive convention and the tension positive convention, respectively. The parameter \( R \) is called the shape ratio, it characterizes the relative magnitudes of the principal stresses. The compression positive (cp) and tension positive (tp) stress tensors are:

\[ \sigma^{\text{cp/tp}} = V\Sigma^{\text{cp/tp}}V^T. \]  \hspace{1cm} (B.19)

![Diagram of fault plane and stress directions](image)

Figure B.A.1: East-west right-lateral strike-slip fault under northwest-southeast maximum compression and southwest-northeast minimum compression. The intermediate stress is vertical.

We consider the outward normal of the northern wall or, equivalently, the inward normal of the southern wall:

\[ \hat{n} = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}. \]  \hspace{1cm} (B.20)

The unitary slip of the southern wall with respect to the northern wall is (right-lateral,
the southern wall moves to the west): 

\[
\hat{d} = \begin{pmatrix} 0 \\ +1 \\ 0 \end{pmatrix}.
\]  

(B.21)

**B.A.1 Compression Positive**

\[
\sigma^{cp} = V\Sigma^{cp}V^T = \begin{pmatrix} 0 & +1 & 0 \\ +1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}
\]  

(B.22)

Compression is positive when defining the elements of the stress tensor for inward pointing normals. Therefore, the normal defined by Equation (B.20) is the inward normal of the southern wall, and \(\sigma^{cp}\hat{n}\) gives the traction on the southern wall:

\[
t(\hat{n}) = \sigma^{cp}\hat{n} = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}.
\]  

(B.23)

The traction on the southern wall is \(-1\hat{y}\), i.e. it points in the east direction, which is opposite to the direction of motion of the southern wall with respect to the northern wall.

**B.A.2 Tension Positive**

\[
\sigma^{tp} = V\Sigma^{tp}V^T = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}
\]  

(B.24)

Tension is positive when defining the elements of the stress tensor for outward pointing normals (common definition of the Cauchy stress tensor in physics). Therefore, the normal defined by Equation (B.20) is the outward normal of the northern
wall, and $\sigma^\text{tp}\hat{n}$ gives the traction on the northern wall:

$$t(\hat{n}) = \sigma^\text{tp}\hat{n} = \begin{pmatrix} 0 \\ +1 \\ 0 \end{pmatrix}. \quad (B.25)$$

The traction on the southern wall is $+1\hat{y}$, i.e. it points in the west direction, which is the direction of motion of the southern wall with respect to the northern wall.

### B.A.3 Implications for the Stress Tensor Inversion

This simple example shows that, given the coordinate system we chose ($x, y, z = \text{north}, \text{west}, \text{upward}$), outward pointing normals of the footwalls and slip directions of the hanging walls with respect to footwalls are consistent with the tension positive convention. The inward pointing normals of the hanging walls and slip directions of the footwall with respect to hanging walls are consistent with the compression positive convention.
Bibliography


