Same-sample offsets between total and aqua-regia determinations in continental soil surveys: element-specific transfer models from harmonized open geochemical data

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

Continental soil surveys routinely determine the same element on the same physical sample by two analytical routes — typically X-ray fluorescence (XRF, total content) and inductively coupled plasma mass spectrometry after aqua-regia digestion (ICP-MS, partial extraction) — yet downstream users frequently pool such numbers as if they described one population. Using a harmonized 191,715-record open geochemical database, we quantify same-sample offsets across the full GEMAS (Geochemical Mapping of Agricultural and Grazing Land Soil) survey, with n=2,224 paired samples for each of six elements, and replicate the design independently on FOREGS (Forum of European Geological Surveys) soils (n=417–420; XRF vs. ICP-OES, inductively coupled plasma optical emission spectrometry). Median XRF/ICP ratios are element-specific and large: Cr 2.94 (95% confidence interval, CI, 2.89–2.99), Zn 1.34, Pb 1.29, Ni 1.27, As 1.23 — and direction-reversed for Cu (0.86), which rules out any single instrument-bias explanation. FOREGS reproduces the Cr offset at 2.54 (topsoil) and 2.50 (subsoil). Log-log regression slopes of 0.74–1.00 (r^2 0.65–0.93) show the offsets are concentration-dependent, so a single correction factor is insufficient in general; we fit per-element power-law transfer functions and validate them out-of-sample with a seeded 80/20 split. The power law reduces the median absolute prediction error for Pb from 18.2% (constant-factor model) to 11.0%, while for elements whose slope is statistically indistinguishable from one (Cu) the constant factor is equivalent, exactly as theory requires. The practical outputs are (i) a transfer-model table that converts the database's certified "zero multi-source poolable cohorts" state into conditionally poolable data with quantified uncertainty, and (ii) a minimum method-metadata standard for future surveys. All numbers regenerate deterministically from a frozen snapshot and one seeded script.

1 Introduction

Legally mandated soil monitoring is converging on cross-border comparability as a first-class requirement: the European Soil Monitoring Law obliges member states to produce comparable soil-health indicators [3], and the Minamata Convention's open-ended scientific group notes that existing soil survey data "cannot be compared across surveys" without explicit harmonization of analytical protocols [7]. The largest obstacle is not units or coordinates — those are mechanical — but analytical method: the same element in the same sample yields a different number depending on whether the determination is total (XRF, fusion methods) or based on a partial extraction such as aqua regia [5, 2].

The magnitude of this method effect is qualitatively well known to survey geochemists. The GEMAS project documented aqua-regia versus total discrepancies in its atlas volumes [5]; the FOREGS atlas likewise reports both total and extractable determinations side by side [6, 2]. What downstream users lack is a *quantitative, per-element, uncertainty-carrying* description of the offset that can be applied programmatically when deciding whether two data sets may be combined — precisely the layer a harmonized multi-survey database needs before it can defensibly pool records from different analytical traditions.

This paper supplies that layer for six elements (As, Cr, Cu, Ni, Pb, Zn) using the one feature of survey design that makes the question answerable without new laboratory work: *both methods were applied to the same physical samples.* Our contributions are: (1) the full same-sample offset matrix on GEMAS agricultural and grazing soils (n=2,224 pairs per element), with bootstrap CIs; (2) evidence that offsets are concentration-dependent (log-log slopes $\neq 1$) and an out-of-sample test quantifying when a power-law transfer model beats a constant factor; (3) an independent replication on FOREGS topsoil and subsoil with a different ICP instrument family; (4) a transfer-model table and a minimum method-metadata standard, designed to be consumed by harmonization pipelines rather than by humans.

The analysis was proposed automatically by the discovery-candidate generator of an autonomous geochemical data skill (candidates dc-012 to dc-016, method-artifact tier) after the same skill's comparability audit certified *zero* multi-source poolable cohorts in its 191,715-record snapshot; this paper is the constructive answer to that finding.

2 Materials and methods

Figure 1 summarizes the design: one physical sample follows two analytical routes, same-sample pairs quantify the offset per element, and a power-law transfer model is validated out of sample.

2.1 Data and pair construction

All records derive from one frozen snapshot (`retest-acquisition-coverage-v12-6c221cf`) of a harmonized open geochemical database (191,715 determinations, 29 sources), in which every record carries normalized concentration (mg/kg), World Geodetic System 1984 (WGS84) coordinates, a method family assignment, and a provenance hash chain to the source file. We use two surveys in which two method families coexist on the same samples:

GEMAS [5]: agricultural (Ap) and grazing-land (Gr) soils sampled 2008–2009 across Europe. Method families in the snapshot: `icp_ms` (aqua-regia digestion, following the GEMAS protocol based on ISO 11466 [4]) and `xrf` (total). For each element and each `sample_id` we take the median of replicate determinations per method, then join the two methods on `sample_id`; this yields exactly n=2,224 same-sample pairs for each of As, Cr, Cu, Ni, Pb, Zn. Mercury is determined by ICP-MS only and therefore has no same-sample counterpart; it is excluded rather than approximated.

FOREGS [6, 2]: European topsoil and subsoil sampled 1997–2001. Method families: `icp_oes` (partial digestion) and `xrf` (total), joined the same way (n=420 topsoil, 417 subsoil pairs for Cr and Zn; other elements lack a same-sample double determination in the harmonized slice and are reported only where pairs exist).

2.2 Statistics

For each survey–element pair we report the median XRF/ICP ratio with a 95% CI from 2,000 seeded bootstrap resamples, the interquartile range (IQR) of the per-sample ratios, and an ordinary least

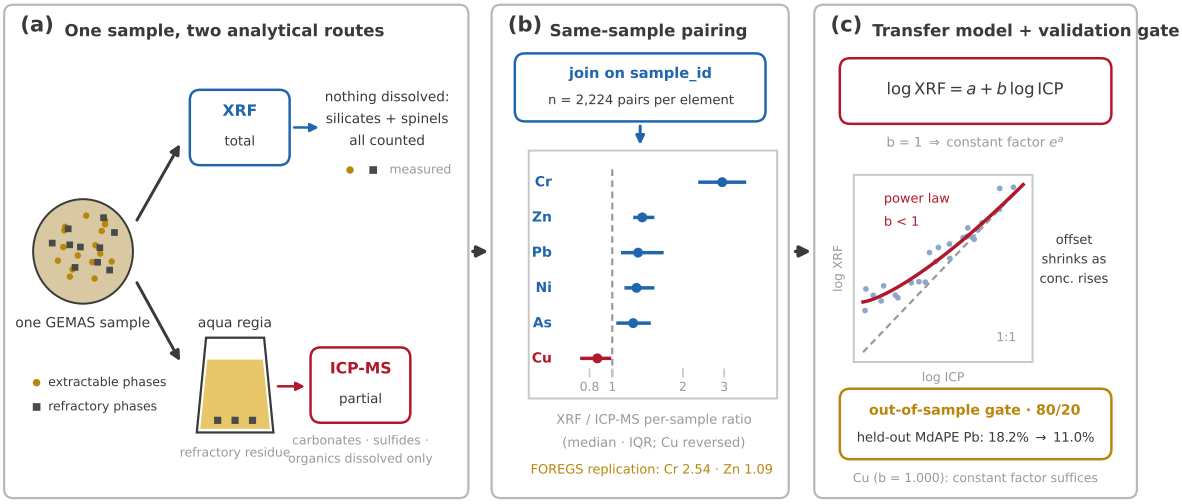

Figure 1: Study design. (a) Each GEMAS sample is determined by two routes: XRF measures total content, refractory silicates and spinels included, while aqua-regia digestion dissolves carbonate-, sulfide- and organically bound phases and leaves a refractory residue, so the subsequent ICP-MS determination is operationally partial. (b) Joining the two determinations on `sample_id` yields n=2,224 same-sample pairs per element; the per-sample XRF/ICP-MS ratio matrix (medians with IQRs) is element-specific and direction-reversed for Cu, with an independent FOREGS XRF/ICP-OES replication. (c) The transfer layer: a log-log power law whose slope b captures the concentration dependence of the offset, gated by a seeded 80/20 out-of-sample comparison against the constant-factor alternative.

squares regression

$$\log \mathrm{XRF} = a + b \log \mathrm{ICP}, \tag{1}$$

whose slope b measures concentration dependence: b=1 reduces Eq. 1 to a constant multiplicative factor e^a, while $b \neq 1$ makes the offset vary with concentration. Model adequacy is tested out-of-sample: a seeded 80/20 train/test split (random number generator seed 20260819), fitting both the power law and the constant-factor model (train-median ratio) on the training set and scoring both by the median absolute percentage error (MdAPE) of XRF predictions on the held-out set. All computation is deterministic; the analysis script and result JSON are archived beside the snapshot.

2.3 Scope of interpretation

Aqua-regia and other partial digestions are *operationally defined*: they dissolve carbonate-, sulfide- and organically bound phases efficiently but leave refractory silicates and spinels largely intact [5, 2]. The offsets quantified here are therefore digestion-and-matrix effects, not instrument defects, and the paper makes no claim about which method is "correct": totals and extractables answer different geochemical questions. Our subject is strictly the *transfer* between them.

Table 1: GEMAS same-sample transfer matrix: XRF (total) vs. ICP-MS (aqua regia), n=2,224 pairs per element. Ratio statistics are of per-sample XRF/ICP-MS ratios; a, b and r^2 refer to the power-law fit of Eq. 1; SE, standard error. All values from the frozen snapshot, seed 20260819.

Element	Median ratio	95% CI	IQR	b (SE)	a	r^2
Cr	2.942	2.893–2.993	2.366–3.660	0.849 (0.009)	1.535	0.805
Zn	1.342	1.335–1.348	1.250–1.486	0.837 (0.008)	0.917	0.833
Pb	1.288	1.269–1.308	1.108–1.628	0.735 (0.012)	1.030	0.647
Ni	1.269	1.256–1.283	1.146–1.486	0.909 (0.006)	0.529	0.925
As	1.228	1.215–1.242	1.058–1.435	0.830 (0.008)	0.495	0.815
Cu	0.864	0.856–0.872	0.742–0.973	1.000 (0.008)	−0.174	0.863

3 Results

3.1 The six-element offset matrix

Table 1 and Figures 2–3 summarize the GEMAS matrix. Median XRF/ICP-MS ratios span 0.864 (Cu) to 2.942 (Cr); every CI excludes unity, and the element ordering Cr $\gg$ Zn $>$ Pb $\approx$ Ni $>$ As $>$ 1 $>$ Cu is stable across the IQR. The Cr offset — total chromium nearly three times the aqua-regia value at the median — is consistent with the known resistance of chromite and other spinels to aqua regia [2]. The Cu reversal (total *below* the extractable determination at the median) rules out any explanation that acts uniformly across elements, such as a calibration scale error in either instrument; element-specific phase partitioning and detection-limit behaviour at low concentrations (visible as horizontal banding in Fig. 2) are the plausible drivers, and their resolution is flagged as follow-up laboratory work rather than asserted.

Figure 4 recasts the same pairs in the Bland–Altman (mean–difference) form standard in method-comparison studies [1]: per-sample $\log_2$ ratios against the geometric mean of the two determinations, with the median offset and nonparametric 95% limits of agreement (LOA). The LOA widths give the per-sample uncertainty of any transfer: for Cr the central 95% of same-sample ratios span 1.49–6.86 — even the 2.5th percentile sits half again above unity — while Zn is comparatively tight (1.06–2.22) and Cu straddles unity (0.37–1.70). The tilted running medians anticipate the concentration dependence quantified next.

3.2 Offsets are concentration-dependent, and when that matters

Slopes b in Table 1 range from 0.735 (Pb) to 1.000 (Cu): for five of six elements the offset shrinks as concentration rises, consistent with a roughly constant refractory or non-extractable component that matters proportionally more in low-concentration samples. The out-of-sample comparison (Table 2) turns this observation into a model-selection rule. For Pb — the element farthest from b=1 — the power law cuts MdAPE from 18.2% to 11.0%, a 40% error reduction; for Zn the reduction is 17%; for Cr, Ni and As the two models are within one percentage point; and for Cu, whose slope is exactly 1.000, the constant factor is marginally *better* (13.5% vs. 14.1%), as expected when the extra parameter fits noise. The practical rule accompanying the transfer table is therefore: use the power law when $|b-1|$ exceeds its standard error severalfold, otherwise the constant factor suffices. Figure 5 shows the comparison directly: on the held-out Pb set the constant-factor predictions fan away from the 1:1 line at both concentration extremes — exactly where a b=0.735 slope must defeat a single multiplier — while the power-law predictions stay centred; the per-element bars restate Table 2 graphically.

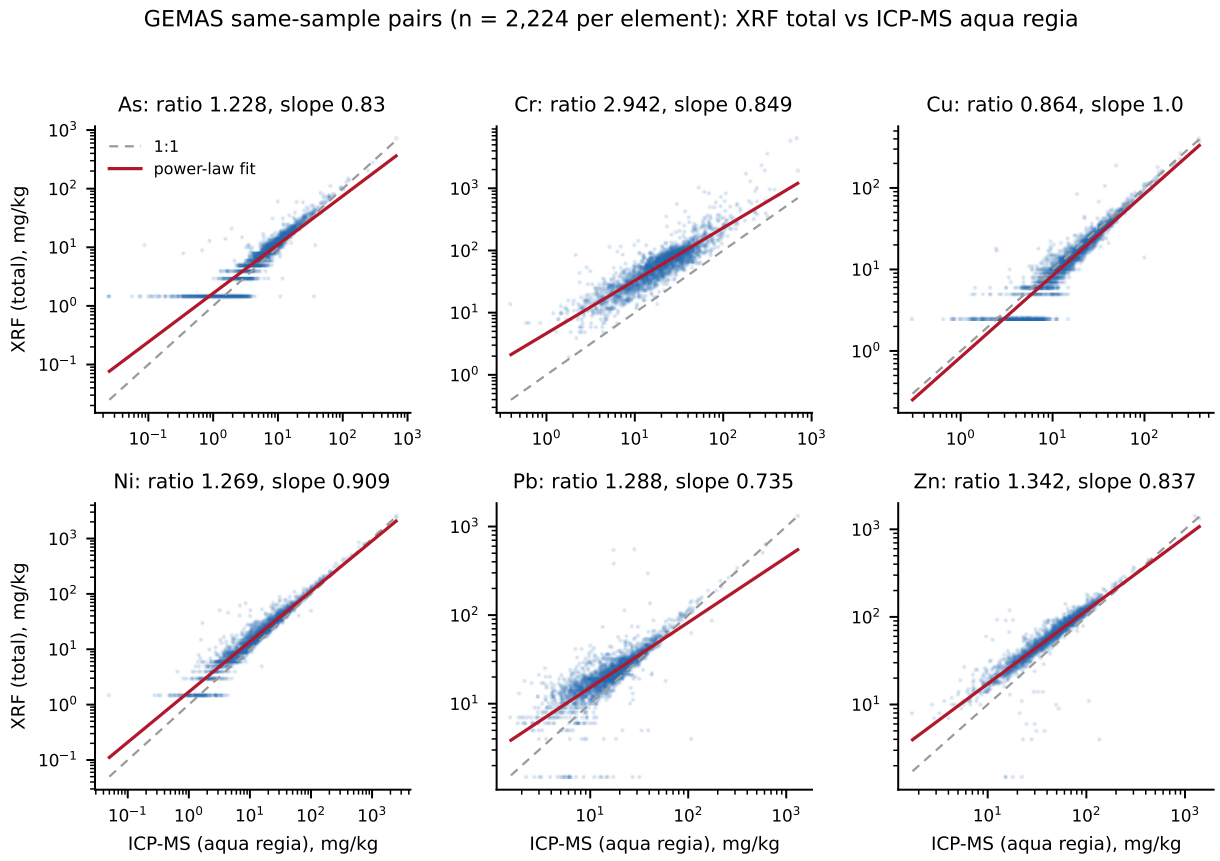

Figure 2: GEMAS same-sample pairs (n=2,224 per element): XRF total against ICP-MS aqua-regia concentration on log-log axes. Dashed grey line: 1:1; red line: fitted power law (Eq. 1, Table 1). Horizontal banding at low concentrations reflects detection-limit quantization of the reported values, retained as-is (no imputation). Points are rasterized at source resolution.

Table 2: Out-of-sample validation (seeded 80/20 split): median absolute percentage error (MdAPE) of predicting the held-out XRF value from the ICP-MS value, for the power-law transfer model vs. a constant-factor model. Gain is the relative MdAPE reduction of the power law.

Element	Power law	Constant factor	Gain
Pb	0.110	0.182	+40%
Zn	0.062	0.075	+17%
Cr	0.203	0.211	+4%
Ni	0.116	0.117	+1%
As	0.173	0.173	0%
Cu	0.141	0.135	−4%

3.3 Independent replication on FOREGS

FOREGS provides a genuinely independent check: different survey (1997–2001 vs. 2008–2009), different sample media (natural topsoil/subsoil vs. agricultural soil), and a different ICP instrument family (optical emission rather than mass spectrometry). Where same-sample double determinations exist (Cr, Zn), the offsets replicate in both magnitude and concentration

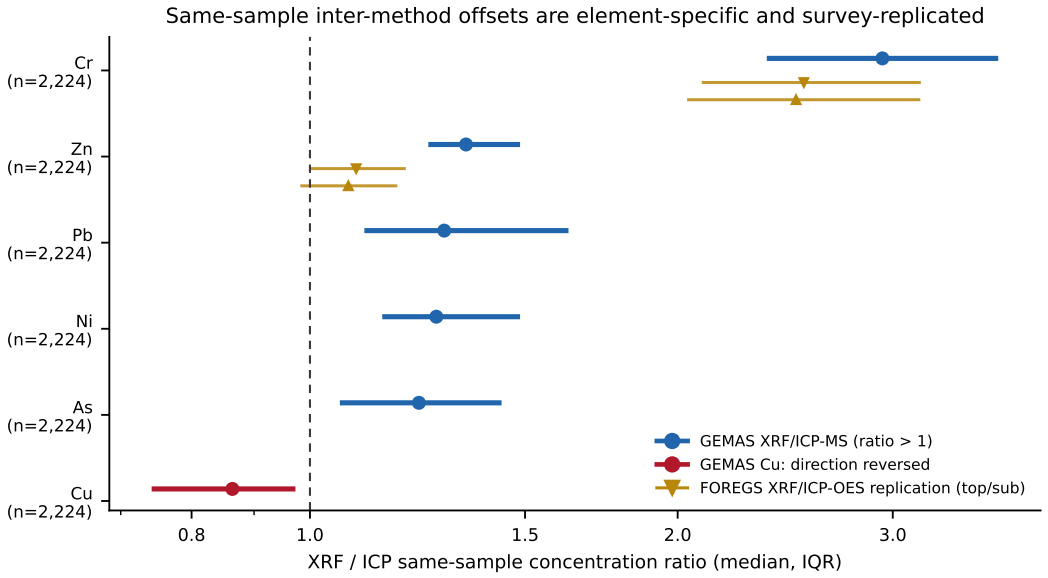

Figure 3: Median same-sample ratios with IQR whiskers. Blue circles: GEMAS XRF/ICP-MS; red circle: the direction-reversed Cu; gold triangles: FOREGS XRF/ICP-OES replication on topsoil (down) and subsoil (up), shown for the elements with same-sample double determinations (Cr, Zn; Table 3). Log-scaled axis.

Table 3: FOREGS replication: XRF (total) vs. ICP-OES (partial digestion), same-sample pairs. Layout as in Table 1.

Panel	El.	n	Median ratio	95% CI	IQR	b (SE)	r^2
Topsoil	Cr	420	2.537	2.476–2.654	2.093–3.163	0.938 (0.023)	0.805
Subsoil	Cr	417	2.500	2.429–2.600	2.036–3.160	0.891 (0.017)	0.864
Topsoil	Zn	420	1.091	1.078–1.107	1.000–1.198	1.131 (0.018)	0.906
Subsoil	Zn	417	1.075	1.062–1.100	0.982–1.179	1.261 (0.019)	0.915

dependence (Table 3): Cr 2.54/2.50 (topsoil/subsoil) against GEMAS 2.94, and Zn 1.09/1.08 against GEMAS 1.34. The Zn difference between surveys (1.09 vs. 1.34) is itself informative: FOREGS ICP-OES follows a stronger multi-acid protocol for some elements [6], so the extractable fraction is survey-specific — exactly why transfer models must be estimated per survey pair rather than assumed universal, and why the metadata standard of §4 requires the digestion protocol, not merely the instrument, to be recorded.

4 Discussion

Consequences for data pooling. If GEMAS aqua-regia and any total-method Cr data were pooled uncorrected, the method split alone would manufacture a factor-3 apparent difference — larger than most genuine regional contrasts in European soil Cr [2]. This is why the harmonized database that hosts these records deliberately certifies zero multi-source poolable cohorts until a transfer layer exists. The matrix in Tables 1–3 is that layer for its two largest soil surveys: applied at the record level it converts “cannot be compared” [7] into “comparable within a quantified

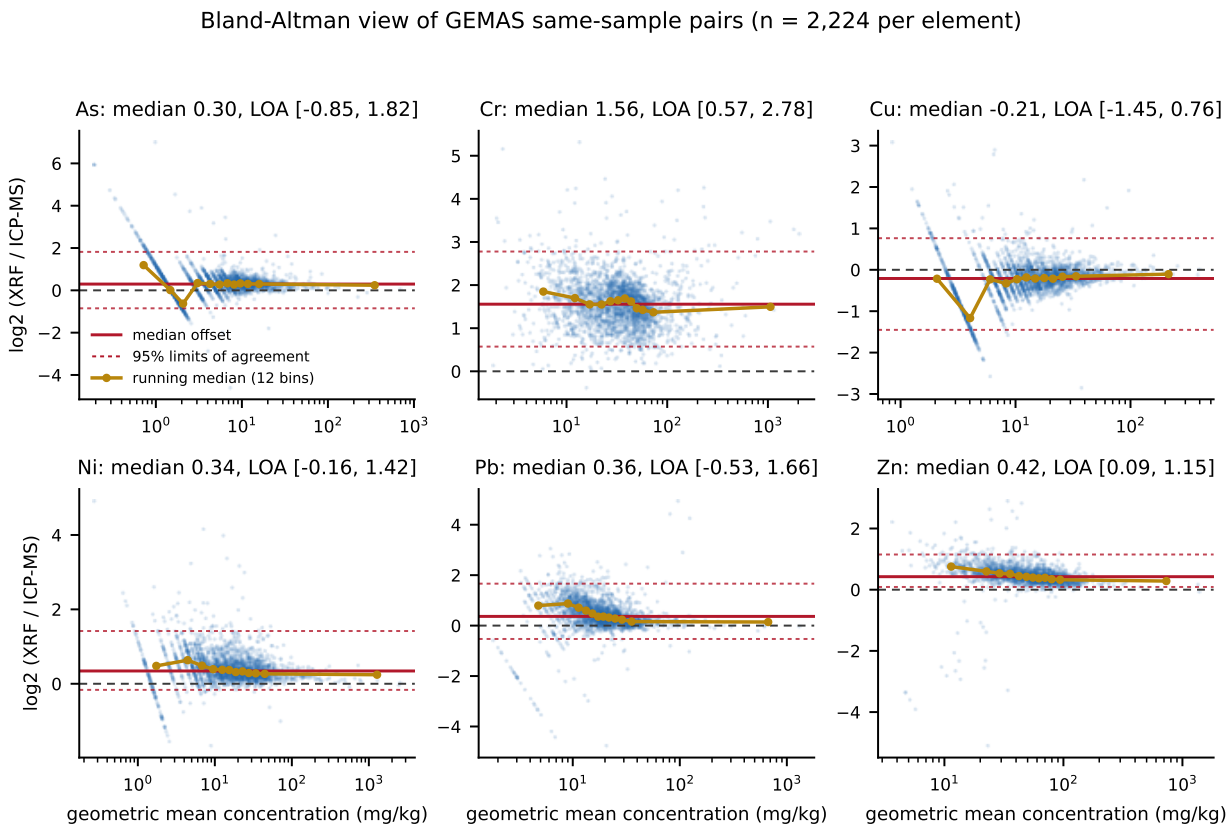

Figure 4: Bland–Altman (mean–difference) view of the same pairs: $\log_2$(XRF/ICP-MS) against the geometric mean of the two determinations. Solid red line: median offset; dashed red lines: nonparametric 95% limits of agreement (2.5th and 97.5th percentiles of the per-sample log-ratios); gold curve: running median in twelve concentration-quantile bins, whose tilt displays the concentration dependence quantified by the slopes b of Table 1. Dashed grey line: zero offset.

uncertainty," with the OOS MdAPE of Table 2 serving as the per-element uncertainty floor.

When a constant factor is enough. The slope criterion gives a defensible, data-driven answer to a recurring practical question. Pb (b=0.735) and the other $b < 1$ elements need the power law at concentration extremes; Cu behaves as a clean proportional offset. Because b and its SE are part of the published table, a pipeline can make this decision mechanically.

Minimum method metadata. Every failure mode encountered in building the matrix traces to missing metadata, not missing chemistry. We therefore propose that survey data releases record, per determination: (1) instrument family (XRF, ICP-MS, ICP-OES, ...); (2) digestion/extraction protocol with standard reference (e.g. aqua regia per ISO 11466 [4]); (3) detection limit and below-limit convention; (4) reference materials analysed in-batch; and (5) a stable sample identifier enabling same-sample joins. GEMAS satisfies all five — which is precisely why this paper could be written from open data; surveys missing item (5) can never contribute to a same-sample matrix, whatever their analytical quality.

Limitations. The matrix covers two European surveys and six elements; transfer functions are survey-pair-specific by construction and must not be extrapolated to other digestion protocols (the FOREGS Zn case demonstrates why). Detection-limit banding inflates ratio dispersion at low concentrations for As and Cu; the OOS errors reflect this dispersion but do not remove it.

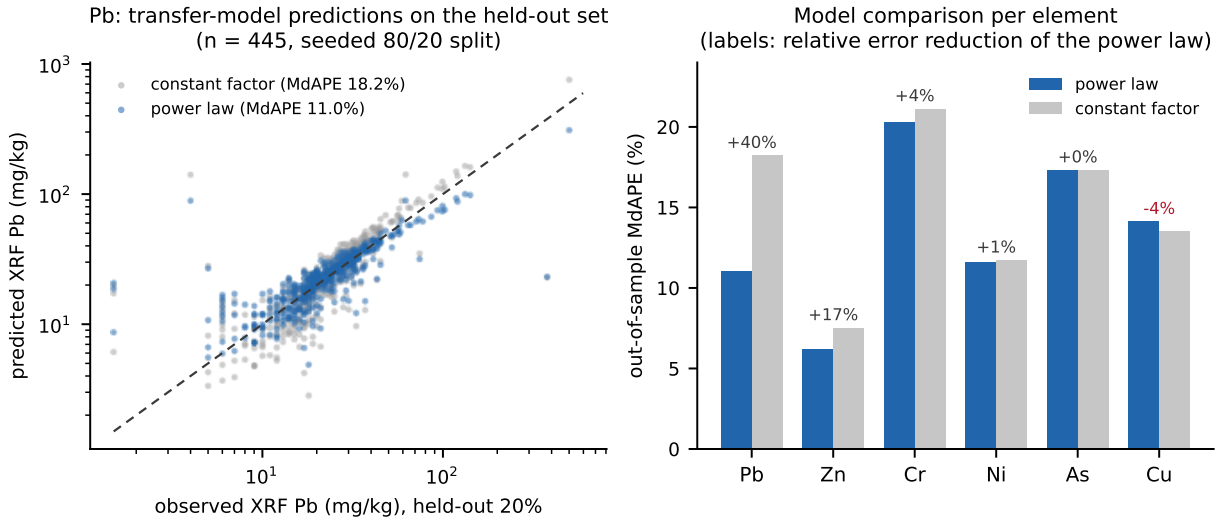

Figure 5: Out-of-sample validation of the transfer models (seeded 80/20 split). Left: predicted versus observed XRF Pb on the held-out set under the power-law (blue) and constant-factor (grey) models; dashed line: 1:1. Right: held-out MdAPE per element for both models; labels give the relative error reduction of the power law, negative for Cu where b=1.000 makes the extra parameter pure noise.

Covariate modelling (pH, organic carbon, clay content) could explain part of the residual spread and is queued as follow-up once those covariates are harmonized into the snapshot.

5 Conclusions

Same-sample double determinations in open continental surveys make the method-comparability problem quantifiable at zero laboratory cost. Across 2,224 GEMAS sample pairs per element, total-to-aqua-regia offsets range from −14% (Cu) to +194% (Cr), are concentration-dependent for five of six elements, and replicate independently in FOREGS. The resulting transfer-model table — medians, CIs, power-law parameters and out-of-sample errors — is published as a machine-readable layer, moving harmonized open geochemistry from a certified "zero poolable cohorts" state to conditional poolability with quantified uncertainty.

Data and code availability

Input surveys: GEMAS [5] and FOREGS [6, 2], ingested with per-record licenses and provenance hashes. The frozen snapshot identifier, analysis scripts (`p345_full.py` and `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-012–dc-016) and has not yet completed human review. Screening-level framing is intentional and load-bearing.

References

[1] Bland, J.M., Altman, D.G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, 327, 307–310. doi:10.1016/S0140-6736(86)90837-8.

[2] De Vos, W., Tarvainen, T. (chief eds.), et al. (2006). *Geochemical Atlas of Europe. Part 2: Interpretation of Geochemical Maps, Additional Tables, Figures, Maps, and Related Publications*. Geological Survey of Finland, Espoo.

[3] European Commission (2025). *Soil Monitoring Law* (Directive on Soil Monitoring and Resilience, in force 16 December 2025). `https://environment.ec.europa.eu/topics/soil-health/soil-monitoring-law_en`.

[4] International Organization for Standardization (1995). *ISO 11466: Soil quality — Extraction of trace elements soluble in aqua regia*. Geneva.

[5] Reimann, C., Birke, M., Demetriades, A., Filzmoser, P., O'Connor, P. (eds.) (2014). *Chemistry of Europe's Agricultural Soils* (GEMAS). Geologisches Jahrbuch B102/B103, Schweizerbart, Stuttgart.

[6] Salminen, R. (chief ed.), et al. (2005). *Geochemical Atlas of Europe. Part 1: Background Information, Methodology and Maps*. Geological Survey of Finland, Espoo.

[7] UNEP Minamata Convention on Mercury, Open-Ended Scientific Group (2025). *Overview of monitoring, emission and release data* (UNEP/MC/OESG.2/2/Rev.1) and OESG scientific report executive summary. `https://minamataconvention.org`.