AIMS Geosciences

AIMS Geosciences, 11(4): 979–992.
DOI: 10.3934/geosci.2025042
Received: 28 May 2025
Revised: 25 October 2025
Accepted: 06 November 2025
Published: 24 November 2025

https://www.aimspress.com/journal/geosciences

Research article

Neural network clustering of a set of geophysical data in comparison with the distribution of earthquake focuses in the middle basin of Lake Baikal

Irina V. Popova*, Daria A. Orekhova and Sergei M. Korotaev

Geoelectromagnetic Research Center of Schmidt Institute of Physics of the Earth RAS, 108840, Moscow, Troitsk, P.O.B. 30, Russian Federation

* **Correspondence**: Email: popov7376@mail.ru: Tel: +7(495) 840-7062.

Abstract: The high degree of earthquake clustering observed in the central part of the Baikal rift zone necessitates an investigation into the conditions conducive to such a distribution of seismic events. The crust peculiar properties, as characterized by geophysical data, affect the course of the major tectonic and seismic processes. To study the contribution of different types of geophysical data to the formation of seismic event grouping structures in the study area, clustering based on the Kohonen neural network was performed. Clustering was performed on parameters such as the propagation velocities of longitudinal, transverse seismic waves and their ratio, resistivity, and temperature. As a result, clusters outlining zones of increased seismic activity were identified through the combination of temperature and resistivity.

Keywords: earthquakes; neural network; Baikal; resistivity; seismic velocity; temperature

1. Introduction

The Baikal Rift Zone is characterized by geophysical field anomalies, the intensive removal of deep heat, and the most recent tectonic activation of the area. Various geophysical methods are employed to address the challenges associated with studying the seismically active Baikal Rift Zone (BRZ). In our study, we consider anomalies of thermal, electric, and wave (elastic wave velocity) fields. The relationship between increased seismic activity and temperature is discussed in [1]. Peaks in temperature anomalies in the rift basins of the Baikal region mainly correspond to tectonically active structures, particularly zones of deep faults where fluids are withdrawn [2]. According to deep seismic sounding data [3], within the framework of the velocity model of the Earth's crust, it was found that a characteristic feature of the distribution of earthquake foci under the central region of Lake Baikal is the clustering of hypocenters into separate, extended zones. Electroprofiling and Magnetotelluric Sounding (MTS) enable conducting zones at various depths to be identified and mapped. These zones are areas of the Earth's crust that have been disturbed due to block interactions and are permeable to fluids. Ancient faults with graphite mineralization, along which modern tectonic movements often occur, are characterized by negative anomalies in the natural electric field [4].

However, when applied separately, geophysical methods reveal only specific features of the weakened zones of the Earth's crust in which earthquakes can occur. We use Kohonen self-organizing maps [5] to perform an integrated analysis of the Earth's interior's different geophysical characteristics. The unique feature of self-organizing maps technology is its ability to transform a multidimensional state of geophysical characteristics into a two-dimensional space of clusters. Kohonen's neural network algorithm was used in [6] to identify areas with similar petrophysical characteristics using seismic velocities and resistivity in the studied subsurface spaces.

Our aim is to identify clusters based on temperature, resistivity, longitudinal, transverse seismic wave velocities and their ratio in different combinations, and to compare them with seismic activity in the middle basin of Lake Baikal. Kohonen maps are also employed to analyze the contribution of various types of data and the comparative effectiveness of different geophysical parameters that are crucial to the formation of potential earthquake foci in this area.

2. Methods

2.1. Geological and geophysical characteristics of the middle basin of Lake Baikal

The Baikal basin structure corresponds to a large graben, with zones of regional faults running along its edges [4]. The poorly articulated relief, caused by movements along the faults, has sedimentary layers up to 4 km thick. Rift structures are defined in the middle of the Lake Baikal basin. Mantle fluids that catalyze phase transitions in the crust and control the rheological properties of the lithospheric mantle may play a key role in the formation of rift structures [1]. The emergence of an area of an abnormal temperature field and fluids is linked to the ascent of plumes to the base of the lithosphere. Mantle plumes spread laterally along the lithosphere's base, exerting a powerful thermal effect on it. Depending on the conditions of their interaction with the lithosphere, they can

lead to rifting [7]. Deformation processes occur in the region, which causes an abnormal temperature field combined with an anomalous gravitational field [1].

Various studies were conducted in the middle part of Lake Baikal. In [8], the data obtained by time-domain electromagnetic sounding was interpreted. According to these data, the upper part of the Earth's crust has a block structure accompanied by faults with a resistivities of 10–55 Ω·m. In [9], the authors used the results of MTS at 5 points also in the middle part of the Baikal water area. A horizontally section with a three conductive layers was reconstructed. In [4], the geoelectric section in the same area was also constructed based on the results of the MTS. According to this section, the crust beneath Lake Baikal has a resistivity of 50–200 Ω·m in the coastal regions, whereas in the axial region there is a conductive sub-vertical zone with a resistivity of 0.5–5 Ω·m, which is most likely associated with the fluidized zone.

According to the seismic studies [10], the thickness of the crust in the Baikal rift basin zone is 36 km (the velocity of longitudinal waves in the basement is 6–6.8 km/s). The sedimentary layer has two levels: The lower level consists of metamorphosed sedimentary formations (velocity of 4.8–5.1 km/s), and the upper level consists of Neogene–Quaternary sediments (velocity of 1.7–3.6 km/s). An upper mantle anomalous zone of seismic wave velocities in the range of 7.7–7.8 km/s is also present under the BRZ.

A transverse wave section was constructed by teleseismic recordings of the international PASSCAL experiment 1991–1992 [11]. Here, there are alternating layers of low and high velocity, most of which may correspond to powerful mylonite zones accompanying large thrusts.

Since 1965, more than 700 measurements of geothermal parameters in the upper sedimentary layer have been carried out with immersion thermographs in the water area of Lake Baikal [12]. Since 1993, measurements have also been carried out in underwater wells. The heat flow through the lake bottom almost everywhere exceeds 50 MW/m^2 (on average 71+−21 MW/m^2). At the framing elevations, however, it varies in the range of 40–50 mW/m^2 [13]. In the zones of underwater faults, increased heat removal is observed.

For our work, we used geoelectric, seismic waves, and geothermal sections. The geoelectric section from [4] was constructed based on the results of the MTS, with at 3 km intervals along the profile and at depths of up to 30 km. These measurements were carried out in the range of periods from 0.003 to 10000 s using the MTU-5 equipment from Phoenyx Geophysics. Data inversion was performed using the LineInterMT software package. A single consolidated seismic section was obtained using data from three published seismic profiles. The velocities of longitudinal waves were considered in all three profiles. The first profile is part of the Ust-Uda–Chita profile, which crosses the Baikal water area [10]. This profile was constructed based on observations from the Novosibirsk experimental and methodological vibroseismic expedition of the Siberian Branch of the Russian Academy of Sciences between 1990 and 1992. There were used pneumatic oscillation sources and bottom oscillation recorders. Data from the second profile can be found on the VSEGEI website (sheet N-48_GK_1: https://webftp.vsegei.ru/GGK1000/2009_N-48/N-48_GK_1.pdf). This profile runs in the same area to the previous one and connects with the third profile located along Lake Baikal [14]. The recordings were made by ocean bottom seismometers, with air guns as the source.

The geothermal model section was calculated under the quasi-stationary assumption of the BRZ thermal regime [1,12].

As the distance between these three sections did not exceed 20 km, they were then projected onto a single plane from Cape Krestovsky to the Bay Proval (see Figure 1). Information about earthquakes between 2003 and 2020 used for Figures 1–8 was obtained from the Russian Earthquakes database of the Unified Geophysical Service of the Russian Academy of Sciences (http://eqru.gsras.ru/events/run/index.php).

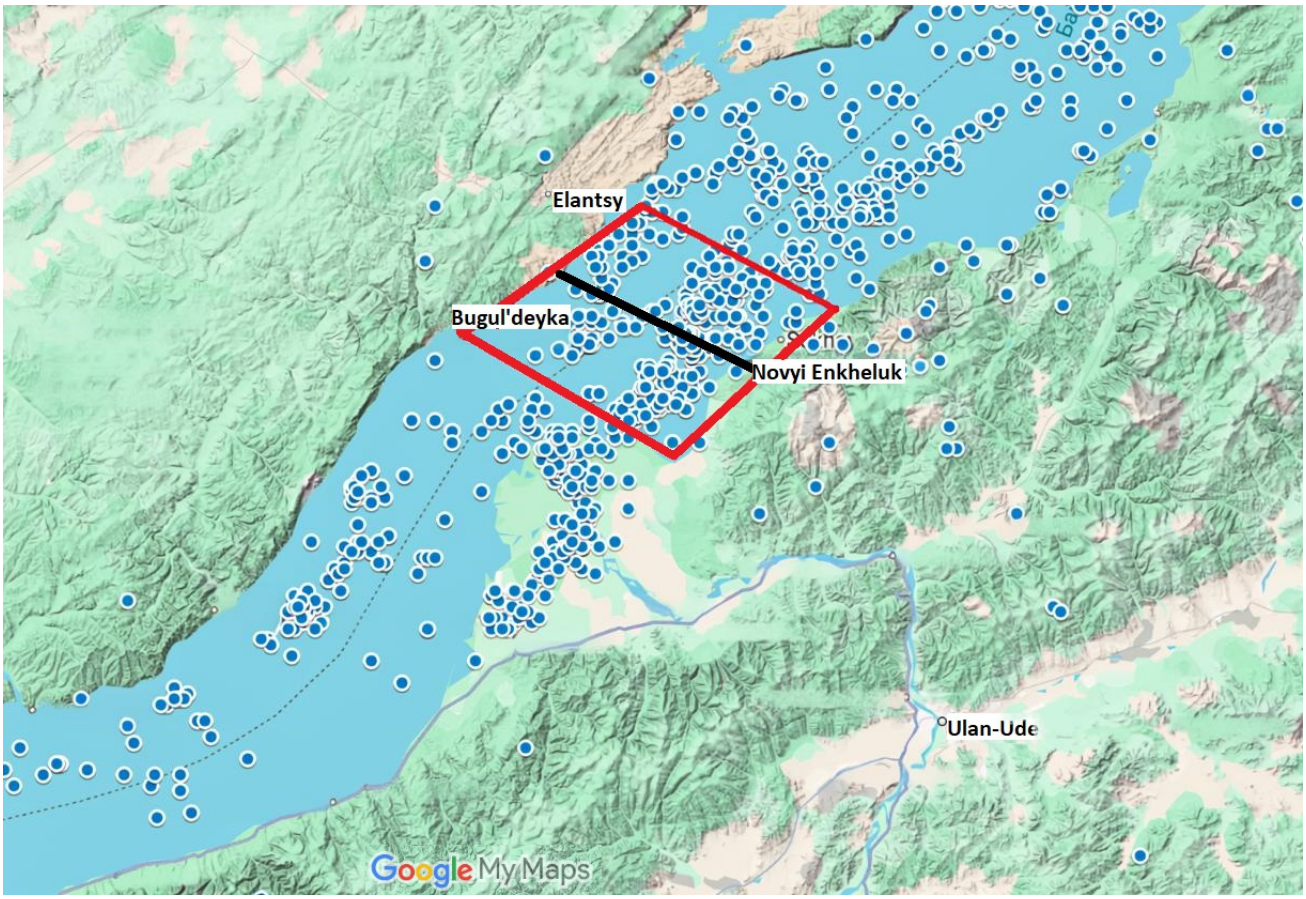

Figure 1. The position of the section under consideration (black line), as constructed from geophysical data. The epicenters of earthquakes are marked with blue circles. The geoelectric, seismic, and thermal sections are in the red area.

Figures 2a, 2b, 2c, and 2d show the geoelectric, longitudinal velocities of seismic waves, and geothermal and transverse velocities of seismic waves sections. These are combined with projections of seismic foci from within a 10 km radius of the sections (earthquake magnitudes $M \geq 5.5$ are marked with a star and the rest with a cross).

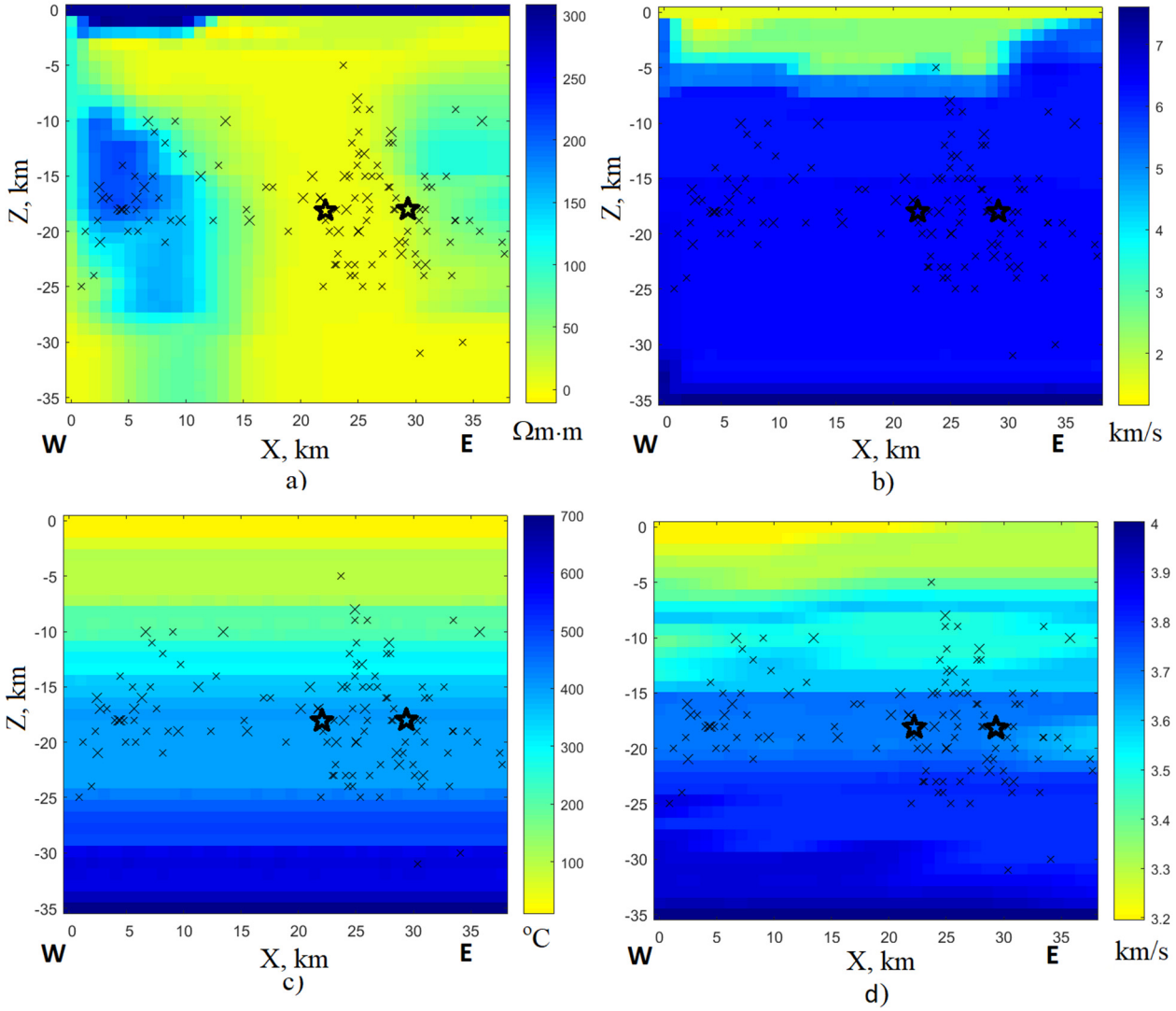

Figure 2. (a) Geoelectric section, (b) longitudinal velocities of seismic waves, (c) geothermal section, and (d) transverse velocities of seismic waves. X – the distance along the profile starting from NW extremity. Earthquake hypocenters with magnitudes M ≥ 5.5 are marked with a star and the rest (earthquakes between 2003 and 2020 years) with a cross.

Based on the presented sections, it can be seen that earthquake foci are mainly at depths of 10–25 kilometers, but no more detailed localization is visible.

2.2. Seismicity of the area

According to [15], earthquakes in central Baikal occur under the influence of horizontal stretching, which appears due to displacement mainly along northeast-southwest planes. In the horizontal plane, earthquake epicenters are grouped along lines that correspond to the stretching of Lake Baikal [3,15]. The largest numbers of epicenters are in a wide strip running along the eastern shore of Lake Baikal (Figure 3).

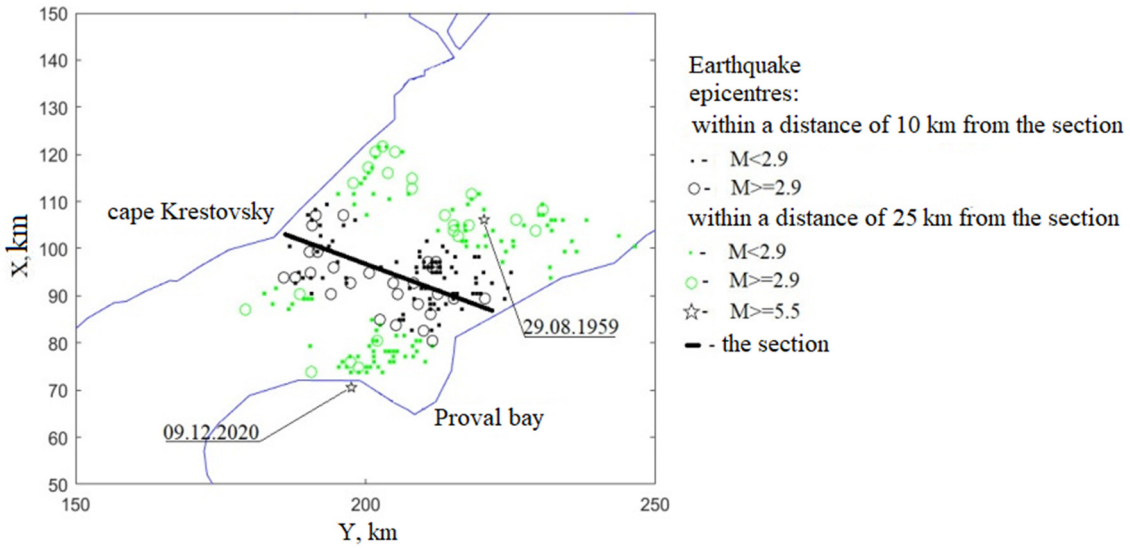

Figure 3. The middle part of Lake Baikal showing earthquake epicenters (with hypocenter depths ranging from 5 to 35 km).

This strip is associated with one of the largest earthquakes that was recorded in the central basin of Lake Baikal on 29 August 1959 (magnitude 6.8). The source's depth was determined to be at 18 km. The epicenter was 17 km to the north from the section under consideration [16]. Another major earthquake occurred in the same area, 25 km to the south from the section, in the Selenga River delta on 9 December 2020 (the Kudarin earthquake), with a magnitude of 5.5. The focus of this earthquake was determined to be at a depth of 18 km [17]. A second strip of epicenters runs along the western shore of Lake Baikal and Olkhon Island. Both strips intersect the section considered in the article.

Information about earthquakes in the middle of Lake Baikal was obtained from the Russian Earthquakes database of the Unified Geophysical Service of the Russian Academy of Sciences (http://eqru.gsras.ru/events/run/index.php) for depths of the seismic source ranging from 3 to 41 km between 2003 and 2020. The average margin of error in determining the depth of the earthquake focus was 4.8 km. The layer with depths from 15 to 25 km is the most seismically active; 70% of events fall within this layer. A further 20% of focuses are at depths between 5 and 15 km. Earthquake magnitudes vary from 1.7 to 5.5; 80% of these are very weak events with $M < 2.9$.

2.3. Clusterization using the Kohonen neural network

Kohonen's neural network is used for clustering the input space. It is based on unsupervised learning. The number of input variables in a neural network equals the number of features that characterize the object under study. Each neuron in the input layer is connected to all the neurons in the output layer. These relationships are described by weight vectors W. All objects in the training set are projected onto a set of rectangular grid nodes of a predefined size. The number of neurons in the output layer corresponds to the expected number of clusters. In the considered two-layer architecture, the signal propagates from the inputs to the outputs. In our case, the input vectors are formed by

geophysical characteristics, such as the propagation velocities of longitudinal, transverse seismic waves and their ratio, resistivity, and temperature, at each point of the section. The researcher sets the number of clusters in the output layer. During the training process, the distance between the weight vector w_{kj} and the input vector X is calculated for each output layer neuron k using the following formula:

$$d_k = \sqrt{\sum_{j=1}^{n} (w_{kj} - x_j)^2} \tag{1}$$

After that, the winning output neuron $k_{\min}$ is searched for, where the distance (1) is minimal. Then, the weights of this neuron and its neighbors are adjusted according to the following rule:

$$w_{kj}(t+1) = w_{kj}(t) + \cdot\eta\,(x_j - w_{kj}(t)), \tag{2}$$

where $w_{kj}(t)$ is the value of the weight coefficient for the connection between input neuron j and output neuron k at time t; η is the learning rate coefficient; and x_j is the value of neuron j in the first layer. The adjustment of weights (2) is then repeated for all vectors in the input dataset. Training continues until the system reaches the desired state: A slight change in the weights, or stabilization of the network output when vectors do not move between clusters. The number of clusters in the research process may vary; it may decrease if larger objects need to be found, or increase if the results of the clustering process need to be detailed. However, the process of cluster number enlargement stops when the clusters' structure stops changing significantly, or when they collapse into small groups of disconnected clusters. During the learning process, the weights are adjusted so that input vectors with similar characteristics activate the same neuron in the output layer. This enables input vectors to be classified into groups of similar elements. Once training is complete, the neural network visually displays multidimensional input data on a plane of output neurons. This representation shows the relationship in the input data.

3. Results

In our research, we combined geophysical data with data on seismic events to determine whether a particular set of physical properties is related to the localization of earthquake focal zones in the middle basin of Lake Baikal. To this end, we constructed self-organizing Kohonen maps using a combination of two geophysical parameters: 1) Resistivity and temperature; 2) resistivity and the ratio of velocities of longitudinal and transverse seismic waves; and 3) temperature and the ratio of velocities of longitudinal and transverse seismic waves. Finally, a cluster map was built for the combination of these three geophysical characteristics. Earthquake foci within ±10 km of the section were projected onto all cluster maps (marked with crosses), as were the locations of the two largest earthquakes: Srednebaikalskoe (M = 6.8, 29/08/1959) and Kudarinskoe (M = 5.5, 09/12/2020) (marked with stars). A total of 131 seismic events are represented in the figures.

Figures 2(a) and 2(c) show the resistivity and temperature separately. Figure 4 shows the Kohonen network clustering results for the two variables combined. The section under consideration

is divided into clusters, and the scale shows the number of each cluster. The range of resistivity and temperature values for each cluster is indicated next to the scale. In this case, the X-axis is oriented along the section.

The method enables clustering to be performed fairly quickly, changing both the set of parameters under consideration and the number of clusters, so hundreds of options were considered in a short time. After a series of experiments with setting a different number of output clusters, it turned out that this set of parameters is most stably and coherently described by eight clusters. As can be seen, most of the earthquakes are grouped into two clusters: Red (№ 7) and light blue (№ 2).

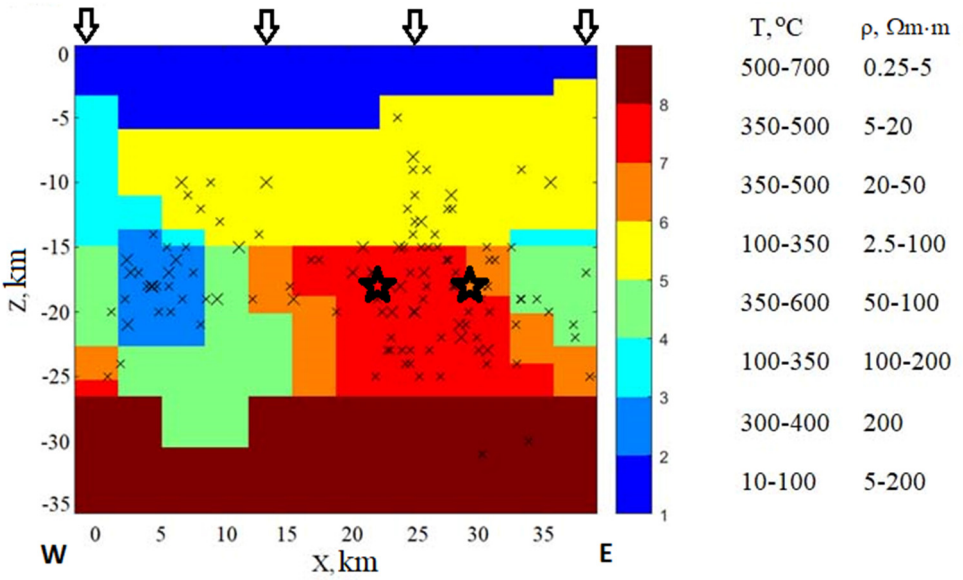

Figure 4. The result of clustering resistivity and temperature data, taking into account the distribution of earthquake foci. Earthquake hypocenters with magnitudes M ≥ 5.5 are marked with a star and the rest (earthquakes between 2003 and 2020) with a cross. Arrows mark the faults crossing the profile.

The red cluster, which has depths ranging from 15 to 25 km, corresponds to a seismically active zone along the eastern coast (see Figure 3). The authors [4] associate this band with a fluidized zone. This cluster corresponds to temperatures of 350–500 °C and low resistivity values of 5–20 Ω·m. The light blue cluster, which has depths ranging from 14 to 22 km, corresponds to a band of earthquake foci near the western shore (see Figure 3). This cluster is more compact, with temperatures ranging from 300 to 400 °C, and much higher resistivity values of around 200 Ω·m compared to cluster 4, which is to the right of the fluidized zone. There are quite a lot of earthquake focuses in the yellow cluster (№ 5). They are not widespread everywhere, but in two extended areas that can be associated with the presence of fault zones. Their location in the horizontal plane is shown in Figure 5. The western zone appears to be associated with the discharge of a Marine fault, and the eastern zone may also correspond to the discharge system. The earthquake focuses near the Gulf of Failure are on the fault. The cluster corresponds to temperatures from 100 to 350 °C and resistivities from 5 to 50 Ω·m.

However, in the zones of active faults, the temperature can be increased (at a depth of 10 km to 300–350 °C [2]), which creates similar conditions to those of the red cluster.

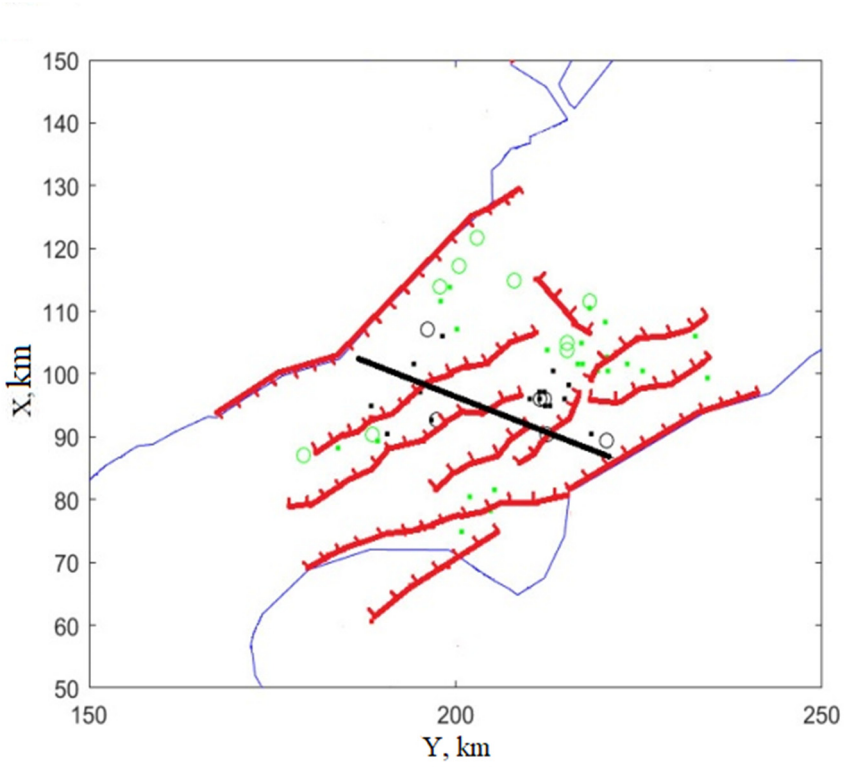

Figure 5. Distribution of earthquake epicenters with focal depths between 5 and 15 km. Based on data from [18].

Figure 6 shows the clustering results obtained using the Kohonen network for the combination of resistivity and ratio of velocity values for longitudinal and transverse seismic waves (Vp/Vs) in a given section. As in the previous case, most earthquake foci fall into two clusters (red and light blue). However, the eastern cluster (red, № 7) does not mark the base of the seismogenic zone at a depth of 25 km.

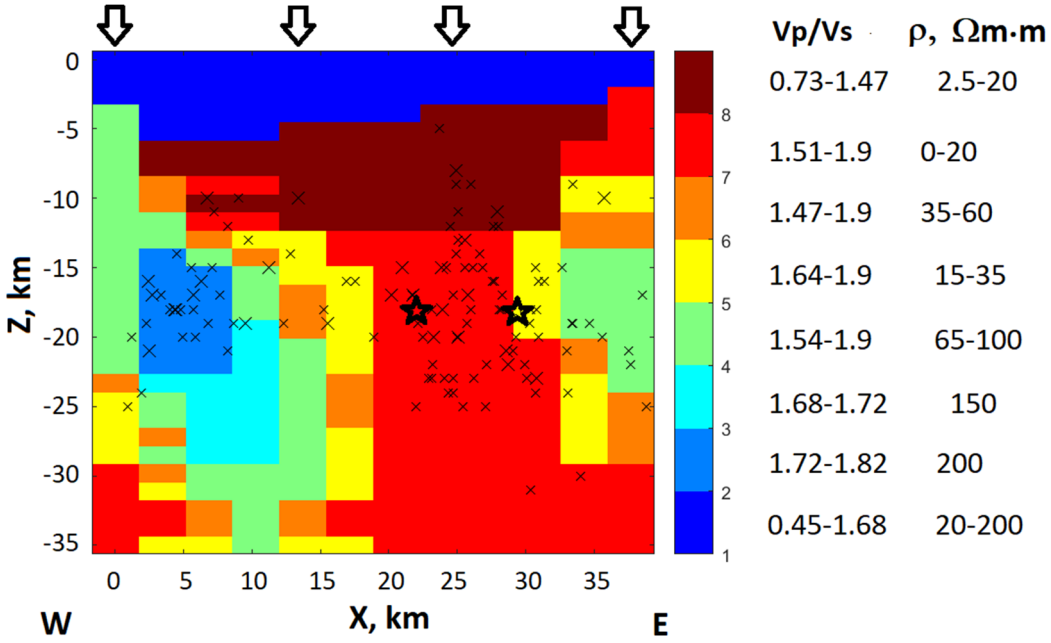

Figure 6. The result of clustering data on the resistivity and the ratio of velocities of longitudinal and transverse seismic waves. Earthquake hypocenters with magnitudes M ≥ 5.5 are marked with a star and the rest (earthquakes between 2003 and 2020) with a cross. Arrows mark the faults crossing the profile.

Figure 7 shows the results of the joint clustering analysis in terms of the ratio of velocities of longitudinal and transverse seismic wave velocity and temperature. In this case, the selected clusters are subhorizontal. Most of the earthquake foci fall into clusters №4-65 (green, yellow, and orange). However, there are sinkholes within which there are no earthquake foci. As with the temperature-resistivity pair, the seismic layer is limited from below by the cluster boundary at a depth of about 25 kilometers for this set of parameters.

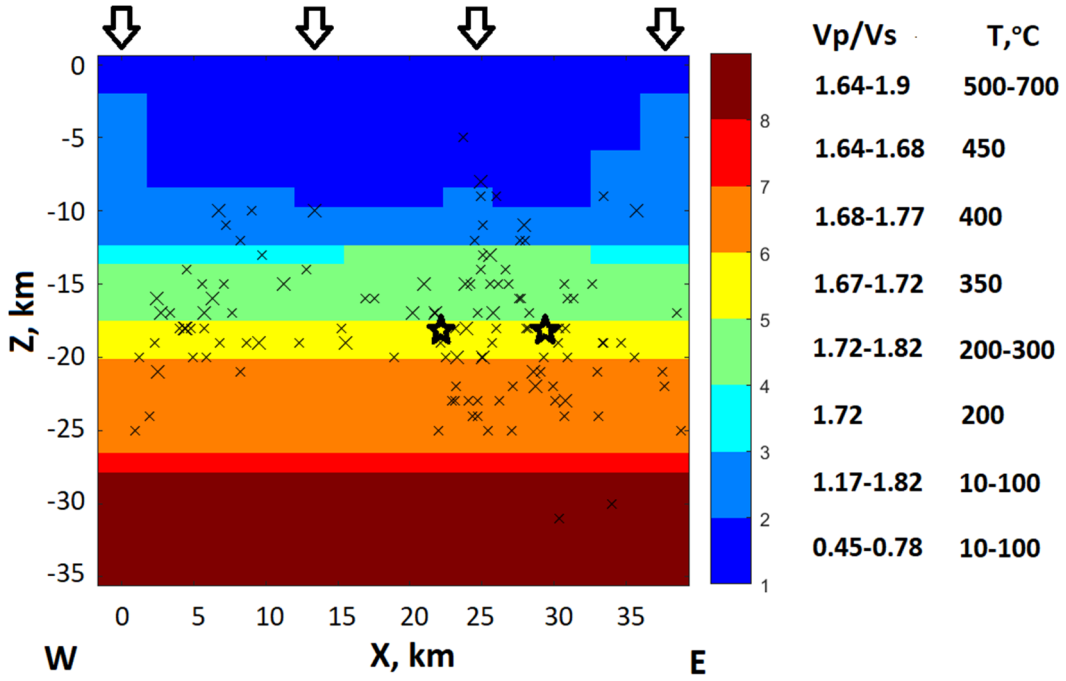

Figure 7. The result of clustering data on the ratio of velocities of longitudinal and transverse seismic waves and temperature. Earthquake hypocenters with magnitudes M ≥ 5.5 are marked with a star and the rest (earthquakes between 2003 and 2020) with a cross. Arrows mark the faults crossing the profile.

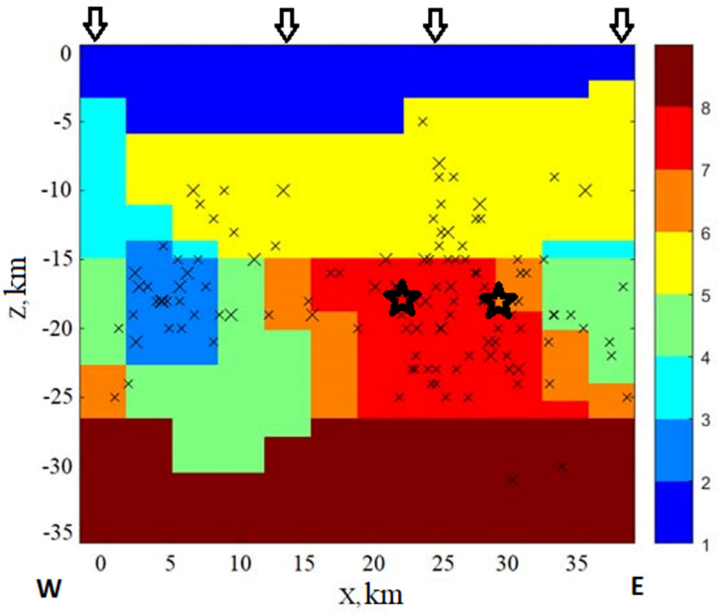

Figure 8. The result of clustering data on resistivity, ratio of velocities of longitudinal and transverse seismic waves, and temperature. Earthquake hypocenters with magnitudes M ≥ 5.5 are marked with a star and the rest (earthquakes between 2003 and 2020) with a cross. Arrows mark the faults crossing the profile.

Finally, clustering was performed on three types of data (temperature, the ratio of velocities of longitudinal and transverse seismic waves, and resistivity) simultaneously (see Figure 8). The resulting clusters, with minor variations, corresponded to those shown in Figure 4 for the combination of two parameters: Temperature and resistivity.

4. Conclusions

The Baikal rift zone belongs to the highly seismic regions of Russia. The grouping of seismic events in space and time is its distinctive feature. There are many methods [19] for isolating clusters of seismic events from regional earthquake catalogs. However, the algorithms used do not take into account the conditions of occurrence of weakened zones and, thus, the possibility of seismic events in such zones in the future. The swarm type of seismic activity can be associated with fault zones, waveguides, mantle inhomogeneities, and structural and material complexes of the Earth's crust, which are characterized by anomalies of geophysical fields. However, individual geophysical methods are limited by the specifics of their application, discussed in the introduction. To understand the conditions for the formation of clusters of seismic events, the integration of geophysical data using the Kohonen neural network was applied.

The results of clustering by the ratio of velocities of longitudinal and transverse seismic waves and temperature simultaneously show that these parameters highlight clusters with subhorizontal boundaries. However, these clusters contain some areas without seismic events.

Combining the ratio of velocities of longitudinal and transverse seismic waves and resistivity has shown that when resistivity is used, not only the subhorizontal, but also the subvertical boundaries of clusters are distinguished. However, the base of the seismogenic zone is not isolated.

Temperature and resistivity is the most informative pair of parameters (Figure 4). Clusters containing most of the seismic events that can be associated with the presence of a fluidized zone and fault zones are outlined. In addition, an upper boundary of up to 5 km without seismic events (Figure 4, dark blue cluster) as well as a low-seismic lower layer starting from about 27 km (Figure 4, brown cluster) are identified. A range of temperature and resistivity values is determined for each cluster.

Temperature and resistivity determine the physical and mechanical properties of the Earth's crust, consequently affecting the course of major tectonic and seismic processes. A quasi-stationary anomalous temperature field is created under the Baikal rift zone. In combination with gravity, this activates the lithospheric layer and initiates deformation in the region [1]. Limiting the depth of the earthquake foci to the 25 kilometers by help of clustering using temperature and resistivity indicates that the disappearance of the rocks' brittle properties and the manifestation of their plasticity probably occurs at these depths.

Author contributions

Irina V. Popova: Conceptualization, Formal analysis, Investigation, Methodology, Software, Writing, Writing – review & editing. Daria A. Orekhova: Data curation, Formal analysis, Investigation, Visualization, Writing, Writing – review & editing. Sergei M. Korotaev: Project administration, Supervision, Validation.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Conflict of interest

The authors declare no conflict of interest.

References

1. Adamovich AN, Ivanova SV (2005) The role of the temperature factor in the evolution of the stress state of the Baikal rift zone (mathematical modeling by the finite element method) (in Russian). In: Topical issues of modern geodynamics of Central Asia. Novosibirsk: Publishing House of the Siberian Branch of the Russian Academy of Sciences, 294.
2. Duchkov AD, Lysak SV, Golubev VA, et al. (1999) Heat flow and geotemperature field of the Baikal region (in Russian). *Geol Geophys* 40: 287–203.
3. Suvorov VD, Tubanov TsA (2008) Distribution of local earthquakes in the crust beneath central lake Baikal (in Russian). *Geol Geophys* 49: 611–620. https://doi.org/10.1016/j.rgg.2007.09.019
4. Seminsky KZh, Kozhevnikov NO, Cheremnykh AV, et al. (2013) Interblock zones of East Siberia: tectonophysical interpretation of geological and geophysical data. *Geodyn Tectonophys* 4: 203–278. https://doi.org/10.5800/GT-2013-4-3-0099
5. Kohonen T (1987) Self-Organization and Associative Memory. Berlin: Springer-Verlag.
6. Spichak VV, Bezruk IA, Popova IV (2008) Deep cluster petrophysical sections construction using geophysical data for assessment of hydrocarbon potential. *J Geophys* 2: 29–38.
7. Lobkovsky LI, Kotelkin VD (2000) Geodynamics of mantle plumes, their interaction with the asthenosphere and lithosphere and their surface manifestation in rift and trap formation (in Russian). In: General issues of tectonics. Tectonics of Russia. Proceedings of the XXXIII Tectonic Conference. Moscow: Geos, 304–308.
8. Morozova GM, Manstein AK, El'tsov IN, et al. (1998) Deep electromagnetic sounding surveys with a controlled source in the Baikal rift zone (in Russian). Geophysical methods of the Earth's crust studies. Publishing studies of SIC UIGGM, SB RAS, Novosibirsk, 52–62.
9. Moroz YF, Moroz TA (2012) Deep geoelectric section of the Baikal rift (in Russian). Vestnik KRAUNTs, Nauki o Zemle 2 (20), 114–126.
10. Sun Y, Krylov SV, Yang B, et al. (1996) Deep seismic sounding of the lithosphere on the Baikal-North Eastern China International transect (in Russian). *Geol Geophys* 37: 3−15.
11. Mordvinova VV, Artemyev AA (2010) The three-dimensional shear velocity structure of lithosphere in the southern Baikal rift system and its surroundings. *Russ Geol Geophys* 51: 694–707. https://doi.org/10.1016/j.rgg.2010.05.010
12. Duchkov AD, Lysak SV, Golubev VA, et al. (1999) Heat flow and geothermal field of the Baikal region (in Russian). *Geol Geophys* 40: 287–303.

13. Lysak SV, Dorofeeva RP (1991) Temperature of the Earth's crust in the southern part of Eastern Siberia (in Russian). In: Geothermal models of geological structures. St. Petersburg: VSEGEI, 99–109.
14. Ten Brink US, Taylor MH (2002) Crustal structure of central Lake Baikal: Insights into intracontinental rifting. *J Geophys Res* 107: 2-1–2-15. https://doi.org/10.1029/2001JB000300
15. Tubanov TsA, Predein PA, Tsydypova LR, et al. (2021) Results and prospects of seismological observations in the central part of the Baikal Rift (in Russian). *Russ Seismol J* 3: 38–57. https://doi.org/10.35540/2686-7907.2021.4.03
16. Misharina LA (1961) Aftershocks of the Srednebaikal earthquake of August 29, 1959. *Geol Geophys* 2: 105–110.
17. Tsydypova LR, Tubanov TsA, Sanzhieva DPD, et al. (2021) Kudarin earthquake on December 9 (10), 2020 with Mw = 5.5, I0 = 7 (Central Baikal, Russia) (in Russian). In: Proceedings of the VI All-Russian Youth Scientific Conference dedicated to the memory of Academician N.L. Dobretsov. Ulan-Ude: BSC SB RAS: 127–128. https://doi.org/10.31554/978-5-7925-0604-6-2021-127-128
18. Lunina OV, Gladkov AS, Sherstyankin PP (2010) A new electronic map of active faults for southeastern Siberia. *Dokl Earth Sci* 433: 1016–1021. https://doi.org/10.1134/S1028334X10080064
19. Radziminovich NA, Ochkovskaya MG (2013) Identification of earthquake aftershock and swarm sequences in the Baikal Rift Zone. *Geodyn Tectonophysics* 4: 169–186. https://doi.org/10.5800/GT-2013-4-2-0096

AIMS Press