Article / Статья

DEVELOPMENT OF MATHEMATICAL MODEL OF GAS-DYNAMIC PHENOMENA FORECASTING METHOD ACCORDING TO GEOLOGICAL DATA IN CONDITIONS OF VERKHNEKAMSKOE POTASH SALT DEPOSIT

S.S. Andreiko

Mining Institute of the Ural Branch of the Russian Academy of Sciences (78a Sibirskaia str., Perm, 614007, Russian Federation)

Article proposes a statistical approach in predicting the gas-dynamic danger in potash mines, which can significantly reduce the dependence of the final prediction results on subjective factors by introducing a model of informative strictly defined criteria. Obtaining the most appropriate forecasting models of reality is only possible through an integrated approach, displaying factors of generation, migration, accumulation and preservation for a long period of geological time gas clusters and centers of gas-dynamic phenomena in the salt rock mass. In the mathematical model development for gas-dynamic phenomena’s prediction method on geological data a parametric mathematical model of forecasting method was used based on discriminant analysis, which is a powerful statistical method and in depth of data analysis and the results significance is among the most effective methods of statistical analysis. In case of compliance with conditions of multidimensional data normal distribution and equality of covariance matrices the parametric model leads to optimal results in forecasting. In case of non-compliance with these restrictions in the use of robust evaluation model allows to compensate the incorrect prediction’s probability growth. A statistical analysis of geological information in the field of gas-dynamic phenomena (GDP) in the potash mines was performed. Learning samples were formed, the mathematical model of forecasting method for zones dangerous because of GDP was developed based on multivariate discriminant analysis with classical and robust statistical procedures. Decision rules to predict the geological data of areas dangerous because of GDP were obtained for conditions of VerkhneKamskoie potash salt deposit. Forecast maps were created for zones dangerous because of GDP for the Ust-Yanvinskii mine field conditions on the base of PJSC "Uralkali" and Polodovodovskii area of VerkhneKamskoie potash salt deposit. The adequacy estimation was given for the method of forecasting mining operations' practice on potash layers dangerous because of GDP in conditions of potash mine fields PJSC "Uralkali".

Received / Получена: 23.11.2015. Accepted / Принята: 27.10.2016. Published / Опубликована: 02.12.2016

Key words: potash mines, gas-dynamic phenomena, geological data, learning sample, mathematical model, discriminant analysis, statistical procedures, robust estimation, decision rules, forecast maps.

Ключевые слова: калийные рудники, газодинамические явления, геологические данные, обучающая выборка, математическая модель, дискриминантный анализ, статистические процедуры, робастные оценки, решающие правила, прогнозные карты.

Sergei S. Andreiko – Doctor of Technical Sciences, Professor, Head of the Laboratory of Geotechnological Processes and Mine Gas Dynamics (tel.: +007 342 216 58 42, e-mail: ssas@miperm.ru).

Andreyko Sergey Semenovich – doctor of technical sciences, professor, head of the laboratory of geotechnological processes and mining gas dynamics (tel.: +007 342 216 58 42, e-mail: ssas@miperm.ru).

Introduction

At the present time about 30 deposits of potassium and magnesium salts is developed by underground method. In the process of underground mining at all the developed deposits there is the problem of intense release of flammable and toxic gases, as well as the problem of various types of gas-dynamic phenomena (GDP). Intensive gas emission leads to mines gas contamination, flares and explosions of flammable gases, which are frequently accompanied by single and multiple accidents, including fatalities. Gas-dynamic phenomena in the form of sudden salt and gas outburst, the destruction of roof rocks and soil of mining, accompanied by gas evolution and the combined type effects by virtue of its suddenness and high power (up to 6.000 tons of shattered rock mass) threaten the lives of miners, incapacitate expensive mining equipment and disrupt the rhythm of the work of potash mines [1-4].

The problem of forecasting of gas-dynamic phenomena in underground mining of potash deposits is studied for a long time. As a rule, to solve the problem is used geological-mathematical methods of forecasting. A statistical approach in predicting the gas-dynamic danger allows significantly reduce the dependence of the final prediction results on subjective factors by introducing a model of informative strictly defined criteria. However, it should be noted that obtaining the most appropriate forecasting models of reality is only possible through an integrated approach, displaying factors of generation, migration, accumulation and preservation for a long period of geological time gas clusters and centers of gas-dynamic phenomena in the salt rock mass.

Research in the field of construction of mathematical models of gas-dynamic localization of the danger areas in the potash mines traditionally were limited to the use of parametric models based on linear discriminant functions. Studies have shown that the parametric mathematical model of forecasting method for zones dangerous because of GDP is a powerful statistical method, both in depth analysis of the data and the value of the results belongs to the most effective methods of statistical analysis. The model is sufficiently simple, concise, well interpreted, convenient to consider the geometric illustrations of separation areas of the array into the dangerous and non-dangerous GDP zones. The parametric mathematical model of forecasting method allows to to select the most informative indicators, which makes it possible to reduce the dimension of the original space of the features and drop uninformative variables. In the case of compliance with the terms of multidimensional data distribution normality and equality of covariance matrices parametric model leads to optimal results in forecasting. In case of non-compliance with these restrictions in the use of robust evaluation model allows to compensate the incorrect prediction’s probability growth. Thus, in dealing with applied problems of forecasting zones dangerous because of GDP, parametric mathematical model based on linear discriminant analysis using robust estimators is the most preferred among the other well-known parametric models.

Statement of the problem

The objective is basing on a set of geological parameters to determine zone as dangerous or non-dangerous because of GDP, the particular area of deposit or a mine field. According to the mathematical model of forecasting method is the decision rule, which is substituting with the values of geological parameters, allows the highest possible chance to assess considered area from the standpoint of the gas-dynamic danger.

The mathematical formulation obtaining a decision rule is reduced to the definition of a finite set of of geological parameters, the intervals of their values, inherent to both dangerous and non-dangerous zones because of GDP, as well as finding the most efficient in terms of the probability of a correct forecasting of statistical procedures.

Formally, each record of available geological data represents a multi-dimensional observation of P-dimensional space with a finite set of attributes X. The set of observations belonging to a class of D₁ forms a "cloud" in the same space. Therefore for successful forecasting is necessary that the "cloud" of D₁ was concentrated in a certain region R₁ of the space X and a small part of D₂, according to another class, got into region R₂. In general, the construction of a decision rule for classifying the multidimensional observation to a particular class can be viewed as the problem of finding k disjoint regions Rᵢ (i = 1, k), satisfying the above conditions. If the P-dimensional observation of space X falls in the region Rᵢ, then we assume that the decision on an object belonging to Dᵢ is accepted. By using such a forecasting rule in some cases, an object can be assigned to another class. So we denote the probability of misclassification of objects through P(i/j), then the criterion of the average probability of an erroneous forecast is as follows:
where \( a_i \) is a priori probability of the object occurrence from the domain \( D_i \).

The solution of this task is carried out at presence of a training set, which should consist of observations relating to both classes (dangerous and non-dangerous GDP zones). Therefore we consider that the training set for \( D_i \) class is a sampling of a general set of objects with the distribution density of \( P(X) \). Forecasting rule that minimizes the loss function \( Q \) is as follows:

\[
x \in D_i, \text{ if } \max_{1 \leq j \leq k} a_j P(X/j) = a_i P(X/i).
\]  

To find the probability meaning of the expression (2), we assume that \( P(X) = \sum_{i=1}^{k} a_i P(X/i) \). The value \( P(j/X) = a_j P(X/j) \times [P(X)]^{-1} \) represents the conditional probability that the \( P \)-dimensional observation \( X \) belongs \( D_i \). Therefore, the expression (2) is called forecasting rule on a maximum of conditional probability, as it gives the smallest value of the loss function (1). In reality, the exact location of the "clouds" of objects and their density \( P(X) \) is unknown, and there are training set that give a rough idea about it. The training set is a set of multivariate observations obtained from research. The matrix of the original data is also included observations belonging to any class that is not clearly established.

At present it is known a large number of pattern recognition methods for constructing decision rules on the training set, or the so-called statistical classification, which are divided into two groups: parametric and nonparametric methods [5-11]. The parametric methods of obtaining the decision rules for predicting traditionally is used the assumption of homogeneity of training samples for each class, measuring all the components of the feature vectors, accessories conditional probability distributions for multivariate Gaussian family, independence of sample values. The most well known of parametric methods is "k-nearest neighbor algorithm" and Rosenblatt-Parzen window method [12-14]. These methods are relatively easy, work well in predicting, but require constant memorization of training samples and are very sensitive to small amounts of training samples in which the risk of misclassification increases dramatically. At the same time robustness factor (sustainability), representing a relative increase of the probability of classification error, with the volume of training data for more than 20 non-parametric methods are 2-4 times greater than for linear discriminant analysis based on the LDF [14]. To obtain the decision rules for predicting zones dangerous because of GDP is used training samples, the volume of which far exceeds the 20, allowing give preference to a parametric model for obtaining decision rules for forecasting zones dangerous because of GDP.

The essence of constructing the decision rule based on the linear discriminant function is as follows. If we denote by \( X_{ij} \) the variable number \( i \) at the observation point number \( j \), taken from the sample, characterizing zones dangerous because of GDP, we obtain the matrix \( V \) of order \( p \times n_1 \) of results of this sample observations:

\[
V = \begin{bmatrix} X_{11} & X_{12} & \ldots & X_{1n_1} \\ X_{21} & X_{22} & \ldots & X_{2n_1} \\ \vdots & \vdots & \ddots & \vdots \\ X_{p_1} & X_{p_2} & \ldots & X_{p_{n_1}} \end{bmatrix}, \tag{3}
\]

Let \( X_{ij} \) as the result of measurement of the variable with index \( i \) at point with index \( j \), taken from the sample characterizing the non-dangerous zone because of GDP. The result is a matrix \( W \) of the order \( p \times n_2 \):

\[
W = \begin{bmatrix} X_{11} & X_{12} & \ldots & X_{1n_2} \\ X_{21} & X_{22} & \ldots & X_{2n_2} \\ \vdots & \vdots & \ddots & \vdots \\ X_{p_1} & X_{p_2} & \ldots & X_{p_{n_2}} \end{bmatrix}, \tag{4}
\]

where \( p \) – number of variables (indicators); \( n_1, n_2 \) – the volume of the indicator samples.

Then using these data for each group are calculated the average and sum of the cross products of deviations from the averages. The average is calculated from the formula.
\[ \bar{X}_{ik} = \left( \sum_{i=1}^{n_k} X_{i(k)} \right) (n_k)^{-1}, \]  

where \( k \) – the number of groups \((k = 2)\); \( n_k \) – the size of the sample in the group \( k \); \( j = 1, 2, \ldots, p \) – variables.

Then find sum of the cross products of the deviations from the average:

\[ S_k = \sum \left( X_{i(k)} - \bar{X}_{jk} \right) \left( X_{i(k)} - \bar{X}_{jk} \right), \]

where \( j = 1, 2, \ldots, p; \ l = 1, 2, \ldots, p \).

The next step is the calculation of the combined variance matrix

\[ D = \left( \sum_{k=1}^{q} S_k \right) \left( \sum_{k=1}^{q} n_k - q \right)^{-1}, \]

where \( q \) – the number of groups \((q = 2)\).

Calculate the overall average for both groups:

\[ \bar{X}_j = \left( \sum_{k=1}^{q} n_k \cdot \bar{X}_{jk} \right) \left( \sum_{k=1}^{q} n_k \right)^{-1}, \]

where \( \bar{X}_j \) – averages of variable \( j \) in group \( k \); \( n_k \) – the size of the sample in group \( k; j = 1, 2, \ldots, p \) – variables.

Then for each group calculated inverse elements of the combined dispersion matrix \( D \):

\[ \left[ d_{ij} \right] = D^{-1}. \]

Discriminant function coefficients are calculated according to the formula

\[ a_{ik} = \sum_{j=1}^{p} d_{ij} \cdot \bar{X}_{jk}, \]

where \( i = 1, 2, \ldots, p; k \) – the number of groups.

Then constant of discriminant functions are calculated:

\[ b_{ik} = -0.5 \sum_{j=1}^{p} d_{ij} \cdot \bar{X}_{jk} \cdot \bar{X}_{jk}. \]

For the case of two classes \( D_1 \) and \( D_2 \) (dangerous and non-dangerous zones because of GDP) we obtain two linear discriminant function of the form

\[ \sum_{i=1}^{p} a_{i} \cdot X_i - b_{ik} = 0. \]

Subtracting less discriminant function from a greater is reduced to a single linear discriminant function of the form

\[ a_1 X_1 + a_2 X_2 + \ldots + a_p X_p - b = 0. \]

For the case of two classes the use of linear discriminant functions based on two assumptions.

The first assumption is that the regions \( D_1 \) and \( D_2 \), containing the most part of objects corresponding to \( D_1 \) and \( D_2 \) classes (non-dangerous and dangerous zones because of GDP) can be divided by \((p-1)\) – dimensional hyperplane. The equation of this hyperplane and is an expression of \( (13) \). Here \( a_0, a_2, \ldots, a_p \) – coefficients characterizing the slope of the hyperplane to the coordinate axes, and \( b \) – parameter characterizing the distance of the hyperplane to the origin. The hyperplane \( (13) \) splits the p-dimensional space into two half-\( R_1 \) and \( R_2 \), which are defined by inequalities

\[ R_1/z(X) < b \quad \text{and} \quad R_2/z(X) > b, \]

where \( z(X) = \sum a_j \cdot X_j \).

Consequently, inequality \( (14) \) gives us a decision rule for classifying the multidimensional vector of observations for one of the two zones – dangerous or non-dangerous because of GDP. Let us agree to refer the observation to \( D_1 \), if \( z(X) \leq b \), and to \( D_2 \) – otherwise. Equation \( (14) \) allows to switch from p-dimensional random variables to the classification of one-dimensional value \( z(X) \), when \( X \in D_i \).

The second assumption provides a method for determining the coefficients of the separating hyperplane such that separation was the best. Suppose that the division will be the better, the farther apart the average values of the random variables \( z_1 \) and \( z_2 \), the distance is measured in terms of the dispersion:

\[ h^2(a) = (m_1 - m_2) D_2^{-1}, \]

where \( m_i = E z_i \); \( D_2 \) – the dispersion of \( z_i \) \((i = 1, 2)\), assumed the same for both classes.

Vector \( \bar{A} \), delivering maximum \( h^2(a) \), given by the equation

\[ \bar{A}_\Sigma = \left[ \sum (\bar{X}_2 - \mu_1) \right]^{-1}, \]

where \( \mu \) – a mean vector for \( D_1 \); \( \Sigma \) – covariance matrix considered to be the same for both classes.

The maximum value of \( h^2(a) \) is called the Mahalanobis distance between classes of \( D_1 \) and \( D_2 \) and equals

\[ D^2 = \sum_{i=1}^{p} \sum_{j=1}^{p} d_{ij} \sum_{k=1}^{q} n_k \left( \bar{X}_{ik} - \bar{X}_j \right) \left( \bar{X}_{jk} - \bar{X}_j \right). \]
Mahalanobis statistics can be used to check the significance of the discriminant function. If we assume a normal multivariate distribution, then $D^2$-statistics is distributed as $\chi^2$ with $p(q-1)$ degrees of freedom. Comparing the calculated value $D^2$ with $\chi^2$ at the 5% significance level, if $D^2 > \chi^2$, one can say that the resulting linear discriminant functions have meaning and can be used as decision rules in forecasting. At the stage of exploratory analysis establish that the objects from $D_i$ ($i = 1, 2$) are distributed according to a multivariate normal distribution with identical covariance matrix. Therefore, the decision rules (14) based on linear discriminant functions (10) and (11) are the best that is delivering to the functional $Q (1)$ the smallest possible value.

It is necessary to note one more very important aspect in finding the decision rules by linear discriminant functions. If at the stage of exploratory analysis found "clogging" of training samples by "emissions" and these anomalous observations could be of interest in studies, then it is advisable to use the distribution parameter estimation procedures that are not sensitive to the data structure. Such evaluation procedure called robust or stable [6, 15-22]. The following robust procedures could be used in the permutation discriminant analysis for obtaining of stable decision rules: winsorized estimators, truncated assessment, Huber weighted estimates, Hampel piecewise-linear M-assessment. If there is distortion in the training samples and using classical estimates of distributions parameters in the decision rules the probability an erroneous forecast is significantly increased, therefore it is necessary to use robust estimates.

If there are "emissions" in the training samples, you can use any of the following stable estimates. If the training sample is formed from distributions with "heavy-tailed" or different from the normal it is better to to apply the weighted estimates of Huber and Hampel. Huber robust iterative estimation is defined the following relations:

$$
\hat{x}_i = \sum_{j=1}^{k} x_{i}w_j / \sum_{j=1}^{k} w_j ,
$$

$$
\hat{\Sigma}_i = \sum_{j=1}^{k} w_j w_k \left( x_{j} - \hat{x}_i \right) \left( x_{k} - \hat{x}_i \right)^T / \sum_{j=1}^{k} w_j ,
$$

$$
w_j = \begin{cases} 
1, & \text{if } d_j \leq d_0, \\
\frac{d_j}{d_0}, & \text{if } d_j > d_0 
\end{cases},
$$

$$
d_j = \left( x_{j} - \hat{x}_i \right) \Sigma^{-1} \left( x_{j} - \hat{x}_i \right), (j = 1, n; i = 1, L).
$$

In general parametric mathematical model of forecasting method for zones dangerous because of GDP, based on the discriminant analysis is a powerful statistical method both in depth analysis of the data and the value of the results belongs to the most effective methods of statistical analysis. The model is sufficiently simple, concise, well interpreted, convenient to consider the geometric illustrations of separation areas of the array into the dangerous and non-dangerous GDP zones. The parametric mathematical model of forecasting method allows to to select the most informative indicators, which makes it possible to reduce the dimension of the original space of the features and drop uninformative variables. In the case of compliance with the terms of multidimensional data distribution normality and equality of covariance matrices parametric model leads to optimal results in forecasting. In case of non-compliance with these restrictions in the use of robust evaluation model allows to compensate the incorrect prediction’s probability growth.

**The procedure for obtaining decision rules**

Taking into account the existing ideas about the formation’s mechanism of the gas-dynamic phenomena’s centers are the salt rock mass of Verkhnekamskoe potash salt deposit [23, 24]. Formation of the training sample was conducted the following way. All documented cases of GDP in mining AB formation in the southern part of Verkhnekamskoe potash salt deposit on the mine fields of Mine-2 and Mine-4 and their corresponding geological indicators was a priori assigned to zone dangerous because of GDP. (Group 1). The rest of the total sample was related to zone non-dangerous because of GDP (Group 2). We investigated 658 observations, 478 are classified as non-dangerous because of GDP (Group 1) and 180 as dangerous (Group 2). To obtain the decision rules we used the following indicators: $X_1$ – formation thickness B, m; $X_2$ – the content in the formation B of potassium chloride (KCl), %; $X_3$ – the content in the formation B of magnesium chloride (MgCl$_2$), %; $X_4$ – the content in the formation B of sodium chloride (NaCl), %; $X_5$ – the content in the formation B of bromine (Br), %; $X_6$ – the content in the formation B of...
calcium sulfate (CaSO₄), %; X₇ – the content in the formation B of insoluble residue, %; X₈ – occurrence depth of formation B roof, m; X₉ – absolute mark of formation B roof, m.

Exploratory analysis of baseline data showed heterogeneity of the total sample, the deviation from the normal distribution, and the presence in the sample of "heavy tails". Due to the fact that not all of used indicators is equally useful for the separation of the formation into non-dangerous and dangerous because of GDP zones, it was necessary to identify a subset of indicators by which one can build "the best" split between the two groups. Stepwise discriminant analysis (SDA) was implemented to solve this problem using the following estimates: classical estimation, Huber weighted estimates and truncated assessment with truncation level of 0.1. SDA used the F-statistic based on one-way ANOVA criteria for the selection of the most informative indicators. In fact, the SDA logic is: initially to determine indicator for which the average values of the two groups "the most different". For each indicator, X₁, X₂, X₃…X₉ difference is measured using the F-statistic and one indicator, which corresponds to a larger value of F, is selected (included). At each step of the procedure is considered conditional distribution of each indicator not included in the subset for given including indicators. Then from among the non-included indicators the one for which the average values of the two groups "the most different" is determined once again. This difference is determined by using F-statistics. The process ends when none of indicators do contribute significantly to the group separation. In the process of implementing SDA were accepted standart value of F-inclusion minimum of variables not included in the sought subset and F-delete the selected variables at a significance level of 0.05. In step 8, the value of F-delete statistic was less than accepted minimum, equal to 3.0, hence "the best" discriminant equation are defined in step 7. Thus indicators X₁, X₂, X₃, X₄, X₅, X₆ and X₇ were included in the sought subset and indicators X₈ and X₉ – not included. Analysis of classification matrices in the implementation of the SDA procedure using classical estimation, Huber weighted estimates and truncated assessment with truncation level of 0.1 showed the following. In the application of classical assessments were correctly classified 62.3 % of the original grouped cases. In case of using Huber weighted estimates and truncated assessment with truncation level of 0.1 were correctly classified 61.7 % and 61.5 % of the original grouped cases.

The validity of the primary classification using both classical and robust procedure is virtually the same, and more than 60 %. So the final choice of the decision rule was held after the reclassification. At the same time raw data were regrouped, and then reclassification procedure was carried out, which resulted in the refinement coefficients of linear discriminant functions for each of the groups, as well as the assessment of the correctness of the classification has been given. For each observation with the use of the obtained discriminant functions and D²-statistics of Mahalanobis posterior probabilities of belonging to a group dangerous and non-dangerous because of GDP were calculated. Based on these probabilities each observation it was classified as belonging to one of the groups and classification matrix was compiled. Table 1 shows the classification results using the classical procedures, using weighting Huber estimates, truncated estimates (truncation level of – 0.1).

### Table 1

<table>
<thead>
<tr>
<th>Group</th>
<th>Classification results</th>
<th>Total %</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>non-dangerous because of GDP</td>
<td></td>
<td>abs.</td>
</tr>
<tr>
<td>GDP non-dangerous (1)</td>
<td>323</td>
<td>91.76</td>
<td>29</td>
</tr>
<tr>
<td>GDP dangerous (2)</td>
<td>18</td>
<td>5.88</td>
<td>288</td>
</tr>
<tr>
<td>Total</td>
<td>341</td>
<td>51.82</td>
<td>317</td>
</tr>
<tr>
<td>Using Huber weighted estimates</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP non-dangerous (1)</td>
<td>334</td>
<td>94.89</td>
<td>18</td>
</tr>
<tr>
<td>GDP dangerous (2)</td>
<td>8</td>
<td>2.61</td>
<td>298</td>
</tr>
<tr>
<td>Total</td>
<td>342</td>
<td>51.98</td>
<td>316</td>
</tr>
<tr>
<td>Using truncated assessment with truncation level of 0.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP non-dangerous (1)</td>
<td>319</td>
<td>90.63</td>
<td>33</td>
</tr>
<tr>
<td>GDP dangerous (2)</td>
<td>24</td>
<td>7.84</td>
<td>282</td>
</tr>
<tr>
<td>Total</td>
<td>343</td>
<td>52.13</td>
<td>315</td>
</tr>
</tbody>
</table>

As seen from Table 1 using classical statistical procedures of linear discriminant analysis was correctly classified 92.86 % of re-grouped observations.

By using linear discriminant analysis with weighted estimates of Huber the probability of correct classification is 96 %.

When used in linear discriminant analysis truncated assessment with truncation level of 0.1 the probability of correct classification is 91.34 %. By the criterion of the highest probability of correct classification we accept for forecasting zones dangerous because of GDP the decision rule obtained using weighting Huber estimates. Table 2 shows constants and coefficients of discriminant functions found using weighted estimates of Huber.

Since in this case there are two groups, two discriminant functions can be reduced to a single subtracting coefficients and constant of smaller
function from the larger. After subtracting we obtain a linear discriminant function of the form:

\[
F_p = -19.98X_1 - 1.16X_2 + 31.18X_3 - 3.64X_4 + 6.87X_5 - 4.57X_6 + 310.72 \geq 0,
\]

where \(X_1\) – formation thickness \(B\), m; – the content in the formation \(B\) of potassium chloride (KCl), %; \(X_3\) – the content in the formation \(B\) of magnesium chloride (MgCl₂), %; \(X_4\) – the content in the formation \(B\) of sodium chloride (NaCl), %; \(X_5\) – the content in the formation \(B\) of bromine (Br), %; \(X_6\) – the content in the formation \(B\) of calcium sulfate (CaSO₄), %; \(X_7\) – the content in the formation \(B\) of insoluble residue, %.

The constants and coefficients of the discriminant functions obtained using Huber weighted estimates

<table>
<thead>
<tr>
<th>Group</th>
<th>Constants of the discriminant functions</th>
<th>Coefficients of the discriminant functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP non-dangerous (1)</td>
<td>–385755.60</td>
<td>(X_1) 7678.37 (X_2) 16092.34 (X_3) 7734.02 (X_4) –11769.6 (X_5) 7829.14 (X_6) 7715.20</td>
</tr>
<tr>
<td>GDP dangerous (2)</td>
<td>–384444.89</td>
<td>(X_1) 7767.20 (X_2) 16123.52 (X_3) 7730.39 (X_4) –12516.62 (X_5) 7836.01 (X_6) 7710.63</td>
</tr>
</tbody>
</table>

Expression (19) is the desired decision rule for classifying AB formation areas to the zones, dangerous or non-dangerous because of GDP. When \(F_p \geq 0\) observation point refers to the area dangerous because of GDP, otherwise to non-dangerous. For the probabilistic interpretation of the discriminant function, which is more convenient in constructing forecast maps, it is advisable to use the expression

\[
P_i = \frac{1}{\sum_{k=1}^{q} \exp(f_k - f_i)},
\]

where \(P_i\) – the probability corresponding to the largest discriminant function; \(q\) – the number of groups (\(q = 2\)); \(f_k, f_i\) – the value of the \(k\)-th and the largest discriminant function correspondingly.

Substituting the values of indicators in discriminant functions, which constants and coefficients are shown in Table. 2, we obtain some values of the discriminant functions. Then, substituting the obtained values of the discriminant functions in the expression (20), we find the probability of classifying the observation point to the zone, which value of the discriminant function larger. The value of \(P_i\) varies from 0 to 1 the boundary is a probability value of 0.5.

Assessing the adequacy of the decision rules and forecast maps of zones dangerous because of GDP

To assess the adequacy of obtained decision rules to practice of mining in the forecast maps of zones dangerous because of GDP were marked the places where under conduction of preparatory and sewage mining operations occurred gas-dynamic phenomena. In developing AB formation of Mine-1 field, 8 out of 98 GDP-cases are not included in forecasting zones dangerous because of GDP (7.8 % of the total). Thus, the adequacy of the zones forecasting results dangerous because of GDP in the formation AB of sylvinite mine

![Fig. Map of zones dangerous because of GDP from a the roof of mine workings at mining AB formation in conditions of Polovodovsky sector of Verkhnekamskoie potash salt deposit](image)
data on the mine fields of Mine-1 and Mine-4 initially not included in the original data. Then examination of new input data classification was carried out, which correctness for formation of AB was 96.7 %. Substituting the values of geological parameters obtained as a result of the wells drilling, detailed exploration from the surface, from underground exploration wells and furrow sampling data in underground mines, into decision rules allowed to construct forecast maps of zones dangerous because of gas-dynamic phenomena.

The figure shows a map of zones dangerous because of GDP from the roof of mine workings at mining AB formation for conditions of Polovodovskii area of Verkhnekamskoie potash salt deposit.

**Conclusion**

As a result of scientific research a mathematical model of the method forecasting zones dangerous because of gas-dynamic phenomena based on linear discriminant analysis using robust (resistant) procedures, allowing to increase the reliability of multivariate forecasting was created. A method for forecasting of zones dangerous because of gas-dynamic phenomena in the exploration and development of AB sylvinite formation using in decision rules the following complex of geological parameters: formation thickness, the content of potassium chloride, the content of magnesium chloride, the content of sodium, the content of bromine, the content of calcium sulfate, the content of insoluble residue in formation rocks. The probability of a correct discrimination of AB sylvinite formation sections into zones dangerous and non-dangerous because of GDP is 0.96. Adequacy of zones forecasting results dangerous because of GDP to practice of mining operations on the formation of AB sylvinite formation is 92 % correspondingly. Forecast maps were created for zones dangerous because of GDP for the Ust-Yaivinskiimine field conditions on the base of PJSC "Uralkali" and Polovodovskii area of Verkhnekamskoie potash salt deposit.

**References**


Список литературы


Please cite this article in English as:

Просьба ссылаться на эту статью в русскоязычных источниках следующим образом: