

Application of Machine Learning Methods for Hydrothermal Alteration Zoning using Remote Sensing Data: a Case Study of the Koldar Massif

Bakhberde Adebijet¹, Elmira Orynassarova¹, Marua Alpysbay¹, Irina Kuznetsova²,
Ainur Yerzhankyzy¹, Aigerim Ilyasova¹

¹ Satbayev University, Almaty, Kazakhstan – elmiraorynbassarova@gmail.com

² International Educational Corporation, Almaty, Kazakhstan

Keywords: Hydrothermal Alteration, ASTER, Porphyry Copper, SVM, SAM, Machine Learning

Abstract

Hydrothermal alteration zones are among the key indicators in the exploration of porphyry copper deposits. In this study, a remote sensing-based approach was implemented to map hydrothermal alteration zones using ASTER satellite data and built-in classification algorithms available in the ENVI software environment. The study area is the Koldar massif, located in southeastern Kazakhstan within the Balkhash–Ili metallogenic belt, known for its intense hydrothermal alteration processes.

Four classification methods were applied: Spectral Angle Mapper (SAM), Support Vector Machine (SVM), Maximum Likelihood (ML), and Minimum Distance (MD). The training samples were generated based on geological maps, lithogeochemical data, and expert visual interpretation. The focus was placed on mapping four types of alteration: argillic, phyllic, propylitic, and potassic zones.

Among the tested algorithms, the SVM method demonstrated the highest performance, achieving an overall classification accuracy of 84.12% and a kappa coefficient of 0.79. Propylitic and phyllic zones were effectively identified, while argillic and potassic zones showed partial spectral confusion due to similar spectral characteristics. The resulting maps show good agreement with geological structures and known mineralized zones of the Koldar intrusion, confirming the applicability of the chosen approach at early stages of exploration in arid environments. This approach provides a reproducible framework for mapping hydrothermal alteration zones and can be adapted for other porphyry systems using medium-resolution multispectral satellite data.

1. Introduction

Porphyry copper deposits are among the most significant sources of copper and molybdenum worldwide and are characterized by the development of extensive zones of hydrothermal alteration, which serve as key indicators in prospecting and exploration activities (Lowell, 1970; Zvezdov 1993). The formation of such deposits is associated with granitoid intrusions within magmatic arcs, where concentric zonation of potassic, phyllic, argillic, and propylitic alterations is typically observed, reflecting the conditions of ore formation and hydrothermal activity (Sillitoe, 2010). The identification and analysis of these alteration zones are critical for enhancing the efficiency of mineral exploration, reducing the cost of detailed fieldwork, and increasing the likelihood of discovering promising targets.

Traditional methods for studying hydrothermal alteration zones involve extensive field surveys, geochemical sampling, and petrographic analysis, which require considerable time and financial resources, particularly in remote and mountainous regions (Mars, 2006). In this context, remote sensing (RS) offers unique opportunities to obtain spatially continuous data on the distribution and morphology of alteration zones over large areas with high temporal efficiency and repeatability. These capabilities make remote sensing approaches increasingly valuable in modern geological exploration systems (Kruse, 2012; Van der Meer, 2012).

The use of multispectral and hyperspectral data (ASTER, Sentinel-2, Landsat-8, Hyperion, PRISMA) significantly enhances the ability to identify minerals characteristic of various types of hydrothermal alterations by analyzing their spectral features in the VNIR and SWIR regions (Hu, 2018; Beiranvand, 2014; Rockwell, 2008). ASTER data have proven effective in detecting Al-OH, Fe-OH, and Mg-OH bearing minerals associated with phyllic, argillic, and propylitic alteration zones (Mars, 2006; Beiranvand, 2014; Testa, 2018). Sentinel-2 data,

with their high spatial resolution and frequent revisit times, are promising for mapping zones of oxidized ores and iron-bearing minerals (Bahrami, 2024; Khaleghi, 2020), while the integration of Landsat-8 and ASTER imagery improves the level of detail and accuracy in delineating alteration zones (Safari, 2018; Tompolidi, 2020). Hyperspectral data, such as those from EO-1 Hyperion, allow for more precise identification of the mineralogical composition of alteration zones, even in areas with vegetation cover and complex topography (Rejith, 2022).

As the conceptual framework for the analysis and interpretation of the spatial structure of hydrothermal alteration zones in porphyry copper systems, this study adopts the classical porphyry copper deposit model proposed by Lowell and Guilbert (Lowell, 1970) (Figure 1). This model illustrates the typical concentric zonation that develops around a magmatic center and includes potassic, phyllic, argillic, and propylitic alteration zones. These zones are associated with specific minerals (e.g., feldspar, sericite, chlorite, kaolinite, epidote, etc.) and ore deposition environments.

The potassic zone, located in the central part of the system, is characterized by the presence of potassium feldspar, quartz, biotite, and anhydrite. It gradually transitions outward into the phyllic zone, dominated by sericite and quartz, and further into the argillic and propylitic zones, where chlorite, epidote, kaolinite, and carbonates are prevalent. These altered rock zones are spatially associated with ore deposition zones, represented by a core of low-grade ores, shells of high- and low-pyrite content mineralization, and peripheral halos of porphyry-style mineralization involving copper, pyrite, molybdenite, and precious metals (Lowell, 1970).

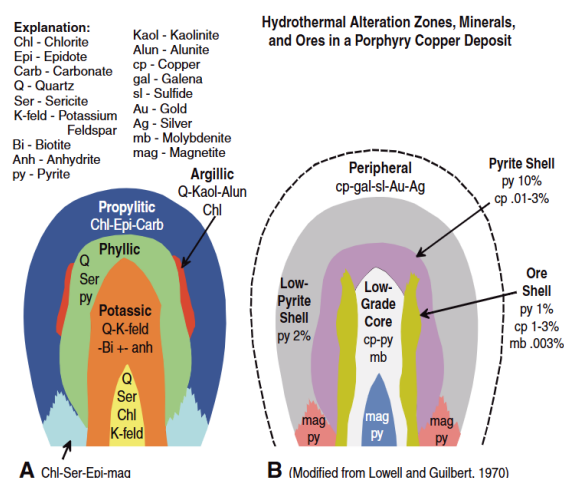


Figure 1. Schematic model of a porphyry copper deposit (after Lowell and Guilbert, 1970), illustrating the spatial zonation of hydrothermal alterations (propylitic, phyllic, argillic, and potassic) and the distribution of ore bodies associated with each alteration zone.

The application of this model in the present study enables the correlation of alteration zones interpreted from remote sensing (RS) data with the geological structure of porphyry copper systems and the localization of ore bodies, forming a basis for predicting ore potential. This approach provides a scientific rationale for delineating target zones for further mineral exploration and contributes to the overall efficiency of integrated porphyry copper exploration strategies in Kazakhstan.

Despite the high informational value of RS data, practical applications are complicated by factors such as spectral similarity among minerals, vegetation cover, atmospheric interference, and pixel mixing effects (Bedini, 2009; Van der Meer, 2012; Beiranvand, 2014). To improve the accuracy of alteration zone mapping and geological interpretation, machine learning (ML) methods have been increasingly adopted in recent years. These methods allow for the processing of large volumes of multidimensional data and enable the detection of subtle spectral features associated with mineralization (Belgiu, 2016; Fu, 2023). Support Vector Machine (SVM) and Random Forest (RF) algorithms have proven to be effective tools for classifying both multispectral and hyperspectral data due to their robustness to noise and ability to handle im-balanced datasets (Pal, 2005; Rodriguez-Galiano, 2012), while the Spectral Angle Mapper (SAM) method continues to be widely used for hyperspectral data analysis (Kruse, 1993; Shahriari, 2015).

The integration of remote sensing data with machine learning techniques enhances the precision and efficiency of hydrothermal alteration zone detection, facilitating more targeted planning of ground-based geological exploration activities and reducing their cost (Beiranvand, 2014; Testa, 2018). Nevertheless, the application of such approaches in Kazakhstan remains limited, despite the region's high ore-forming potential and favorable conditions for remote sensing, including an arid climate, sparse vegetation, and extensive rock exposures.

Despite the high efficiency of remote sensing (RS) and machine learning (ML) methods in prospecting and exploration, their application in Kazakhstan remains limited. However, the arid climate, widespread rock outcrops, and high mineral potential provide favorable conditions for the implementation of such approaches. A promising area is the Koldar Massif in Central Kazakhstan, where the presence of porphyry copper and polymetallic deposits is assumed, yet no systematic studies using

RS data and classification algorithms have been conducted to date.

The aim of this study is to adapt and test an integrated methodology for the detection and mapping of hydrothermal alteration zones associated with porphyry copper systems, using ASTER satellite data in combination with automated classification algorithms.

To achieve this aim, the following objectives were pursued:

1. Identification of phyllic, argillic, and propylitic alteration zones using the Minimum Distance (MD), Maximum Likelihood (ML), Spectral Angle Mapper (SAM), and Support Vector Machine (SVM) classification methods.

2. Analysis of the spatial relationship between the identified alteration zones and tectonic structures to delineate prospective areas.

3. Validation of results based on geological data.

The findings of this research will help assess the effectiveness of integrating remote sensing data with machine learning algorithms for mineral prospecting in Central Asia. The proposed approach can be applied to enhance accuracy, reduce exploration costs, and optimize targeting strategies in Kazakhstan and other regions with similar geological settings.

1.1 Study Area

The study area is the Koldar intrusive massif, located within the Aktogay ore field in the East Kazakhstan region. The massif is a multiphase intrusive body of laccolithic form, elongated in a sublatitudinal direction for 17–18 km and covering an area of approximately 75 km². It is composed of rocks ranging from gabbro-diorites to granites, including diorites, granodiorites, plagiogranite-porphyrines, and porphyritic granites, which were formed in several magmatic phases.

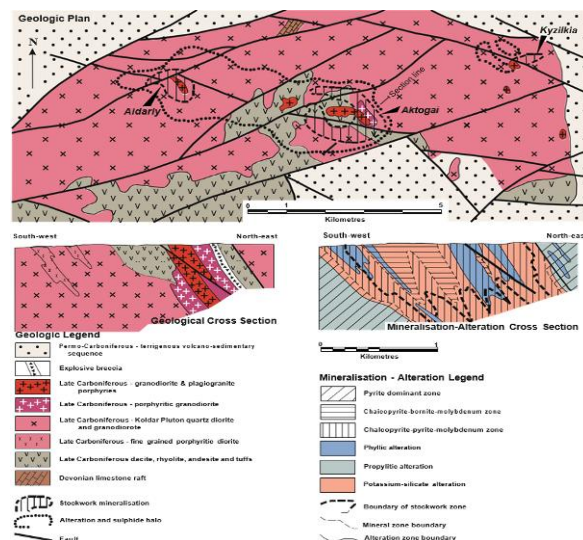


Figure 2. Geology, hydrothermal alterations, and mineralization in the Koldar massif

The Koldar massif intrudes into the volcanogenic-sedimentary deposits of the Keregetas and Koldar suites and is structurally associated with the Koldar horst-anticline. Geophysical data indicate the presence of a feeding magmatic conduit extending to a depth of up to 4.5 km, confirming the massif's active magmatic evolution. The massif is controlled by a system of regional faults (Aktogay, Koldar, Ikbass), which play a significant role in the localization of ore bodies and the development of hydrothermal alterations.

The intrusive rocks of the massif are significantly altered due to processes such as quartzification, K-feldspar alteration, biotitization, and other types of metasomatism. The Koldar massif is a source of magmatic and hydrothermal activity that led to the formation of stockwork-type porphyry copper mineralization, represented by zones of propylitization, quartz-sericite, K-feldspar–biotite, and other hydrothermal alterations. It is spatially and genetically associated with the Aktogay, Aidarli, and Kyzylkiya deposits, which together form a single porphyry copper cluster in the region.

2. Materials and methods

The primary data source used in this study was an ASTER Level 1A image acquired on August 14, 2004. This satellite product includes: VNIR (3 bands, 15 m spatial resolution), SWIR (6 bands, 30 m), and TIR (5 bands, 90 m) (Rouskov, 2005).

The main emphasis was placed on the SWIR bands, which are most sensitive to spectral features of hydrothermal alteration minerals (Al-OH, Fe-OH, Mg-OH).

Data processing was performed using ENVI 5.6 software, and included the following steps:

1. Radiometric and geometric correction,
 2. Atmospheric correction using the FLAASH module,
 3. Creation of a VNIR + SWIR image stack for further analysis.
- To eliminate the influence of vegetation cover, masking was applied based on the NDVI values calculated using the standard formula:

$$NDVI = (NIR - RED) / (NIR + RED)$$

where RED corresponds to ASTER Band 2 and NIR to Band 3N (Huang, 2021). Pixels with NDVI > 0.3 were excluded from analysis to focus the classification on exposed surface areas.

The formation of training data is a critical step in the task of automated classification of hydrothermal alteration zones. In this study, training samples were derived from a combination of geological and remote sensing data obtained from published sources (Orynassarova, 2025), including maps of hydrothermal alteration zones. Expert knowledge was additionally employed to refine the spatial boundaries of classes and to adapt the data to the geological context of the Koldar massif.

Initially, regional geological maps at a scale of 1:200,000 were used, containing information on lithology, structural features, and previously identified alteration zones. These maps enabled delineation of preliminary contours of ore-bearing areas and provided orientation within structural-tectonic blocks subjected to hydrothermal activity.

A key role in the interpretation of alteration zones was played by visual analysis of color-synthesized ASTER composites, particularly those derived from SWIR bands, where hydrothermally altered rocks exhibit distinctive spectral signatures. Complex band combinations (e.g., SWIR6/SWIR8, SWIR5/SWIR6) were effective in highlighting zones with high concentrations of clay and sericitic minerals. These images were used for expert interpretation—conducted jointly by specialists in remote sensing and geology.

Based on the resulting interpretations, training points were delineated and digitized. Each point represented a homogeneous area in terms of spectral and geological-structural characteristics, corresponding to one of four alteration classes:

Argillic – zones of clay alteration, containing kaolinite, montmorillonite, and other aluminosilicate minerals, indicative of low-temperature ore-forming conditions.

Phyllic – sericitic-quartz zones, characterized by high concentrations of sericite and quartz, typical for intermediate fluid-alteration zones.

Propylitic – outer alteration zones, represented by epidote, chlorite, and carbonates, resulting from weak metasomatic overprinting of host rocks.

Potassic – potassium alteration zones, including biotite, potassium feldspar, and sericite, typically localized near the central parts of intrusions.

Digitization was performed manually using all available data sources, followed by verification of spectral signature homogeneity and spatial consistency. As a result, a representative training dataset was constructed and used uniformly across all classification methods (SAM, SVM, ML, MD), ensuring comparability of the obtained results.

This approach enabled high-quality sample preparation, ensuring reliable identification of hydrothermal alteration zones during subsequent automated interpretation of satellite imagery.

To delineate hydrothermal alteration zones within the Koldar Massif, four classification methods were applied, reflecting both traditional approaches and modern machine learning algorithms. The use of multiple classifiers allowed for a comparative analysis of their effectiveness under conditions of complex spectral noise and overlapping mineralogical associations.

The SAM (Spectral Angle Mapper) method is based on comparing the spectrum of each image pixel with reference spectra defined from training polygons. The algorithm calculates the cosine of the angle θ between reflectance vectors.

$$\theta = \cos^{-1} \left(\frac{\vec{t} \cdot \vec{r}}{\|\vec{t}\| \|\vec{r}\|} \right)$$

where:

$\vec{t}$ – spectral vector of the pixel

$\vec{r}$ – reference spectrum (mean value for the class)

$\|\vec{t}\|$ and $\|\vec{r}\|$ – euclidean norms of the vectors.

If the angle θ is less than a given threshold, the pixel is classified as belonging to the corresponding class. This method is robust to illumination variations but is sensitive to spectral similarity between classes. The comparison is performed by measuring the angle between spectral reflectance vectors in a multidimensional space. The smaller the angle, the closer the pixel spectrum is to the reference. This method is particularly resistant to illumination variations, which is especially important when processing images over heterogeneous terrain (Liu, 2013). The SAM classification was performed using the tools provided in the ENVI Classification module.

The Support Vector Machine (SVM) method is one of the most powerful machine learning algorithms used in remote sensing applications. In this study, an SVM implementation with a Radial Basis Function (RBF) kernel was used, enabling nonlinear separation of classes in the feature space:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$$

where:

x_i, x_j – feature vectors of pixels

γ – is the kernel parameter (a hyperparameter) that controls the degree of nonlinearity of the model (Zhang, 2012).

The SVM algorithm is implemented in ENVI, followed by validation through visual comparison and statistical metrics.

The Maximum Likelihood (ML) method is based on a Bayesian probabilistic model that assumes a normal distribution of spectral values for each class. ML is a probabilistic classifier that assumes the spectral features of each class follow a multivariate normal distribution (Sisodia, 2014). ENVI implements maximum likelihood classification by calculating the following

discriminant functions for each pixel in the image (Richards, 1999):

$$g_i(x) = \ln p(\omega_i) - \frac{1}{2} \ln |\Sigma_i| - \frac{1}{2} (x - m_i)^T \Sigma_i^{-1} (x - m_i)$$

where:

i = class

x = n -dimensional data (where n is the number of bands)

$p(\omega_i)$ = probability that class ω_i occurs in the image and is assumed the same for all classes

$|\Sigma_i|$ = determinant of the covariance matrix of the data in class ω_i

Σ_i^{-1} = its inverse matrix

m_i = mean vector

Classification is performed by calculating the probability of each pixel belonging to a particular class based on the estimated means and covariance matrices. This method is sensitive to the quality of the training data statistics and requires enough training data to produce stable results. It was implemented in the ENVI environment with automatic calculation of the necessary statistical parameters.

The Minimum Distance (MD) method determines the class membership of a pixel based on the minimum Euclidean distance D between the pixel's spectrum $\vec{x}$ and the mean spectrum of a class $\vec{\mu}_k$:

$$D(\vec{x}, \vec{\mu}_k) = \sqrt{\sum_{i=1}^n (x_i - \mu_{k,i})^2}$$

where:

x_i - the pixel's spectral value in the i -th band,

$\mu_{k,i}$ - the mean value for class k in the i -th band.

The method is simple to implement but performs poorly in the presence of strong inter-band correlation or overlapping classes (Wacker, 1972).

To quantitatively assess the reliability of hydrothermal alteration maps obtained using different classification methods (SAM, SVM, ML, MD), statistical validation was conducted using a confusion matrix.

This approach allows not only for the evaluation of overall accuracy but also for identifying classification errors specific to each class.

A confusion matrix is an $N \times N$ table, where N is the number of classified classes. Each row represents the actual class, and each column corresponds to the predicted class. The diagonal shows the number of correctly classified pixels.

Accuracy metrics. Overall Accuracy (OA): indicates the proportion of correctly classified pixels relative to the total number of reference (ground truth) pixels (Fitzgerald, 1994):

$$OA = \frac{\sum_{i=1}^n x_{ii}}{N} * 100\%$$

where x_{ii} – number of correctly classified pixels of class i , N – total number of validated (reference) pixels.

Kappa Coefficient (κ) – a metric that takes into account the possibility of agreement occurring by chance (Kerr, 2015):

$$\kappa = \frac{OA - Pe}{1 - Pe}$$

where Pe – the expected probability of random classification.

Producer's Accuracy (PA) indicates how well a class on the map represents the real-world objects (Stehman, 2013):

$$PA_i = \frac{x_{ii}}{\sum_{j=1}^n x_{ij}} * 100\%$$

where $\sum_{j=1}^n x_{ij}$ – the total number of pixels that belong to class i according to the reference data.

User's Accuracy (UA) indicates the probability that a pixel classified into a given class actually belongs to that class (Barsi, 2018):

$$UA_i = \frac{x_{ii}}{\sum_{j=1}^n x_{ji}} * 100\%$$

where $\sum_{j=1}^n x_{ji}$ – the total number of pixels classified as class i .

Omission Error is the percentage of pixels that belong to a class but were not recognized as such (Congalton, 2019):

$$Omission_i = 100\% - PA_i$$

Commission Error is the percentage of pixels that were incorrectly assigned to a class (Congalton, 1991):

$$Commission_i = 100\% - UA_i$$

For each method (SAM, SVM, ML, MD), the same training and testing datasets were used, based on interpretation and geological data. This ensured a standardized and objective comparative assessment of the classification algorithms' effectiveness. The test dataset consisted of independent points that were not involved in the training process but were evenly distributed across the study area and among the classes.

The classification results were visualized as maps and presented in the form of accuracy metric tables, error diagrams, and comparative performance indicators.

3. Results and discussion

The classification of hydrothermal alteration zones employing four distinct methods – Spectral Angle Mapper (SAM), Support Vector Machine (SVM), Maximum Likelihood (ML), and Minimum Distance (MD) – resulted in the generation of thematic maps delineating the spatial distribution of argillic, phyllic, propylitic, and potassic alteration zones across the Koldar Massif. The study area encompasses approximately 386 km², with a total perimeter of 78.7 km. All classification procedures were confined to this delineated polygon (Figure 3).

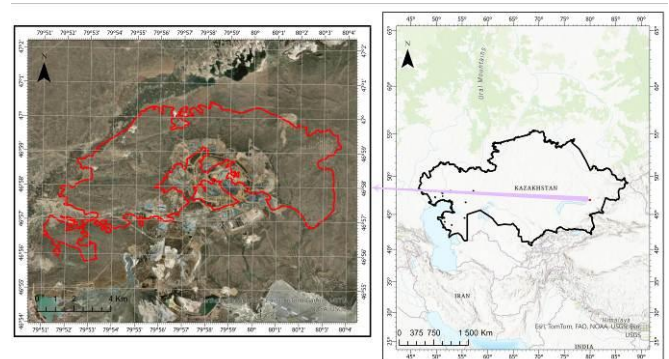


Figure 3. Study area

The accuracy assessment of the classification was performed using confusion matrices and includes the following metrics: Overall Accuracy, Kappa Coefficient, and Producer's Accuracy for each class. For clarity and comparison, the accuracy indicators are presented in Table 1 below.

Method	Overall Accuracy	Kappa	PA (Propylitic)	PA (Potassic)
SAM	51.13%	0.3754	77.27%	26.17%
SVM	84.12%	0.7914	92.00%	82.55%
MD	62.89%	0.4984	87.30%	69.13%
ML	82.80%	0.7759	95.24%	79.87%

Method	PA (Argillic)	PA (Phyllic)
SAM	29.06%	82.80%
SVM	76.07%	86.02%
MD	19.66%	74.19%
ML	80.30%	74.19%

Table 1. Comparison of Classification Accuracy Metrics

The best results were demonstrated by the SVM method (Figure 4), which achieved an overall accuracy of 84.12% and a Kappa coefficient of 0.7914. The highest classification accuracy was observed for propylitic (PA = 92.0%) and phyllic alteration zones (PA = 86.02%), while argillic zones were classified somewhat less accurately (PA = 76.07%). The SVM method provided balanced accuracy across all classes, with relatively low omission and commission errors.

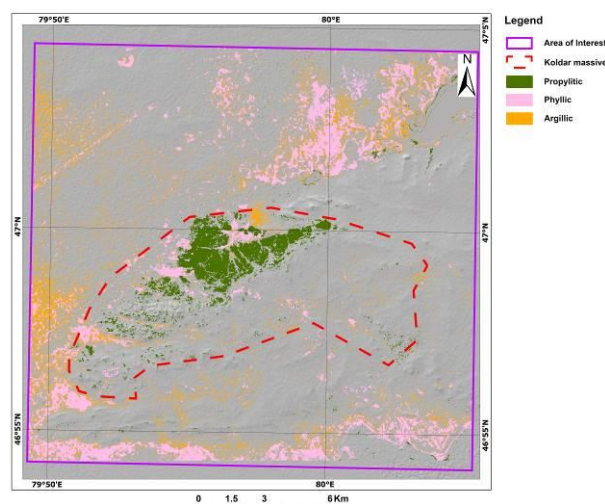


Figure 4. SVM Classification Result

The MD method (Figure 5) also demonstrated high performance (OA = 82.8%, Kappa = 0.7759), achieving the highest classification accuracy for propylitic zones (PA = 95.24%). However, the accuracy for argillic zones was lower (PA = 80.30%) compared to SVM.

The MD method (Figure 5) also demonstrated high performance (OA = 82.8%, Kappa = 0.7759), achieving the highest classification accuracy for propylitic zones (PA = 95.24%). However, the accuracy for argillic zones was lower (PA = 80.30%) compared to SVM.

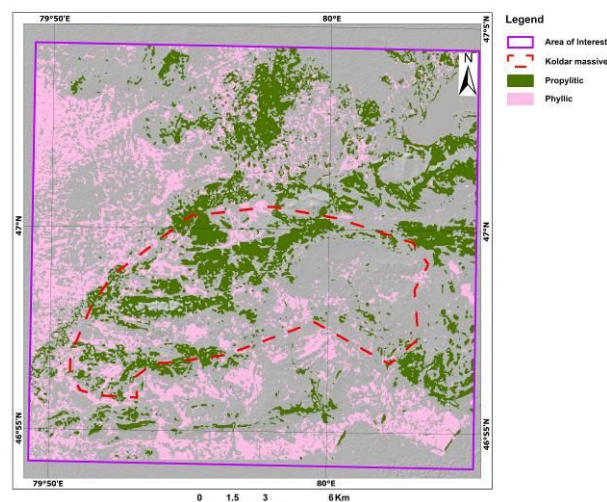


Figure 5. Minimum Distance Classification Result

The Maximum Likelihood algorithm demonstrated (Figure 6) moderate performance (OA = 62.89%, Kappa = 0.4984), with the highest classification accuracy observed for propylitic (PA = 87.3%) and phyllic zones (PA = 74.2%). However, the argillic alteration zone showed low values for both producer's accuracy (19.7%) and user's accuracy (69.7%).

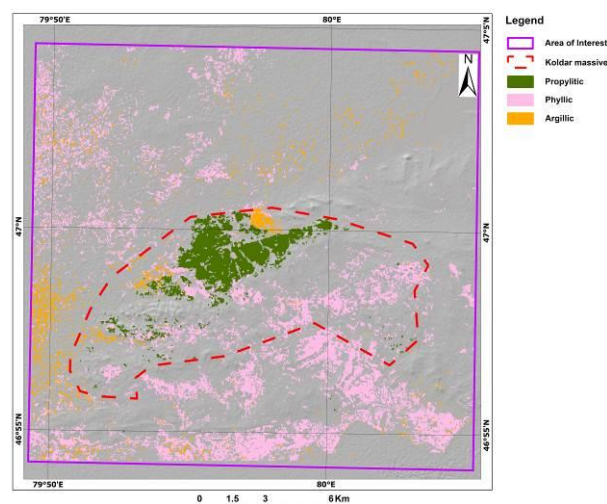


Figure 6. Maximum Likelihood Classification Result

The least accurate results were obtained using the Spectral Angle Mapper (SAM) method (Figure 7) (OA = 51.13%, Kappa = 0.3754), which can be attributed to its limited robustness to noise and the spectral similarity between classes. Nevertheless, the phyllic alteration zone (PA = 82.8%) was recognized as significantly better than the argillic (PA = 29.1%) and potassium alteration zones (PA = 26.2%).

The least accurate results were obtained using the Spectral Angle Mapper (SAM) method (Figure 7) (OA = 51.13%, Kappa = 0.3754), which can be attributed to its limited robustness to noise and the spectral similarity between classes. Nevertheless, the phyllic alteration zone (PA = 82.8%) was recognized as significantly better than the argillic (PA = 29.1%) and potassium alteration zones (PA = 26.2%).

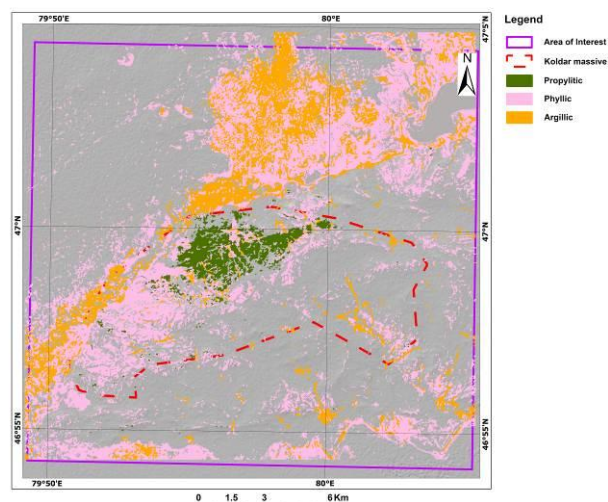


Figure 7. SAM (Spectral Angle Mapper) Classification Result

To assess the reliability of the classification results, a comparison was made with the geological map (Figure 2) and cross-sections of mineralization and hydrothermal alteration, which illustrate a characteristic spatial zonation: from peripheral propylitic zones to argillic and phyllic alterations, with possible potassium alteration zones in the central part.

Propylitic zones (marked in green on the classification maps) show the highest spatial consistency with the outer parts of the intrusive body and the contact zones with surrounding rocks, which aligns well with the geological data (see Figures 9 and 10). These zones were consistently identified across all methods, most clearly using SVM and SAM (Figures 3 and 5), with high classification accuracy (User Accuracy = 83–95%, Producer Accuracy = 87–95%).

Phyllic zones (pink color) are most distinctly delineated by the ML and SVM methods. Their spatial distribution is concentrated along fault zones and in the central parts of the massif, which is supported by geological cross-sections indicating that phyllic alterations are closely associated with ore bodies and zones of intense hydrothermal activity. In contrast, the MD method shows significant fragmentation and a high rate of false positives for this class (Commission Error > 48%).

Argillic alterations (orange color) are reliably detected only by the SAM and ML methods. These zones are mainly located on the periphery of phyllic alterations and partially overlap with areas associated with clay minerals according to geological sections. The SVM method tends to overestimate the extent of this class, as indicated by the inflated User Accuracy (>98%) and low Producer Accuracy (19.66%) observed in the MD method.

Potassic zones are poorly distinguishable across all methods, likely due to the limited spectral signature in the ASTER data and their minimal representation in the training polygons. Nevertheless, some weak localization of these zones in the central part of the massif – near the presumed apical part of the intrusion – was achieved using the SVM and ML classifications. A comparative analysis of the spatial distribution of hydrothermal zones produced by different classification algorithms allows for several general conclusions regarding each method's sensitivity to specific alteration types. The SVM algorithm demonstrated the best correspondence with the known geological zonation, consistent identification of all major alteration types, and the fewest artifacts. This method was particularly effective in delineating phyllic and propylitic alteration zones, as confirmed by geological maps and lithochemical data.

Probabilistic and statistical methods (ML and MD) demonstrated moderate effectiveness: although they identified the main alteration zones, their results were accompanied either by excessive fragmentation (MD) or inflated accuracy metrics due to underrepresentation of certain classes (ML). The SAM method, despite its traditional applicability to hyperspectral data, proved to be the least robust when applied to ASTER multispectral imagery, likely due to the limited spectral information and fewer available bands.

Thus, machine learning – particularly using the Support Vector Machine (SVM) approach – shows strong potential for mapping hydrothermal alteration zones based on multispectral data. To further improve classification accuracy, the use of ensemble methods, integration of additional geological indicators, and inclusion of textural and topographic parameters is recommended.

In summary, the SVM and ML methods demonstrated the highest efficiency in detecting complex hydrothermal alteration zones, owing to their ability to capture both linear and non-linear relationships in the spectral features of ASTER data. The obtained results show a high degree of consistency with geological maps and lithochemical data, confirming the applicability of these methods for preliminary geological mapping. These findings can serve as a foundation for planning follow-up fieldwork, geochemical surveys, and refining the boundaries of potentially prospective ore-bearing zones.

4. Conclusion

In this study, a methodology for interpreting hydrothermal alteration zones associated with porphyry copper systems was tested using ASTER remote sensing data and machine learning algorithms. The Koldar Intrusive Complex in Central Kazakhstan was selected as the study area due to its geological prospectivity, pronounced hydrothermal activity, and mineralization potential. The application of classification methods including SVM, SAM, ML, and MD made it possible to produce thematic maps showing the spatial distribution of the main types of metasomatic alterations: argillic, phyllic, propylitic, and potassic. Comparative analysis of the results showed that the Support Vector Machine (SVM) algorithm achieved the highest accuracy, with an overall accuracy of 84.12% and a kappa coefficient of 0.79. The MD and ML methods also demonstrated satisfactory results, whereas SAM proved less accurate when a limited training dataset was used.

The resulting alteration maps correlate well with the tectonic and lithological structure of the region and allow for the identification of areas most promising for further geological exploration. The use of NDVI and vegetation masking improved classification accuracy by excluding vegetation-covered areas.

The proposed methodology is recommended for use in the early stages of mineral exploration in arid and semi-arid environments, where surface exposures and lighting conditions are favorable for remote sensing applications. Moreover, it can be adapted for analyzing other types of ore systems and underexplored areas with limited geological mapping.

In future work, we plan to integrate hyperspectral data (e.g., PRISMA) and expand the training dataset using geochemical data to improve the predictive value and reliability of the results.

Acknowledgements

This research was funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. BR21882179).

References

- Bahrami, H., Esmaceli, P., Homayouni, S., Pour, A. B., Chokmani, K., & Bahrroudi, A. (2024). Machine learning-based lithological mapping from ASTER remote-sensing imagery. *Minerals*, 14(2), 202. <https://doi.org/10.3390/min14020202>
- Barsi, Á., Kugler, Z., László, I., Szabó, G., & Abdulmutalib, H. M. (2018). Accuracy dimensions in remote sensing. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 42, 61–67. <https://doi.org/10.5194/isprs-archives-XLII-3-61-2018>
- Bedini, E. (2009). Mapping lithology of the Sarfartoq carbonatite complex, southern West Greenland, using HyMap imaging spectrometer data. *Remote Sens. Environ.*, 113(6), 1208–1219. <https://doi.org/10.1016/j.rse.2009.02.007>
- Belgiu, M., & Drăguț, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.*, 114, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>
- Fitzgerald, R. W., & Lees, B. G. (1994). Assessing the classification accuracy of multisource remote sensing data. *Remote Sens. Environ.*, 47(3), 362–368. [https://doi.org/10.1016/0034-4257\(94\)90103-1](https://doi.org/10.1016/0034-4257(94)90103-1)
- Fu, Y., Cheng, Q., Jing, L., Ye, B., & Fu, H. (2023). Mineral prospectivity mapping of porphyry copper deposits based on remote sensing imagery and geochemical data in the Duolong ore district, Tibet. *Remote Sens.*, 15(2), 439. <https://doi.org/10.3390/rs15020439>
- Hu, B., Xu, Y., Wan, B., Wu, X., & Yi, G. (2018). Hydrothermally altered mineral mapping using synthetic application of Sentinel-2A MSI, ASTER and Hyperion data in the Duolong area, Tibetan Plateau, China. *Ore Geol. Rev.*, 101, 384–397. <https://doi.org/10.1016/j.oregeorev.2018.07.017>
- Huang, S., Tang, L., Hupy, J. P., Wang, Y., & Shao, G. (2021). A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. *J. For. Res.*, 32(1), 1–6. <https://doi.org/10.1007/s11676-020-01155-1>
- Kerr, G. H., Fischer, C., & Reulke, R. (2015, July). Reliability assessment for remote sensing data: Beyond Cohen's kappa. In 2015 *IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)* (pp. 4995–4998). IEEE. <https://doi.org/10.1109/IGARSS.2015.7326954>
- Khaleghi, M., Ranjbar, H., Abedini, A., & Calagari, A. A. (2020). Synergetic use of the Sentinel-2, ASTER, and Landsat-8 data for hydrothermal alteration and iron oxide minerals mapping in a mine scale. *Acta Geodyn. Geomater.*, 17(3). <https://doi.org/10.13168/AGG.2020.0023>
- Kruse, F. A. (2012). Mapping surface mineralogy using imaging spectrometry. *Geomorphology*, 137(1), 41–56. <https://doi.org/10.1016/j.geomorph.2010.09.032>
- Kruse, F. A., Lefkoff, A. B., Boardman, J. W., Heidebrecht, K. B., Shapiro, A. T., Barloon, P. J., & Goetz, A. F. (1993). The spectral image processing system (SIPS) – interactive visualization and analysis of imaging spectrometer data. *Remote Sens. Environ.*, 44(2–3), 145–163. [https://doi.org/10.1016/0034-4257\(93\)90013-N](https://doi.org/10.1016/0034-4257(93)90013-N)
- Liu, X., & Yang, C. (2013, December). A kernel spectral angle mapper algorithm for remote sensing image classification. In 2013 6th Int. Congr. Image Signal Process. (CISP) (Vol. 2, pp. 814–818). IEEE. <https://doi.org/10.1109/CISP.2013.6745277>
- Lowell, J. D., & Guilbert, J. M. (1970). Lateral and vertical alteration-mineralization zoning in porphyry ore deposits. *Econ. Geol.*, 65(4), 373–408. <https://doi.org/10.2113/gsecongeo.65.4.373>
- Mars, J. C., & Rowan, L. C. (2006). Regional mapping of phyllic- and argillic-altered rocks in the Zagros magmatic arc, Iran, using ASTER data and logical operator algorithms. *Geosphere*, 2(3), 161–186. <https://doi.org/10.1130/GES00044.1>
- Orynassarova, E., Ahmadi, H., Adebijet, B., Bekbotayeva, A., Abdullayeva, T., Pour, A. B., ... & Bermukhanova, A. (2025). Mapping alteration minerals associated with Aktogay porphyry copper mineralization in Eastern Kazakhstan using Landsat-8 and ASTER satellite sensors. *Minerals*, 15(3), 277. <https://doi.org/10.3390/min15030277>
- Pal, M., & Mather, P. M. (2005). Support vector machines for classification in remote sensing. *Int. J. Remote Sens.*, 26(5), 1007–1011. <https://doi.org/10.1080/01431160512331314083>
- Rejith, R. G., Sundararajan, M., Gnanappazham, L., Kaliraj, S., & Chandrasekar, N. (2022). Exploring beach placer minerals in the east coast of Tamil Nadu, India, using EO-1 Hyperion data. *J. Appl. Remote Sens.*, 16(1), 012017. <https://doi.org/10.1117/1.JRS.16.012017>
- Richards, J. A. (1999). Remote sensing digital image analysis. Berlin, Alemania.
- Rockwell, B. W., & Hofstra, A. H. (2008). Identification of quartz and carbonate minerals across northern Nevada using ASTER thermal infrared emissivity data – Implications for geologic mapping and mineral resource investigations in well-studied and frontier areas. *Geosphere*, 4(1), 218–246. <https://doi.org/10.1130/GES00126.1>
- Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J. P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote Sens.*, 67, 93–104. <https://doi.org/10.1016/j.isprsjprs.2011.11.002>
- Rouskov, K., Popov, K., Stoykov, S., & Yamaguchi, Y. (2005, June). Some applications of the remote sensing in geology by using of ASTER images. In Sci. Conf. SPACE, ECOLOGY, SAFETY with Int. Participation (pp. 167–173).
- Safari, M., Maghsoudi, A., & Pour, A. B. (2018). Application of Landsat-8 and ASTER satellite remote sensing data for porphyry copper exploration: A case study from Shahr-e-Babak, Kerman, south of Iran. *Geocarto Int.*, 33(11), 1186–1201. <https://doi.org/10.1080/10106049.2017.1334834>
- Shahriari, H., Honarmand, M., & Ranjbar, H. (2015). Comparison of multi-temporal ASTER images for hydrothermal alteration mapping using a fractal-aided SAM method. *Int. J. Remote Sens.*, 36(5), 1271–1289. <https://doi.org/10.1080/01431161.2015.1011352>
- Sillitoe, R. H. (2010). Porphyry copper systems. *Econ. Geol.*, 105(1), 3–41. <https://doi.org/10.2113/gsecongeo.105.1.3>

Sisodia, P. S., Tiwari, V., & Kumar, A. (2014, May). Analysis of supervised maximum likelihood classification for remote sensing image. In *Int. Conf. Recent Adv. Innov. Eng. (ICRAIE-2014)* (pp. 1–4). IEEE. <https://doi.org/10.1109/ICRAIE.2014.6909319>

Stehman, S. V. (2013). Estimating area from an accuracy assessment error matrix. *Remote Sens. Environ.*, 132, 202–211. <https://doi.org/10.1016/j.rse.2013.01.016>

Tompolidi, A. M., Sykioti, O., Koutroumbas, K., & Parcharidis, I. (2020). Spectral unmixing for mapping a hydrothermal field in a volcanic environment applied on ASTER, Landsat-8/OLI, and Sentinel-2 MSI data: The Nisyros (Greece) case study. *Remote Sens.*, 12(24), 4180. <https://doi.org/10.3390/rs12244180>

Van der Meer, F. D., Van der Werff, H. M., Van Ruitenbeek, F. J., Hecker, C. A., Bakker, W. H., Noomen, M. F., ... & Woldai, T. (2012). Multi- and hyperspectral geologic remote sensing: A review. *Int. J. Appl. Earth Obs. Geoinf.*, 14(1), 112–128. <https://doi.org/10.1016/j.jag.2011.08.002>

Wacker, A. G., & Landgrebe, D. A. (1972). Minimum distance classification in remote sensing. *LARS Tech. Rep.*, 25.

Wang, H., & Liu, X. (2013). Mapping mineral resources using ASTER images: Application to the Xizang region in Tibet, China. *Int. J. Remote Sens.*, 34(14), 4985–5000. <https://doi.org/10.1080/01431161.2013.786153>

Zhang, Q., & Li, Z. (2010). Remote sensing technology for mineral exploration. *Ore Geol. Rev.*, 38(2), 69–79. <https://doi.org/10.1016/j.oregeorev.2009.11.001>

Zhang, Y. (2012, September). Support vector machine classification algorithm and its application. In *Int. Conf. Inf. Comput. Appl.* (pp. 179–186). Springer. https://doi.org/10.1007/978-3-642-34041-3_27

Zvezdov, L. P., Erokhin, Yu. V., & Ivanov, V. V. (1993). Metasomatic zoning of porphyry copper deposits and its significance in prospecting. *Geol. Ore Deposits*, 35(6), 441–453.