

# AUTOMATED STRUCTURAL ANALYSIS AND DIRECTIONAL VALIDATION USING PCA-TRANSFORMED SENTINEL-2 IMAGERY: A CASE STUDY FROM THE TSAGAAN-UUL REGION, SOUTHERN MONGOLIA

Gáspár Albert<sup>1</sup>, Munkhsuren Badrakh<sup>2</sup>

<sup>1</sup> ELTE Eötvös Loránd University, Institute of Cartography and Geoinformatics, 1117 Budapest, Pázmány Péter prom. 1/A, Hungary; albert@ludens.elte.hu

<sup>2</sup> ELTE Eötvös Loránd University, Doctoral School of Earth Sciences, Faculty of Science, 1117 Budapest, Pázmány Péter prom. 1/A, Hungary; muujuub@student.elte.hu

## Abstract

*Geological mapping in remote, semi-arid regions often relies on medium-scale legacy maps that lack the resolution to resolve complex tectonic histories. This study introduces an automated, high-resolution workflow for lineament mapping and validation in the Tsagaan-Uul Block, Southern Mongolia. Using multispectral Sentinel-2 imagery, we applied Principal Component Analysis (PCA) to enhance structural and lithological features. To ensure objectivity, automated edge detection was implemented using directional Sobel filters aligned with regional tectonic trends (NW–SE, E–W, and NE–SW), followed by automated vectorization. A key contribution is the development of a spatial validation method that compares satellite-derived lineaments against digitized legacy maps within a regular 1 km grid. This grid-based statistical approach enables the identification of overprinted deformation phases and distinguishes tectonic fractures from lithological contacts. The proposed workflow offers a reproducible, objective alternative for structural analysis in complex terrains where high-quality field data is scarce, providing new insights into the region's multi-phase structural evolution.*

**Keywords:** Lineament mapping, Sentinel-2, Principal Component Analysis (PCA), Automated edge detection, Spatial validation, Tsagaan-Uul Block, Structural geology.

## INTRODUCTION

The term 'lineament' was first introduced to geology by W. H. Hobbs in 1904 to describe prominent structural features that are observable on the Earth's surface (Hobbs, 1904). These features include ridge crests, drainage lines, and boundaries between different rock types. Later, the definition was refined, describing lineaments as linear or slightly curved features distinguishable from surrounding patterns that reflect subsurface geological structures (O'Leary et al., 1976). Geological lineaments are particularly important as they typically correspond to brittle deformation features, such as faults and fractures, which form under specific stress conditions in the Earth's crust (Fossen, 2016; Ahmadi & Pekkan, 2021). Faults, which are large shear fractures with visible displacement, and fractures formed by stress, are critical indicators of a region's tectonic history.

Starting with the launch of Landsat-1 in 1972, advances in remote sensing have revolutionised structural geology by enabling the large-scale analysis of the Earth's surface features (Gabrielsen & Olesen, 2024). Early studies used multispectral imagery to map tectonic features in unprecedented detail. More recently, the combination of digital elevation models (DEMs) and multispectral data, such as Landsat and Sentinel-2, has become essential for capturing terrain variations and lithological boundaries (Al-Obeidat et al., 2016; Soliman & Han, 2019). Although manual interpretation of these datasets can produce detailed results, it is often subject to inconsistency and bias and is highly time-consuming. Consequently, automated approaches using algorithms such as edge detection and the Hough transform offer a faster, more reproducible alternative that significantly reduces human bias (Farahbakhsh et al., 2020; Mohammadpour et al., 2020).

Although automated methods are efficient, they can misclassify non-geological features, necessitating robust validation procedures (Ahmadi & Pekkan, 2021). This is particularly important in remote, semi-arid regions, where geological mapping often relies on legacy medium-scale maps (e.g., 1:50,000). Produced decades ago through aerial photography and fieldwork, these older maps often lack the resolution required to unravel complex tectonic histories and are limited by generalisation (e.g. Badrakh et al., 2025). Furthermore, subsequent deformations can reorient older structural patterns, making it difficult to identify distinct deformation phases.

The Tsagaan-uul area in southern Mongolia is a region of significant geological complexity, comprising diverse sedimentary, metamorphic, and igneous rock formations dating from the Neoproterozoic to the Paleozoic era (Badarch et al., 2002; Badamtulga et al., 2005). The region is characterised by significant topographic variation and major structural features, such as the northeast-trending Tsagaan-Uul Fault (Badamtulga et al., 2005). Despite the unique geological setting, high-resolution automated lineament mapping and systematic validation against legacy data have been limited in Mongolia. This study addresses this gap by introducing a high-resolution, automated workflow that utilises Sentinel-2 imagery and principal component analysis (PCA). PCA is used to enhance the contrast of lithological boundaries and structural features, making them more noticeable and distinguishable than in the original spectral bands (Drury, 1987). Automated edge detection using directional Sobel filters is implemented to ensure objectivity, followed by vectorisation.

The primary objective is to present a new spatial validation method that compares satellite-derived lineaments with digitised legacy maps within a regular 1 km grid. This approach enables the objective identification of overprinted deformation phases and offers a reproducible alternative for structural analysis in areas where high-quality field data is scarce.

## GEOLOGICAL SETTING

The Tsagaan-uul study area is located in Umnugobi Province, southern Mongolia. Geologically, it lies within the South Gobi Zone and is assigned to the Tsagaan Uul terrane of the southern Mongolian accretionary collage (Badarch et al., 2002). This terrane has been interpreted as a Precambrian basement domain and is commonly correlated with the Khatanbulag Ancient Massif (Fig. 1/b). The region has undergone a remarkably complex geological evolution from the Neoproterozoic to the Quaternary, characterised by cycles of supercontinent assembly and rifting, passive margin development, and subsequent accretionary and collisional orogenesis.

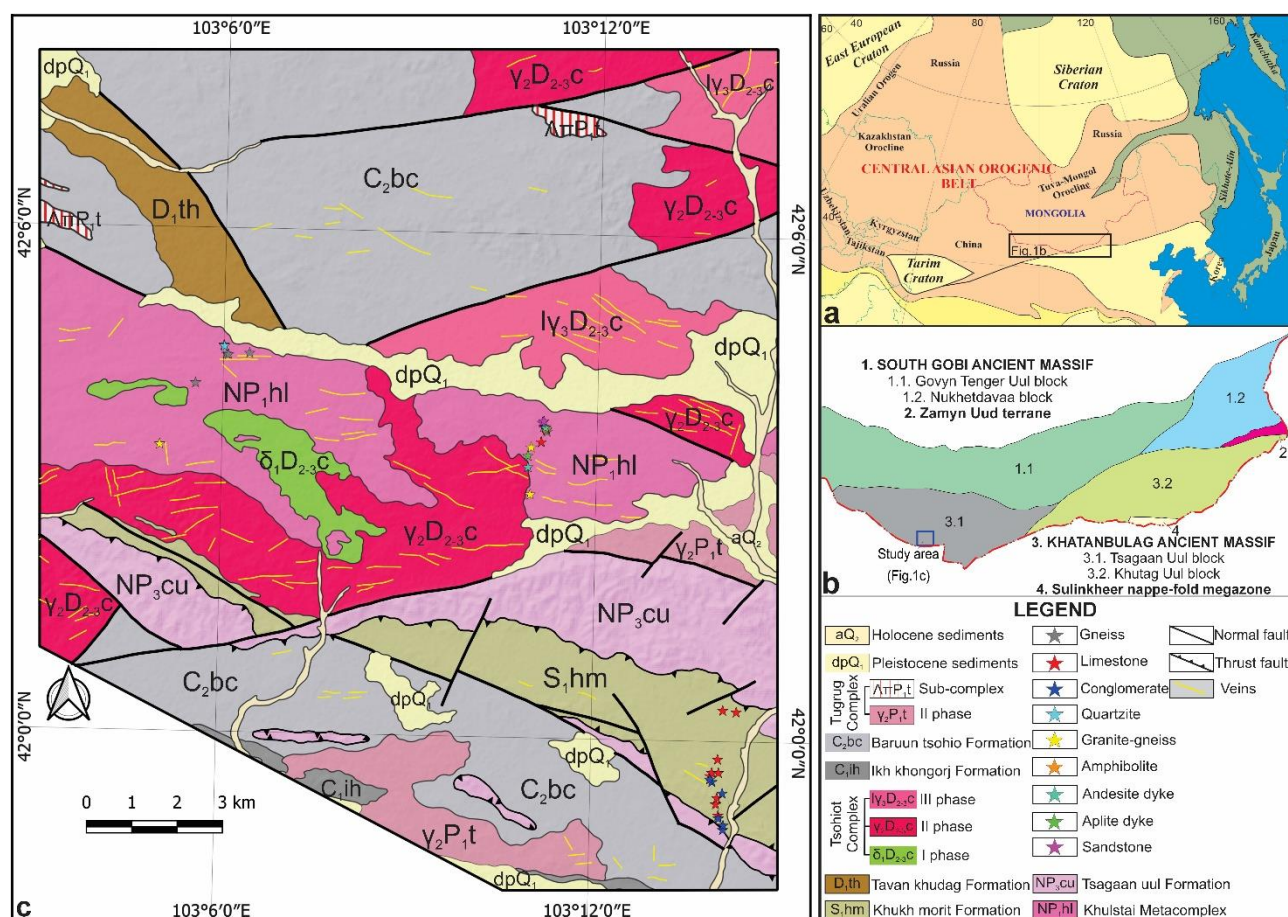


Figure 1. Geological settings of the study area. (a) A schematic map of the Central Asian Orogenic Belt (Xiao et al., 2017); (b) Tectonic subdivision of southern Mongolia, modified from Tomurtogoo (2017), showing the South Gobi Ancient Massif (1.1, 1.2), the Zamyn Uud terrane (2), and the Khatanbulag Ancient Massif (3.1 and 3.2), which includes the Tsagaan-Uul block (3.1) where the study area is located, along with the Sulinkheer nappe-fold megazone (4); (c) Geological map of the Tsagaan-uul area and sample locations (modified after Badamtulga et al., 2005; and Badrakh et al., 2025)

The study area covers an area of approximately 274 square kilometres, bounded by latitudes of 41° 58' N to 42° 08' N and longitudes of 103° 02' E to 103° 15' E (Fig. 1/c). The northern and central regions are characterised by rugged, mountainous terrain associated with the Tsagaan uul Formation (NP<sub>3cu</sub>), with elevations reaching 1,433 metres. In contrast, the southern region consists of lower-lying areas covered by Quaternary alluvial and proluvial deposits.

### **Neoproterozoic Basement and Rodinia Cycle (D1–D2)**

The oldest structural level in the region is represented by the Early to Middle Neoproterozoic crystalline basement. The basement's central feature is the Khulstai Metacomplex (NP<sub>1hl</sub>), comprising ortho- and paragneisses, schists, amphibolites, quartzites, and marbles. This complex is interpreted as a remnant of a microcontinental fragment involved in the formation of the Rodinia supercontinent during the Grenvillian period, ~1,000–900 million years ago (Badarch et al., 2002; Kovach et al., 2021). This initial deformation phase (D1) resulted in high-grade metamorphism and complex folding with northeast-oriented axes. Geochronological studies of detrital zircons from paragneisses in this unit have yielded ages between 886 and 970 Ma, with a weighted mean age of  $925 \pm 12$  Ma constraining the maximum depositional age (Narantsetseg et al., 2022).

Following the consolidation of the basement, the Late Neoproterozoic (~800–700 Ma) was characterised by lithospheric extension and rifting (D2), which was associated with the breakup of Rodinia (Levashova et al., 2010). This phase resulted in the development of an epicontinental basin along the southern margin of the basement. The Tsagaan uul Formation (NP<sub>3cu</sub>) was deposited within this subsiding basin. This 800–1000 m thick succession of stromatolite-bearing limestones, marbles, and quartzites represents the sedimentary cover of the older crystalline basement (Badamtulga et al., 2005). The initial rifting was of an extensional nature; however, the formation is presently influenced by intricate folding with NE-trending axes, which is indicative of subsequent compressional overprinting.

### **Early Paleozoic Evolution: Uplift and Passive Margin (D3–D4)**

The transition into the Paleozoic era was marked by a significant tectonic shift, which is a well-documented occurrence in the geological record. A period of regional uplift and inversion occurred during the Cambrian (~540–500 Ma), designated as phase D3. This event resulted in a significant stratigraphic hiatus, as Cambrian sedimentary rocks are entirely absent from the study area (Badamtulga et al., 2005; Wang et al., 2009).

Subsequently, renewed subsidence began in the Ordovician-Silurian (~485–420 Ma), as the region evolved into a stable passive continental margin setting (D4). This phase was driven by post-rift thermal subsidence and potential distal extension as the Paleo-Asian Ocean widened (Badarch et al., 2002). In this tectonically stable environment, the Khukh Morit Formation (S<sub>1hm</sub>) accumulated on a shallow-marine continental shelf. This formation, which has a thickness of up to 2000 m, consists of limestones, marbles, and quartzites that are rich in fossils. The presence of diverse marine fossils, including corals (*Favosites*, *Heliolites*) and brachiopods, indicates a warm, stable carbonate platform environment (Narantsetseg et al., 2022). After this, the rocks were subject to NW-SE trending synclinal folding during events of Devonian-Carboniferous accretion (Badamtulga et al., 2005).

### **Devonian Active Margin and Accretion (D5–D6)**

By the Early Devonian period (~410–390 Ma), the tectonic regime shifted to an active margin. Phase D5 is distinguished by the occurrence of back-arc rifting, accompanied by the formation of graben-like or pull-apart basins that exhibit a northwest trend (Windley et al., 2007). The Tavan Khudag Formation (D<sub>1th</sub>), composed of pelitic schists, siltstones, and sandstones with basaltic and andesitic volcanic components, records this renewed lithospheric stretching and bimodal volcanism. This extension indicates a rotation in the regional stress field in comparison to the Neoproterozoic trends.

During the Mid-to-Late Devonian period (~380–360 Ma), a gradual convergence of tectonic plates resulted in the closure of these back-arc basins and the subsequent onset of accretionary collision (D6). This phase was characterised by significant crustal thickening and the emplacement of arc-related magmatic rocks belonging to the Tsohiot Complex ( $\gamma_2D_2-3c$ ) (Badamtulga et al., 2005). The complex displays a typical calc-alkaline evolution from early diorites to later biotite-hornblende granites and leucogranites.

## Late Paleozoic Transpression and Consolidation (D7)

The Late Paleozoic era (~320–270 Ma) was characterised by the final closure of the Paleo-Asian Ocean and the formation of the Central Asian Orogenic Belt. Phase D7 is characterised by regional sinistral transpression. During this period, the Ikh Khongorj Formation (C<sub>1</sub>ih) was deposited in remnant marine conditions, while thick terrigenous-volcanogenic successions formed in the Baruun Tsokhio Formation (C<sub>2</sub>bc). This stage culminated in the emplacement of the Tugrug Complex granites, likely along major strike-slip fault zones. Sinistral strike-slip movement and large-scale block rotations occurred as part of the final continental consolidation (Badamtulga et al., 2005; Wu et al., 2014).

## Mesozoic and Cenozoic Reactivation (D8–D12)

After the consolidation in the Paleozoic, the region shifted to an intracontinental regime. During the Mesozoic, the study area experienced multiple reactivation pulses. Late Triassic to Jurassic contraction (D8), which was associated with the closure of the Mongol-Okhotsk Ocean, resulted in the refolding of earlier structures and the formation of east-west trending folds (Guy et al., 2014). This was followed by an extensional phase (D9) from the Late Jurassic to the Early Cretaceous, which created large-scale rift basins and intermontane depressions across southern Mongolia (Johnson et al., 2001). A final compressional pulse in the Mesozoic era, known as the Yanshanian orogeny (D10), caused minor folding and uplift in the Late Cretaceous period (Lin et al., 2013).

Since the Late Cenozoic, the collision between India and Asia has transmitted far-field stresses into southern Mongolia (D11). This has reactivated inherited Paleozoic basement structures (i.e., those of D7) via NE-directed compression and transpressional strike-slip faulting (Cunningham, 2010). Many sinistral faults from the Permian were reactivated during this period. These tectonic processes have shaped the formation of the modern intermontane basins, which are filled with Pliocene, Pleistocene, and Quaternary alluvial, slope, and aeolian sediments.

## Summary of Tectonic Evolution

The following table (Tab. 1) summarizes the key deformation phases and geological processes that have shaped the Tsagaan-Uul area.

Table 1. Tectonic Evolution Summary of the Tsagaan-uul Area (based on Badamtulga et al., 2005; Narantsetseg et al., 2022)

| Phase | Time Period                              | Tectonic Event                                | Geological Process & Result                                                                                               | Key Formations/Complexes                                                                  |
|-------|------------------------------------------|-----------------------------------------------|---------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------|
| D1    | Early-Mid Neoproterozoic (~1,000–900 Ma) | Grenvillian Orogeny                           | Amalgamation of microcontinents into the Rodinia supercontinent; high-grade metamorphism and complex NE-oriented folding. | Khulstai Metacomplex (NP <sub>1</sub> hl): Orthogneiss, paragneiss, amphibolite.          |
| D2    | Late Neoproterozoic (~650–540 Ma)        | Continental Rifting                           | Lithospheric extension related to Rodinia breakup; development of an epicontinental carbonate platform basin.             | Tsagaan uul Formation (NP <sub>3</sub> cu): Stromatolite limestone and marble.            |
| D3    | Cambrian (~540–500 Ma)                   | Regional Uplift / Inversion                   | Post-rift isostatic uplift or basin inversion; significant hiatus (unconformity) with no sedimentary record in the area.  | (Unconformity / Missing Section)                                                          |
| D4    | Ordovician-Silurian (~485–420 Ma)        | Post-rift Thermal Subsidence (Passive Margin) | Thermal subsidence and development of a stable continental shelf; accumulation of thick carbonate-terrigenous sequences.  | Khukh Morit Formation (S <sub>1</sub> hm): Fossil-bearing limestone, marble, quartzite.   |
| D5    | Early Devonian (~410–390 Ma)             | Back-Arc Rifting                              | Formation of NW-trending extensional grabens with bimodal volcanism.                                                      | Tavan Khudag Formation (D <sub>1</sub> th): Basalt, andesite, tuffs, and pelitic schists. |
| D6    | Mid-Late Devonian (~380–360 Ma)          | Accretionary Collision                        | Closure of extensional basins; transition to an active continental margin; crustal thickening and arc-related magmatism.  | Tsohiot Complex (γ <sub>2</sub> D <sub>2-3</sub> c): Diorites and leucogranites.          |

|                  |                                     |                           |                                                                                                                                   |                                                                                                               |
|------------------|-------------------------------------|---------------------------|-----------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------|
| <b>D7</b>        | Carboniferous-Permian (~320–270 Ma) | Sinistral Transpression   | Final closure of the Paleo-Asian Ocean; sinistral strike-slip faulting and syn-tectonic granite emplacement; final arc volcanism. | Baruun Tsohio (C <sub>2bc</sub> ), Ikh Khongorj (C <sub>1ih</sub> ), and Tugrug Complex (γ <sub>2P1t</sub> ). |
| <b>Meso-Ceno</b> | Mesozoic – Present (<250 Ma)        | Continental Consolidation | Shift to a purely continental regime; rifting (Jurassic), followed by far-field Cenozoic reactivation and erosion.                | Pleistocene/Holocene (Q): Alluvial and proluvium sediments.                                                   |

## MATERIALS AND METHODS

The primary dataset for this study consists of multispectral imagery from the Sentinel-2A satellite, which provides a strategic balance between spectral diversity and spatial resolution (10 m for visible and near-infrared bands) (ESA, 2015). To ensure consistency across all spectral channels, the bands were resampled to a uniform 10 m spatial resolution using a bilinear interpolation method. The study area (AOI) was then subset to cover the 274 km<sup>2</sup> extent of the Tsagaan-uul region. For validation purposes, a legacy 1:50,000-scale geological map (Badamtulga et al., 2005) was digitised, with particular focus on previously mapped faults, lithological contacts, and structural lines. Figure 2 illustrates the main steps of the process.

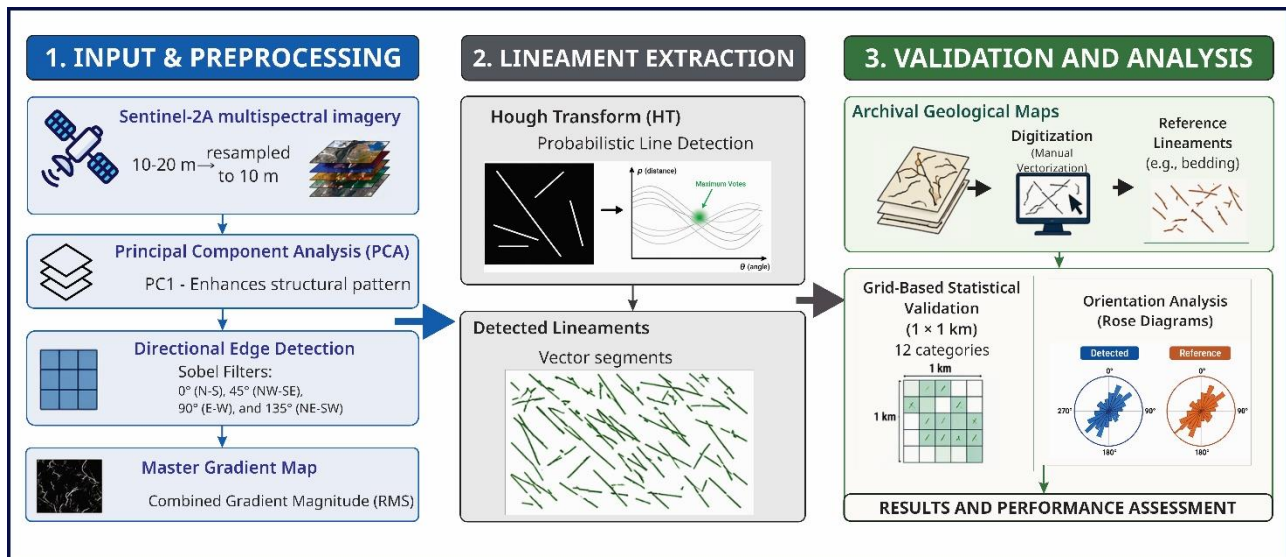


Figure 2. Workflow of the automated lineament extraction and validation process, including PCA-based enhancement, directional Sobel filtering, Hough Transform line detection, and validation using grid-based statistical and orientation analyses

### Image Enhancement: Principal Component Analysis (PCA)

To maximize the contrast of geological structures and lithological boundaries within the Neoproterozoic basement and Paleozoic cover sequences, Principal Component Analysis (PCA) was applied to the multispectral stack. PCA effectively compresses the information from multiple bands into a few uncorrelated components, significantly enhancing features that are often subtle in original spectral bands.

The first principal component (PC1) was selected as the primary input for structural analysis. In semi-arid terrains like Tsagaan-uul, PC1 typically captures over 90% of the total variance, predominantly representing topographic textures and albedo variations, which are critical for identifying morphological expressions of subsurface structures.

## Directional Edge Detection and Raster Fusion

Unlike standard smoothing or non-directional filters, directional Sobel filters were implemented within SAGA GIS to preserve fine structural details while targeting specific tectonic trends. To capture the region's dominant fault orientations, four distinct filter kernels were applied at 0° (N-S), 45° (NW-SE), 90° (E-W), and 135° (NE-SW).

To create a comprehensive "Master Gradient Map," the output rasters from these four directions were normalized to a 0–1 scale and merged using a Root Mean Square (RMS) raster fusion technique via the Grid Calculator. This process ensures that the final gradient image contains the strongest structural signals from all primary tectonic directions while minimizing random noise.

## Automated Lineament Extraction using Hough Transform

The vectorization of the detected edges was performed using the Probabilistic Hough Transform (HT) algorithm (Figure 3). The HT is a voting-based system that identifies lines by transforming image-space pixels into a parameter space ( $\rho$ ,  $\theta$ ), allowing for the detection of linear features even when they are fragmented or partially obscured by erosion (Duda & Hart, 1972).

Based on a sensitivity analysis to balance detection completeness with structural realism, the HT parameters were optimized for the 10 m resolution Sentinel-2 imagery (Table 2). A critical step in the extraction process was the application of a minimum line length filter of 250 m (25 pixels), along with a maximum line gap setting of 5 pixels, to reconnect segmented features that belong to the same geological structure.

Table 2. Optimized Hough Transform parameters for Sentinel-2 PC1 imagery.

| Parameter          | Value     | Impact on Extraction                               |
|--------------------|-----------|----------------------------------------------------|
| Rho ( $\rho$ )     | 1 pixel   | Ensures fine-scale distance precision              |
| Theta ( $\theta$ ) | 1°        | Balanced angular resolution for directional trends |
| Threshold          | 20 votes  | Minimizes spurious detections from noise           |
| minLineLength      | 25 pixels | Filters out segments shorter than 250 m            |
| maxLineGap         | 5 pixels  | Links fragmented edges into continuous lineaments  |

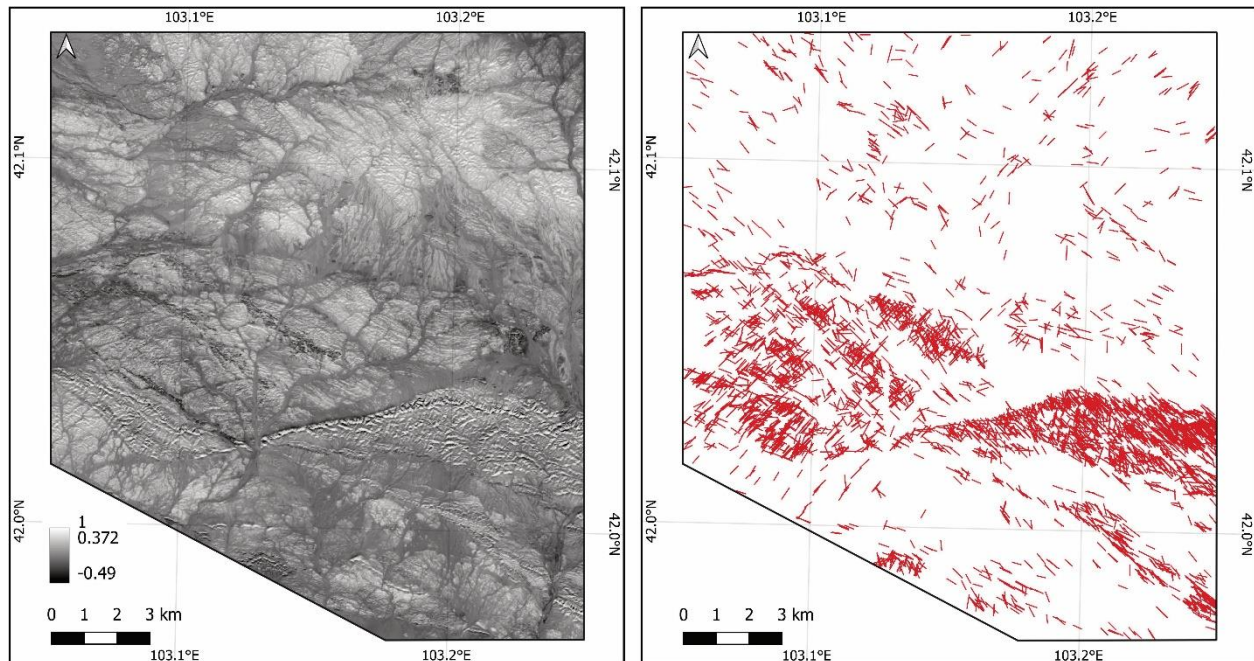


Figure 3. Lineament extraction results from the first principal component (PC1) of Sentinel-2A data. (a) PC1 grayscale image used as input for lineament detection. (b) Extracted lineaments obtained using the Hough Transform algorithm with the optimized parameter set.

## Spatial Validation Methodology: The Grid-based Approach

A critical component of this workflow is the spatial validation of automated lineaments against archived geological data. The key innovation of this study is a grid-based statistical validation method designed to distinguish between tectonic structures, such as faults and fractures, and lithological features, such as stratification and facies boundaries. To this end, the study area has been divided into a regular grid of 1 km squares (298 pcs), matching the data density of the legacy 1:50,000 scale geological mapping.

To establish a baseline for non-tectonic linear features, stratification data from archive maps, including dip-direction symbols and mapped lithological contacts, were digitised. To ensure statistical comparability with the high-density automated lineaments, the digitised features were 'exploded' into segments of approximately 200 m. This process represents the prevailing orientation of the rock sequences as a quasi-continuous vector field across the grid.

We assigned each vector (including both the automated HT lineaments and the exploded stratification lines) to an orientation category (1–12) using a custom QGIS expression with a 15° step interval. We then performed a cell-by-cell correlation analysis using a 2 x 12 frequency matrix. The Pearson correlation coefficient ( $r$ ) was calculated for each category to determine the degree of overlap between the automated results and the known lithology. If the correlation coefficient for a specific orientation category exceeded the threshold of  $r > 0.2$ , the lineaments in that category were primarily attributed to lithological stratification rather than tectonic deformation. This filtered out the 'noise' caused by erosion-resistant bedding planes. Table 3 summarises the relationship between the identified deformation phases and their corresponding lineament orientation categories.

Table 3. Mapping of deformation phases to directional categories based on HT lineament distribution.

| Phase   | Activation     | Stress ( $\sigma_1$ ) | Lineament Trend        | Azimuth Range | Category  | Appearance                                                                    |
|---------|----------------|-----------------------|------------------------|---------------|-----------|-------------------------------------------------------------------------------|
| D1 / D3 | Proterozoic    | NW-SE (Comp.)         | NE-SW (Folding)        | 30–60°        | 3, 4      | Dominant in the Khulstai Metacomplex and Tsagaan uul Fm. But only in the west |
| D5      | Devonian Rift  | NE-SW (Ext.)          | NW-SE (Grabens)        | 135–150°      | 10        | Dominant in the Khulstai Metacomplex and Tsagaan uul Fm. But only in the east |
| D6      | Devonian       | NE-SW (Comp.)         | NW-SE Folding/SS       | 120–180°      | 9, 10, 11 | It is found all over the area                                                 |
| D7      | Late Paleozoic | NNE-SSW (Comp.)       | WSW-ENE (Sinistral SS) | 45–75°        | 4, 5      | Occurs only in the western part                                               |
| D8      | Mesozoic       | N-S (Comp.)           | E-W structures         | 75–115°       | 6, 7      | All over the area                                                             |
| D11     | Cenozoic       | NE-SW (Comp.)         | Reactivated D7 trends  | 45–75°        | 4, 5      | Occurs only in the western part                                               |
| N/A     | Undefined      | N/A                   | New Features*          | Various       | 1, 2, 12  | All over the area affecting mainly the Tsagaan uul Fm.                        |

\*Explained in the results section

Analysis of the correlation matrix (Table 4) revealed that Categories 7 (90–115°) and 8 (115–120°) show the strongest correlation with mapped stratification. Consequently, lineaments in these directions were treated with caution during tectonic interpretation, as they likely reflect bedding rather than brittle failure.

Table 4. Correlation coefficients ( $r$ ) between stratification and automated HT lineaments per category.

| C1    | C2    | C3    | C4    | C5   | C6   | C7  | C8   | C9   | C10   | C11  | C12   |
|-------|-------|-------|-------|------|------|-----|------|------|-------|------|-------|
| -0.03 | -0.03 | -0.04 | -0.03 | 0.13 | 0.13 | 0.2 | 0.44 | 0.01 | -0.08 | -0.1 | -0.07 |

## RESULTS AND DISCUSSION

A quantitative comparison between the automated HT-derived lineaments and the legacy map features was conducted across the 12 directional categories (Figure 4). The automated workflow detected a total of 3,517 unique lineament segments, providing a significantly higher structural density compared to the 1,729 unique features digitised from the 1:50,000 legacy maps. However, for the grid-based directional analysis, the cumulative count across all cells reached 5,241 observations for the HT-derived lineaments and 2,183 observations for the map-based lines. This discrepancy is a result of the spatial partitioning methodology: lineaments spanning multiple grid cells were recorded in each intersecting 1 km cell to ensure that regional structural trends are accurately represented in the local directional frequencies.

The frequency distribution reveals three major structural peaks. The automated HT extraction identified its primary peak in Category 9 (120°–135°), with 1,247 observations, and two smaller peaks occurred in Category 3 (30°–45°) and Category 7 (90°–115°) for the HT dataset (612 and 673 observations, respectively). The legacy map's stratification data is heavily dominated by Category 7 (90°–115°), with 866 observations. In both Category 7 and 8, the two datasets show similar patterns, supporting the statistical observations that the HT-derived lineaments are connected to the stratigraphic features. For the dominant categories of HT-derived lines and lines digitized from maps, the rose diagrams also provide a clear indication, supporting the statistical correlations (Figure 5).

The discrepancy between the peak frequencies of the two datasets (Category 7 for the map and Category 9 for the HT) provides a critical insight into the limitations of automatic lineament detection in complex terrains. The high count in Category 7 on the legacy map reflects the prominent morphological expression of the Paleozoic bedding. If we had not detected this category based on the map layer symbols as a stratigraphic trend, we would have concluded that a very strong structural phase was present based on the HT-derived data. The correlation analysis further confirms this distinction. The high correlation coefficient ( $r=0.44$ ) in Category 8, where both datasets show moderate counts (203 for HT, 112 for Map), suggests that this specific orientation is the most reliable proxy for lithological stratification. The successful identification of these lithological proxies (stratification) demonstrates the robustness of the grid-based validation method. By filtering these "non-tectonic" signals, we can more accurately isolate the brittle deformation patterns.

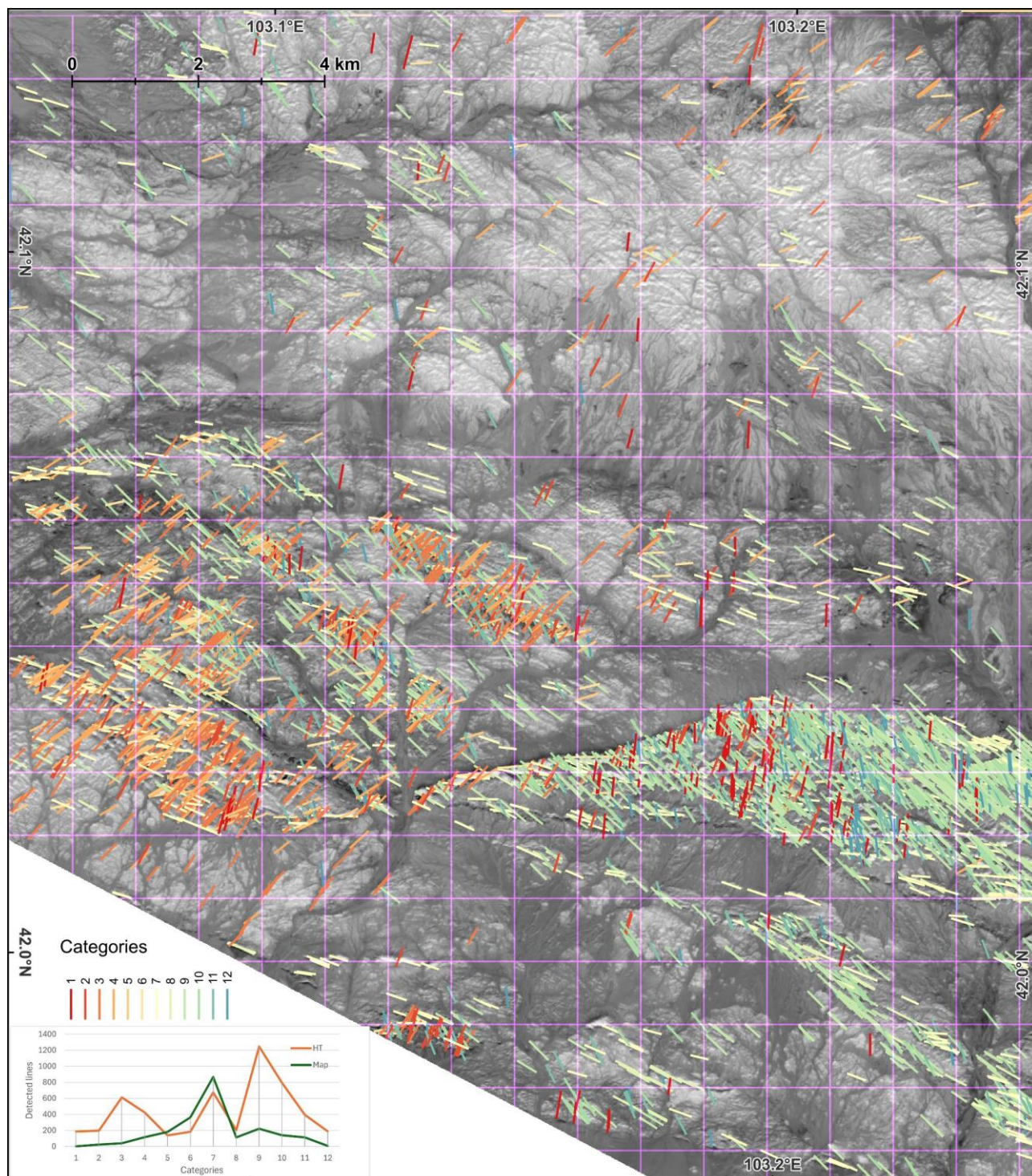


Figure 4. Map of the automatically detected lineaments (HT) showing the 12 direction categories and the diagram comparing the directional frequency between automated HT lineaments and legacy map features. The 1 km grid used for the statistics is shown in the background (purple).

Concerning Category 9 for the HT, we showed that it aligns with the D6 Devonian tectonic phase, making it responsible for the region's main structural framework. There was notable spatial heterogeneity regarding Categories 3 and 4 (30°–60° azimuth), which correspond to the NE–SW structural trend. These lineaments are almost exclusively concentrated in the western part of the Tsagaan-Uul block. Geologically, this distribution indicates that the effects of the earliest deformation phases — the D1 Neoproterozoic consolidation and the D3 Cambrian inversion — are most clearly evident in the western structural basement. In contrast, they may have been significantly altered or buried in the eastern regions by subsequent Paleozoic events.

In contrast, lineaments within Category 10 (with an azimuth of 135–150°) show a strong tendency to occur in the eastern part of the study area. These north-west to south-east trending features are linked to the D5 Devonian rifting and D6 accretionary phases. Their high density in the east suggests that Devonian back-arc basin development and associated crustal extension were most intense in this area, resulting in a pervasive structural fabric that remains the dominant tectonic signal in the eastern subdomains.

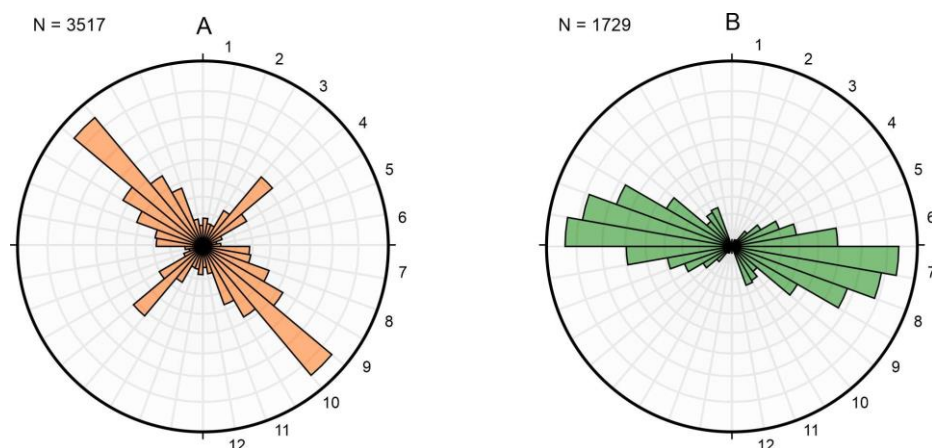


Figure 5. Rose diagrams indicating the major directions and the 12 categories. A: the automatically detected lineaments (HT); B: the stratigraphic linear symbols on the legacy map after segmentation.

The automated extraction identified a distinct set of north-south trending lineaments spanning Categories 1, 2, and 12 (Figure 4: red and blue lines). Notably, this ~N-S orientation is not linked to any of the 11 deformation phases (D1–D11) that have been documented in previous regional studies (e.g., Badamtulga et al., 2005). The presence of these features, which were confirmed by our high-resolution, PCA-based analysis but were absent from legacy tectonic models, may suggest an unrecognised deformation event or a basement structure that has propagated to the surface. This discovery highlights the potential of automated workflows in revealing subtle structural patterns that traditional mapping techniques cannot detect. Further field investigation is needed to determine the nature of these structures.

## CONCLUSIONS

This study presents a high-resolution, automated workflow for lineament mapping in geologically complex semi-arid terrains. The integration of PCA enhancement with directional Sobel filtering and Probabilistic Hough Transform enabled the objective extraction of over 3,500 structural features. The application of a 1 x 1 km spatial validation grid provided a reproducible means to distinguish between tectonic fractures and lithological stratification. Our results demonstrate that while automated lineament extraction is effective, it can underestimate the complexity of the terrain and may conflate structural and stratigraphic features. The proposed method provides an objective way for improving structural geological analysis in remote areas where previous field data is available as legacy maps.

Another key finding of this study is the identification of "New Features" in Categories 1, 2, and 12 that are absent from the literature. These lineaments likely represent subtle fractures or obscured faults that were not resolvable through traditional aerial photo interpretation but are clearly captured in the PCA-transformed Sentinel-2 data.

## REFERENCES

- Ahmadi, H., & Pekkan, E. (2021). Fault-Based Geological Lineaments Extraction Using Remote Sensing and GIS—A Review. *Geosciences*, 11(5), 183. <https://doi.org/10.3390/geosciences11050183>
- Al-Obeidat, F., Feltrin, L., & Marir, F. (2016). Cloud-based Lineament Extraction of Topographic Lineaments from NASA Shuttle Radar Topography Mission Data. *Procedia Computer Science*, 83, 1250–1255. <https://doi.org/10.1016/j.procs.2016.04.260>
- Badamtulga, L., Damdinjav, B., Sumya, T., & Altantogos, T. (2005). *1:50000 scale geological mapping and general prospecting results report carried out in the Tsagaan-Uul unit in Noyon, Bayandalai and Khurmen sub-districts of Umnogobi province* (No. 5626).
- Badarch, G., Dickson Cunningham, W., & Windley, B. F. (2002). A new terrane subdivision for Mongolia: Implications for the Phanerozoic crustal growth of Central Asia. *Journal of Asian Earth Sciences*, 21(1), 87–110. [https://doi.org/10.1016/S1367-9120\(02\)00017-2](https://doi.org/10.1016/S1367-9120(02)00017-2)

- Badrakh, M., Tserendash, N., Choindonjams, E., & Albert, G. (2025). Potential of random forest machine learning algorithm for geological mapping using PALSAR and Sentinel-2A remote sensing data: A case study of Tsagaan-uul area, southern Mongolia. *Journal of Asian Earth Sciences: X*, 14, 100204. <https://doi.org/10.1016/j.jaesx.2025.100204>
- Cunningham, D. (2010). Tectonic setting and structural evolution of the Late Cenozoic Gobi Altai orogen. *Geological Society, London, Special Publications*, 338(1), 361–387. <https://doi.org/10.1144/SP338.17>
- Drury, S. A. (1987). *Image interpretation in geology*. Allen & Unwin.
- Duda, R. O., & Hart, P. E. (1972). Use of the Hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1), 11–15.
- ESA. (2015). *European Space Agency (ESA). Sentinel-2 User Handbook* (p. 64p) [ESA Standard Document]. [https://sentinels.copernicus.eu/documents/247904/685211/Sentinel-2\\_User\\_Handbook](https://sentinels.copernicus.eu/documents/247904/685211/Sentinel-2_User_Handbook)
- Farahbakhsh, E., Chandra, R., Olierook, H. K. H., Scalzo, R., Clark, C., Reddy, S. M., & Müller, R. D. (2020). Computer vision-based framework for extracting tectonic lineaments from optical remote sensing data. *International Journal of Remote Sensing*, 41(5), 1760–1787. <https://doi.org/10.1080/01431161.2019.1674462>
- Fossen, H. (2016). *Structural Geology*. Cambridge University Press. <https://books.google.hu/books?id=86GzCwAAQBAJ>
- Gabrielsen, R. H., & Olesen, O. (2024). The Concept of Lineaments in Geological Structural Analysis; Principles and Methods: A Review Based on Examples from Norway. *Geomatics*, 4(2), 189–212. <https://doi.org/10.3390/geomatics4020011>
- Guy, A., Schulmann, K., Clauer, N., Hasalová, P., Seltmann, R., Armstrong, R., Lexa, O., & Benedicto, A. (2014). Late Paleozoic–Mesozoic tectonic evolution of the Trans-Altai and South Gobi Zones in southern Mongolia based on structural and geochronological data. *Gondwana Research*, 25(1), 309–337. <https://doi.org/10.1016/j.gr.2013.03.014>
- Hobbs, W. H. (1904). Lineaments of the Atlantic Border region. *Geological Society of America Bulletin*, 15(1), 483–506. <https://doi.org/10.1130/GSAB-15-483>
- Johnson, C. L., Webb, L. E., Graham, S. A., Hendrix, M. S., & Badarch, G. (2001). Sedimentary and structural records of late Mesozoic high-strain extension and strain partitioning, East Gobi basin, southern Mongolia. In M. S. Hendrix & G. A. Davis, *Paleozoic and Mesozoic Tectonic Evolution of Central and Eastern Asia: From Continental Assembly to Intracontinental Deformation*. Geological Society of America. <https://doi.org/10.1130/0-8137-1194-0.413>
- Kovach, V., Kozakov, I., Kröner, A., Salnikova, E., Wang, K.-L., Lee, H.-Y., Plotkina, J., Gorokhovskiy, B., Adamskaya, E., Tolmacheva, E., & Shpakovich, L. (2021). Early Neoproterozoic crustal growth and microcontinent formation of the north–central Central Asian Orogenic Belt: New geological, geochronological, and Nd–Hf isotopic data on the Mélange Zone within the Zavkhan terrane, western Mongolia. *Gondwana Research*, 91, 254–276. <https://doi.org/10.1016/j.gr.2020.12.022>
- Levashova, N. M., Kalugin, V. M., Gibsher, A. S., Yff, J., Ryabinin, A. B., Meert, J. G., & Malone, S. J. (2010). The origin of the Baydaric microcontinent, Mongolia: Constraints from paleomagnetism and geochronology. *Tectonophysics*, 485(1–4), 306–320. <https://doi.org/10.1016/j.tecto.2010.01.012>
- Lin, W., Faure, M., Chen, Y., Ji, W., Wang, F., Wu, L., Charles, N., Wang, J., & Wang, Q. (2013). Late Mesozoic compressional to extensional tectonics in the Yiwulüshan massif, NE China and its bearing on the evolution of the Yinshan–Yanshan orogenic belt. *Gondwana Research*, 23(1), 54–77. <https://doi.org/10.1016/j.gr.2012.02.013>
- Mohammadpour, M., Bahroudi, A., & Abedi, M. (2020). Automatic Lineament Extraction Method in Mineral Exploration Using CANNY Algorithm and Hough Transform. *Geotectonics*, 54(3), 366–382. <https://doi.org/10.1134/S0016852120030085>
- Narantsetseg, T., Enkhdalai, B., Orolmaa, D., Munkhjinn, D., Erdenejargal, C., Munkhsuren, B., Delgerzaya, P., Uugantsetseg, B., Tulga, A., & Enkh-Orshikh, O. (2022). *The basement and cover complexes of the Khatanbulag and South Gobi Massifs: Geological development and mineralization*.
- O’Leary, D. W., Friedman, J. D., & Pohn, H. A. (1976). Lineament, linear, lineation: Some proposed new standards for old terms. *Geological Society of America Bulletin*, 87(10), 1463–1469. [https://doi.org/10.1130/0016-7606\(1976\)87<1463:LLSPN>2.0.CO;2](https://doi.org/10.1130/0016-7606(1976)87<1463:LLSPN>2.0.CO;2)
- Soliman, A., & Han, L. (2019). Effects of vertical accuracy of digital elevation model (DEM) data on automatic lineaments extraction from shaded DEM. *Advances in Space Research*, 64(3), 603–622. <https://doi.org/10.1016/j.asr.2019.05.009>
- Tomurtogoo, O. (2017). *Tectonic subdivision of Mongolia: scale 1:4,500,000*
- Wang, C., Sun, Y., Li, N., Zhao, G., & Ma, X. (2009). Tectonic implications of Late Paleozoic stratigraphic distribution in Northeast China and adjacent region. *Science in China Series D: Earth Sciences*, 52(5), 619–626. <https://doi.org/10.1007/s11430-009-0062-7>

Windley, B. F., Alexeiev, D., Xiao, W., Kröner, A., & Badarch, G. (2007). Tectonic models for accretion of the Central Asian Orogenic Belt. *Journal of the Geological Society*, 164(1), 31–47. <https://doi.org/10.1144/0016-76492006-022>

Wu, G.-Y., Liu, M., & Wang, Y. (2014). Palinspastic reconstruction and geological evolution of Permian residual marine basins bordering China and Mongolia. *Journal of Palaeogeography*, 3(2), 219–232. <https://doi.org/10.3724/SP.J.1261.2014.00053>

Xiao, W., Shen, P., & Chen, X. (2017). Tectonics, magmatism, and mineralization of Circum–Balkash–Junggar area in the Central Asian Orogenic Belt. *International Geology Review*, 59(9), 1047–1052. <https://doi.org/10.1080/00206814.2017.1326181>

## BIOGRAPHY

Dr. Gáspár Albert is a geologist, cartographer and GIS specialist. He currently works at the Eötvös Loránd University as a lecturer teaching basic- and geoscientific mapping and modelling. He works on several research and exploration projects requiring geoinformatics.

Munkhsuren Badrakh is a PhD student at Eötvös Loránd University. Her research focuses on geological and mineral prospectivity mapping using remote sensing data, as well as machine learning and advanced geospatial analysis, particularly in structural interpretation and lineament mapping in complex terrains.