ABSTRACT

Dissertation submitted for the degree of Doctor of Philosophy (PhD) in the educational program 8D07201 «Geology and Exploration of Mineral Deposits»

BEKISHEV YERKEBULAN TEMIRKHANOVICH

«CREATION OF A GIS DATABASE OF RARE-METAL DEPOSITS OF THE KALBA-NARYM METALLOGENIC BELT»

General characteristics of the work.

In the context of steadily growing global demand for mineral resources, including rare, rare earth, and critical metals, the development of scientifically grounded methods to improve the efficiency of geological exploration is of particular importance. The current stage of mineral resource base development is characterized by the depletion of traditional sources, increasing complexity of geological structures, and the necessity to develop new types of mineralization, including concealed and deep-seated deposits. Under these conditions, a key priority is the implementation of integrated digital technologies that ensure the integration and interpretation of heterogeneous geological data.

The Central Kalba region of the Kalba–Narym rare-metal belt (Eastern Kazakhstan) represents one of the promising metallogenic areas, distinguished by a high degree of exploration and, at the same time, significant potential for discovering new deposits. Over more than a century of geological research, a vast amount of data has been accumulated, including geological maps of various scales, fieldwork results, and geochemical and geophysical materials. However, these data differ in reliability, acquisition methods, and formats, which leads to fragmentation and complicates their integrated use in mineralization forecasting.

Moreover, modern studies continuously generate new data, such as results of high-precision geochemistry, scanning electron microscopy, and satellite imagery

obtained using advanced space technologies. These datasets must also be integrated into a unified information framework for rare-metal objects.

In this regard, the study substantiates the need to develop a unified, continuously updated geographic information system (GIS) that ensures standardization, harmonization, and integration of diverse geological data. The implementation of such an approach enables consistent identification of information of different origins and acquisition times, improves the reliability of interpretation, and provides a basis for spatial analysis of metallogenic patterns.

Significant attention in the study is given to the application of remote sensing methods as an effective tool for regional geological mapping. The developed methodology for spectral processing of Landsat-8 satellite data, using principal component analysis and band ratio techniques, makes it possible to identify major lithological complexes and zones of spectral anomalies of Fe-oxide, Si–O, and Al–OH types. The spatial correlation of these anomalies with zones of hydrothermal alteration and ore mineralization confirms their diagnostic significance and demonstrates the potential of this approach for preliminary prediction of prospective mineralized areas.

The study also implements an approach based on the application of machine learning methods for interpreting high-resolution remote sensing data. The use of WorldView-3 satellite imagery in combination with machine learning algorithms, particularly the Random Forest method, enhances the accuracy of geological object classification, identification of alteration zones, and recognition of potentially mineralized structures. The integration of these methods with GIS technologies forms a modern digital platform for comprehensive analysis of geological systems.

Thus, the presented research is aimed at the development and testing of an integrated approach that combines geoinformation technologies, remote sensing methods, and machine learning algorithms to improve the efficiency of mineralization forecasting within the Central Kalba region. The obtained results provide a scientific and methodological basis for improving geological exploration, reducing investment risks, and expanding the mineral resource base of the region.

Relevance of the study.

The relevance of this study is determined by the need to improve the efficiency of geological exploration under conditions of depletion of traditional deposits and increasing demand for rare and critical metals. The Central Kalba region of the Kalba–Narym rare-metal belt, which possesses significant metallogenic potential, is characterized by a large volume of heterogeneous geological, geochemical, and geophysical data accumulated over a long period. However, these data are inconsistent and fragmented, which complicates their integrated use.

In this context, particular importance is attached to the development of an integrated approach based on the creation of a unified geographic information system (GIS) that enables the unification and systematization of long-term datasets, as well as the application of modern remote sensing methods and machine learning algorithms. The use of spectral analysis of Landsat-8 data and high-resolution WorldView-3 imagery in combination with machine learning algorithms (Random Forest) makes it possible to identify zones of hydrothermal alteration and potentially mineralized structures.

Thus, the study is aimed at developing scientifically grounded methods for the integrated interpretation of geological data, which is a key prerequisite for improving the reliability of mineralization forecasting and the efficient development of the region’s mineral resource base.

Therefore, the relevance of this research is defined by the necessity to develop an integrated approach combining GIS technologies, remote sensing methods, and machine learning algorithms to enhance the reliability of mineralization prediction and to identify prospective areas within Central Kalba. The implementation of this approach will improve the efficiency of geological exploration, reduce subsurface use risks, and ensure the scientifically justified expansion of the region’s mineral resource base.

Object of the study: The Asubulak–Belogorskoye ore node of the Kalba–Narym ore belt (Eastern Kazakhstan).

Subject of the study: The relationships between geological geospatial data, spectral characteristics of rocks and minerals, and their indicators identified using remote sensing data.

Purpose of the study: The aim of the study is to develop and test an integrated approach for predicting ore-bearing structures of the Kalba–Narym rare-metal belt based on the integration of geoinformation technologies, remote sensing methods (Landsat-8, WorldView-3), and machine learning algorithms (Random Forest), in order to improve the reliability of identifying prospective areas and enhance the efficiency of geological exploration.

Objectives of the study:

1) Digitization of systematized geospatial materials;

2) Creation of a web GIS based on digital maps;

3) Conducting geostatistical analysis using digital maps;

4) Calculation of spectral indices-indicators of mineralization (iron oxides, hydroxyl minerals, calcification) according to Landsat-8 data and their comparison with geological data;

5) Development of machine learning models to identify pegmatite bodies and distinguish them from granites according to WorldView-3 data.

Research methodology.

The study is based on a comprehensive analysis of geological and remote sensing materials of various scales aimed at identifying spatial patterns of rare-metal mineralization within the Kalba–Narym belt.

The research included the collection of cartographic data, their georeferencing, digitization, and integration into a unified GIS database. Based on digital maps, spatial correlation and geostatistical methods were applied.

At the regional level, Landsat-8 data were subjected to atmospheric correction, followed by the calculation of spectral indices and band combinations serving as indicators of mineralization.

High-resolution WorldView-3 imagery underwent preliminary spectral processing, after which spectral information was enhanced with geometric

characteristics. A Random Forest machine learning model, trained on sample datasets, was used to identify pegmatite bodies, allowing the recognition of spectral features associated with mineralization.

All obtained results were integrated within a GIS environment, resulting in the development of an interactive cartographic system implemented as a modern multi-level analytical platform.

Main provisions submitted for defense:

1) The integration of multi-scale geological, remote sensing, and attribute data into a unified web-based geographic information system (web GIS) of the Kalba–Narym ore belt ensures their unification and standardization within a single spatial database. This enables the comparison of heterogeneous geological information, identification of spatial patterns in ore object distribution, and the creation of a unified digital database suitable for use by subsoil users in ranking prospective areas, planning exploration activities, and reducing geological risks.

2) Spectral processing of Landsat-8 data using principal component analysis and band ratio techniques allows, at the regional level, the differentiation of major lithological complexes and the identification of spectral anomaly zones of Fe-oxide, Si–O, and Al–OH types. These anomalies spatially correlate with zones of hydrothermal alteration and ore mineralization in Central Kalba, making this approach an effective tool for preliminary geological mapping and prediction.

3) Object-based analysis of high-resolution WorldView-3 satellite data combined with the Random Forest algorithm provides more reliable identification of pegmatite bodies and their discrimination from host granitoids compared to pixel-based classification, due to the joint use of spectral, textural, and geometric features of the objects.

Scientific novelty:

The scientific novelty of the conducted research lies in the development and implementation of an integrated geoinformation approach to ore potential forecasting, based on a synthesis of classical metallogenic concepts, Earth remote sensing methods, and machine learning algorithms:

1) For the first time within the Kalba-Narym Rare Metal Belt, using Central Kalba as an example, a new-generation integrated geoinformation system has been created. This system ensures the unification, standardization, and joint interpretation of heterogeneous geological, geochemical, geophysical, and remote sensing data accumulated over more than 100 years of research. The developed system implements the principles of multi-scale analysis and spatial data consistency and served as the foundation for creating an interactive digital map of the Kalba–Narym Belt.

2) A methodology for spectral interpretation of Landsat-8 data has been developed and scientifically substantiated, incorporating principal component analysis and spectral band ratios. This methodology enables the differentiation of lithological complexes at the regional level and identifies zones of spectral anomalies of Fe-oxide, Si–O, and Al–OH types. Their spatial and genetic association with zones of hydrothermal alteration and ore mineralization has been established, confirming the diagnostic significance of spectral indicators in mineralization forecasting.

3) Machine learning methods (Random Forest) were adapted for the first time for geological mapping and rare metal mineralization forecasting using high-resolution WorldView-3 satellite data. An approach based on the integration of spectral, textural, and geometric features has been proposed, enhancing the classification accuracy of granitoid complexes and rare metal pegmatites at the regional level.

4) A conceptual model for integrating GIS, remote sensing data, and artificial intelligence algorithms has been developed. This model facilitates a transition from fragmentary analysis to systematic spatial modeling of geological objects and metallogenic patterns. It has been demonstrated that the integrated use of these methods significantly increases the reliability of forecasting ore-bearing structures.

5) New patterns in the spatial distribution of hydrothermally altered rocks and rare metal mineralization have been identified, as reflected in predictive maps derived from the integration of spectral, geological, and machine learning models. These findings expand the understanding of the factors controlling mineralization localization within the Kalba-Narym Belt, using Central Kalba as a case study.

Practical significance:

The practical significance of the research is determined by the implementation of the developed integrated approach, which combines geoinformation technologies, Earth remote sensing methods, and machine learning algorithms, as well as its applicability for solving a wide range of geological exploration problems and mineral resource management tasks.

The created geoinformation system represents a unified digital environment that ensures the collection, storage, unification, and analysis of heterogeneous geological, geochemical, geophysical, and remote sensing data. Its use significantly enhances the efficiency of geological mapping, structural analysis, and ore potential forecasting at various scales.

The developed methodology for spectral processing of Landsat-8 data enables the rapid delineation of hydrothermal alteration zones (Fe-oxide, Si–O, and Al–OH types) and can be used as an effective tool for preliminary geological mapping and early-stage exploration activities, reducing field investigation costs.

The application of machine learning algorithms (Random Forest) to WorldView-3 data improves the accuracy of recognizing lithological complexes, granitoids, and rare metal pegmatites, which significantly increases the reliability of predictive models and the effectiveness of identifying promising areas.

The developed predictive maps and models of spatial mineralization distribution can be used for planning exploration and appraisal activities, as well as for ranking territories according to their investment attractiveness.

The research results have been implemented into production at GeoProekt Corporation LLP, as well as into the educational process of D. Serikbayev East Kazakhstan Technical University within the educational program 6B07302 «Geodesy and Cartography» for the discipline «Aerospace Survey Methods».

Overall, the research results form a scientific and methodological basis for the transition to digital technologies in geological exploration and can be used to address the challenges of replenishing and expanding the mineral resource base of the Republic of Kazakhstan.

Data sources and author’s contribution.

The work is based on materials obtained by the author during the course of study and within the framework of the scientific project BR24992854 «Development and implementation of competitive science-based technologies to ensure sustainable development of mining and metallurgy industry East Kazakhstan region» carried out at D. Serikbayev East Kazakhstan Technical University.

The author completed an international research internship at Masaryk University (Department of Geography, Brno, Czech Republic).

During the preparation of the dissertation, the author reviewed domestic and international scientific publications and reference materials on the Kalba–Narym belt and its rare metal deposits. The study includes materials obtained during fieldwork at deposits, as well as data provided by colleagues.

Medium-resolution remote sensing data (Landsat-8) were obtained free of charge from the EarthExplorer and Copernicus Data Space Ecosystem portals. High-resolution WorldView-3 satellite imagery was acquired on a commercial basis within the framework of the scientific project BR24992854.

Approbation of the work and publications.

The main results of the dissertation research have been published in scientific journals and have also been presented and discussed at international scientific conferences.

Publications in international scientific journals indexed in Scopus / Web of Science databases:

1) Mizernaya M.A., Aitbayeva S.S., Kotler P.D., Dolgopolova A.V., Seltmann, R., Bekishev Y., Kuzmina O.N., Oitseva T.A., Shayakhmetova Z.A., Akbarov Y.Y., Baisalova A.O. Pegmatites of the Kalba–Narym Batholith (East Kazakhstan): Origin and Classification. Minerals, 2026, 16, 187.

Publications in scientific journals recommended by the Committee for Quality Assurance in Science and Higher Education of the Ministry of Science and Higher Education of the Republic of Kazakhstan:

1) Bekishev Ye.T., Rakhymberdina M. Y., Mizernaya M.A., Mataibayeva I. Ye. Kuzmina, O.N. Remote sensing in the search foe rare metal deposits of East

Kazakhstan // Bulletin of D. Serikbayev East Kazakhstan Technical University. – 2023. – №3. – P. 86–98.

2) Bekishev Ye.T., Rakhymberdina M. Ye., Levin E., Kapasov A.K. Creation of a database of rare metal deposits in the Kalba-Narym ore belt // The Mining Journal of Kazakhstan. – 2025. – №9 (245). – P. 33–39.

3) Assylkhanova Zh. A., Rakhymberdina M. Ye., Bekishev Ye.T., Levin E., Grokhotov Y.V. Creation of a digital geological database of the Kalba-Narym ore belt in East Kazakhstan // Universitet Enbekteri – University Proceedings. – 2025. – №4 (101).

4) Bekishev Ye., Levin E., Mataibayeva I., Orazbekova G. Study and Prospects of the Red Cordon Rare Metal Ore Occurrence in the Asubulak Ore Field // Universitet Enbekteri – University Proceedings. – 2025. – №4 (101).

Participation in international scientific conferences:

1) Asian Conference on Remote Sensing 2024 (ACRS 2024). IOP Conference Series: Materials Science and Engineering, 913(5), 052050;

2) 9th International Conference on Cartography and GIS, 16–21 June 2024, Nessebar, Bulgaria. ISSN: 1314-0604;

3) XX International Scientific and Practical Conference, 22–24 May 2023, Munich, Germany. ISBN: 978-9-40368-892-3;

4) APRSEI – PHEDCS 2025, 23–25 September 2025, Tashkent, Uzbekistan // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-5/W3-2025.

Structure and scope of the dissertation.

The dissertation consists of 122 pages and includes an introduction, 6 chapters, a conclusion, a list of 124 references, 40 figures, 17 tables, and 2 appendices.

Chapter One of the dissertation examines rare metals, their role and significance in modern industrial production, and analyzes global trends in growing demand for this group of minerals. Special attention is paid to the expanding applications of rare metals in high-tech industries, including electronics, energy, and the production of innovative materials.

Additionally, this chapter provides a review of rare metal deposits in the Republic of Kazakhstan, discussing their geological and economic characteristics, the degree of their study, and their development prospects.

Chapter Two of the dissertation outlines the research methodology. It describes the stages of collecting and analyzing geological and cartographic materials, as well as Earth remote sensing data, including satellite imagery from Landsat 8 and WorldView-3.

Information is provided on the geological field work conducted, including routes, observations, and the collection of rock samples for subsequent X-ray diffraction (XRD) analysis. The software tools used for data processing and analysis, including ENVI and QGIS, are described.

Furthermore, the methodology for processing satellite remote sensing data is discussed in detail, including preprocessing, spectral analysis, and image interpretation. Approaches to digitizing geological maps and integrating the obtained data into a geoinformation environment for subsequent analysis are described.

Chapter Three of the dissertation presents a geological description of the Kalba-Narym ore belt. The features of its tectonic structure, including the main structural elements that determine the formation and distribution of ore bodies, are examined.

The magmatic complexes of the Kalba-Narym batholith, their composition, age, and geodynamic formation conditions are discussed. Special attention is paid to the metallogenic specialization of the region, highlighting the main types of mineralization.

A separate section is devoted to rare metal pegmatites, their genetic characteristics, formation conditions, and role in the localization of rare metals. A description of the main deposits of Central Kalba, their geological structure, and mineral resource potential is also provided.

Chapter Four of the dissertation discusses geoinformation systems and their capabilities in solving geological problems. An overview of global geological databases, GIS platforms, and web resources used for storing, analyzing, and

visualizing spatial data is presented, along with the current state and brief history of geoinformation technology development in Kazakhstan.

The methodology for collecting, systematizing, and digitizing geological maps, followed by integration into a unified geoinformation environment, is described. Special attention is paid to data preparation processes and their publication in a web-GIS format, which provides convenient access to and visualization of geological information.

Based on the created digital maps, correlation, cluster, and geostatistical analyses were performed to identify patterns in the spatial distribution of geological objects and the factors controlling the localization of rare metal mineralization.

Chapter Five of the dissertation addresses regional lithological mapping and the identification of indicators of hydrothermal rock alteration based on Landsat 8 remote sensing data.

Principal component analysis (PCA) was performed to determine the information content of individual spectral bands and to interpret the contribution of each principal component to the delineation of geological objects. Classification maps reflecting the distribution of lithological units were constructed based on the obtained results.

Geological indices characterizing the content of iron oxides, hydroxyl minerals, and silicification zones were calculated using the spectral band ratio method. False-color composites were generated based on these indices, enabling the identification and mapping of hydrothermally altered rock zones.

Chapter Six of the dissertation discusses the development of machine learning models for identifying pegmatite bodies and distinguishing them from granites using high-resolution WorldView-3 satellite data.

In the first stage, models were trained using only the spectral characteristics of the images. Their effectiveness in classifying geological objects and identifying pegmatite bodies was evaluated.

In the second stage, an approach utilizing hybrid data, comprising both spectral and geometric features of the objects, was implemented. A comparative analysis of the

results showed improved classification accuracy and more reliable differentiation between pegmatites and granites when using this integrated approach.