

Explainable Random Forest Gold Prospectivity Mapping Integrating Till and Bedrock Geochemistry in Western Karelia, Finland

Dzulfigar Fawwaz¹, Muhammad Taher²

¹King Abdullah University of Science and Technology, Thuwal, Kingdom of Saudi Arabia, ²Aalto University, Espoo, Finland

Glacial geochemical surveys record displaced mineral signatures that systematically deviate from their bedrock source, complicating prospectivity modelling in ice-covered shield terrains. Integrating surficial till geochemistry with in-situ bedrock data, while maintaining geological interpretability, is therefore central to effective targeting. This study develops an explainable Random Forest workflow for gold prospectivity mapping in the Karelian greenstone-belt terrain of Western Karelia, Finland, using publicly available Geological Survey of Finland (GTK) geochemical databases. Till geochemistry from two analytical suites (511p base-metal ICP-AES and 519u Au-Pd-Te GFAAS) and bedrock geochemistry with encoded geological attributes are integrated through four models evaluated under spatial-block cross-validation (LeaveOneGroupOut on a 4×4 grid) to mitigate spatial autocorrelation leakage.

The till model uses 33 features comprising multi-element concentrations and eight pathfinder ratios (Cu/Zn, Ni/Cr, Co/Ni, Cu/Ni, La/Y, Th/Sc, K/Ca, Fe/Mg). Two bedrock models compare 18 geochemical pathfinder elements alone versus augmented with 11 ordinal-encoded geological attributes (alteration intensity, metamorphic grade, deformation style, ore mineralogy, among others). A fourth combined model fuses till and bedrock features via anisotropic inverse-distance weighting aligned with the dominant NW-SE Weichselian ice-flow direction (~315° azimuth, 3,000 m up-ice search radius), explicitly modeling glacial transport dispersion. Hyperparameter tuning uses RandomizedSearchCV (150 iterations) scored on average precision under spatial-block CV throughout. Preliminary results indicate that the combined till-bedrock model yields stronger prospectivity discrimination than individual models, suggesting that directional data fusion capturing glacial transport geometry adds predictive value.

SHAP (SHapley Additive exPlanations) analysis, complemented by Gini and permutation importance, is applied to assess predictor contributions and verify geological coherence. Prospectivity rasters at 2.5 km resolution are generated to delineate high-potential zones for comparison against known gold occurrences. This workflow demonstrates that glacial-transport-aware data fusion, combined with explainable machine learning and rigorous spatial validation, can provide geologically interpretable decision support for mineral exploration in glaciated shield terrains.