International Journal of Geology, Earth & Environmental Sciences ISSN: 2277-2081
An Open Access, Online International Journal Available at http://www.cibtech.org/jgee.htm
2025 Vol. 15, pp. 204-212/Ergashov and Muminkulova
Research Article

STUDY OF GEOCHEMICAL SPECIALIZATION AND PREDICTION OF ANOMALOUS ZONES OF RARE EARTH ELEMENT MINERALIZATION IN BLACK SHALES USING TRADITIONAL METHODS AND ARTIFICIAL INTELLIGENCE: FINAL RESULTS

[1]Z.Y. Ergashov* and [2]N.G. Muminkulova
National University of Uzbekistan named after Mirzo Ulugbek[1], LLC "Topfin Result"[2]
**Author for Correspondence*

ABSTRACT
This article presents the results of research conducted to study the geochemical characteristics of rare earth elements (REE) mineralization in black shales of the Ustuk area, located in the central part of the Northern Nurata mountain range, and to develop search criteria for their exploration. The research involved analytical results (ICP-ms) and geological data from more than 6,000 samples collected from drill holes and surface mine workings during geological exploration work in the area. The obtained data were prepared in «MS Excel» and subjected to statistical analysis using the TIBCO Statistica 14 program. To perform the geostatistical stage of the research, a digital three-dimensional (3D) geological-geochemical model of the study area was developed in GIS environments Datamine Studio RM and Leapfrog Geo. Geochemical anomalous zones of REE were predicted using artificial intelligence (AI) technologies. Within the framework of this research, as a result of studying geological and geochemical data, the main and accompanying geochemical associations of REE were determined and interpreted together with the conditions of their formation. Metasomatic and skarn processes developing along a hinge structure that intersects the black shales of the lower Suvliksay subformation ($R_3\,sv_1$) at a gentle angle under the influence of the Sentob intrusive were identified as the main factors of mineralization. In the scientific work, a modern approach to identifying geochemical anomalies and the application of AI technologies as a positive-practical solution found confirmation.

Keywords: *Black Shales, Rare Earth Elements (REE), Geochemical Search Criteria, 3D Geological-Geochemical Model, Artificial Intelligence (AI), Machine Learning Models (GLM, SVM)*

INTRODUCTION
As a result of the scientific and technological progress of recent decades, there has been a steady increase in the demand for rare metals and rare earth elements (REE), which, in turn, enhances the relevance of searching for new sources of these raw materials.
At the initiative of the Ministry of Mining Industry and Geology of the Republic of Uzbekistan, exploration works are being conducted in promising areas to identify industrial mineralization of REE. Specifically, the Ustuk area, located in the northern part of the Nuratau mountain range, is one of the key objects of geological research concerning its potential for REE.
Historically, uranium mineralization has been established in this area since the mid-20th century, occurring alongside associated elements such as molybdenum and vanadium in black shales. During geological exploration from 2011 to 2020, a joint occurrence of uranium mineralization and rare earth elements was identified in the Ustuk area. However, the geochemical characteristics of REE mineralization in black shales, as well as the geochemical search criteria, have not been sufficiently studied. In this regard, there is a need for in-depth analysis of the geochemical aspects of REE mineralization and the development of criteria for their exploration.

In this study, traditional statistical methods, such as descriptive statistics, correlation matrix, and factor analysis, as well as artificial intelligence technologies for predicting geochemical anomalous zones of REE, were used to study the specialization of REE in various rock types, patterns of their concentration, relationships with uranium-molybdenum-vanadium mineralization, and other chemical elements. The identification of these zones using geostatistical methods is complex due to phenomena such as spatially uneven sample distribution and noise effects.

Geological Setting of the Ustuk area

The Ustuk area is located in the central part of the Northern Nuratau mountain range, which stretches in a northwest direction for more than 100 km and has a width of 7–40 km.

The area lies within the Zarafshan-Turkestan structural-formational zone (SFZ), in the uranium-bearing region of Ustuk-Fozilmon (95 km^2) of the Uchmola-Ustuk uranium-bearing structural node, in the watershed area with absolute elevations ranging from 1,600 to 2,114 meters.

The central and western parts of the Northern Nuratau mountain range are composed of Proterozoic, Paleozoic, and Cenozoic formations. The studied area is primarily composed of Proterozoic (lower Suvliksay subformation – R_3 sv_1), Paleozoic (Konsoy formation - O_{1-2} kn) formations, and intrusive formations of the Shurak subcomplex (γδC3 s) that cut through them, as well as eluvial-deluvial Quaternary deposits resulting from weathering processes.

The black shales of the lower Suvliksay subformation, which are associated with the REE mineralization, are widespread in the central part of the area along the main watershed from southeast to northwest, forming the core of the anticline. Since the rocks of the subformation are predominantly dark gray and black in color, they are difficult to distinguish with the naked eye due to alteration caused by metamorphism, and the main composition consists of carbonaceous-siliceous shales, the term "black shales" is commonly used, although their composition includes various types of rocks. The lower part of the formation consists of gray, dark-gray, banded aphanitic quartzites, dark-gray massive fine-grained dolomites and dolomitized limestones, as well as dark-gray, sometimes greenish-gray banded mica-quartzites and carbonaceous shales. Dolomites and dolomitized limestones appear in the section as lenses and lens-shaped layers, reaching up to 1.5 km in length and thickness ranging from 0.5 m to 9–11 m. Between the layers of quartzites and dolomites, mica-quartz, mica-quartz-feldspar, and siliceous shales are found in lens-shaped layers, extending 2.5–3.0 km in length and ranging from 0.1 m to 3–4 m in thickness. The thickness of the subformation reaches up to 900 m.

The deposits of the Konsoy formation form the wings of the anticline, on which the studied area is located, and are divided into two lower and upper subformations based on lithological, mineralogical-petrological, and textural-structural characteristics.

The area under study and its surroundings are covered by formations of the Shurak Late Carboniferous adamellite-granodiorite subcomplex. These consist of the Ustuk intrusive massif in the southern part of the area, the Sentob massif in the northern part, and dike formations.

Tectonically, the central and western parts of the Northern Nuratau mountain range exhibit features of both Caledonian folding of the uplift-fault type and magmatogenic dome-like Hercynian folding. The Sharyazh fault, which runs almost parallel to the main axis of the anticline structure and extends in a northwest direction, divides the geological units of the area into two blocks.

As a result of recent work conducted in the Ustuk area, a REE mineralization zone has been identified in the central part of the area along the main axis of the anticline, from southeast to northwest. This zone has a width ranging from 5 to 70 meters and extends up to 4,000 meters.

MATERIALS AND METHODS

Geostatistical Studies

The geological-geochemical studies of the rare earth element mineralization developed in the Ustuk area are based on the analysis of data from 6,797 different samples collected during geological exploration works. The studies were carried out in the following stages:

The initial stage involved preparing the sample data. Samples that lacked analytical results were identified from the database, and 5,910 samples with ICP-MS analysis results were retained for the next stage. The chemical element quantities in the laboratory analysis results, where different notations were present, were converted into numerical form.

In the final stage of data preparation, after studying the distribution patterns of chemical element concentrations, the main dataset was accepted as the set of values within the 6 σ interval on the histogram of normal distribution. The 6 σ interval, adopted as the main dataset, refers to the interval that is smaller and larger by 3 standard deviations from the median value (Me±3σ). Here, Me represents the median of the dataset, and σ represents the standard deviation.

In the final stage of database preparation, all samples were divided into 14 distinct groups of rock types.

Based on the results of the descriptive statistical analysis and normal distribution histograms, median values (Me) were adopted as the local geochemical background. Minimum anomalous values were calculated using the formula Me + σ, where Me + σ represents the threshold for minimum anomalous concentrations.

To understand the distribution patterns of chemical elements across various rock groups in the Ustuk area, different plots were constructed based on upper percentile values (Me + 2σ, approximately the 97.7th percentile), derived from descriptive statistical analysis, and interpreted as follows:

RESUTLS AND DISCUSSION

High concentrations of REEs, along with associated elements such as U, V, and Mo, are recorded in dolomitized limestones, limestones, quartzites, and in metasomatic formations developed within the terrigenous rocks of the lower Suvliksay subformation.

In most cases, the total content of light rare earth elements (LREE) exceeds that of heavy rare earth elements (HREE), with this contrast being most pronounced in intrusive formations. An exception is observed in quartzites and microquartzites of the lower Suvliksay subformation, where the total HREE content slightly exceeds that of the LREE (Figure 1).

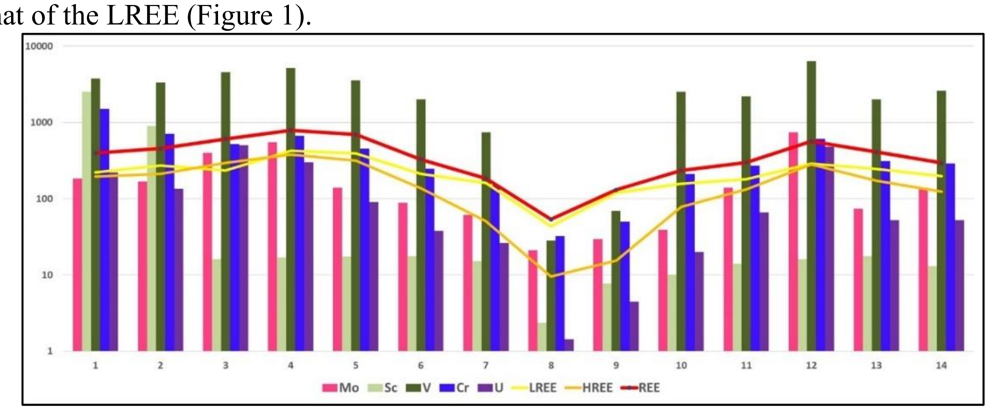

Figure 1. Comparative plot of the upper percentile values of REEs and their major associates in various rock types

Source: based on the research results obtained during the dissertation work of Z.Y. Ergashov.

The major contribution to the total content of light rare earth elements (LREE) is represented by cerium (Ce), lanthanum (La), and neodymium (Nd). The sum of the upper percentile values of LREE varies from 44 to 426 g/t, while the upper percentile values of Ce, La, and Nd range from 20 to 155 g/t, 11 to 136 g/t, and 9 to 96 g/t, respectively. Relatively high concentrations of samarium (Sm) and praseodymium (Pr) are observed in dolomitized limestones, limestones, and metasomatic rocks formed within the terrigenous deposits of the Lower Suvliksay subformation. In addition, another zone of elevated Pr content is associated with biotite–muscovite granite dikes.

Within the total content of heavy rare earth elements (HREE), yttrium (Y) constitutes the predominant share across all rock groups. The upper percentile values of Y vary from 5 to 281 g/t, while the sum of the upper

percentile values of HREE ranges from 10 to 379 g/t. The highest concentrations of lutetium (Lu), europium (Eu), terbium (Tb), holmium (Ho), and thulium (Tm) are recorded in carbonaceous–siliceous shales of the Lower Suvliksay subformation. Relatively high concentrations of Ho and Tm occur in all terrigenous and carbonate formations of the subformation, as well as in metasomatic rocks developed within its terrigenous deposits.
The total upper percentile values of REEs vary from 53 to 790 g/t, with the highest overall concentrations corresponding to dolomitized limestones of the Lower Suvliksay subformation, followed by limestones of the subformation, quartzites, and metasomatic formations developed in terrigenous rocks. The lowest concentrations are characteristic of adamellites and granitoids of the first phase of the Shurak subcomplex.

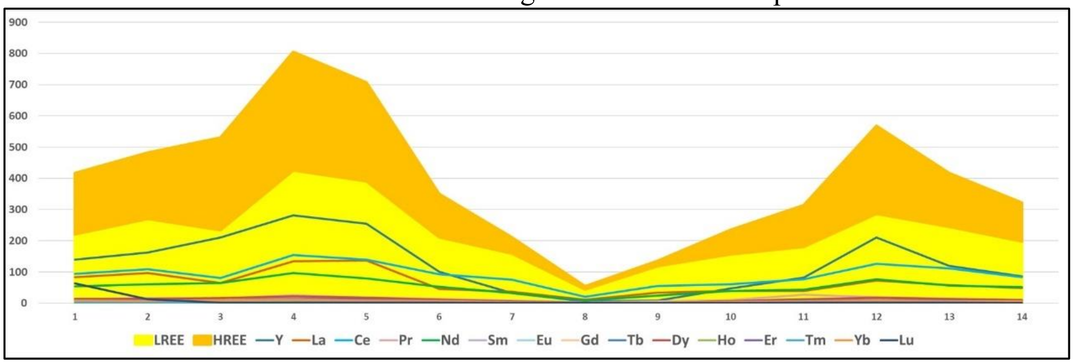

Figure 2. Comparative plot of the percentile values of REEs in various rock types
Source: based on the research results obtained during the dissertation work of Z.Y. Ergashov.

Methodology and results of geostatistical studies
Using geological data (geological maps, drilling and surface excavation data), a 3D digital geological model of the study area was constructed with the software packages Datamine Studio RM and Leapfrog Geo.
At the initial stage, separate wireframe models were created for different formations, i.e., for suites of sedimentary–metamorphic rocks and for phases of intrusive bodies. In constructing the wireframe models in Datamine Studio RM, the relationships between formations based on their geological age and structural position were taken as the foundation.

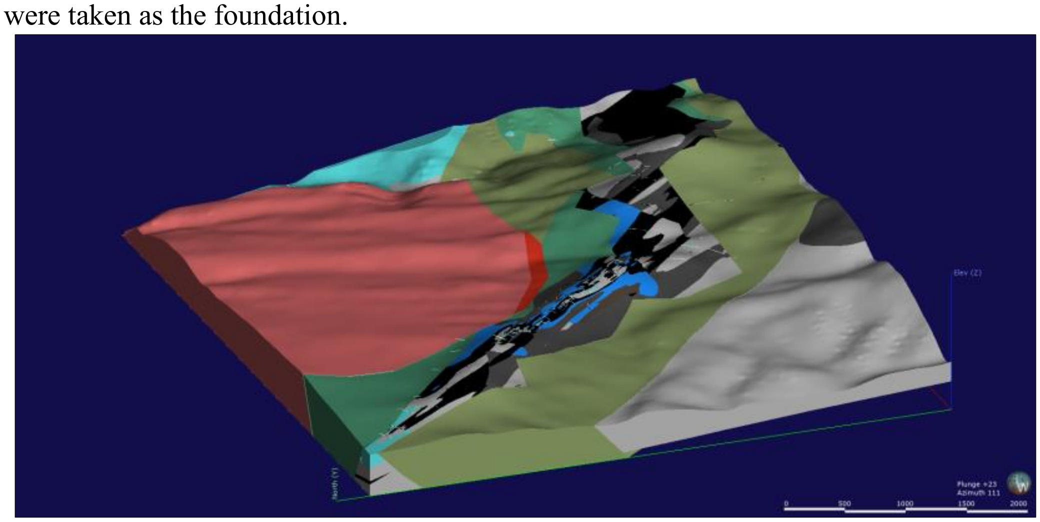

Figure 3. Digital 3D geological model of the Ustuk area constructed using Leapfrog Geo software
Source: based on the research results obtained during the dissertation work of Z.Y. Ergashov.

Since the rocks of the Suvliksay subformation—composed of interlayers of quartz–sericite schists, quartzites, limestones, dolomitized limestones, as well as carbonaceous–siliceous shales, together with the tectonic fractures and metasomatic alteration zones developed within them—are associated with mineralization, and since geological exploration within these rocks was conducted in a relatively systematic manner, a digital geological model of the various layers forming the subformation was constructed in Leapfrog Geo based on lithological sample descriptions and dip/strike data (see Fig. 3).

For the purpose of studying the spatial relationships of rare earth elements and their associated elements, data on the REE group (Y, La, Ce, Pr, Nd, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb, Lu, REE), as well as companion elements such as U and Cr, identified through statistical analysis, were interpolated in three-dimensional space using *Leapfrog Geo* software. As a result, a digital 3D geochemical model of the Ustuk area was constructed (Fig. 4).

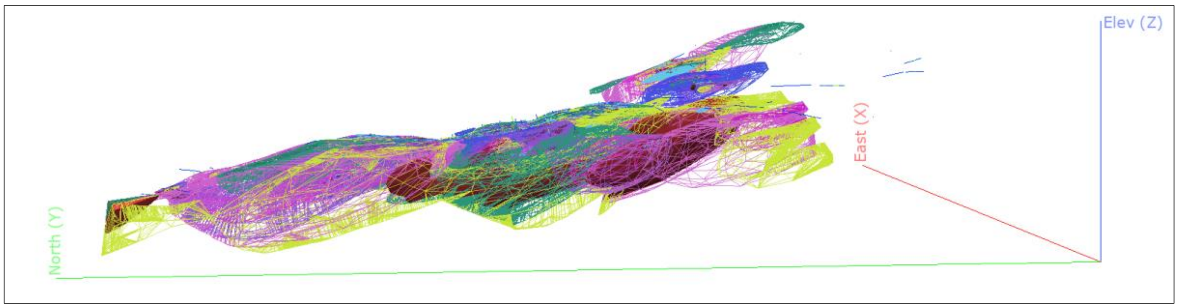

Figure 4. Digital 3D geochemical model of the Ustuk area

(Source: based on the research results obtained during the dissertation work of Z.Y. Ergashov)

After completing the construction of the 3D geological–geochemical model of the Ustuk area, the following conclusions were drawn, based on different geological cross-section orientations, regarding the spatial distribution of primary halos of REE dispersion and their relationship to geological–structural, tectonic, and magmatic factors:

Anomalous REE concentrations are confined to the carbonate–terrigenous rocks of the Lower Suvliksay subformation and, in areas more distal from the intrusion, are primarily associated with carbonaceous–siliceous shales (Fig. 5a, b, c), while in proximal zones they occur in dolomitized limestones and metasomatic rocks developed within the terrigenous deposits of the subformation. Elevated REE contents in metasomatic rocks formed near the intrusion indicate the important role of magmatic processes, which induced metasomatism and skarn alterations, in the formation of mineralization.

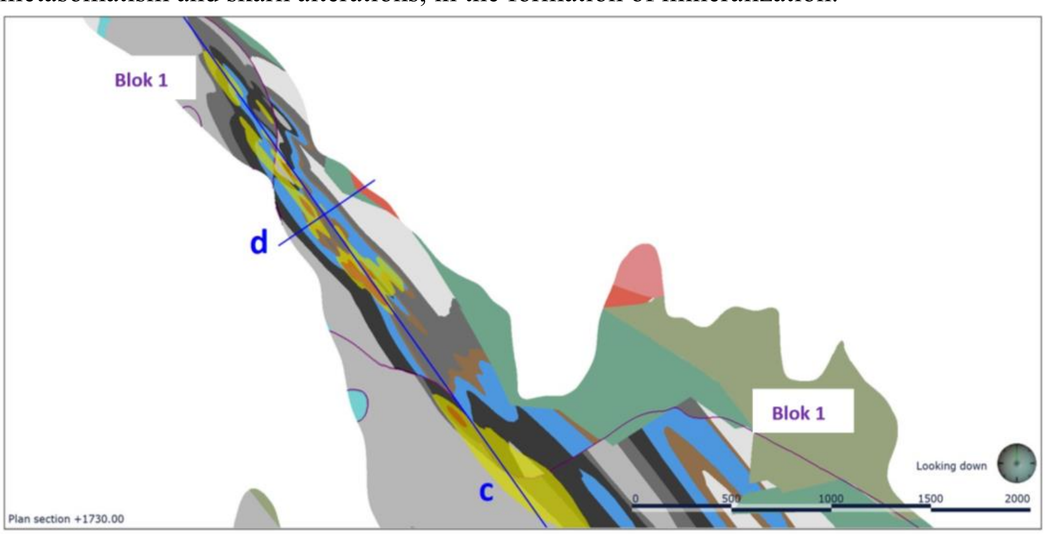

Figure 5a) Horizontal geological cross-section at the +1,730 m level

The localization of anomalous zones of primary REE halos along the hinge fault that transects the central part of the area and divides the core of the Ustuk anticline into two blocks confirms the significance of tectonic factors in the mineralization process.
The higher intensity of mineralization in the second block compared to the first, as well as the predominance of anomalous REE zones within the second block, confirm the control of mineralization by the geological–structural position of the Lower Suvliksay subformation (Fig. 5b, c).

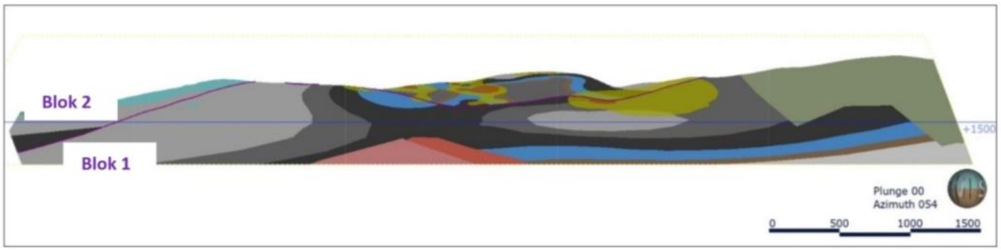

Figure 5b) Vertical geological cross-section along line c

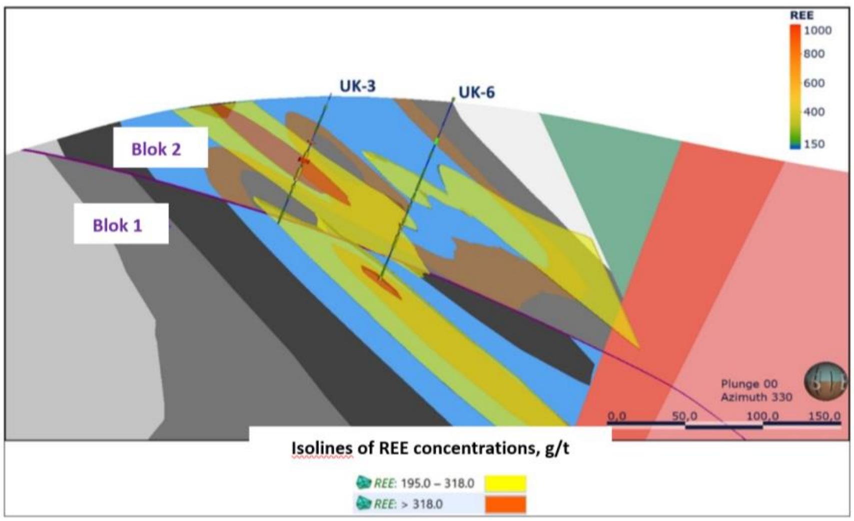

Figure 5c) Horizontal geological cross-section along line d

Figure 5. Diagrams illustrating REE mineralization in various rock types based on the digital 3D geological–geochemical model of the Ustuk area
Source: based on the research results obtained during the dissertation work of Z.Y. Ergashov.

Methodology and results of processing geological–geochemical data using artificial intelligence technologies
Although in traditional practice such methods as ordinary interpolation, kriging, and inverse distance weighting (IDW) are commonly applied for constructing contour lines in 2D and iso-volumes in 3D space of geochemical halos, in a number of cases the reliability of interpretation proves to be unsatisfactory.
If the concentration of an interpolated chemical element in the database is unevenly distributed across the geochemical field—i.e., if variable values differ sharply even over short distances—the so-called "nugget effect" (random dispersion) is observed. In datasets with such dispersion, the kriging method produces smoothed (biased) results. In these cases, high values are underestimated and low values are overestimated, which hinders the identification of anomalous areas, disrupts the integrity of geochemical anomalies, and negatively affects the quality of geochemical maps, as can be seen in Fig. 5c.

Such issues may reduce the effectiveness of interpreted geochemical (cartographic) results, meaning that in some cases anomalous concentrations that actually exist may not be encompassed by the corresponding iso-volumes. As a result, geochemical anomalies may not be fully represented, complicating the evaluation of geochemical exploration criteria developed from the study area data.

Below are scientific works by several foreign specialists dedicated to addressing this problem, which also confirms its relevance in international practice:

"The smoothing effect of data interpolation could cause useful information loss in geochemical mapping…" (Cheng Li *et al.,* 2021).

"The use of kriging is correct when the data density is sufficient with respect to heterogeneity of the spatial distribution of the geochemical parameters. However, if anomalous geochemical values are focused in hotspots of which boundaries are insufficiently densely sampled, kriging could provide misleading maps with the real contours of hotspots blurred by data smoothing…" (Stojdl *et al.,* 2017).

To address the above-mentioned issues and improve the reliability of interpolated geochemical data, the interpretation of geological–geochemical information was carried out using a non-traditional approach—artificial intelligence technologies.

At the first stage, based on the 3D geological–geochemical model of the Ustuk area, a block model was created in *Leapfrog Geo* software. This model incorporated the values of primary geochemical halos of REEs and associated elements, geological characteristics of the rock types, and spatial coordinates.

Each block was treated as one interpolated value ("estimated sample"), and its statistical parameters (mean, median, lower and upper quartiles, etc.) were compared with the corresponding parameters of 5,910 measured samples ("measured sample") collected during geological exploration. It was shown that the values of most chemical elements in the "estimated sample" were lower than those in the "measured sample."

To identify interrelationships in the "estimated sample," correlation and cluster analyses were performed, and their results were compared with those obtained from the "measured sample." Overall, it was established that REE group elements display similar interrelationships among themselves and with associated elements such as U, V, and Cr.

For initiating research using artificial intelligence technologies, the "measured sample" database was adopted as the training dataset, while the "estimated sample" database was used as the test dataset. In AI-based data analysis, it is important that "training data" and "test data" share analogous attributes and patterns of interrelationships, as this directly affects prediction accuracy. In this regard, databases were created that included chemical element values, coordinates, and lithological characteristics, all exhibiting comparable relationships.

Since the concentrations of chemical elements calculated for the "measured sample" were higher than those for the "estimated sample," all numerical attributes in both datasets were normalized separately. During normalization, the minimum values were set to 0 and the maximum values to 1, with all other values scaled within the range of 0 to 1.

At the next stage, in order to select a machine learning model most suitable for the nature of the interrelationships between attributes, the "training data" were randomly divided into two subsets: 60% for model training and 40% for prediction. Initial testing was then performed using various machine learning models.

Across all error categories—except for relative error—as well as in terms of the shortest prediction time, the Generalized Linear Model (GLM) demonstrated the best performance. By relative error values, the advantage was shown by the Support Vector Machine (SVM). However, SVM proved to be the most resource-intensive model in terms of prediction time.

The graph below compares the REE values with the predicted values obtained using the two models that demonstrated the best results.

Research Article

Table 1: Preliminary evaluation results of artificial intelligence models

№	**Machine learning model name**	**Abbreviation**	**RMSE**	**Absolute error**	**Relative error**	**Squared error**	**Total Time (sec)**
1	Generalized Linear Model	**GLM**	3,14	1,89	2,1%	14,40	2
2	Deep learning	**DL**	14,13	7,63	6,2%	202,22	7
3	Decision Tree	**DT**	61,07	16,41	7,4%	4 017,99	4
4	Random Forest	**RF**	61,19	18,61	9,6%	3 852,35	13
5	Support Vector Machine	**SVM**	53,63	5,70	1,4%	3 598,73	105

Source: based on the research results obtained during the dissertation work of Z.Y. Ergashov.

The comparison graph of predicted REE contents obtained using artificial intelligence from the “Test data” shows that the SVM model produces significant errors when predicting high REE values (900 g/t and above), whereas the GLR model provides nearly accurate predictions across the entire value range.
The GLR model was selected as the machine learning model for prediction, and forecasting was carried out based on interpolated data (“Estimated sample”). In the “Estimated sample,” the mean REE content is 134.36 g/t, with a variation range from 19.69 to 928.81 g/t. In the “Predicted sample,” the mean value is 243 g/t, with a variation range from 15.91 to 2,568.78 g/t. For comparison, in the “Measured sample,” the

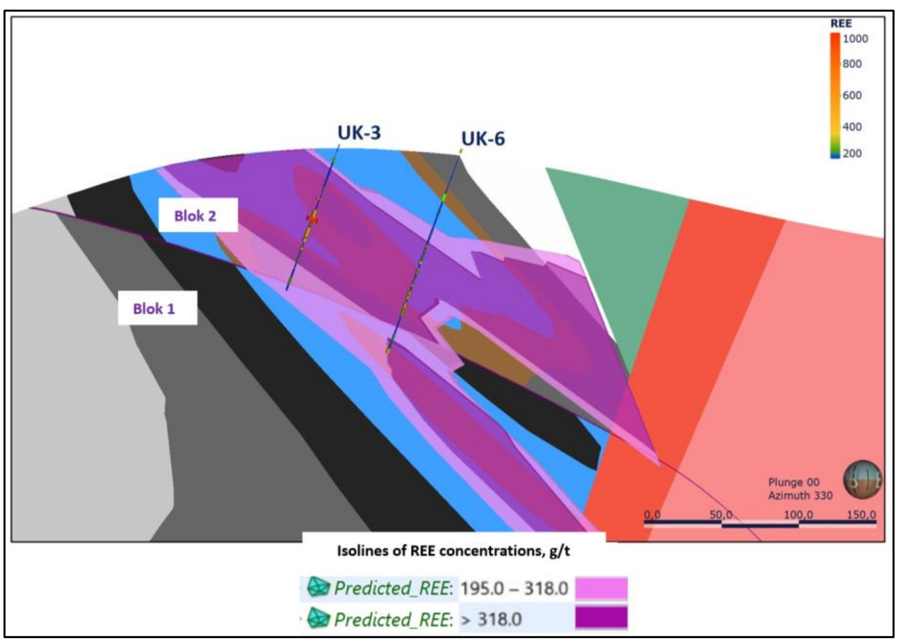

Figure 6: Vertical geological cross-section showing REE anomalies obtained based on prediction using artificial intelligence
Source: based on the research results obtained during the dissertation work of Z.Y. Ergashov.

mean REE content is 164 g/t, with a variation range from 3.0 to 3,190.0 g/t. This is explained by the fact that the REE values in the "Estimated sample" were excessively smoothed during interpolation. The higher mean value in the "Predicted sample" compared to the "Measured sample" is explained by the fact that the former dataset mainly covers the Lower Suvluksay sub-suite, where high REE concentrations are concentrated, whereas the latter covers all lithologies of the study area.

The prediction results were integrated into the "Leapfrog Geo" software, and the REE values in the "Predicted sample" were interpolated. As a result, the obtained iso-volumes of minimum anomalous REE values cover most of the spatial areas containing anomalous values from the "Measured sample" and are arranged in an orderly pattern along the main geological structures.

The iso-volumes of predicted REE contents forming a stable trend along the principal hinge once again confirm the critical role of structural-tectonic factors in mineralization, as well as the role of post-magmatic hydrothermal fluids that circulated through these structures. The increased intensity of iso-volumes in the central part of the area, near the intrusive body, indicates a significant influence of magmatic factors.

CONCLUSION

The specialization of the Proterozoic black shales of the Ustuk area in rare earth elements (REE) is determined by several factors, including the initial sediment accumulation under deep-marine conditions, as well as the subsequent redistribution of chemical elements under the influence of magmatic activity.

The research has confirmed that the modern approach to identifying geochemical anomalies through the use of artificial intelligence technologies may serve as a promising practical solution.

REFERENCES

Bhattacharya S., Sharma S., Agrawal V., Dix M. C., Zanoni G., Birdwell J. E., Wylie A. S., Wagner Jr. W., Wagner T. (2024). Influence of Organic Matter Thermal Maturity on Rare Earth Element Distribution: A Study of Middle Devonian Black Shales from the Appalachian Basin, USA. *Energies*, **17**(9), 2107. *https://doi.org/10.3390/en17092107*.

Cheng Li, Bingli Liu, Ke Guo, Binbin Li, Yunhui Kong (2021). Regional Geochemical Anomaly Identification Based on Multiple-Point Geostatistical Simulation and Local Singularity Analysis – A Case Study in Mila Mountain Region, Southern Tibet. *Minerals*, **11**(10), 1037.

McManus J., Berelson W. M., Severmann S., Poulson R. L., Hammond D. E., Klinkhammer G. P., Holm C. (2006). Molybdenum and uranium geochemistry in continental margin sediments: Paleoproxy potential. *Geochimica et Cosmochimica Acta,* **70**(18), 4643–4662. *https://doi.org/10.1016/j.gca.2006.06.1564.*