Estimation of Probability Distributions of Geoacoustic Signal Characteristics

M I Gapeev¹, Yu I Senkevich¹, O O Lukovenkova¹

¹Laboratory of Acoustic Research, Institute of Cosmophysical Research and Radio Wave Propogation, Mirnaya 7, Paratunka, 684034, Kamchatka Region, Russia

E-mail: gm16493@gmail.com

Abstract. The paper presents an estimation of probability distributions of geoacoustic signal characteristics. The studied signals have a pulsed nature. The ones have been recording at the geodynamic test site of the IKIR FEB RAS (Kamchatka Peninsula) for more than 20 years. To estimate the distribution of characteristics, such time intervals were determined in which histograms of the distribution did not change. The following characteristics were chosen for the estimation: maximum amplitude, the position of pulse envelope maximum, duration, filling frequency, and pulse-to-pulse interval. The obtained estimates made it possible to develop an empirical model of the geoacoustic emission signal. The model can help to test new and existing algorithms for the processing and analysis of geoacoustic signals. The paper also shows that the formalization of the selected characteristics makes it possible to search for anomalies, including those associated with seismic events, by the characteristic variations.

1. Introduction

Acoustic emission in solids is an appearance of elastic vibrations generated by dislocation changes in the medium. In the case when rocks are used as a medium, such radiation is called “geoacoustic emission” (GAE). GAE arises as a result of changes in the stress-strain state of the geomedium in seismically active regions.

The IKIR FEB RAS has been studying of geoacoustic emission and its relationship with the seismic activity of the region since 1999 in the seismically active region of the Kamchatka Peninsula. GAE signals are pulse streams. In [1–8] it has been established that GAE anomalies are observed before strong earthquakes. These anomalies are expressed in changes in various signal characteristics: pulse repetition rate, their total number, frequency composition, shape, etc.

Abnormal changes in GAE signals are difficult to formalize due to their great variability. The variability is explained by the nonlinearity and chaotic state of a seismic process. To identify the signal anomalies, it is necessary to understand the typical GAE behavior. To solve this problem, the authors propose an empirical model of the GAE signal, built using statistical methods. In turn, the obtained model made it possible to develop a generator of GAE signals corresponding to a certain seasonal and weather activity. The generator is used in researches conducted in IKIR FEB RAS for testing algorithms of GAE signal processing and analysis.
2. Method to estimate the geoacoustic signal characteristics

The GAE signal pulses are preliminarily separated from the stream using a detector that implements the adaptive threshold detection algorithm. The detected pulses constitute the original sample for the estimation [9, 10].

The authors selected the following signal characteristics to construct an empirical model of the geoacoustic emission signal: maximum amplitude, the position of pulse envelope maximum, duration, filling frequency, and pulse-to-pulse interval. A probability distribution for each characteristic was estimated. The pulse frequency composition was analysed using the sparse approximation method. The authors considered the frequency of the function that has the highest correlation with the pulse. As shown in [11–16], this function carries the most information about the pulse frequency composition. Figure 1 schematically shows the estimated characteristics of the GAE signal.

![Figure 1. GAE signal characteristics.](image)

The pulse duration

A max is maximum amplitude, p max is the position of pulse envelope maximum (it is converted into a percentage of the pulse duration).

Samples of the characteristics were taken from the following observation periods: January, April, July, and September 2017. The histograms were compared to estimate the duration of the time interval in which the probability distributions of the signal characteristics do not change significantly. The method works as follows: estimation of GAE pulse characteristics of the initial signal and construction of the probability distribution histogram; then, increasing signal duration and reconstruction of the histogram for this signal; comparison of the resulting histograms. The comparison measure is the Euclidean norm of the histogram difference.

\[
E = \left( \sum_{i=1}^{N} |x_i - y_i|^2 \right)^{1/2},
\]

where \(x_i, y_i\) are the corresponding relative frequencies of the histograms, \(N\) is the number of histogram bins.

Then the process is repeated until the threshold of the Euclidean norm is reached. The authors choose the initial signal duration and the Euclidean norm threshold empirically. The duration of the initial time interval for estimation of the signal characteristics is 15 min. The step of increasing the time interval is 15 min. The threshold value of the norm (equation (1)) is 0.001.

Reaching the threshold value indicates that a further increase in the sample used to estimate the distribution does not significantly change this distribution. It means that we find an interval where the distribution of signal characteristics does not conditionally change. If we analyse the distributions of the pulse amplitudes and the pulse-to-pulse intervals, we can speak about the parametric stationarity of the GAE signal at the selected time interval.

The following time intervals for the probability distribution estimation were obtained: 1.5 days for the pulse amplitude, 17.5 hours for the pulse duration, 5 hours for the pulse-to-pulse interval, 1.5 days
for the pulse frequency, and 1.3 days for the position of pulse maximum. Figure 2 shows the histograms of the characteristic distributions built on the obtained intervals.

3. Geoacoustic signal model
Based on the obtained statistical estimates of the above distributions, the authors propose the following model of the GAE signal:

\[ s(t) = \sum_{i=1}^{N} A_i \cdot g_i(t + \tau_i), \]  

where \( N \) is the number of pulses in the signal; \( A_i \) is the amplitude of the \( i \)-th pulse; \( g_i(t) \) is an analytical model of the \( i \)-th pulse; \( \tau_i \) is the generation time of the \( i \)-th pulse, calculated based on the distribution of pulse-to-pulse intervals. Each pulse can be analytically represented by the Berlage function:

\[ g(t) = A \cdot t^{n(t_{end}, P_{max})} \cdot \exp\left( \frac{n(t_{end}, P_{max}) \cdot \Delta}{P_{max} \cdot t_{end}} \right) \cdot \sin(2\pi ft), \]

where \( P_{max} \) is the position of the pulse maximum, \( t_{end} \) is the pulse duration, \( f \) is the pulse frequency, \( n(t_{end}, P_{max}) \) and \( \Delta \) are the parameters affecting the pulse envelope steepness.

To generate a sequence of random numbers according to the obtained estimates of the distributions (figure 3), the stepwise approximation method was used [17–20].

4. Discussion of the practical research results
The empirical model of the characteristic signal of the GAE as in equation (2) allowed to determine the value of the deviation from the corresponding parameters using the criterion represented in equation (1). We can draw conclusions about the appearance of abnormal changes in the signal by comparing these deviations with previously defined norms. For example, histograms of the probability distribution of pulse duration in the background and abnormal periods and a graph of the change in the norm of histogram difference are set in in figure 2(a) and figure 2(b), respectively. The norm calculated by criterion in a sliding window with a width of 15 minutes begins to grow from a certain moment, this may indicate the occurrence of an anomaly in the GAE signal.

![Figure 2](image-url)

**Figure 2.** An example of histograms of the pulse duration distribution in the background (blue) and abnormal (orange) periods (a), as well as a graph of the change in the histogram difference (b).
Figure 3. Estimating the distribution of GAE signal characteristics: the amplitude maximum value (a), the pulse duration (b), the frequency (c), the duration of pulse-to-pulse interval (d), the position of maximum (e).

By introducing an empirical threshold of the difference norm of distribution histograms, we can determine the deviation degree of the current signal from the background one and use this result to identify anomalies.
5. Conclusion
As a result of the study, the main GAE signal characteristics were analysed, the probability distributions of these characteristics were estimated on selected time intervals. The geoacoustic pulse signal model and algorithm for simulating the GAE background signals were developed.

The authors formulate the operating principle of the anomaly detector for GAE signals. This detector analyses changes in the probability distributions of signal characteristics. The development of an automatic anomaly detector and empirical rules for anomaly classification are topics for further research.

6. References
[1] Gregori G, Poscolieri M, Paparo G, De Simone S, Rafanelli C and Vetrice G 2010 “Storms of crustal stress” and AE earthquake precursors *Natural Hazards and Earth System Sciences* **10** 319–337
[2] Marapulets Yu and Shevtsov B 2012 Mesoscale acoustic emission (Vladivostok: Dalnauka) p 126
[3] Marapulets Yu, Solodchuk A and Shcherbina 2016 A Changes of geoacoustic emission directivity at “Mikizha” site associated with earthquakes in Kamchatka *E3S Web of Conferences* **11** 00014
[4] Kuptsov A 2005 Variations in the geoacoustic emission pattern related to earthquakes on Kamchatka *Izvestiya. Physics of the Solid Earth* **41** 825–831
[5] Marapulets Yu, Shevtsov B, Larionov I, Mishchenko M, Shcherbina A and Solodchuk A 2012 Geoacoustic emission response to deformation processes activation during earthquake preparation *Russian Journal of Pacific Geology* **6**(6) 457–464
[6] Larionov I, Marapulets Yu and Mishchenko M 2019 Results of atmospheric-lithospheric observations of acoustic radiation in Kamchatka *E3S Web of Conferences* **127** 02023
[7] Shcherbina A, and Solodchuk A 2019 Estimation of the power of geoacoustic emission sources registered in Mikizha Lake, Kamchatka region *E3S Web of Conferences* **127** 03003
[8] Lukovenkova O, Malkin E, Mishchenko M and Solodchuk A 2021 Anomalies in electromagnetic and geoacoustic emission signals before Kamchatka earthquakes (ML ≥ 4.75) in 2013 *Vestnik KRAUNC. Fiz.-mat. nauki* **34**(1) 137–149
[9] Tristanov A, Lukovenkova O and Klionskiy D 2018 Adaptive approach to signal analysis (through the example of geoacoustic emission) *Izvestiya SPbGETU “LETI”* **7** 79–97
[10] Marapulets Yu, Senkevich Y, Lukovenkova O, Solodchuk A, Larionov I, Mishchenko M, Malkin E, Shcherbina A and Gapeev M 2020 Complex analysis of acoustic and electromagnetic signals to assess seismic hazard level (Vladivostok: Dalnauka) p 120
[11] Lukovenkova O, Senkevich Y, Solodchuk A, and Shcherbina A 2020 Overview of processing and analysis methods for pulse geophysical signals *E3S Web of Conferences* **196** 02023
[12] Kim A, Marapulets Yu, Tristanov A and Lukovenkova O 2017 Methods of registration and time-frequency analysis of geoacoustic emission signals emission (Vladivostok: Dalnauka) p 148
[13] Ebenezer S, Papandreou-Suppappola A and Suppappola S 2004 Classification of acoustic emissions using modified matching pursuit *EURASIP Journal on Applied Signal Processing.* **3** 347–357
[14] Lukovenkova O, Marapulets Yu, Tristanov A, Kim A and Kashutina I 2018 Optimization of adaptive matching pursuit method to analyze geoacoustic emission signals *Vestnik KRAUNC. Fiz.-mat. Nauki* **4**(24) 197–207
[15] Mallat S and Zhang Z 1993 Matching pursuits with time-frequency dictionaries *IEEE Transactions on Signal Processing* **41**(12) 3397–3415
[16] Lukovenkova O and Solodchuk A 2020 Analysis of geoacoustic emission and electromagnetic radiation signals accompanying earthquake with magnitude Mw = 7.5 *E3S Web of Conferences* **196** 03001
[17] Bykov V 1971 Digital modeling in statistical radio engineering (Moscow: Sovetskoe radio) p 328
[18] Kramer C 1975 Mathematics methods of statistics (Moscow: MIR) p 648
[19] Kobzar A 2006 Applied mathematical statistics For engineers and scientists (Moscow: Fizmatlit) p 816
[20] Ivchenko G and Medvedev Yu 2014 Mathematical statistic (Moscow: LIBROCOM) p 352