Supplementary material for "Geochemical Anomaly Detection via Supervised Learning: Insights from Interpretable Techniques for a Case Study in Pangxidong Area, South China"

1. The permutation importance for evidence maps based on RF modeling

The core logic of permutation importance is to quantify a variable's contribution by measuring the drop in a model's prediction accuracy when that variable's link to the outcome is broken. This is done by randomly shuffling the values of predictor X_j, which severs its association with Y. A large decrease in accuracy upon using the shuffled data indicates that X_j was originally important, making the accuracy difference a practical importance score [38].



Figure S1. The comparison of Gini importance and Permutation importance based on RF modeling with different training datasets (a, b, c, d for training datasets R,A,B,C).

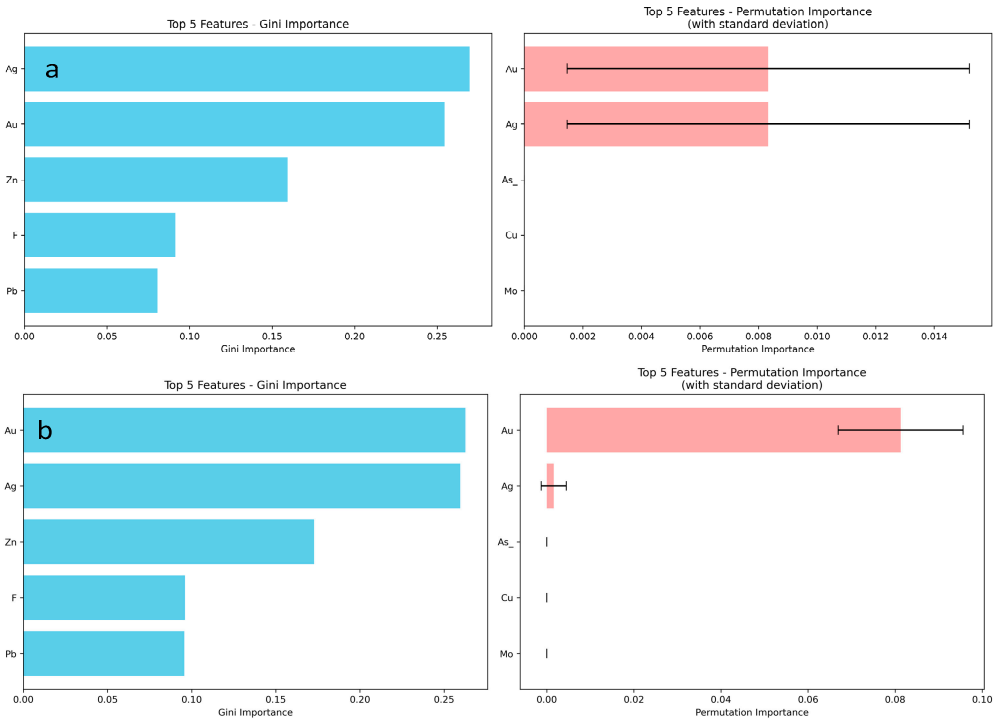


Figure S2. Feature importance based on RF modeling with different test_size=0.4 (a: training dataset B; b: training dataset C).

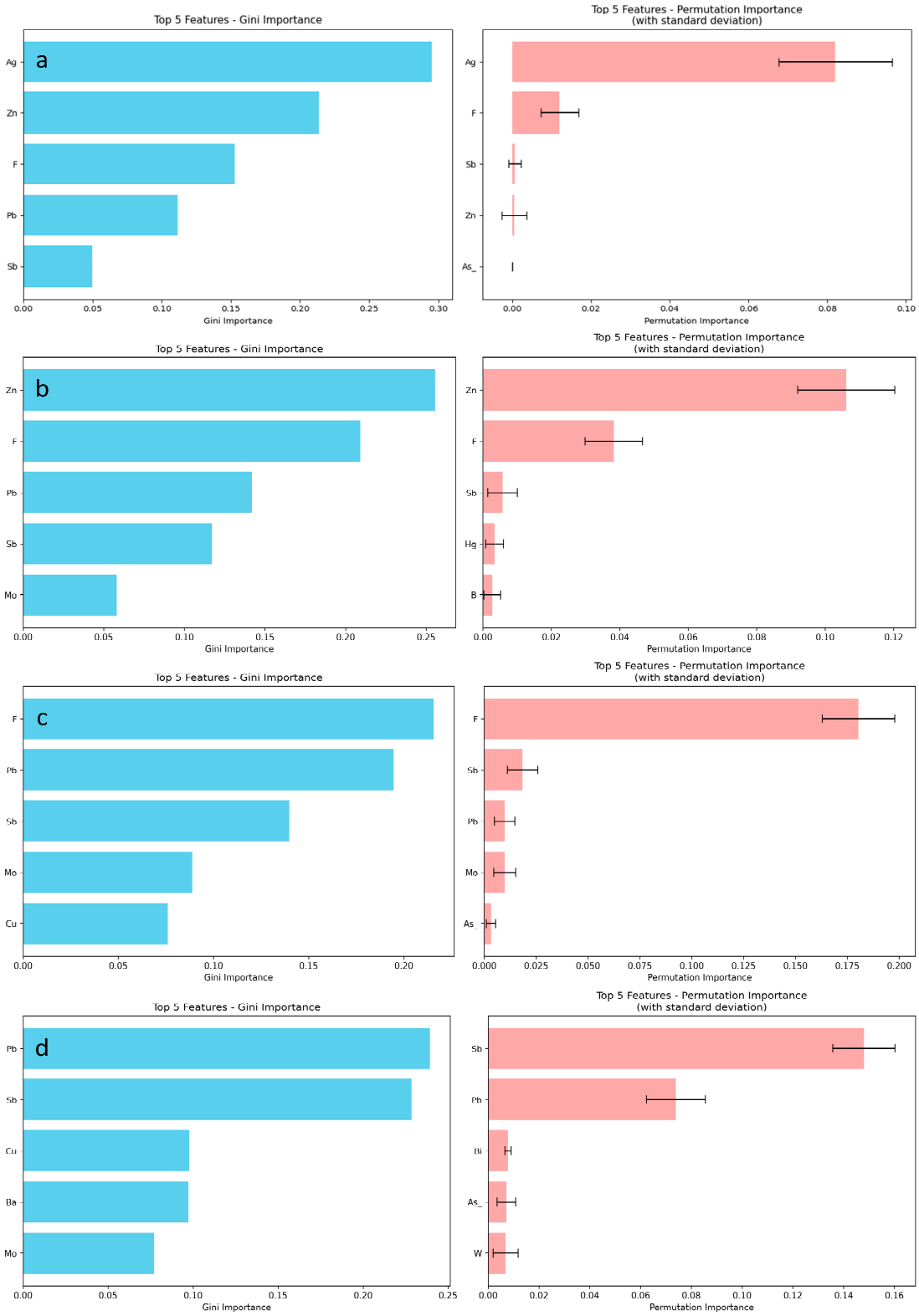


Figure S3. Feature importance based on RF modeling with training dataset (**a**, **b**, **c**, **d** for without Au, Au-Ag, Au-Ag-Zn and Au-Ag-Zn-F).

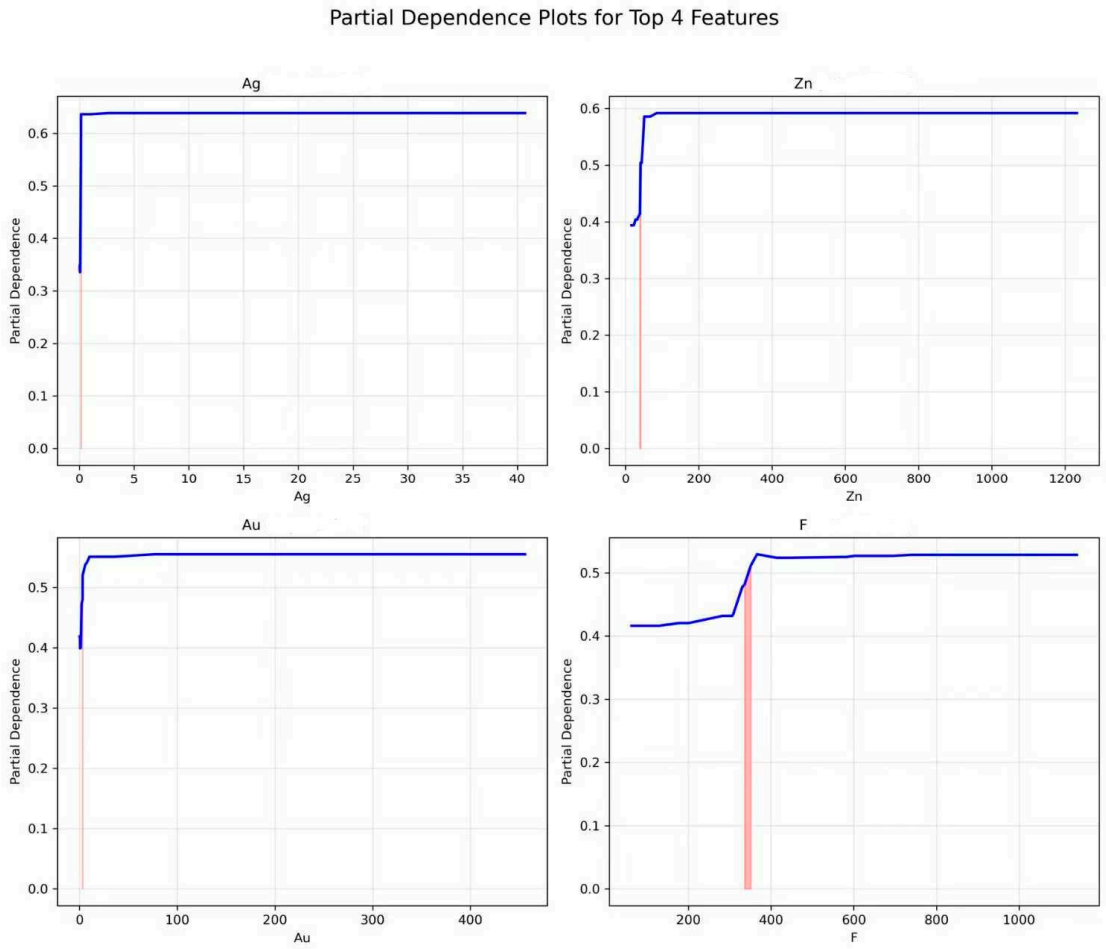


Figure S4. The Partial dependence plots of the evidence maps used in RF modelling with training data R (Red indicate active area).

Reference

38. Strobl, C., A. L. Boulesteix, A. Zeileis and T. Hothorn (2007). "Bias in random forest variable importance measures: Illustrations, sources and a solution." BMC Bioinformatics 8: 25 - 25.