JOURNAL OF VOLCANOLOGY AND SEISMOLOGY, 2025, No. 1, pp. 92–110

APPLICATION OF THE NEAREST NEIGHBOR METHOD FOR ANALYZING VOLCANIC SWARMS BASED ON DATA FROM ERUPTIONS OF BÁRÐARBUNGA AND FAGRADALSFJALL VOLCANOES IN ICELAND

© 2025 E. M. Grekov

Institute of Earthquake Prediction Theory and Mathematical Geophysics of the Russian Academy of Sciences (IEPT RAS), Moscow, Russia
Department of Earth Physics, Faculty of Physics, Lomonosov Moscow State University, Moscow, Russia
**e-mail: grekov.em16@physics.msu.ru*

Received September 06, 2024
Revised November 11, 2024
Accepted November 11, 2024

Abstract. The paper is devoted to the analysis of grouping of volcanic seismicity events, especially in volcanic swarms. Volcanic swarms observed during the eruptions of the Bárðarbunga (2014) and Fagradalsfjall (2021) volcanoes in Iceland were analyzed. In the paper, an attempt was made to apply the nearest neighbor method for the stated goal. The method allows identifying groups with different scales of generalized distances. For example, it typically reveals two groups of events in tectonic seismicity and is widely used to identify aftershocks. As a result of the work, significant differences were observed in the shape of the distributions of generalized distances to the nearest neighbor for volcanic seismicity compared to tectonic seismicity. Namely, two types of unimodal distributions were found, one of them is observed mainly before the eruption, and the other during the eruption. The first type is probably caused by the merging of two close distribution modes and reflects the internal heterogeneity of seismicity during such periods. However, the unimodality of distributions makes it difficult to identify events in terms of related (clustered) or independent (background). Based on the results obtained, it can be assumed that before the eruption, the proportion of background seismicity fluctuates around 70%, and during the eruption from 90 to 100%. This may indicate different sources of seismicity at one or another stage of the eruption.

Keywords: *Iceland, seismology, volcanic seismicity, volcanic swarms, seismicity clustering, nearest neighbor method*

DOI: 10.31857/S02030306250106e5

INTRODUCTION

The Ben-Zion–Zaliapin nearest neighbor method [Zaliapin et al., 2008; Zaliapin, Ben-Zion, 2013] allows establishing levels of statistical connections between seismic events according to the parent-offspring scheme, where the parent of an event is its nearest neighbor that occurred earlier in time. This method is widely used for declustering tasks in tectonic seismicity [Baranov, Shebalin, 2019]. It is known that the distribution of generalized distances to the nearest neighbor has a bimodal form in most cases of tectonic seismicity. It is generally accepted that the left mode, corresponding to the group of related events with closer generalized distances, represents aftershock activity, while the right mode

corresponds to unrelated events (sometimes called background events). It should be noted that this is essentially just an interpretation of the presence of two modes in the distribution, although it is well-confirmed for tectonic seismicity. In a more general sense, different modes of such distribution simply reflect the heterogeneity of the seismicity under consideration, which may, for example, indicate different initiation mechanisms for one group or another, or their different nature. That is, in the general case, such a distribution can have more than two modes. For example, in [Malyutin, 2023], a distribution with three modes was observed when studying California seismicity; it was shown that one of the modes corresponded to the seismicity of the geyser area. It is also worth noting that such separation is relative, and if we observed a unimodal distribution, it would be difficult to interpret it unambiguously since there would be nothing to compare it with.

The question of why there are often precisely two groups of events in tectonic seismicity is actually ambiguous. One hypothesis is the division of events by nature into exogenous/endogenous; in this case, some events are considered induced by sources external to the system, while others are internal, self-induced [Sornette, Helmstetter, 2003]. The second process in this case depends on some internal state of the medium, its "preparedness"; sometimes it is assumed that the degree of this "preparedness" can be characterized by the number of offspring initiated by other earthquakes. For example, in [Nandan et al., 2021], the "branching" parameter of the process n is investigated, which in the context of seismicity has the meaning of the average number of direct offspring per parent. A similar parameter is the delta productivity of earthquakes [Baranov, Shebalin, 2019].

From this perspective, it is interesting to consider volcanic seismicity, as it has a different nature of initiation; it would be interesting to see whether such a seismic regime also shows two groups of events similar to those observed for tectonic seismicity, or whether several modes corresponding to different volcanic processes can be observed. Science already knows a number of seismic signals associated with volcanic activity [Minakami, 1960; Gordeev, 2007]. It is worth noting that not all such signals are medium ruptures analogous to tectonic events. Special attention should be paid to volcano-tectonic signals — high-frequency tectonic signals with hypocenters localized at a depth of several tens of kilometers under the volcano, associated with the destruction of the medium under magma pressure [Gordeev, 2007].

Attempts have already been made to analyze volcanic seismicity in the designated context, for example, in the work [Traversa, Grasso, 2010]. Authors evaluated the proportion of background seismicity during periods of quiescence and during magma intrusions along dikes for the volcanoes Etna and Vesuvius, assuming the presence of two groups similar to tectonic seismicity. For this purpose, distributions of inter-event times were used, which for tectonic seismicity is believed to be approximated by a gamma distribution [Traversa, Grasso, 2010], where one of the parameters of this distribution characterizes the proportion of unrelated events [Molchan, 2005]. For periods of quiescence on volcanoes, as well as for the "reference" tectonic seismicity of California, the same gamma distribution was obtained, and its shape did not depend on either the duration of the period under consideration or changes in the level of activity, while the proportion of background seismicity ranges from 20 to 40%. However, for sections of dike intrusions, it turned out that the distribution of times between events does not correspond to the gamma distribution at all [Traversa, Grasso, 2010]. This effect was then successfully reproduced on model data. The effect was observed both with an increase in the level of background activity and with an artificial decrease in time resolution.

In this work, for the same task, we tried to apply another method – the nearest-neighbor method of Ben-Zion–Zaliapin [Zaliapin et al., 2008; Zaliapin, Ben-Zion, 2013].

DATA

The paper analyzes data from the local seismic network of Iceland (the catalog is available on the website of the Icelandic Meteorological Service hraun.vedur.is/ja/viku/). In particular, sections of the catalog associated with the volcanic processes

of two eruptions of the Bárðarbunga (2014) and Fagradalsfjall (2021) volcanoes are studied.

In the first case, a set of events was selected, which represents a cloud of earthquakes migrating from the central caldera to the Holuhraun volcanic plateau. This was a major fissure eruption accompanied by the collapse of the central caldera. However, here we will consider events related specifically to the migration of the seismicity cloud, presumably caused by the intrusion of magma into the rocks (Fig. 1a). By the time the cloud reached the plateau, lava began to flow to the surface there, and the cloud's trail coincides with the location of the volcano's fissure system. This is evidenced by eruption reports from the Icelandic Meteorological Office (IMO). Confirmation of this can also be seen in the works [Einarsson, Brandsdóttir, 2021; Sigmundsson et al., 2015]. In [Sigmundsson et al., 2015], the formation of the dike is analyzed in detail, with segments of the path identified, each being the result of breaking through a barrier under magma pressure. The dike formation was completed around September 4, 2014, and the eruption began on August 29, 2014 [Sigmundsson et al., 2015].

In the second case, distinct migration is not observed, however, seismicity is also concentrated in the area of the volcano's fissure system (see Fig. 1b), and a similar pattern is observed in the distribution of events over time in both cases (Fig. 2). It should also be noted that two "branches" of seismicity are expressed in the spatial distribution of seismicity (see Fig. 1b). According to the study [Fischer et al., 2022], one of them corresponds to seismicity associated with the rift zone, while the other corresponds to seismicity in the volcanic system of channels and fissures. After the 2021 eruption, a short eruption also occurred in 2022. These two events are considered as two episodes of the same volcanic process. The surge in seismicity that began on February 24, 2021, apparently, is also associated with the intrusion of magma into the rocks [Fischer et al, 2022]. The authors associate the events from approximately February 24 to March 19, 2021, with the intrusion of magma along a dike preceding the eruption.

In order to observe the differences in distribution forms as a whole, data from selected areas of the catalog from 1995 to 2022 were analyzed, including periods of quiescence.

Boundaries of the first area: from 64.56 to 64.92 in latitude, from –17.25 to –16.6 in longitude. The magnitude range in the selected area: from –0.72 to 4.63.

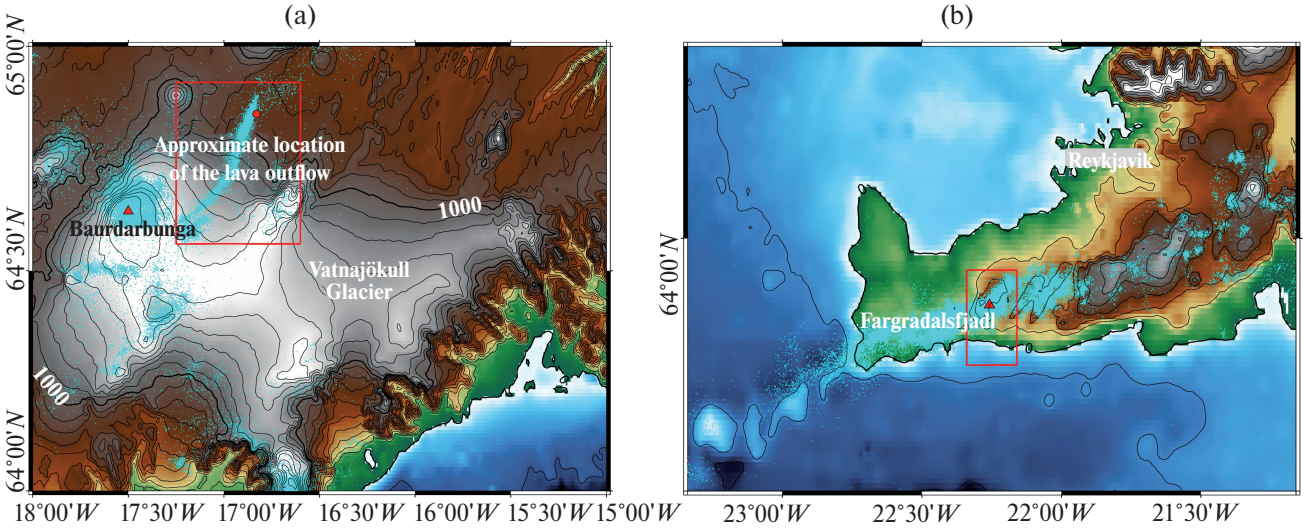

Fig. 1. Distribution of studied earthquakes in space.
a – events from the selected catalog area from 2014 to 2021 for the Bárðarbunga volcano eruption on the Holuhraun plateau 2014–2015; b – catalog events from 2020 to 2022 for the Fagradalsfjall volcano eruption (2021).
Turquoise dots – earthquakes, red triangle – volcano location. The studied areas are highlighted with red rectangles.

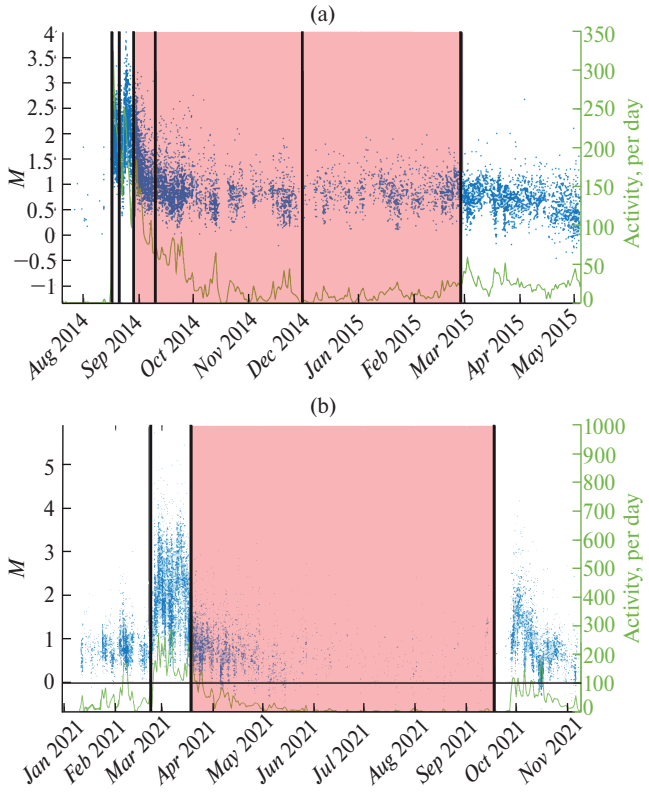

Fig. 2. Distribution of studied earthquakes in time.
a – events from the selected catalog area for the Bárðarbunga volcano eruption on the Holuhraun plateau 2014–2015; b – events from the selected catalog area for the Fagradalsfjall volcano eruption (2021).
Blue dots – earthquakes, light green line – seismic activity (number of events per day), vertical gray lines mark the start and end times of several periods considered in Table 1 and Table 2. Red areas highlight the eruption periods.

Boundaries of the second area: from 63.8 to 63.95 in latitude, from –22.34 to –22.16 in longitude. The magnitude range in the selected area: from –0.8 to 5.72.

Further, we will refer to these two cases under consideration as BAR for the Bárðarbunga volcano and FAG for the Fagradalsfjall volcano.

METHODS

One of the main problems in analyzing volcanic seismicity is the high heterogeneity of the magnitude of completeness, which is probably caused by the network's property of getting "jammed" with a too large flow of events. Moreover, eruptions consist of multiple processes and stages, each of which likely has a different seismicity regime, so first and foremost, it is worth analyzing temporal variations in seismicity parameters.

For this purpose, the studied seismicity is divided into time intervals. The boundaries of the intervals are selected based on three factors: probable process boundaries (for example, before an eruption there is a segment with sharply increased activity, it is logical to assume that some separate process occurs during this period), by the homogeneity of the magnitude of completeness, by the number of events (there need to be enough events in the interval to estimate parameters).

Estimation of the magnitude of completeness for volcanic seismicity presents particular difficulties due to the presence of nonlinear frequency-magnitude distributions, sometimes "rounded horseshoe-shaped", sometimes with double slope, sometimes bimodal (for example, similar in the work [Jacobs, Mcnutt, 2010]). Presumably, this is caused by high heterogeneity of the data. At the same time, the most popular methods for estimating the representative magnitude (for example, MAXC – maximum curvature method) are designed for "classic" type graphs – a linear segment with one "sharp" bend. For the anomalies described above, such methods typically give an underestimated value.

Three methods for estimating the magnitude of completeness have been implemented in this work: MAXC (Maximum curvature), GFT (Goodness-of-Fit Test), MBS (M_C by *b-value* stability). These methods are described in detail in the paper [Mignan, Woessner, 2012]. It is impossible to definitively choose the best method, therefore it is probably most reliable to use all three estimates and then select the maximum one from them.

As a result, the following method is used to estimate the magnitude of completeness (M_C) for the selected time intervals: variations of M_C over time are constructed using the three methods described above, then the maximum value from the variations belonging to the studied interval is selected. Ultimately, some of the selected intervals are excluded because they contain insufficient number of representative events.

The estimation of M_C variations over time is conducted as follows.

1) A certain initial width of the time window is set (for example, $W = 10$ days).

2) Then this window slides through events in time, and M_C and the number of events above the completeness threshold in the current window are estimated.

3) If the number of events in the window turns out to be less than a certain specified threshold (for example, $N_{ther} = 100$), then the window expands to the right until the condition is met.

4) The next window starts sliding from the expanded right boundary of the previous one.

This method allows estimating the magnitude of completeness over time with highly varying activity: where activity is low, the windows will be wider, and the picture will be smoother; where activity is high, the windows will be narrow, and the picture will be more detailed. Otherwise, we would have to skip points where there were insufficient events for estimation.

However, in some intervals, due to their specificity, some methods give knowingly high estimates; in such cases, the most appropriate of the remaining ones is selected.

Next, for the selected intervals with the obtained estimates M_C, the value of the magnitude-frequency distribution slope parameter (*b-value)* is estimated. For this purpose, the Bender method is applied [Bender, 1983]. The value of the fractal dimension presumably does not vary significantly over time, and besides, a large number of events is required for its correct estimation. Therefore, the same value calculated for all periods in the aggregate is used for all intervals. For the BAR case, the value $d_f = 1.2$, is obtained, and for the FAG case $d_f = 1.3$.

Further, with the obtained estimates of M_C, *b-value* and d_f, the distributions of distances to nearest neighbors in the selected intervals are constructed.

Nearest Neighbor Method

The essence of the method consists in introducing a metric – a generalized distance between events in the space-time-magnitude domain (proximity function). For each event, this generalized distance is calculated and the nearest neighbor is found, that is, the event most likely connected to it [Zaliapin et al., 2008; Zaliapin, Ben-Zion, 2013].

The following expression is chosen as the proximity function [Shebalin et al., 2020]:

$$\eta_{ij} = \begin{cases} t_{ij} r_{ij}^{d_f} 10^{-bm_i}, & t_{ij} > 0 \\ +\infty, & t_{ij} \le 0 \end{cases},$$

where t_{ij} is the time between events i and j, r_{ij} is the spatial distance between them, m_i is the magnitude of event i. b is the parameter of the Gutenberg-Richter law; d_f is the fractal dimension of the earthquake epicenter distribution. The earlier of the pair of events is called the "parent", while the later one is its "offspring".

The theoretical justification for why such an expression can be chosen as a proximity function was provided in the work by [Baiesi, Paczuski, 2004].

In most cases for tectonic seismicity, events are divided into two groups based on generalized distances; the first group with smaller distances is associated with clustered (connected) events, the second with background (unconnected) events [Baranov, Shebalin, 2019]. This was also illustrated by a model experiment in the work [Zaliapin et al., 2008].

In real catalogs of tectonic seismicity, both background events and clustered ones are present, so typically the distribution of distances to the nearest neighbor is bimodal and represents a combination of two distributions. To separate these distributions, a slightly modified method from [Baranov, Shebalin, 2019] is used. First, the distribution of distances to the nearest neighbor is built for the complete catalog $\rho_{real}(\eta)$, then the left (clustered) peak is roughly cut off with a variable threshold (for example, with a step of 0.5), the events of the remaining right part of the distribution are shuffled: for each earthquake time, the hypocenter coordinates and magnitudes of another event from the catalog are randomly selected. Then, the distribution $\rho_{random}(\eta)$ for the nearest neighbors in the resulting catalog is

constructed. It is assumed that in the case of tectonic seismicity, such a distribution approximates the background peak of the real distribution. For each preliminary threshold, the shuffling procedure is applied several times and the best option is selected, as shuffling is a random procedure. Then, the best option is also selected from all preliminary thresholds. The criterion for selecting the best option is minimizing the sum of squared differences between the histogram columns of the real distribution and the shuffled distribution, multiplied by the coefficient k, which is found using linear regression on the right slope of the two distributions. The minimization of the sum of squares is carried out in the range from the right maximum of the real distribution to the right. Then, the bimodal distribution can be decomposed into two parts:

$$\rho_{real}(\eta) = (1-k)\rho_{clustred}(\eta) + k\rho_{random}(\eta).$$

To optimize the weight k, the best match $k\rho_{random}(\eta)$ with the right branch $\rho_{real}(\eta)$ along the right slope of the distributions is sought.

The threshold value η_0 is determined from the condition of equality of intensities (number of events per unit time) of clustered earthquake flows with nearest neighbors $\eta \le \eta_0$ and non-clustered events with nearest neighbors $\eta > \eta_0$:

$$1 - F_{clustered}(\eta_0) = 1 - \frac{F_{real}(\eta_0) - kF_{random}(\eta_0)}{1-k} = F_{random}(\eta_0).$$

The input data for the method are: a seismic catalog containing information about the time, coordinates, and magnitude of events; magnitude of completeness M_C; the value of the slope of the magnitude-frequency distribution b; fractal dimension of the catalog d_f.

The steps are described in more detail in the book [Baranov, Shebalin, 2019].

It should be noted that on the plots with the distribution of distances to the nearest neighbor shown below, strictly speaking, only the histogram of all events (blue) is a probability density. The red and yellow histograms (background and clustered events) make up the complete probability density only in sum. And the histogram for the clustered part (yellow) can in some cases have negative values on the vertical axis, which has no mathematical meaning and is the result of inaccurate approximation of the background peak. At the same time, the coefficient k we will further call the background seismicity fraction, and $1-k$ – the degree of clustering, although the meaning of these parameters may be more general than they have in tectonic seismicity.

Criterion of asymmetry in seismic energy release

Additionally, to distinguish swarm activity from large aftershock series, the criterion of asymmetry in seismic energy release is applied [Roland, McGuire, 2009; Passarelli et al., 2018], this is especially important for short time periods that may include such a series entirely, giving a burst of events similar to swarm activity, especially if no migration of the swarm in space is observed. The method consists in calculating the seismic moment for each event using the moment magnitude according to the formula from [Kanamori, 1977]:

$$M_0 \sim 2 * 10^4 * 10^{1,5M_w + 11,8}.$$

Then the centroid time is calculated [Jordan, 1991]:

$$t_{centroid} = \bar{t} = \frac{\sum_i^n t_i * M_{0i}}{M_{0tot}},$$

where t_i – event time, M_{0i} – seismic moment of the event, n – number of events, M_{0tot} – total seismic moment of all events.

Then the asymmetry coefficient is calculated [Roland, McGuire, 2009]:

$$skewness = \frac{\sum_i^n (t_i - \bar{t})^3 * M_{0i}}{M_{0tot} * \sigma^3}.$$

It has been shown [Roland, McGuire, 2009] that this coefficient has values much greater than 8 for aftershock series, i.e., it has left asymmetry,

as most of the seismic moment is released by the mainshock at the beginning of the sequence. For seismic swarms, energy release is more distributed in time, giving small positive or negative parameter values; in [Passarelli et al., 2018], for six swarms the values lie in the range from –4 to 4.

RESULTS

BAR case

Fig. 3 shows the distributions of generalized nearest neighbor distances obtained for the selected periods for the BAR case, as well as their possible decomposition into clustered and background parts. The distribution numbers correspond to the row numbers in Table 1 . Yellow highlights periods with presumed volcanic processes preceding the eruption, while red indicates the period of volcanic activity; triangles of the same colors also mark the corresponding graphs in Fig. 3. For periods presumably associated with volcanic activity, the asymmetry coefficient values of the released seismic moment are also provided. Additionally, the average activity (average number of events per day) is shown for periods of interest.

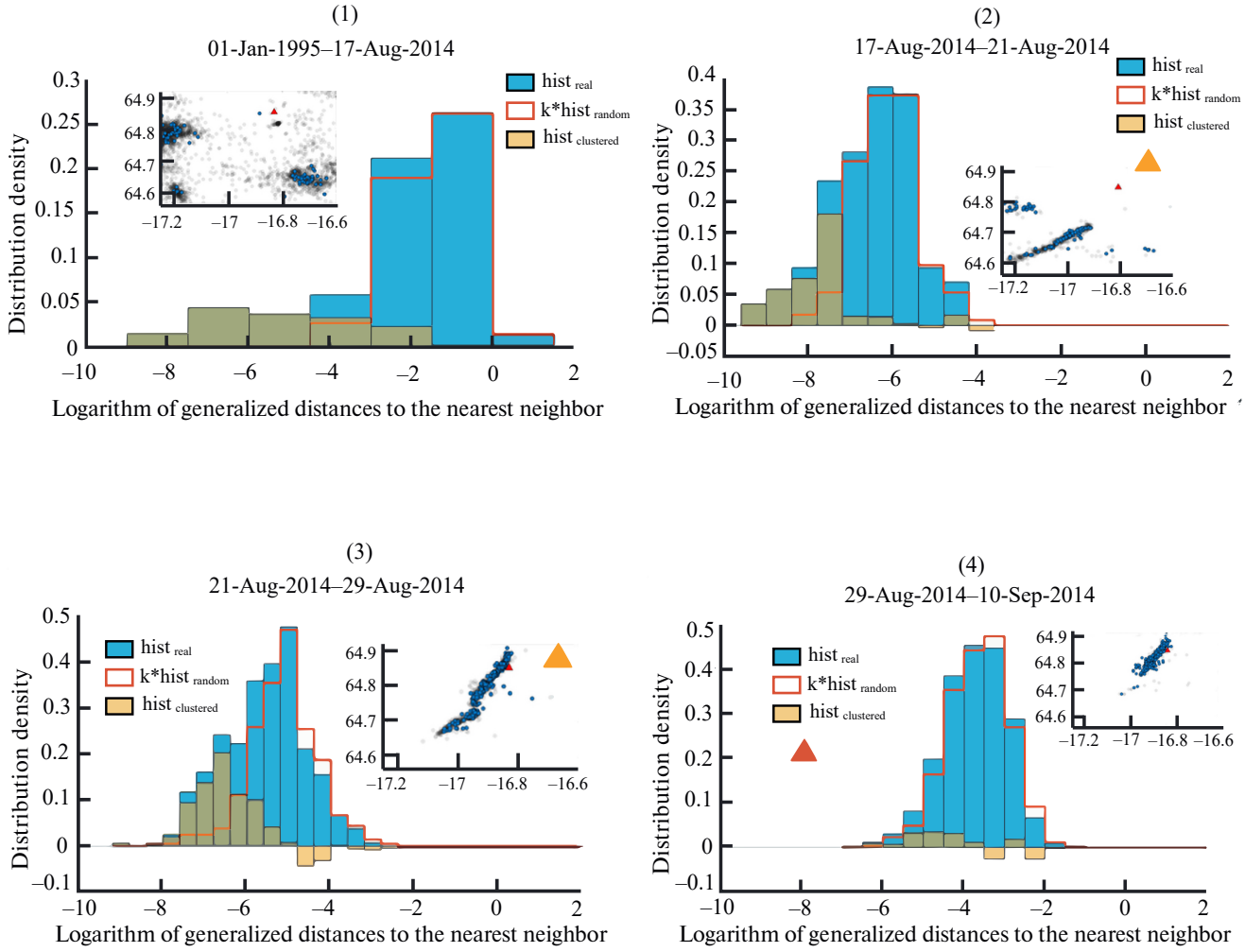

Fig. 3. Distributions of generalized distances to the nearest neighbor in selected time periods for the BAR case. Blue histogram – actual real distribution, red histogram – distribution of randomized catalog (approximation of background peak), yellow histogram – obtained distribution of the clustered part. In the insets – distribution of corresponding events in space, X-axis – longitude, Y-axis – latitude, black semi-transparent circles – events below the representativeness level, blue circles – events above the representativeness level. Graph numbers correspond to row numbers in Table 1. Yellow triangles mark distributions for presumed eruption preparation periods, red triangles – distributions for periods during eruptions.

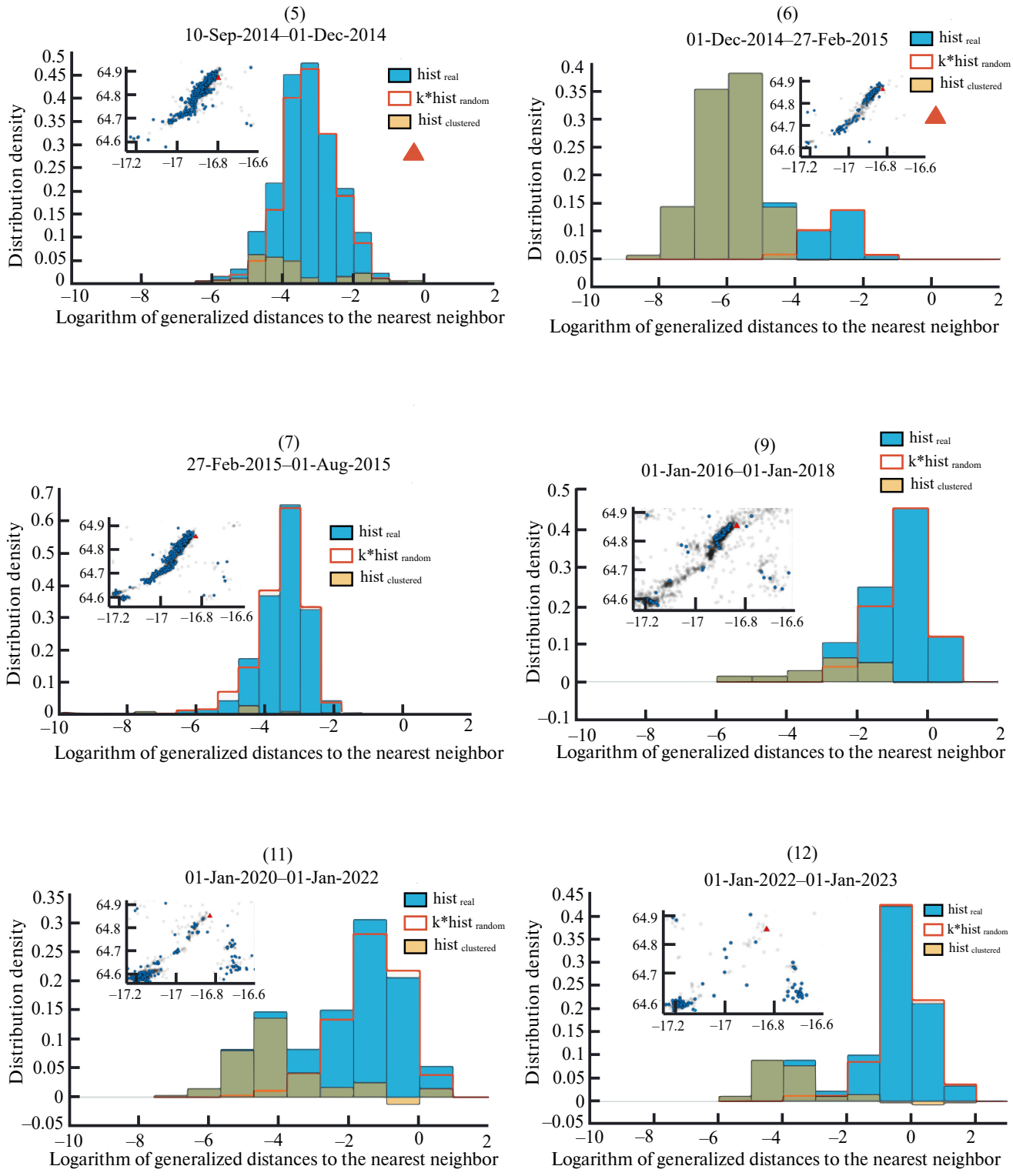

Fig. 3. *Continued*

Let's examine each period in detail. It can be seen that during the quiet period long before the eruption (see Fig. 3(1)), weak seismicity is observed, most likely of tectonic nature, not concentrated in the volcano conduit area. The distribution of generalized nearest neighbor distances for this

Table 1. Seismicity parameters of identified time periods for the BAR case

	Period	M_C	N_C	b	Activity, days^{-1}	k	Energy skewness
1	01.01.1995–17.08.2014	2.39	91	1.28	0.73	0.73	2
2	17.08.2014–21.08.2014	2.29	142	1.63	35.500	0.75	2.869
3	21.08.2014–29.08.2014	2.29	403	1.12	50.375	0.74	0.032
4	29.08.2014–10.09.2014	1.23	662	0.81	55.167	0.95	-0.627
5	10.09.2014–01.12.2014	0.9	863	1.37	10.524	0.86	1.885
6	01.12.2014–27.02.2015	1.19	138	2.19	1.568	0.16	5.671
7	27.02.2015–01.08.2015	0.95	497	2.20	3.207	0.99	0.548
8	01.08.2015–01.01.2016	1.39	16	–	–	–	–
9	01.01.2016–01.01.2018	1.22	69	1.48	0.80	0.80	10
10	01.01.2018–01.01.2020	1.37	44	–	–	–	–
11	01.01.2020–01.01.2022	0.64	357	1.36	0.69	0.69	12
12	01.01.2022–01.01.2023	0.7	90	1.07	0.79	0.79	Note.

Note. M_C — magnitude of completeness; N_C — number of events above the completeness threshold; b — slope of the recurrence graph; k — proportion of background seismicity; energy skewness — coefficient of asymmetry of released energy.

period has a bimodal form characteristic of tectonic seismicity.

In the next period, starting from 08.17.2014, an intrusion process begins, and the distribution shape becomes more unimodal. As mentioned above, when observing a unimodal distribution, difficulties arise because there is nothing to compare the scale of distances with. However, in this case, the shuffled catalog can be used for comparison; its distribution shows what the generalized distances would be if events occurred independently in time (in this case, all events are shuffled without a preliminary threshold). By comparing the observed distribution with the shuffled one, conclusions can be drawn about the presence or absence of clustering. In this period, the observed distribution differs from the shape of the shuffled catalog distribution, allowing the separation of a small "tail" of clustered seismicity in the left part. This pattern is even more pronounced in the next period (see Fig. 3(3)), in the continuation of the intrusion process up to the beginning of lava flow on 08.29.2014.

With the onset of the eruption (see Fig. 3(4), 3(5)), seismicity concentrates in the area of lava outflow, while the distribution shape changes and begins to more closely match the distribution after mixing, with the coefficient k ranging from 85 to 95%.

The sharp change to a bimodal graph shape and decrease in background seismicity during the BAR eruption in period No. 6 (see Fig. 3(6)) is apparently explained by a single strong event (Fig. 4). Since it has a magnitude significantly greater than all other events in this period, many events are connected to it, forming a large aftershock series and creating the bimodal graph shape.

Further, in the period immediately after the eruption (see Fig. 3(7)), seismicity is still concentrated in the area where the intrusion occurred, and the distribution still shows a low degree of clustering. After which, over time (see Fig. 3(9), 3(11), 3(12)), the distribution shape gradually returns to the "classic" bimodality.

For period No. 3, likely corresponding to magma intrusion along a dike, delta productivity values were also calculated [Baranov, Shebalin, 2019]: $\Lambda_{0.5} = 0.2$ — delta productivity with $\Delta M = 0.5$ and $\Lambda_{1.0} = 0.87$ — delta productivity with $\Delta M = 1.0$.

FAG case

Fig. 5 shows the distributions of generalized distances to the nearest neighbor obtained for the selected periods for the FAG case, as well as their possible decomposition into clustered and background parts. The distribution numbers correspond to the row numbers in Table 2. Yellow highlights periods with presumed volcanic processes preceding the eruption, while red highlights periods of volcanic activity, with triangles of the same colors also marking the corresponding graphs in Fig. 5. For periods presumably associated with volcanic activity, the values of the seismic moment release asymmetry coefficient are also provided. For several periods in the table, delta productivity values are also given.

It can be seen that in periods long before the eruption (see Fig. 5(1)–5(4)), seismicity is mainly concentrated within the area highlighted by the green ellipse in the insets of Fig. 5, there is a well-pronounced bimodality in the distribution of generalized distances to the nearest neighbor with a high degree of clustering (from 70 to 90%). This seismicity likely has a tectonic nature [Fischer et al., 2022].

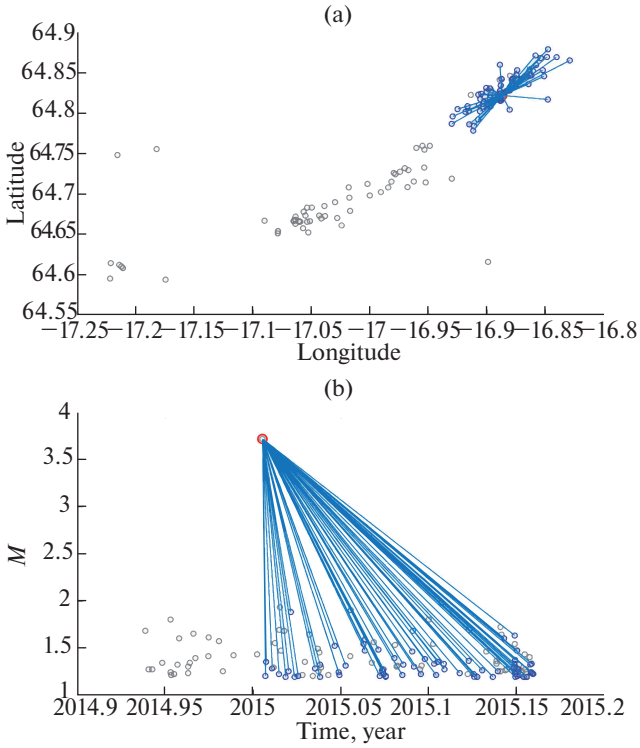

Fig. 4. Aftershock series, which is the presumed cause of the surge in the proportion of background seismicity during the eruption in the BAR case.
a – aftershock series in space; b – aftershock series in time; gray circles – background unrelated events, blue circles – aftershocks, red circle – mainshock of the aftershock series, light blue lines – connections established by the nearest neighbor method.

Table 2. Seismicity parameters of the selected time periods for the FAG case

	Period	M_C	N_C	b	Activity, days^{-1}	k	Energy skewness	$\Lambda_{0.5}$	$\Lambda_{1.0}$
1	01.01.1995–01.01.2019	1.3	1971	0.851	0.179	0.179	0.316	0.316	0.938
2	01.01.2019–01.01.2020	0.91	714	0.776	0.107	0.107	–	–	3
3	01.01.2020–01.07.2020	1.17	223	1.114	0.296	0.296	–	–	4
4	01.07.2020–22.02.2021	1.61	535	0.678	0.198	0.198	–	–	5
5	22.02.2021–19.03.2021	2.84	473	0.901	18.92	0.533	0.826	0.176	0.415
6	19.03.2021–18.09.2021	0.8	1207	1.055	6.596	0.822	2.649	0.129	0.511
7	18.09.2021–20.12.2021	0.89	1179	0.767	12.677	0.044	5.697	0.331	0.990
8	20.12.2021–27.12.2021	2.6	153	0.898	21.860	–	–	–	–
9	27.12.2021–29.07.2022	0.82	626	1.304	2.930	0.678	–	–	10
10	29.07.2022–03.08.2022	3.17	43	–	–	–	–	–	–
11	03.08.2022–21.08.2022	0.75	231	0.818	12.830	0.728	22.841	–	–
12	21.08.2022–01.01.2023	0.54	339	1.288	2.550	0.578	–	–	–

Note. M_C — representative magnitude; N_C — number of events above the threshold of representativeness; b — slope of the recurrence graph; k — proportion of background seismicity; energy skewness — coefficient of asymmetry of released energy; $\Lambda_{0.5}$ — delta productivity with $\Delta M = 0.5$; $\Lambda_{1.0}$ — delta productivity with $\Delta M = 1.0$.

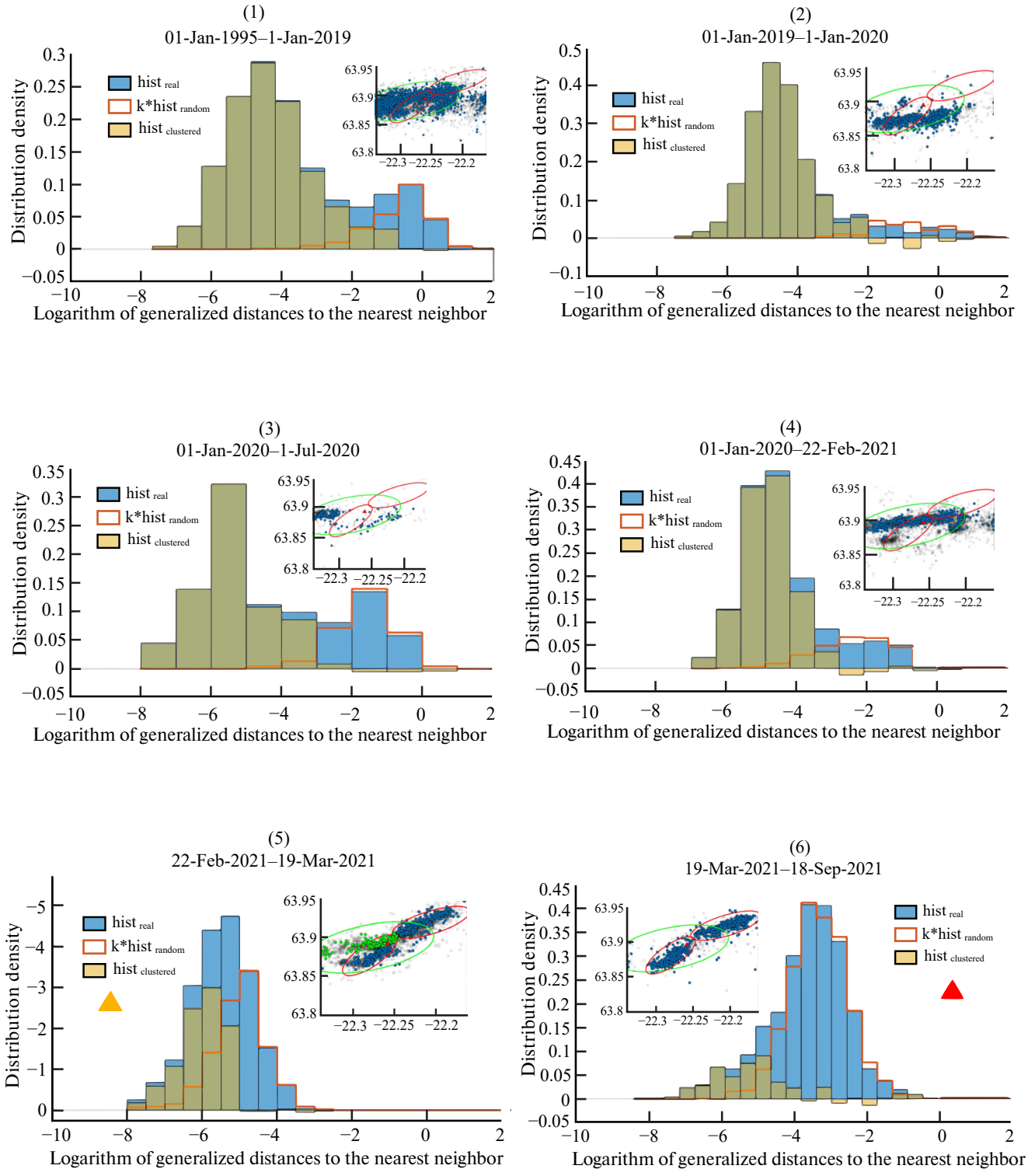

Fig. 5. Distributions of generalized distances to the nearest neighbor in selected time periods for the FAG case.
Blue histogram – actual real distribution, red histogram – distribution of randomized catalog (approximation of background peak), yellow histogram – obtained distribution of the clustered part. In the insets – distribution of corresponding events in space, X-axis – longitude, Y-axis – latitude, black semi-transparent circles – events below the representativeness level, blue circles – events above the representativeness level. The green ellipse shows the area where rift seismicity is concentrated, red ellipses show areas of volcanic seismicity concentration for illustrative purposes. Graph numbers correspond to row numbers in Table 2. Yellow triangles mark distributions for presumed eruption preparation periods, red triangles – distributions for periods during eruptions.

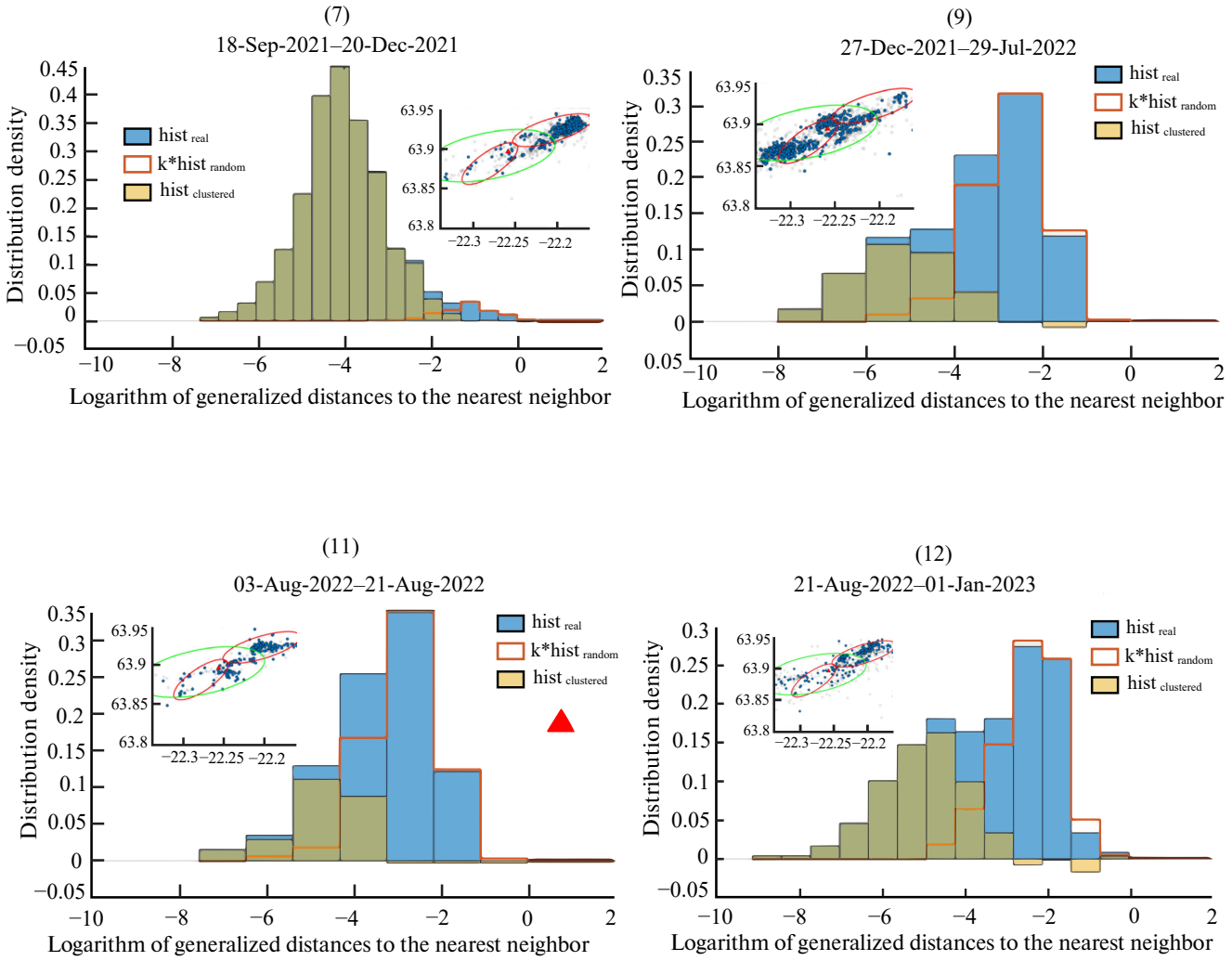

Fig. 5. *Continued*

By the time of the 2021 eruption, events begin to concentrate in the area of the volcanic fissure system (see Fig. 5, highlighted by red ellipses in the insets). The transition period is No. 5 (see Fig. 5(5)), therefore part of the events that may relate to tectonic seismicity was removed from consideration (see Fig. 5(5), removed events are highlighted in green in the inset). In this period, corresponding to the intrusion process before the eruption, a unimodal distribution is observed, similar to what was observed for the BAR case (see Fig. 3(2), 3(3)). Further, during the eruption (see Fig. 5(6)), a unimodal distribution is also observed, but with a pronounced "tail" in the left part, unlike the BAR case, however, the degree of clustering in this period also decreases compared to other periods.

Period No. 7 (see Fig. 5(7)) is ambiguous, the right peak is almost completely absent, and the graph shows a decomposition obtained by manual method, based on the assumption that this period contains an aftershock series, this is also indirectly indicated by the relatively large value of the energy release asymmetry parameter.

Further, in the interval between the two eruption episodes (see Fig. 5(9)), the distribution again acquires bimodality. Then during the second eruption episode (see Fig. 5(11)), it has a rather unimodal form, however, this episode is too short to draw any definitive conclusions.

In the period after the eruption, the distribution again tends toward a bimodal form.

In general, in both cases, there is a difference in the seismicity regime immediately before the eruption, during the eruption, and long before or after the eruption (see Fig. 3(2)–3(5), 3(7)), 5(5), 5(6), 5(11)). At the same time, for tectonic seismicity in the same regions, a "classic" bimodal graph is observed.

It can also be noted that even after the eruption, seismicity continues to concentrate on the same structures where the intrusion occurred for quite a long time. This may be explained by some effect of medium relaxation after intensive impact.

DISCUSSION OF RESULTS

It is worth noting that in the previous section, in all cases for convenience, we refer to the degree of clustering as the proportion that the left (yellow) part of the distribution makes up of the total, although in cases of tectonic and volcanic seismicity, the presence of this part of the distribution may have different origins.

Two types of unimodal distributions

Among the detected unimodal graphs, there appear to be two types of distributions. Before the eruption, a unimodal distribution is observed, which, when all events are shuffled, does not coincide with itself, apparently indicating the presence of connections between events that are broken by the randomization procedure. During the eruption, distributions are observed that, when shuffled, almost completely coincide with themselves, suggesting that this is completely random seismicity, consisting mainly of background events.

It is worth noting that the first type of unimodal distribution also appears to correspond to aftershock series (for example, Fig. 6). The energy asymmetry parameter for these events is: 11.10.

Ambiguity of decompositions

It should be mentioned that for the first type of unimodal distributions, it is difficult to unambiguously decompose into two parts, and it is partially subjective in nature. For example, an alternative decomposition can be presented for period No. 3 of the BAR case (Fig. 7).

Moreover, it is worth keeping in mind that, assuming the two parts of the distribution overlap each other so much that they become indistinguishable, by separating them with one strict threshold, we incorrectly identify a significant portion of events.

The delta productivity values for the alternative decomposition are as follows: $\Lambda_{0.5} = 0.3$ and $\Lambda_{1.0} = 1.13$, and the proportion of background seismicity is about 40%.

To try to solve the problem with unimodal distributions of the first type, one can divide the selected period into even smaller ones (Table 3) and track changes in the distribution over time (Fig. 8).

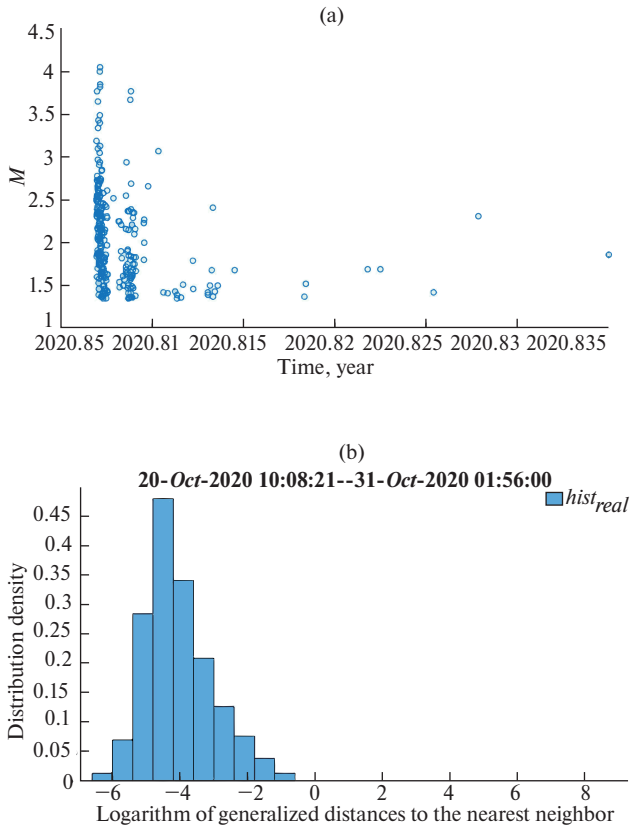

Fig. 6. Example of an aftershock series from the region for the FAG case.
a – distribution of series events in time; b – distribution of generalized distances to the nearest neighbor of the aftershock series.

Table 3. Parameters for a set of consecutive overlapping windows

	Period	Activity, day^{-1}	Energy skewness
1	18.08.2014 14:37:25–22.08.2014 08:06:39	75.631	0.262
2	19.08.2014 13:04:54–24.08.2014 10:52:52	50.120	–1.385
3	21.08.2014 00:56:32–24.08.2014 11:36:40	70.547	–2.725
4	22.08.2014 06:39:52–25.08.2014 19:12:23	56.493	–0.031
5	23.08.2014 13:25:44–26.08.2014 19:25:45	103.384	0.222
6	24.08.2014 12:15:30–28.08.2014 14:30:14	54.964	–0.399
7	25.08.2014 22:41:39–30.08.2014 21:41:18	42.557	1.409
8	27.08.2014 23:13:11–01.09.2014 10:00:54	57.531	–0.618
9	29.08.2014 18:53:29–03.09.2014 01:12:21	92.913	2.693

Figure 8 shows the transition of the distribution form from the first type to the second. The graphs covering the period from August 21 to 30, 2014 collectively represent period No. 3 (see Table 1), in all of them the proportion of background seismicity k ranges approximately from 60 to 70%. Based on this, preference can be given to the first variant of decomposition for this section (see Fig. 3(3)), which also gives a value of k = 74%, although this is not strict proof.

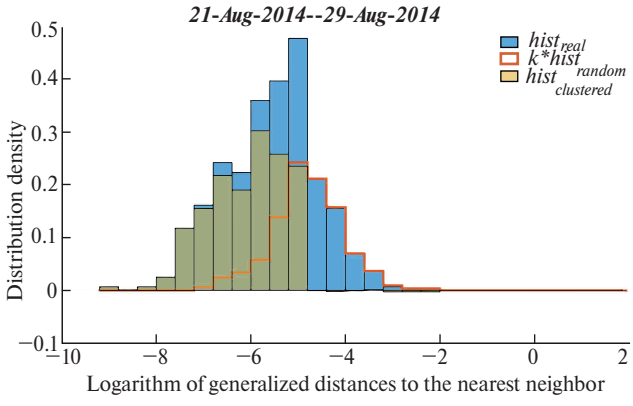

Fig. 7. Example of an alternative variant of decomposition into background and clustered parts for period No. 3 (see Table 1).
Blue histogram – actual real distribution, red histogram – distribution of randomized catalog (approximation of background peak), yellow histogram – obtained distribution of clustered part.

Overall, from the perspective of the nearest neighbor distance distribution shape, three types of situations are observed.

1. "Classical" bimodal distribution, easily separable into two components.

2. Unimodal distribution, which after randomization almost completely coincides with itself.

3. Unimodal distribution with a small "hump" or tail on the left slope of the distribution. With careful parameter selection, the mode of the randomized distribution approaches the mode of the original distribution, and also well approximates its right slope, allowing decomposition into two components.

Practice shows that the first case is observed for "classical" tectonic seismicity in most regions of the Earth. The second case is also observed in some situations, for example, for geyser seismicity [Malyutin, 2023], as well as in laboratory experiments [Matochkina, 2023]. The third situation, apparently, is observed in cases of high activity, due to which both components are too close to each other and become poorly distinguishable; examples of this have also been obtained in laboratory experiments [Matochkina, 2023] and for induced seismicity [Baranov et al., 2020]. Presumably, the authors of the article [Traversa, Grasso, 2010] encountered the same effect.

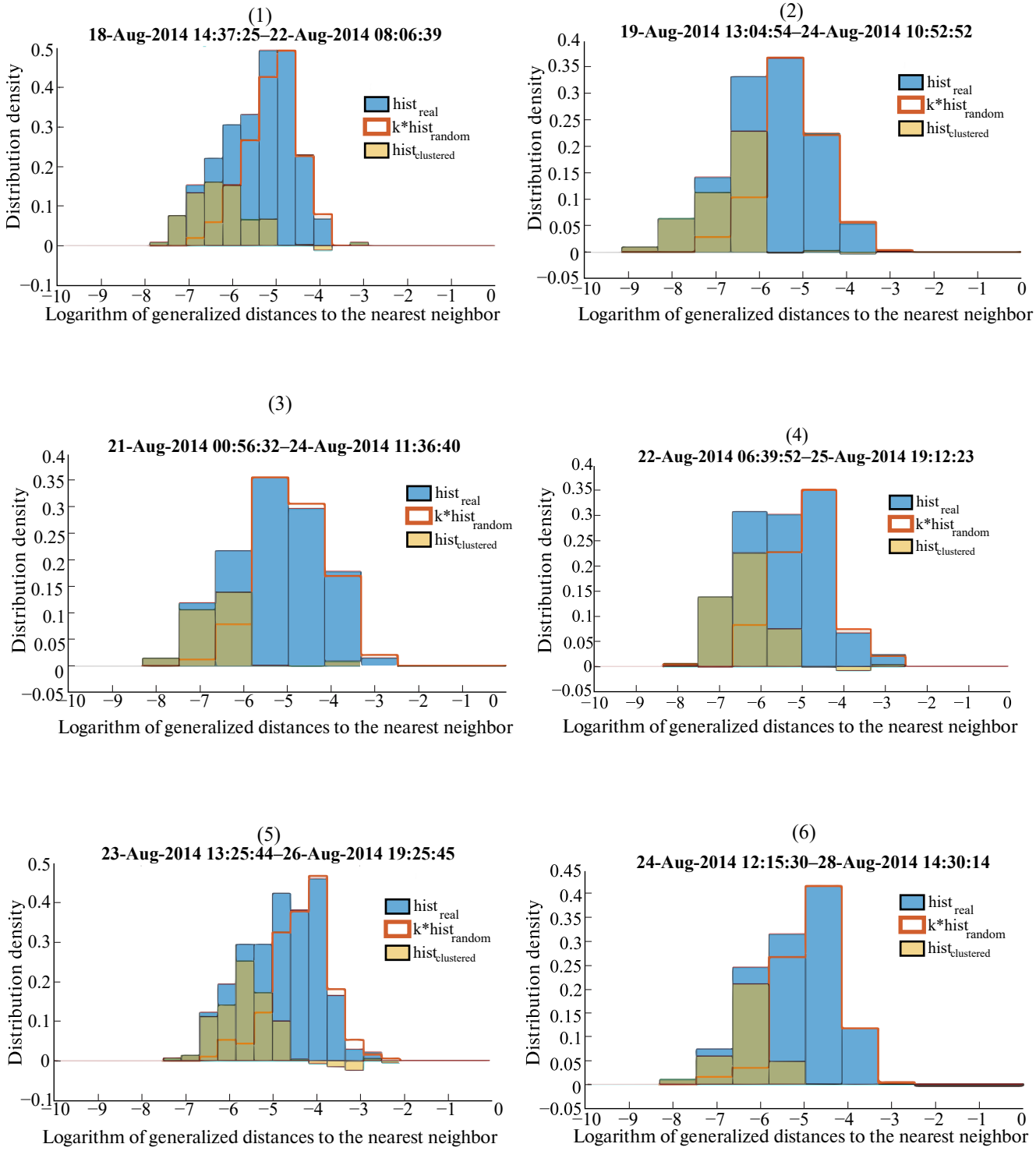

Fig. 8. Obtained distributions of generalized distances to the nearest neighbor for the BAR case for a set of consecutive time windows in the period before the eruption.
Blue histogram – actual real distribution, red histogram – distribution of randomized catalog (approximation of background peak), yellow histogram – obtained distribution of clustered part. The numbers of the graphs correspond to the row numbers from Table 3.

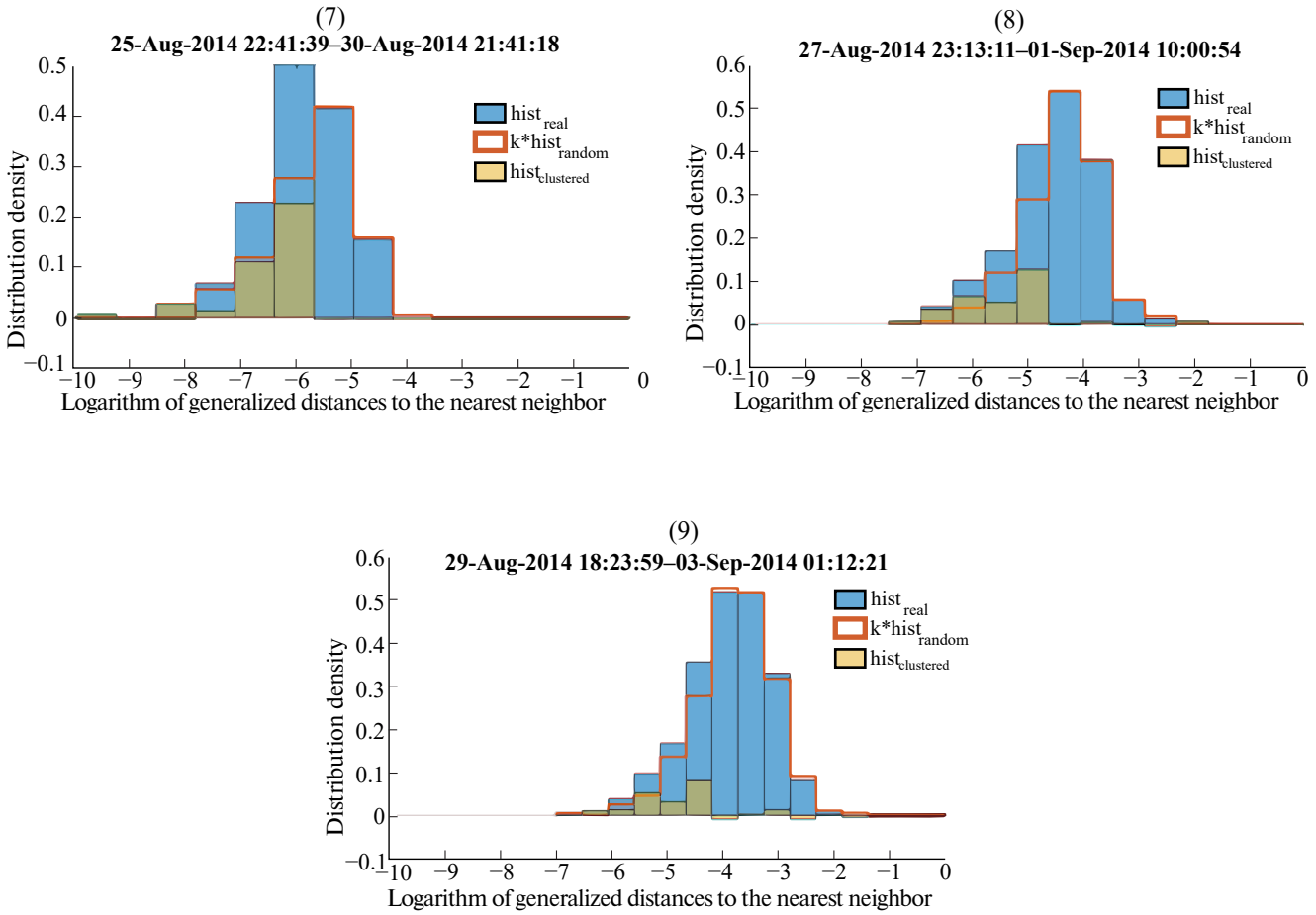

Fig. 8. *Continued*

Thus, the distribution observed during the intrusion period likely exhibits some internal heterogeneity of seismicity.

MAIN CONCLUSIONS

1) For seismicity associated with eruptions, unimodal forms of nearest neighbor distance distributions are observed, in contrast to the "classical" bimodal distribution known for tectonic seismicity.

2) Two types of unimodal distributions are observed:

a) a unimodal distribution that, after randomization, almost completely coincides with itself;

b) a unimodal distribution with a small "hump" or tail on the left slope of the distribution. The mode of the randomized distribution may coincide with the mode of the original distribution, and also well approximates its right slope, allowing decomposition into two components. The fact that such a distribution does not coincide with itself upon randomization may indicate that these events do not occur completely randomly.

3) The second type is observed, presumably, for volcanic swarms preceding an eruption, likely caused by magma intrusion into dikes. Additionally, a similar distribution shape is observed for single aftershock series; to distinguish one from the other, the parameter of seismic moment release asymmetry over time was analyzed.

4) The unimodality of the distribution in the case of volcanic swarms can be hypothetically explained by the fact that due to the high concentration of events in space and time, the background and clustered peaks are too close to each other and become poorly distinguishable. In this case,

attempts to separate the distribution into two give estimates of the background seismicity proportion from 60 to 90%. Such a distribution shape likely indicates internal heterogeneity of such seismic regimes. It is also worth noting that a very similar distribution is observed in laboratory studies for acoustic emission in rock samples [Matochkina, 2023], and was also observed for anthropogenic seismicity [Baranov et al., 2020].

5) The first type of unimodal distributions manifests itself during the eruption. It is worth noting that a very low degree of clustering is observed, for example, for seismicity in geothermal zones, in geyser valleys [Malyutin, 2023]. And this type of unimodal distributions (complete coincidence of the randomized distribution with the real one) is observed for events initiated by water injection into rocks [Malyutin, 2023].

6) Attempts were made to decompose events into background and clustered parts for unimodal distributions of volcanic swarms preceding eruptions. And further, based on this separation, to calculate delta productivity, due to the small number of events and their low magnitude, it is not possible to select values ΔM greater than 1. The values turn out to be similar to those known for tectonic seismicity (about 0.5), for example, [Baranov, Shebalin, 2019]. However, for the BAR case before the eruption (see Table 1, period No. 3), the value is abnormally large at 0.87, or 1.13 depending on the chosen decomposition option.

CONCLUSION

Volcanic seismicity of two eruptions of volcanoes in the Iceland region was analyzed for grouping and homogeneity. In this work, the nearest neighbor method of Ben-Zion–Zaliapin was applied for this purpose. It is shown that the distribution of generalized distances to the nearest neighbor, which usually has a bimodal form in tectonic seismicity, becomes unimodal during periods associated with volcanic activity. Moreover, it has been demonstrated that this feature is specifically related to eruption periods, rather than to the selection of the regions under consideration. However, even among unimodal distributions, two types were identified by comparison with the distribution of a randomized catalog. There is reason to believe that one of the types, observed predominantly during intrusion periods before an eruption, is rather a composition of two close distributions, indicating the heterogeneity of seismicity, while it should be kept in mind that the nature of such heterogeneity may differ from that observed in tectonic seismicity. During eruptions, this heterogeneity decreases or almost disappears. In general, both types of distributions have already been observed in other studies: in laboratory experiments [Matochkina, 2023], for technogenic seismicity [Baranov et al., 2020], when studying events initiated by water injection into rocks [Malyutin, 2023]. Establishing the nature of heterogeneity during intrusion periods may be an important direction for further research.

ACKNOWLEDGMENTS

The author expresses gratitude to the Icelandic Meteorological Service IMO (https://en.vedur.is) for open access to the earthquake catalog of the local seismic network.

FUNDING

The work was supported by the Russian Science Foundation Project No. 20-17-00180P.

CONFLICT OF INTEREST

The author of this work declares that he has no conflict of interest.

REFERENCES

1. *Baranov S.V., Zhukova S.A., Korchak P.A., Shebalin P.N.* Productivity of technogenic seismicity // Physics of the Earth. 2020. Pp. 40–51. DOI: 10.31857/S0002333720030011.
2. *Baranov S.V., Shebalin P.N.* Patterns of post-seismic processes and forecasting the hazard of strong aftershocks. M.: RAS, 2019. 218 p.
3. *Gordeev E.I.* Seismicity of volcanoes and control of volcanic activity // Bulletin of the Far Eastern

Branch of the Russian Academy of Sciences. 2007. No. 2. Pp. 38–45.

4. *Malyutin P.A.* Impact of fluid regimes on variations in earthquake productivity according to field experiment data // Problems of comprehensive geophysical monitoring of seismically active regions // Proceedings of the Ninth All-Russian Scientific and Technical Conference with International Participation, September 24–30, 2023. Petropavlovsk-Kamchatsky, 2023. Pp. 156–162.
5. *Matochkina S.D.* Verification of the earthquake productivity law in laboratory rock destruction experiments // III All-Russian Scientific Conference with International Participation "Modern Methods of Seismic Hazard Assessment and Earthquake Prediction" (October 25–26, 2023, ITPZ RAS, Moscow). M.: ITPZ RAS, 2023. Pp. 160–164.
6. *Baiesi M., Paczuski M.* Scale-free networks of earthquakes and aftershocks // Phys. Physical Rev. E // Statistical, nonlinear, and soft matter physics. 2004. Vol. 69. Iss. 066106 .
DOI: 10.1103/PhysRevE.69.066106
7. *Bender B.* Maximum likelihood estimation of *b* values for magnitude grouped data // Bull. of the Seismological Society of America. 1983. Vol. 73. Pp. 00–00.
8. *Einarsson P., Brandsdóttir B.* Seismicity of the Northern Volcanic Zone of Iceland // Front. Earth Sci. 2021. Vol. 9. 628967.
DOI: 10.3389/feart.2021.628967
9. *Fischer T., Hrubcova P., Salama A., Doubravová J., Agustsdottir T., Gudnason E., Horalek J., Hersir G.P.* Swarm seismicity illuminates stress transfer prior to the 2021 Fagradalsfjall eruption in Iceland // Earth and Planet. Sci. Lett. 2022. Vol. 594. 117685.
DOI: 10.1016/j.epsl.2022.117685
10. *Jacobs K., Mcnutt S.* Using seismic b-values to interpret seismicity rates and physical processes during the preeruptive earthquake swarm at Augustine Volcano 2005–2006. // US Geological Survey Professional Paper. 2010. Pp. 59–75.
11. *Jordan T.H.* Far-field detection of slow precursors to fast seismic ruptures // Geophys. Res. Lett. 1991. Vol. 18. Pp. 2019–2022.
12. *Kanamori H.* Energy release in great earthquakes // J. Geophys. Res. 1977. Vol. 82(20). Pp. 2981–2987.
13. *Mignan A., Woessner J.* Estimating the magnitude of completeness for earthquake catalogs // Community Online Resource for Statistical Seismicity Analysis. 2012.
DOI: 10.5078/corssa-00180805. Available at http://www.corssa.org
14. *Minakami T.* Fundamental research for predicting volcanic eruptions. Part 1 // Bull. Earthq. Res. Inst. Univ. Tokyo. 1960. Vol. 38. Pp. 497–544.
15. *Molchan G.* Interevent Time Distribution in Seismicity: A Theoretical Approach // Pure and Applied Geophysics. 2005. Vol. 162.
DOI: 10.1007/s00024-004-2664-5
16. *Nandan S., Ram S., Ouillon G., Sornette D.* Is Seismicity Operating at a Critical Point? // Phys. Rev. Lett. 2021. Vol. 126.
DOI: 10.1103/PhysRevLett.126.128501
17. *Passarelli L., Rivalta E., Jónsson S., Hensch M., Metzger S., Jakobsdóttir S.S., Maccaferri F., Corbi F., Dahm T.* Scaling and spatial complementarity of tectonic earthquake swarms. // Earth and Planet. Sci. Lett. 2018. Vol. 482. P . 62–70.
DOI: http://doi.org/10.1016/j.epsl.2017.10.052
18. *Roland E., Jeffrey J.* McGuire. Earthquake swarms on transform faults // Geophys. J. International. 2009. Vol. 178. Pp. 1677–1690.
19. *Shebalin P.N., Narteau C., Baranov S.V.* Earthquake productivity law // Geophys. J. International. 2020. Vol. 222. Issue 2. Pp. 1264–1269.
https://doi.org/10.1093/gji/ggaa252
20. *Sigmundsson F., Hooper A., Hreinsdottir S., Vogfjörd K., Ofeigsson B., Heimisson E., Dumont S., Parks M., Spaans K., Gudmundsson G., Drouin V., Árnadóttir T., Jonsdottir K., Gudmundsson M., Högnadóttir T., Fridriksdottir H., Hensch M., Einarsson P., Magnússon E., Eibl E.* Segmented lateral dyke growth in a rifting event at Bárðarbunga volcanic system, Iceland // Nature. 2015. Vol. 517. Pp. 191–195.
21. *Sornette D., Helmstetter A.* Endogeneous Versus Exogeneous Shocks in Systems with Memory // Phys. A: Statistical Mechanics and its Applications. 2003. Vol. 318. 577–591.
DOI: 10.1016/S0378-4371(02)01371-7
22. *Traversa P., Grasso Jean-Robert* . How is Volcano Seismicity Different from Tectonic Seismicity? // Bull. of the Seismological Society of America. 2010. Vol. 100.
DOI: 10.1785/0120090214

23. *Zaliapin I., Ben-Zion Y.* Earthquake clusters in southern California I: Identification and stability // J. Geophys. Res. Solid Earth. 2013. Vol. 118. Pp. 2847–2864.
DOI: 10.1002/jgrb.50179

24. *Zaliapin I., Gabrielov A., Keilis-Borok V.I., Wong H.* Clustering analysis of seismicity and aftershock identification // Phys. Rev. Lett. 2008. Vol. 101. 018501.
DOI: 10.1103/PhysRevLett.101.018501