Classification analysis of the meteotropic reactions of people depending on the intensity of the electric field of the atmosphere

K Kh Mashukov, A A Tashilova*, L A Kesheva, I Kh Mashukov
High-mountain Geophysical Institute, 2 Lenin Ave., Nalchik, Russian Federation

E-mail: hazret53@rambler.ru

Abstract. The results of the study of the influence of the electric field intensity in the atmosphere on the frequency of exacerbations of cardiovascular diseases are presented on the example of the population of Nalchik. The seasonal trends of reducing both the average tension and the number of ambulance calls for diseases has been revealed. It was obtained that the challenges for disease are grouped into homogeneous factor groups. The group combining hypertensive crisis, cardiovascular diseases and angina pectoris (component 1) has a positive statistically significant dependence on the magnitude of the average daily intensity of the electric field of the atmosphere near the surface of the earth. According to the results of cluster analysis, the days with the maximum dependence between factor 1 and variations of the electric field intensity in the atmosphere were revealed.

Introduction
Meteorological factors, including electromagnetic fields, which are generated by various meteorological processes and propagate almost without attenuation over long distances, are one of the important environmental factors that influence people's well-being.

This paper presents the results of statistical studies that, in the chaotic and unpredictable effect of the electric field of the atmosphere on people with cardiovascular diseases, will help to detect some general patterns that are important from a medical-physical point of view.

The electric field of the atmosphere is an important component of the atmosphere, which is physically a quasi-static electric field created by the ionospheric charge and the Earth’s own charge. With the passage of clouds, due to the electrical processes occurring in them, sharp changes in the electric field value are recorded, up to a change in the direction of the vector, the electric field intensity can reach 1000 V/m or more, which adversely affects the adaptive properties of a living organism.

In the works devoted to this research topic, attention has already been drawn to the fact that people react sensitively to the slightest changes in the synoptic situation and to the contrast in the change of weather. In the works [1-3] the results of studies on the surface air temperature, pressure and humidity of the atmosphere are given. The authors noted that the risk of developing exacerbations of hypertension increases with a decrease in air temperature and an increase in atmospheric pressure and relative humidity.
Materials and methods
Most of the relationships between diseases and field intensity are non-linear and multifactorial, which requires the use of adequate methods of analysis. In our study to identify the dependence of the frequency of exacerbations of diseases on changes in the electric field intensity, the study was carried out in several stages using regression factor and cluster analysis. The calculations were carried out using the statistical packages MINITAB 14.12 and SPSS 21.0 [4].

In the period from 2006 to 2008 in Nalchik (Kabardino-Balkarian Republic, Russia), monitoring was conducted to record the gradient of the potential of the electric field of the atmosphere at the Earth's surface using the recording apparatus of the FSBU “VGI” (Figure 1). The complex included: electrostatic fluxmeter, secondary converter, analog-to-digital converter (ADC), computer, a set of cables. As a measuring instrument in the complex, the “Field-2” electric field intensity sensor, developed at the Main Geophysical Observatory named after A.I. Voeikov (MGO), is used. The sensor of the electric field “Field-2” is a measuring transducer and, complete with recording equipment, is designed to measure the electric field of the atmosphere near the surface [5].

Figure 1. Block - diagram of the complex apparatus for measuring the intensity of electric field
1- electrostatic fluxmeter; 2- secondary converter;
3 - analog-to-digital converter (ADC); 4 - computer

During the same period, average daily data on the number of visits for the following diseases were collected and systematized: hypertension, hypertensive crisis, general myocardial infarction, cardiovascular pathology, angina pectoris, arrhythmia (recorded calls from the ambulance station of Nalchik).

Results
The authors in [6, 7] began studies on the identification of hidden patterns of the effect of changes in the intensity of the electric field of the atmosphere on the exacerbation of hypertensive and cardiovascular diseases in the period 2005–2006. This paper presents the results that continue the research begun.

In a preliminary analysis based on the collected data for a specified period, the time course of calls and intensity from January to June 2007 was obtained, trends with regression equations were constructed, from which it followed that there was a seasonal decrease in intensity, both the average tension and the number of ambulance calls help for diseases.

Next, descriptive statistics were obtained for the number of calls for each disease, the electric field intensity (Table 1), and also a correlation analysis was made to identify the relationship between the quantities studied.

Table 1. Descriptive statistics of the intensity of the electric field of the atmosphere and the frequency of exacerbations of diseases according to ambulance data for January-June 2007
As can be seen from table 1, the greatest average daily number of calls falls to hypertension - 33, the smallest - to myocardial infarction, an average of 1 call per day. The total sum of calls is maximum for hypertension 5931 and minimum for myocardial infarction 179 calls in six months. Mode characterizes the most common number of calls per day, for hypertension, 29 calls, for myocardial infarction, 0 calls, which means that there are many days without calls for this disease. The coefficients of asymmetry (Skewness) and excess (Kurtosis) are rather small (<1), which indicates the normal distribution of the series in the number of calls and the possibility of applying the Pearson correlation analysis. The exception was calls for vascular diseases: they have right-sided asymmetry and great excess.

Let us find the correlation coefficients between the average daily intensity of the electric field of the atmosphere (V/m) and the number of calls per day for the above-mentioned diseases. The obtained correlation coefficients between the average daily intensity and the number of calls per day are very small and vary from 0.026 to 0.257. However, there is a significant correlation (at a significance level of p = 0.05) between some diseases, and this is natural, since they belong to the same type of diseases.

From Figure 2 it can be seen that on some days, for example, February 7, 12, and 20, there is a maximum of calls for certain groups of diseases, and vice versa, on some days there is a minimum number of calls.
Suppose that there is a hidden pattern of dependence of the number of calls of certain groups of diseases on the variation of the electric field of the atmosphere these days. We will try to group diseases into homogeneous factors according to the existing connections between them. To accomplish this task, we apply factor analysis using the capabilities of the statistical program SPSS 21.0.

Factor analysis is a procedure by which a large number of variables are reduced to a smaller number of independent quantities (factors). In one factor, several variables are combined that have a tight correlation between themselves and a weak correlation with variables that are combined by other factors. After the array is preliminarily checked for suitability of the source data (tests KMO and Barlett), standardization of the specified values of variables is performed and the correlation coefficients between the considered variables are calculated, which serves as the initial matrix for grouping the parameters into factors. To construct the correlation matrix, the eigenvalues of characteristic numbers and eigenvectors are determined, for the determination of which diagonal elements of the matrix are used. To solve this problem, the principal component method is used. The eigenvalues are sorted in descending order, for which as many factors are usually selected as there are eigenvalues exceeding one. In Table 2, the eigenvalue of the characteristic number, equal to 1.013, corresponds to the optimal number of components in the factor model 3. Respectively, the factor model consisting of 3 factors saves 65.068% of the initial information in three components.

Table 2. Total Variance Explained

| Component | Total | % of Variance | Cumulative [%] | Total | % of Variance | Cumulative [%] | Total | % of Variance | Cumulative [%] |
|-----------|-------|---------------|----------------|-------|---------------|----------------|-------|---------------|----------------|
| 1         | 1.631 | 27.19         | 27.19          | 1.631 | 27.19         | 27.19          | 1.585 | 26.417        | 26.417         |
| 2         | 1.260 | 20.99         | 48.187         | 1.260 | 20.999        | 48.187         | 1.172 | 19.530        | 45.947         |
| 3         | 1.013 | 16.88         | 65.068         | 1.013 | 16.881        | 65.068         | 1.147 | 19.121        | 65.068         |
Extraction Method: Principal Component Analysis.

The next step in presenting the results of the factor analysis is a rotated matrix of components obtained using the method of rotating factors (varimax) (Table 3). After rotation, the matrix becomes noticeably more convenient for analysis - in it almost all factor loadings are either small or large and, therefore, such a matrix is easier to interpret. From Table 4 it can be seen that the number of calls is grouped according to three factors, within each factor there are diseases with maximum factor loadings. In factor 1 - hypertensive crisis, cardiovascular pathology and angina pectoris (factor loadings 0.706; 0.559; 0.775), in factor 2 - arrhythmia (factor load 0.878) and in factor 3 - hypertension and myocardial infarction (factor loads - 0.572; 0.860).

|               | Component 1 | Component 2 | Component 3 |
|---------------|-------------|-------------|-------------|
| Hypertension  |             |             |             |
| Hypertensive crisis | 0.706       |             | -0.572      |
| Cardiovascular pathology |             | 0.860       |             |
| Myocardial infarction | 0.559       |             |             |
| Angina pectoris    | 0.775       |             |             |
| Arrhythmia        |             |             | 0.878       |

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

Let us now find out the connection of the intensity of the atmospheric field with the obtained factors. The results of the correlation analysis for the three factors and the electric field intensity are presented in Table 4.

Table 4. Correlations of the components of the factor model with the average daily electric field intensity of the atmosphere

|       | Component 1 | Component 2 | Component 3 |
|-------|-------------|-------------|-------------|
| Intensity |             | Pearson Correlation |             |             |
|       | 1.000       | 0.218       | 0.094       | 0.066       |
|       | Sig. (2-tailed) | 0.000*   | 0.041       | 0.382       | 0.542       |
| N     | 88          | 88          | 88          | 88          |

* Correlation is significant at the 0.05 level (2-tailed).

When the significance level was p = 0.05, a statistically significant correlation coefficient r = 0.218 (with Sig. = 0.041 <0.05) occurred for factor 1 (hypertensive crisis, cardiovascular pathology and angina pectoris) and average daily electric field intensity. It should be noted that the correlation coefficient did not exceed 0.218, which indicates a moderate relationship between the compared values.

To further determine the conditions for the maximum dependence of the meteotropic response of patients (factor 1) on changes in the electric field, it is necessary to classify objects, namely: divide the set of days into homogeneous subsets with peaks in the number of calls and changes in the atmospheric field near the earth. The study of the potential gradient of the electric field of the atmosphere at the surface of the earth gives us this opportunity, since the registration of changes in
tension was made every 3 minutes during the day for three years (2006–2008). To identify such days it is necessary to conduct a cluster analysis (CA) for the period January-June 2007.

According to the results of CA, three data sets were formed for the first, second and third clusters. The first cluster included 29 days, the second cluster - 33 days and the third cluster - 25 days with the desired characteristics (the maximum correlation of exacerbations of diseases from the intensity of the electric field of the atmosphere).

It should be noted that the direct use of variables in the analysis can lead to the fact that the classification will determine the variables with the greatest variation of values, as in our case: the magnitude of the electric field varies from 0 to 1500 V/m, and the value of the number of calls has the maximum value - 64 calls per day (hypertension). Therefore, with differently scaled compared values of two or more rows, Z-standardization (Z-Scores) is used for analysis. The standardized value is the deviation of \( x_i \) from its average value divided by the standard deviation:

\[
Z_{x_i} = \frac{x_i - \bar{x}}{\sigma} \quad \text{.(1)}
\]

With an acceptable error level of \( p = 0.01 \) (99% confidence interval), the standardized value of \( Z_{x_i} \) varies in the range: \(-3 < Z_{x_i} < 3\). Accordingly, it is easy to compare different-scale values:

1) at \( Z_{x_i} \rightarrow 0 \), the value of \( x_i \) is close to the average \( \bar{x} \);
2) at \( Z_{x_i} \rightarrow -3 \), the value of \( x_i \) is much smaller \( \bar{x} \);
3) at \( Z_{x_i} \rightarrow 3 \), the \( x_i \) value is much larger \( \bar{x} \).

For further analysis, standardized values of the group of calls for disease factor 1- “zf1_1” and field intensity “zint1_1” (cluster 1) are presented in formed clusters; “Zf1_2” and “zint1_2” (cluster 2); “Zf1_3” and “zint1_3” (cluster 3). (Figure 3).

| internat | date | f_a_j |这是因为 | clusters | var000008 | mtl1_1 | date1_1 | ztl1_1 | cL1 | zint1_1 |
|-----------|------|-------|---------|----------|-----------|--------|----------|--------|-----|--------|
| 1         | -676,33 | 1 Feb | -371,74 | 1       | -676,33   | 1 Feb  | -0,37   | 1,51   |
| 2         | -381,28 | 2 Feb | 2,16779 | 2       | -305,86   | 4 Feb  | -0,88   | 1,42   |
| 3         | -296,90 | 3 Feb | 1,76522 | 1       | -300,77   | 5 Feb  | -0,76   | 1,41   |
| 4         | -305,96 | 4 Feb | -60003  | 1       | -275,81   | 13 Feb | -0,80   | 1,53   |
| 5         | -330,77 | 5 Feb | -1,10123 | 1      | -425,35   | 15 Feb | -1,03   | 1,50   |
| 6         | -282,86 | 6 Feb | -5,44322 | 2      | -204,19   | 17 Feb | -0,70   | 1,80   |
| 7         | -301,52 | 7 Feb | 4,54258 | 2       | -291,85   | 20 Feb | -0,74   | 1,47   |
| 8         | -214,81 | 8 Feb | 1,69977 | 2       | -200,15   | 21 Feb | -0,81   | 1,81   |
| 9         | -129,26 | 9 Feb | 63597   | 2       | -220,39   | 1 Apr  | -0,70   | 1,74   |
| 10        | -212,91 | 10 Feb | 0,0864  | 2       | -67,21    | 2 Apr  | -0,76   | 1,04   |
| 11        | -162,65 | 11 Feb | -7,7544 | 2       | -69,51    | 10 Apr | -1,36   | 1,86   |
| 12        | -220,93 | 12 Feb | -3,8059 | 2       | -247,26   | 11 Apr | -1,86   | 1,63   |
| 13        | -275,81 | 13 Feb | -8,0185 | 2       | -62,75    | 12 Apr | -1,24   | 1,57   |
| 14        | -404,00 | 14 Feb | 1,76973 | 2       | -181,12   | 13 Apr | -0,76   | 1,89   |
| 15        | -231,05 | 15 Feb | 2,23003 | 2       | -404,26   | 14 Apr | -0,76   | 1,25   |
| 16        | -404,96 | 16 Feb | 0,07085 | 2       | -65,29    | 15 Apr | -0,76   | 1,17   |
| 17        | -304,18 | 17 Feb | 1,76965 | 1       | -68,31    | 24 Apr | -2,29   | 1,03   |
| 18        | -687,70 | 18 Feb | 4,01034 | 3       | -619,93   | 3 Apr  | -2,73   | 1,73   |
| 19        | -652,94 | 19 Feb | 1,89676 | 2       | -425,89   | 30 Apr | -0,76   | 1,03   |
| 20        | -391,05 | 20 Feb | 0,0137  | 1       | -15,51    | 1 Jun  | -1,90   | 1,51   |
| 21        | -300,15 | 21 Feb | 1,87713 | 1       | -319,56   | 2 Jun  | -0,14   | 1,38   |
| 22        | -666,70 | 22 Feb | 4,59105 | 3       | -423,56   | 13 Jun | -1,49   | 1,80   |
| 23        | -375,91 | 23 Feb | 3,67109 | 3       | -261,03   | 14 Jun | 1,00    | 1,51   |
| 24        | -690,67 | 24 Feb | 4,67661 | 2       | -705,60   | 15 Jun | -1,29   | 1,72   |
| 25        | -661,71 | 25 Feb | 4,56695 | 3       | -187,73   | 16 Jun | -0,51   | 1,86   |
| 26        | -605,90 | 26 Feb | 1,84001 | 2       | -160,19   | 17 Jun | -0,51   | 1,96   |
| 27        | -237,20 | 27 Feb | 2,74864 | 2       | -1005,65  | 22 Jun | 1,41    | 1,25   |
| 28        | -372,90 | 28 Feb | 0,55695 | 2       | -845,45   | 25 Jun | -0,56   | 1,10   |

**Figure 3.** Fragment of an array of cluster analysis results (standardized values)

Then, for each cluster (1, 2, 3), a correlation analysis was performed between the number of calls for diseases included in factor 1 and the electric field intensity (table 5).
Table 5. Correlations of "electric field intensity - factor 1" pairs in clusters 1, 2, 3

| Claster 1  | F1_1  | Claster 2  | F1_2  | Claster 3  | F1_3  |
|------------|-------|------------|-------|------------|-------|
| INT1_1     | Correlation | -0.590*    | INT1_2     | Correlation | 0.048*    | INT1_3     | Correlation | -0.315*    |
| Sig.       | 0.001            | 0.792            | Sig.       | 0.117            |
| N          | 29                     | 33                     | N          | 26                     |

* Correlation is significant at the 0.01 level (2-tailed).

From table 6 it can be seen that the maximum dependence of the number of calls for diseases from the 1st factor (hypertensive crisis, vascular diseases, angina) to changes in the average daily intensity of the electric field of the atmosphere \( r = -0.590, \text{Sig.} = 0.001 \) with a permissible error \( p < 0.01 \) was observed inside the 1-st cluster. According to cluster analysis, the 1st cluster includes 29 days with dates: February 2007 - 1, 4, 5, 13, 16, 17, 20 and 21; April 2007 - 1, 2, 10, 11, 12, 13, 14, 18, 23, 24 and 30; June 2007 - 1, 2, 13, 14, 15, 16, 17, 22, 25 and 26.

Thus, having the initial data set for 181 days with calls for diseases, we obtained that only a group of diseases - hypertensive crisis, vascular diseases, angina pectoris, combined in factor 1, has a statistically significant correlation with changes in the intensity of the atmospheric field on certain days.

Based on the standardized values of factor 1 (Fig. 3), we find their maximum and minimum values with the corresponding values of standardized values of tension and dates for cluster 1. The largest values of the number of calls for diseases from the 1st factor group \( (zf1_1 > 0) \) were recorded 23 April \( (zf1_1 = 2.29) \) and June 22 \( (zf1_1 = 1.43) \) with a standardized value of the field intensity \( (INT1_1 = -1.63) \) and \( (INT1_1 = -2.55) \), respectively. The smallest number of calls \( (zf1_1 < 0) \) were February 21 \( (-1.68) \), April 12 \( (-1.24) \), April 18 \( (-1.46) \) and June 1 \( (-1.9) \) with a standardized field intensity value these days \( INT1_1 = 0.81; 1.37; -0.17; 1.51 \) respectively.

Further, as a result of the regression analysis, a regression equation was obtained for the standardized value of the frequency of calls for factor 1 (ZF1) diseases versus the standardized value of the atmospheric electric field per day (Zint):

\[
Zf1 = -0.298 - 0.523Zint
\]  

(2)

It can be seen from the equation that at zero standardized intensity, the standardized number of calls is \( -0.298 \), which, based on the definition of the standardized value, characterizes the call value as less than the average. Positive standardized voltage values (above average initial intensity) lead to a reduction in the number of calls \( (Zf1 < 0) \). Negative standardized voltage values (below average initial intensity) lead to an increase in the number of calls \( (Zf1 > 0) \).

Summary

The impact of changes in the electric field on the human body is actually a multifactorial process, the nature of which varies in time in a complex way. According to the results of the work, the following conclusions can be made:

1. Seasonal dynamics of reduction are observed, both average intensity and the number of ambulance calls for diseases from January to June.
2. According to the results of factor and correlation analysis, it was found that there is a moderate, statistically significant relationship between factor 1 (hypertensive crisis, cardiovascular pathology and angina pectoris) and the average daily electric field intensity.
3. According to the results of cluster analysis, it was found that the exacerbations of diseases from factor 1 correlate as much as possible with cluster 1, which includes days with certain characteristics of the electric field intensity.
4. According to the results of the regression analysis, it was found that the values of the electric field above the daily average lead to a reduction in the number of disease calls from factor 1, and below the average daily voltage - to an increase in the number of calls.
References
[1] Alyabina O V, Vasiliev V P, Maximov A V 2007 The influence of climatic factors on the exacerbation of hypertension of city residents (Proceedings of the Altai State University) 11-14.
[2] Bobina I V, Kobzeva O O 2010 The Influence of meteorological factors on frequency of an arterial hypertension’s aggravations (Proceedings of the Altai State University) 13-16.
[3] Shagdurova E A 2011 Influence of meteorological factors on the occurrence of acute cerebral circulatory disorders (Siberian Medical Journal) 131-133.
[4] Moosmüller G, Rebecca N N 2009 Market research with SPSS (Moscow, Infra-M).
[5] Mashukov I Kh, Shapovalov V A, Adzhieva A A 2018 EFM550 sensor network for studying the characteristics of the electric field intensity in the atmospheric surface layer, Technosphere Safety Systems (Materials of the Vth All-Russian Conference and School for Young Scientists (with international participation)) 204-206.
[6] Adzhieva O A, Bzhekshiev S L, Mashukov I Kh, Mashukov Kh Kh 2007 Studying the relationship between changes in the gradient of potential electric field of the atmosphere in the earth and meteotropic reactions in cardiological and hypertensive patients (Proceedings: VI Russian conference on atmospheric electricity) 263-264.
[7] Tashilova A A, Mashukov Kh Kh, Mashukov I Kh, Kesheva L A 2012 Study of the effect of variations in the electric field near the ground on hypertensive patients with cardiovascular diseases using statistical analysis methods 2012 (Proceedings of the 1-st International Conference "Youth in the formation of an innovative economy and the transition to a knowledge society", Russia, Nalchik, September 19-20, 2012) 353-356.