# Microseismic Constraints on the State of the North Anatolian Fault Thirteen Years after the 1999 M7.4 Izmit Earthquake

Eric Beaucé[1,1], Robert D van der Hilst[2,2], and Michel Campillo[3,3]

[1]Columbia University
[2]Massachusetts Institute of Technology
[3]Université Joseph Fourier, Grenoble

November 30, 2022

**Abstract**

The 17 August 1999 $M_{w}$7.4 Izmit earthquake ruptured the western section of the North Anatolian Fault Zone (NAFZ) and strongly altered the fault zone properties and stress field. Consequences of the co- and post-seismic stress changes were seen in the spatio-temporal evolution of the seismicity and in the surface slip rates. Thirteen years after the Izmit earthquake, in 2012, the dense seismic array DANA was deployed for 1.5 years. We built a new catalog of microseismicity (M < 2) by applying our automated detection and location method to the DANA data set. Our method combines a systematic backprojection of the seismic wavefield and template matching. We analyzed the statistical properties of the catalog by computing the Gutenberg-Richter b-value and by quantifying the amount of temporal clustering in groups of nearby earthquakes. We found that the microseismicity mainly occurs off the main fault and that the most active regions are the Lake Sapanca step-over and near the Akyazi fault. Based on previous studies, we interpreted the b-values and temporal clustering \textit{i}) as indicating that the Akyazi seismicity is occurring in high background stresses and is driven by the Izmit earthquake residual stresses, and \textit{ii}) as suggesting evidence that intricate seismic and aseismic slip was taking place on heterogeneous faults at the eastern Lake Sapanca, near the brittle-ductile transition. Geodesy shows enhanced north-south extension around Lake Sapanca following the Izmit earthquake, therefore, the seismicity supports the possibility of slow slip at depth in the step-over.

# Microseismic Constraints on the Mechanical State of the North Anatolian Fault Zone Thirteen Years after the 1999 M7.4 Izmit Earthquake

**Eric Beaucé[1,3], Robert D. van der Hilst[1], Michel Campillo[2,1]**

[1]Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, USA
[2]Institut des Sciences de la Terre, Université Grenoble Alpes, France
[3]Lamont-Doherty Earth Observatory, Columbia University, NY, USA

**Key Points:**

- We built an extensive earthquake catalog using our new automated method along the western NAFZ.
- We observe variable statistical properties, b-value, and temporal clustering along the fault.
- The properties of the Sapanca seismicity support the possibility of slow slip in the step-over.

Corresponding author: Eric Beaucé, ebeauce@ldeo.columbia.edu

**Abstract**

The 17 August 1999 $M_w$7.4 Izmit earthquake ruptured the western section of the North Anatolian Fault Zone (NAFZ) and strongly altered the fault zone properties and stress field. Consequences of the co- and post-seismic stress changes were seen in the spatio-temporal evolution of the seismicity and in the surface slip rates. Thirteen years after the Izmit earthquake, in 2012, the dense seismic array DANA was deployed for 1.5 years. We built a new catalog of microseismicity (M $<$ 2) by applying our automated detection and location method to the DANA data set. Our method combines a systematic backprojection of the seismic wavefield and template matching. We analyzed the statistical properties of the catalog by computing the Gutenberg-Richter b-value and by quantifying the amount of temporal clustering in groups of nearby earthquakes. We found that the microseismicity mainly occurs off the main fault and that the most active regions are the Lake Sapanca step-over and near the Akyazi fault. Based on previous studies, we interpreted the b-values and temporal clustering *i*) as indicating that the Akyazi seismicity is occurring in high background stresses and is driven by the Izmit earthquake residual stresses, and *ii*) as suggesting evidence that intricate seismic and aseismic slip was taking place on heterogeneous faults at the eastern Lake Sapanca, near the brittle-ductile transition. Geodesy shows enhanced north-south extension around Lake Sapanca following the Izmit earthquake, therefore, the seismicity supports the possibility of slow slip at depth in the step-over.

## Plain Language Summary

On 17 August 1999, a large M7.4 earthquake struck near the city of Izmit, in western Turkey, and caused important human and material losses. The earthquake resulted from the large and sudden displacement of crustal blocks along the North Anatolian Fault Zone (NAFZ). Transient changes in the crustal and fault properties are commonly observed following such large events. In this study, we analyzed the statistical properties of microearthquakes, that is, of small earthquakes (M $<$ 2) typically too small to affect the surrounding population, to gain knowledge about the state of the NAF more than a decade after the Izmit earthquake. First, we addressed the challenge of locating microearthquakes, in space and time, by applying our automatic earthquake detection and location algorithm. Then, the statistical analysis allowed us to characterize physical properties of the NAF and, thus, to highlight the peculiar properties of faults near Lake Sapanca. We interpreted that these faults are heterogeneous and slip both seismically and aseismically. Our study calls for taking a closer look at the extension across Lake Sapanca with different, complementary geophysical methods.

## 1 Introduction

The North Anatolian Fault Zone (NAFZ) 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 (Figure 1A). The fault slips, overall, in a right-lateral manner to accommodate the westward motion of Anatolia with respect to Eurasia due to the combination of 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 & Angelier, 1979; McClusky et al., 2000; Reilinger et al., 2006). Near the Gulf of Izmit, in western Turkey, the NAFZ splits into a northern strand and a southern strand. These two strands bound the Almacik mountains in the east and the Armutlu block in the west, and separate the Istanbul Zone in the north from the Sakarya Terrane in the south, which are the remains of the passive margin of the Intra-Pontide Ocean (see Figure 1B, *e.g.* Akbayram et al., 2013). Most of the deformation is accommodated on the northern strand (*e.g.* Meade et al., 2002; Reilinger et al., 2006).

The 17 August 1999 M7.4 Izmit earthquake and the 12 November 1999 Düzce M7.2 earthquake are the most recent (as of the time of writing) events of a series of westward migrating M>7 earthquakes that ruptured almost entirely the North Anatolian Fault Zone (*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, or five, 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 1B). To the west, the rupture propagated along the Gölcük segment and stopped on the Yalova segment (Langridge et al., 2002), increasing the probability of major failure further west beneath the Marmara Sea (Parsons et al., 2000). The Düzce earthquake nucleated near the eastern termination of the Izmit earthquake, likely due to increased Coulomb stress (Parsons et al., 2000; Utkucu et al., 2003). The co- and post-seismic stress changes and the transient changes of the fault's mechanical properties caused by the Izmit earthquake affected the local seismicity patterns and the focal mechanisms of microearthquakes (*e.g.* Bohnhoff et al., 2006; Pınar et al., 2010; Ickrath et al., 2015). GPS and interferometric synthetic aperture radar (InSAR) observations suggest that fast and rapidly decaying deep afterslip occurred in the middle-to-lower crust in the months following the Izmit-Düzce earthquake sequence (*e.g.* Reilinger et al., 2000; Bürgmann et al., 2002), then relayed by slower post-seismic slip at depth (Ergintav et al., 2009; Hearn et al., 2009). Patterns of surface displacement also suggest the existence of shallow creep along the Izmit-Sapanca and the Sapanca-Akyazi segments (*e.g.* Çakir et al., 2012; Hussain et al., 2016). Transient creep episodes have been identified more than a decade after the Izmit earthquake (Aslan et al., 2019).

Despite the overall good understanding of the east-west motion along the western NAFZ, smaller scale, north-south extension at some locations remains enigmatic. Co- and post-seismic slip on vertical fault segments seems unable to reproduce the patterns of north-south extension observed in geodetic data (*e.g.* Ergintav et al., 2009; Hearn et al., 2009). Even though refining the geometry of the main fault segments of the NAFZ helps explain the observations (*e.g.* slightly north dipping faults, Cakir et al., 2003), models of the post- and inter-seismic deformation along the NAFZ would benefit from taking into account secondary structures, such as the faults in step-overs. Microseismicity (M $<$ 2) provides information at small length scales at seismogenic depths and thus is complimentary to geodetic data in building a better understanding of slip along the NAFZ (aseismic vs seismic, distributed vs localized), that is, of its mechanical state.

The abundant number of microearthquakes makes them well-suited for statistical analyses. Of interest to this study are the b-value of the Gutenberg-Richter law (Gutenberg & Richter, 1941) that describes the frequency-magnitude distribution of a population of earthquakes, and the fractal dimension $D$ of the earthquake occurrence time series (Smalley Jr et al., 1987; Beaucé et al., 2019) that quantifies the strength of temporal clustering. The b-value acts as a stressmeter (Amelung & King, 1997; C. H. Scholz, 2015), and the fractal dimension $D$ is related to the density of fractures and seismic-aseismic slip partitioning (C. Scholz, 1968; Dublanchet et al., 2013).

The dense seismic array DANA (Dense Array for North Anatolia DANA , 2012, see Figure 1C) was deployed around the rupture trace of the 1999-08-17 Izmit earthquake, it operated about thirteen years later 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). Here, we study microearthquakes in order to improve our understanding of the mechanical state of the North Anatolian Fault Zone more than a decade after the 1999 M7.4 Izmit earthquake. First, we briefly describe our automated earthquake detection and location method (Section 2), and then present the earthquake catalog (Section 3.1) and a statistical analysis of collective properties of earthquakes (b-value, Section 3.2, and temporal clustering, Section 3.3). These observations allow a characterization of the physical environment in which seismicity takes

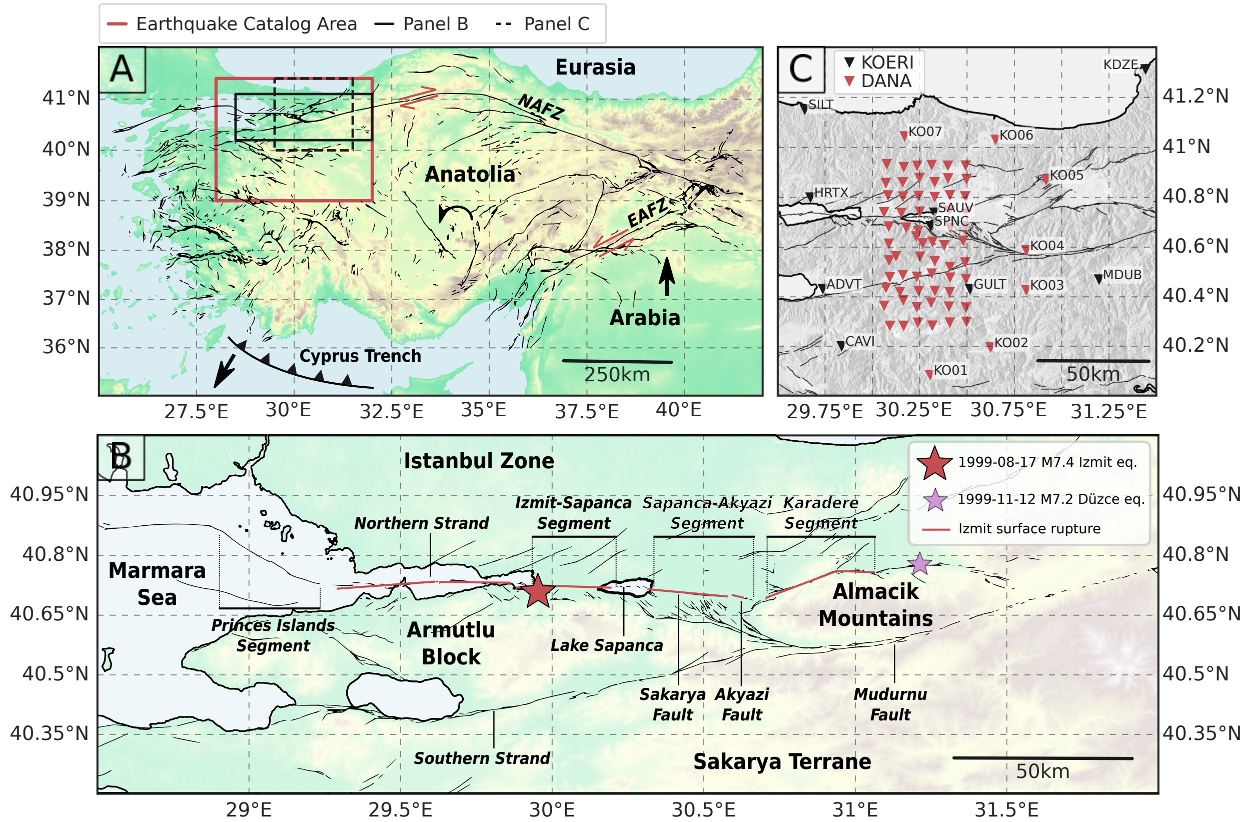


**Figure 1.** **A:** Large scale view of the North Anatolian Fault Zone. Abbreviations: NAFZ - North Anatolian Fault Zone, EAFZ - East Anatolian Fault Zone. The red arrows indicate the direction of coseismic motion. Our study region is located at the western end of the NAFZ. **B:** Magnified view of 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 NAFZ: 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, ...).

place. We interpret and discuss our results to question the role of secondary structures in the dynamics of NAFZ (Section 4).

# 2 Methodology

## 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 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 all stations but SAUV, which samples at 100 Hz. We bandpass filtered the data between 2 Hz and 12 Hz to eliminate low frequency noise and to allow us to downsample the time series to 25 Hz in order to make the computation less intensive.

## 2.2 Earthquake Detection, Location, and Magnitude Estimation

We analyzed the 2012-05-04/2013-09-20 time period with a fully automatized earthquake detection and location method. The core of the workflow, summarized in Figure 2, consists of three stages:

1. Backprojection (Section 2.2.1): The energy of the seismic wavefield is continuously backprojected onto a 3D grid of potential sources to detect coherent (earthquake) sources.
2. Relocation (Section 2.2.2): The P- and S-wave first arrivals of the previously detected events are identified with the automatic phase picker PhaseNet (Zhu & Beroza, 2019), and the picks are used in the NonLinLoc earthquake location software (Lomax et al., 2000, 2009).
3. Template matching (Section 2.2.3): The successfully relocated earthquakes are used as template earthquakes in a matched-filter search to detect other, smaller earthquakes in the same region using the Fast Matched Filter software (Beaucé et al., 2018).

The detection method is discussed in detail in Beaucé et al. (2019), but the relocation is now fully automated and includes PhaseNet and NonLinLoc. In an extra step, we further characterized the detected earthquakes by relocating them with the double-difference method (Section 2.2.4) and estimating their magnitude (Section 2.2.5).

### *2.2.1 Backprojection and Location*

We continuously backprojected the energy of the seismic wavefield recorded at the array of seismic stations onto a 3D grid of potential sources beneath the study region, searching for the space-time locations of coherent sources. Backprojection is now a widely used earthquake detection and source imaging method (*e.g.* Ishii et al., 2005; Walker et al., 2005; Honda & Aoi, 2009; W. Frank & Shapiro, 2014). We computed the composite network response (CNR):

$$\mathrm{CNR}(t) = \max_{k} \{\mathrm{NR}_k(t)\}\,; \qquad \mathrm{NR}_k(t) = \sum_{s,c} \mathrm{env}\left(u_{s,c}(t + \tau_{s,c}^{(k)})\right). \tag{1}$$

In this equation, $t$ is the detection time and $\mathrm{NR}_k(t)$ is the network response for source location indexed by $k$ at time $t$. $\mathrm{NR}_k(t)$ is the sum 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 from Karabulut et al. (2011) (see Table S1). We note that 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 NAFZ (*e.g.* Karahan et al., 2001; Kahraman et al., 2015; Papaleo et al., 2018). This velocity model produced a visually satisfying agreement between earthquake epicenters and fault surface traces, and allowed consistency with a

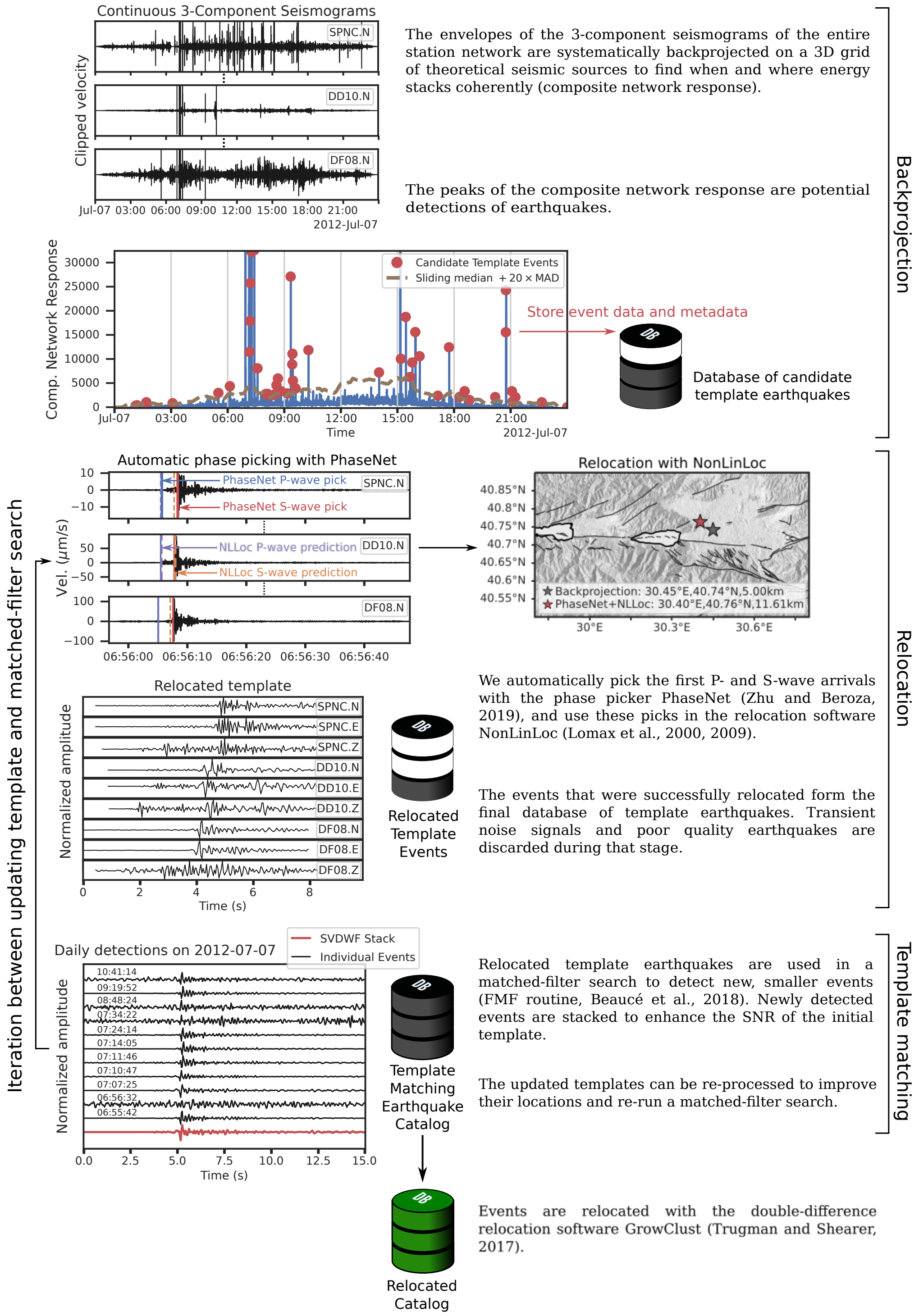


**Figure 2.** Summary flowchart of the earthquake detection and location method. For clarity, only a subset of stations are shown in the above panels, but all the analysis is carried on the 79 stations together. Template matching is performed on the 10 stations closest to the source and the detection threshold is set to 8×RMS of the correlation coefficients in a 30-minute sliding window. See Data and Resources for code availability.

previous study on the same data set (Poyraz et al., 2015). The backprojection method naturally provides an estimate of the location of each detected events. However, the network response finds the times that aligned the envelope maxima rather than the P- and S-wave arrivals, which results in approximate locations.

#### *2.2.2 Relocation*

All the events detected through the CNR were processed with the deep neural network PhaseNet (Zhu & Beroza, 2019) to automatically pick the P- and S-wave first arrivals. These picks were then used by the location software NonLinLoc (Lomax et al., 2000, 2009) to get the earthquake locations and their uncertainties given as 1-$\sigma$ intervals. We required at least four P- and S-wave picks and a total minimum of 15 picks to relocate an event. Requiring both P- and S-wave picks helps constrain the earthquake depth, and imposing at least 15 picks efficiently reduced the number of solutions with very large uncertainties. Events that could not be successfully relocated with NonLinLoc (*e.g.* noisy picks, multiple sources recorded at the same time) were discarded. More information about the input parameters used by PhaseNet and NonLinLoc can be found in Supplementary Material (Section 1).

#### *2.2.3 Template Matching*

Successfully relocated events were kept as templates and used 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 & Ringdal, 2006; Shelly et al., 2007; Ross et al., 2019). It consists in searching for all earthquakes with similar waveforms and moveouts to a known earthquake, that is, earthquakes sharing a similar location and focal mechanism. The similarity is measured by the network-averaged correlation coefficient (CC) between the template waveforms $T_{s,c}$ and the seismograms $u_{s,c}$ shifted by the template moveout $\tau_{s,c}$:

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

where $w_{s,c}$ is the weight attributed to station $s$, component $c$, and $N$ is the length of the template waveforms. We ran 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 used a template window of 8 seconds starting 4 seconds before the S wave on the horizontal components and 1 second before the P wave on the vertical components. We used a detection threshold of 8 times the root mean square (RMS) of the CC time series in a 30-minute sliding window ($8\times\mathrm{RMS}\,\{\mathrm{CC}(t)\}$). The 8 s template duration is adequate given the signal duration of small magnitude earthquakes at $\sim$ 10-50 km source-receiver distances. The $8\times$RMS threshold is in the conservative range of commonly used threshold in template matching studies (*e.g.* Shelly et al., 2007; Ross et al., 2019). Note that $8 \times$ RMS is about 12 times the median absolute deviation (MAD) for a gaussian distribution.

After a matched-filter search over the whole study period, each template earthquake has detected potentially many new similar earthquakes. The similarity of the detected events can be leveraged to form higher SNR waveforms of the template earthquake by summing them. We used the Singular Value Decomposition and Wiener Filtering method (Moreau et al., 2017) for the efficient extraction of coherent signal in the recordings of similar earthquakes. The new template earthquakes with higher SNR waveforms were in turn used to refine the locations, and run another iteration of the matched-filter search. This detection/stacking/relocation workflow is commonly iterated several times in template matching studies. However, stacking the waveforms of similar earthquakes cancels out their differences at high frequencies, and thus acts as a low-pass filter that removes the short-scale information contained in the exact location of an individual event. In order to trade-off the SNR improvement with the loss of short-scale information, we iterated only once in this study.

Neighboring templates often detect the same events, therefore we kept a single event out of all detections occurring within three seconds of each other, from templates whose uncertainty ellipsoids were separated by less than 5 km, and with average waveform similarity greater than 0.33. These thresholds were chosen based on physical considerations (the time threshold 3 sec assumes location errors of up to 10-15 km, the space threshold 5 km accounts for coherency of waves at 2 Hz, etc) and empirically by inspecting the output catalog for duplicated events.

#### 2.2.4 Double-Difference Relative Relocation

We refined the earthquake locations in the region of interest, near the NAFZ and beneath the stations, with the double-difference relative relocation method (*e.g.* Poupinet et al., 1984; Waldhauser & Ellsworth, 2000). P- and S-wave differential arrival times were computed by finding the lag times that maximize the inter-event correlation coefficients and summing them to the travel time differences. The differential times were then processed by the relocation software GrowClust (Trugman & Shearer, 2017, additional information on parameters are given in Supplementary Material). GrowClust estimates location uncertainties with the non-parametric bootstrap resampling method (Efron & Tibshirani, 1986). The original data set is perturbed multiple times by randomly sampling the differential travel times and repeating the location procedure on each such replicated data set. Variations in the locations thus obtained give the position errors.

#### 2.2.5 Magnitude Estimation

Local magnitudes were computed from the amplitude ratios of peak velocities. This required estimating the magnitude of at least one event per template to calibrate our local magnitude scale. Therefore, we computed the moment magnitude $M_w$ by fitting the Brune model (Equation (3), Brune, 1970) to the multi-station average displacement spectra that satisfied an SNR criterion (see details in Section 1.6 and Figure S1).

$$|u_{\mathrm{Brune}}(f)| = \frac{\Omega_0}{\left(1 + \frac{f}{f_c}\right)^2}. \tag{3}$$

In Equation (3), $\Omega_0$ is the low-frequency plateau, which is proportional to the seismic moment $M_0$, and $f_c$ is the corner frequency. Additional information on how we corrected the spectra for geometrical spreading and attenuation to compute $M_0$ from $\Omega_0$ is given in Section 1.6. The moment magnitude $M_w$ is:

$$M_w = \frac{2}{3}\left(\log M_0 - 9.1\right). \tag{4}$$

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

$$M_{L,i} = M_{\mathrm{ref}} + \underset{s,c}{\mathrm{Median}} \left\{ \log \frac{A^i_{s,c}}{A^{\mathrm{ref}}_{s,c}} \right\}, \tag{5}$$

or more generally if there are several reference events:

$$M_{L,i} = \underset{k}{\mathrm{Median}} \left\{ M_{\mathrm{ref},k} + \underset{s,c}{\mathrm{Median}} \left\{ \log \frac{A^i_{s,c}}{A^{\mathrm{ref},k}_{s,c}} \right\} \right\}. \tag{6}$$

Using Equation (6) to compute a local magnitude $M_L$ for every event with a moment magnitude $M_w$, we measured the scaling between $M_w$ and $M_L$ and built the calibration first-order relationship $M_w = A + BM_L$ (see Figure S1B).

### 2.3 Identifying Mining-Related Seismicity

Template matching lends itself particularly well to identifying sources of mining-related earthquakes. We identified these by analyzing the distribution of detection times within the day. Templates that detected more than 80% of events between 6am and 6pm were categorized as mining-related templates (see Figure S2 in Supplementary), since we do not expect natural seismicity to occur within preferred times.

### 2.4 Gutenberg-Richter b-value

The frequency-magnitude distribution of earthquakes typically follows the Gutenberg-Richter law (Gutenberg & Richter, 1941):

$$\log N(M) = a - bM. \tag{7}$$

In Equation (7), $N(M)$ is the number of earthquakes exceeding magnitude $M$, the a-value depends on the total number of observed events, and the b-value controls how frequent larger earthquakes are (typically $b \approx 1$). We estimated the b-value with the maximum likelihood technique (Aki, 1965):

$$b = \frac{1}{\ln(10)\left(\bar{M} - M_c\right)}. \tag{8}$$

Equation (8) is derived for continuous magnitudes $M$ (no bias from binned magnitudes). $M_c$ is the magnitude of completeness, *i.e.* the magnitude above which all events are detected. We computed $M_c$ with the maximum curvature technique (*e.g.* Wiemer & Katsumata, 1999), that is, taking the mode of the (non-cumulative) frequency-magnitude distribution as the magnitude of completeness. We used the kernel density method to estimate the probability density function (pdf) of the frequency-magnitude distribution. We computed $M_c$ on the pdf instead of the raw histogram to mitigate the bin-size dependence of the $M_c$ estimate. The estimation of b-value and magnitude of completeness is illustrated on two earthquake populations in Figure 3.

At each template location, we selected all the templates within a 5 km-radius and used the events they detected to compute $b$ and $M_c$. Following Tormann et al. (2013), we imposed a minimum of 50 events to compute the b-value and, in addition, requested a minimum of 30 events above the magnitude of completeness. As these numbers are still low, we carefully estimated the uncertainties to assess the statistical significance of b-value differences between different groups following Utsu (1966). The confidence interval for a single b-value can be derived from its probability density function $\rho$:

$$\rho(\hat{b}) = \frac{n^n}{\Gamma(n)} \left(\frac{b}{\hat{b}}\right)^{n+1} e^{-n\frac{b}{\hat{b}}} \frac{1}{b}, \tag{9}$$

where $\hat{b}$ is the b-value random variable, $b$ is the estimate as given by Equation (8), $n$ is the number of earthquakes with magnitude $M > M_c$, and $\Gamma$ is the gamma function. Confidence intervals were derived from the percentiles of the cumulative distribution function (see Figure 3C).

Utsu (1966) also noted that the b-value ratio between two populations 1 and 2 follows the F distribution with degrees of freedom $2n_1$ and $2n_2$, where $n_1$ and $n_2$ are the numbers of earthquakes with $M > M_c$ in groups 1 and 2, respectively:

$$\frac{\hat{b_2} b_1}{\hat{b_1} b_2} \sim F(2n_1, 2n_2). \tag{10}$$

In Equation (10), the groups are indexed such that $b_1 > b_2$. The confidence level at which the two b-values are different is equal to the probability that $\hat{b}_1 > \hat{b}_2$:

$$\mathbb{P}\left(\hat{b}_1 > \hat{b}_2\right) = \mathbb{P}\left(\frac{\hat{b}_2 b_1}{\hat{b}_1 b_2} < \frac{b_1}{b_2}\right) \equiv \mathrm{cdf}_{F(2n_1, 2n_2)}\left(\frac{b_1}{b_2}\right). \tag{11}$$

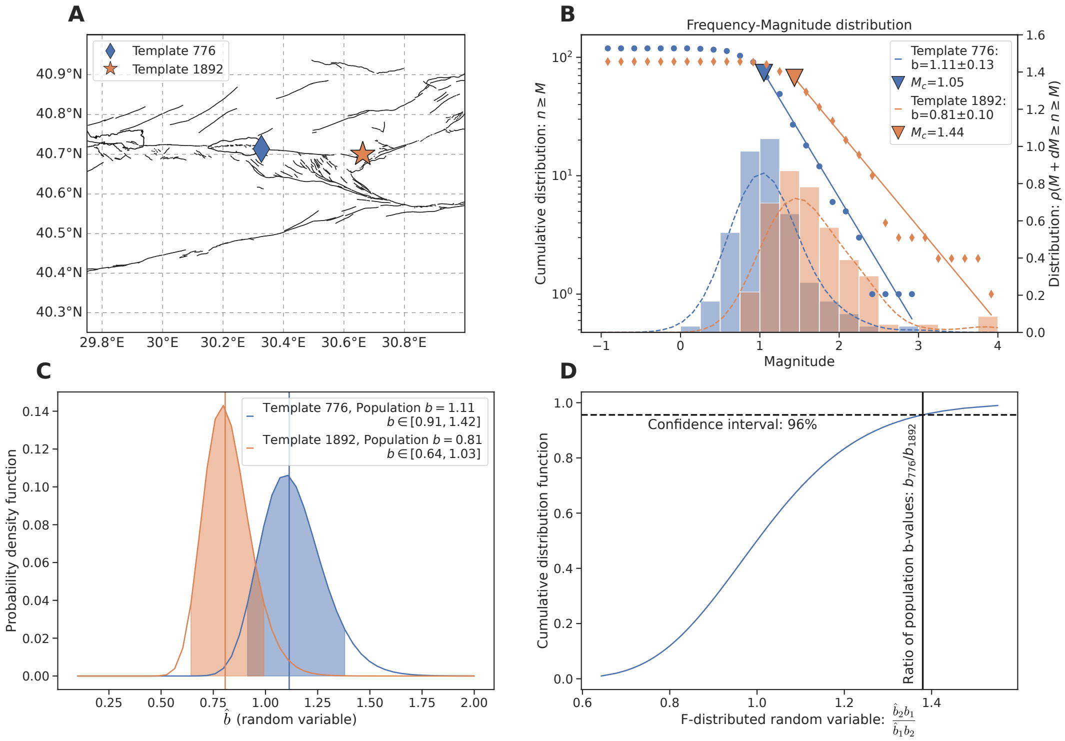


**Figure 3.** Estimation of the b-value and its uncertainties on two earthquake populations (two template families). **A:** Location of the two population centroids. Template 776 is located at the eastern side of Lake Sapanca and template 1892 is located near Akyazi. **B:** Cumulative (scatter plot) and non-cumulative (histogram) frequency-magnitude distributions. Th dashed curves are the kernel density estimate of the non-cumulative probability density functions (pdf). The mode of the pdf is used as the magnitude of completeness (maximum curvature method). The b-value is computed with the maximum likelihood estimate (MLE, Equation (8)). **C:** The b-value pdf computed with Equation (9). The shaded area is the 90% confidence interval, also given in the legend. The b-value population is taken as the MLE (also shown with the vertical bars). **D:** Significance of the b-value difference between two populations using Equation (11). In this example, the b-value of template 776's event family is greater than template 1892's at the 96% confidence level.

In Equation (11), the right-hand term is the cumulative distribution function (cdf) of the F distribution. This method is exemplified at Figure 3D.

### 2.5 Temporal Clustering

In this manuscript, we refer to temporal clustering as the property of earthquake sequences in which events influence the timings (advance or delay) of future earthquakes, that is, the non-randomness of earthquake sequences (*e.g.* Gardner & Knopoff, 1974; Marsan & Lengline, 2008). We quantified the strength of temporal clustering in earthquake sequences by analyzing the statistical properties of the number of earthquakes per unit time, which we refer to as the earthquake occurrence time series. It is given by:

$$e(t) = \text{Number of eventsin}\,[t; t + \Delta t]\,, \tag{12}$$

where $\Delta t$ is a user-defined time bin duration, and $t$ is the calendar time. An example is given in Figure 4A. Burst-like sequences covering wide intervals of recurrence times are not random (see Figure 4B,C) but clustered in time. Time clustered seismicity exhibits time scale invariant characteristics. The spectrum of the earthquake occurrence $e(t)$ (as computed by Equation (12)) follows a power law of frequency ($\propto f^{-\beta}$, see Figure 4D), and the time series $e(t)$ shows a fractal statistics (Figure 4E). We measured the fractal dimension of the 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 were occupied by at least one earthquake (Smalley Jr et al., 1987; Lowen & Teich, 2005). For a certain range of time bin sizes $\tau$, we observe:

$$x \propto \tau^{1-D}. \tag{13}$$

In Equation (13), $D$ is the fractal dimension of the time series. The fractal dimension varies between the two end-members $D = 0$ for a point process (*e.g.* Poisson point process for the background seismicity), and $D = 1$ for a line (uninterrupted seismicity). A large fractal dimension ($D > 0.2$) characterizes cascade-like activity where past events strongly influence the timings of future events. Fractal analysis has been used in multiple studies to characterize earthquake clustering (Smalley Jr et al., 1987; Lee & Schwarcz, 1995; Beaucé et al., 2019). Note that periodic seismicity does not follow a fractal behavior and cannot be characterized by this method. Building the $x(\tau)$ curve (Equation (13), Figure 4E) is computationally more simple than estimating the spectrum (Figure 4D). Likewise, it is simpler to fit $x(\tau)$. Therefore, we chose to compute the fractal dimension $D$ to characterize temporal clustering in the rest of this study.

The method described in Section 2.5 does not explicitly deal with space. However, we applied this analysis to subsets of the earthquake catalog containing neighboring earthquakes (as described for the b-value, see Section 2.4), and thus obtained a fractal dimension for each template.

# 3 Results

## 3.1 The Earthquake Catalog

### *3.1.1 Regional Seismicity*

Following the method described in Section 2.2, we built a database of 3,546 templates and with them detected 35,437 events, including both natural and anthropogenic seismicity. We applied our analysis between 38.50°N-41.50°N and 28.00°E-32.00°E (see Figure 1A). Figure 5 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. We purposely present an earthquake catalog for this region that extends far beyond the NAFZ to provide a comprehensive description of the earthquake signals found in the data set. We found that most of the 1,972 events detected with templates located 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. Furthermore, we found that about half of the detected seismicity was due to mining activity (see Section 2.3): among the 31,329 earthquakes detected with the 3,320 templates, we identified 16,674 natural earthquakes and 14,655 mining-related earthquakes. The locations of mining activity that we identified (see Figure 5) agree well with the analysis of Poyraz et al. (2015) (their Figure 3) whereas the Kandilli catalog (see Data and Resources) tends to report less explosions, in particular beneath the DANA array (see Figure S3).

The majority of earthquakes occurred outside the station array and not in the NAFZ itself, that is, north of 40.80°N or south of 40.30°N (see Figure 5B). Location uncertainties increase with increasing distance from the DANA array: inside 40.30°N-41.00°N and 30.00°E-30.50°E, the average horizontal uncertainty is $\bar{h}_{\max} = 0.97$ km and the average

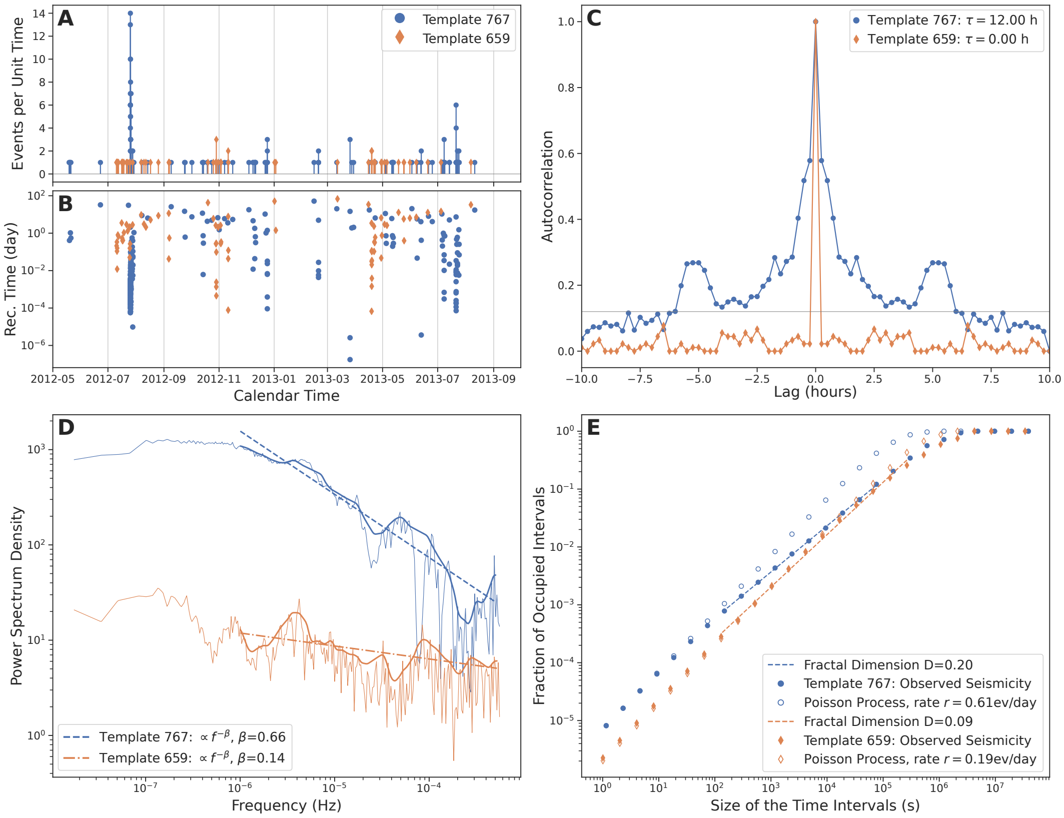


**Figure 4.** Quantifying the strength of temporal clustering in a strongly clustered sequence (Template 767, blue dots) and a weakly clustered sequence (Template 659, orange diamonds). **A:** Number of earthquakes per unit time (referred to as earthquake occurrence, see Equation (12)). **B:** Recurrence times vs. origin times. **C:** Autocorrelation of the earthquake occurrence time series. The horizontal black line is the arbitrary threshold used to define the correlation time $\tau$. **D:** Power spectral density of the earthquake occurrence. The linear trend, in the log-log space, is the exponent of the power-law that indicates a scale invariant process. **E:** Fractal analysis of the earthquake occurrence (see text and Equation (13)). We measure the slope between $dt_{\mathrm{min}}$=100 s and $dt_{\mathrm{max}} = 1/r$, where $r = N/T$ is the average seismic rate (number of events $N$ divided by time span $T$). For reference, for each template we simulate the seismicity from a Poisson point process with average rate $r$. The slope of the Poisson point process gives a fractal dimension $D = 0$ (*i.e.* dimension of a point).

vertical uncertainty is $\bar{v}_{\mathrm{max}} = 0.74$ km, whereas these uncertainties increase to $\bar{h}_{\mathrm{max}} = 8.53$ km and $\bar{v}_{\mathrm{max}} = 4.57$ km outside this box (see Figure 5C,D).

Since accurate moment magnitude estimation rely on correct source-receiver distances (see Section 2.2.5), we only computed moment magnitudes for events with $h_{\mathrm{max}} < 5$ km (see Figure 5D). After the SNR criterion, we could estimate moment magnitudes within 168 template families of natural seismicity, from which we computed 1,929 local magnitudes. These magnitudes range from -1 to 4, and we obtained $b = 0.85$ and $M_c = 1.18$ (see Section 2.4 and Figure S4). We computed a $M_w$-$M_L$ calibration close to identity, $M_w = 0.15 + 0.93 M_L$ (see Section 2.2.5 and Figure S1B). The magnitude of completeness of our catalog indicates that we were not able to estimate magnitudes below $M \approx 1$ but still detected them: only 12% of the detected seismicity has a magnitude estimate. Therefore, $M_c = 1.18$ is only an upper bound to the magnitude of complete-

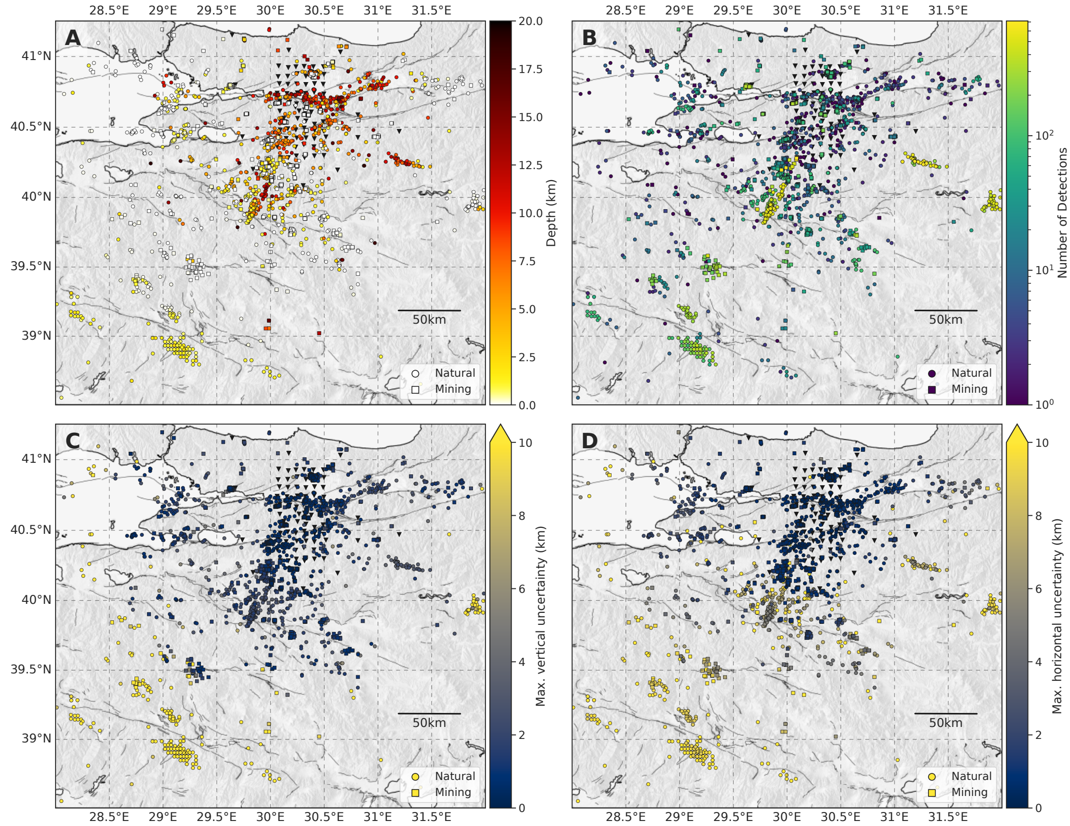


**Figure 5.** Map view of the locations of the template earthquakes detected and used in this study. Only templates with maximum horizontal uncertainty less than 15 km and depth less than 20 km are shown (total of 3,320 templates). Filled dots are for natural earthquakes (1,471 templates), and squares 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 detected earthquakes actually originate from outside the North Anatolian Fault Zone. **C:** Maximum vertical uncertainty $v_{\max}$, i.e. depth range spanned by the projection of the uncertainty ellipse onto a vertical plane. **D:** Maximum horizontal uncertainty $h_{\max}$, i.e. length of the major semi-axis of the projection of the uncertainty ellipse onto the horizontal plane.

ness of the whole catalog. For reference, we estimated a b-value and magnitude of completeness of $b = 0.91$ and $M_c = 1.05$ with the catalog published in Poyraz et al. (2015), with magnitudes ranging from 0 to 4 (see Figure S4). Our magnitudes seemed to be systematically larger than theirs for smaller events, with an average difference of 0.5 unit over all compared events (see Figure S4C). Detailed b-values and magnitudes of completeness are presented in Section 3.2.

We present the spatio-temporal distribution of the seismicity in Figure 6. An overall decaying activity of natural earthquakes is superimposed to a uniform mining-related activity (compare Figure 6A vs. B). We observe two sequences of slowly decaying activity below 39°N and around 40°N. The southernmost earthquake sequence (39°N) is part of the aftershock activity of the M5.1 2012-05-03 39.18°N/29.10°E/5.4 km earthquake (just before the deployment of DANA). The 40°N sequence is not featured in the Kandilli nor in the United States Geological Survey catalog.

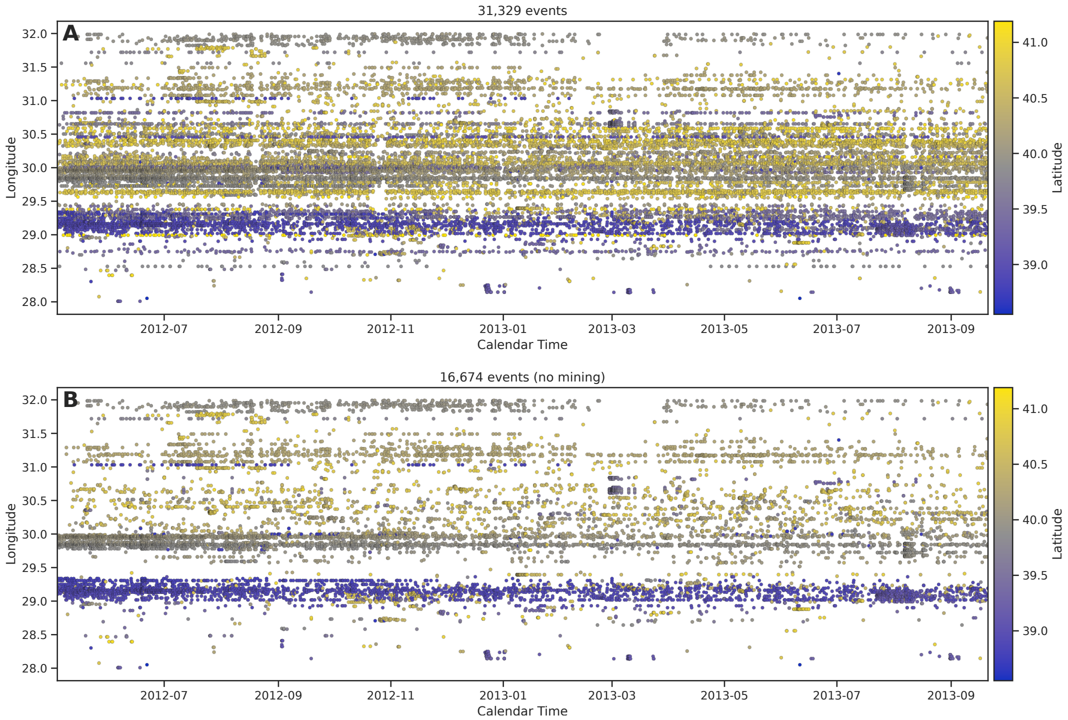


**Figure 6.** Spatio-temporal distribution of the earthquake activity in the study region. The longitude of each event is shown against its origin time, and the color codes the latitude. **A:** We detected 31,329 events with the 3,320 template earthquakes presented in Figure 5 from 2012-05-04 to 2013-09-20. **B:** The templates due to natural seismicity detected 16,674 earthquakes. The seismic activity taking place on the NAFZ (latitudes 40.35°N-40.80°N) represents a small amount of the total seismicity (∼ 2,000 events).

Our earthquake catalog and detection and location codes are available online (see Data and Resources, and see Supplementary Material for additional information about the structure of the catalog file). This analysis of the regional seismicity shows that most of the detected seismic activity occurred outside the North Anatolian Fault Zone, which may be a feature of this section of the NAFZ being early in its earthquake cycle (Ben-Zion & Zaliapin, 2020). In the following, we focus on the template earthquakes located in the vicinity of the NAFZ and near the station array (40.25°N-41.00°, 29.80°E-31.00°E).

#### *3.1.2 Seismicity of the North Anatolian Fault Zone*

Figure 7 shows the locations of the template earthquakes in the fault zone, as well as the 2,141 earthquakes relocated with the double-difference method (see Section 2.2). The median horizontal and vertical errors on relative locations are 73 m and 91 m, respectively, meaning that they can reliably be interpreted in terms of active structures. Earthquake hypocenters reveal a complex network of faults, with much of the seismicity occurring on secondary faults rather than on the NAFZ itself. We divided the fault zone into nine subregions (*cf.* Figure 7A) whose names we will keep referring to in this manuscript. These are organized into four along-strike sections: Izmit-Sapanca, fault-parallel Sapanca-Akyazi, Karadere, and the entire southern strand, and six fault-perpendicular sections: Lake Sapanca west and east, fault perpendicular Sapanca-Akyazi, Akyazi, and the southern strand west and east. The northern strand is overall more active than the southern strand, and the Sapanca-Akyazi segment hosts the densest activity. In particular, both terminations of the segment, the eastern side of Lake Sapanca and the area around the Akyazi fault, host strong seismicity. The Akyazi region features the deepest seismicity in the vicinity of the NAFZ (down to 20 km). The group of earthquakes located at the northermost of the Sapanca-Akyazi region (Figure 7A-B) are part of the 2012-07-07 $M_L$4.1 Serdivan earthquake sequence. Most of the seismicity along the southern strand occurs in areas where surface fault traces indicate more structural complexity. Note that the relocated seismicity tends to be distributed in patches, 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 7C show the events' depth distribution. The seismicity is enhanced in the lower half of the seismogenic zone: 7-15 km along the northern strand, and even deeper than 15 km around the Akyazi fault, and 5-10 km depth along the southern strand. The main exception to that depth distribution are the earthquakes at the western side of Lake Sapanca, with hypocenters clustered around 5 km depth. The Lake Sapanca W. fault perpendicular cross-section (see Figure 7C) shows that this shallow seismicity seems restricted to the southern side of the fault, namely the Armutlu Bloc.

The map views and cross-sections in Figure 7 suggest 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. We emphasize that the detected microseismicity illuminates the deformation zone associated with the NAFZ rather than the fault itself. The Sapanca-Akyazi and Akyazi fault perpendicular cross-sections could indicate a north dipping deformation zone, although these mostly show almost horizontally aligned earthquakes. Under the assumption that the deformation zone does dip towards the north, we approximately measure a 60° dip angle in the middle of the Sapanca-Akyazi segment, and 85° near the Akyazi fault. Identifying a global dip direction of the deformation zone along the southern strand is equally ambiguous. In the east, one could either identify slightly south dipping structures ($\sim 85°$) or more strongly north dipping structures ($\sim 70°$).

We present the temporal distribution of the seismicity in Figure 8. The recurrence times are given against their detection times for each of the nine cross-sections introduced above. The recurrence time is the time interval between two consecutive co-located earthquakes. In practice, recurrence times are computed as the time intervals between consecutive events detected by a same template. The most striking feature of Figure 8 is the organization of some earthquake sequences into bursts of seismicity with recurrence times spanning many orders of magnitude. These sequences are time clustered (*e.g.* W. B. Frank et al., 2016; Beaucé et al., 2019) and recurrence times are power-law distributed (*e.g.* Utsu et al., 1995). These bursts are usually associated with sequences of foreshocks-mainshock-aftershocks, although in general earthquake sequences can have no clear mainshock (that is, 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 and near Akyazi is almost exclusively organized into such sequences of burst-like seismicity, whereas the southern strand hosts much less of these burst-like episodes. Figure 8 also reports the local magnitudes (see Section 2.2.5). 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, including 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). The area around the Akyazi fault also produced four $M_L > 3$ earthquakes, whereas earthquakes near Lake Sapanca did not exceed $M_L = 3$.

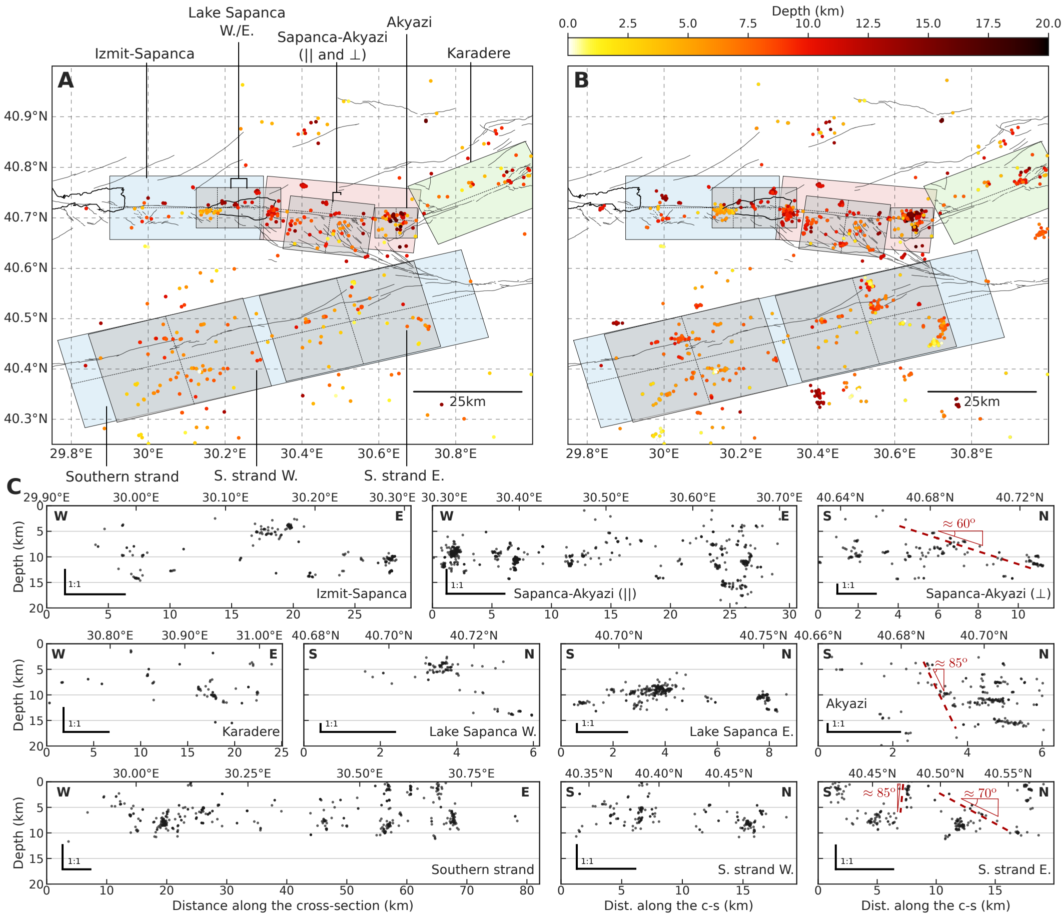


**Figure 7.** Earthquakes in the North Anatolian Fault Zone. **A:** Locations of the template earthquakes with color coded depths. We define nine subregions along the different segments of the fault. Only in this figure the Sapanca-Akyazi region is subdivided into a fault parallel and a fault perpendicular sections. The thin black dotted 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. 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). Note that the x scales and the aspect ratio across cross-sections vary. The 1:1 aspect ratio is drawn in the lower left corner of each cross-section. The dashed red lines and angles are given for reference but are not our take-away message.

#### *3.1.3 Comparison with Past Seismicity*

We combined different earthquake catalogs to compare the 2012-2013 detected seismicity with the pre-Izmit, Izmit-Düzce, and early post-Düzce seismicity (Bulut et al., 2007; Ickrath et al., 2015; Bohnhoff et al., 2016, and see Figure 9). We note that the Izmit-Düzce earthquake catalog is more complete in the west (around the Izmit-Sapanca segment) than the pre-Izmit and early post-Düzce catalogs due to the higher number of sta-

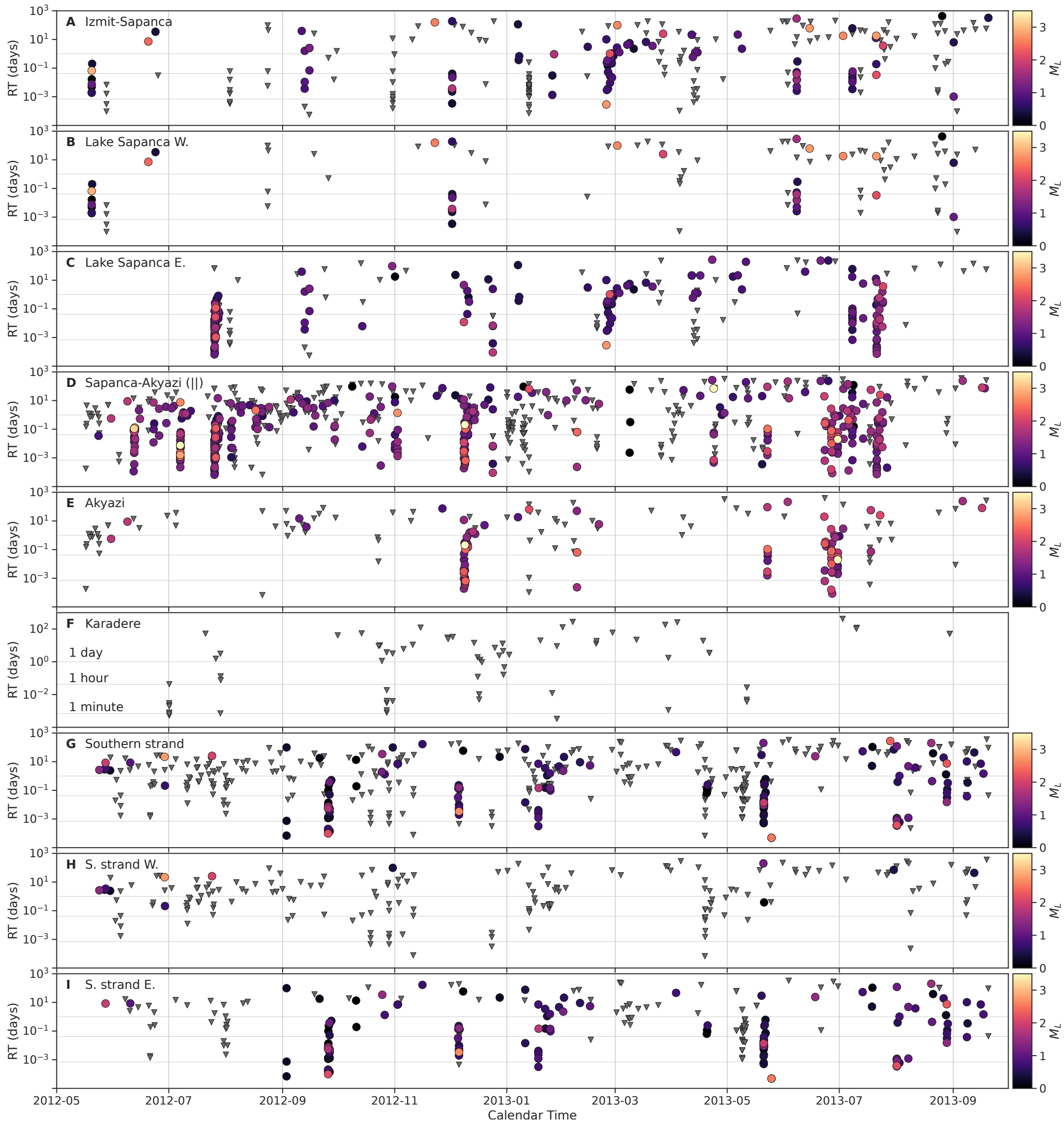


**Figure 8.** Time evolution of the earthquake recurrence times for different subsets of the earthquake catalog (refer to Figure 7 for the name of the areas). The recurrence time is the time between two consecutive events detected by a same template. Note that the y-axis is in log scale and that some seismic episodes span many orders of magnitude of recurrence time. These episodes are characteristic of burst-like, or cascade activity (see text). The color scale indicates the local magnitude, and inverted grey triangles are events for which no reliable estimates were obtained.

tions used in this time period (see, *e.g.* Ickrath et al., 2015, and Figure 9B). It is also worth mentioning that these three catalogs show both natural and mining-related seismicity whereas we have discarded the man-made seismicity to the best of our ability (see Section 2.3). During these three time periods, the (moment) magnitudes of completeness of these catalogs are $M_c$ =1.56, 1.69, and 1.44, respectively, and few $M_w < 1$ earthquakes are reported (see Figure S5). Using our $M_w$-$M_L$ calibration to convert our local magnitudes to effective moment magnitudes, we obtained $M_c = 1.18$ and 27% of the earthquakes contributing to the frequency-magnitude distribution have $M_w < 1$ (see

Figure S5D). Therefore, also recalling that most of the detected events are too small to be characterized by their magnitude, our catalog reports smaller events than the catalogs we compare it with.

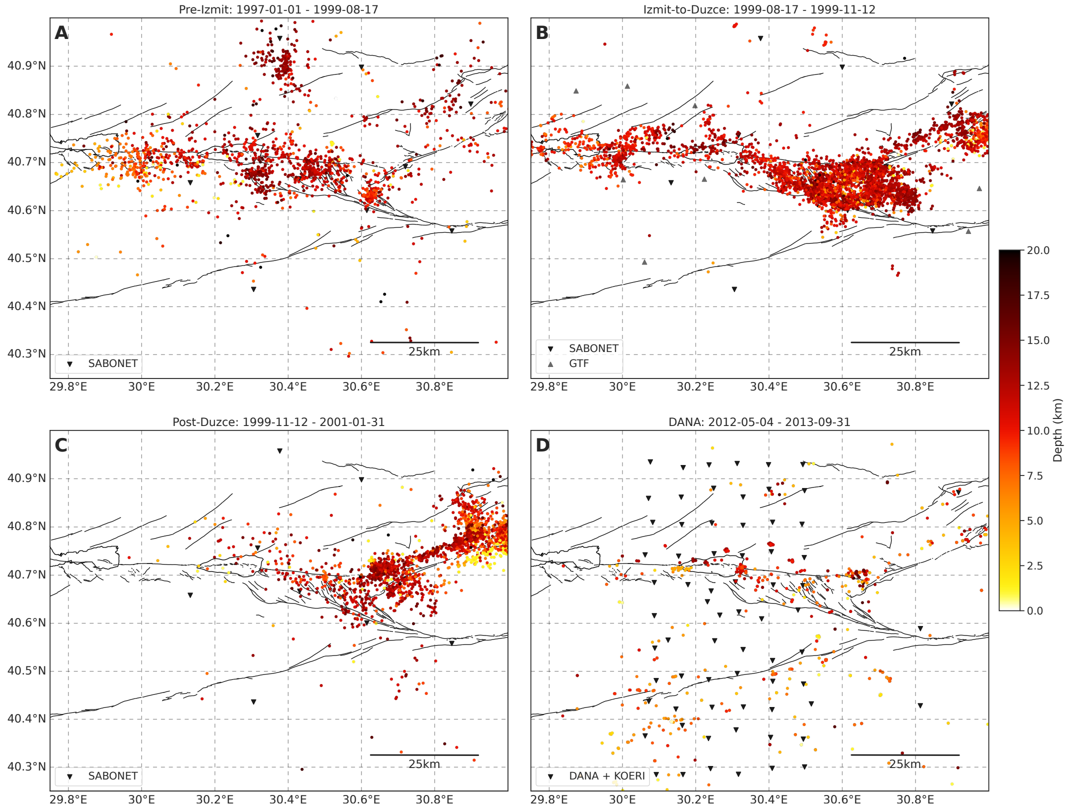


**Figure 9.** Comparison of the **A:** pre-Izmit, **B:** Izmit-Düzce, the SABONET stations were complemented by 21 temporary stations from the German Task Force (GTF, grey triangles), **C:** early post-Düzce, and **D:** late post-Düzce seismicity. The inverted black triangle are the seismic stations and the colored dots are the earthquake locations.

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 (Figure 9A). The Izmit earthquake is known to have nucleated near a swarm of seismicity that was active before the M7.4 event (Crampin et al., 1985; Lovell et al., 1987; Ito et al., 2002). In the three months between the Izmit earthquake and the Düzce event, the seismic activity was strongest in the area around the triple junction between the Sapanca-Akyazi segment, the Karadere segment, and the Mudurnu fault (Figure 9B). The Izmit hypocentral region remained active and, comparatively, little activity was detected near Lake Sapanca. After the Düzce earthquake, most activity along the Izmit-Sapanca and Sapanca-Akyazi terminated, and seismicity concentrated along the Karadere segment (Figure 9C). The Akyazi region, where little coseismic slip was observed (Ozalaybey et al., 2002; Bohnhoff et al., 2006, 2008), hosted a cluster of strong activity, possibly driven by the Izmit residual stresses. Note that no seismicity was detected near Lake Sapanca. About 13 years after the Izmit and Düzce earthquakes, we detected the strongest activity at the eastern side of Lake Sapanca, and near the Akyazi fault (Figure 9D). If not due to the absence of $M < 1$ earthquakes in these catalogs, the lack of intense seismicity near Lake Sapanca in the early post-Düzce period suggests that faults near Lake Sapanca did not slip during the afterslip-driven aftershock sequence with Omori-like decaying seismicity (Perfettini & Avouac, 2004). More-over, the Omori law predicts a seismicity rate about four orders of magnitude lower 13 years after the mainshock (using Omori law parameters from Bayrak & Öztürk, 2004), therefore the seismic activity near Lake Sapanca should have been high after the Izmit earthquake if the 2012-2013 seismicity were to be remnants of aftershocks. The 2012-2013 Lake Sapanca seismicity also appears much stronger than the pre-Izmit seismicity (Figure 9A).

### 3.2 Observed b-values

Computed b-values and magnitudes of completeness (see Section 2.4) are presented in Figure 10. Of most interest to this study, we see that earthquakes at the eastern side of Lake Sapanca exhibit high b-values ($b \approx 1.1$), whereas earthquakes near the Akyazi fault show low b-values ($b \approx 0.8$). The magnitude of completeness varies from $M_c \approx 1.3$ near Akyazi to $M_c \approx 1.0$ near Lake Sapanca. We note that visual checking of the frequency-magnitude distributions showed that, in general, they follow the Gutenberg-Richter law well, except for the Serdivan earthquakes where a peak around $M_L \approx 2.5$ can be observed. The significance of the b-value difference between the eastern Lake Sapanca and Akyazi was assessed by applying the statistical test presented in Section 2.4 (see Equation (11)). We found that the difference was significant at the 96% confidence level (see Figure 3D).

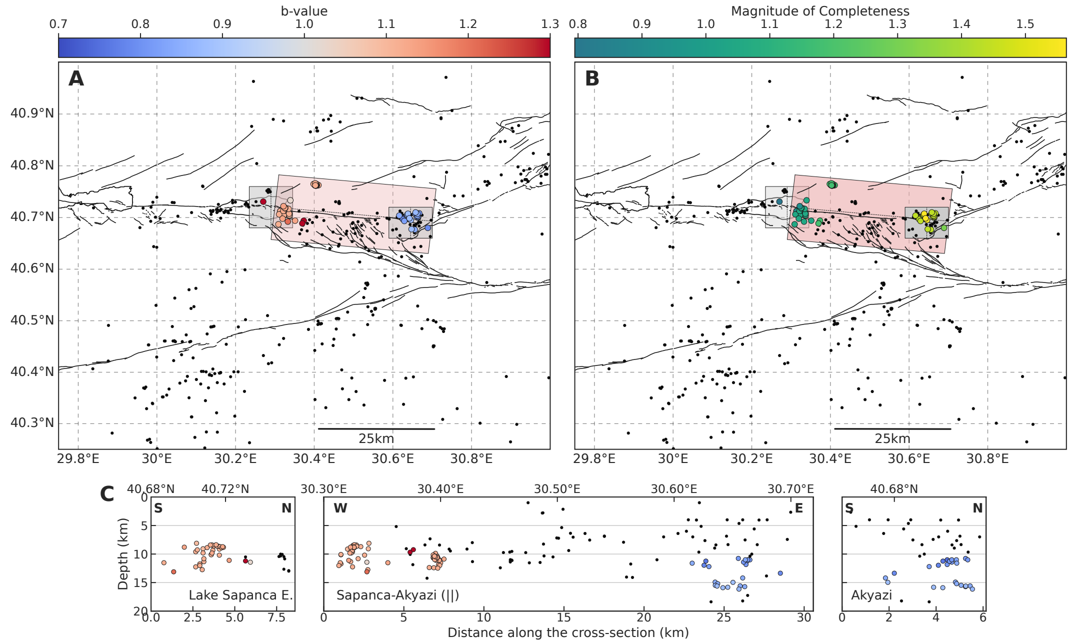


**Figure 10.** **A:** Map view of template earthquakes with color coded Gutenberg-Richter b-value. Smaller black dots are event families for which we could not compute moment magnitudes (see text). **B:** Map view of template earthquakes with color coded magnitude of completeness. In both top panels, the shaded areas refer to the regions introduced in Figure 7. **C:** Template earthquakes with color coded b-value on fault parallel and fault perpendicular cross-sections. Hypocenters are projected along the dotted axes shown on the map view.

### 3.3 Observed Temporal Clustering

We characterized temporal clustering as a function of space (see Figure 11) following the method described in Section 2.5. The strongest temporal clustering (fractal dimension $D > 0.20$) is observed on the eastern side of Lake Sapanca, beneath the so-called Rangefront trace. Other areas of strong activity, like the Serdivan earthquakes (around 30.404°E/40.763°N) and the Akyazi area, only show small-to-moderate temporal clustering ($D < 0.14$), thus confirming the outstanding character of the eastern Lake Sapanca. We note that while the temporal organization of recurrence times shown in Figure 8 indicated burst-like seismicity in all of the above mentioned areas, this quantitative analysis was necessary to distinguish between strongly and moderately time clustered sequences. A few other isolated locations exhibit strong temporal clustering, and seem to be systematically occurring near the bottom of the seismogenic zone (*cf.* Figure 11C). Comparing the cumulative number of detections per template and their fractal dimension shows that there is no trivial correlation between the two (see Figure 11A vs. B). We note that we did the same fractal analysis on all templates of the study region and found another region of strong temporal clustering on the NAFZ, in the eastern Marmara Sea, where the 1999 Izmit earthquake arrested (see Figure S6).

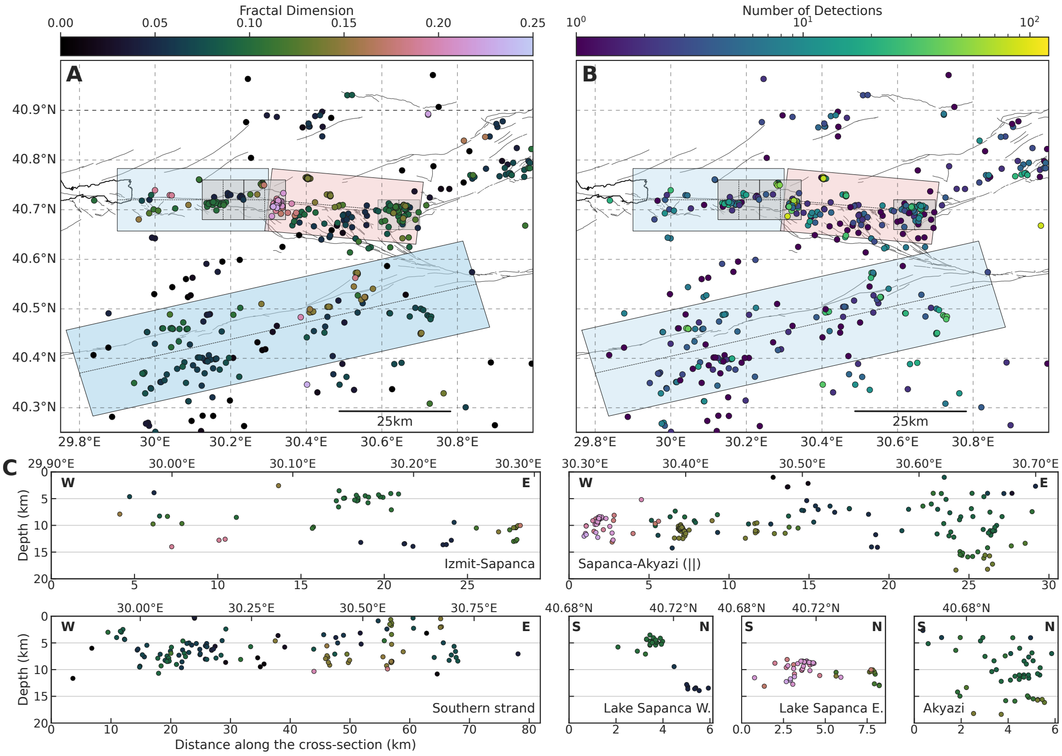


**Figure 11.** **A:** Map view of template earthquakes with color coded fractal dimension (*cf.* Equation (13)) showing the strength of temporal clustering. **B:** Map view of template earthquakes with color coded cumulative number of detections. In both top panels, the shaded areas refer to the regions introduced in Figure 7. **C:** Template earthquakes with color coded fractal dimension on fault parallel and fault perpendicular cross-sections. Hypocenters are projected along the dotted 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 Interpretation and Discussion

In the present section, we discuss implications of the observed spatial earthquake distribution for the mechanical state of the NAFZ (Section 4.1), we interpret the b-values in terms of low and high stresses (Section 4.2), and we explain how temporal clustering can be related to fault rheology (Section 4.3).

### 4.1 Spatial Distribution of Seismicity

Apart from the Karadere segment, the seismicity is taking place off the main fault on a complex network of secondary faults, similarly to the Izmit-Düzce aftershocks (*e.g.* Ozalaybey et al., 2002; Bulut et al., 2007, and see Figure 7). This feature is in stark contrast with the simplicity of the Izmit and Düzce earthquakes, which occurred on simple fault segments (Barka et al., 2002; Langridge et al., 2002). Off-fault seismicity has also been observed to be a characteristic of fault zones early in their seismic cycle (Ben-Zion & Zaliapin, 2020) and might be due to off-fault, distributed deformation contributing to accommodate slip deficits resulting from heterogeneous slip along the fault (Dolan & Haravitch, 2014).

Shallow creep has been observed along the Izmit-Sapanca and the Sapanca-Akyazi segments (*e.g.* Çakir et al., 2012; Hussain et al., 2016; Aslan et al., 2019). The creep rates and creep locations along the Izmit rupture have evolved with time (*e.g.* Bürgmann et al., 2002). Aslan et al. (2019) use 2011-2017 InSAR data and 2014-2016 GPS data and, thus, is the closest study to ours in time. The authors' model shows shallow creep down to 5 km along the Izmit-Sapanca segment and down to 2 km at the western end of the Sapanca-Akyazi segment. Such shallow creep should drive microseismicity in the vicinity of the creeping fault sections (*e.g.* Lohman & McGuire, 2007). The depth cross-sections (Figure 7C) only show shallow seismicity at the western Lake Sapanca ($\approx$ 5 km depth). Although these depths are consistent with the creep depth given in Aslan et al. (2019), a direct causality link to the shallow creep is not straightforward because it is taking place off-fault (Figure 7B). Seismicity along the Sapanca-Akyazi segment do not support creep-driven activity at shallow depths ($\approx$ 2 km). However, hypocenters suggest that, at the time of the study, the base of the seismogenic zone is around 10-15 km, which is in good agreement with Aslan et al. (2019).

Comparing our catalog with the seismicity in the past (see Section 3.1.3) showed that the eastern Lake Sapanca did not appear to be a particularly active area, either before or right after the Izmit earthquake. However, this comparison relies on catalogs made with different methods and station coverage, and, consequently, different magnitudes of completeness. The seismic activity at the eastern Lake Sapanca may be a permanent feature of the step-over that can only be observed with low magnitude of completeness catalogs ($M_c \lesssim 1.0$). We further discuss in Section 4.4 whether the seismicity at the eastern Lake Sapanca is a new feature of the fault zone caused by the post-Izmit deformation or is a constant phenomenon that could not be observed in such detail in the past. Variations in seismicity along the southern strand are harder to interpret because the lack of earthquakes in previous catalogs (see Figure 9) is partly due to the absence of stations in the past.

### 4.2 Gutenberg-Richter b-value

Laboratory experiments have shown that the b-value seems to be controlled by the state of stress, specifically that $b$ decreases with increasing differential stress (*e.g.* C. H. Scholz, 1968; Amitrano, 2003). Decreasing b-value with depth (Mori & Abercrombie, 1997; Wiemer & Wyss, 1997) and high b-value along creeping sections (*e.g.* Amelung & King, 1997; Wiemer & Wyss, 1997) also support the negative correlation of $b$ with stress. Thus, the b-value can be used as a stressmeter.

Our results (Figure 10) show a clear difference in b-values between the eastern Lake Sapanca ($b \approx 1.1$) and the Akyazi ($b \approx 0.8$) seismicity. We recall that this difference is significant at the 96% confidence level (see Section 3.2). We interpret the higher b-values at Lake Sapanca as an indication of low background stresses, while we interpret the lower b-values at Akyazi as indicating high background stresses. Low stress at the eastern Lake Sapanca suggests that aseismic slip might play a role in driving the seismicity, implying that, there, faults have weak sections. High stress near Akyazi can be understood as resulting from the stress concentration that occurred during the Izmit earthquake, when little co-seismic slip occurred along the Akyazi fault and the Akyazi gap.

## 4.3 Temporal Clustering, Earthquake Interactions, and Fault Mechanical Properties

Strongly time clustered seismicity with a wide range of recurrence times, as presented in Section 3.3, cannot be explained only by fluctuations of the background seismicity rate, for example due to the injection of fluids at depth. Indeed, a Poisson point process with a transient increased rate only shifts the distribution of recurrence times towards shorter times but does not widen the distribution and does not have a large fractal dimension (see Figure S7). Temporal clustering, that is, cascading of events, emerges when different faults or sections of a fault interact (*e.g.* Burridge & Knopoff, 1967; Marsan & Lengline, 2008; Fischer & Hainzl, 2021). Earthquakes can trigger each other due to the static stress changes induced by the co- and postseismic displacements (*e.g.* King & Cocco, 2001), but also due to the dynamic stress changes induced by the elastic waves radiated by the rapid coseismic motions (*e.g.* Fan & Shearer, 2016). Furthermore, because of the stress redistribution following any slip motion (not necessarily at seismic speeds), interaction can occur between a seismogenic asperity and its creeping surroundings: accelerated creep (*e.g.* afterslip) increases the stressing rate on the asperity (*e.g.* Cattania, 2019; Cattania & Segall, 2021). In realistic, complex conditions where seismic and aseismic slip co-occurs on short length scales (*e.g.* Collettini et al., 2011), numerical models show that both co-seismic and creep mediated stress changes are important factors controlling the clustering of earthquakes (Dublanchet et al., 2013; Cattania & Segall, 2021). The contribution of creep mediated stress transfers to temporal clustering might even be more important than static stress changes due to the breaking of asperities (Dublanchet, 2019). In fact, this means that both seismic and aseismic events can cluster in time, but that earthquake catalogs only capture the seismic signature of temporal clustering. Effectively, these interacting stress fields result in a clock advance or delay in the cycle of the earthquake sources (*e.g.* Harris et al., 1995; Gomberg et al., 1998) and thus in non-random earthquake sequences.

Figure 12 sketches different earthquake interaction scenarios explaining temporal clustering: in a locked fault Figure 12A, and with creep mediated stress transfers Figure 12B. Note that remote creep acting on a sparse asperity population (Figure 12C) would produce Poissonian seismicity (*e.g.* Lohman & McGuire, 2007). Thus, areas of strong temporal clustering (see Figure 11) indicate faults with intrinsic properties: heterogeneous rheology resulting in juxtaposed seismic and aseismic slip, and rough or densely fractured fault zone providing many seismogenic asperities. These properties enhance interaction-driven seismicity, that is, driven by the redistributed stresses of past events. However, the long time-scale behavior of clustered seismicity may be modulated by time-dependent remote forcing.

Where rheology transitions from brittle to ductile, for example at the base of the seismogenic zone, faults are likely to host both unstable, seismic slip and stable, aseismic slip (C. H. Scholz, 1998; Skarbek et al., 2012). Therefore, seismicity near the bottom of the seismogenic zone would be expected to display temporal clustering because, there, interacting asperities are likely to be embedded in a creeping fault (*cf.* Figure 12B, Dublanchet et al., 2013). We investigated the relationship between temporal clustering and the proximity to the bottom of the seismogenic zone to elucidate the role of fault stability in our observations (*i.e.* scenario Figure 12A vs. 12B). The results, in Figure 13, indicate that, as expected, seismicity tends to get more time clustered as it gets closer to the brittle-ductile transition and that strong clustering almost always happens at the bottom of the seismogenic zone. Exceptions are at the western side of Lake Sapanca (Figure 13B) where results might be biased due to the absence of significant seismicity at

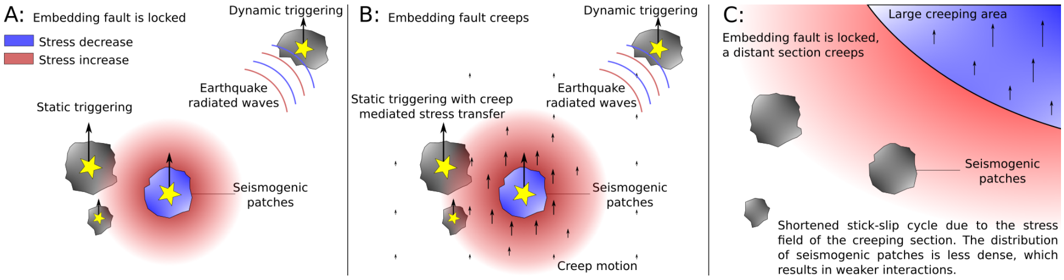


**Figure 12.** Sketch of different earthquake interaction scenarios. **A:** Seismogenic asperities embedded in a locked fault. **B:** Seismogenic asperities embedded in a creeping fault. In **A** and **B**, the color shows the stress change due to rupture of the seismogenic patch. The triggered ruptures occur with some delay. **C:** Seismogenic asperities embedded in a locked fault, but stressed by a remote creeping section of the fault. The asperities are not close enough to the creeping patch to strongly interact via static stress changes. The spatial configuration of asperities does not promote strong interactions.

depth, and along the Karadere segment (Figure 13F) where large source-receiver distances yield poor hypocentral depth resolution and thus low confidence results.

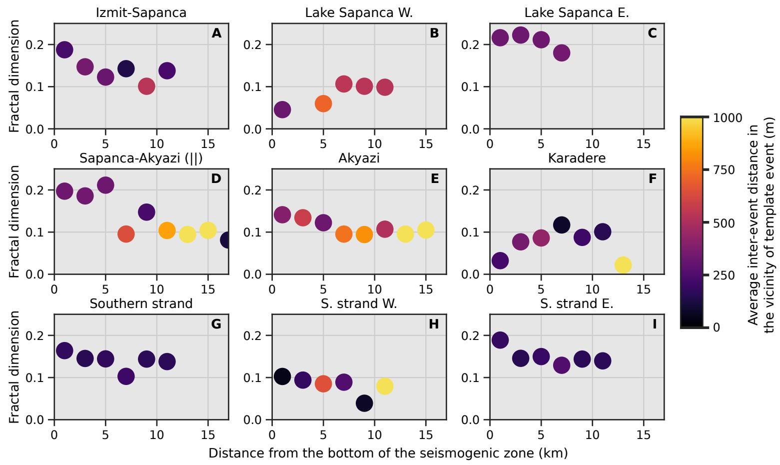


**Figure 13.** Clustering vs. depth vs. event density. Inside each region, templates are binned per distance from the bottom of the seismogenic zone and the fractal dimension is averaged among the 10% largest values, resulting in a "soft" maximum of each bin. The location of the bottom of the seismogenic zone is approximated by the depth of the locally deepest template. Dots are colored according to the average inter-event distance within the neighboring earthquake subcatalogs; this is a proxy for asperity density. Darker colors mean higher density. Strongest clustering tends to occur at the bottom of the seismogenic zone, *i.e.* at the transition zone between unstable (brittle) and stable (ductile) sliding.

We also investigated a possible correlation between the proximity to the brittle-ductile transition and the density of seismic sources, which could as well explain the increase in temporal clustering. We took the average inter-event distance within neighboring earthquake subcatalogs as a proxy for asperity density. We note that this measure of asperity density is imperfect because a single asperity can break repeatedly. The smaller number of detected earthquakes along the southern strand might also be insufficient to compute a meaningful average inter-event distance. We do not observe a clear systematic increase in asperity density with decreasing distance from the bottom of the seismogenic zone, but the observational limits mentioned above prevent us from drawing definite conclusions. Figure 13 rather shows that both the proximity to the brittle-ductile transition and a large event density favor temporal clustering. Our observations therefore support that dense asperity populations along with creep mediated stress transfers do promote strong temporal clustering (*cf.* Figure 12B, Dublanchet et al., 2013). Thus, this study suggests that faults at the eastern side of Lake Sapanca are in heterogeneous stability regimes allowing unstable (seismic) and stable (aseismic) slip.

### 4.4 Implications for the Lake Sapanca Step-Over

In summary, the Gutenberg-Richter b-values (see Section 4.2) and temporal clustering (see Section 4.3) point to the role of different rheological properties in producing earthquakes between the two sides of Lake Sapanca. At the western side, the shallow active sections seem incapable of producing strongly time clustered seismicity. At the eastern side, the depth distribution, the strong temporal clustering (Figure 11C), and the high b-values (Figure 10) suggest that thin along-dip fault sections slip in a mixed seismic and aseismic mode.

Heterogeneous faults near the brittle-ductile transition have stable and unstable sections (*e.g.* Collettini et al., 2011). Weakly unstable sections may produce transient episodes of slow slip (*e.g.* Bürgmann, 2018). The temporal distribution of earthquakes at the eastern Lake Sapanca (see Figure 8C) suggests that faults are slipping during intermittent episodes of deformation. Thus, the weakest sections of the faults at the eastern Lake Sapanca might be intermittently driven to slowly slip and, in turn, activate the seismogenic asperities (Skarbek et al., 2012; Cattania & Segall, 2021). The lack of seismicity along the up-dip sections suggests they are either fully locked or fully creeping, but there is no evidence for such a large creeping section in geodetic data (*e.g.* Aslan et al., 2019). Geologic data suggest that the so-called Sapanca Complex, constituted of weak serpentinites and strong metabasites (Akbayram et al., 2013, and references therein), might reach the southeastern side of Lake Sapanca at depth where we observe the highly clustered seismicity. Such lithology is consistent with the scenario of strong asperities embedded in a weak, stable fault. How much seismic moment is released through (partial) aseismic slip during these episodes of strong microseismicity remains an open question.

Whether these intermittent episodes of deformation are a permanent feature of Lake Sapanca or result from the mechanical changes that faults underwent because of co- and post-seismic stress changes is hard to elucidate entirely since our comparison with the past seismicity relies on unequal catalogs (see Section 3.1.3). However, we know from geodetic data that north-south extension around the Lake Sapanca step-over accelerated considerably following the Izmit earthquake (Ergintav et al., 2009; Hearn et al., 2009). Stress analyses have also shown that the NAFZ weakened after the Izmit-Düzce earthquake sequence (*e.g.* Pınar et al., 2010; Ickrath et al., 2015). Given that deformation in the step-over is faster than before the Izmit earthquake, (micro)seismicity should also be stronger. Early post Izmit-Düzce seismicity (before early 2001) is either lacking from the catalogs due to insufficient detection capability, or the increase of seismicity occurred later due to postseismic relaxation processes such as enhanced slip rates below seismogenic depths.

The postseismic response of at least two releasing step-overs of the NAFZ, Lake Sapanca and another one in the eastern Marmara Sea, has been shown to produce substantial north-south extension following the Izmit earthquake (Ergintav et al., 2009; Hearn et al., 2009). Ergintav et al. (2009) have shown that models of postseismic slip on the main fault do not account well for the north-south extension in these two step-overs, in particular after the first three years. We further compared these step-overs by extending our temporal clustering analysis further along the NAFZ and found that the eastern Marmara Sea was also hosting clustered seismicity at the eastern termination of the Princes Islands segment (*cf.* Figure S6). Large earthquake location uncertainties prevented us from carrying the same detailed study but this section has been identified as an area of high b-value (Raub et al., 2017). We can hypothesize that Lake Sapanca and the eastern Marmara Sea behave similarly. In both cases, fault heterogeneities, and perhaps their stress history, could explain an hybrid seismic and aseismic slip regime (Collettini et al., 2011). As to how much slip is accommodated seismically vs. aseismically and whether the aseismic part is related to the deformation missing from the current models has to be addressed by the means of geodesy.

## 5 Summary and Concluding Remarks

We processed 1.5 years of continuous data collected during the DANA experiment (May 2012 - September 2013, see Data and Resources) with an automated earthquake detection and location method (Beaucé et al., 2019, and see Section 2) and produced an earthquake catalog with 31,329 events between 38.50°N-41.50°N and 28.00°E-32.00°E, with depths shallower than 20 km, and horizontal location uncertainty less than 15 km (see Section 3.1). We found that 14,655 detected events were induced or triggered by mining activity against 16,674 natural earthquakes, the latter mostly occurring outside of the North Anatolian Fault Zone itself. We focused our analysis on about 2,000 relocated earthquakes in the NAFZ and near the station array.

We analyzed the earthquake catalog to investigate collective properties of earthquakes: the b-value of the Gutenberg-Richter law (see Section 3.2), which we related to the level of background stresses driving the ruptures (see Section 4.2), and the strength of temporal clustering (see Section 3.3), which we interpreted in terms of interacting stress fields and fault rheology (see Section 4.3). We showed that strongest temporal clustering almost systematically occurred in the brittle-ductile transition zone, suggesting that a mixed seismic-aseismic slip regime enhances temporal clustering (see Section 4.3).

We found that the patterns of seismicity have changed durably after the Izmit-Düzce earthquake sequence (see Section 3.1.3). The region near the Akyazi fault, where the co-seismic displacement was noticeably low, was still one of the most active areas some thirteen years later. This seismicity indicate a low b-value ($b \approx 0.8$, *cf.* Sections 3.2 and 4.2) and weak-to-moderate time clustering (see Sections 3.3 and 4.3), suggesting that the high residual stresses left by the absence of co-seismic release are driving the seismicity.

We also detected strong seismicity around Lake Sapanca. At the western side, we observed shallow ($\approx 5$ km, see Section 3.1.2) and weakly-to-moderately time clustered seismicity (see Section 3.3). Although it is occurring off the main fault, the depths and weak clustering are consistent with creep-driven seismicity (see Section 4.3, Figure 12C). The seismicity at the eastern side takes place in a narrow depth interval at the bottom of the seismogenic zone ($\approx$ 10-13 km depth, see Section 3.1.2), has a high b-value ($b \approx 1.1$, see Section 3.2), and is strongly time clustered (see Section 3.3). We suggested that these are the characteristics of mixed seismic and aseismic slip on heterogeneous faults at the brittle-ductile transition. Such east-west differences over a short distance likely reflect the heterogeneous geology of the region (*e.g.* Akbayram et al., 2013). The proposed rheology of the faults at the eastern Lake Sapanca could indicate that these intersect the so-called Sapanca complex at depth, which is made of weak and strong materials. The Lake Sapanca seismicity is a major feature of our earthquake catalog but might have been missed in the past due to insufficient detection capability (see Section 4.1).

The results of our study emphasize the important role of secondary structures in the late postseismic stage of the NAFZ, and possibly through the interseismic phase. The structural complexity of these structures appears in stark contrast to the relatively simple co-seismic dynamics of the Izmit earthquake (*i.e.* rupture on almost straight and vertical fault segments Barka, 1999; Langridge et al., 2002). The north-south extension across the Lake Sapanca step-over accelerated following the Izmit earthquake (Ergintav et al., 2009; Hearn et al., 2009), and we therefore question whether the proposed seismic-aseismic heterogeneous slip regime could be related to this deformation (see Section 4.4). The seismicity supports, but not prove, the possibility of slow slip in the step-over. We suggested that the releasing step-over in the Marmara Sea, with similar temporal clustering and accelerated extension following the Izmit earthquake, could behave analogously to the Lake Sapanca step-over. The present study does not provide the means to relate the observed surface deformation to slip on specific faults, but it does encourage the search for slow slip on normal faults in these step-overs. Finally, our study emphasizes that slip may not always happen in well separated seismic and aseismic sections but, instead, may happen over complex, intricate unstable and stable domains.

## 6 Data and Resources

The earthquake catalog is available at the Zenodo data set repository (DOI: `10.5281/zenodo.6362973`). We used the version 1.0.1 of our BPMF Python package for earthquake detection and location, which is stored at `https://doi.org/10.5281/zenodo.6780316` (last accessed December 2021). The last version is maintained on Github 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/`, last accessed December 2021). The maps were made with the Cartopy Python library (version 0.18.0, last accessed December 2021, Met Office, 2010 - 2015). 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/KO`).

## 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.

## References


Akbayram, K., Okay, A. I., & Satır, M. (2013). Early cretaceous closure of the intra-pontide ocean in western pontides (northwestern turkey). *Journal of Geodynamics*, *65*, 38–55.

Aki, K. (1965). Maximum likelihood estimate of b in the formula log n= a-bm and its confidence limits. *Bull. Earthq. Res. Inst., Tokyo Univ.*, *43*, 237–239.

Amelung, F., & King, G. (1997). Earthquake scaling laws for creeping and non-creeping faults. *Geophysical research letters*, *24*(5), 507–510.

Amitrano, D. (2003). Brittle-ductile transition and associated seismicity: Experimental and numerical studies and relationship with the b value. *Journal of Geophysical Research: Solid Earth*, *108*(B1).

Aslan, G., Lasserre, C., Cakir, Z., Ergintav, S., Özarpaci, S., Dogan, U., ... Renard, F. (2019). Shallow creep along the 1999 Izmit earthquake rupture (Turkey) from GPS and high temporal resolution interferometric synthetic aperture radar data (2011–2017). *Journal of Geophysical Research: Solid Earth*, *124*(2), 2218–2236.

Barka, A. (1999). The 17 august 1999 izmit earthquake. *Science*, *285*(5435), 1858–1859.

Barka, A., Akyuz, H., Altunel, E., Sunal, G., Cakir, Z., Dikbas, A., ... others (2002). The surface rupture and slip distribution of the 17 august 1999 Izmit earthquake (m 7.4), North Anatolian fault. *Bulletin of the Seismological Society of America*, *92*(1), 43–60.

Bayrak, Y., & Öztürk, S. (2004). Spatial and temporal variations of the aftershock sequences of the 1999 izmit and düzce earthquakes. *Earth, planets and space*, *56*(10), 933–944.

Beaucé, E., Frank, W. B., Paul, A., Campillo, M., & van der Hilst, R. D. (2019). Systematic detection of clustered seismicity beneath the Southwestern Alps. *Journal of Geophysical Research: Solid Earth*, *124*(11), 11531–11548.

Beaucé, E., Frank, W. B., & Romanenko, A. (2018). Fast matched filter (FMF): An efficient seismic matched-filter search for both CPU and GPU architectures. *Seismological Research Letters*, *89*(1), 165–172.

Ben-Zion, Y., & Zaliapin, I. (2020). Localization and coalescence of seismicity before large earthquakes. *Geophysical Journal International*, *223*(1), 561–583.

Bohnhoff, M., Bulut, F., Görgün, E., Milkereit, C., & Dresen, G. (2008). Seismotectonic setting at the north anatolian fault zone after the 1999 mw= 7.4 izmit earthquake based on high-resolution aftershock locations. *Advances in Geosciences*, *14*, 85–92.

Bohnhoff, M., Grosser, H., & Dresen, G. (2006). Strain partitioning and stress rotation at the North Anatolian fault zone from aftershock focal mechanisms of the 1999 Izmit m w= 7.4 earthquake. *Geophysical Journal International*, *166*(1), 373–385.

Bohnhoff, M., Ickrath, M., & Dresen, G. (2016). Seismicity distribution in conjunction with spatiotemporal variations of coseismic slip and postseismic creep along the combined 1999 izmit-düzce rupture. *Tectonophysics*, *686*, 132–145.

Bouchon, M., Bouin, M.-P., Karabulut, H., Toksöz, M. N., Dietrich, M., & Rosakis, A. J. (2001). How fast is rupture during an earthquake? new insights from the 1999 Turkey earthquakes. *Geophysical Research Letters*, *28*(14), 2723–2726.

Bouchon, M., Karabulut, H., Aktar, M., Özalaybey, S., Schmittbuhl, J., & Bouin, M.-P. (2011). Extended nucleation of the 1999 mw 7.6 Izmit earthquake. *science*, *331*(6019), 877–880.

Brune, J. N. (1970). Tectonic stress and the spectra of seismic shear waves from earthquakes. *Journal of geophysical research*, *75*(26), 4997–5009.

Bulut, F., Bohnhoff, M., Aktar, M., & Dresen, G. (2007). Characterization of aftershock-fault plane orientations of the 1999 izmit (turkey) earthquake using high-resolution aftershock locations. *Geophysical Research Letters*, *34*(20).

Bürgmann, R. (2018). The geophysics, geology and mechanics of slow fault slip. *Earth and Planetary Science Letters*, *495*, 112–134.

Bürgmann, R., Ergintav, S., Segall, P., Hearn, E. H., McClusky, S., Reilinger, R. E., ... Zschau, J. (2002). Time-dependent distributed afterslip on and deep below the izmit earthquake rupture. *Bulletin of the Seismological Society of America*, *92*(1), 126–137.

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

Cakir, Z., Chabalier, J.-B. d., Armijo, R., Meyer, B., Barka, A., & Peltzer, G. (2003). Coseismic and early post-seismic slip associated with the 1999 izmit earthquake (turkey), from sar interferometry and tectonic field observations. *Geophysical Journal International*, *155*(1), 93–110.

Çakir, Z., Ergintav, S., Özener, H., Dogan, U., Akoglu, A. M., Meghraoui, M., & Reilinger, R. (2012). Onset of aseismic creep on major strike-slip faults. *Geology*, *40*(12), 1115–1118.

Cattania, C. (2019). Complex earthquake sequences on simple faults. *Geophysical Research Letters*, *46*(17-18), 10384–10393.

Cattania, C., & Segall, P. (2021). Precursory slow slip and foreshocks on rough faults. *Journal of Geophysical Research: Solid Earth*, *126*(4), e2020JB020430.

Collettini, C., Niemeijer, A., Viti, C., Smith, S. A., & Marone, C. (2011). Fault structure, frictional properties and mixed-mode fault slip behavior. *Earth and Planetary Science Letters*, *311*(3-4), 316–327.

Crampin, S., Evans, R., & Üçer, S. B. (1985). Analysis of records of local earthquakes: the turkish dilatancy projects (tdp1 and tdp2). *Geophysical Journal International*, *83*(1), 1–16.

DANA . (2012). *Dense array for north anatolia.* International Federation of Digital Seismograph Networks. Retrieved from `http://www.fdsn.org/doi/10.7914/SN/YH_2012` doi: 10.7914/SN/YH_2012

Dolan, J. F., & Haravitch, B. D. (2014). How well do surface slip measurements track slip at depth in large strike-slip earthquakes? the importance of fault structural maturity in controlling on-fault slip versus off-fault surface deformation. *Earth and Planetary Science Letters*, *388*, 38–47.

Dublanchet, P. (2019). Scaling and variability of interacting repeating earthquake sequences controlled by asperity density. *Geophysical Research Letters*, *46*(21), 11950–11958.

Dublanchet, P., Bernard, P., & Favreau, P. (2013). Interactions and triggering in a 3-d rate-and-state asperity model. *Journal of Geophysical Research: Solid Earth*, *118*(5), 2225–2245.

Efron, B., & Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical science*, 54–75.

Ergintav, S., McClusky, S., Hearn, E., Reilinger, R., Cakmak, R., Herring, T., ... Tari, E. (2009). Seven years of postseismic deformation following the 1999, m= 7.4 and m= 7.2, izmit-düzce, turkey earthquake sequence. *Journal of Geophysical Research: Solid Earth*, *114*(B7).

Fan, W., & Shearer, P. M. (2016). Local near instantaneously dynamically triggered aftershocks of large earthquakes. *Science*, *353*(6304), 1133–1136.

Fischer, T., & Hainzl, S. (2021). The growth of earthquake clusters. *Frontiers in Earth Science*, *9*, 79.

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

Frank, W. B., Shapiro, N. M., Husker, A. L., Kostoglodov, V., Gusev, A. A., & Campillo, M. (2016). The evolving interaction of low-frequency earthquakes during transient slip. *Science advances*, *2*(4), e1501616.

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

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

Gomberg, J., Beeler, N., Blanpied, M., & Bodin, P. (1998). Earthquake triggering by transient and static deformations. *Journal of Geophysical Research: Solid Earth*, *103*(B10), 24411–24426.

Gutenberg, B., & Richter, C. (1941). *Seismicity of the earth* (Vol. 34). Geological Society of America.

Harris, R. A., Simpson, R. W., & Reasenberg, P. A. (1995). Influence of static stress changes on earthquake locations in southern california. *Nature*, *375*(6528), 221–224.

Hearn, E., McClusky, S., Ergintav, S., & Reilinger, R. (2009). Izmit earthquake postseismic deformation and dynamics of the north anatolian fault zone. *Journal of Geophysical Research: Solid Earth*, *114*(B8).

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

Hussain, E., Wright, T. J., Walters, R. J., Bekaert, D., Hooper, A., & Houseman, G. A. (2016). Geodetic observations of postseismic creep in the decade after the 1999 Izmit earthquake, Turkey: Implications for a shallow slip deficit. *Journal of Geophysical Research: Solid Earth*, *121*(4), 2980–3001.

Ickrath, M., Bohnhoff, M., Dresen, G., Martinez-Garzon, P., Bulut, F., Kwiatek, G., & Germer, O. (2015). Detailed analysis of spatiotemporal variations of the stress field orientation along the izmit-düzce rupture in nw turkey from inversion of first-motion polarity data. *Geophysical Journal International*, *202*(3), 2120–2132.

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

Ito, A., Uçer, B., Bariş, Ş., Nakamura, A., Honkura, Y., Kono, T., … Işikara, A. M. (2002). Aftershock activity of the 1999 izmit, turkey, earthquake revealed from microearthquake observations. *Bulletin of the Seismological Society of America*, *92*(1), 418–427.

Kahraman, M., Cornwell, D. G., Thompson, D. A., Rost, S., Houseman, G. A., Türkelli, N., … Gülen, L. (2015, Nov). Crustal-scale shear zones and heterogeneous structure beneath the North Anatolian Fault Zone, Turkey, revealed by a high-density seismometer array. *Earth Planet. Sci. Lett.*, *430*, 129–139. doi: 10.1016/j.epsl.2015.08.014

Kandilli Observatory And Earthquake Research Institute, Boğaziçi University. (1971). *Bogazici university kandilli observatory and earthquake research institute.* International Federation of Digital Seismograph Networks. Retrieved from `http://www.fdsn.org/doi/10.7914/SN/KO` doi: 10.7914/SN/KO

Karabulut, H., Schmittbuhl, J., Özalaybey, S., Lengline, O., Kömeç-Mutlu, A., Durand, V., … Bouin, M. (2011). Evolution of the seismicity in the eastern Marmara Sea a decade before and after the 17 August 1999 Izmit earthquake. *Tectonophysics*, *510*(1-2), 17–27.

Karahan, A. E., Berckhemer, H., & Baier, B. (2001). Crustal structure at the western end of the North Anatolian Fault Zone from deep seismic sounding.

King, G., & Cocco, M. (2001). Fault interaction by elastic stress changes: New clues from earthquake sequences. In *Advances in geophysics* (Vol. 44, pp. 1–VIII). Elsevier.

Langridge, R., Stenner, H. D., Fumal, T., Christofferson, S., Rockwell, T., Hartleb, R., . . . Barka, A. (2002). Geometry, slip distribution, and kinematics of surface rupture on the sakarya fault segment during the 17 august 1999 izmit, turkey, earthquake. *Bulletin of the Seismological Society of America*, *92*(1), 107–125.

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

Le Pichon, X., & Angelier, J. (1979). The hellenic arc and trench system: a key to the neotectonic evolution of the eastern mediterranean area. *Tectonophysics*, *60*(1-2), 1–42.

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

Lomax, A., Michelini, A., Curtis, A., & Meyers, R. (2009). Earthquake location, direct, global-search methods. *Encyclopedia of complexity and systems science*, *5*, 2449–2473.

Lomax, A., Virieux, J., Volant, P., & Berge-Thierry, C. (2000). Probabilistic earthquake location in 3D and layered models. In *Advances in seismic event location* (pp. 101–134). Springer.

Lovell, J., Crampin, S., Evans, R., & Üçer, S. B. (1987). Microearthquakes in the tdp swarm, turkey: clustering in space and time. *Geophysical Journal International*, *91*(2), 313–330.

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

Marsan, D., & Lengline, O. (2008). Extending earthquakes' reach through cascading. *Science*, *319*(5866), 1076–1079.

McClusky, S., Balassanian, S., Barka, A., Demir, C., Ergintav, S., Georgiev, I., . . . others (2000). Global positioning system constraints on plate kinematics and dynamics in the eastern mediterranean and caucasus. *Journal of Geophysical Research: Solid Earth*, *105*(B3), 5695–5719.

Meade, B. J., Hager, B. H., McClusky, S. C., Reilinger, R. E., Ergintav, S., Lenk, O., . . . Ozener, H. (2002). Estimates of seismic potential in the marmara sea region from block models of secular deformation constrained by global positioning system measurements. *Bulletin of the Seismological Society of America*, *92*(1), 208–215.

Met Office. (2010 - 2015). Cartopy: a cartographic python library with a matplotlib interface [Computer software manual]. Exeter, Devon. Retrieved from `http://scitools.org.uk/cartopy`

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

Mori, J., & Abercrombie, R. E. (1997). Depth dependence of earthquake frequency-magnitude distributions in california: Implications for rupture initiation. *Journal of Geophysical Research: Solid Earth*, *102*(B7), 15081–15090.

Ozalaybey, S., Ergin, M., Aktar, M., Tapirdamaz, C., Biçmen, F., & Yörük, A. (2002). The 1999 izmit earthquake sequence in turkey: seismological and tectonic aspects. *Bulletin of the Seismological Society of America*, *92*(1), 376–386.

Papaleo, E., Cornwell, D., & Rawlinson, N. (2018). Constraints on North Anatolian Fault Zone width in the crust and upper mantle from S wave teleseismic tomography. *Journal of Geophysical Research: Solid Earth*, *123*(4), 2908–2922.

Parsons, T., Toda, S., Stein, R. S., Barka, A., & Dieterich, J. H. (2000). Heightened odds of large earthquakes near istanbul: An interaction-based probability calculation. *Science*, *288*(5466), 661–665.

Perfettini, H., & Avouac, J.-P. (2004). Postseismic relaxation driven by brittle creep: A possible mechanism to reconcile geodetic measurements and the decay rate of aftershocks, application to the chi-chi earthquake, taiwan. *Journal of*

*Geophysical Research: Solid Earth*, *109*(B2).

Pınar, A., Üçer, S., Honkura, Y., Sezgin, N., Ito, A., Barış, Ş., ... Horiuchi, S. (2010). Spatial variation of the stress field along the fault rupture zone of the 1999 Izmit earthquake. *Earth, planets and space*, *62*(3), 237–256.

Poupinet, G., Ellsworth, W., & Frechet, J. (1984). Monitoring velocity variations in the crust using earthquake doublets: An application to the calaveras fault, california. *Journal of Geophysical Research: Solid Earth*, *89*(B7), 5719–5731.

Poyraz, S. A., Teoman, M. U., Türkelli, N., Kahraman, M., Cambaz, D., Mutlu, A., ... others (2015). New constraints on micro-seismicity and stress state in the western part of the North Anatolian Fault Zone: Observations from a dense seismic array. *Tectonophysics*, *656*, 190–201.

Raub, C., Martínez-Garzón, P., Kwiatek, G., Bohnhoff, M., & Dresen, G. (2017). Variations of seismic b-value at different stages of the seismic cycle along the north anatolian fault zone in northwestern turkey. *Tectonophysics*, *712*, 232–248.

Reilinger, R., Ergintav, S., Bürgmann, R., McClusky, S., Lenk, O., Barka, A., ... others (2000). Coseismic and postseismic fault slip for the 17 august 1999, m= 7.5, izmit, turkey earthquake. *Science*, *289*(5484), 1519–1524.

Reilinger, R., McClusky, S., Vernant, P., Lawrence, S., Ergintav, S., Cakmak, R., ... others (2006). GPS constraints on continental deformation in the Africa-Arabia-Eurasia continental collision zone and implications for the dynamics of plate interactions. *Journal of Geophysical Research: Solid Earth*, *111*(B5).

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

Scholz, C. (1968). Microfractures, aftershocks, and seismicity. *Bulletin of the Seismological Society of America*, *58*(3), 1117–1130.

Scholz, C. H. (1968). The frequency-magnitude relation of microfracturing in rock and its relation to earthquakes. *Bulletin of the seismological society of America*, *58*(1), 399–415.

Scholz, C. H. (1998). Earthquakes and friction laws. *Nature*, *391*(6662), 37–42.

Scholz, C. H. (2015). On the stress dependence of the earthquake b value. *Geophysical Research Letters*, *42*(5), 1399–1402.

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

Skarbek, R. M., Rempel, A. W., & Schmidt, D. A. (2012). Geologic heterogeneity can produce aseismic slip transients. *Geophysical Research Letters*, *39*(21).

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

Stein, R. S., Barka, A. A., & Dieterich, J. H. (1997). Progressive failure on the North Anatolian fault since 1939 by earthquake stress triggering. *Geophysical Journal International*, *128*(3), 594–604.

Taylor, G., Rost, S., Houseman, G. A., & Hillers, G. (2019). Near-surface structure of the North Anatolian Fault zone from Rayleigh and Love wave tomography using ambient seismic noise. *Solid Earth*, *10*(2), 363–378.

Toksoz, M., Reilinger, R., Doll, C., Barka, A., & Yalcin, N. (1999). Izmit (Turkey) earthquake of 17 August 1999: first report. *Seismological Research Letters*, *70*(6), 669–679.

Toksöz, M., Shakal, A., & Michael, A. (1979). Space-time migration of earthquakes along the North Anatolian fault zone and seismic gaps. *Pure and Applied Geophysics*, *117*(6), 1258–1270.

Tormann, T., Wiemer, S., Metzger, S., Michael, A., & Hardebeck, J. L. (2013). Size distribution of parkfield's microearthquakes reflects changes in surface creep rate. *Geophysical Journal International*, *193*(3), 1474–1478.

Trugman, D. T., & Shearer, P. M. (2017). GrowClust: A hierarchical clustering algorithm for relative earthquake relocation, with application to the Spanish Springs and Sheldon, Nevada, earthquake sequences. *Seismological Research Letters*, *88*(2A), 379–391.

Utkucu, M., Nalbant, S. S., McCloskey, J., Steacy, S., & Alptekin, Ö. (2003). Slip distribution and stress changes associated with the 1999 november 12, düzce (turkey) earthquake (mw= 7.1). *Geophysical Journal International*, *153*(1), 229–241.

Utsu, T. (1966). A statistical significance test of the difference in b-value between two earthquake groups. *Journal of Physics of the Earth*, *14*(2), 37–40.

Utsu, T., Ogata, Y., et al. (1995). The centenary of the omori formula for a decay law of aftershock activity. *Journal of Physics of the Earth*, *43*(1), 1–33.

Waldhauser, F., & Ellsworth, W. L. (2000). A double-difference earthquake location algorithm: Method and application to the northern hayward fault, california. *Bulletin of the Seismological Society of America*, *90*(6), 1353–1368.

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

White, M. C., Fang, H., Nakata, N., & Ben-Zion, Y. (2020). PyKonal: A Python Package for Solving the Eikonal Equation in Spherical and Cartesian Coordinates Using the Fast Marching Method. *Seismological Research Letters*, *91*(4), 2378–2389.

Wiemer, S., & Katsumata, K. (1999). Spatial variability of seismicity parameters in aftershock zones. *Journal of Geophysical Research: Solid Earth*, *104*(B6), 13135–13151.

Wiemer, S., & Wyss, M. (1997). Mapping the frequency-magnitude distribution in asperities: An improved technique to calculate recurrence times? *Journal of Geophysical Research: Solid Earth*, *102*(B7), 15115–15128.

Zhu, W., & Beroza, G. C. (2019). PhaseNet: a deep-neural-network-based seismic arrival-time picking method. *Geophysical Journal International*, *216*(1), 261–273.

# Supporting Information for "Microseismic Constraints on the Mechanical State of the North Anatolian Fault Zone Thirteen Years after the 1999 M7.4 Izmit Earthquake"

Eric Beaucé[1,3], Robert D. van der Hilst[1], Michel Campillo[2,1]

[1]Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, USA

[2]Institut des Sciences de la Terre, Université Grenoble Alpes, France

[3]Lamont-Doherty Earth Observatory, Columbia University, NY, USA

**Contents of this file**

1. Table S1.
2. Figures S1 to S7.
3. Additional information about method parameters.

## 1. Earthquake Catalog

### 1.1. Velocity Model

See Table S1.

### 1.2. Automated Phase Picking with PhaseNet

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

---

### 1.3. Absolute Earthquake Location with NonLinLoc

NonLinLoc (NLLoc, Lomax et al., 2000, 2009) 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.

### 1.4. Double-difference Relative Relocation with GrowClust

GrowClust (Trugman & 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).

### 1.5. 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.

### 1.6. Magnitude Estimation

Within each family of earthquakes detected by a same template, we computed the S-wave spectra with the multi-taper method (Prieto et al., 2009). The SNR was 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 was used to compute the multi-channel weighted average of the S-wave spectra (see Equation (1) and Figure S1A).

$$\bar{v}(f) = \frac{1}{W(f)} \sum_{s,c} w_{s,c} \alpha_{s,c} v_{s,c}(f), \qquad W(f) = \sum_{s,c} w_{s,c}(f). \tag{1}$$

In Equation (1), $v_{s,c}(f)$ is the velocity spectrum of station $s$, component $c$ at frequency $f$, $w_{s,c}$ is the corresponding weight (see Figure S1A) and $\alpha_{s,c}$ is the factor that corrects for geometric spreading and attenuation (see Equation (5)). The average spectra were converted to displacement spectra $u(f)$ and fitted with the Brune model (Equation (2), Brune, 1970):

$$|u_{\mathrm{Brune}}(f)| = \frac{\Omega_0}{\left(1 + \frac{f}{f_c}\right)^2}, \tag{2}$$

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 gave a seismic moment estimate using Equation (3) (Richards, 1971).

$$\begin{aligned}
|u^S(f)| &= \frac{R^S}{2\rho\beta^3 r} \frac{M_0}{1 + \left(\frac{f}{f_c}\right)^2} \exp\left(-\frac{\pi f t^S_{s,c}}{Q^S(f)}\right), && (3)\\
\Longrightarrow M_0 &= \frac{\Omega_0 2\rho\beta^3 r}{R^S} \exp\left(\frac{\pi f t^S_{s,c}}{Q^S(f)}\right), && (4)\\
\Longrightarrow \alpha_{s,c} &= \frac{2\rho\beta^3 r_{s,c}}{R^S} \exp\left(\frac{\pi f t^S_{s,c}}{Q^S(f)}\right). && (5)
\end{aligned}$$

In Equation (3-5), we used 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 & Richards, 2002). The source-receiver distance $r_{s,c}$ and the S-wave travel time $t^S_{s,c}$ were computed from the source location and velocity model. Finally, a frequency dependent quality factor was obtained from Izgi, Eken, Gaebler, Eulenfeld, and Taymaz (2020). The moment magnitude $M_w$ is:

$$M_w = \frac{2}{3} \left( \log M_0 - 9.1 \right). \tag{6}$$

### 1.7. Identifying Mining Templates

Mining seismicity is identified by looking at the statistics of the detected events' time of the day within each family of events detected with a same template. See Figure S2. We compared the locations of mining-related seismicity identified by our analysis with the explosions (quarry blasts) reported in the Kandilli catalog (see Figure S3).

### 1.8. Comparison with the Frequency-Magnitude Distribution of the Poyraz et al. 2015 Catalog

Comparison of the frequency-magnitude distributions of the hand-made catalog in Poyraz et al. (2015) and our catalog. See Figure S4.

### 1.9. Comparison with the Frequency-Magnitude Distributions of the Past Seismicity

Comparison of the frequency-magnitude distributions of the pre-, co-, and post-Izmit seismicity with the 2012-2013 seismicity. See Figure S5.

## 2. Temporal Clustering

Extra information on temporal clustering:

- Extended temporal clustering analysis, see Figure S6.

- A transient increase in the Poisson rate of a Poisson point process does not produce temporal clustering (see Figure S7). The increase itself may have a power-law time dependence, but we argue that, in this case, it must be caused by an interaction-driven mechanism.

## References

Aki, K. (1965). Maximum likelihood estimate of b in the formula log n= a-bm and its confidence limits. *Bull. Earthq. Res. Inst., Tokyo Univ.*, *43*, 237–239.

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

Bohnhoff, M., Ickrath, M., & Dresen, G. (2016). Seismicity distribution in conjunction with spatiotemporal variations of coseismic slip and postseismic creep along the combined 1999 izmit-düzce rupture. *Tectonophysics*, *686*, 132–145.

Brune, J. N. (1970). Tectonic stress and the spectra of seismic shear waves from earthquakes. *Journal of geophysical research*, *75*(26), 4997–5009.

Bulut, F., Bohnhoff, M., Aktar, M., & Dresen, G. (2007). Characterization of aftershock-fault plane orientations of the 1999 izmit (turkey) earthquake using high-resolution aftershock locations. *Geophysical Research Letters*, *34*(20).

Ickrath, M., Bohnhoff, M., Dresen, G., Martinez-Garzon, P., Bulut, F., Kwiatek, G., & Germer, O. (2015). Detailed analysis of spatiotemporal variations of the stress field orientation along the izmit-düzce rupture in nw turkey from inversion of first-motion polarity data. *Geophysical Journal International*, *202*(3), 2120–2132.

Izgi, G., Eken, T., Gaebler, P., Eulenfeld, T., & Taymaz, T. (2020). Crustal seismic attenuation parameters in the western region of the north anatolian fault zone. *Journal of Geodynamics*, *134*, 101694.

Karabulut, H., Schmittbuhl, J., Özalaybey, S., Lengline, O., Kömeç-Mutlu, A., Durand, V., ... Bouin, M. (2011). Evolution of the seismicity in the eastern Marmara Sea a decade before and after the 17 August 1999 Izmit earthquake. *Tectonophysics*, *510*(1-2), 17–27.

Lomax, A., Michelini, A., Curtis, A., & Meyers, R. (2009). Earthquake location, direct, global-search methods. *Encyclopedia of complexity and systems science*, *5*, 2449–2473.

Lomax, A., Virieux, J., Volant, P., & Berge-Thierry, C. (2000). Probabilistic earthquake location in 3D and layered models. In *Advances in seismic event location* (pp. 101–134). Springer.

Poyraz, S. A., Teoman, M. U., Türkelli, N., Kahraman, M., Cambaz, D., Mutlu, A., ... others (2015). New constraints on micro-seismicity and stress state in the western part of the North Anatolian Fault Zone: Observations from a dense seismic array. *Tectonophysics*, *656*, 190–201.

Prieto, G. A., Parker, R., & Vernon Iii, F. (2009). A fortran 90 library for multitaper spectrum analysis. *Computers & Geosciences*, *35*(8), 1701–1710.

Richards, P. G. (1971). An elasticity theorem for heterogeneous media, with an example of body wave dispersion in the earth. *Geophysical Journal International*, *22*(5), 453–472.

Trugman, D. T., & Shearer, P. M. (2017). GrowClust: A hierarchical clustering algorithm for relative earthquake relocation, with application to the Spanish Springs and Sheldon, Nevada, earthquake sequences. *Seismological Research Letters*, *88*(2A), 379–391.

Wiemer, S., & Katsumata, K. (1999). Spatial variability of seismicity parameters in aftershock zones. *Journal of Geophysical Research: Solid Earth*, *104*(B6), 13135–

13151.

Zhu, W., & Beroza, G. C. (2019). PhaseNet: a deep-neural-network-based seismic arrival-time picking method. *Geophysical Journal International*, *216*(1), 261–273.

| Depth (top of the layer, km) | $v_P$ (km/s) | $v_S$ (km/s) |
|---|---|---|
| -2 | 2.900 | 1.670 |
| 0 | 3.000 | 1.900 |
| 1 | 5.600 | 3.150 |
| 2 | 5.700 | 3.210 |
| 3 | 5.800 | 3.260 |
| 4 | 5.900 | 3.410 |
| 5 | 5.950 | 3.420 |
| 6 | 6.050 | 3.440 |
| 8 | 6.100 | 3.480 |
| 10 | 6.150 | 3.560 |
| 12 | 6.200 | 3.590 |
| 14 | 6.250 | 3.610 |
| 15 | 6.300 | 3.630 |
| 20 | 6.400 | 3.660 |
| 22 | 6.500 | 3.780 |
| 25 | 6.700 | 3.850 |
| 32 | 8.000 | 4.650 |
| 77 | 8.045 | 4.650 |

**Table S1.** 1D velocity model due to Karabulut et al. (2011) used in this study.

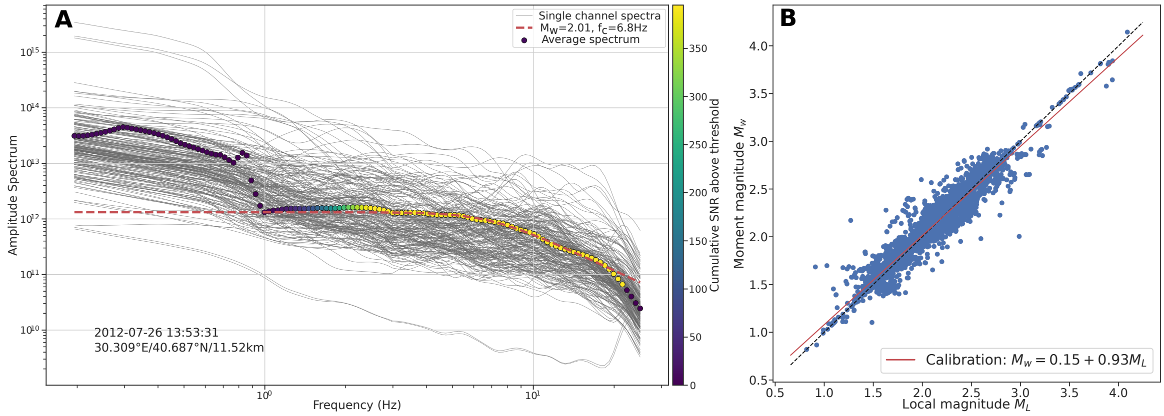


**Figure S1.** **A:** Average S-wave spectrum fitted with the Brune model (red curve). This is a weighted average of all single-channel S-wave spectra (thin grey spectra, Equation (1) of the supplementary material). The weight of each frequency bin of each channel is proportional to the excess signal-to-noise ratio (SNR) defined as $w(f) = \mathrm{SNR}(f) - \mathrm{SNR}_t(f)$, where $\mathrm{SNR}_t(f)$ 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 $M_0$. **B:** Scaling between moment magnitude $M_w$ and local magnitude $M_L$. All events with a moment magnitude estimate also have a local magnitude computed with Equation (6) in the main text. The calibration is close to identity: $M_w = 0.15 + 0.93M_L$.

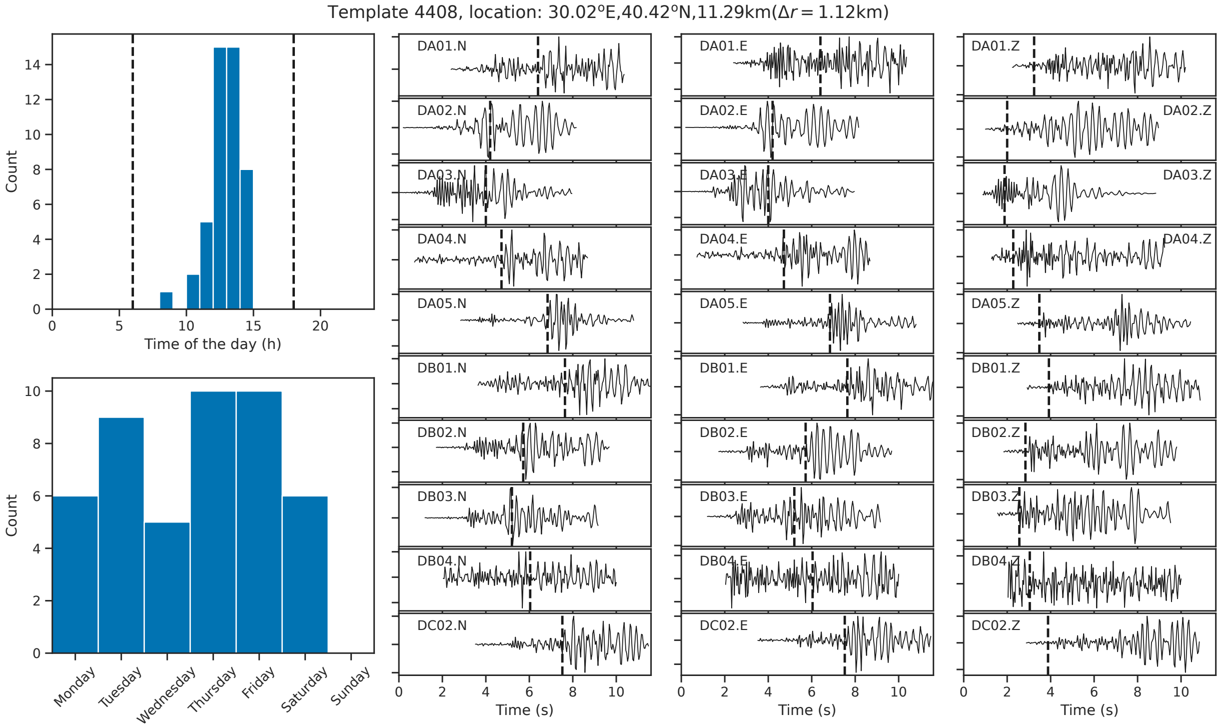


**Figure S2.** **Top left panel:** Mining-related seismicity is characterized by predominantly diurne seismicity, whereas we expect no preferred time for natural seismicity. In fact, natural seismicity shows slightly more events at night because noise is generally lower, and earthquake detection is easier. **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.

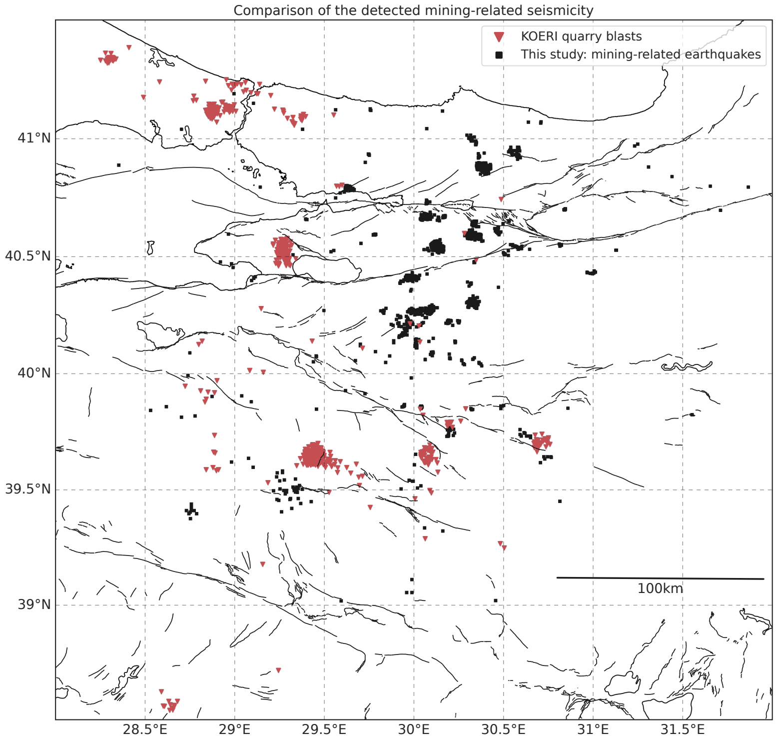


**Figure S3.** Comparison of the identified locations of mining-related seismicity in our catalog (black squares) with the reported explosions (red triangles) of the Kandilli catalog. Largest discrepancies appear beneath the DANA array. Discrepancies in the south are most likely due to the large location uncertainties ($h_{\text{max}}$ >10 km) in our catalog.

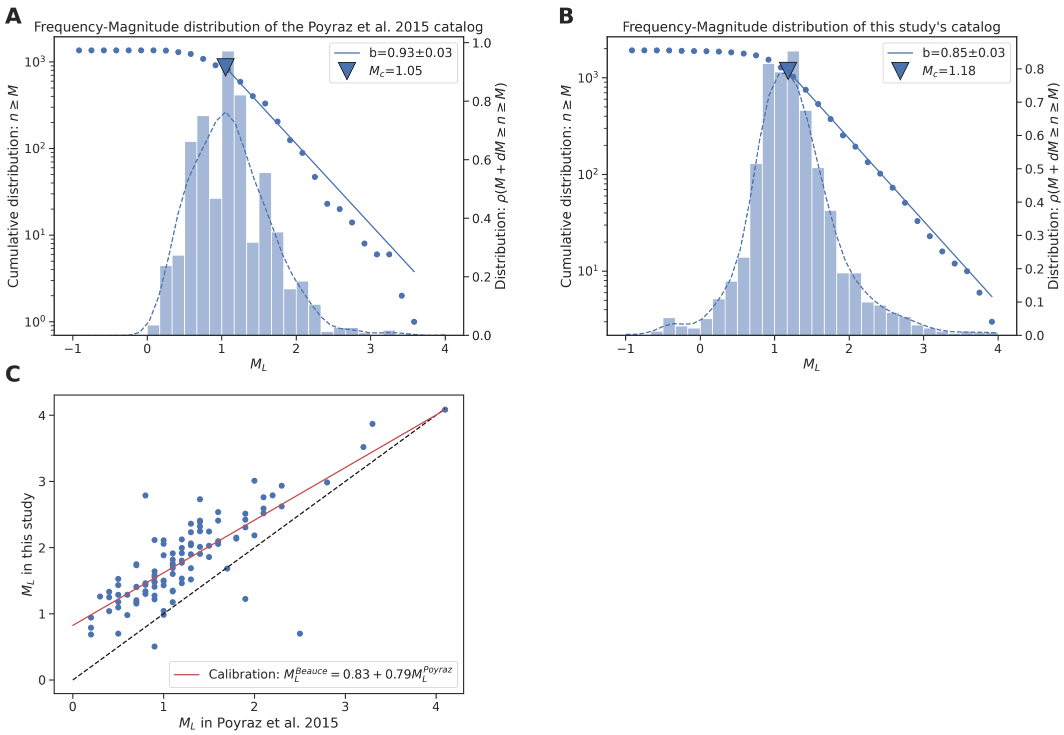


**Figure S4.** The b-value is computed with the maximum likelihood method (Aki, 1965). The magnitude of completeness $M_c$ is computed with the maximum curvature method (Wiemer & Katsumata, 1999). **A:** Frequency-magnitude distribution of the catalog published in Poyraz et al. (2015). The total number of events is 1371. **B:** Frequency-magnitude distribution of this study's catalog without mining-related seismicity. The total number of natural earthquakes for which we could estimate a magnitude is 1929. Both b-values and magnitude of completeness are similar across catalogs. **C:** Comparison of the magnitudes computed in Poyraz et al. (2015) (x-axis) and in our study (y-axis) for events that were detected and characterized in both catalogs. Our magnitudes are larger for small events and the average magnitude difference is 0.57.

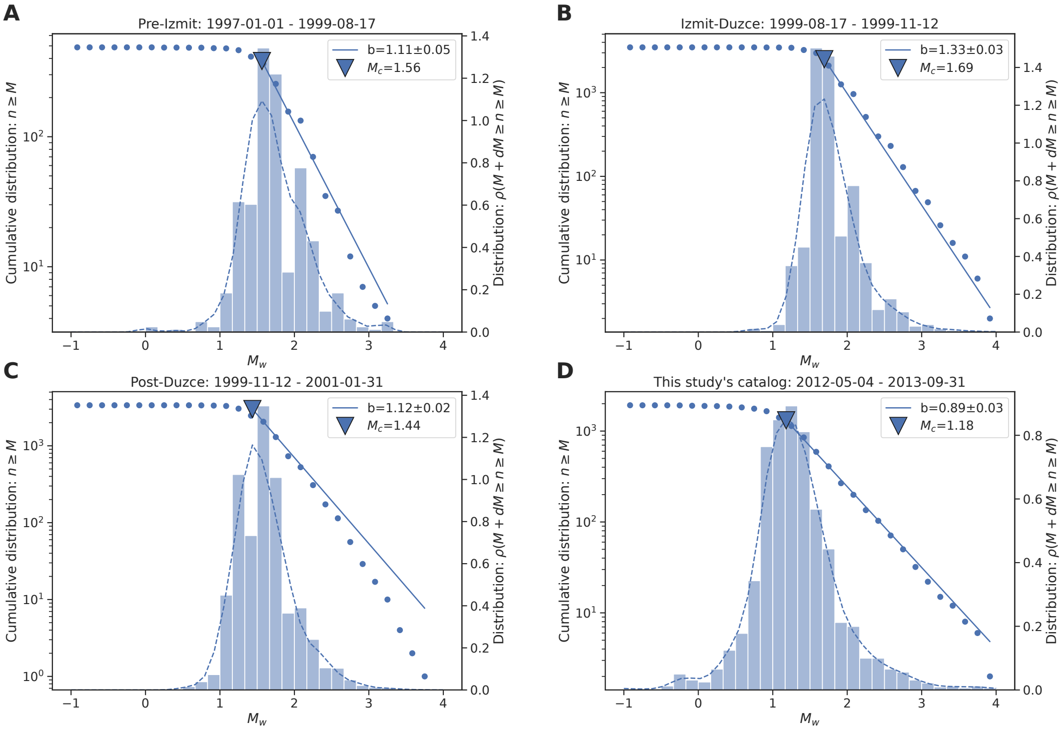


**Figure S5.** The b-value is computed with the maximum likelihood method (Aki, 1965). The magnitude of completeness $M_c$ is computed with the maximum curvature method (Wiemer & Katsumata, 1999). **A:** Frequency-magnitude distribution of the pre-Izmit seismicity (Ickrath et al., 2015; Bohnhoff et al., 2016). **B:** Frequency-magnitude distribution of the Izmit-Düzce seismicity (Bulut et al., 2007; Bohnhoff et al., 2016). **C:** Frequency-Magnitude distribution of the post-Düzce seismicity (Ickrath et al., 2015). **D:** Frequency-Magnitude distribution of this study's catalog (without mining seismcity). We used our $M_L$-$M_w$ calibration (see Figure S1B) to convert our local magnitudes to moment magnitudes.

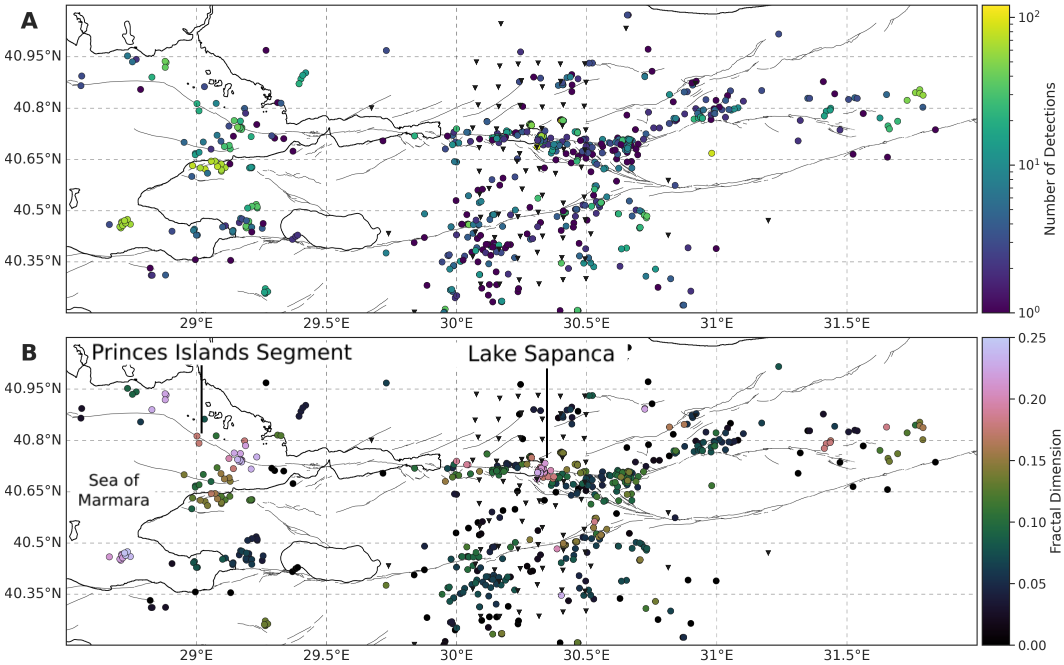


**Figure S6.** Earthquake clustering along the North Anatolian Fault Zone. **A:** Cumulative number of detections per template. **B:** Fractal dimension (as introduced in Figure 4 of the main manuscript). The eastern Sea of Marmara and Lake Sapanca show the strongest clustering along the NAF.

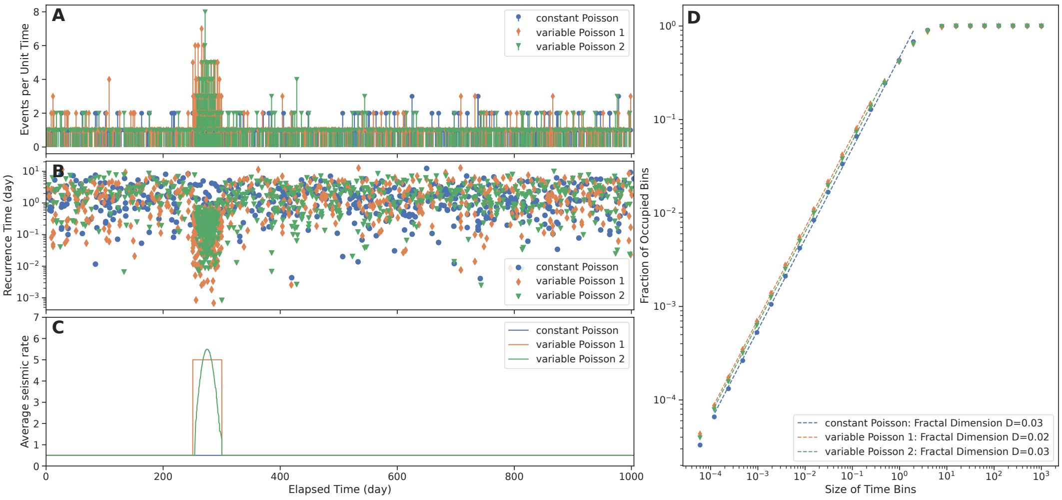


**Figure S7.** A Poisson point process cannot produce clustered seismicity, even with varying rate. **A:** Number of earthquakes per unit time. **B:** Recurrence time vs. origin time. **C:** Average number of earthquakes per unit time of the random Poisson process. **D:** Fractal analysis (see main manuscript) of the number of events per unit time. A transient increase in average seismicity rate does not reproduce a clustered seismicity with $D \neq 0$.