Creating Forecast Maps of the Spatial Distribution of Dangerous Geodynamic Phenomena Based on the Principal Component Method

V Kiselev¹, N Guseva¹, A Kuranov¹

¹Saint-Petersburg Mining University, 2, 21st Line, St Petersburg 199106, Russia

E-mail: Kuranov_AD@pers.spmi.ru

Abstract. Subject matter of the research is mining and geological subsurface objects use in the conditions of the DGDE possibility. Scope of research is the regularities of spatial distribution of geological objects properties, causing DGDE in the mining enterprise. The purpose of research is to map the mining enterprise territory subject to DGDE using the method of objective determination of weight coefficients and influencing factors. The main tasks: establish the size of the operational territorial units (OTU) and discretization of the research area; determination of the weight coefficients of the factors influencing the process of the DGDE occurrence formation on the subsurface territory; create a mathematical expression for a generalizing value determining that characterizes the degree of adverse factors coincidence, and therefore the risk rank of the DGDE occurrence.

1. Introduction

The current state of the mining industries in Russia and other countries is characterized by the complexity of mining and geological conditions [6, 10, 16, 32]. This, in some cases, leads to the occurrence of dangerous geodynamic events (rock bursts, sudden outbursts, rock collapse etc.) that disrupt the production process or cause accidents. The way out of this situation is to develop measures to prevent the occurrence of these events [3, 9, 33], the effectiveness of which, in many respects, depends on the results of forecasting the place and time of dangerous geodynamic events occurrence. In all mines, as a rule, local forecast is performed, which is carried out in dangerous zones and near the mine workings during operation [19].

Due to the fact that dangerous geodynamic events (DGDE) occur with a certain combination of natural and technological causes [18, 27, 36, 37], methods are used to make the forecast that allow taking into account the entire set of influencing factors [7, 13, 14, 15, 17, 22]. One of these methods is to create a forecast map (plan), which shows the location of areas that have not yet been identified but are potentially dangerous for rock bursts.

Currently, forecast maps (plans) are mainly created in the environment of geoinformation systems (GIS) based on maps-layers of spatial variability of the influencing factors. In accordance with GIS-technologies, factor layers are subjected to overlay operations, the results of which are used to create thematic maps [5, 19, 27]. At the same time, the degree of influence of each factor is taken into account by applying the weighted overlay method [4, 28, 35]. However, in this method, the values of weight coefficients are determined by expert methods [12, 20, 24, 26] or using artificial neural networks (ANN) [23, 25]. In the first case, the results are subjective, in the second - it need to have a
sufficient number of DGDE occurrences, which is often difficult to ensure. This circumstance creates a contradictory situation that requires its resolution. Thus, the development of a method for determining the objective values of weight coefficients, influencing factors for creating forecast maps of the spatial distribution of DGDE occurrence zones is relevant.

2. Research methods
The occurrence of dangerous geodynamic events is caused by many factors, each of which has a unique character of variability in space. To more accurately reflect this on plans or maps, the entire territory under research is divided into operational territorial units (OTU), which are primary, conditionally indivisible cells, for which information is collected for subsequent analysis and zoning [29]. The OTU uses point and special area objects, the size of which is set depending on the tasks to be solved. In our case, the size of the OTU is predetermined by the size of the development workings for which the forecast is made. Since the development workings have a width of 5-6 m, the size of the OTU was assumed to be 1 m. However, since weighted overlay tools use only integer rasters as input data [35], it was decided to use vector point objects located 1 m apart instead of OTU raster models. In this case, each OTU can have continuous (floating-point) feature values.

The following two tasks: determining the values of factor weights and determining the generalizing value are to the problem of reducing the multidimensional data dimensionality, which is structured as an object-criterion matrix [1]. The object is the OTU, the criterion is the value of the factor for the OTU. Taking into account the fact that the determination of weighting factors by expert methods is subjective and the use of artificial neural networks requires a sufficient number of fixed DGDE, the most acceptable method for our case, which ensures the loss of the least amount of information, is the principal component analysis [2, 6, 9].

As inputs is an object-criterion matrix $X$ with dimension $(I \times J)$, where $I$ is the number of OTU (rows), a $J$ is the number of independent variables (columns), $(J >> 1)$. In the principal component analysis uses new, formal variables $y_a$ ($a = 1, \ldots, A$), which are linear combination of the original variables $x_j$ ($j = 1, \ldots, J$)

$$y_a = p_{a1}x_1 + \ldots + p_{aJ}x_J$$

(1)

With these new variables the matrix $X$ is decomposed into the product of two matrices $T$ and $P$

$$X = TP^t + E = \sum_{a=1}^{A} y_a p_a^t + E,$$

(2)

where $T$ – is the matrix of scores with dimension $(I \times A)$;
$P$ – is the matrix of loadings with dimension $(J \times A)$;
$E$ – is the matrix of residuals with dimension $(I \times J)$;
$y_a$ – the principal components.

The numbers of columns – $y_a$ in the matrix $T$ and $p_a$ in the matrix $P$, is equal to $A$, which is called the number of principal components (PC). This value is obviously less than the number of variables $J$ and the number of objects $I$.

The Statistics software package was used as a tool for mathematical processing of initial data (section Multivariate Exploratory Techniques – Factor Analysis).

3. The results
In order to test the effectiveness of the proposed approach, the initial data consisting of 12 factors were combined into two groups:

- mining and geological: geological faults (TECT), the depth of occurrence (SURF), natural seismic (density (SS_DN) and energy of seismic events (SS_EN)), rock jointing ((FR_CD), (FR_RQD), ore body (ORE);
- mining and engineering: impact of development workings (MINE) and stoping front, both on the forecast horizon (FRONT_XXX) and on the higher horizon (FRONT_XXX+), seismic, caused by blasting (density (BVR_DN) and energy of seismic events (BVR_EN)).

Layers of graphical and semantic information were created for each factor in the ArcGIS 10.0 environment. After that, the original graphical information was converted to a grid of point objects (OTU) with an interval of 1x1 m. As an example, figure 1 shows the maps fragments of the point objects location that reflect the zones and the degree of their influence on the dangerous geodynamic events possibility on the forecast horizon. In this case, the values of factors were standardized into a set of criterion $x(1)$, $x(2)$, $x(3)$, $x(4)$, ..., $x(12)$ according to the formula

$$x(i) = \frac{x_i - \bar{x}}{\sigma_x},$$

where $x(i)$ – is the standardized value; $x_i$ – criterion value; $\bar{x}$ – average value; $\sigma_x$ – standard deviation.

![Figure 1](image1.png)

**Figure 1.** The maps fragments of the point objects location that reflect the zones and the degree of their influence: a- tectonic disturbances, b- stoping front on the dangerous geodynamic events possibility on the forecast horizon.

The size of zones and the degree of influence of tectonic disturbances were determined by statistical processing of the distances between: a) tectonic disturbances; b) stoping front and recorded dangerous geodynamic events.

At the next stage of data preparation, the above factors were gathered in a single layer, where all semantic information was collected in a single table using the spatial connection of tables based on the coincidence of point objects coordinates (figure 2).

![Figure 2](image2.png)

**Figure 2.** A fragment of the summary table of semantic information.
To move from the initial indicators to the principal components, the proportion of variance provided by the first two principal components \((ya)\) was estimated and the objects were shown in the space of new features. For this, a sample correlation matrix was determined based on the criterion data (figure 3).

Eigenvalues \(\lambda_i\) and their cumulative eigenvalues is shown in figure 4.

\[
\begin{align*}
\lambda_1 &= 3.491064; \\
\lambda_2 &= 1.955513; \\
\lambda_3 &= 1.315542; \\
\lambda_4 &= 1.198957; \\
\lambda_5 &= 0.847178; \\
\lambda_6 &= 0.536240; \\
\lambda_7 &= 0.415249; \\
\lambda_8 &= 0.375578; \\
\lambda_9 &= 0.347383; \\
\lambda_{10} &= 0.310596;
\end{align*}
\]

They are listed in descending order of their values.

The results of calculations show that almost all information (79.6 \%) about the entropy characteristics variability of the observed sample is contained in the first 4 principal components. However, all factors were left to get the final result. The eigenvectors values, which represent the initial values of weight coefficients, are shown in figure 5.
As the principal component for the first factor, a mathematical expression was compiled:

\[ Z(1) = 0.474778 \cdot \text{SURF} + 0.113865 \cdot \text{FRONT} \_\text{XXX} - 0.382329 \cdot \text{TECT} + 0.1733 \cdot \text{SS} \_\text{DN} + 0.05167 \cdot \text{SS} \_\text{EN} - 0.443264 \cdot \text{BVR} \_\text{DN} - 0.180195 \cdot \text{BVR} \_\text{EN} + 0.318495 \cdot \text{FR} \_\text{CD} + 0.246897 \cdot \text{FR} \_\text{RQD}. \]

The same was done for the second and subsequent columns of table 2. However, in the case under consideration, it was decided to limit to only the first principal component.

Based on the results of the calculation \( Z(1) \), the rank values \( R = Z(1) \cdot 10 + 50 \) and \( PR \) (unit fractions) were calculated. The obtained results were included in the table of risk rank values (figure 6), which characterize the degree of adverse factors coincidence in each OTU.

| №  | POINT_X | POINT_Y | R   | PR   |
|----|---------|---------|-----|------|
| 41935 | 38837,5 | 31663,5 | 85,3 | 0,85 |
| 41936 | 38838,5 | 31663,5 | 85,2 | 0,85 |
| 41937 | 38839,5 | 31663,5 | 84,6 | 0,85 |
| ...   | ...     | ...     | ... | ...  |
| 41561 | 38963,5 | 31702,5 | 29,1 | 0,29 |
| 41562 | 38964,5 | 31702,5 | 29,1 | 0,29 |
| 41963 | 38961,5 | 31701,5 | 28,5 | 0,29 |
| ...   | ...     | ...     | ... | ...  |

**Figure 6.** A table fragment of risk rank.

Based on table 4 in the ArcGIS 10.0 environment, a map of the spatial distribution zone of the degree of adverse factors coincidence (figure 7).

**Figure 7.** Diagram of the zones location of the degree of adverse factors coincidence on the forecast horizon, based on the results of using the principal component analysis.
For testing (verification) of the proposed method, a forecast map was constructed using ANN. The same set of factors was used as the initial data. The initial data was prepared in the ArcGIS 10.0 environment. For training ANN, data about recorded DGDE were used. Training and forecast were carried out in the Advangeo software package environment [34]. The forecast results are shown in figure 8.

![Figure 8](image)

**Figure 8.** Diagram of the zones location of the degree of adverse factors coincidence on the forecast horizon, based on the results of using an artificial neural network.

### 4. The discussion of results

As shown in figures 7 and 8, the most dangerous areas identified by both the ANN and the principal component analysis are well correlated with the recorded DGDE. This indicates the validity and reliability of the method. However, in contrast to the results obtained with the ANN help, the proposed method does not require prior training on previously recorded DGDE and can be applied at the stage of preliminary prediction.

The proposed method allows to make a list of mine working sites that require further investigation by other methods to quantify the DGDE risks and to plan a more rational mining development.

### 5. Conclusions

The purpose of the performed work: first, to create a forecast map of the mining enterprise territory exposed to dangerous geodynamic events using the principal component analysis in the absence of information about the occurred DGDE; second, to test the proposed approach by comparing the results of mapping with the results obtained using artificial neural networks.

The initial data was converted by sampling the territory into separate point OTU arranged in a grid of point objects with an interval of 1×1 m. The transformations were due to the fact that weighted overlay tools in modern GIS (for example, ArcGIS) only work with integer raster values. The principal component analysis was used to calculate the objective values of the weight coefficients of each factor and the thematic map construction function.
The obtained results showed a good coincidence of areas with a high probability of DGDE occurrence and recorded DGDE. This indicates that the proposed methodical approach is reasonable and can be considered reliable.

The proposed approach should be considered as the first step in predicting the DGDE occurrence on the territory. In the identified areas, more detailed and "highly specialized" studies should be conducted to clarify the results of the first stage. In addition, the proposed approach provides a working tool for decision makers, designers and engineering specialists to create plans for the enterprise development and will allow to solve the forecasting problems not only in a 2D model, but also in three-dimensional model.

6. Reference
[1] Aivazyan S A, Bezhaeva Z I, Staroverov O V 1974 Classification of multidimensional observations (Moscow: Statistics) 240 p
[2] Aivazyan S A, Bukhshtaber V M, Enyukov I S, Meshalkin L D 1989 Applied statistics Classification and reduction of dimension (Moscow: Finance and statistics) 607 p
[3] Bryson J F Method of Eliminating Coal Bumps or Minimizing Their Effect Transactions of the American Institute of Mining Metallurgical and Petroleum Engineers 119:40-57
[4] Buttinelli M, Procesi M, Pignone M, Quattrocchi F 2013 Spatial analysis and weighted overlay models as decision-making tools for geothermal resources exploration and exploitation: The Southern Italy case study European Geothermal Congress (Pisa)
[5] Kapralov E G, Koshkarev A V, Tikunov V S etc. 2010 Geoinformatics: in 2 books The first book: textbook for students of higher educational institutions edited by V Tikunov 3rd ed. revised and supplemented (Moscow: publishing center "Academy")
[6] 1997 Mining science: development and conservation of the earth's interior Edited by K N Trubetskoy (M.: publishing house Of the Academy of mining Sciences) 478 p
[7] Ereemenko V A Natural and technogenic factors of rock bursts occurrence in the iron ore deposits development in Western Siberia Mining information and analytical Bulletin (scientific and technical journal)
[8] Gao W, Han J 2020 Prediction of Destroyed Floor Depth Based on Principal Component Analysis (PCA)-Genetic Algorithm (GA)-Support Vector Regression (SVR) GeotechGeolEng https://doi.org/10.1007/s10706-020-01227-3
[9] Iannacchione A T and DeMarco M J 1992 Optimum Mine Design to Minimize Coal Bumps: A Review of Past and Present US Practices Paper in New Technology in Mine Health and Safety, Proceedings of the Society of Mining Engineers Annual Meeting, Phoenix, AZ, Chapter 24 pp 235-247
[10] Iannacchione A T, Tadolini S C 2008 Coal Mine Burst Prevention Controls Proceedings of the 27th International Conference on Ground Control in Mining (Morgantown, West Virginia)
[11] Jolliffe I T 2002 Principal Component Analysis, Series: Springer Series in Statistics, 2nd ed., (Springer, NY) XXIX 487 p
[12] Klyushnikova E V, Shitova E M 2016 Methodological approaches to calculating the integral indicator, ranking methods Electronic scientific and practical journal "InnoCenter" 1(10) TverskoyInnoCenter pp 4-18
[13] Kopytov A I, Voityov M D, Morozov S S Natural and technogenic factors ofrock bursts occurrencein the development of iron ore deposits in GornayaShoria and Khakassia Bulletin of the Kuzbass state technical University 10.0000/cyberleninka.ru/article/n/природныйе-i-tehnogennyе-faktory-vozniknoveniya-gornyh-udarov-pri-razrabotke-zhelezonorudnyh-mestorozhdeniy-gornoy-shorri-i-haka.
[14] Korshunov G I, Loginov A K, Shik V M, Artemev V B 2011 Geomechanics at coal mines (Moscow: publishing house "Gornoedelo", LLC "Kimmeriyskiy center") 388 p
[15] Kuznetsov S V, Odintsev V N, Slonim M E, Trofimov V A 1981 Methodology for calculating mining pressure (Moscow: Nauka) 104 p
[16] Kurlenya M V, Seryakov V M, Eremenko A A 2005 Technogenic geomechanical stress fields (Novosibirsk: Nauka) 264 p
[17] Lapin E S, Sapozhnikov M G 2011 Ranking of factors affecting the rock burst occurrence in the underground development of coal seams GornyZhurnal 2 pp 4-8
[18] Li Wu, Chang Mao Xu, Zhi Gang Li, Qing Jun Zuo 2014 Comprehensive Prediction of Rockburst in High-Speed Railway Tunnel Applied mechanics and materials ISSN: 1662-7482 Vols 477-478 pp 553-557 doi: 10.4028/www.scientific.net/AMM.477-478.553 (Nras Publications, Switzerland)
[19] Murthy K S R, Amminedu E, Venkateswara Rao V 2003 Integration of thematic maps through GIS for identification of groundwater potential zones Journal of the Indian Society of Remote Sensing vol 31 pp 197–210
[20] Orlov A I 1996 Expert assessments Magazine "Zavodskaya laboratoriya" Vol. 62 1 pp 54-60
[21] Petukhov I M, Smirnov V A, Vinokur V Sh et al. 1975 Geophysics of rock burst (Moscow: Nedra) 134 p
[22] Petukhov I M, Batugina I M 1999 Subsoil geodynamics (Moscow: Nedra) 288 p
[23] Raš T 1993 Method of hierarchies analysis (M.: "Radio isvyaz") 314 p
[24] Saati T 1990 Method of hierarchies analysis (M.: "Radio isvyaz") 314 p
[25] Stepanov P P 2017 Artificial neural networks Young scientist 14(138) pp 185-187
[26] Surkov A A 2019 Combining economic forecasts using expert information Statistics and Economics Vol. 16 5 pp 4-14
[27] Shatov V, Guriev G, Yakubchuk A, Dolgopolova A GIS package on mineral deposits database and thematic maps of Central Eurasia Mineral Deposit Research: Meeting the Global Challenge pp 1331-1334
[28] Shit P K, Bhunia G S, Maiti R 2016 Potential landslide susceptibility mapping using weighted overlay model (WOM) Model. Earth Syst. Environ 2 21 https://doi.org/10.1007/s40808-016-0078-x
[29] Trofimov A M 1985 Computer use in zoning the territory: Methods of automatic zoning (Kazan: publishing house of Kazan University) 72 p
[30] Trubetskoy K N, Viktorov S D, Osokin A A, Shlyapin A V 2017 "Rock burst forecast based on control of submicron particle emission during deformation and destruction of rockmasses” Gorny Zhurnal 6
[31] Turchaninov I A 1978 About the relationship of the stress state and rock mass properties Journal of Mining Science 2 pp 20-25
[32] 1991 Instructions for safe mining operations in the deposits of Gornaya Shoria that are prone to rock bursts Vostnigri, VNIMI (Novokuznetsk) 90 p
[33] Whyatt J K, Blake W, Williams T J 2002 White Mining Publication: 60 Years of Rockbursting in the Coeur D'Alene District of Northern Idaho, USA: Lessons Learned and Remaining Issues Proceedings of the 109th Annual Exhibit and Meeting, Society for Mining, Metallurgy, and Exploration Phoenix, AZ Preprint 02-164 10 p
[34] www.beak.de/advangeo
[35] https://desktop.arcgis.com/ru/arcmap/10.3/tools/spatial-analyst-toolbox/how-weighted-overlay-works.htm
[36] Sidorov D V, Potapchuk M I, Sidlyar A V, Kursakin G A Assessment of rock burst hazard in the development of deep horizons of the Nikolaev Deposit Journal of Mining Institute 1(238) pp 392
[37] Tyupin V N Assessment of the critical depth of deposits under the rock burst hazard condition Journal of Mining Institute 1(236) p 167