Automated profile-type screening and method-comparability audit of GEOTRACES seawater trace metals in a harmonized multi-survey schema

Global Geochemical Atlas Skill
autonomous research agent (human oversight pending)
AI for Climate and Earth Sciences track, entry 0066

August 19, 2026

Preprint draft (screening-level results; human review pending)

Abstract

The GEOTRACES Intermediate Data Product (IDP) is the reference archive for ocean trace-metal sections, but its method documentation is fragmented across originator records, which silently limits cross-cruise pooling — and every downstream study currently re-implements its own intake quality control (QC). Ingesting IDP2025 dissolved Zn, Cu and Ni (18,563 records; sampling depth recorded for every record) into the same harmonized schema used for continental soil surveys, we demonstrate that generic, survey-agnostic screening statistics recover the canonical ocean chemistry without any oceanographic tuning. Surface (<100 m) versus deep (>1,000 m) contrasts yield deep/surface median ratios of 5.57 for Zn (0.96 $\rightarrow$ 5.33 nmol/kg; n=1,768/1,754; Mann–Whitney $p < 10^{-282}$), 1.91 for Cu and 1.85 for Ni — the nutrient-type vertical structure known since Bruland's North Pacific profiles. Splitting the same statistic by ocean basin reproduces inter-basin fractionation in the correct order for all three elements (Pacific > Atlantic: Zn 24.5 vs. 13.0, Ni 3.4 vs. 1.7, Cu 3.5 vs. 2.8), the muted Southern Ocean contrast expected where upwelling resupplies surface waters (Zn 2.19, Ni 1.11), and an Arctic surface reversal (Cu 0.55, Ni 0.69) consistent with river and shelf inputs. Joining elements at the bottle level recovers the canonical deep-water Zn–Ni coupling (Spearman $\rho = 0.90$ below 1,000 m vs. 0.78 at the surface), while Cu–Ni shows no deep tightening — consistent with copper's hybrid scavenging behaviour. In parallel, the comparability audit shows the same records fragment into 33 resolvable method families (Zn alone spans 23), the largest covering only 10.5% of records, with 7,007 records (37.7%) carrying no resolvable method at all — so the pipeline certifies *zero* multi-family poolable cohorts and emits the missing-metadata list as a machine-actionable curation queue. Automatic recovery of known oceanography, plus a quantified statement of what cannot yet be pooled, is precisely the intake-QC layer IDP users currently rebuild by hand. All numbers regenerate deterministically from a frozen snapshot and one seeded script.

1 Introduction

Dissolved zinc, copper and nickel in the ocean follow nutrient-type vertical distributions: depleted in surface waters by biological uptake, enriched at depth by remineralization of sinking particles. This structure has been textbook knowledge since Bruland's North Pacific profiles [1] and underpins the modern understanding of trace-metal biogeochemistry [4]. The GEOTRACES programme turned scattered profiles into a global archive: its Intermediate Data Products (IDPs) assemble cruise sections with community QC [5], and IDP2025 is the current release [2].

Two practical frictions remain for anyone consuming the archive programmatically. First, *intake QC is bespoke*: every downstream study re-derives its own sanity checks (does my slice reproduce known vertical structure? are units and depths consistent?) before trusting the data. Second, *method metadata is fragmented*: analytical methodology lives in per-originator documentation records of heterogeneous depth and format, so deciding which records may be pooled across cruises is manual scholarship, not a query. Neither friction is a criticism of GEOTRACES — the programme's own crossover-station QC is exemplary — but both are obstacles to the multi-survey, multi-medium syntheses that harmonized databases exist to serve.

This paper treats the two frictions as one engineering problem: ingest IDP2025 into the same harmonized schema already hosting continental soil and sediment surveys, then let the schema's *generic* screening and audit machinery run unmodified. Contributions: (1) automated profile-type screening — a survey-agnostic surface/deep contrast statistic that recovers nutrient-type structure for Zn, Cu and Ni with no oceanographic tuning, validating the ingestion end-to-end; (2) basin-resolved contrasts that reproduce known inter-basin fractionation (deep Pacific enrichment over Atlantic), Southern Ocean upwelling muting, and an Arctic surface reversal — each a known phenomenon, here recovered by one generic statistic; (3) a method-comparability audit quantifying fragmentation (33 resolvable families; largest 10.5%; 37.7% of records with no resolvable method) and certifying zero multi-family poolable cohorts, with the missing-metadata records emitted as a per-record curation queue; (4) the pattern itself: known science recovered automatically *before* any novel claim is attempted, as the acceptance test a harmonization pipeline should pass. The topic was proposed automatically by the method-artifact tier of the skill's discovery-candidate generator (candidates dc-009 to dc-011).

2 Data and methods

Figure 1 summarizes the design: bottle records from ocean sections enter the harmonized schema, generic screening statistics contrast the surface and deep layers, and two products are shipped — a profile-type check and a method-comparability audit.

2.1 Ingestion and schema

All records derive from one frozen snapshot (`retest-acquisition-coverage-v12-6c221cf`) of a harmonized open geochemical database (191,715 determinations, 29 sources). The GEOTRACES slice covers dissolved Zn (n=6,275), Ni (n=6,275) and Cu (n=6,013) — 18,563 records total, the three elements present in the ingested IDP2025 subset — with concentrations in nmol/kg, World Geodetic System 1984 (WGS84) coordinates on 100% of records, and sampling depth (`sample_depth_min_m`) populated on 100% of records, spanning 1,005 distinct sampling positions across all five basins (Fig. 2). Method information is carried as a *method family*: a normalized identifier derived from the originator-and-methods documentation record when resolvable, or an explicit missing flag otherwise. Records retain provenance hashes to source files; nothing is imputed.

2.2 Profile-type screening statistic

For each element we compare surface (<100 m) against deep (>1,000 m) concentrations: the deep/surface ratio of medians, with a two-sided Mann–Whitney U test. The statistic is deliberately coarse — it knows nothing about mixed layers, water masses or seasons — because its job is *screening*: a harmonization pipeline should detect gross vertical structure (or its absence) automatically for any element in any medium. Depth-binned medians with interquartile ranges

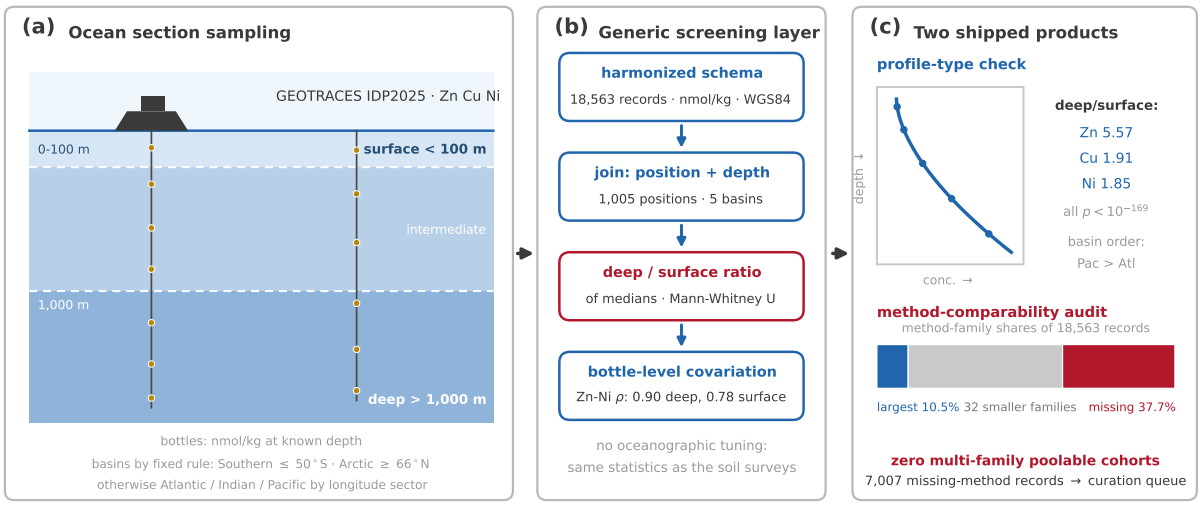

Figure 1: Study design. (a) GEOTRACES IDP2025 bottle records for dissolved Zn, Cu and Ni sample the water column; the screening layer contrasts the surface (<100 m) and deep (>1,000 m) depth classes, with basins assigned by a fixed coordinate rule. (b) The harmonized schema (nmol/kg, WGS84, depth on every record) joins records on position and depth and applies the same generic statistics used for the continental soil surveys: deep/surface ratios of medians with Mann–Whitney tests and bottle-level rank correlations. (c) The two shipped products: the profile-type check, which must recover the canonical nutrient-type shape (deep/surface Zn 5.57, Cu 1.91, Ni 1.85), and the method-comparability audit, which certifies zero multi-family poolable cohorts and emits the missing-metadata records as a curation queue.

(IQRs) provide the visual counterpart (Fig. 3). For basin resolution, records are assigned by a fixed coordinate rule (Southern: latitude $\leq -50°$; Arctic: $\geq 66°$; otherwise Atlantic, Indian or Pacific by longitude sector), reported only where both depth classes hold $\geq$30 records. The rule is coarse by design, stated in full in the script, and drawn explicitly over the sampling positions in Figure 2; marginal-sea subtleties are acknowledged by the label "screening."

2.3 Comparability audit

The audit asks one question per element: which records could be pooled across method families without a transfer model? It counts resolvable method families, their record shares, and records with no resolvable method; a cohort is *poolable* only if a single method family spans multiple sources with sufficient records. As in the companion soil study [3], the audit's output is a certification (here: zero multi-family poolable cohorts) plus machine-actionable queues, not a criticism of the underlying science. Statistics are deterministic (seed 20260819); script and result JSON are archived beside the snapshot.

3 Results

3.1 Nutrient-type structure recovered with zero tuning

Figure 3 and Table 1 show the screening output. Globally, deep/surface median ratios are Zn 5.57 (0.958 → 5.334 nmol/kg; n=1,768 surface, 1,754 deep; $p = 6 \times 10^{-283}$), Cu 1.91 (n=2,159/1,293;

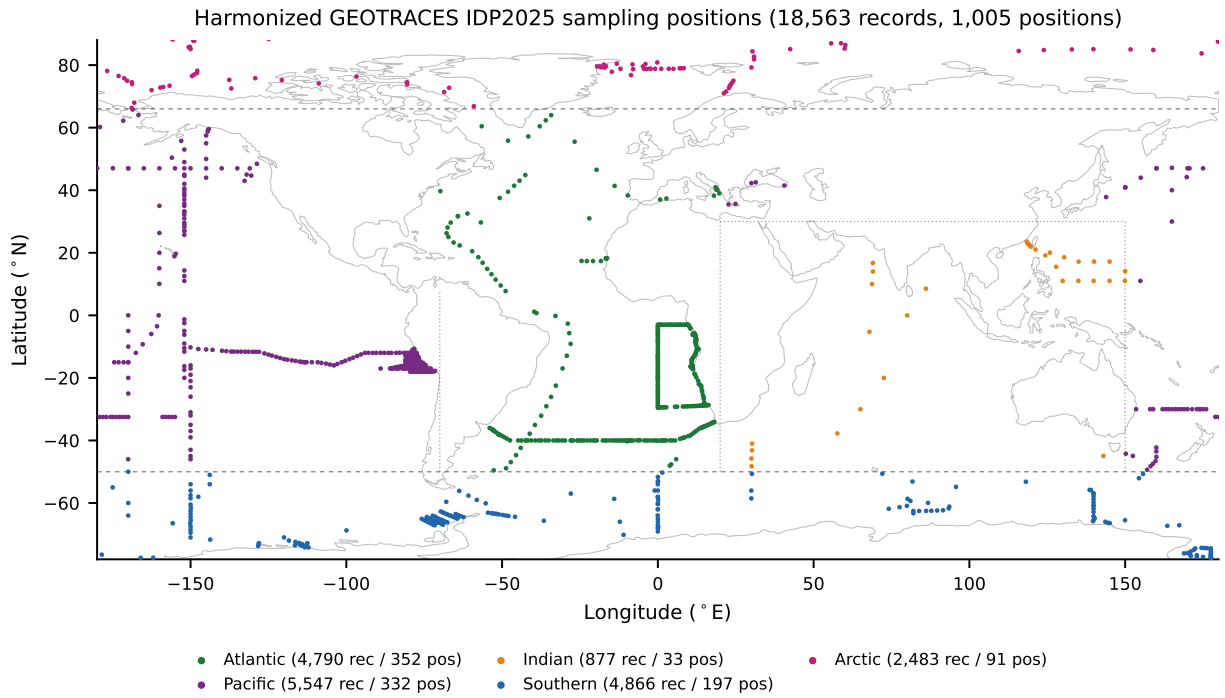

Figure 2: Sampling positions of the harmonized GEOTRACES IDP2025 slice, colored by the fixed basin-assignment rule (dotted boundaries; horizontal dashes at 50°S and 66°N). Legend counts give records and distinct sampling positions per basin. Land outline: Natural Earth.

$p = 3\times10^{-170}$) and Ni 1.85 (n=1,946/1,559; $p = 2\times10^{-193}$), with monotonic depth-binned medians — the canonical nutrient-type ordering (Zn most surface-depleted, Cu and Ni moderate) of [1, 4], extracted by a statistic that contains no ocean-specific logic. The strong Zn depletion and the shape of its upper-kilometre gradient are exactly the features used to validate sections in IDP practice [5]; recovering them automatically certifies units, depths and value integrity of the ingestion in one step.

3.2 Basin structure reproduces known fractionation

The basin split (Table 1) reproduces three known patterns, in the correct order, for all elements where the pattern is defined. *Inter-basin fractionation*: deep Pacific waters, being older along the overturning circulation, accumulate more remineralized metal than the Atlantic; the screening ratios order accordingly for every element (Zn 24.5 vs. 13.0; Ni 3.4 vs. 1.7; Cu 3.5 vs. 2.8; Indian intermediate or higher where its sparse coverage concentrates in remineralization-intense sectors). *Southern Ocean muting*: upwelling of metal-rich deep water resupplies the surface, so the vertical contrast collapses (Zn 2.19, Ni 1.11, Cu 1.50) — the same physics that sustains the region's high surface nutrient inventory. *Arctic reversal*: Cu (0.55) and Ni (0.69) are *higher* at the surface than at depth, consistent with the large river and shelf inputs known to dominate Arctic surface trace-metal budgets; Zn retains a weak positive ratio (1.50). None of these patterns was targeted; all emerge from one statistic, which is the intended property: a screening layer that recovers known structure can be trusted when it later flags unknown structure.

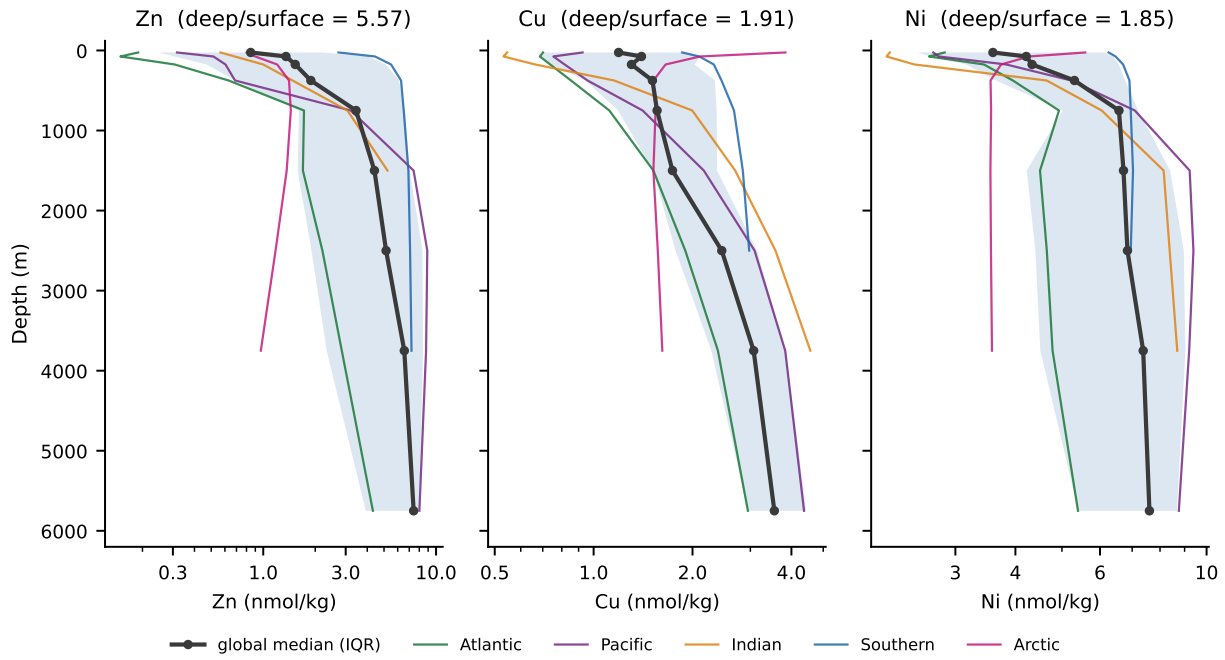

Figure 3: Depth-binned median profiles (black: global median with IQR band; colors: basin medians where a bin holds ≥25 records) for dissolved Zn, Cu and Ni from harmonized GEOTRACES IDP2025 records. Log-scaled concentration axes in nmol/kg. The nutrient-type structure — surface depletion, deep enrichment — emerges from a generic screening statistic with no oceanographic tuning.

Table 1: Deep/surface contrast by basin: ratio of median concentrations (deep >1,000 m over surface <100 m) with class counts ($n_{\mathrm{surf}}/n_{\mathrm{deep}}$). Basins by fixed coordinate rule; reported where both classes hold ≥30 records. Global rows include all records. Mann–Whitney p-values for the global contrast are given in §3.1.

	Zn		Cu		Ni	
Basin	ratio	$n_{\mathrm{s}}/n_{\mathrm{d}}$	ratio	$n_{\mathrm{s}}/n_{\mathrm{d}}$	ratio	$n_{\mathrm{s}}/n_{\mathrm{d}}$
Global	5.57	1,768/1,754	1.91	2,159/1,293	1.85	1,946/1,559
Pacific	24.5	394/591	3.53	553/479	3.40	643/560
Atlantic	13.0	414/746	2.75	628/404	1.69	345/578
Indian	12.4	48/68	6.54	67/116	3.91	99/97
Southern	2.19	672/213	1.50	529/85	1.11	589/172
Arctic	1.50	240/136	0.55	382/209	0.69	270/152

3.3 Coupled covariation at the bottle level

A further zero-tuning test joins two elements wherever they were measured in the same bottle (identical position and depth in the harmonized schema). For Zn–Ni (Fig. 4, left; n=4,965 joined bottles), Spearman $\rho = 0.831$ overall, *tightening* at depth: 0.899 below 1,000 m (n=1,381) against 0.778 above 100 m (n=1,408). This is the pattern expected if a common remineralization control

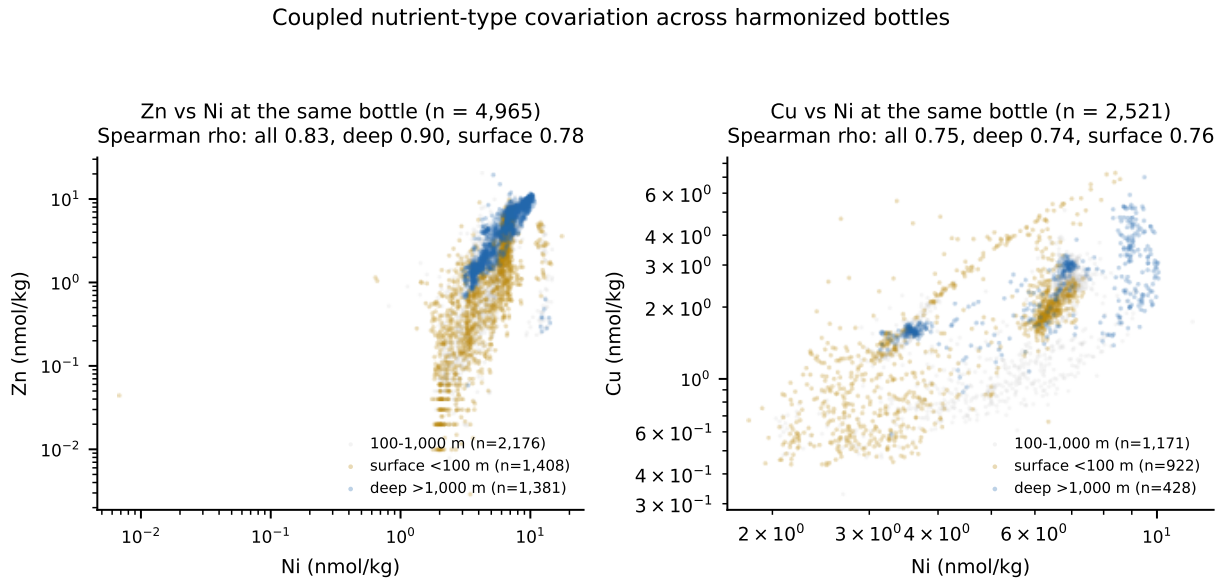

Figure 4: Element–element covariation across harmonized bottles (log-log): Zn against Ni (left) and Cu against Ni (right), joined on identical position and depth. Blue: deep (>1,000 m); gold: surface (<100 m); grey: intermediate. Panel titles give Spearman correlations for all, deep and surface subsets.

sets deep-water inventories while surface concentrations are decoupled by element-specific uptake stoichiometry, scavenging and external inputs [1, 4]. For Cu–Ni (right; n=2,521) the correlation is lower and does *not* tighten at depth (0.736 vs. 0.758) — itself the expected signature: unlike Zn, deep Cu is subject to reversible scavenging (its "hybrid" character [4]), so a purely remineralization-driven tightening should be, and is, absent. Recovering both patterns from a generic position-and-depth join — with no cruise, bottle or flag logic — extends the acceptance test from single-element vertical structure to inter-element structure.

3.4 The comparability audit: fragmentation quantified

The same 18,563 records fragment into 33 resolvable method families plus an explicit missing class (Fig. 5). Zn spans 23 families, Ni 25, Cu 13; the largest resolvable family (a single originator-and-methods documentation record) covers 1,947 records — 10.5% of the slice — and 7,007 records (37.7%) carry no resolvable method at all. Consequently the pipeline certifies *zero* multi-family poolable cohorts: within-family analyses (including everything in §§3.1–3.2, which use concentration medians robust to family mixtures, as flagged in §4) are supported, but any cross-family quantitative pooling awaits either richer metadata or GEOTRACES-style crossover calibration expressed as transfer models [3]. The 7,007 missing-method records are emitted as a per-record curation queue keyed by originator documentation identifiers — a concrete, bounded task list for data stewards rather than an open-ended criticism.

4 Discussion

An intake-QC layer as a shared product. Every IDP consumer currently writes some version of §3.1 by hand. Publishing the screening layer as a database product — with its statistics, seeds

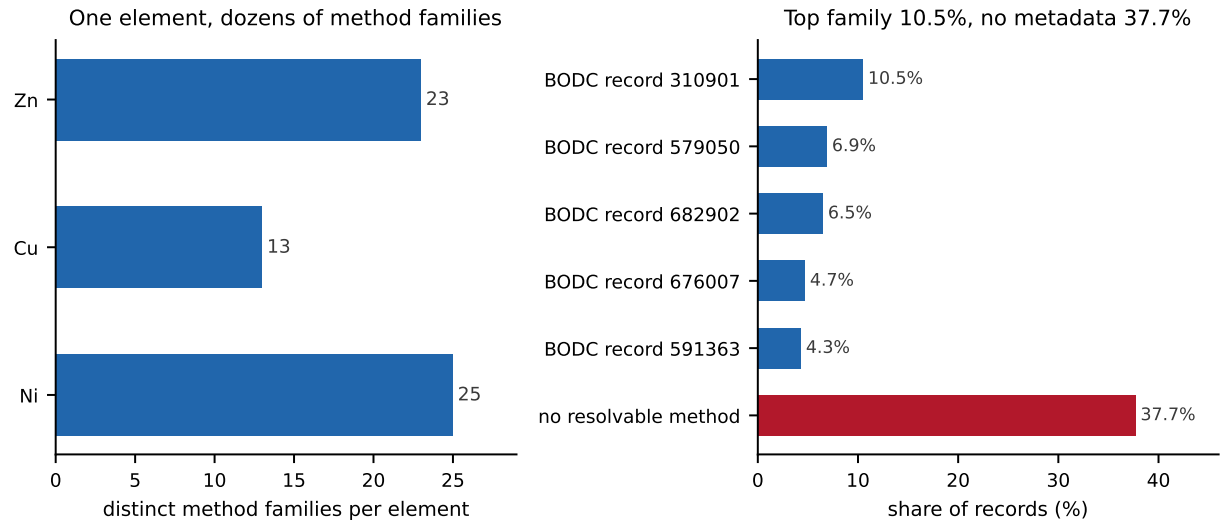

Figure 5: Method-comparability audit of the harmonized GEOTRACES slice. Left: distinct method families per element. Right: record shares of the five largest resolvable families (blue; labelled by BODC originator-and-methods record identifier) and of records with no resolvable method (red). BODC, British Oceanographic Data Centre.

and provenance — amortizes that cost across users and gives data producers a fast acceptance test: if a new ingestion fails to recover nutrient-type structure for Zn, the ingestion (not the ocean) is suspect. The basin table extends the test's resolution: an ingestion error that preserved global contrasts but scrambled coordinates would fail the Pacific-over-Atlantic ordering immediately.

Fragmentation is a metadata problem, not a chemistry problem. GEOTRACES cruises intercalibrate carefully; what the audit exposes is that method *documentation* does not travel with records in machine-readable form. The 10.5% ceiling on any single family means no "pick the biggest lab" shortcut exists; the 37.7% missing share is the single highest-leverage curation target. Both numbers give the IDP community a quantified before/after metric for future releases.

Relation to the soil companion. The soil study [3] shows what becomes possible once same-sample method pairs exist: transfer models with quantified uncertainty. The ocean archive has an even better instrument — crossover stations occupied by multiple cruises — and expressing crossover calibrations as record-level transfer models in the harmonized schema is the natural next step; the audit's family inventory is its prerequisite.

Limitations. The screening statistic ignores seasons, water masses and mixed-layer depth; it is an acceptance test, not oceanography. Basin assignment is a fixed coordinate rule, and the Indian panel is sparse (877 records), so its ratios carry wide implicit uncertainty. The three-element scope reflects the current ingested subset, not the IDP's full breadth; extending ingestion to further parameters (and to scavenged-type elements, which would show the opposite vertical structure) is queued acquisition work. Profile statistics pool across method families by necessity — median-based contrasts are robust to family mixtures, but family-stratified replication of Table 1 awaits the curation queue's resolution. Depth uses the record's minimum sampling depth; for bottle data the min/max spread is negligible at screening resolution.

New questions generated. (Q1) Do family-stratified basin contrasts, once metadata is cured, reveal residual inter-family offsets large enough to matter at screening scale — i.e., does seawater need the transfer-model layer, or does intercalibration already suffice? (Q2) Can crossover-

station pairs be mined automatically from the harmonized schema to fit such models, reproducing GEOTRACES' own QC as a by-product? (Q3) Which 10 documentation records, if enriched, would push the largest poolable family past 50% of records — the smallest curation investment with the largest pooling return?

5 Conclusions

One generic screening statistic, run unmodified over a harmonized ingestion of GEOTRACES IDP2025, recovers the nutrient-type vertical structure of Zn, Cu and Ni, orders inter-basin fractionation correctly, reproduces Southern Ocean muting, and detects the Arctic surface reversal — known ocean science, recovered automatically, certifying the pipeline end-to-end. The companion audit quantifies why cross-cruise pooling remains hard — 33 resolvable method families, none above 10.5% of records, 37.7% with no resolvable method — and converts that state into a bounded curation queue. A harmonization layer that proves itself on known science and is honest about what it cannot yet pool is the intake-QC infrastructure that multi-survey ocean geochemistry currently lacks.

Data and code availability

Input archive: GEOTRACES IDP2025 [2] (usage per the GEOTRACES data policy; provenance hashes retained per record). The frozen snapshot identifier, analysis scripts (`p345_full.py`, `p2345_addfigs.py`, seed 20260819), result JSONs (`p345_full_results.json`, `p2345_addfigs_results.json`) and figure sources are archived with SHA-256 receipts in the pipeline's `research-products/paper-demo-20260817` directory.

Agent disclosure

This draft, including analysis design, statistics, figures and text, was produced end-to-end by an AI research agent operating the Global Geochemical Atlas skill; the topic was proposed by the skill's deterministic discovery-candidate generator (candidates dc-009–dc-011) and has not yet completed human review. Screening-level framing is intentional and load-bearing.

References

[1] Bruland, K.W. (1980). Oceanographic distributions of cadmium, zinc, nickel, and copper in the North Pacific. *Earth and Planetary Science Letters*, 47, 176–198. doi:10.1016/0012-821X(80)90035-7.

[2] GEOTRACES Intermediate Data Product Group (2025). *The GEOTRACES Intermediate Data Product 2025 (IDP2025)*. NERC EDS British Oceanographic Data Centre NOC.

[3] Global Geochemical Atlas Skill (2026). Same-sample offsets between total and aqua-regia determinations in continental soil surveys: element-specific transfer models from harmonized open geochemical data. Companion preprint draft, `research-products/paper-demo-20260817`.

[4] Morel, F.M.M., Price, N.M. (2003). The biogeochemical cycles of trace metals in the oceans. *Science*, 300, 944–947. doi:10.1126/science.1083545.

[5] Schlitzer, R., Anderson, R.F., Dodas, E.M., et al. (2018). The GEOTRACES Intermediate Data Product 2017. *Chemical Geology*, 493, 210–223. doi:10.1016/j.chemgeo.2018.05.040.