Open Journal of Geology, 2026, 16(8), 544-562
https://www.scirp.org/journal/ojg
ISSN Online: 2161-7589
ISSN Print: 2161-7570

Optimizing Percentile Thresholds and Buffer Distances for Stream-Sediment Geochemical Anomaly Delineation: A Spatial-Efficiency Analysis from the Zhongdian Porphyry-Skarn Metallogenic Belt, Southwestern China

Ohouo Rebecca Mondah*, Fossou Jean Luc Herve Kouadio, Martial Pohn Koffi Adingra, Kouamelan Serger Kouamelan

Unité de Formation et de Recherche des Sciences de la Terre et des Ressources Minères, Laboratoire de Géologie, Ressources Minérales et Énergétiques, Université Félix Houphouët-Boigny Abidjan-Cocody, Abidjan, Côte d'Ivoire
Email: *rbecca.mondah@gmail.com, *mondah.rebecca74@ufhb.edu.ci

How to cite this paper: Mondah, O.R., Kouadio, F.J.L.H., Adingra, M.P.K. and Kouamelan, K.S. (2026) Optimizing Percentile Thresholds and Buffer Distances for Stream-Sediment Geochemical Anomaly Delineation: A Spatial-Efficiency Analysis from the Zhongdian Porphyry-Skarn Metallogenic Belt, Southwestern China. *Open Journal of Geology*, **16**, 544-562.
https://doi.org/10.4236/ojg.2026.168027

Received: July 23, 2026
Accepted: August 28, 2026
Published: August 31, 2026

Abstract

We jointly evaluated percentile threshold and buffer distance for stream-sediment geochemical anomaly mapping in the Zhongdian porphyry-skarn belt of northwestern Yunnan. The dataset comprises 4307 geochemical cells on a 500 m lattice and 13 mineralized targets. Seven elements and three ratios were tested at P85, P90, and P95 using six buffer radii from 250 to 2000 m. For each of the 180 configurations, buffers were dissolved, clipped to a fixed 1076.75 km^2 boundary, and measured directly. Performance was assessed by target capture and spatial efficiency, with Wilson intervals, 1000 spatial-block bootstrap replicates, blocked validation, and 999 max-statistic spatial permutations. Cu-P95 at 250 m gave the highest full-data efficiency (SE = 11.63), capturing 6 of 13 targets within 3.97% of the area. Pb-P95, Ag-P95, and Cu/(Pb + Zn)-P95 at the same radius ranked next. Seven configurations exceeded the search-corrected 5% familywise threshold of SE = 5.87. Capture favored a broader setting: Cu-P95 at 750 m captured 12 of 13 targets within 18.35% of the area, whereas extending the buffer to 1000 m added footprint without adding capture. The resulting eight-configuration efficiency frontier separates selective and higher-recall operating rules. Spatially blocked selection recovered P95 and 250 m as the modal setting, but global out-of-fold capture was 23.1% (95% CI: 8.2% - 50.3%), showing that the full-data maximum overstates transferable performance. The selected 250 m and 750 m configurations were finally converted into prospectivity-target footprints. The workflow is reproducible, but the numerical settings are specific to the gridded Zhongdian dataset and should be re-optimized locally.

Keywords

Stream-Sediment Geochemistry, Percentile Threshold, Buffer Distance, Anomaly Footprint, Spatial Efficiency, Efficiency Frontier, Spatial Cross-Validation, Prospectivity Targeting, Zhongdian Arc

1. Introduction

Stream-sediment geochemistry remains one of the most practical reconnaissance methods in mountainous mineral belts. Active sediment integrates material eroded from an upstream catchment and can therefore indicate mineralized sources that are poorly exposed at the surface [1]-[3]. The value of a regional survey, however, depends on how convincingly high concentrations are converted into areas for follow-up.

Two operational choices control that conversion. The first is the threshold above which a value is treated as anomalous. Percentile thresholds are attractive for strongly right-skewed geochemical distributions, but a percentile identifies a proportion of observations, not a proportion of land area [4]. The second choice is the distance over which an anomalous observation is taken to indicate proximity to mineralization. That distance is influenced by catchment geometry, downstream dilution, relief, sampling density and source size [2] [5]-[7]. Threshold and radius must therefore be considered together: stricter thresholds reduce the area selected but can miss deposits, whereas larger buffers increase capture at the cost of a larger footprint.

Prediction-area analysis provides a direct way to express this trade-off by comparing the proportion of deposits captured with the proportion of the study area selected [8] [9]. Its implementation is straightforward only if the map footprint is measured correctly. Once anomalous locations are buffered, overlapping polygons must be dissolved and clipped to a fixed boundary. Replacing this unioned area with the selected-sample fraction fixes the denominator even as the radius changes, making any apparent buffer optimum internally inconsistent.

A second problem arises when many configurations are searched against a small inventory of known targets. The largest value among 180 element-threshold-radius combinations is not an unbiased estimate of the selected rule. Spatial autocorrelation further weakens random cross-validation because nearby targets can occur in both training and test sets [10] [11]. Spatially separated validation, bootstrap assessment of rank stability and a max-statistic permutation benchmark are consequently needed before an optimum is interpreted.

The Zhongdian porphyry-skarn belt provides an appropriate test case. It contains a Cu-dominant Late Triassic porphyry system together with Pb-Zn-Ag and Au-bearing occurrences distributed along the southern Yidun arc. We use the available 500 m geochemical grid and a fixed inventory of 13 mineralized targets to address four questions: which threshold-radius combinations are spatially efficient; how does the answer change when capture rather than selectivity is prioritized; which results survive search-corrected and spatially blocked validation; and how can the selected configurations be translated into prospectivity targets?

2. Geological Setting and Analytical Dataset

2.1. Regional Tectonic Framework

The Zhongdian belt lies in the southern Yidun terrane, between the Ganzi-Litang suture to the east and the Jinshajiang suture to the west (**Figure 1(a)**, **Figure 1(b)**). The terrane forms part of the eastern Tethyan domain and records westward subduction of the Ganzi-Litang ocean beneath the Yidun block during the Late Triassic [12] [13]. Arc magmatism produced a north-northwest-trending belt of intermediate to felsic intrusions and associated hydrothermal systems.

Upper Triassic volcanic, volcaniclastic, carbonate and clastic rocks, commonly assigned to the Tumugou Formation and related units, form the principal country-rock succession. These rocks were intruded by quartz diorite, granodiorite, monzodiorite and quartz monzonite porphyries (**Figure 1(b)**). Carbonate-bearing intervals favored skarn and replacement mineralization, whereas porphyry-style stockworks and disseminations are concentrated within and around the intrusive centers [13]-[15].

2.2. Magmatism and Mineralization

Ore formation was concentrated in the Late Triassic. Published geochronology places the Pulang and Songnuo intrusive systems broadly near 216 Ma, while the Xuejiping porphyry system is approximately 213 Ma [14]-[17]. Pulang is the largest known porphyry Cu-Au district in the belt and comprises several intrusive phases with potassic, phyllic and propylitic alteration and Cu-Au-bearing stockwork mineralization [15] [16]. Smaller Cu-Mo, Cu-Au-Ag, Pb-Zn and polymetallic occurrences are distributed along intrusive and structural corridors.

The 13 validation targets include the Cu-dominant Pulang, Lannitang, Gaochiping, Chundu and Qiansui occurrences; Cu-Mo associations at Relin and Hongshan; Cu-Au-Ag associations at Xuejiping and Pulang; and Pb-Zn-Ag-Au-bearing targets such as Langdu, Yaza, Disuga and Songnuo (Supplementary **Table S1**). Their locations are clustered along several NNW-trending corridors, with a central group and more isolated northern and southern targets (**Figure 1(c)**). This spatial clustering is one reason why random validation folds would be misleading and why spatially separated testing was used below. Because the available inventory reports metal associations rather than verified genetic classes, the analysis does not assign individual targets to porphyry, skarn, or epithermal subsets.

2.3. Geochemical Grid and Validation Extent

The analytical surface contains 4307 grid cells at an exact 500 m spacing. It is a gridded geochemical product, not the original stream-sediment sample table. A

500 × 500 m square centered on each node produces a non-overlapping coverage of 1076.75 km². All 13 targets lie within this coverage, at distances of 108.0 - 322.3 m from the nearest grid node.

Coordinate values are consistent with a projected UTM zone 47N system. WGS 84/UTM zone 47N (EPSG:32647) was used as the working reference for display and area calculations, but the source CSV files do not store CRS metadata. The definitive datum and the interpolation history of the grid should therefore be confirmed from the originating GIS archive. This limitation does not affect relative distances and areas because the grid and target coordinates share the same meter-based system.

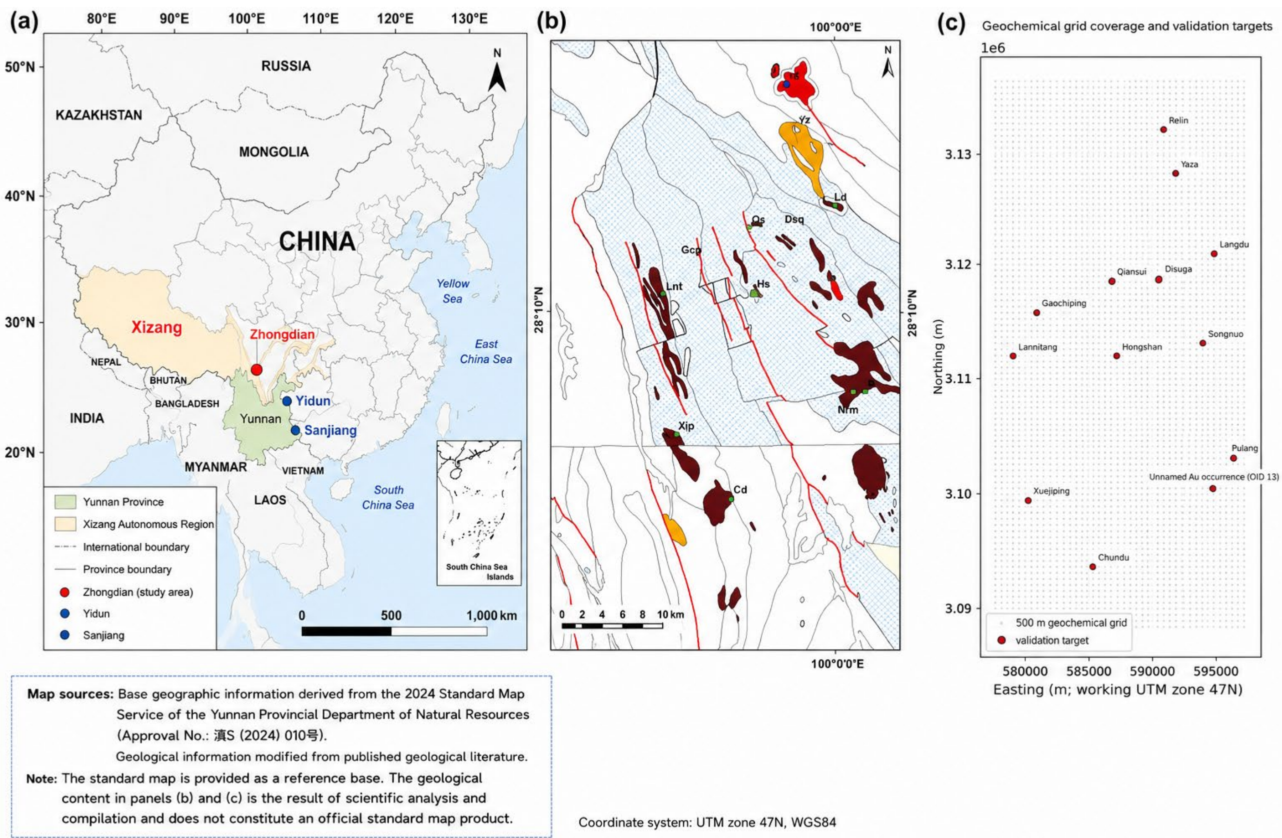

Figure 1. Regional geology and analytical coverage of the Zhongdian porphyry-skarn belt. (a) Regional location of the study area in southwestern China, based on the 2024 Standard Map of Yunnan Province issued by the Yunnan Provincial Department of Natural Resources. (b) Simplified geology, major intrusions, structures, and representative mineral occurrences of the Zhongdian-Yidun region, modified from Mondah *et al*., 2021 [18]. (c) Extent of the 4307-cell geochemical grid and the 13 validation targets in the working UTM zone 47N system.

3. Materials and Methods

3.1. Geochemical Variables and Ratios

Cu, Pb, Zn, Ag, Au, Mo and W are available for every grid cell. Cu, Pb, Zn, Ag, Mo and W are expressed in micrograms per gram, whereas Au is expressed in nanograms per gram. No coordinates are missing or duplicated. Three ratios were calculated on positive concentrations: RCu = Cu/(Pb + Zn), RMo = Mo/(Cu + W), and RAu = Au/Ag. Au was converted to micrograms per gram before RAu was calculated, so numerator and denominator share common mass units. Ratios were formed on the concentration scale, not by dividing logarithms. Each of the ten variables was thresholded at P85, P90, and P95.

3.2. Validation-Target Inventory

The target table contains 13-point locations. One Au occurrence is unnamed and was retained as OID 13 so that the validation set was not altered. Eight records contain national catalogue codes and size-class attributes; the remaining five contain names, commodity associations and coordinates only. Discovery status, positional uncertainty and verified genetic type are not recorded. **Supplementary Table S1** reports on the available fields and nearest-grid distances. Commodity groups were not analyzed separately because the Pb-Zn, Au-only and polymetallic strata contain only two, one and one targets, respectively.

3.3. Boundary, Buffers and True Anomaly Footprint

The study boundary B was generated as the union of the 4307 square cells and has an area of 1076.75 km^2. For every variable-threshold combination, selected cell centroids were buffered at 250, 500, 750, 1000, 1500 and 2000 m. Overlapping polygons were dissolved and intersected with B. The resulting union-buffer footprint therefore varies with the spatial arrangement of anomalous cells and with buffer radius.

3.4. Performance Measures and Efficiency Frontier

For a configuration capturing nc of $N = 13$ targets within a clipped union-buffer area U_B, capture C, footprint A, and spatial efficiency SE were calculated as:

$$C = 100 n_c / N \quad (1)$$

$$A = 100\, Area\left(U_B \cap B\right) / Area\left(B\right) \quad (2)$$

$$SE = C / A \quad (3)$$

SE = 1 is the geometric expectation for a single configuration under uniformly distributed targets. Because 180 configurations were searched and the targets are clustered, statistical evidence was based on the max-statistic permutation test rather than on SE > 1 alone. Wilson 95% intervals were used for capture proportions [19].

Capture and footprint were also evaluated as a two-objective problem. A configuration was considered non-dominated when no alternative captured more targets within an equal or smaller footprint. The set of non-dominated configurations defines an efficiency frontier and separates selective first-pass rules from higher-recall rules.

3.5. Spatial Validation, Bootstrap and Permutation Testing

Two spatial validation schemes were used. In spatial-exclusion leave-one-target-out validation, each target was withheld, and training targets within 5 km were excluded before the best configuration was selected. In the second scheme, 5 km spatial blocks were held out in turn. Selection was repeated globally across all ten variables and within each variable separately. Rank stability was assessed with 1000 spatial-block bootstrap replicates [20].

The search-corrected benchmark was obtained from 999 random sets of 13 points generated uniformly within the study boundary. Every random set was scored against all 180 configurations, and only the maximum SE was retained. The familywise p-value used the standard plus-one correction [21]. This procedure controls the fact that the reported maximum was selected from a search rather than specified in advance.

3.6. Prospectivity-Target Delineation and Reproducibility

The selected Cu-P95 configurations were converted into two prospectivity products. The 250 m union-buffer footprint represents the most spatially selective screen, whereas the 750 m footprint represents the higher-recall alternative on the efficiency frontier. These polygons are follow-up target zones, not newly discovered deposits; their geological significance requires independent mapping, geophysics, geochemistry and field verification.

Processing was implemented in Python using pandas, GeoPandas, Shapely, NumPy, Matplotlib and PyProj, with the random seed fixed at 20260730. The input files, analysis script, metadata, and complete configuration table accompany the revision. The full analytical sequence is summarized in Figure 2.

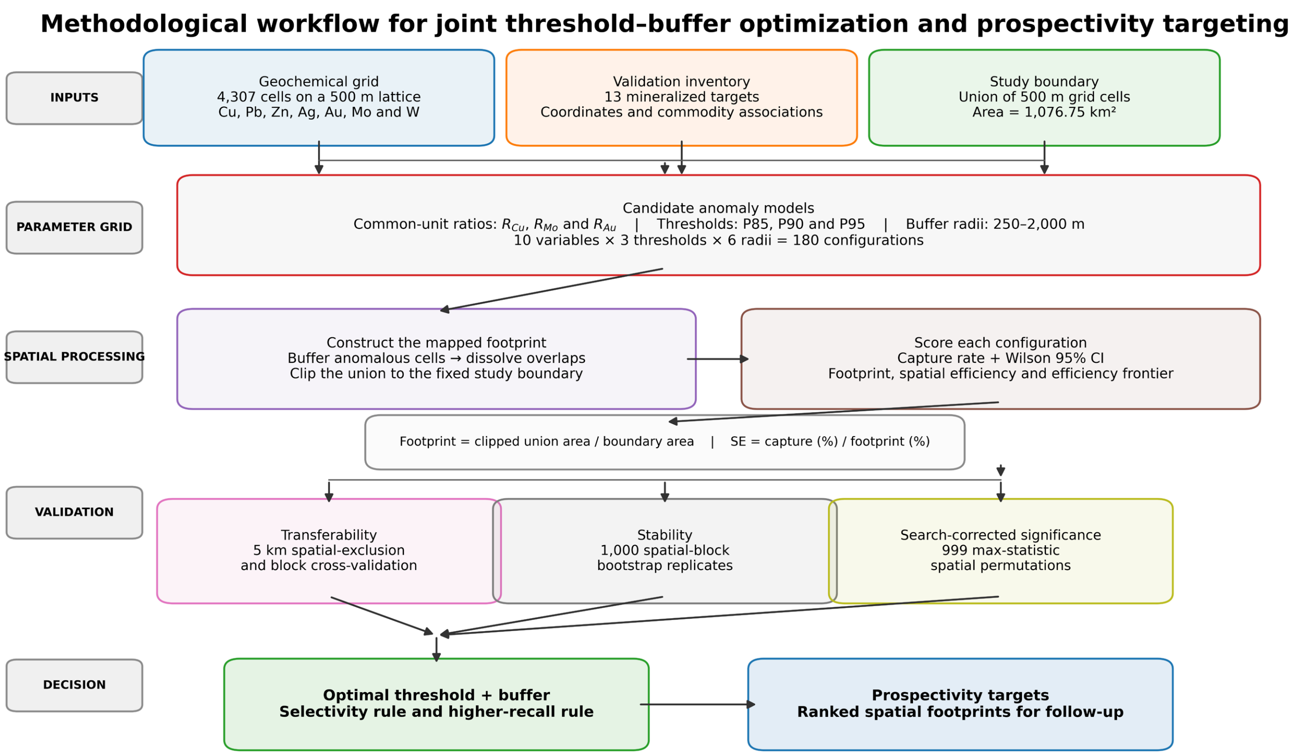

Figure 2. Methodological workflow. The analysis proceeds from common-unit variables, percentile thresholds, and buffer radii to dissolved, boundary-clipped footprints; all 180 configurations are then evaluated by blocked validation, bootstrap stability, and max-statistic permutation testing. The workflow ends with an optimal threshold-plus-buffer rule and the corresponding prospectivity targets.

4. Results

4.1. True Footprints and the Effect of the Denominator

P85, P90 and P95 select approximately 15%, 10% and 5% of the grid cells, but these fractions are not the mapped areas produced after buffering. At 1000 m, P95 footprints range from 19.40% of the boundary for W to 35.89% for RAu; Cu occupies 25.57% (Table 1). At P85 and the same radius, single-element footprints range from 46.74% for Pb to 68.21% for Ag.

Using the selected-cell fraction in place of the true footprint inflates SE by 3.80 - 7.16 times at 1,000 m and P95, with a mean inflation factor of 5.48 (Table 2). For Cu, the approximation changes SE from 3.61 to 18.24. The difference is large enough to alter both variable ranking and the selected radius.

Table 1. Thresholds, anomalous-cell counts, clipped union-buffer footprints, target capture and spatial efficiency at the 1000 m reference radius. Concentrations are μg/g except Au (ng/g); ratios are dimensionless after common-unit conversion. CI = Wilson 95% confidence interval.

Variable	Pct.	Threshold	Anomalous cells n (%)	Footprint km² (%)	Capture n/13	Capture % (95% CI)	SE
Cu	85	65.31	646 (15.00)	610.97 (56.74)	13/13	100.0 (77.2 - 100.0)	1.76
Cu	90	79.80	433 (10.05)	445.41 (41.37)	13/13	100.0 (77.2 - 100.0)	2.42
Cu	95	126.00	218 (5.06)	275.36 (25.57)	12/13	92.3 (66.7 - 98.6)	3.61
Pb	85	102.00	651 (15.11)	503.28 (46.74)	11/13	84.6 (57.8 - 95.7)	1.81
Pb	90	149.00	432 (10.03)	391.35 (36.35)	10/13	76.9 (49.7 - 91.8)	2.12
Pb	95	282.70	216 (5.02)	265.12 (24.62)	8/13	61.5 (35.5 - 82.3)	2.50
Zn	85	151.00	649 (15.07)	665.85 (61.84)	10/13	76.9 (49.7 - 91.8)	1.24
Zn	90	181.00	433 (10.05)	505.38 (46.94)	10/13	76.9 (49.7 - 91.8)	1.64
Zn	95	252.00	217 (5.04)	304.76 (28.30)	6/13	46.2 (23.2 - 70.9)	1.63
Ag	85	0.400	667 (15.49)	734.41 (68.21)	13/13	100.0 (77.2 - 100.0)	1.47
Ag	90	0.500	457 (10.61)	579.55 (53.82)	13/13	100.0 (77.2 - 100.0)	1.86
Ag	95	0.800	223 (5.18)	338.19 (31.41)	12/13	92.3 (66.7 - 98.6)	2.94
Au	85	3.70	657 (15.25)	657.32 (61.05)	13/13	100.0 (77.2 - 100.0)	1.64
Au	90	5.00	440 (10.22)	508.39 (47.21)	13/13	100.0 (77.2 - 100.0)	2.12
Au	95	8.00	218 (5.06)	284.33 (26.41)	11/13	84.6 (57.8 - 95.7)	3.20
Mo	85	2.00	690 (16.02)	698.36 (64.86)	10/13	76.9 (49.7 - 91.8)	1.19
Mo	90	2.70	437 (10.15)	517.62 (48.07)	9/13	69.2 (42.4 - 87.3)	1.44
Mo	95	4.07	216 (5.02)	308.38 (28.64)	9/13	69.2 (42.4 - 87.3)	2.42
W	85	3.60	657 (15.25)	641.96 (59.62)	11/13	84.6 (57.8 - 95.7)	1.42
W	90	4.20	432 (10.03)	450.87 (41.87)	10/13	76.9 (49.7 - 91.8)	1.84
W	95	6.80	220 (5.11)	208.93 (19.40)	4/13	30.8 (12.7 - 57.6)	1.59
RCu	85	0.380	646 (15.00)	676.24 (62.80)	13/13	100.0 (77.2 - 100.0)	1.59

Continued

RCu	90	0.443	431 (10.01)	516.60 (47.98)	11/13	84.6 (57.8 - 95.7)	1.76
RCu	95	0.599	216 (5.02)	271.53 (25.22)	11/13	84.6 (57.8 - 95.7)	3.36
RMo	85	0.051	646 (15.00)	731.15 (67.90)	8/13	61.5 (35.5 - 82.3)	0.91
RMo	90	0.062	431 (10.01)	575.92 (53.49)	6/13	46.2 (23.2 - 70.9)	0.86
RMo	95	0.090	216 (5.02)	339.04 (31.49)	5/13	38.5 (17.7 - 64.5)	1.22
RAu	85	0.030	646 (15.00)	763.46 (70.90)	10/13	76.9 (49.7 - 91.8)	1.08
RAu	90	0.036	437 (10.15)	604.03 (56.10)	9/13	69.2 (42.4 - 87.3)	1.23
RAu	95	0.048	216 (5.02)	386.40 (35.89)	6/13	46.2 (23.2 - 70.9)	1.29

Table 2. Effect of the footprint denominator at 1000 m and P95. The selected-cell approximation divides capture by the fraction of selected cells; the corrected index uses the clipped union-buffer footprint.

Variable	Selected-cell fraction (%)	True footprint (%)	Approximate SE	Corrected SE	Inflation
Cu	5.06	25.57	18.24	3.61	5.05×
Pb	5.02	24.62	12.27	2.50	4.91×
Zn	5.04	28.30	9.16	1.63	5.62×
Ag	5.18	31.41	17.83	2.94	6.07×
Au	5.06	26.41	16.72	3.20	5.22×
Mo	5.02	28.64	13.80	2.42	5.71×
W	5.11	19.40	6.02	1.59	3.80×
RCu	5.02	25.22	16.87	3.36	5.03×
RMo	5.02	31.49	7.67	1.22	6.28×
RAu	5.02	35.89	9.20	1.29	7.16×
Mean	5.05	27.69	-	-	5.48×

4.2. Footprint Growth and Buffer Coalescence

Footprint growth explains why the denominator must respond to radius. For Cu-P95, 218 separate 250 m discs would cover 42.8 km^2 and the measured union covers 42.74 km^2, an overlap loss of only 0.2%. The loss rises to 30.3% at 500 m, 48.7% at 750 m, 59.8% at 1000 m, 72.8% at 1500 m and 79.9% at 2000 m. The same pattern occurs across all ten variables (Figure 3). At 250 m, the buffer diameter equals the grid spacing; beyond that distance, successive increments mainly fill gaps between neighboring anomalous cells.

4.3. Capture Response to Buffer Radius

Target capture rises monotonically with radius, but at different rates among variables and thresholds (Figure 4). Cu-P95 captures 46.2% of targets at 250 m, 61.5% at 500 m, and 92.3% at 750 m. The 1000 m radius adds no Cu-P95 target, and complete capture is reached only at 1500 m. Ag-P95 rises from 38.5% at 250 m to 92.3% at 1000 m, whereas Au-P95 rises from 23.1% to 84.6% over the same interval. W-P95 remains at or below 30.8% through 1500 m and reaches 61.5% only at 2000 m.

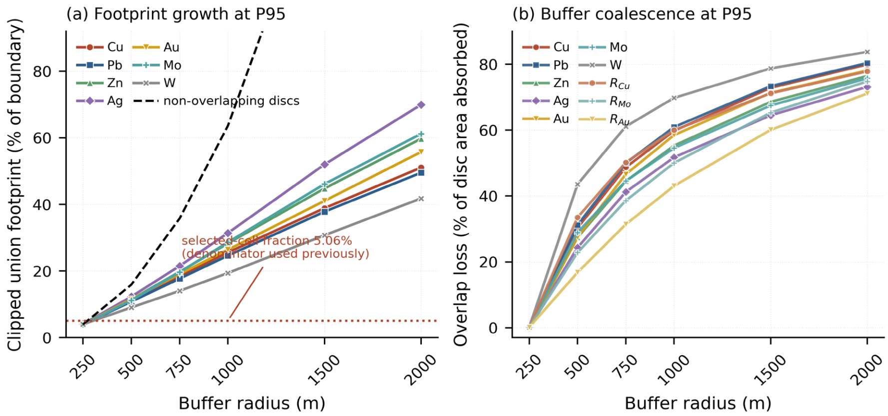

Figure 3. (a) Clipped union-buffer footprint against radius at P95. The dashed curve shows the nominal area of separate Cu buffers; the horizontal line marks the selected-cell fraction. (b) Overlap loss, expressed as the proportion of nominal disc area absorbed by buffer coalescence, for all ten variables at P95.

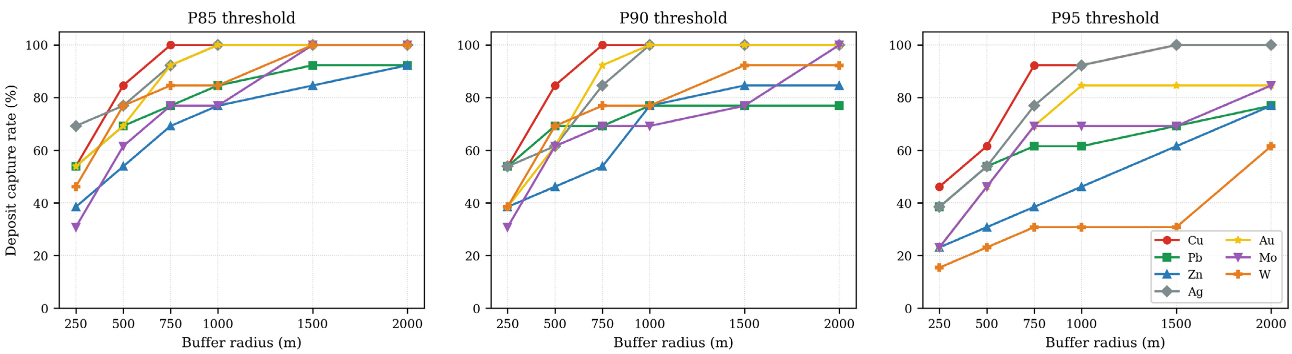

Figure 4. Target capture against buffer radius for the seven single elements at P85, P90 and P95. With 13 validation targets, capture changes in increments of 7.7 percentage points.

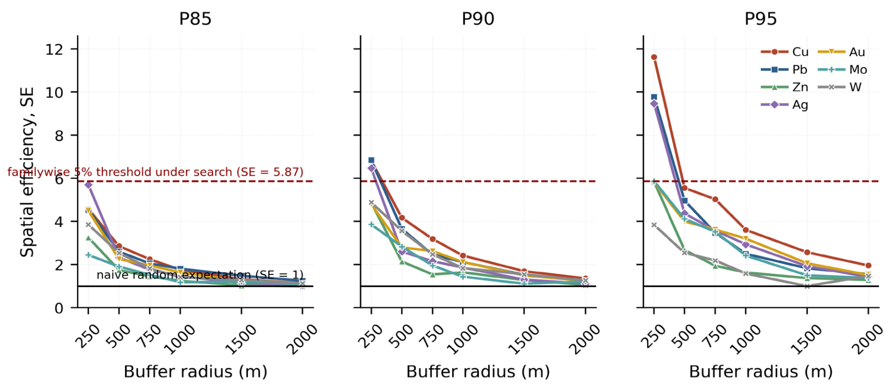

Figure 5. Spatial efficiency against buffer radius. The upper horizontal line is the 5% familywise threshold from the max-statistic permutation test (SE = 5.87); the lower line is the conventional single-configuration expectation (SE = 1).

4.4. Spatial Efficiency and the Radius Optimum

Spatial efficiency changes in the opposite direction. Because footprint expands faster than capture, SE decreases with radius for every variable and threshold (Figure 5). Cu-P95 at 250 m is the maximum among all 180 configurations (SE = 11.63), capturing 6 of 13 targets inside 3.97% of the boundary. Pb-P95 (SE = 9.78), Ag-P95 (9.47), and RCu-P95 (7.82) at 250 m follow. The design contains no interior efficiency maximum: the statistical optimum lies at the shortest radius tested.

4.5. Joint Threshold-Radius Efficiency Frontier

A single maximum does not summarize the operational trade-off because capture and efficiency favor different settings. Eight configurations are non-dominated: no alternative captures more targets within an equal or smaller footprint (Figure 6; Table 3). The frontier extends from Pb-P95 and Cu-P95 at 250 m, which are highly selective, to Cu-P90 at 750 m, which captures all 13 targets within 31.36% of the boundary. Cu-P95 at 750 m is the most attractive high-recall compromise, capturing 12 of 13 targets within 18.35% of the area.

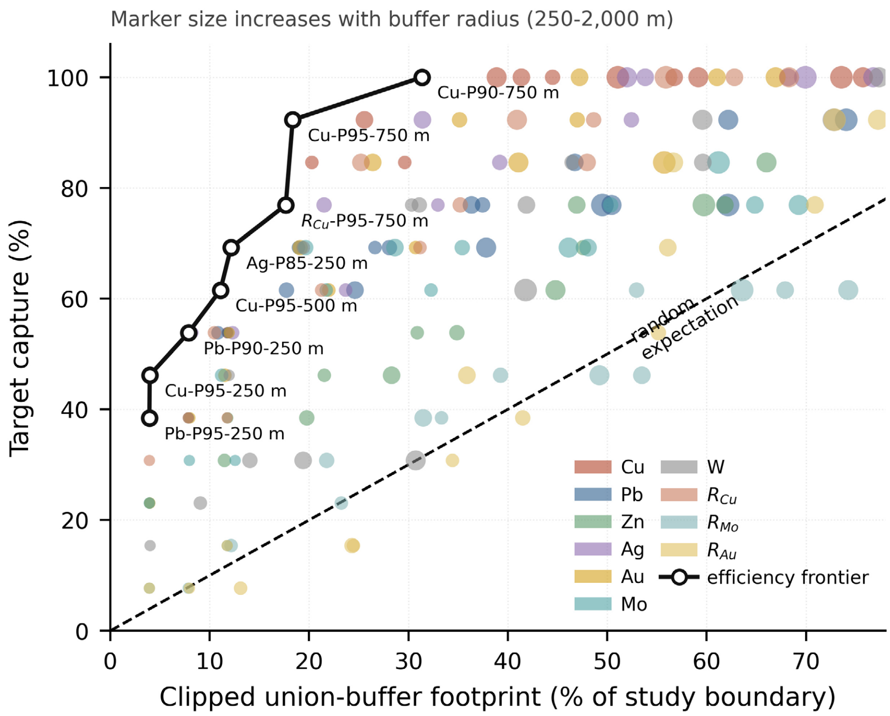

Figure 6. Capture versus clipped footprint for all 180 configurations. Marker size increases with buffer radius. The heavy line joins the eight non-dominated configurations that form the efficiency frontier; the diagonal is the conventional random expectation.

Table 3. Efficiency frontier. Each row is non-dominated: no alternative captures more targets within an equal or smaller footprint.

Configuration	Capture n/13	Capture % (95% CI)	Footprint (km²)	Footprint (%)	SE
Pb-P95-250 m	5/13	38.5 (17.7 - 64.5)	42.3	3.93	9.78
Cu-P95-250 m	6/13	46.2 (23.2 - 70.9)	42.7	3.97	11.63
Pb-P90-250 m	7/13	53.8 (29.1 - 76.8)	84.7	7.87	6.85
Cu-P95-500 m	8/13	61.5 (35.5 - 82.3)	119.3	11.08	5.55

Continued

Ag-P85-250 m	9/13	69.2 (42.4 - 87.3)	130.8	12.14	5.70
RCu-P95-750 m	10/13	76.9 (49.7 - 91.8)	190.1	17.66	4.36
Cu-P95-750 m	12/13	92.3 (66.7 - 98.6)	197.6	18.35	5.03
Cu-P90-750 m	13/13	100.0 (77.2 - 100.0)	337.7	31.36	3.19

4.6. Performance at the 1000 m Reference Radius

At 1000 m and P95, the corrected ranking is Cu (SE = 3.61), RCu (3.36), Au (3.20), Ag (2.94), Pb (2.50), Mo (2.42), Zn (1.63), W (1.59), RAu (1.29) and RMo (1.22) (Figure 7; Table 1). None reaches the familywise threshold of 5.87. This radius is therefore useful only as a comparison with earlier practice; it is neither the efficiency optimum nor the most economical high-recall rule.

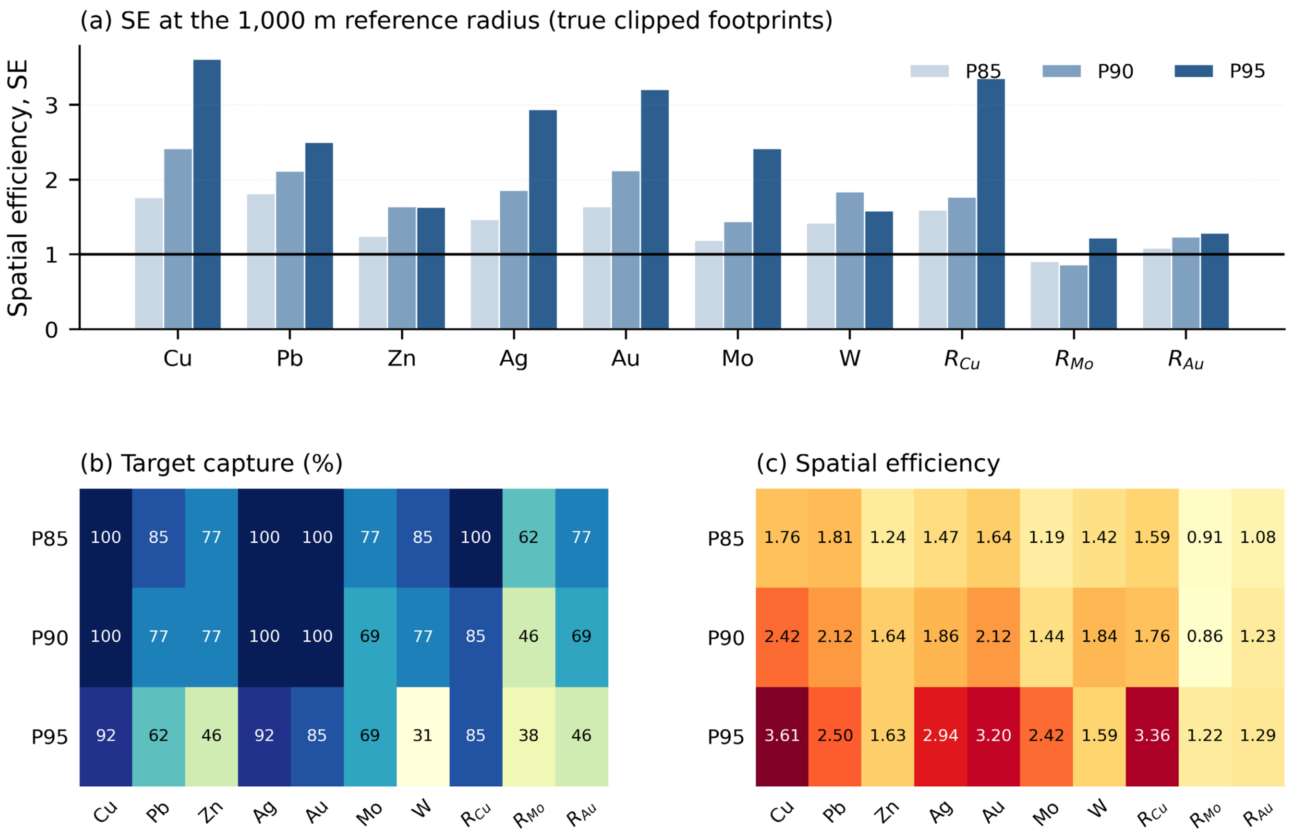

Figure 7. (a) Spatial efficiency at the 1000 m reference radius for all ten variables, calculated from true clipped footprints. (b) Target capture and (c) spatial-efficiency heat maps. No value reaches the familywise threshold of SE = 5.87.

4.7. Search-Corrected Significance and Stability

The distribution of the maximum SE from 999 spatial permutations has a median of 2.93 and a 95th percentile of 5.87 (Figure 8(a)). The observed maximum of 11.63 yields a familywise p-value of 0.001. Seven configurations exceed the 5% familywise threshold; all occur at 250 m and involve Cu, Pb, Ag, or RCu (Table 4). The effect of correcting the footprint denominator and re-optimizing the configuration is summarized in Figure 8(d).

Table 4. Configurations exceeding the 5% familywise threshold (SE = 5.87) from 999 max-statistic spatial permutations.

Configuration	Capture n/13	Capture (%)	Footprint (%)	SE	SE/Threshold
Cu-P95-250 m	6/13	46.2	3.97	11.63	1.98×
Pb-P95-250 m	5/13	38.5	3.93	9.78	1.67×
Ag-P95-250 m	5/13	38.5	4.06	9.47	1.61×
RCu-P95-250 m	4/13	30.8	3.93	7.82	1.33×
Pb-P90-250 m	7/13	53.8	7.87	6.85	1.17×
Cu-P90-250 m	7/13	53.8	7.88	6.83	1.16×
Ag-P90-250 m	7/13	53.8	8.32	6.47	1.10×

Across 1000 spatial-block bootstrap replicates, Cu-P95-250 m is the global winner in 45.5% of replicates, Pb-P95-250 m in 33.6%, and Ag-P95-250 m in 9.5% (Figure 8(b)). Within Cu alone, the same configuration is selected in 97.8% of replicates. Its bootstrap median SE is 11.63 with a 95% interval of 5.81 - 17.44. Threshold and radius are stable; the leading variable is less certain because Pb and Ag together win more than 40% of global replicates.

4.8. Spatially Out-of-Fold Performance

Spatial-exclusion and regular-block validation recover P95 and 250 m as the modal setting for most variables; W favors P90 at 250 m. When the pathfinder variable is fixed in advance, out-of-fold capture is 46.2% for Cu, 38.5% for Pb, Ag, and W, and 30.8% for RCu. When selection is allowed across all ten variables, out-of-fold capture falls to 3 of 13 targets, or 23.1% (95% CI: 8.2% - 50.3%) (Figure 8(c); Table 5). The setting is stable between folds, but the full-data capture of the selected winner is optimistic.

Table 5. Spatially out-of-fold performance and configuration stability. Capture is from 5 km spatial-exclusion leave-one-target-out validation; the final column reports within-variable selection across 1000 spatial-block bootstrap replicates.

Selection scope	Modal configuration	OOF capture n/13 (%)	95% CI (%)	Mean footprint (%)	Bootstrap selection (%)
Cu	Cu-P95-250 m	6/13 (46.2)	23.2 - 70.9	3.97	97.8
Pb	Pb-P95-250 m	5/13 (38.5)	17.7 - 64.5	3.93	91.0
Zn	Zn-P95-250 m	3/13 (23.1)	8.2 - 50.3	5.16	57.8
Ag	Ag-P95-250 m	5/13 (38.5)	17.7 - 64.5	4.06	86.0
Au	Au-P95-250 m	3/13 (23.1)	8.2 - 50.3	3.97	60.7
Mo	Mo-P95-250 m	3/13 (23.1)	8.2 - 50.3	3.93	70.4
W	W-P90-250 m	5/13 (38.5)	17.7 - 64.5	7.87	56.6
RCu	RCu-P95-250 m	4/13 (30.8)	12.7 - 57.6	3.93	82.8
RMo	RMo-P95-250 m	1/13 (7.7)	1.4 - 33.3	5.30	48.2
RAu	RAu-P95-250 m	2/13 (15.4)	4.3 - 42.2	12.04	49.0
Global	Cu-P95-250 m	3/13 (23.1)	8.2 - 50.3	3.95	-

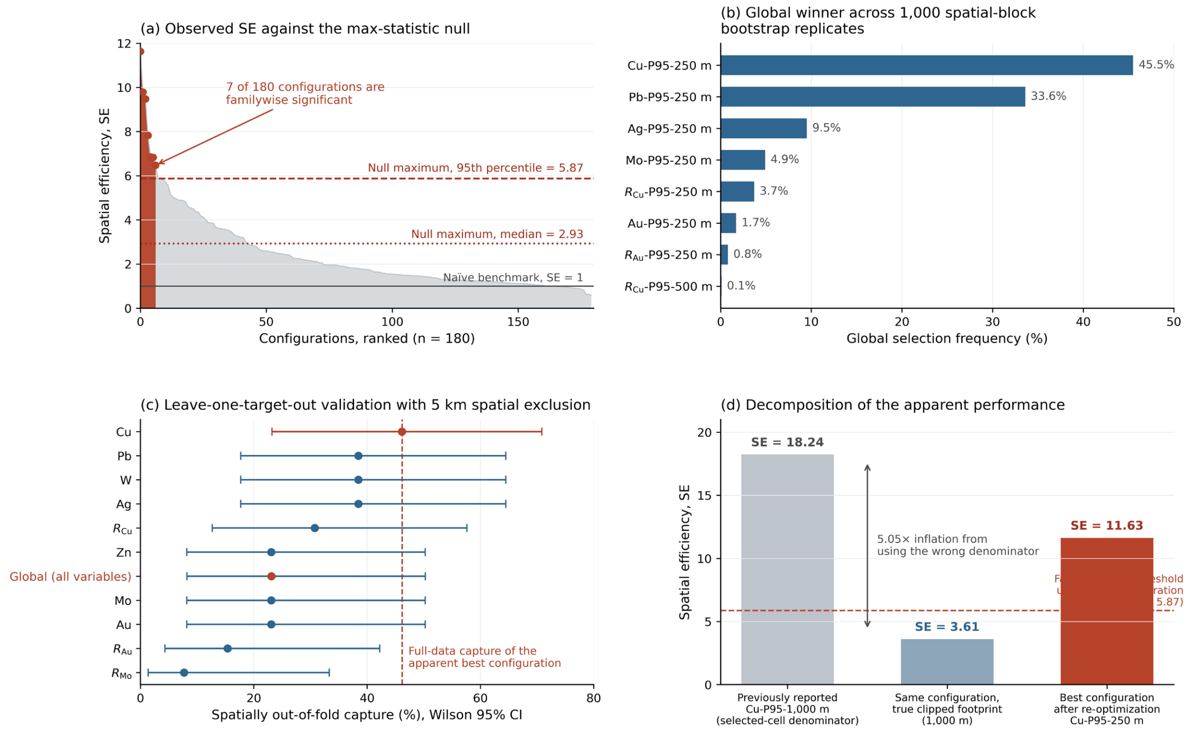

Figure 8. Validation and uncertainty. (a) Observed spatial efficiencies ranked against the max-statistic null. (b) Global winner frequencies across 1000 spatial-block bootstrap replicates. (c) Spatial-exclusion out-of-fold capture with Wilson 95% intervals. (d) Decomposition of the apparent performance into denominator choice, corrected 1000 m performance and the re-optimized maximum.

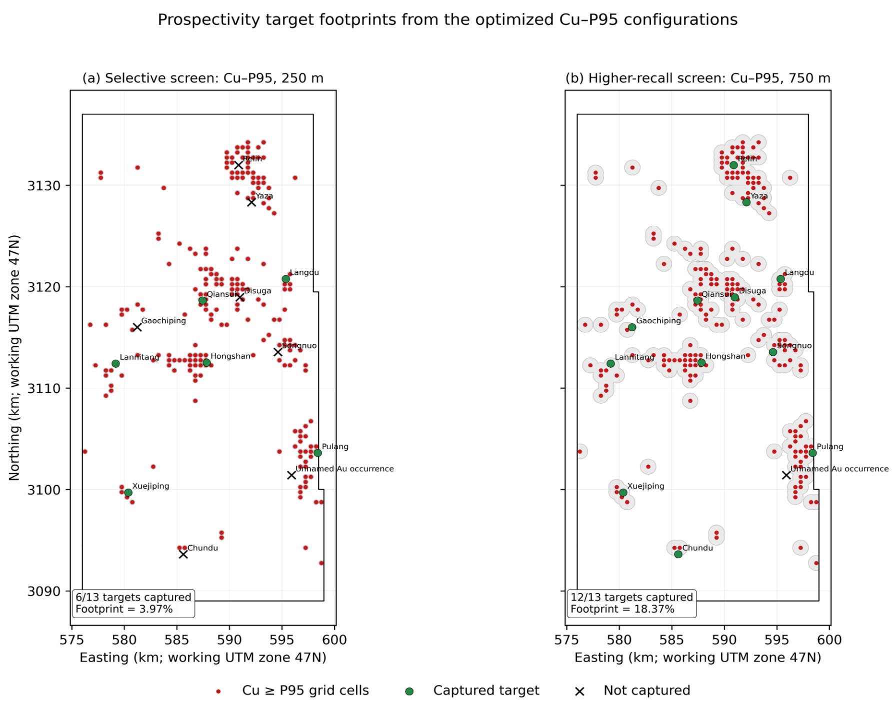

Figure 9. Prospectivity target footprints derived from Cu-P95. (a) The 250 m selective screen captures 6 of 13 targets within 3.97% of the boundary. (b) The 750 m higher-recall screen captures 12 of 13 within 18.35%. Filled circles mark captured validation targets; crosses mark targets outside the selected footprint.

4.9. Prospectivity Targets from the Selected Configurations

The optimal settings were converted into two prospectivity footprints (**Figure 9**). The Cu-P95-250 m screen selects 3.97% of the belt and captures Langdu, Lannitang, Hongshan, Xuejiping, Pulang and Qiansui. The Cu-P95-750 m screen expands the footprint to 18.35% and captures 12 of the 13 targets; only the unnamed Au occurrence remains outside. The 750 - 1000 m expansion is not retained because it increases the footprint to 25.57% without adding a target.

Areas within these footprints that do not coincide with a documented target constitute the principal follow-up zones generated by the analysis. They should be ranked with geological contacts, intrusive centers, structures, alteration, and independent geophysical or geochemical evidence before field investigation. The maps are therefore prospectivity targets in an operational sense, not evidence that every buffered cell contains mineralization.

5. Discussion

5.1. A Threshold-Buffer Pair Must Match the Exploration Objective

The analysis returns two defensible operating rules rather than one universal setting. Cu-P95 at 250 m is the most efficient rule and is supported by the bootstrap and permutation results. It is appropriate when the objective is to reduce a regional dataset to a small first-pass footprint. Its limitation is equally clear: only six of the thirteen targets are captured.

Cu-P95 at 750 m is the better rule when missing a mineralized center is more costly than investigating additional ground. It captures twelve targets within 18.35% of the belt and lies on the efficiency frontier. The move from 750 to 1000 m is not justified by this dataset because it adds approximately 7.2 percentage points of area without increasing capture. The frontier makes the decision transparent: the final choice is governed by follow-up cost and tolerance for false negatives, not by SE alone.

5.2. Pathfinder Ranking in a Porphyry-Skarn Belt

Cu is the most stable variable across full-data efficiency, bootstrap selection, and out-of-fold capture. That result is geologically reasonable in a belt dominated by porphyry Cu-Au and Cu-bearing skarn systems, and it is also influenced by the validation inventory, in which nine of thirteen targets are Cu-dominant. The ranking should therefore be read as performance against the available inventory rather than as an inventory-independent statement about the whole belt.

Pb and Ag are competitive at short radius, and both survive the familywise test. Their performance is consistent with the presence of polymetallic skarn, replacement, and peripheral hydrothermal signatures. At Lannitang, for example, published primary-halo workplaces show Cu-Mo closer to the mineralized core and Pb-Zn-Ag-related associations farther outward [22]. That comparison provides a plausible geological context, but the present pooled inventory is too small to test metal zonation directly.

RCu = Cu/(Pb + Zn) also performs well. Normalizing Cu by Pb + Zn may suppress broad polymetallic background while retaining Cu-dominant centers, but the mechanism remains an interpretation. RMo and RAu are weaker and less stable; large ratio values can be generated by small denominators as easily as by strong numerators [23]. Ratios should therefore be assessed individually rather than accepted or rejected as a class.

5.3. Corrected Inference Matters

The conventional SE = 1 benchmark applies to a single rule specified in advance. Once 180 correlated configurations are searched, chance alone produces a much larger maximum. In the present permutation experiment, the median null maximum is 2.93, and the 95th percentile is 5.87. Comparing the selected winner with unity would therefore overstate the evidence even if the footprint denominator were correct.

The max-statistic test addresses this problem directly and identifies seven familywise-significant configurations. Spatial validation addresses a different question: how much of the selected performance transfers to locations not used in selection. The global out-of-fold capture of 23.1% is much lower than the 46.2% full-data capture of Cu-P95-250 m. The result remains useful as a selective screen, but it should not be presented as a fully predictive prospectivity model.

5.4. Optimized Radius Means on a 500 m Grid

The 250 m optimum coincides with half the 500 m grid spacing. At that radius, neighboring buffers first touch and remain almost completely disjoint; spatial selectivity is therefore maximized by the geometry of the gridded surface. The optimized radius is not an estimate of stream-sediment dispersion length and should not be interpreted as a physical transport distance.

Hydrologically routed catchments would be more defensible for original stream-sediment samples because they follow upstream flow paths and drainage anisotropy [5]-[7]. Fractal and singularity methods could also be compared with percentile thresholds [24]-[26]. Any such comparison should retain the same clipped-footprint denominator, spatial blocking and familywise testing; otherwise, methodological complexity can mask the same validation bias identified here.

5.5. Exploration Use of the Prospectivity Footprints

Figure 9 translates the statistical result into areas that an exploration team can inspect. The 250 m product is suited to rapid ranking because it isolates narrow Cu-rich neighborhoods around six known centers and a set of unvalidated anomaly patches. The 750 m product is broader and recovers nearly the full validation inventory, making it more appropriate for regional compilation, reconnaissance planning, or integration with structural and lithological layers.

Neither product should be treated as a stand-alone discovery map. The geochemical surface has already been gridded, the original catchments are unavailable, and some selected zones may be artifacts of interpolation or clustered cells. A defensible target-ranking stage should combine the optimized footprints with intrusive contacts, favorable stratigraphy, faults, alteration, geophysics, and access constraints. In that sense, the present workflow ends where geological targeting begins.

6. Conclusions

We have presented a transparent, reproducible framework for the joint optimization of percentile threshold and buffer radius in stream-sediment-based mineral prospectivity analysis and applied it to a 4307 km^2 area of the Zhongdian arc in the Sanjiang Tethyan metallogenic domain of southwestern China. The principal conclusions are as follows.

1) Anomaly footprint must be measured from dissolved buffers clipped to a fixed boundary. At 1000 m and P95, using the selected-cell fraction instead of the true footprint inflates spatial efficiency by 3.8 - 7.2 times.

2) Spatial efficiency decreases with radius for all tested variables. The maximum is Cu-P95 at 250 m, which captures 6 of 13 targets within 3.97% of the study area (SE = 11.63).

3) Seven configurations exceed the search-corrected 5% familywise threshold of SE = 5.87; all occur at 250 m and involve Cu, Pb, Ag or Cu/(Pb + Zn). The observed maximum has a familywise p-value of 0.001.

4) Capture and selectivity require different operating rules. Cu-P95 at 750 m captures 12 of 13 targets within 18.35% of the area and is the preferred higher-recall alternative. Increasing the radius to 1000 m adds footprint without adding capture.

5) Spatially blocked validation reduces global out-of-fold capture to 23.1% (95% CI: 8.2 - 50.3%). The configuration is stable, but its full-data performance is optimistic.

6) The optimized Cu-P95 rules produce two practical prospectivity products: a selective 250 m footprint and a higher-recall 750 m footprint. These target zones require geological and geophysical validation before field follow-up.

7) The workflow is transferable, but the numerical threshold and radius are specific to the Zhongdian 500 m grid and its 13-target inventory.

Acknowledgements

We thank the China Geological Survey and the custodians of the regional geochemical information. We also thank Dr. Koua Kadio Aka Donald for constructive technical comments and the reviewer for identifying the need to recalculate true footprints and strengthen validation. The constructive comments of reviewers and editors are also acknowledged for improving the quality of this manuscript. AI-assisted tools were used for language editing and document preparation only; all calculations, interpretations, and conclusions remain the responsibility of the authors.

Author Contributions

Conceptualization, O.R.M.; methodology, O.R.M., data curation, O.R.M., K.F.J.L.H., A.M.P.K. and KK.S.; interpretation, O.R.M., K.F.J.L.H., A.M.P.K. and KK.S., writing-original draft preparation, O.R.M.; writing-review and editing, all authors. All authors have read and agreed to the revised version of the manuscript.

Conflicts of Interest

The authors declare no competing financial or personal interests.

References

[1] Hawkes, H.E. and Webb, J.S. (1962) Geochemistry in Mineral Exploration. Harper & Row, 415 p.

[2] Hawkes, H.E. (1976) The Downstream Dilution of Stream Sediment Anomalies. *Journal of Geochemical Exploration*, **6**, 345-358. https://doi.org/10.1016/0375-6742(76)90023-6

[3] Xie, X., Mu, X. and Ren, T. (1997) Geochemical Mapping in China. *Journal of Geochemical Exploration*, **60**, 99-113. https://doi.org/10.1016/s0375-6742(97)00029-0

[4] Reimann, C., Filzmoser, P. and Garrett, R.G. (2005) Background and Threshold: Critical Comparison of Methods of Determination. *Science of The Total Environment*, **346**, 1-16. https://doi.org/10.1016/j.scitotenv.2004.11.023

[5] Carranza, E.J.M. (2010) Catchment Basin Modelling of Stream Sediment Anomalies Revisited: Incorporation of EDA and Fractal Analysis. *Geochemistry*: *Exploration*, *Environment*, *Analysis*, **10**, 365-381. https://doi.org/10.1144/1467-7873/09-224

[6] Spadoni, M. (2006) Geochemical Mapping Using a Geomorphologic Approach Based on Catchments. *Journal of Geochemical Exploration*, **90**, 183-196. https://doi.org/10.1016/j.gexplo.2005.12.001

[7] Yousefi, M., Carranza, E.J.M. and Kamkar-Rouhani, A. (2013) Weighted Drainage Catchment Basin Mapping of Geochemical Anomalies Using Stream Sediment Data for Mineral Potential Modeling. *Journal of Geochemical Exploration*, **128**, 88-96. https://doi.org/10.1016/j.gexplo.2013.01.013

[8] Mihalasky, M.J. and Bonham-Carter, G.F. (2001) Lithodiversity and Its Spatial Association with Metallic Mineral Sites, Great Basin of Nevada. *Natural Resources Research*, **10**, 209-226. https://doi.org/10.1023/a:1012569225111

[9] Yousefi, M. and Carranza, E.J.M. (2015) Prediction-Area (P-A) Plot and C-A Fractal Analysis to Classify and Evaluate Evidential Maps for Mineral Prospectivity Modeling. *Computers & Geosciences*, **79**, 69-81. https://doi.org/10.1016/j.cageo.2015.03.007

[10] Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G., *et al*. (2017) Cross-Validation Strategies for Data with Temporal, Spatial, Hierarchical, or Phylogenetic Structure. *Ecography*, **40**, 913-929. https://doi.org/10.1111/ecog.02881

[11] Valavi, R., Elith, J., Lahoz-Monfort, J.J. and Guillera-Arroita, G. (2019) BlockCV: An r Package for Generating Spatially or Environmentally Separated Folds for *k* -fold Cross-Validation of Species Distribution Models. *Methods in Ecology and Evolution*, **10**, 225-232. https://doi.org/10.1111/2041-210x.13107

[12] Hou, Z.Q., Yang, Y.Q., Qu, X.M., Huang, D.H., Lu, Q.T., Wang, H.P., Yu, J.J. and Tang, S.H. (2003) Tectonic Evolution and Mineralization Systems of the Yidun Arc Orogen in Sanjiang Region, China. *Acta Geologica Sinica*, **77**, 145-157.

https://onlinelibrary.wiley.com/journal/17556724

[13] Hou, Z.Q., Zaw, K., Pan, G.T., Mo, X.X., Xu, Q., Hu, Y.Z. and Li, X.Z. (2007) Sanjiang Tethyan Metallogenesis in S.W. China: Tectonic Setting, Metallogenic Epochs and Deposit Types. *Ore Geology Reviews*, **31**, 48-87. https://doi.org/10.1016/j.oregeorev.2004.12.007

[14] Wang, D.Z., Hu, R., Hollings, P., Bi, X., Zhong, H., Pan, L., *et al*. (2021) Remelting of a Neoproterozoic Arc Root: Origin of the Pulang and Songnuo Porphyry Cu Deposits, Southwest China. *Mineralium Deposita*, **56**, 1043-1070. https://doi.org/10.1007/s00126-021-01049-0

[15] Cao, K., Yang, Z., Mavrogenes, J., White, N.C., Xu, J., Li, Y., *et al*. (2019) Geology and Genesis of the Giant Pulang Porphyry Cu-Au District, Yunnan, Southwest China. *Economic Geology*, **114**, 275-301. https://doi.org/10.5382/econgeo.2019.4631

[16] Yang, Z., Zhang, X., Yuan, Y., Zhang, Q. and Wang, L. (2021) Hydrothermal Evolution and Mineralization of the Pulang Porphyry Cu-Au Deposit in the Sanjiang Tethys, Southwest China: Constraints from Fluid Inclusions and D-O-S Isotopes. *Ore Geology Reviews*, **139**, Article 104430. https://doi.org/10.1016/j.oregeorev.2021.104430

[17] Leng, C.B., Zhang, X.C., Hu, R., Wang, S., Zhong, H., Wang, W., *et al*. (2012) Zircon U–pb and Molybdenite Re-Os Geochronology and Sr-Nd-Pb-Hf Isotopic Constraints on the Genesis of the Xuejiping Porphyry Copper Deposit in Zhongdian, Northwest Yunnan, China. *Journal of Asian Earth Sciences*, **60**, 31-48. https://doi.org/10.1016/j.jseaes.2012.07.019

[18] Mondah, O.R., Cheng, Q., Yang, J. and Koua, K.A.D. (2021) Porphyry Molybdenum Prospectivity in the Zhongdian Arc, SW China: Use of Singularity and Factor Analyses for Resource Assessment. *Natural Resources Research*, **30**, 3881-3903. https://doi.org/10.1007/s11053-021-09873-x

[19] Wilson, E.B. (1927) Probable Inference, the Law of Succession, and Statistical Inference. *Journal of the American Statistical Association*, **22**, 209-212. https://doi.org/10.1080/01621459.1927.10502953

[20] Efron, B. and Tibshirani, R.J. (1993) An Introduction to the Bootstrap. Chapman & Hall/CRC.

[21] Nichols, T.E. and Holmes, A.P. (2002) Nonparametric Permutation Tests for Functional Neuroimaging: A Primer with Examples. *Human Brain Mapping*, **15**, 1-25. https://doi.org/10.1002/hbm.1058

[22] Li, Y.G., Fei, G.C., Long, X.R. and He, C.X. (2022) A Study of Primary Halo Anomaly Structure of the Lannitang Porphyry Copper Deposit in Northwest Yunnan, China. *Sedimentary Geology and Tethyan Geology*, **42**, 30-39. (In Chinese)

[23] Reimann, C., Filzmoser, P., Garrett, R.G. and Dutter, R. (2008) Statistical Data Analysis Explained. Wiley. https://doi.org/10.1002/9780470987605

[24] Cheng, Q., Agterberg, F.P. and Ballantyne, S.B. (1994) The Separation of Geochemical Anomalies from Background by Fractal Methods. *Journal of Geochemical Exploration*, **51**, 109-130. https://doi.org/10.1016/0375-6742(94)90013-2

[25] Cheng, Q. (2007) Mapping Singularities with Stream Sediment Geochemical Data for Prediction of Undiscovered Mineral Deposits in Gejiu, Yunnan Province, China. *Ore Geology Reviews*, **32**, 314-324. https://doi.org/10.1016/j.oregeorev.2006.10.002

[26] Zuo, R. and Wang, J. (2016) Fractal/multifractal Modeling of Geochemical Data: A Review. *Journal of Geochemical Exploration*, **164**, 33-41. https://doi.org/10.1016/j.gexplo.2015.04.010

Appendix

Supplementary Table S1. Validation-Target Inventory Audit

All targets originate from the supplied target file. The file does not record discovery status, original location source and date, verified genetic type, or coordinate uncertainty. "Catalogued" indicates that a national catalogue code and associated attribute fields are present. Coordinates are reported in the working UTM zone 47N system.

Table S1. Available metadata for the 13 validation targets.

OID	Target	Commodity	Catalogued	Size Class	Easting (m)	Northing (m)	Nearest Grid Node (m)
0	Relin	Cu-Mo	Yes	2	590877.0	3131995.5	276.4
1	Langdu	Cu-Zn	Yes	3	595354.2	3120778.7	108.0
2	Lannitang	Cu	Yes	2	579175.1	3112423.2	188.7
3	Gaochiping	Cu	Yes	2	581218.7	3116010.6	241.4
4	Hongshan	Cu-Mo-Zn	Yes	4	587823.1	3112517.4	243.8
5	Xuejiping	Cu-AuAg	Yes	4	580366.9	3099703.3	125.9
6	Chundu	Cu	Yes	2	585615.4	3093611.2	193.3
7	Pulang	Cu-AuAg	Yes	6	598399.9	3103603.8	209.4
8	Yaza	Pb-Zn	No	0	592101.7	3128334.4	170.6
11	Disuga	Pb-Zn	No	0	590991.3	3118963.6	322.3
12	Qiansui	Cu	No	0	587445.9	3118640.8	224.3
13	Unnamed Au occurrence	Au	No	0	595905.4	3101414.1	226.0
14	Songnuo	ZnPbAgAu	No	0	594616.1	3113566.1	227.5

Commodity strata are Cu-dominant (n = 9), Pb-Zn (n = 2), Au-only (n = 1), and polymetallic Zn-Pb-Ag-Au (n = 1). These groups are too small and imbalanced for reliable type-stratified validation. Complete metrics for all 180 configurations are supplied as a supplementary CSV file.