CLUSTER ANALYSIS OF LIGHTNING DISCHARGES: BASED ON VEREYA-MR NETWORK DATA

I.D. Tkachev
Institute of Solar-Terrestrial Physics SB RAS, Irkutsk, Russia, tid007@iszf.irk.ru

R.V. Vasilyev
Institute of Solar-Terrestrial Physics SB RAS, Irkutsk, Russia, roman_vasilyev@iszf.irk.ru
Irkutsk State University, Irkutsk, Russia

E.P. Belousova
Institute of Solar-Terrestrial Physics SB RAS, Irkutsk, Russia, elenapbel@iszf.irk.ru
Irkutsk State University, Irkutsk, Russia

Abstract. Monitoring thunderstorm activity can help you solve many problems such as infrastructure facility protection, warning of hazardous phenomena associated with intense precipitation, study of conditions for the occurrence of thunderstorms and the degree of their influence on human activity, as well as the influence of thunderstorm activity on the formation of near-Earth space. We investigate the characteristics of thunderstorm cells by the method of cluster analysis. We take the Vereya-MR network data accumulated over a period from 2012 to 2018 as a basis. The Vereya-MR network considered in this paper is included in networks operating in the LF-VLF range (long and super-long radio waves). Reception points equipped with recording equipment, primary information processing systems, communication systems, precision time and positioning devices based on global satellite navigation systems are located throughout Russia. In the longitudinal-latitude thunderstorm distributions of interest, the dependence on the location of recording devices might be manifested. We compare the behavior of thunderstorms on the entire territory of the Russian Federation with those in the Baikal natural territory. We have established the power of thunderstorms over the Baikal region is lower. The daily variation in thunderstorm cells we obtained is consistent with the data from other works. There are no differences in other thunderstorm characteristics between the regions under study. This might be due to peculiarities of the analysis method. On the basis of the work performed, we propose sites for new points of our own lightning location network, as well as additional methods of cluster analysis.

Keywords: thunderstorms, cluster analysis, lightning discharges.

INTRODUCTION

Forecasting extreme weather events, associated with the development of convective instability in the atmosphere, remains relevant to this day in terms of the problem of global climate change [Konstantinova, Gorbatenko, 2016]. The formation of thunderstorms and showers is the result of a complex interaction of atmospheric processes with the underlying surface. Despite a fairly developed network of meteorological stations equipped with modern equipment, the current level of forecasting such dangerous phenomena as thunderstorms is still at the forecast success rate 50–70 %. Horizontal dimensions of convective zones are many times smaller than distances between stations at which they are recorded. Coverage of Earth’s surface by observation stations is far from uniform.

The development and use of a lightning location network will be of interest for solving practical problems of monitoring and preventing dangerous phenomena associated with heavy rainfall and atmospheric electricity discharges: forest fires, floods, mud flows, damage to energy and transport infrastructure, etc. Monitoring lightning activity is important for timely protection from damage to power lines with the aid of lightning rods, for the safety of air travel, marine guidance, and for research in such areas as meteorology, hydrology, geology, etc. [Panyukov et al., 2003]. The range of scientific problems that can be solved by analyzing data from networks of lightning direction finders is quite wide, which is relevant for scientific research carried out by various organizations. These devices can be used to examine thunderstorms, conditions of their occurrence, the degree of their influence on human activity, as well as properties of convective clouds exposed to artificial influence [Sinkevich et al., 2018], the influence of lithospheric processes on radio wave propagation [Argunov, 2018], etc. Thunderstorm activity plays an essential role in the formation of near-Earth space. Lightning location networks provide factual data on the influence of processes caused by solar-terrestrial relations, on daily and seasonal atmospheric electric field variations in regions. There are a number of papers showing that after lightning discharges electric fields of unipolarly charged clouds cause the electron density to increase at altitudes 95–100 km, which should lead to a
noticeable increase in positioning errors in global navigation satellite systems (GNSS) [Mareev et al., 2019]. The atmospheric electricity phenomenon occupies a significant place in space weather research [Yang et al., 2018], and there are proposals to use lightning location networks for studying the effect of electromagnetic pulses of lightning discharges on solar wind electrons trapped in the geomagnetic field [Blioh, 1997; Ripoll et al., 2019].

Lightning location networks generally consist of pulsed radio signal receivers of the same type, located on Earth’s surface at some distance from each other [Panyukov et al., 2003; Kononov, Yusupov, 2004; Shabaganova, Kozlov, 2010; Shabaganova et al., 2012; Moskovenko et al., 2012]. The distance depends on the frequency band of radio waves in hand — VLF or VHF. Networks exploiting the VLF band are composed of nodes spaced hundreds and thousands of kilometers apart. This allows monitoring of thunderstorm activity over large areas without significant equipment costs. However, significant limitation on the use of these networks is the low sensitivity to intercloud and intracloud discharges, which are quite successfully recorded in the VHF band. Due to strong attenuation and distortion of the phase picture of radio emission in this range when propagating over long distances, lightning location points comprising a VHF network should be in the direct visibility. VHF devices can record intracloud and intercloud discharges and perform precise positioning of discharges, and yet a great number of network nodes are required to cover a large area. An important factor for successful work is a system for synchronizing network nodes. Detecting a short (millisecond) electromagnetic pulse requires time counters in the recording devices to be synchronized with an accuracy that significantly (at least by an order of magnitude) exceeds the propagation time of the electromagnetic signal between individual network nodes ($3\times 10^8$ s). In modern networks, this problem is addressed using satellite navigation systems [Tarabukina, Kozlov, 2018].

The lightning location network Vereya-MR [Moskovenko et al., 2012] is a network operating in the LF/VLF range (long and super long radio waves). Reception points equipped with recording equipment, primary information processing systems, communication systems, precision time and positioning devices using GNSS are located throughout Russia (see Figure 2 in [Moskovenko et al., 2012]). The work is coordinated by a single center, which amasses on-line data from reception points to calculate the location of lightning discharges and to maintain the database of discharges for onward transmission to users.

The objectives of this work are to study the possibility of using the Vereya-MR network data to obtain physical characteristics of thunderstorm cells (propagation velocity, duration, direction of thunderstorm movement, etc.); to collect information about thunderstorm activity in the Baikal natural territory; to search for local features in order to determine the optimal sites for points of the new lightning location network.

**METHOD AND DATA IN USE**

Since thunderstorms are natural phenomena, to some extent localized in time and space, it will be useful to analyze not individual lightning discharges recorded by the network, but their combination observed within a certain spatial region for a finite time interval both to reduce the amount of information processed and to obtain characteristics of thunderstorms. For this purpose, it is convenient to analyze clusters of lightning discharges. Cluster analysis is one of the methods for generalized description of an assembly of many random events. It is employed to isolate and describe events grouped according to some characteristic, in this case according to the concentration of electric discharges in a certain bounded spatial region within a thunderstorm formation (cluster, thunderstorm cell). Observing a cluster thus identified, we can determine its parameters — power, area, and speed. Study of thunderstorms with cluster analysis is not new and has been carried out in papers by Kononov I.I. [Kononov, Yusupov, 2004], Shabaganova S.N. [Shabaganova, Kozlov, 2010; Shabaganova et al., 2012], and others, from which we have taken the main ideas for conducting the cluster analysis based on the Vereya-MR network data. Initial parameters of the Vereya-MR network are geographical coordinates of a discharge, the time of its occurrence up to a millisecond, and the maximum strength amplitude of the electrical component of the electromagnetic pulse from the lightning discharge. This information is sufficient to localize groups of events of this kind in space and time.

The key point of the study is to choose the most appropriate maximum distance between discharges, required to determine their belonging to the same cluster. Given that the mean speed of thunderstorm cells is within 30–40 km/hr, and the average radius of a thunderstorm cell is ~20–40 km [Scholkmann et al., 2012; Sinkevich et al., 2018], a condition was formulated that at a current moment of time a thunderstorm cell is a combination of lightning discharges, in which at least six discharges occur for an hour, with a maximum distance between them of no more than 75 km. Since we explore general statistical characteristics without requiring accurate coordinate reference (up to a kilometer), there is no need to use an accurate model of Earth (geoid). It is assumed that Earth is spherical, so the distance between discharges is calculated using the well-known formula

$$\cos(d) = \sin(\phi_1)\sin(-\phi_2) + \cos(\phi_1)\cos(-\phi_2)\cos(\lambda_1 - \lambda_2),$$

where $\phi$ is the latitude; $\lambda$ is the longitude; $d$ is the distance between points in radians, which in the case of spherical Earth can be quite simply converted into kilometers: $b = dR$, where $R$ is the Earth radius. Analyzing thereby a set of lightning discharges, which occurred for an hour relative to one discharge, and repeating this procedure for each discharge from the entire database, we can identify thunderstorm clusters that live from several minutes to several hours. The resulting clusters can be characterized in space and time according to their location — the global center of mass of the cluster.
and the times of their occurrence and end — the times of occurrence of the first and last discharges in the cluster respectively. Moreover, we can introduce the concept of the power of a thunderstorm cluster as the number of thunderstorm discharges per unit time recorded during its lifetime. The difference in the location of the center of mass of the cluster during its lifetime characterizes its displacement.

DATA PROCESSING RESULTS

The available Vereya-MR database of lightning discharges contains records from 2012 to 2018. For this period, we have obtained a set of thunderstorm clusters by the method described above. We assume that each cluster in this set represents a separate thunderstorm cell. While the cluster selection algorithm may be insufficiently effective and may contain some inaccuracies, the information collected on its basis should correctly characterize properties of thunderstorm cells. Characteristics of the thunderstorm cells obtained were analyzed both for the entire coverage area of the lightning location network and for the Baikal region — this is a region with coordinates 100°–113° E and 50°–57° N.

Longitude-latitude dependence

Spatial location of a thunderstorm cell can be defined as the mean of coordinates of all lightning discharges in a given cluster. For the regions of interest, we plotted latitude and longitude distributions of clusters (Figure 1).

Over longitude, there is an uneven distribution of thunderstorm clusters. Such a situation may be associated with different density of space coverage by receiving antennas of the lightning location network, which may cause continuous operation of the system to disrupt. At the same time, the uneven distribution of thunderstorm clusters may be due to features of the relief: at longitudes, where the number of clusters decreases, there are mountain ranges: the Ural Mountains (60° N, 60° E) and Eastern Sayan (53°49′ N, 97°35′ E), which can influence the formation of thunderstorm cells. The coverage area of the Vereya-MR network is localized in the range from 43° N (Vladivostok) to 64° N (Arkhangelsk) (the location of receiving stations is shown in more detail in Figure 2 in [Moskovenko et al., 2012]), which leads to the latitudinal gaussian distribution of clusters observed (Figure 1, a) with a maximum around 55° N. For the same reason, it is impossible to see an increase in the number of clusters as they move equatorward. If we look at the map of the coordinate dependence of the number of thunderstorm clusters (Figure 2, a), we can distinguish the following features:

1. When moving to the east, the maximum of thunderstorm clusters shifts toward lower latitudes: from 60° in the west in the vicinity of the Moscow region to 50° in the Far East.

2. As for the longitude distribution of clusters (Figure 2, a), there are three regions where the number of clusters increases: at 40°, from 80° to 90°, and around 130° E.

3. In the vicinity of Moscow and the Moscow region, the number of thunderstorm cells is maximum. This may be due to the largest number of recording devices as well as their long uptime.

For the Baikal region, the longitude distribution of thunderstorm clusters repeats the trend of distribution of clusters across the country (Figure 1 b, blue contour). In the latitude distribution, there are deviations from the general picture: Across the country, clusters are evenly distributed from 50° N to 60° N, while in the Baikal region to the north the number of thunderstorms decreases. Figure 2, b shows that in the Baikal region there is an area 50°–53° in latitude and 100°–104° in longitude, where the number of clusters is greater than in the rest of the region. The Sayan foothills are mainly located in this area. In the region where the number of clusters is lower or they are complete absent is Lake Baikal (51°–55° N, 103°–109° E). This may indicate that thunderstorms are formed more actively in foothills than over water bodies. World statistics of thunderstorms shows that they are formed over oceans by 30 % less frequently than over land [Rycroft et al., 2000].

Figure 1. Distribution of thunderstorm clusters over latitude (a) and longitude (b) for the period 2012–2018: green contour — over the entire territory of the Russian Federation (Y-axis on the left), blue contour — over the Baikal natural territory (Y-axis on the right); henceforward, N is the number of clusters
Cluster analysis of lightning discharges

Figure 2. Distribution of clusters throughout the Russian Federation (a) and over the Baikal natural territory (b) for 2012–2018. The color intensity indicates the number of clusters.

Determining territorial maxima of thunderstorm activity more accurately and analyzing these areas in detail require:

a) reconstruction of regular observations of thunderstorms in the region;

b) an effective network of several lightning location receivers located in the region.

Temporal characteristics of thunderstorm clusters, size, movement

To estimate the size of a thunderstorm cell, it is necessary to calculate the distance between lightning discharges and the cell’s center of mass. The center of mass means a point with averaged coordinates of all lightning discharges over longitude and latitude for a given cluster. The cluster size is defined as the mean of these distances. From the distribution of values (Figure 3, a) we can determine that thunderstorm cells have a radius from 30 to 40 km, which is consistent with the data obtained in [Shabaganova et al., 2012]. Cluster displacement is the distance between mean coordinates of the first five and the last five discharges respectively. The mean displacement of a thunderstorm cluster, calculated by Formula 1, was about 40 km (Figure 3, b).

Figure 3, c shows local time distribution of thunderstorm clusters. The resulting distribution is consistent with the main temporal features of the dynamics of thunderstorms: low thunderstorm activity is observed...
I.D. Tkachev, R.V. Vasileyev, E.P. Belousova

During the pre-dawn hours; and increased one, in the afternoon. Most thunderstorm cells in the territory of Russia are formed in the interval from 15 to 17 hrs. The mean duration of thunderstorms is ~1 hr (Figure 3, d), which is also in line with current views. At the same time, there are thunderstorm clusters with a lifetime from 4 to 6 hrs, whose phenomenon for the database and clustering algorithm in hand calls for additional studies.

In the Baikal region, no special differences in the characteristics under study are observed. This suggests that in terms of the formation and course of thunderstorms in the region they do not differ much from thunderstorms in the entire territory of the Russian Federation. It should be borne in mind, however, that such similarity between the characteristics may appear if a clustering algorithm depends on cluster parameters, so additional statistical studies of the database of lightning discharges are required, which would focus on the choice of optimal values of cluster parameters (duration, size, and number of lightning discharges).

Determining the direction of thunderstorm clusters

Velocity and direction of a thunderstorm cell depend on the direction of the wind and, first of all, on the interaction of the upward and downward streams of the cloud with the carrier air currents in middle atmospheric layers, where the thunderstorm develops. In the mid-latitude troposphere and stratosphere, the west-to-east air transport prevails — westerlies. The direction of thunderstorm cells is determined on a plane where coordinates of the beginning and end of the cluster are calculated for each cluster (as in the case with the determination of cluster displacements). Then, we can find projections of cluster displacement onto the latitude and longitude axes, and determine the movement angle relative to the north (geographic azimuth) in the resulting triangle. The spatial movements of the thunderstorm cell are small enough in comparison with the Earth radius, so a good approximation for solving this problem would be the use of simple trigonometric identities. Further, the directions of movement of clusters are averaged for one day. Figure 4 shows velocity distributions of clusters, as well as the wind rose for directions of clusters. The direction of thunderstorm clusters is affected by westerlies; and their prevailing direction is northeastward and eastward (Figure 4, b). For the Baikal region, the peculiarity is that the eastward direction of thunderstorm cells prevails (Figure 4, a). At the same time, there are thunderstorm clusters moving southwestward, but they make up a small fraction of the total.

Power of thunderstorm clusters

In this study, by the power of thunderstorm clusters is meant the number of lightning discharges per unit time. Features of power distribution of thunderstorm cells are illustrated in Figure 5. A characteristic feature of the distributions is the power spectrum slope. For the Baikal natural territory, the distribution slope is more pronounced (Figure 5, b); this may indicate that the power of thunderstorms is significantly lower there than in the rest of the Russian Federation. According to the map (Figure 5, c), we can say that the location of severe thunderstorms coincides with the zones of an increased number of thunderstorm formations. It is necessary to understand why, according to our findings, thunderstorms are less powerful in the Baikal region, given that their duration is comparable to the duration of thunderstorms throughout the country. Answering this question also requires additional statistical studies of the database of lightning discharges, focusing on the choice of optimal values of parameters of a cluster, namely its lifetime.

Figure 4. Azimuth distribution of thunderstorm cells over the Baikal natural territory (a) and throughout the Russian Federation (b), as well as velocity of thunderstorm cells (c): green color — throughout the Russian Federation (Y-axis on the left); blue, over the Baikal natural territory (Y-axis on the right). Data is for the period 2012–2018.
DISCUSSION

In this work, we have demonstrated that the Vereya-MR network data can be used to study thunderstorm cells. The cluster parameters obtained are consistent with the modern understanding of physical parameters of thunderstorms. The established spatial dependences of the location of thunderstorm cells can be associated with both the relief features and the location of network stations. Local daily thunderstorm activity, duration, speed, size, and movement of thunderstorm clusters for the entire territory of Russia and for the Baikal region do not differ greatly. Differences are observed only in the power of thunderstorm formations. At the same time, the study has shown the need for additional verification of the clustering method in order to confirm that the physical characteristics of thunderstorms are independent of clustering parameters.

One of the ways to solve this problem is to conduct cluster analysis, varying the temporal parameter (half an
hour, an hour, two hours, etc.). Then, we have to compare the results to exclude or neutralize the effect of such dependence. An alternative study is to conduct a spatial-temporal correlation analysis of the database of lightning discharges to determine characteristics of thunderstorm clusters. Studying the asymmetry of clusters and the dynamics of this asymmetry in terms of the dynamics of atmospheric fronts, as well as the size and shape distribution of clusters to determine the relationship between frontal and intramass thunderstorms — these and other expected problems are addressed using the cluster analysis and are directly related to hazardous weather events in the Baikal natural territory.

We are planning follow-up studies to examine thunderstorm activity over the Baikal natural territory in more detail, to compare several methods of cluster analysis, to study characteristics of thunderstorm cells more thoroughly, as well as experiments on observing thunderstorms, using fast photometric devices, in order to search for high-altitude lightning discharges. To address the formulated problems, it is necessary to develop an effective lightning location network in the Irkutsk and Chita regions and in the Republic of Buryatia. The location of points for detecting lightning discharges, proposed on the basis of the findings of this study, is shown in Figure 6.

![Figure 6. Estimated location of lightning location points in the Irkutsk and Chita regions and in the Republic of Buryatia](image-url)
CONCLUSION

The cluster analysis of data from the lightning location network Vereya-MR we have carried out allowed us to obtain the main characteristics of thunderstorm clusters. The study has shown that the most active genesis region of thunderstorms in the Baikal natural territory are mountainous areas, mostly the foothills of the Eastern Sayan, located in the southwest of the Baikal natural territory. The spatial and temporal characteristics of thunderstorm cells in the Baikal natural territory obtained with the aid of the Vereya-MR lightning location network do not differ from the characteristics of thunderstorm cells throughout the Russian Federation. The power of thunderstorms is lower however, which suggests that thunderstorm characteristics depend on the method of identifying thunderstorm clusters in the database of lightning discharges. Further research is required before applying this method to processing of data from the new lightning location network deployed in the Baikal natural territory. It is impossible to use the Vereya-MR network for this purpose because of the rare location of lightning location points and its decommissioning in 2019. From the analysis of existing lightning location networks, we have found out that for successful monitoring of thunderstorm activity in the Baikal natural territory lightning location points can be located at geophysical observatories of ISTP SB RAS (Irkutsk) and IPM SB RAS (Ulan-Ude). It would be most optimal to develop a multipoint lightning location network of broadband VLF radio receivers since such devices are the most simple and low-maintenance. Modern software and hardware enable us to create a simple and reliable receiver capable of pre-processing oscillograms of pulses from lightning discharges of all possible types and transmitting information via existing mobile networks to the information processing center.

The work was financially supported by the Ministry of Science and Higher Education of the Russian Federation (Subsidy No. 075-GZ/Ts3569/278) in terms of research on processing and analysis of lightning discharge data and by grant No. 075-15-2020-787 for implementation of a large scientific project in priority areas of scientific and technological development (the project “Fundamentals, Methods and Technologies for Digital Monitoring and Forecasting of the Ecological Situation of the Baikal Natural Territory”) in terms of research on the development and location of estimated lightning location points in the Irkutsk Region.

REFERENCES


How to cite this article