Martinez Garzon, P., Ben-Zion, Y., Zaliapin, I., Bohnhoff, M. (2019): Seismic clustering in the Sea of Marmara: Implications for monitoring earthquake processes. - *Tectonophysics*, *768*.

DOI: http://doi.org/10.1016/j.tecto.2019.228176

Seismic clustering in the Sea of Marmara: Implications for monitoring earthquake processes

Patricia Martínez-Garzón[1], Yehuda Ben-Zion[2], Ilya Zaliapin[3], Marco Bohnhoff[1, 4]

1.Helmholtz Centre Potsdam, GFZ German Research Centre for Geosciences, Section 4.2: Geomechanics and Scientific Drilling, Telegrafenberg, 14473 Potsdam, Germany

2. University of Southern California, Department of Earth Sciences, Los Angeles, CA 90089-0740, United States

3. University of Nevada, Reno, Department of Mathematics and Statistics, Reno, NV 89557

4. Institute of Geological Sciences, Free University of Berlin, Berlin, Germany.

Corresponding author:
Patricia Martínez-Garzón, Helmholtz Centre Potsdam, GFZ German Research Centre for Geosciences, Section 4.2: Geomechanics and Scientific Drilling, Telegrafenberg, 14473 Potsdam, Germany (**patricia@gfz-potsdam.de**)

Tectonophysics, 2019

Short title. Earthquake clusters in the Sea of Marmara

Keywords. Seismicity clusters, foreshock detection, earthquake repeaters, North Anatolian Fault, fault segmentation, earthquake dynamics

1. Introduction

Laboratory rock deformation experiments show typically foreshocks and other signals associated with preparation of large events (e.g. Goebel et al., 2013; Selvadurai et al., 2017; Renard et al., 2018). Foreshock activity has been observed before some large earthquakes such as the August 1999 M_W 7.4 Izmit earthquake along the North Anatolian fault (e.g. Bouchon et al., 2011; Ellsworth & Bulut, 2018). However, other large events including the November 1999 Mw 7.1 Düzce earthquake to the east of the Izmit event were not preceded by clear foreshocks (e.g. Wu et al., 2014). Analysis of pre-shock activity along the North Anatolian fault and other major faults has not been done systematically, in part because of the lack of high quality seismic catalogs. Refined hypocenter catalogues offering improved spatial resolution and lower magnitude of completeness allow for detailed studies of foreshocks. This is of particular importance for fault segments near densely populated regions, such as the Marmara section of the North Anatolian Fault Zone in Turkey, that are late in their seismic cycle. Below the eastern Sea of Marmara close to the Istanbul metropolitan region, several foreshocks have recently been observed preceding a M_W 4.4 event (Malin et al., 2018).

Earthquake cluster identification is essential for understanding the dynamics of seismicity. Systematic analysis of earthquake clusters in a region can provide a context for local variations of foreshocks and other informative patterns of seismicity. The number and structure of earthquake clusters can vary in space and time on a range of scales (e.g., Ben-Zion, 2008; Zaliapin & Ben-Zion, 2016a). Analytical and numerical results in a viscoelastic damage rheology models suggest that basic properties of earthquake clustering are controlled by the effective viscosity of the deforming medium (Ben-Zion & Lyakhovsky, 2006). This implies that heat flow and the presence of fluids should play an important role in determining key properties of earthquake clustering (Zaliapin & Ben-Zion, 2013b).

Repeating earthquakes representing overlapping rupture areas and similar earthquake magnitudes are also important for quantifying regional seismic hazard, and are seen as indicators for fault creep. Observations of repeating earthquakes along the Western High and Central Basin of the Sea of Marmara suggested that aseismic slip may occur at these locations (Schmittbuhl et al., 2016a; Bohnhoff et al., 2017). Earthquake repeaters are commonly identified by employing waveform cross-correlation to find highly similar seismic waveforms (e.g., .g. Poupinet et al., 1984; Nadeau & McEvilly, 2004; Peng and Ben-Zion, 2005). Recently, using analysis of earthquake clusters, it was found that fluid induced seismicity tends to display an unusually high concentration of events characterized by a relatively short distance and long time to the events initiating the clusters (Schoenball et al., 2015; Zaliapin & Ben-Zion, 2016b). Such events share some key features with the classical earthquake repeaters; however, the precise relation between these two types of events requires further exploration.

In this study we utilize a recently derived high-resolution seismicity catalog (Wollin et al., 2018) and nearest-neighbor cluster identification and classification techniques (Zaliapin & Ben-Zion, 2013a, 2013b) to analyze clusters of seismicity in the Sea of Marmara region of the North Anatolian Fault, Turkey. Our main goals are to (1) estimate the spatial distribution of mainshock and aftershock rates and use it to infer the proximity to failure on different fault segments, (2) test the potential of the nearest-neighbor cluster approach to identify areas with enhanced occurrence of earthquake repeaters, and (3) characterize areas with enhanced foreshock activity. In the next section we describe the state-of-the-art knowledge on the seismotectonics and crustal properties of the analyzed fault segments in the Sea of Marmara. The examined seismicity catalogs, the nearest-neighbor methodology and the statistical approach employed are described in Section 3. The main results of the analysis that concerns the spatial distribution of clusters and the relative proportions of foreshocks, mainshocks, and aftershocks are described in Section 4. The implications of the results are discussed in the final Section 5.

2. Fault Segmentation in the Sea of Marmara region

The North Anatolian Fault Zone (NAFZ) is a major dextral strike-slip plate-boundary that spans more than 1200 km across the northern boundary of the Anatolian Plate from east to west (Barka, 1992; Sengör, 2005; Bohnhoff et al., 2016). The eastern and central portions of the NAFZ are composed of a single well-developed fault. In the west the NAFZ splits into at least two or three main branches forming a horse-tail structure.

The Marmara section is the only portion of the NAFZ that was not activated in a M>7 earthquake during the 20th century and thus constitutes a major seismic gap (Bohnhoff et al., 2013). Given the average recurring interval on the order of 250 years and its last activation in 1766 it is considered late in its seismic cycle with high probability to generate a major earthquake in the next decades (Parsons, 2004; Murru et al., 2016). In this region, the pure strike-slip system observed along most of the NAFZ is gradually converted into a transtensional setting due to the rollback of the Hellenic subduction zone superposing a NS-extensional stress field on top of the dextral strike-slip system (e.g. Flerit et al., 2004; Le Pichon et al., 2015). The Sea of Marmara hosts two of the major fault branches of the horse-tail structure. The northern branch, here named "Marmara Section" (in accordance with Wollin et al., 2018) runs directly along the Sea of Marmara accommodating the largest deformation rates (e.g. Hergert & Heidbach, 2010; Ergintav et al., 2014). It is composed of several fault segments combined with extensional basins (Armijo et al., 1999; Le Pichon et al., 2015). We focus our analysis on six pronounced seismicity spots in the

Marmara region displaying different seismotectonic characteristics. We summarize the main features of these spots below.

2.1 Western Sea of Marmara region

The westernmost analyzed area extends along the Ganos section and the Tekirdag Basin (TB, Fig. 1). The Ganos section represents a well-defined fault segment with a relatively narrow deformation zone. It last ruptured in a M7.4 event in 1912. It is not known how far this earthquake ruptured offshore into the Terkirdag Basin. This Basin currently hosts the largest cumulative moment release of the entire Sea of Marmara region (Schmittbuhl et al., 2016b).

Directly to the east of the Tekirdag Basin, there is the Western High and Central Basin (WH, Fig. 1). There, evidence for earthquake repeaters was found, suggesting that the fault is releasing a substantial portion of its accumulated strain aseismically through creep (Schmittbuhl et al., 2016a; Bohnhoff et al., 2017). These observations have recently been evidenced from ocean-bottom geodesy (Yamamoto et al., 2019). In addition to tectonic loading, degassing in the ocean floor from underground hydrocarbon reservoirs has been suggested as additional mechanism driving the seismicity (Géli et al., 2018). According to the fault mapping, the fault zone is broader and composed of several sub segments. However, the seismicity tends to concentrate in a narrower section directly on top of the main mapped fault segment (Wollin et al., 2018). Lastly, the Central High – Kumburgaz Basin is located directly to the east of the Central Basin in the central Sea of Marmara (KB, Fig. 1). Seismicity rates from this region are comparatively lower than immediately to the East. Seafloor acoustic techniques revealed that this segment is currently fully locked (Sakic et al., 2016).

2.2 Eastern Sea of Marmara region

To the east of the Kumburgaz Basin, the Princess Island segment is of special relevance because of its vicinity to the Istanbul metropolitan region (PI, Fig. 1). This fault segment appears currently locked and accumulating strain, as evidenced by a gap in seismicity and lack of fault slip indicated by the GPS observations (Bohnhoff et al., 2013; Ergintav et al., 2014). The microseismicity within this region tends to accumulate on both edges of the fault as well as below 10 km depth where the segment tends to merge with the Cinarcik branch of the fault to a single master fault (Bohnhoff et al., 2013). In comparison to the Western High – Central Basin, a substantial portion of the micro-seismicity in this area appears to occur off-fault.

The Cinarcik Basin, constitutes a pull-apart structure bounded by the Princess Island segment to the north and the Cinarcik Fault to the south (CF, Fig. 1). The Cinarcik Fault runs approximately parallel to the coast of the Armutlu Penisula. This fault segment could have hosted the largest earthquake in the Sea of Marmara region recorded in the instrumental era (1963, M 6.3 earthquake, Bulut & Aktar, 2007) and it also could represent the western end of the rupture of the 1999 M 7.1 Izmit earthquake. The Armutlu Peninsula is a high temperature hydrothermal system which is rich in fluids. It is sensitive to earthquake triggering and some of the most vigorous Izmit aftershocks occurred here (Durand et al., 2010). In 2016, a M_W 4.4 earthquake occurred offshore near the town of Yalova. At least 18 foreshocks were identified during the 40h preceding the rupture (Malin et al., 2018). Following this earthquake, a 50-day lasting strain release was detected, indicating that some of the accommodated tectonic strain could have been released aseismically (Martínez-Garzón et al., 2019).

The fault segment in the Gulf of Gemlik is part of the southern fault branch bounding the southern Sea of Marmara shore (GG, Fig. 1). This fault segment is possibly connecting the Iznit

Lake section of the NAFZ with the southern Marmara branch towards the Bursa region. The Gemlik area has generated several M > 4 events in the last decade. This fault segment is also relevant for the seismic hazard as it is in direct vicinity to Bursa city with more than 3 million inhabitants.

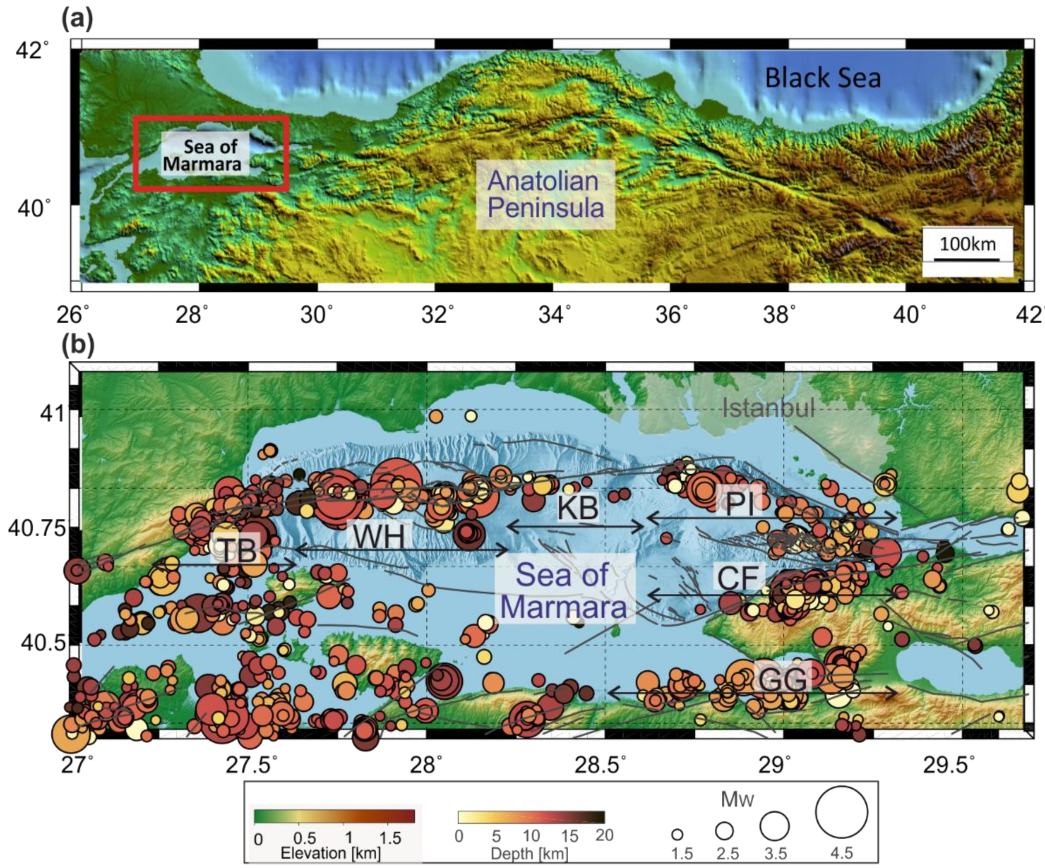

Figure 1. (a) Regional map framing the studied area (red rectangle) on the western portion of the North Anatolian Fault Zone. (b) Map of the Sea of Marmara region with epicenter locations from the Wollin et al., (2018) catalog (for the period January 2006 to March 2016 and with a magnitude of completeness $M_C^{WOLL} = 2.1$) color encoded with hypocentral depth. The different analyzed are Tekirdag Basin and Ganos section (TB), Western High – Central Basin (WH), Kumburgaz Basin (KB), Princess Islands (PI), Cinarcik Fault and Armutlu Peninsula (CF) and Gulf of Gemlik (GG).

3. Data selection and Methodology

3.1 Earthquake catalog

We analyze two seismicity catalogs of different quality containing different number of events. The main seismicity catalog is a ten-year (January 2006 – March 2016) catalog containing the seismicity from the region around the Sea of Marmara (Wollin et al., 2018). The catalog covers the region within 26.5°-30.5°E and 40°-41°N and includes 4,744 relocated events. After removing areas of suspected quarry activities, 3,974 events are identified as earthquakes (Fig. 1, see Wollin et al., 2018 for details on the quarry identification). The median area of the horizontal error ellipse for the relocated events is 2.5 km^2 and the mean vertical error is 3.8 km. The M_W magnitude range of the events in the catalog is [0 – 4.5]. Using the estimations of Wollin et al., (2018), we examine 1,625 events with magnitude above completeness $M \geq M_C^{WOLL}$=2.1.

We additionally use the KOERI seismicity catalog (http://www.koeri.boun.edu.tr/sismo/2/earthquake-catalog/, last accessed 01/03/2019) between the years 2000 and 2018 (Fig. S1). The catalog covers the same region and the provided magnitudes are in the range M_W [1, 5.7]. The events are located using the absolute location method Hypoinverse (https://earthquake.usgs.gov/research/software/#HYPOINVERSE, last accessed 01/03/2019). Horizontal and vertical uncertainties are not specified for individual events. After removing suspected quarries following Wollin et al., (2018), a total of 12,739 are selected for further analysis. We assume that for the small events in the Sea of Marmara region $M_W \approx M_L$ (Kılıç et al., 2017) and convert all the magnitudes in the catalog to M_W. We utilize the maximum curvature technique and a method based on a goodness-of-fit technique (Woessner & Wiemer, 2005) to estimate the temporal evolution of the magnitude of completeness M_C^{KOER} using a sliding window of 100 events

(Fig S2). This results in an estimation of $M_C^{KOER} = 2.1$, representative for the examined time period. Finally, a total of 8,566 earthquakes with $M \geq M_C^{KOER}$ are used.

3.2 Earthquake cluster identification

In each examined catalog, we identify seismicity clusters according to their space-time-magnitude nearest-neighbor proximity (Zaliapin et al., 2008; Zaliapin & Ben-Zion, 2013a, 2013b). This technique is selected because of its soft parametrization and robustness with respect to incompleteness, event location errors, and parameter values. The proximity η_{ij} of event j to an earlier event *i* in the space-time and magnitude domain can be defined as (Baiesi & Paczuski, 2004):

$$\eta_{ij} = \begin{cases} t_{ij}\left(r_{ij}\right)^{d} 10^{-bm_i}, & t_{ij} > 0, \\ \infty, & t_{ij} \leq 0, \end{cases} \quad (1)$$

where $t_{ij} = t_j - t_i$ [in years] and r_{ij} [in kilometers] are the temporal and spatial distances between the earthquakes *i* and *j*, respectively, *d* is the fractal dimension of the hypocenter (or epicenter) distribution, *b* is the *b*-value of the Gutenberg-Richter relation and m_i is the magnitude of the (earlier) event *i*. The scalar proximity η_{ij} between events can be expressed as the product of its temporal and spatial components normalized by the magnitude of the earlier event *i*:

$$\eta_{ij} = T_{ij} \cdot R_{ij} \quad (2)$$

$$T_{ij} = t_{ij} 10^{-qbm_i}, R_{ij} = \left(r_{ij}\right)^{d} 10^{-(1-q)bm_i}, 0 < q < 1. \quad (3)$$

We fix $q = 0.5$, providing equal weights to the temporal and spatial distances. To estimate the spatial distance between events we used epicentral locations, since the vertical location accuracy from these catalogs is lower than the horizontal. The parameter used values are $b = 1$ and $d = 1$, representing the epicentral distribution of seismicity as approximately linear, in agreement with the seismicity distribution from fault structures. This method for identifying seismicity clusters is generally not sensitive to moderate variations in these parameters (see Zaliapin and Ben-Zion, 2013a for details), and equivalent results are obtained using for example *b*=1.2.

We denote by η_j the shortest of the proximities between event *j* and all earlier events. The event at which this minimal value is attained is called the parent of *j*. The distribution of the nearest-neighbor proximities η_j in observed catalogs is generally bimodal (e.g. Fig 2, Fig S3, Zaliapin et al., 2008; Zaliapin & Ben-Zion, 2013a, 2016a). The long-proximity mode (representing rescaled times and distances larger that the estimated separation threshold between the two seismicity modes) roughly corresponds to *background* Poissonian-like seismicity, while the short-proximity mode is associated with *clustered* earthquakes (i.e. foreshocks and aftershocks).

Individual clusters are formed by earthquakes that are connected by *short* proximity links. Each earthquake connected to the parent by a long link is considered a *background* event and starts a new cluster. A *single* is a cluster that consists of one background event with no associated foreshocks or aftershocks, while the multiple-event clusters are called *families*. The largest event in each cluster is called *mainshock*; all events within the cluster and prior to/after the mainshock are called *fore/after-shocks* (see Fig. 6 of Zaliapin and Ben-Zion, 2013a).

In the Sea of Marmara region, the seismicity rates vary among the fault segments, and the station coverage is not uniform since many of the fault segments run offshore. Therefore, the seismicity can be represented as a non-homogeneous Poisson process in space. To account for this effect in our cluster identification, we calculate the nearest-neighbor proximity η_j for each event using the entire catalog, and implement a space-dependent threshold for each event that separates short and long proximities in identifying individual earthquake clusters. The results of this analysis are illustrated in Fig. 2. For each event, we start with a circular region of 2 km surrounding the event and we iteratively increase the radius taking intervals of 1 km until the number of events contained in the region is larger than 5% of the entire catalog (81 and 477 events for the Wollin and KOERI catalog, respectively). The distribution of rescaled times and distances for these events is used to estimate the separation threshold between short and long proximities from the initial event in the center of the circular region. Using a distribution with a larger number of events (i.e. more than 5% of the catalog) to estimate the event-based threshold results in smoothing the threshold variations (therefore, decreasing the detection of non-homogeneous distributions). Conversely, decreasing the number of events contained in the distribution to calculate the threshold allows detecting more effectively non-homogeneities in the distribution of rescaled times and distances. The utilized proportion of 5% was selected as optimal to effectively detect changes in the distributions of the analyzed areas, but the mainresults are not affected when varying the proportion to within +-10%.

To separate the *short* and *long* proximity modes of the seismicity in each window, we fit a Gaussian mixture model with two modes to the logarithmic proximities $\log_{10} \eta_j$ (Zaliapin & Ben-Zion, 2016); the threshold is defined as the point of equal density of the two estimated modes. Note that the proportion of events used to estimate the event-based threshold between short and long proximity modes does not affect or limit the number of events contained in each individual cluster. Figs. 2a,b show the distribution of the nearest-neighbor proximity values, its rescaled components, and the estimated space-dependent threshold. For most of the examined events, there is a clear separation between the background and cluster mode, which is best seen in the 2D plot Fig. 2a. The threshold values are concentrated around the value – 4; the threshold distribution is left-skewed with some extreme values as low as –7 and as high as – 3.8. Therefore, although it is more correct to account for the effect of potential non-homogeneities in the distribution, this effect is not large in our catalog and the main results are preserved using also a homogeneous threshold.

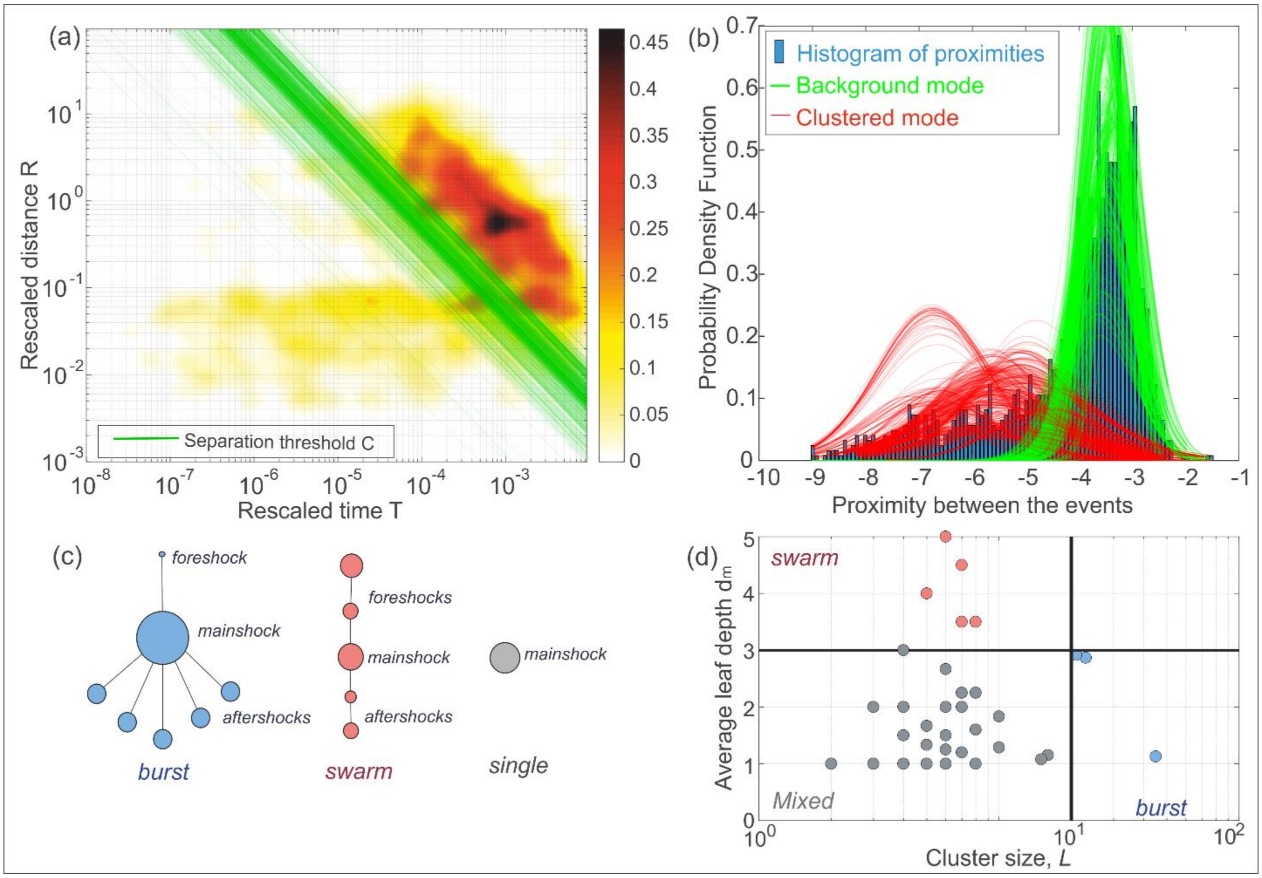

Figure 2. Cluster identification using Wollin et al. (2018) seismicity catalog. (a) Joint distribution of the rescaled time and space components (T, R) of the nearest neighbor proximity. Green lines show the separation thresholds obtained for different spatial portions of the catalog as discussed in Sect. 3.2. (b) Histogram of the nearest neighbor proximities (blue bars) showing a bimodal distribution of background and clustered events. The green and red lines show the result of fitting a Gaussian mixture model that identifies the background and cluster modes, respectively. (c) Conceptual sketch showing the topological structure of typical burst-like and swarm-like clusters as well as a single. The size of the circles is proportional to event magnitude. (d) Average leaf depth d_m vs cluster size L. This diagram guides in identifying swarm-like and burst-like clusters.

3.3 Generalized linear regression models

To quantify differences in earthquake cluster properties among the analyzed fault segments we use generalized linear models, which are an extension of ordinary regression that allows one to work with non-normal data (Agresti, 2018). We examine three cluster statistics: The proportion of earthquake families among the identified clusters (Section 4.1); the proportion of mainshocks that are preceded by at least one foreshock (Section 4.3); and the duration of the foreshock sequences -- the time between the first event in the sequence and the mainshock (Section 4.3). The latter analysis is only performed in two regions -- the western and eastern Sea of Marmara. In all experiments, the examined statistic is used as the model response and the region (as a categorical variable) is a single model predictor.

The first two statistics are analyzed using the logistic regression model. Specifically, each mainshock i is associated with a Bernoulli random variable Y_i that equals 1 if the mainshock has at least one offspring (for the first model) or at least one foreshock (for the second model), and 0 otherwise. Furthermore, each mainshock is associated with region indicator (dummy) variables $x_1,\ldots,x_p$ such that $x_j = 1$ if the examined mainshock belongs to region j, and $x_j = 0$ otherwise. The model fits the values $\pi(\boldsymbol{x}) = P(Y=1)$ as a function of the region indicators:

$$\pi(x) = \frac{exp(\alpha+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_p x_p)}{1+exp(\alpha+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_p x_p)}, \quad (4)$$

where α is the model intercept and $\beta = (\beta_1, \ldots, \beta_p)$ are region coefficients. To avoid redundancy, β_1 is set to 0. The null hypothesis H_0:"The probability of success, P(Y=1), is the same in all regions" corresponds to $\beta_i = 0$ for all i. The model is equivalent to a linear expression for the logarithmic odds of success:

$$ln\left(\frac{P(Y=1|x)}{P(Y=0|x)}\right) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p.$$

The large-sample distribution of the estimated coefficients in this generalized linear model is Normal (Agresti, 2018), which facilitate inference. The model also allows making inference about the equality of proportions between two selected regions. Specifically, the logarithm of the conditional odds ratio between two regions equals the difference between the estimated coefficients:

$$ln\left(\frac{P(Y=1|x_i=1)}{P(Y=0|x_i=1)}\frac{P(Y=0|x_j=1)}{P(Y=1|x_j=1)}\right) = \beta_i - \beta_j, \quad (5)$$

with zero difference corresponding to the null hypothesis: H_0:"The probability of success is the same in the two examined regions". Similarly, the third model fits the average sequence duration $\mu(x)$ as a function of region indicator:

$$\mu(x) = \alpha + \beta_1 x_1 + \beta_2 x_2. \quad (6)$$

The data for the different regions as well as the estimated coefficients in the three models are provided in Table S1.

4. Results

In the following, we present results obtained for the higher quality catalog from Wollin et al. (2018). A comparison with results for the KOERI catalog (documented in supplementary materials) is provided in the discussion.

4.1 Spatio-temporal properties of mainshocks and aftershocks

The nearest neighbor proximities show a bimodal distribution emphasizing the background and clustered seismicity modes (Figs 2a, b). According to the respective cluster identification (Sec. 3.2), 70% of the events in this catalog are classified as *background* seismicity (i.e., 70% of earthquakes are mainshocks). The highest background rates are observed in the Tekirdag Basin and in the Cinarcik Fault– northern portion of the Armutlu Peninsula (Fig 3a). The remaining 30% of the dataset forms *clustered* seismicity, out of which 24% are *aftershocks* and 6% are *foreshocks*. The largest concentration of aftershocks appears around the Western High, coinciding with the location of the three largest events reported in the catalog (M_W 4.5, 4.5, 4.3), as well as in the Cinarcik Fault (Fig 3b).

Each cluster identified as discussed in Sect. 3.2 is represented as a tree graph. We use the following statistics of individual clusters (families): (i) cluster size L is the number of events in a cluster; L=1 for singles, and L>1 for families, and (ii) topological leaf depth d_m, which is the average distance from the cluster leaves to the root. When the earthquake *families* are sufficiently large (i.e., the family size L > 10), two end-member family types have been previously identified (Fig 2c). *Burst-like* sequences are characterized by a small value of d_m; they are mostly comprised of conventional mainshock-aftershock sequences. Such families are typical for regions of relatively low heat flow and reduced fluid content. S*warm-like* sequences, which are chains of events of similar magnitude with no clear mainshocks, are characterized by a larger average leaf depth d_m and are typical of regions with relatively high heat flow and/or high fluid content. Based on the distribution of average leaf depths d_m and size L of our clusters, we identified some burst-like and swarm-like clusters in our catalog (Fig 2d). The three *burst-like* sequences are located on the Western High and they are related to the largest mainshocks contained in the catalog (Fig 3a). Several swarm-like clusters were also identified, concentrating in the Tekirdag Basin and the Cinarcik Fault – Armutlu Peninsula (Fig 3a). This suggests that these two regions could have larger heat flow and/or presence of fluids than their surroundings.

We divide the study region into six areas containing one or more different fault segments and calculate the proportion of mainshocks with associated *family* (e.g. foreshocks and/or aftershocks) with respect to the total population of background events, which is the proportion of families with respect to the total number of families and singles. The selected areas are (1) Ganos Fault - Tekirdag Basin, (2) Western High-Central Basin, (3) Kumburgaz Basin, (4) Princess Islands segment, (5) Cinarcik Fault - Armutlu Peninsula, (6) Gulf of Gemlik. Interestingly, clear differences are visible in the proportion of families within each area. With about 25% of the background events having family, the Western High-Central Basin and Cinarcik Fault contain the largest proportion of families in the Sea of Marmara region (Fig 3c). Different proportion of *family* mainshocks and *singles* among different regions could reflect either larger stress transfer (for example due to the occurrence of larger earthquake magnitudes) or, alternatively, it could reflect the proximity to failure of each of the regions.

We fit a logistic regression model to the six regions of the Sea of Marmara to check how significant the differences between family proportions within examined regions are (Section 3.3). Selecting the Tekirdag Basin (TB) region as a reference with $\beta_1 = 0$, larger values of the coefficients $\beta_2 = 0.67$ (WH) and $\beta_5 = 0.75$ (CF) agree with the larger family proportions found in these two regions (Table S1). These are also the only two coefficients with p-value < 0.05, thus indicating that they are statistically different from the reference region TB. A complete pairwise comparison of the estimated proportions, based on the odds ratio estimation of Eq. (5) and Fisher exact test in a 2x2 table is illustrated in Table 1 (elements above diagonal). Recall that the null hypothesis H_0: “The probabilities of success are the same in both regions” corresponds to the odds ratio equal to unity. The odds ratio above (below) one suggests that the probability of success is higher (lower) in the first of the two examined regions. The results suggest two groups of regions having statistically different proportion of mainshocks with families: regions WH and CF show a higher proportion of families (24.00% and 25.42%, respectively), while the other four regions have a smaller proportion of approximately 13%.

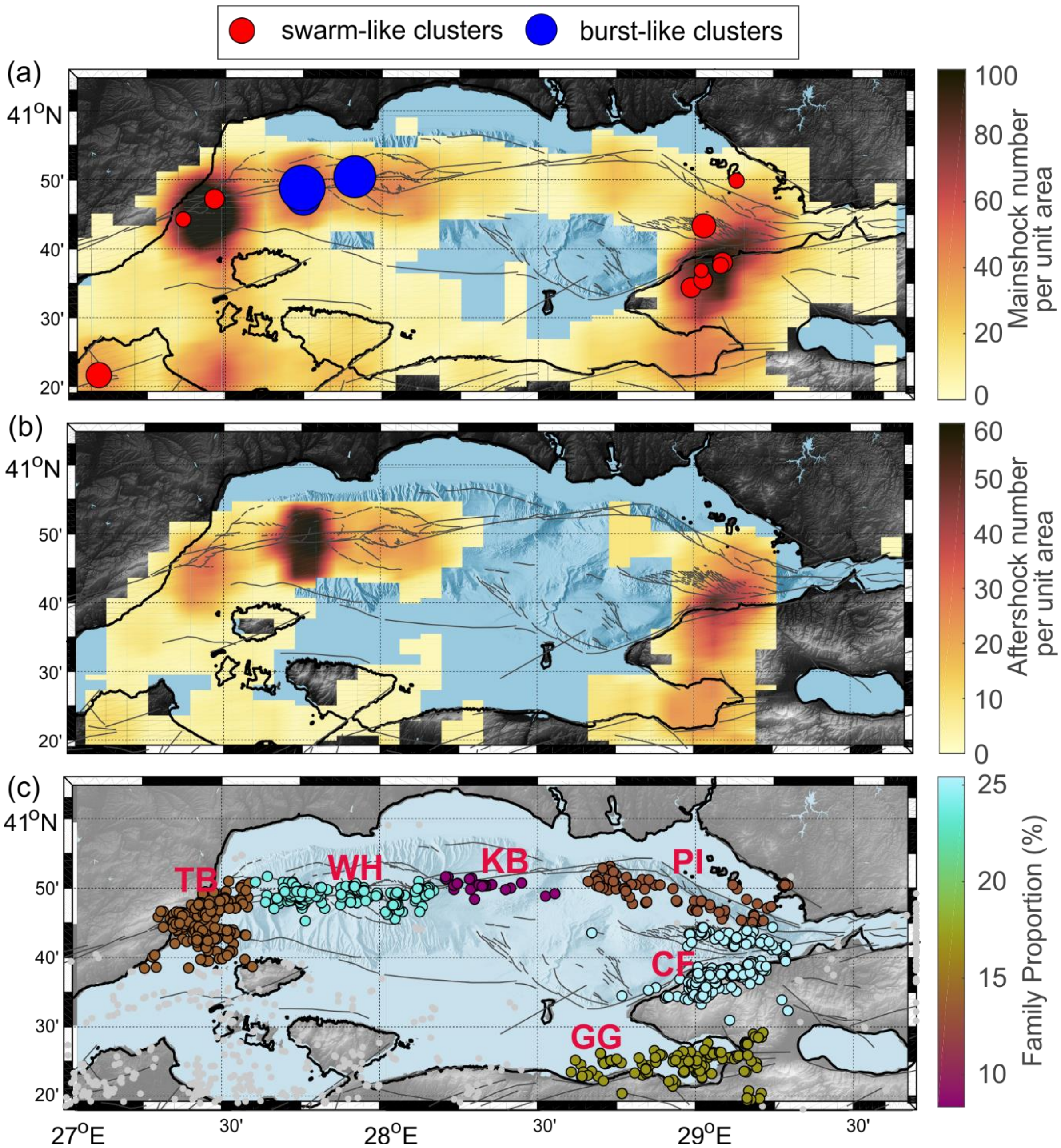

Figure 3. (a) Number of mainshocks per unit area. Blue and red color circles represent the mainshocks from identified bursts and swarms, respectively, according to the threshold displayed in Figure 2d. Size of the circle is encoded with mainshock magnitude. (b) Number of aftershocks per unit area. (c) Proportion of families among different sections of the Sea of Marmara. The initials beside each fault region represent the following: (TB) Ganos section-Tekirdag Basin, (WH)

Western High-Central Basin, (KB) Kumburgaz Basin, (PI) Princess Islands segment, (CF) Cinarcik Fault - Armutlu Peninsula, (GG) Gulf of Gemlik.

	TB	WH	KB	PI	CF	GG
TB		0.51 (0.03)	1.77 (0.74)	1.06 (1)	0.47 (<0.01)	0.79 (0.50)
WH	0.48 (0.15)		3.45 (0.11)	2.06 (0.09)	0.93 (0.79)	1.54 (0.20)
KB	-	-		0.6 (0.72)	0.27 (0.07)	0.45 (0.37)
PI	1.01 (1.00)	2.09 (0.39)	-		0.45 (0.04)	0.75 (0.53)
CF	0.39 (0.03)	0.80 (0.70)	-	0.38 (0.14)		1.67 (0.11)
GG	0.99 (1.00)	2.07 (0.21)	-	0.99 (1.00)	2.57 (0.08)	

Table 1: Pairwise regional comparison of the proportion of mainshocks with offspring (top part of the table) and mainshocks with foreshocks (bottom part of the table). Each cell shows the estimated odds ratio and the respective *p*-value (in parentheses), according to the Fisher exact test. Cells with *p*-value below 0.1 are shown in black, the rest in gray. The odds ratio above (below) one suggests that the probability of success is higher (lower) in the region indicated in the first column.

4.2 Nearest neighbor distributions to identify earthquake repeaters

The distribution of the nearest-neighbor proximity differs substantially among the six analyzed sections (Fig. 4). In the Terkirdag and Kumburgaz Basins, nearly no clustered seismicity is observed (Figs, 4a 4c), implying that the majority of the seismicity correspond to background. The Western High-Central Basin display an unusual clustered mode with lower rescaled distances *R* than any of the other regions (Fig 4b), suggesting that the events tend to occur closer to each other than in other fault regions. Both Western High-Central Basin and the Armutlu Peninsula display the largest density in the clustered mode area (Figs 4b, 4e).

The Western High – Central Basin and Gulf of Gemlik areas have a larger proportion of events displaying relatively low rescaled distance *R* and high rescaled time *T* (Figs 4b, 4f). We refer to events with these features as "*earthquake repeaters*", since they occur after a long time with respect to its parent event but in a very similar location (Schoenball et al., 2015; Zaliapin & Ben-Zion, 2016b). This purely statistical definition is somewhat different from that of classical *earthquake repeaters*, which are events whose source locations overlap and recurrence statistics are relatively periodic (e.g. Poupinet et al., 1984; Nadeau & McEvilly, 2004). We calculated for each event the ratio T/R, where increased value corresponds to earthquake repeaters. Individual earthquakes with the largest T/R (Fig 5a) as well as regions with the largest average T/R (Fig 5b) are located at both sides of the Central Basin- Western High. Interestingly, these regions have documented traditional earthquake repeater sequences (Schmittbuhl et al., 2016a; Bohnhoff et al., 2017). This indicates that the nearest-neighbor analysis could provide insight on classical repeater-prone regions. Furthermore, the analysis indicates that the Gulf of Gemlik region also display large T/R. This suggests that classical earthquake repeaters may also be found in that region.

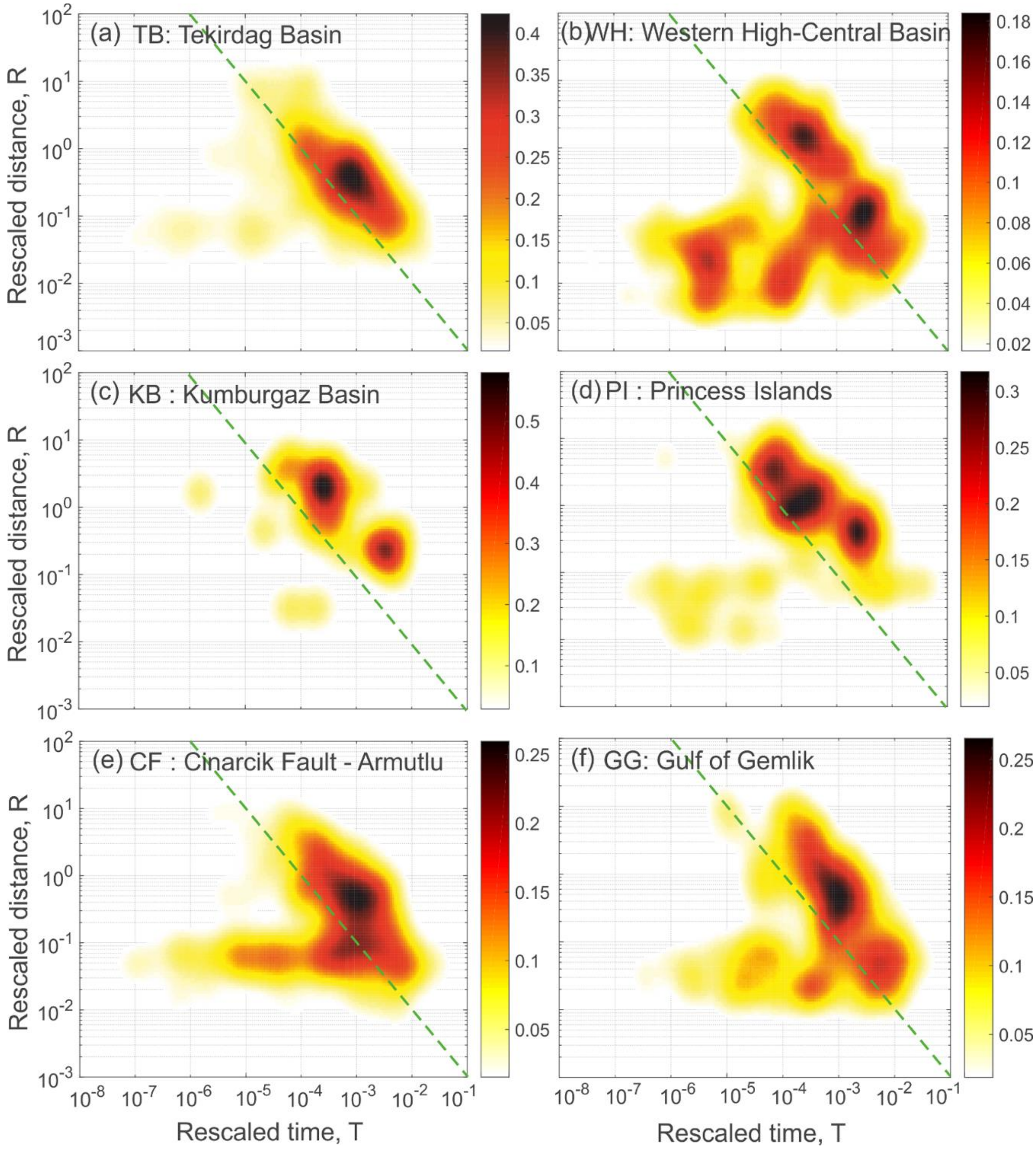

Figure 4. Joint distribution of the rescaled components (T, R) of the nearest-neighbor proximity in the six examined regions. (a) Ganos section – Tekirdag Basin, (b) Western High – Central Basin, (c) Kumburgaz Basin, (d) Princess Island segment, (e) Cinarcik Fault – Armutlu Peninsula, (f) Gulf of Gemlik. Note that a green line corresponding to $\eta = -4$ has been added to all panels only to facilitate visual comparison.

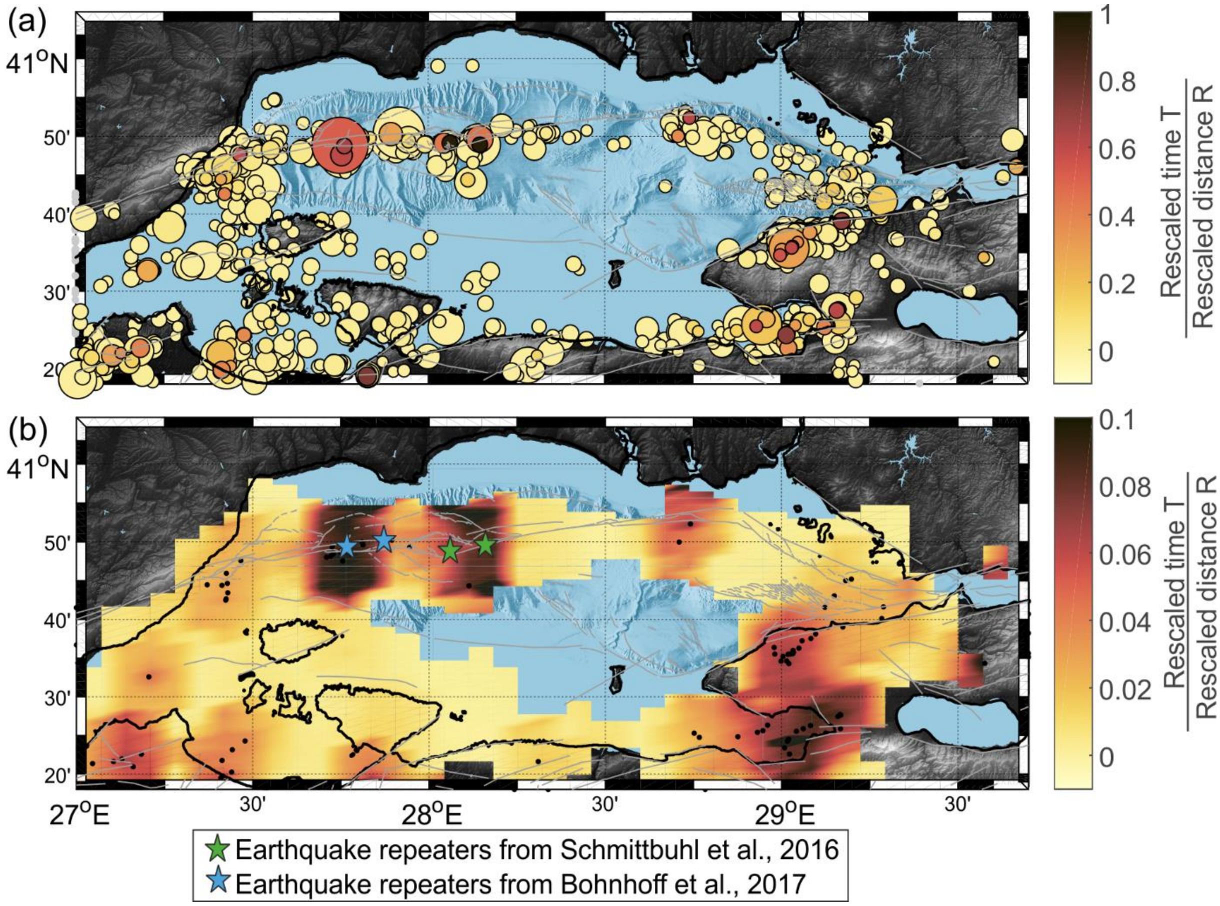

Figure 5. (a) Map of the background seismicity in the Sea of Marmara using the catalog of Wollin et al. (2018). Colors represent the ratio T/R between rescaled time and distance from the parent event . (b) Similar map as in (a) with color code corresponding to the ratio T/R smoothed by a kernel density estimation.

4.3 Characterization of foreshock properties

The cluster analysis indicates that 6% of the events in the examined catalog are *foreshocks*. The density of foreshocks peaks around the Cinarcik Fault - Armutlu Peninsula (Fig 6a), where several foreshocks were detected prior to the 2016 M_W 4.4 earthquake (Malin et al., 2018). To provide a better context for such studies, we calculate the proportion of mainshocks that have at least one foreshock in the six analyzed regions. The CF has the highest proportion (12%) of such mainshocks (Fig 6b). It is followed by the WH with 8%. The smallest proportion (0%) is found in the KB, where no foreshocks are identified (Fig 6b). However, the number of events in this region is rather small.

We use the generalized logistic regression to quantify significance of the differences in the proportions of mainshocks preceded by at least one foreshock among the different segments in the Sea of Marmara (Tables S1). Using the TB region as reference, the highest proportion of mainshocks with foreshocks is found for the CF, the respective coefficient β_5 = 0.95 is significantly different from the reference value $\beta_1 = 0$ (Figure 6b). A complete pairwise comparison of the estimated proportions, analogous to that performed in section 4.1 is illustrated in Table 1 (values below diagonal). These pairwise comparisons do not include KB, which shows no foreshocks. The most significant differences are seen when comparing the CF to the TB and the GG, with the CF region having significantly higher proportion of foreshocks (10.73%) than the other two regions (4%). The differences between the other segments do not appear statistically significant, which in some cases might be due to small sample sizes.

We also quantify the median time between the first foreshock of the sequence and the corresponding mainshock. Interestingly, the duration of the foreshock sequences appears to be different between the east and west of the Sea of Marmara. The east region (including Princess Islands, Cinarcik Fault and Gulf of Gemlik) displays a median time between first foreshock and mainshock of 6.83 hours (Fig 6c, Table S1). In contrast, the west (including Ganos Fault, Tekirdag Basin, Western High, Central Basin, Kumburgaz Basin) has an overall foreshock duration of only 2.90 hours (Fig 6c, Table S1). In the next step, we fit a generalized linear model to the data from the duration of the foreshock sequences of the western and eastern Sea of Marmara. The large β_2 coefficient indicates that the available sequences from eastern Marmara have a larger duration of the foreshock sequence. However, the differences suggested by the data are not significant (p-value of 0.1), which might be due to small sample sizes. Further data are needed to statistically confirm the two-fold increase in the foreshock sequence duration observed in the examined regions.

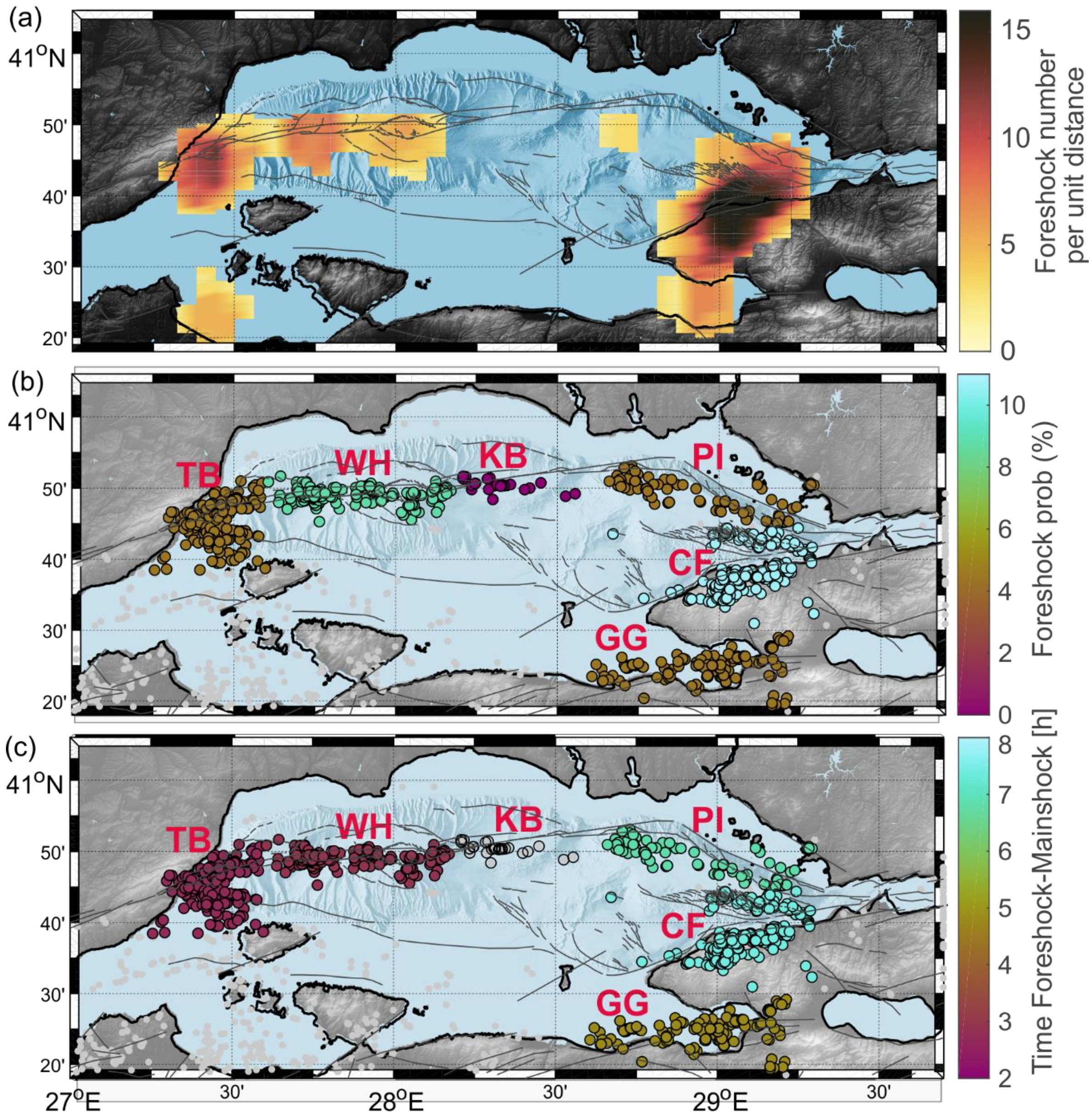

Figure 6. (a) Number of foreshocks per unit of area. (b) Proportion of mainshocks that have at least one associated foreshock or more. (c) Median duration of foreshock sequences. Initials beside the section are defined in Fig 3d.

5. Discussion

5.1 Consistency of results between the catalogs

We analyzed two earthquake catalogs of varying quality in the Sea of Marmara region to investigate the consistency of the clustering features between the catalogs. In Section 4 we focused on the Wollin et al. (2018) catalog of larger quality. Here we first compare these results with those obtained using the KOERI catalog.

The obtained proportion of background seismicity and aftershocks as well as their spatial distributions are very similar in the KOERI catalog (Fig S3, Fig S4). Setting an analogous threshold as in the case of the Wollin et al. catalog to identify burst and swarm-like clusters, shows that the epicentral locations of burst and swarms are also consistent.

Analysis of the parameter T/R using the KOERI catalog shows also similar features; however, the differences between the segments are less clear (Fig S1, S5). Using the T/R parameter, the same maximum around the Western High – Central Basin is found, but the maximum around the Gulf of Gemlik is more diffuse and it covers the entire eastern Sea of Marmara (Fig S1). This is interpreted as a signature of the comparatively lower quality of the catalog, which could smear the results that appear sharper with the Wollin et al. catalog.

The proportion of 6% of foreshocks is also found by using the KOERI catalog. However, the foreshock statistics are not consistent within the two catalogs (Fig S6). These features may be more sensitive to various factors such as poor epicentral locations or inaccurate magnitude estimation and they may only be recovered with higher-quality seismicity catalogs.

5.2 Background, aftershock and foreshock rates with respect to other faults

About 70% of the seismicity catalog represents background seismicity (i.e. *mainshocks*), while only 24% and 6% represent aftershock and foreshock sequences, respectively. Interestingly, a similar proportion of foreshocks to the one found here (6%) was reported in early studies of the seismicity catalog in southern California (Jones, 1985), as well as at global scale (Zaliapin & Ben-Zion, 2016a). A closer look indicates that areas of lower and higher heat flow tend to display lower and higher foreshock rates, respectively (Zaliapin & Ben-Zion, 2013b). The high proportion of background seismicity is comparable to that found in the San Jacinto strike-slip fault in California (Zaliapin & Ben-Zion, 2016b), but lower than the background proportion found at global scale (Zaliapin & Ben-Zion, 2016a). Similarly, the encountered proportion of aftershocks (24%) is also lower than at global scale (41%, see Zaliapin & Ben-Zion, 2016a). The reduced proportion of aftershocks in the Sea of Marmara could be partially due to the small range of magnitudes (M_W [2.1 4.5]) included in the analyzed catalog, or could reflect an incompleteness of detected events in the lower magnitude range.

The majority of the observed *swarm*-like clusters tend to concentrate around the Cinarcik Fault - Armutlu Peninsula. This area is known to have relatively higher heat flow than the surroundings as well as enhanced presence of fluids (e.g. Kinscher et al., 2013). These factors tend to reduce the effective viscosity of the crust and were found to promote the existence of swarms in southern California (Zaliapin & Ben-Zion, 2013b) and worldwide (Zaliapin & Ben-Zion, 2016a). Therefore, although a detailed map of heat in the Sea of Marmara is not available, the obtained results are in agreement with similar findings worldwide.

5.3 Foreshock distribution and potential for monitoring earthquake nucleation

The Sea of Marmara region is considered as a seismic gap that can rupture in a M>7 earthquake during this century (Bohnhoff et al., 2013; Ergintav et al., 2014). Monitoring and identifying potential earthquake preparation processes that may give some information about the increased probability of occurrence for a larger earthquake remains of uttermost importance, especially in the light of the adjacent Istanbul Metropolitan area. The occurrence of foreshocks preceding a mainshock is of importance, because of their potential use as an alert of the activation of the corresponding region. However, the main challenge in operational analysis of premonitory foreshocks is that the very definition of this event type is conditioned on the occurrence of a later mainshock. There are no criteria to classify an earthquake as a foreshock prior to the mainshock occurrence.

Our results show that the largest proportions of mainshocks preceded by foreshock activity occur on the Cinarcik Basin – Armutlu Peninsula and the western high – Central Basin area. Together with pre-seismic slip, foreshocks are one the few indications of an upcoming larger earthquake. The results of this study provide information on the overall likelihood of foreshocks in different fault sections in the Sea of Marmara region. In addition, the duration of the foreshock sequences, and consequently, the available time to detect and identify the preparation process is observed to be larger in the eastern than in the western fault segments.

5.4 Are nearest neighbor distributions useful to identify earthquake repeaters?

Characteristic repeating earthquakes rupturing the same fault patch over quasi-periodic time intervals can improve detection of aseismic slip sources as well as an estimation of the creeping rates (e.g. Poupinet et al., 1984; Nadeau & McEvilly, 2004). In the Sea of Marmara, classical earthquake repeater sequences have been identified in the Western High and the Central Basin (Schmittbuhl et al., 2016a; Bohnhoff et al., 2017). Here, we have utilized a nearest neighbor approach to search for areas where the rescaled distance and time of a given event to its parent is anomalously small and large, respectively. The areas displaying relatively larger T/R coincide well with previously mapped locations of earthquake repeaters in the Sea of Marmara. Therefore, the nearest neighbor technique appears to provide simple indications of the areas where earthquake repeaters in combination with aseismic slip could be present. Furthermore, the analysis reported additional indication for earthquake repeaters in the Gulf of Gemlik. A search for classical earthquake repeaters in this area should to be done in a future study.

5.5 Proportion of earthquake families and proximity of a fault segment to failure

The proportion of earthquake families within a population allows quantifying the role of earthquake interaction within a certain fault segment. Since the stress transfer from the occurrence of $M_W < 4.5$ earthquakes is in the order of few kPa in the surrounding mainshock area (e.g. Rothert & Shapiro, 2007), this small stress transfer should be sufficient to bring the crust to failure and trigger aftershocks. Therefore, larger proportion of earthquake families may be typical of areas where the crust is closer to failure, or alternatively, where the stress transfer is larger.

The Western High-Central Basin and the Armutlu Penisula display the largest proportion of earthquake families within the Sea of Marmara region (Fig 3c), suggesting that they are more susceptible to earthquake triggering. This possibly indicates that these segments are closer to failure than the other analyzed segments in the Sea of Marmara. Other small perturbations of the same magnitude as the earthquake interaction may also trigger seismicity in these segments. For example, the passing of surface waves from a large regional or teleseismic event have been observed to trigger seismicity in fault segments closer to failure (e.g. Aiken et al., 2015). Indeed, the largest triggering of aftershocks after the M_W 7.1 1999 Izmit earthquake occurred in the Armutlu Peninsula (Durand et al., 2010). It is therefore expected that these two regions may also be susceptible to triggering from other small stress perturbations, such as tidal oscillations or seasonal changes in the level of the water mass.

6. Conclusions

We analyzed clusters of seismicity in the Sea of Marmara region, NW Turkey, utilizing a high-quality relocated hypocenter catalog and the nearest neighbor earthquake distance approach. The main conclusions of our analysis are as follows:

(1) About 70% and 24% of the hypocenter catalog are identified as mainshocks and aftershocks, respectively. Largest background rates are observed around the Tekirdag Basin and the Cinarcik Fault. The largest density of aftershocks is observed around the Western High, coinciding with the location of the largest events in the catalog.

(2) About 6% of the events in the hypocenter catalog are identified as foreshocks. The largest proportion of foreshocks is found in the Cinarcik Fault and Armutlu Peninsula, a region known to have elevated heat flow and hydrothermal systems.

(3) Significant differences in selected cluster statistics are observed among the examined fault segments. The technique also successfully identifies regions where earthquake repeaters have been observed, and suggests additional repeaters in the Gulf of Gemlik.

(4) The Western High and Cinarcik Fault – Armutlu Peninsula display the largest proportion of earthquake families, which might be an indicator that these segments are closer to failure. This suggest a higher susceptibility of earthquake triggering from teleseismic earthquakes in these two regions.

Acknowledgements

We thank Robert Shcherbakov, an anonymous reviewer and Editor Kelin Wang for thoughtful and constructive comments. PMG acknowledges funding from the Helmholtz Association in the frame of the Helmholtz Young Investigators Group SAIDAN (VH-NG-1323). YBZ and IZ acknowledge support from the Earthquake Hazards Program of the USGS (grants G17AP00086 and G17AP00087) and the National Science Foundation (grants EAR-1723033 and EAR-1722561).

References

Agresti, A. (2018). *An Introduction to Categorical Data Analysis*. John Wiley & Sons.

Aiken, C., Zimmerman, J. P., Peng, Z., & Walter, J. I. (2015). Triggered Seismic Events along the Eastern Denali Fault in Northwest Canada Following the 2012 Mw 7.8 Haida Gwaii, 2013 Mw 7.5 Craig, and Two Mw>8.5 Teleseismic EarthquakesTriggered Seismic Events along the Eastern Denali Fault in Northwest Canada. *Bulletin of the Seismological Society of America*, *105*(2B), 1165–1177. https://doi.org/10.1785/0120140156

Armijo, R., Meyer, B., Hubert, A., & Barka, A. (1999). Westward propagation of the North Anatolian fault into the northern Aegean: Timing and kinematics. *Geology*, *27*(3), 267–270. https://doi.org/10.1130/0091-7613(1999)027<0267:WPOTNA>2.3.CO;2

Baiesi, M., & Paczuski, M. (2004). Scale-free networks of earthquakes and aftershocks. *Physical Review E*, *69*(6), 066106. https://doi.org/10.1103/PhysRevE.69.066106

Barka, A. A. (1992). The North Anatolian fault zone. *Annales Tectonicae, Spec. Iss.*, *VI*, 164–195.

Ben-Zion, Y. (2008). Collective behavior of earthquakes and faults: Continuum-discrete transitions, progressive evolutionary changes, and different dynamic regimes. *Reviews of Geophysics*, *46*(4), RG4006. https://doi.org/10.1029/2008RG000260

Ben-Zion, Y., & Lyakhovsky, V. (2006). Analysis of aftershocks in a lithospheric model with seismogenic zone governed by damage rheology. *Geophysical Journal International*, *165*(1), 197–210. https://doi.org/10.1111/j.1365-246X.2006.02878.x

Bohnhoff, M., Bulut, F., Dresen, G., Malin, P. E., Eken, T., & Aktar, M. (2013). An earthquake gap south of Istanbul. *Nature Communications*, *4*. https://doi.org/10.1038/ncomms2999

Bohnhoff, M., Martínez-Garzón, P., Bulut, F., Stierle, E., & Ben-Zion, Y. (2016). Maximum earthquake magnitudes along different sections of the North Anatolian fault zone. *Tectonophysics*, *674*, 147–165. https://doi.org/10.1016/j.tecto.2016.02.028

Bohnhoff, M., Wollin, C., Domigall, D., Küperkoch, L., Martínez-Garzón, P., Kwiatek, G., et al. (2017). Repeating Marmara Sea earthquakes: indication for fault creep. *Geophysical Journal International*, *210*(1), 332–339. https://doi.org/10.1093/gji/ggx169

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. https://doi.org/10.1126/science.1197341.

Bulut, F., & Aktar, M. (2007). Accurate relocation of İzmit earthquake (Mw = 7.4, 1999) aftershocks in Çınarcık Basin using double difference method. *Geophysical Research Letters*, *34*(10), n/a–n/a. https://doi.org/10.1029/2007GL029611

Durand, V., Bouchon, M., Karabulut, H., Marsan, D., Schmittbuhl, J., Bouin, M.-P., et al. (2010). Seismic interaction and delayed triggering along the North Anatolian Fault. *Geophysical Research Letters*, *37*(18). https://doi.org/10.1029/2010GL044688

Ellsworth, W. L., & Bulut, F. (2018). Nucleation of the 1999 Izmit earthquake by a triggered cascade of foreshocks. *Nature Geoscience*, *11*(7), 531. https:/doi.org/10.1038/s41561-018-0145-1

Ergintav, S., Reilinger, R. E., Çakmak, R., Floyd, M., Çakir, Z., Doğan, U., et al. (2014). Istanbul's earthquake hot spots: Geodetic constraints on strain accumulation along faults in the Marmara seismic gap. *Geophysical Research Letters*, 2014GL060985. https://doi.org/10.1002/2014GL060985

Géli, L., Henry, P., Grall, C., Tary, J.-B., Lomax, A., Batsi, E., et al. (2018). Gas and seismicity within the Istanbul seismic gap. *Scientific Reports*, *8*(1), 6819. https://doi.org/10.1038/s41598-018-23536-7

Goebel, T. H., Schorlemmer, D., Becker, T. W., Dresen, G., & Sammis, C. G. (2013). Acoustic emissions document stress changes over many seismic cycles in stick-slip experiments. *Geophysical Research Letters*. https://doi.org/10.1002/grl.50507

Hergert, T., & Heidbach, O. (2010). Slip-rate variability and distributed deformation in the Marmara Sea fault system. *Nature Geoscience*, *3*(2), 132–135. https://doi.org/10.1038/ngeo739

Jones, L. M. (1985). Foreshocks and time-dependent earthquake hazard assessment in southern California. *Bulletin of the Seismological Society of America*, *75*(6), 1669–1679.

Kinscher, J., Krüger, F., Woith, H., Lühr, B. G., Hintersberger, E., Irmak, T. S., & Baris, S. (2013). Seismotectonics of the Armutlu peninsula (Marmara Sea, NW Turkey) from geological field observation and regional moment tensor inversion. *Tectonophysics*, *608*, 980–995. https://doi.org/10.1016/j.tecto.2013.07.016

Kılıç, T., Ottemöller, L., Havskov, J., Yanık, K., Kılıçarslan, Ö., Alver, F., & Özyazıcıoğlu, M. (2017). Local magnitude scale for earthquakes in Turkey. *Journal of Seismology*, *21*(1), 35–46.

Le Pichon, X., Şengör, C., Kende, J., İmren, C., Henry, P., Grall, C., & Karabulut, H. (2015). Propagation of a strike slip plate boundary within an extensional environment: the westward propagation of the North Anatolian Fault. *Canadian Journal of Earth Sciences*. https://doi.org/10.1139/cjes-2015-0129

Malin, P. E., Bohnhoff, M., Blümle, F., Dresen, G., Martínez-Garzón, P., Nurlu, M., et al. (2018). Microearthquakes preceding a M4.2 Earthquake Offshore Istanbul. *Scientific Reports*, *8*(1), 16176. https://doi.org/10.1038/s41598-018-34563-9

Murru, M., Akinci, A., Falcone, G., Pucci, S., Console, R., & Parsons, T. (2016). $M \geq 7$ earthquake rupture forecast and time-dependent probability for the Sea of Marmara region, Turkey. *Journal of Geophysical Research: Solid Earth*, 2015JB012595. https://doi.org/10.1002/2015JB012595

Nadeau, R. M., & McEvilly, T. V. (2004). Periodic Pulsing of Characteristic Microearthquakes on the San Andreas Fault. *Science*, *303*(5655), 220–222. https://doi.org/10.1126/science.1090353

Poupinet, G., Ellsworth, W. L., & 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. https://doi.org/10.1029/JB089iB07p05719

Parsons, T. (2004). Recalculated probability of $M \geq 7$ earthquakes beneath the Sea of Marmara, Turkey. *Journal of Geophysical Research: Solid Earth*, *109*(B5), B05304. https://doi.org/10.1029/2003JB002667

Peng, Z. and Y. Ben-Zion, 2005. Spatio-temporal variations of crustal anisotropy from similar events in aftershocks of the 1999 M7.4 İzmit and M7.1 Düzce, Turkey, earthquake sequences, *Geophys. J. Int.*, 160(3), 1027-1043, doi: 10.1111/j.1365-246X.2005.02569.x.

Renard, F., J. Weiss, J. Mathiesen, Y. Ben Zion, N. Kandula and B. Cordonnier, (2018). Critical evolution of damage towards system-size failure in crystalline rock, *J. Geophys. Res.*, 123, 1969-1986, doi: 10.1002/2017JB014964.

Rothert, E., & Shapiro, S. A. (2007). Statistics of fracture strength and fluid-induced microseismicity. *Journal of Geophysical Research: Solid Earth*, *112*(B4), B04309. https://doi.org/10.1029/2005JB003959

Sakic, P., Piété, H., Ballu, V., Royer, J.-Y., Kopp, H., Lange, D., et al. (2016). No significant steady state surface creep along the North Anatolian Fault offshore Istanbul: Results of 6 months of seafloor acoustic ranging. *Geophysical Research Letters*, *43*(13), 2016GL069600. https://doi.org/10.1002/2016GL069600

Schmittbuhl, J., Karabulut, H., Lengliné, O., & Bouchon, M. (2016a). Long-lasting seismic repeaters in the Central Basin of the Main Marmara Fault. *Geophysical Research Letters*, 2016GL070505. https://doi.org/10.1002/2016GL070505

Schmittbuhl, J., Karabulut, H., Lengliné, O., & Bouchon, M. (2016b). Seismicity distribution and locking depth along the Main Marmara Fault, Turkey. *Geochemistry, Geophysics, Geosystems*, *17*(3), 954–965. https://doi.org/10.1002/2015GC006120

Schoenball, M., Davatzes, N. C., & Glen, J. M. G. (2015). Differentiating induced and natural seismicity using space-time-magnitude statistics applied to the Coso Geothermal field. *Geophysical Research Letters*, *42*(15), 6221–6228. https://doi.org/10.1002/2015GL064772

Selvadurai, P. A., Glaser, S. D., & Parker, J. M. (2017). On factors controlling precursor slip fronts in the laboratory and their relation to slow slip events in nature. *Geophysical Research Letters*, 2017GL072538. https://doi.org/10.1002/2017GL072538

Sengör, A. M. C. (2005). The North Anatolian Fault: a new look. *Ann. Rev. Earth Panet. Sci.*, *33*, 37–112.

Woessner, J., & Wiemer, S. (2005). Assessing the Quality of Earthquake Catalogues: Estimating the Magnitude of Completeness and Its Uncertainty. *Bulletin of the Seismological Society of America*, *95*(2), 684–698.

Wollin, C., Bohnhoff, M., Martínez-Garzón, P., Küperkoch, L., & Raub, C. (2018). A unified earthquake catalogue for the Sea of Marmara Region, Turkey, based on automatized phase picking and travel-time inversion: seismotectonic implications. *Tectonophysics*. https://doi.org/10.1016/j.tecto.2018.05.020

Wu, C., Z. Peng, X. Meng and Y. Ben-Zion, 2014. Lack of spatio-temporal localization of foreshocks before the 1999 Mw7.1 Duzce, Turkey earthquake, *Bull. Seism, Soc. Am.*, 104, 560–566, doi: 10.1785/0120130140

Yamamoto, R., Kido, M., Ohta, Y., Takahashi, N., Yamamoto, Y., Pinar, A., et al. (2019). Seafloor Geodesy Revealed Partial Creep of the North Anatolian Fault Submerged in the Sea of Marmara. *Geophysical Research Letters*, *0*(0). https://doi.org/10.1029/2018GL080984

Zaliapin, I., & Ben-Zion, Y. (2013a). Earthquake clusters in southern California I: Identification and stability. *Journal of Geophysical Research: Solid Earth*, *118*(6), 2847–2864. https://doi.org/10.1002/jgrb.50179

Zaliapin, I., & Ben-Zion, Y. (2013b). Earthquake clusters in southern California II: Classification and relation to physical properties of the crust. *Journal of Geophysical Research: Solid Earth*, *118*(6), 2865–2877. https://doi.org/10.1002/jgrb.50178

Zaliapin, I., & Ben-Zion, Y. (2016a). A global classification and characterization of earthquake clusters. *Geophysical Journal International*, *207*(1), 608–634. https://doi.org/10.1093/gji/ggw300

Zaliapin, I., & Ben-Zion, Y. (2016b). Discriminating Characteristics of Tectonic and Human-Induced Seismicity. *Bulletin of the Seismological Society of America*, *106*(3), 846–859. https://doi.org/10.1785/0120150211

Zaliapin, I., Gabrielov, A., Keilis-Borok, V., & Wong, H. (2008). Clustering Analysis of Seismicity and Aftershock Identification. *Physical Review Letters*, *101*(1), 018501. https://doi.org/10.1103/PhysRevLett.101.018501