Through the ceiling: a joint isotope–chemistry attribution of the Uluburun tin bulk

iddin annakam, part II — working title

Conor Reid*

working draft of August 1, 2026

Abstract

Part I established a ceiling: on the Sn-isotope channel the Uluburun cargo bulk's two leading candidates, Cornwall and the Erzgebirge, are unidentifiable in principle — ten batch means carry $N_{\text{eff}} \approx 3$ pseudo-batches of information where ≈420 would be needed — and its value-of-information analysis priced the trace-element channel as the exit. This paper builds that exit and walks through it. We assemble a license-tracked trace-element reference database (8,718 element-analyses across 223 ore specimens, with measurement basis and censoring carried per row), derive per-element smelting-partition priors from published experiments on the disputed ores themselves rather than assumption, and join the chemical channel to the isotope model. Our first fit answered decisively and wrongly — its verdict rode on the two most censoring-corrupted elements through the same smallest-variance pathology part I diagnosed in the field's methods, a failure we document because it is the methods lesson — and the rebuilt model replaces all imputation with Bayesian censored likelihoods and propagates reference-parameter uncertainty end to end. Its results lead, as part I's method demands, with exclusions. Against every candidate source anyone has chemically measured, the cargo's chemistry excludes the alternatives — pairwise Bayes factors of 8–12 against the Erzgebirge (as sampled, at $n = 7$), 7–12 against the Slavkov Forest, 20–42 against Portugal, and effectively no overlap for Spain, nor for Mushiston, whose measured Cu–Sn ores are the natural-bronze argument, priced — and the exclusions hold across every treatment of the ignorance term that stands in for unmeasured sources, including a hierarchical exchangeable-sources default. Cornwall is the only measured candidate that survives contact with its own reference data; its unconditional share of the posterior runs from 0.07 (the narrowest data-centered ignorance) to 0.63 (ore-anchored; 0.52–0.63 under the hierarchical default), and we decline to assert the ordering among survivors, because it is a fact about a prior. The conditional verdict, on the isotope channel's deadlocked pair, is **Cornwall over the Erzgebirge at ≈10:1 per batch** ($P = 0.87$–0.93 across tail-shape and correlation treatments; ≈13:1 under the superseded draft mass-balance anchor); a drawn posterior predictive check places the observed batches at the 84th–89th percentile of what genuine Cornish tin generates, and neither the reference-size asymmetry (Cornwall cut to seven specimens, or to any single granite district, still wins every cell), nor any fixed mass-balance shift in $[0, 2]$ log-units, nor a full log-unit of independent per-source systematics flips it. What remains is what no one has measured: Serbia and the Taurus — the map of that gap is this paper's measurement agenda.

*From the *iddin-annakam* analysis repository (Reid, 2026a). Companion to "Identifiable or not?" (Reid, 2026b). Contact: conor.reid@hey.com.

1 Introduction: the ceiling, and the priced exit

Part I of this series (Reid, 2026b) reanalyzed the Sn-isotope case for the origin of the Uluburun shipwreck's one-ton tin cargo (Pulak, 1998; Powell et al., 2022) and ended at a ceiling. The cargo bulk's ten batch means, after accounting for the correlations that batching induces, carry the information of roughly three independent observations ($N_{\text{eff}} \approx 3$); separating the two leading candidates, Cornwall–Devon and the Erzgebirge, at their observed isotopic separation would require on the order of 420. The per-batch posterior sits at $P(\text{Cornwall}) = 0.53$–$0.54$ — a coin flip — and part I's aggregation analysis showed that no quantity of ingots from the same batching process can move it far. On that channel the question is not unresolved; it is unidentifiable.

The same analysis priced the exits. Among the channels evaluated in part I's value-of-information section, the trace-element chemistry of the ingots — measured by Powell et al. (2022) but never joined to the isotope model — promised an 18–69% cut in the Cornwall-vs-Erzgebirge posterior entropy using data already published. This paper builds that channel with the same review standard part I applied to the field: every modeling convenience priced, every failure documented, every number regenerable from the public repository (Reid, 2026a).

The paper makes six contributions. (1) A license-tracked trace-element reference database (8,718 element-analyses across 223 ore specimens) that carries measurement basis and detection-limit censoring per row (§2). (2) Per-element smelting-partition priors derived from published experiments on the disputed ores themselves, not from assumption (§3). (3) A joint isotope–chemistry model whose first, imputation-based fit answered decisively and wrongly — documented in full because the failure mode is the methods lesson (§4). (4) A censored-likelihood rebuild under which the verdict survives honest uncertainty (§5). (5) A battery of checks — the mass-balance profile, differential per-source systematics, element decomposition, a drawn posterior predictive check, tail shape, reference-size symmetry, correlation structure — each of which costs certainty and none of which flips the answer (§5). (6) An extension to the full twelve-component candidate field, with pairwise Bayes factors against every chemically-measured alternative, that converts the two-source conditional into a set of unconditional exclusions and an explicit map of what remains unmeasured (§6).

2 The trace-element reference database

The database joins three published sources into one long-format table (one row per element per analysis): the ore trace-element compilation of Williams et al. (2025) (8,057 rows — the only multi-district ore reference in existence, including the 100-sample Cornwall–Devon suite and the 11-sample "Erzgebirge" group), the Mushiston ore table from Berger et al. (2022) (431 rows, EDXRF), and the Mushiston mineral-separate chemistry of Konopelko et al. (2022) (230 rows, SEM-EDS, grain basis). Every row carries its source dataset, measurement method, measurement basis (bulk ore vs. mineral separate vs. grain), license, and — critically — its censoring status: a below-detection measurement is stored as a detection limit, never as a value. Rows whose source license permits redistribution ship as CSV with the repository; the remainder are regenerated by parser scripts from the cited supplements.

Basis heterogeneity is not a nuisance column; it is a firewall. A grain-scale SEM-EDS analysis of a mineral separate and a bulk-ore EDXRF measurement are not the same quantity, and part I's isotope library showed what silently pooling bases does to a reference. The model below therefore uses bulk-basis rows only, and the working references reduce to the Williams et al.

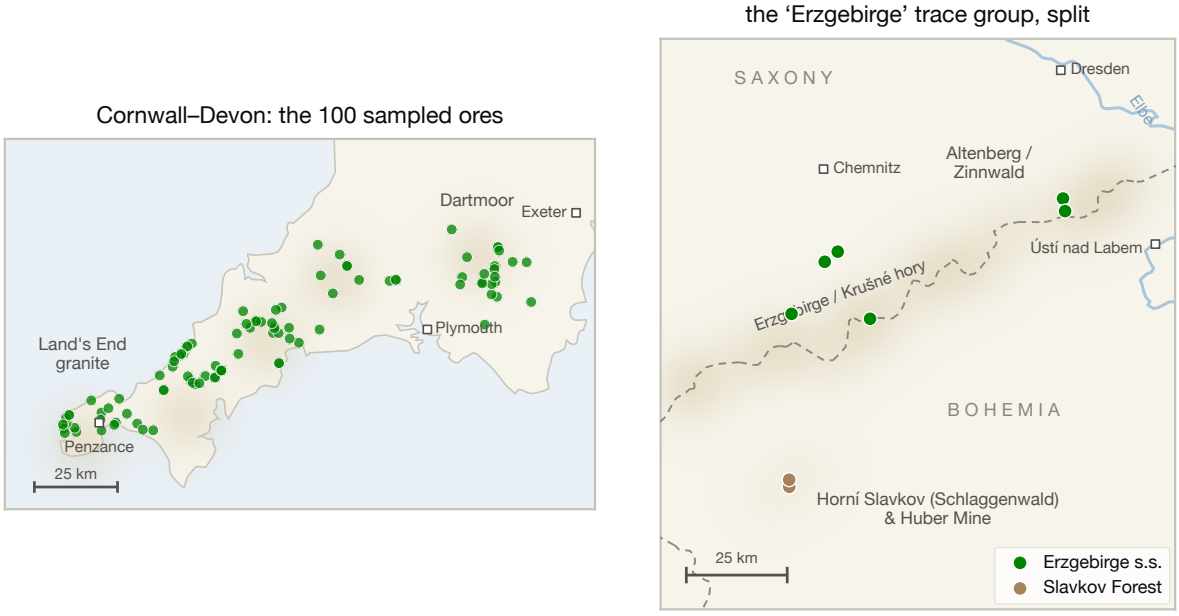

Figure 1: The reference data at the scale it was collected. Left: the 100 Cornwall–Devon ore samples of Williams et al. (2025), strung along the Cornubian granite spine from Land's End to Dartmoor. Right: the eleven samples published as one "Erzgebirge" group, split. The four Slavkov-district analyses (labelled Horní Slavkov, Schlaggenwald, and Huber Mine; the first two are one town's Czech and German names) sit on a separate granite massif ∼40 km southwest of the range proper — the soft wash marks the crest, which the modern Saxon–Bohemian border (dashed) follows. It is the same granularity blend part I corrected in the isotope library, recurring on the trace side.

(2025) suites: Cornwall–Devon at $n = 100$ and the Erzgebirge at $n = 11$ — which is where the first correction lands. Williams's "Erzgebirge" group blends in four analyses from the Slavkov Forest, a separate granite massif ∼40 km southwest of the range whose ores are isotopically distinct (Fig. 1) — the same granularity blend part I documented and split in the isotope library. (The four are labelled Horní Slavkov, Schlaggenwald, and Huber Mine in the source table; the first two are the Czech and German names of the same town, at identical coordinates, so the localities are two.) We split it here too. Erzgebirge *sensu stricto* is $n = 7$ — two specimens from the eastern range (Altenberg, Zinnwald), five from the western (Geyer ×2, Ehrenfriedersdorf, Černý Potok, and Pöhla, spelled "Pohia" in the source table) — so the entry spans the range, thinly. Slavkov is carried as its own component in the full-field model of §6. One mismatch must be named now: part I's isotope entry for the Erzgebirge pooled $n = 32$ analyses across the range, so this paper tests an isotope-defined regional candidate against a trace reference of seven specimens. Part I's own finding — a verdict's strength must be read against its entry's aggregation level — applies to every Erzgebirge statement below, and §6 scopes the exclusion accordingly.

The ingot side is the Powell et al. (2022) panel — In, Sb, Pb, Bi, Cu on 105 ingots, fully quantified (no censoring; verified 105/105 positive values on all five elements), aggregated to the five bulk batches part I's models operate on. The reference side is nothing like as clean. On the decisive elements the references are heavily censored: 29 of 100 Cornwall bismuth values and 4 of 7 Erzgebirge *s.s.* bismuth values sit below detection. How that censoring is treated turns out

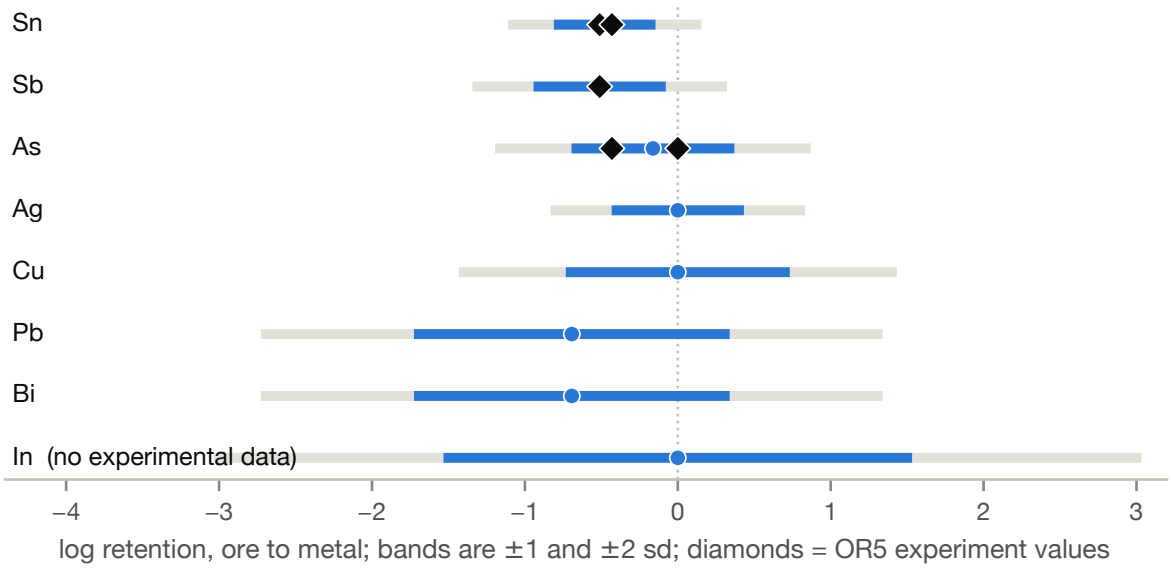

Figure 2: Per-element smelting-partition priors (log effective partition ratio, ore → metal; > 0 means enrichment), built from Berger et al. (2022)'s smelting experiments on Mushiston ores rather than assumption. Diamonds mark values the experiments state numerically. Pb and Bi are wide because the experiments found them *matrix-dependent*; In has no experimental data anywhere — a measurement gap this paper inherits and names.

to decide what the model says (§4), which is why the database stores it as structure rather than resolving it at ingest.

Coverage is the database's loudest fact, and it is graded. Two components have references worth the name: Cornwall ($n = 100$) and the Erzgebirge *s.s.* ($n = 7$) — the pairing §§4–5 can adjudicate. Four more have references thin enough that only a censoring-honest model can use them at all: the Slavkov Forest ($n = 4$, from the split above), Spain and Portugal ($n = 9$ and $n = 8$ — Williams's Iberia group, split by the table's own country column, with most panel values below detection), and Mushiston (Berger's ore table: Sb/Bi/Cu by EDXRF on the Cu–Sn ore subset, $n = 30$, no indium). All of them enter the full-field model of §6. And two candidate sources at the center of this dispute — Serbia and the Taurus — have no trace-element data anywhere in the literature. Part I's coverage map marked them as open circles; the claim structure below is built so that this limitation is carried explicitly rather than assumed away.

3 Partition priors from measured experiments

An ore reference cannot be compared to ingot metal directly: smelting changes every concentration. The trace channel therefore needs an analog of the isotope channel's fractionation offset δ, and we give it the same treatment part I gave δ — a prior measured on the disputed ores, not assumed. For element e the ore→metal log-concentration shift is decomposed as

$$\lambda_e \;=\; \kappa \;+\; \log r_e,$$

where κ is a shared mass-balance term (the metal phase is a small fraction of the charge, concentrating everything that stays) and r_e is an *effective partition ratio* (metal over ore, net of the

process) — values above 1 mean enrichment, as the experiments observed for Bi in the tin smelt, so "retention" is used loosely below for $r_e < 1$. κ is identical for every element and every source — but shared is not the same as harmless, and an earlier draft of this section claimed it cancels. It does not: a shift common to both sources still moves a likelihood ratio when the predictive variances differ (the derivative of the log-LR in κ is a difference of inverse-variance-weighted residuals, zero only by coincidence), just as part I's δ traded against source location rather than cancelling. So κ gets the treatment δ got: a measured anchor, a propagated prior, a profile. Mass balance fixes the anchor — metal mass is ore mass × grade × tin retention, so a fully-retained element concentrates by $\kappa = -\log(g\, r_{\mathrm{Sn}}) \approx 0.91$ at the experiments' $g \approx 0.65$, $r_{\mathrm{Sn}} \approx 0.62$. (An earlier version of our prior module anchored at $-\log g = 0.43$, an error this paper's first review caught; both anchors are carried through §5.) The prior is $\kappa \sim \mathcal{N}(0.91, 0.5^2)$, propagated through every classification, and the verdict is additionally profiled over fixed $\kappa \in [0, 2]$. The alternative that makes κ cancel exactly — log-ratio coordinates — is unavailable honestly here: ratios of censored values are interval-censored, and the Erzgebirge side has roughly two complete Bi/Sb pairs; complete-case log-ratios would silently drop the censored mass, the exact sin §5 exists to avoid. The discrimination burden falls on the r_e, which is where the experiments come in.

The priors (Fig. 2) come from the smelting experiments of Berger et al. (2022) (their Online Resource 5): two successful smelts of Mushiston oxidized ores — a bronze co-smelt and a tin smelt — with element-wise before/after chemistry. Tin itself loses 35–40% of its mass to slag and fume ($r_{\mathrm{Sn}} \sim \log\mathcal{N}(\log 0.62, 0.30^2)$); antimony "behaved similarly to tin" in both experiments; arsenic and silver are retained. The two informative failures are lead and bismuth: heavily depleted in the bronze co-smelt, slightly *enriched* in the tin smelt — matrix-dependent, in the experimenters' own reading — so both get wide priors ($\sigma = 1.0$ log-units) by honesty rather than convenience. And indium, the element on which the field's headline trace-element claims have leaned hardest, has *no experimental partition data at all*: no smelting experiment anywhere has measured what smelting does to indium. It gets an effectively uninformative prior ($\sigma = 1.5$) and a standing entry in the measurement agenda of §6. The same source anchors the isotope side: the experiments' measured ore→metal fractionation sits where the complete-reduction prior part I took from Berger et al. (2018) expects it — +0.13‰ inside the 0.09–0.18‰ interval, +0.20‰ just above it, both far from the ≤ 0.88‰ incomplete-reduction regime — the model's δ prior, exercised on the disputed deposit's own ores. (All δ values in this series are on the $\delta^{124/116}$Sn NIST 3161a scale; part I documents the per-source conversions, and two scale errors it caught by cross-checking.)

All of this rests on one experimental campaign, one furnace type, and one ore matrix; the priors are correspondingly generous, and the sensitivity of every downstream claim to their width is reported where it matters. We would rather carry wide, measured priors than narrow, assumed ones; part I is one long documentation of what the latter cost the field.

4 The joint model, and the first fit's instructive failure

The joint model is deliberately plain. Per bulk batch g the data are $y_g = (m_g, t_g)$: the batch-mean $\delta^{124/116}$Sn and the batch-mean log-trace vector over the working panel. Under source k,

$$m_g \sim \mathcal{N}(\mu_k^{\mathrm{iso}} + \delta,\ \sigma_k^{\mathrm{iso}\,2}), \qquad t_g \sim \mathcal{N}(\mu_k^{\mathrm{trace}} + \lambda,\ \Sigma_k^{\mathrm{trace}}),$$

with δ the smelting fractionation prior of part I ($\mathcal{N}(0.14, 0.05^2)$), λ the partition priors of §3, and the source parameters estimated from the references. Classification is by prior-propagated Monte Carlo over (δ, λ, reference parameters) with a uniform two-source prior — no MCMC is needed for the membership computation itself, which keeps every sensitivity run cheap enough to actually run. The two channels are treated as conditionally independent given the source — stated rather than buried: we know of no pathway by which a deposit's Sn-isotope composition constrains its trace-element inventory at fixed source, but it is an assumption, not a theorem.

Two bookkeeping points, stated precisely. The batches are Powell's own group labels, and the five bulk batches used here (P1, MC1A/B, MC2A/B) are the cargo-bulk subset of part I's ten batch units (eight named groups plus two ungrouped ingots); each contributes a paired isotope mean and trace mean computed over the identical ingot membership — all 105 ingots carry both channels, so there is no coverage asymmetry between the two. Second, t_g is the mean of per-ingot log concentrations, not the log of a batch bulk assay; the two differ by a Jensen term that scales with within-batch dispersion, which the source-level predictive variance we carry — a deliberately conservative choice — dominates.

The first fit (all five elements, references summarized by their sample moments after the field-standard $x/2$ substitution for below-detection values) answered immediately and decisively: $P(\text{Cornwall}) = 0.98$–$1.00$ in every batch (1.000 exactly with λ zeroed). Part I taught us what to do with a decisive answer: decompose it. The per-element log-likelihood-ratio decomposition finds the verdict carried almost entirely by lead (+3 to +13 nats per batch) and bismuth (+9 to +10), while indium and antimony — the elements the literature's qualitative trace argument had celebrated (Powell et al., 2022) — contribute approximately nothing at batch level (indium is slightly *pro*-Erzgebirge). And lead and bismuth are exactly the two elements the model should trust least: the widest, matrix-dependent partition priors (§3), the heaviest reference censoring (§2), and an imputation rule that both mislocates the censored mass and *shrinks the reference variance* — manufacturing precisely the artificial tightness that part I diagnosed, in the field's isotope methods, as the smallest-variance pathology. Our own first fit reproduced the field's central error on a different channel. Dropping the suspect pair collapses the verdict to $P = 0.57$–0.64:

panel (imputed references, diagonal)	P(Cornwall), bulk batches
all five elements	0.975–0.999
drop Bi	0.73–0.99
drop Pb and Bi	0.57–0.64

A decisive verdict resting on the two least-trustworthy elements is not a result; it is a diagnosis. We report it in full because the repair — treating censoring as likelihood structure rather than a data-cleaning step — is the paper's methodological core, and because the episode is a live demonstration that the review standard of part I applies to us.

5 The censored rebuild

The rebuild removes imputation everywhere. Each (source, element) reference is fit as a Bayesian censored normal in log space: observed rows contribute $\mathcal{N}(\log v;\, \mu, \sigma^2)$, below-detection rows contribute $\Phi((\log x - \mu)/\sigma)$ — the probability of falling below the limit, not a made-up value at half of it. The full posterior of (μ, σ) is then propagated into the classification, which also delivers, for free, the reference-parameter uncertainty that the $n = 7$ Erzgebirge side always needed.

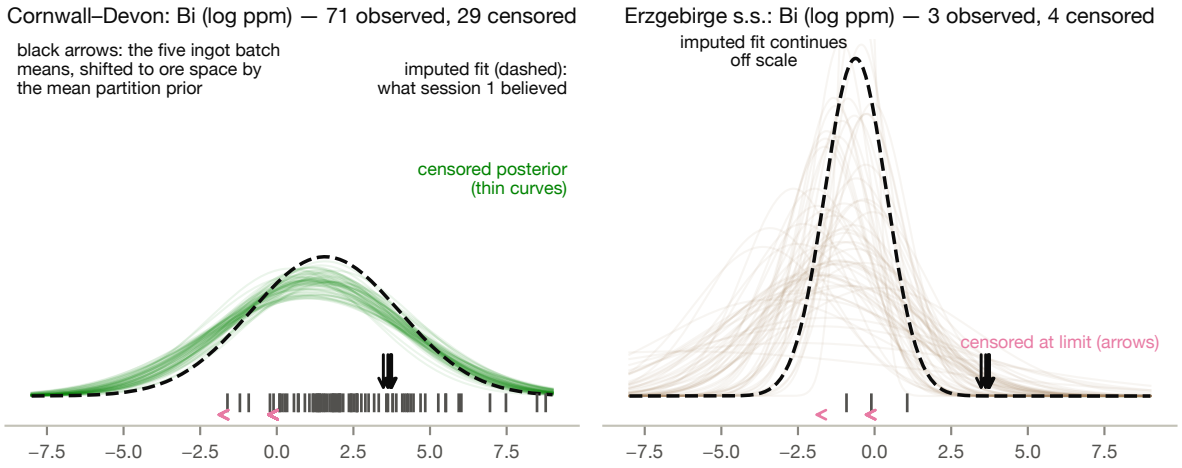

Figure 3: The pivotal correction. Bismuth references under the session-1 imputation (dashed) versus Bayesian censored fits (thin posterior curves); ticks are observed values, arrows are below-detection limits. On the Cornwall side imputation mislocates modestly; on the Erzgebirge side — three observed values, four censored — it manufactures a confident spike where honesty leaves wide uncertainty. Session 1's decisive verdict rode on exactly this artifact. Black arrows mark the five ingot batch means, shifted to ore space by the mean partition prior: inside Cornwall's posterior curves ($z \approx +0.7$), far up the Erzgebirge's ($z \approx +1.8$) — the discrimination is an Erzgebirge-side mismatch, not a Cornwall-side tail.

The likelihood is per-element (diagonal) for both sources — stated openly, because estimating a censored full covariance from seven specimens is not honest work — and the imputed-but-diagonal variant is reported alongside so the censoring effect is isolated from the diagonalization effect.

The correction moves the references exactly as the arithmetic predicts (Fig. 3): Cornwall's bismuth reference widens (σ 2.43 → 2.97), and the Erzgebirge's — three observed values, four censored — relocates and honestly explodes (μ −0.62 → −1.47, σ 1.00 → 2.18 ± 1.1). The imputed fit had been both mislocating the reference and manufacturing confidence. And the verdict holds anyway:

panel (censored, honest λ)	P(Cornwall), bulk batches
all five elements	0.932–0.986
drop Bi	0.771–0.946
drop Pb and Bi	0.658–0.709
drop Pb, keep honest Bi	**0.895–0.921**

The bolded row is the one this paper stands behind. Lead is excluded on part I's own grounds — the ingots' lead is externally contaminated, and if the lead isotopes are corrupted the lead concentrations are suspect too — while bismuth stays, because its censoring is now modeled rather than papered over. Conditional on Cornwall-or-Erzgebirge: ≈10:1 for Cornwall, per batch — and per batch only: the five batches are not independent draws (part I's $N_{\text{eff}} \approx 3$ was computed for exactly this cargo's batching), so the per-batch odds must not be multiplied. The cargo-level claim, stated with bounds rather than a point: under full batch dependence the cargo carries one batch's worth of evidence (≈10:1); under independence it would compound to the fifth power; part I's $N_{\text{eff}} \approx 3$ sits between. We decline to pick the point — the dependence

structure is part of what part I showed is unidentified — and note only that every point on the interval favors Cornwall.

The rest of this section is the battery of ways we tried to break that number (Fig. 4).

Correlations. The diagonal likelihood is an approximation the Cornwall side's $n = 98$ complete rows can test. A Gaussian copula — for Gaussian marginals, simply a multivariate normal with correlation matrix R estimated from the complete uncensored rows ($\max |r| = 0.25$ on the working panel) — moves the verdict to $P = 0.902$–0.932. Slightly up, not down.

The mass-balance shift. The corrected anchor ($\kappa \sim \mathcal{N}(0.91, 0.5^2)$, §3) is the baseline for every number in this paper, including the table above; the superseded draft anchor ($-\log g = 0.43$) gives 0.921–0.940 (≈13:1) and is retained in the profile for the record. Profiled at fixed values: $\kappa = 0$ gives 0.948–0.959; 1.0 gives 0.897–0.928; 1.5 gives 0.852–0.900; 2.0 gives 0.788–0.860. No κ in $[0, 2]$ flips the verdict; "what if $\kappa = 1.5$?" has a computed answer.

Differential systematics — the assumption the rest of the ladder cannot touch. λ is shared across sources, so its uncertainty (and κ's) is common-mode; the dangerous mode is differential. If bismuth partitions differently out of Cornish and Erzgebirge ore matrices, the separation itself is noised — and Bi is both the anchor element and the one the experiments flagged as matrix-dependent. Two defenses. The textual one: the matrix dependence Berger observed was between a bronze co-smelt and a tin smelt — a process difference, not a source-geology difference — both candidates are mineralogically similar cassiterite systems (unlike Mushiston's stannite), and the cargo is a single product class, so the process is common-mode. The computed one: add independent per-source, per-element offsets $\eta_{k,e} \sim \mathcal{N}(0, \sigma^2_{\text{sys}})$ and sweep. At $\sigma_{\text{sys}} =$ 0.3/0.5/1.0/1.5/2.0/3.0 log-units the verdict (re-anchored κ) is 0.89–0.92 / 0.89–0.92 / 0.86–0.90 / 0.83–0.87 / 0.79–0.83 / 0.71–0.75: it decays gracefully toward the isotope-only floor and *never inverts*. A full log-unit of independent systematics — larger than the entire ore-to-metal shift the experiments measured for most elements — costs three points. The same rung prices the channel the censored likelihoods cannot: modern reference specimens versus Bronze Age exploited ore. Censoring repairs sampling noise, not sampling bias; if Williams's Erzgebirge specimens are deep-workings hypogene material while LBA exploitation was placer and supergene, the mobile elements (Bi, Pb) are where a directional offset would live — and it is priced here, not assumed away.

Element decomposition, redone honestly. Leave-one-out and element-alone runs on the safe panel: bismuth is the anchor (alone with the isotopes it reaches $P = 0.84$–0.86; without it the verdict falls to 0.66–0.71), copper and antimony are the supporting cast (alone, 0.63–0.69 and 0.56–0.58), and indium alone is statistically indistinguishable from isotopes alone (0.52–0.53 vs. 0.53–0.54) — at batch level, the celebrated discriminator contributes nothing. The verdict no longer rides on distrusted elements: it rides on honestly-censored bismuth with independent support, and it survives bismuth's complete removal at ≈2:1. The full lead-free lattice — all fifteen element subsets — confirms the surface rather than four points on it: every bismuth-bearing panel sits at 0.84–0.92, every bismuth-free panel at 0.52–0.71, and the working panel (0.895–0.921) is statistically tied with Sb+Bi+Cu (0.893–0.923) at the top of the fifteen. One selection choice must be dated honestly: excluding lead was decided before any fit, on part I's contamination finding; keeping bismuth was decided after session 1, as the product of the censored rebuild. The lattice is reported so the reader holds the whole surface, not our path through it.

Indium, with its prior tightened. Indium's non-contribution could have been our own prior's artifact — we gave λ_{In} $\sigma = 1.5$ because no experiment has measured it (§3), and a wide prior

can wash an element out by itself. It is not the prior: tightened to $\sigma = 0.3$, indium alone still yields $P = 0.502$–0.511, because the two references are simply indistinguishable on indium (predictives $+3.73 \pm 1.95$ vs. $+3.66 \pm 1.93$ log-ppm). No prior can conjure discrimination from identical references; the literature's favourite discriminator is uninformative here as a fact about the data.

Predictive location, and the posterior predictive check. Part I's standard is that a model must survive its own PPC before its attributions are believed. Three computations, in increasing strength. *Single-point controls:* a batch built at the Erzgebirge's own posterior mean comes back Erzgebirge (P(Cornwall) = 0.37); at Cornwall's, Cornwall (0.69) — both recovered, neither a hair trigger. *Predictive location:* against Cornwall's predictive distributions the batch means sit at $z \approx -0.2$ (In), $+0.0$ to $+0.2$ (Sb), $+0.7$ (Bi), $+0.2$ to $+0.9$ (Cu) — inside the reference on every element — while against the Erzgebirge's they sit at $+1.7$ to $+1.8$ on bismuth (Fig. 3). The verdict is not a tail artifact of Cornwall's distribution; it is a $\approx 1.8\sigma$ mismatch on the Erzgebirge side, compounded by consistent smaller ones. *The drawn control — the PPC proper:* 2,000 synthetic batches drawn from each source's full predictive (not one batch at its mean), classified identically. True-Cornwall batches produce P(Cornwall) with median 0.64 and 90th/99th percentiles 0.92/0.96; true-Erzgebirge batches, median 0.35. The real batches (0.894–0.922) sit at the 84th–89th percentile of the true-Cornwall distribution — the upper quintile, not beyond it — and at the 98th–99th of the true-Erzgebirge one. The observed chemistry is the kind genuine Cornish tin generates and genuine Erzgebirge tin essentially never does. (The upper quintile is also where a correctly-specified model should put it: the synthetic batches carry single-specimen dispersion, while a real batch mean averages ≈ 21 ingots — the same conservative variance choice the likelihood itself makes.)

Tail shape. At $n = 7$ with $\sigma = 2.18 \pm 1.1$ the Gaussian tail is an assumption doing work, so the trace channel is refit with Student-t predictives: $P = 0.867$–0.894 at $\nu = 4$, 0.877–0.905 at $\nu = 7$. Fat tails cost about three points of probability and spare the verdict.

Reference-size symmetry. The obvious worry — raised by every reader of a 100-vs-7 comparison, including this paper's author — is that the data-rich source wins by volume. It does not, and the test is direct: subsample Cornwall's reference to seven specimens (preserving per-element censoring status), refit everything, reclassify; repeat for twelve random subsets. Cornwall wins all sixty subset×batch cells — median $P = 0.853$, minimum 0.729. Volume buys the difference between 6:1 and 10:1, not the verdict. The mechanism, stated correctly: once reference-parameter uncertainty is propagated, sample size enters through the predictive width — the $n = 7$ side gets a fatter, heavier-tailed predictive, which makes it *harder to exclude*, not easier. The asymmetry runs against our conclusion, and the conclusion survives it. The sterner, granularity-matched version — a random 7 of Cornwall's 100 still spans the whole granite belt, so ask instead whether any *single district* could carry the verdict alone — refits Cornwall's reference from one granite district at a time: Land's End ($n = 20$) gives 0.87–0.89, Carnmenellis ($n = 30$) 0.88–0.92, St Austell ($n = 22$) 0.83–0.85, Bodmin ($n = 7$) 0.94–0.96, Dartmoor ($n = 21$) 0.87–0.92. Any single Cornish district, alone, still beats the Erzgebirge entry at ≥ 5:1.

6 What the conditional means, and what would widen it

Everything above is conditional on the bulk being Cornwall-or-Erzgebirge — the pairing the isotope channel left mutually unidentifiable, and the only pairing with references worth the name. The honest question is what happens when the other ten components of part I's source model

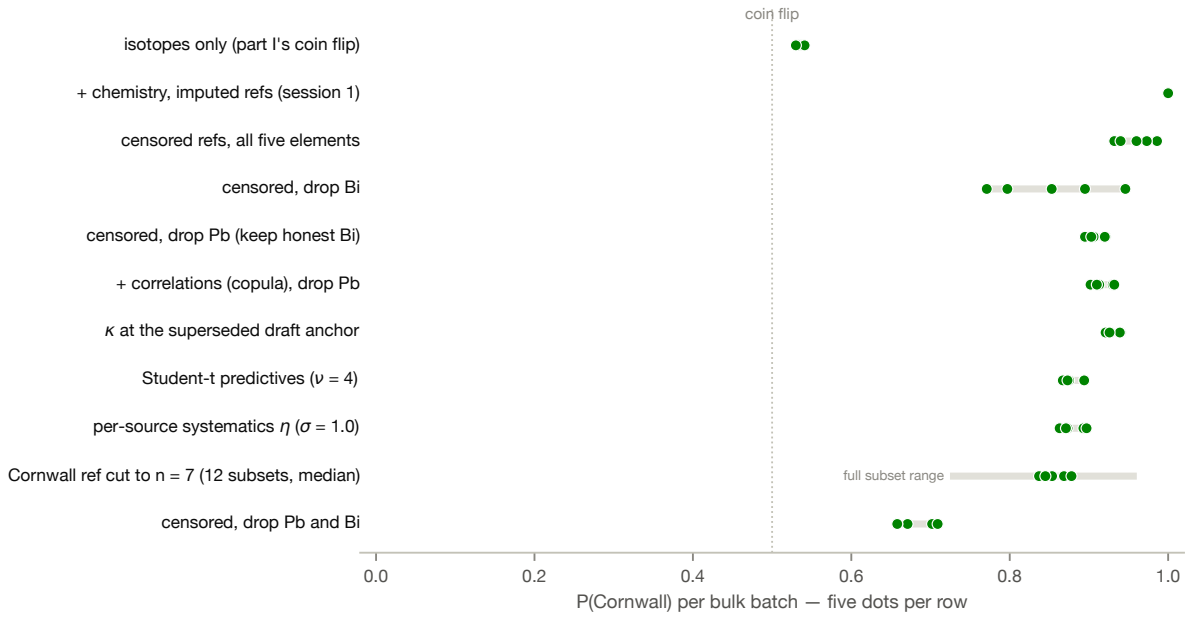

Figure 4: The whole argument as a ladder. Five dots per row (the bulk batches). Isotopes alone are part I's coin flip; imputed chemistry says 1.000 for the wrong reason; each honesty measure costs certainty and the verdict survives them all — the contamination-safe, censoring-honest, correlation-aware rows agree at $P(\text{Cornwall}) \approx 0.90$–$0.93$; Student-$t$ predictives and a full log-unit of independent per-source systematics (η) each cost ≈3 points, and the superseded draft κ anchor ran ≈3 points higher. The size-symmetry row cuts Cornwall's reference to seven specimens, the Erzgebirge's own data poverty, across twelve random subsets: the median drops to ≈ 0.85 and no subset–batch cell falls below 0.73 — the verdict is carried by where the chemistry sits, not how much of it exists. Dropping both suspect elements shows the floor (≈ 0.70).

re-enter. We extend the model to the full field with partial-channel likelihoods, per source *and per element*: where a real reference exists it enters as its censored posterior (Cornwall with the copula, the rest diagonal); where none exists, that element gets a deliberately broad ignorance factor (centered on the pooled ingot mean, $\sigma = 3.0$ log-units) — so likelihood dimensionality is equal across sources, and no component is excluded by fiat. The graded coverage of §2 now pays off: six components carry measured chemistry into this run — Cornwall, the Erzgebirge, the Slavkov Forest, Spain, Portugal, and Mushiston (Sb/Bi/Cu, with an ignorance factor standing in for its unmeasured indium).

The result (per-batch ranges, baseline ignorance width):

component	trace reference	P, bulk batches
Cornwall–Devon	$n = 100$, censored, copula	0.26–0.33
Erzgebirge *s.s.*	$n = 7$, censored	0.021–0.032
Slavkov Forest	$n = 4$, censored	0.026–0.036
Spain	$n = 9$, censored	$< 10^{-3}$
Portugal	$n = 8$, censored	0.006–0.013
Mushiston	$n = 30$, Sb/Bi/Cu (EDXRF)	$< 10^{-4}$
Serbia	*none exists*	0.23–0.31
Taurus	*none exists*	0.17–0.22
Afghan orbit + outlier class	*none exists*	0.17–0.18

Because full-field shares depend on what else is in the field, the exclusions are also stated as pairwise Bayes factors — Cornwall against each measured alternative, invariant to the rest of the roster. Under the re-anchored κ: Erzgebirge 8–12, Slavkov Forest 7–12, Portugal 20–42, Spain $>$ 1,300, Mushiston beyond the model's meaningful range (read "no overlap," not the literal number; Gaussian tails are not to be trusted at that magnitude). Under the draft anchor the factors run higher (Erzgebirge 12–16, Portugal 39–78) — and the Erzgebirge pairwise factor is, by construction, the same computation as §5's headline odds, so the two quote the same number under the same anchor. The pairwise factor is the quantity that means "excluded," and it will not move when someone adds a thirteenth component.

(The twelve components, for the record: Mushiston, Dara-i-Pech, Paprok, and Kunar — the Afghan orbit — plus Cornwall–Devon, Erzgebirge *s.s.*, Slavkov Forest, West Serbia, Spain, Portugal, the Taurus, and the outlier class.)

Three readings. *First*, the chemistry eliminates every measured alternative. Cornwall is the only component in the candidate field that survives contact with its own reference data — every other source somebody has actually measured is pushed below 0.04 per batch in the baseline full field (below 0.075 under every ignorance treatment tried), and below 1:7 in the pairwise factors. For the Erzgebirge, "measured" means the seven-specimen entry of §2, two eastern and five western sites: the exclusion is of the range as Williams sampled it, Slavkov's ($n = 4$) is the weakest of the set and should be held loosely, and part I's aggregation-level warning applies to both. *Second*, the Mushiston number is a statement about the ore population that has actually been measured: the hypogene stannite and supergene Cu–Sn mixtures that define the deposit's type mineralization (Konopelko et al., 2022) — copper-bearing down to its mineralogy (mushistonite is a Cu–Sn hydroxide) and carrying Cu at 9–35 mass% against a clean-tin cargo, a figure that is partly true by the subset's construction, which is why the claim is scoped to that population. This prices the "Mushiston smelts to natural bronze" argument of Berger et al. (2023) for the measured ores; it does not rule out an undocumented Cu-poor ore body — though we note the experiments that supplied our partition priors smelted Mushiston's *oxidized* ore and titled the result "shiny bronze" (Berger et al., 2022). Part I's isotope-side exclusion (6.6σ) stands on its own grounds and does not need this one. (The reference is EDXRF normalized to 100 mass%, a different method family from the Williams suites; the offset such a comparison risks is orders of magnitude smaller than this signal.) *Third*, the surviving posterior mass sits almost entirely on sources no one has measured — Serbia, the Taurus, and (diffusely) the Afghan orbit — held up by ignorance, not evidence. Part I's coverage map drew these as open circles; the joint posterior now literally is that map.

The ignorance width is a judgment, so we sweep it. Across $\sigma = 2.0$–4.0 every exclusion is stable — the Erzgebirge never exceeds 0.056, Slavkov 0.065, Portugal 0.023, Spain and Mushiston

10^{-3} — while the *ordering among survivors* is not: Serbia leads at 2.0 (0.32–0.43), Cornwall at 4.0 (0.48–0.56).

Width, moreover, is not even the live parameter; the *center* is. Centering unmeasured sources on the pooled ingot mean — as above — centers them on the data: it guarantees they fit, so measuring a source can only hurt it relative to leaving it unmeasured. That is the mirror image of part I's closed-world finding, and it deserves the name. Re-centering the ignorance term on the measured ore references instead swings the survivor ordering wholesale — Cornwall 0.56, Serbia 0.13, Taurus 0.09 — while every exclusion persists (Erzgebirge ≤ 0.060, Slavkov ≤ 0.073, Portugal ≤ 0.013, Spain and Mushiston ≈ 0).

A single defensible default exists between those endpoints, and we ran it: an exchangeable-sources prior, with unmeasured source means drawn from the fitted population of the measured ones (per element: population mean and spread across the measured references, plus mean within-source variance). It lands near the ore-centered endpoint — Cornwall 0.52–0.63, Serbia 0.12–0.13, Taurus ≈ 0.09 — and confirms every exclusion (Erzgebirge ≤ 0.064, Slavkov ≤ 0.074, Portugal ≤ 0.026, Spain and Mushiston ≈ 0), with or without Mushiston's extreme chemistry in the population. Between two centerings, the hierarchical default, and the width sweep, the survivors reorder freely and the excluded never return. That asymmetry defines the claim structure: we assert the exclusions, and we decline to assert the ordering — the ordering is a fact about a prior, and it will remain one until somebody measures an ore.

Which is the point of the measurement agenda, in value order. *First*, trace-element measurement of the West Serbian placer ores: part I's value-of-information analysis already ranked Serbia the single highest-value measurement in the dispute on the isotope channel, and the full-field run now shows it carrying the largest unmeasured share of the joint posterior; the existing surveys (e.g. Mathur et al., 2025) report isotopes without trace panels. *Second*, the same for the Taurus deposits — a conspicuous gap given how much of this dispute began there (Powell et al., 2022; Mathur et al., 2025). *Third*, seven more Erzgebirge ores — from LBA-plausible workings specifically, placer and supergene material rather than deep hypogene specimens: the fastest honest way to overturn this paper is a resampled Erzgebirge reference that lands somewhere else, since the exclusion rests on the range as sampled at $n = 7$, and the drawn PPC of §5 cannot test what a mis-sampled reference hides. We would rather name the falsification path than have it named for us. *Fourth*, one smelting experiment with an indium-doped charge: the literature's most-celebrated discriminator has no measured partition behavior at all (§3), and at batch level it currently contributes nothing (§5) — whether it could contribute at ingot level is unanswerable until its smelting behavior is measured. *Fifth*, release of the field's per-ingot analytical compilations: batch means are what the published record supports, and part I showed how much resolution batching destroys; the per-ingot rebuild is designed and waiting. (Portugal once belonged on this list; the database build of §2 measured it, and the model excluded it. The agenda shrinks as it should: by measurement.)

None of these require new method. The model above accepts a new reference the day it is published: a Serbian or Taurus trace panel that matches the cargo's chemistry would immediately and properly restructure this posterior, and the framework would report it without amendment. That is what it is for.

Reproducibility

Every number in this paper, including every entry in every table and every figure, is generated by a script in the public repository (Reid, 2026a), from the reference database build through the sensitivity batteries (`scripts/phase8*.py`); the repository records the provenance and license of every data source and ships the redistributable subset as CSV. The conventions are part I's: no number appears in the text that a reader cannot regenerate.

References

Daniel Berger, Elin Figueiredo, Gerhard Brügmann, and Ernst Pernicka. Tin isotope fractionation during experimental cassiterite smelting and its implication for the isotope composition of ancient tin. *Journal of Archaeological Science*, 92:73–86, 2018. doi: 10.1016/j.jas.2018.02.006.

Daniel Berger, Gerhard Brügmann, Ronny Friedrich, Joachim Lutz, Hans-Peter Meyer, and Ernst Pernicka. Shiny bronze in glassy matter: an inconspicuous piece of slag from the Bronze Age mining site of Mušiston (Tajikistan) and its significance for the development of tin metallurgy in Central Asia. *Archaeological and Anthropological Sciences*, 14:150, 2022.

Daniel Berger, Kai Kaniuth, Gerhard Brügmann, and Ernst Pernicka. Why Central Asia's Mushiston is not a source for the Late Bronze Age tin ingots from the Uluburun shipwreck. *Frontiers in Earth Science*, 11:1211478, 2023. doi: 10.3389/feart.2023.1211478.

Dmitry Konopelko et al. The Mushiston Sn deposit in Tajik Tien Shan as the type locality for stannite–cassiterite–hydrostannate mineralization: new mineral chemistry data and genetic constraints. *Journal of Geochemical Exploration*, 239:107017, 2022. doi: 10.1016/j.gexplo.2022.107017.

Ryan Mathur, Wayne Powell, Junming Yao, Frederico Guimaraes, Yanbo Cheng, Linda Godfrey, Fernando Tornos, David Killick, Jay Stephens, Jingwen Mao, Mingguang Sun, and Bernd Lehmann. Global Sn isotope compositions of cassiterite identify the magmatic–hydrothermal evolution of tin ore systems. *Geosciences*, 15(1):28, 2025. doi: 10.3390/geosciences15010028.

Wayne Powell, Michael Frachetti, Cemal Pulak, et al. Tin from Uluburun shipwreck shows small-scale commodity exchange fueled continental tin supply across Late Bronze Age Eurasia. *Science Advances*, 8(30):eabq3766, 2022. doi: 10.1126/sciadv.abq3766.

Cemal Pulak. The Uluburun shipwreck: an overview. *International Journal of Nautical Archaeology*, 27(3):188–224, 1998.

Conor Reid. iddin-annakam: a Bayesian mixture-model and identifiability reanalysis of the Uluburun tin cargo. Analysis repository: `https://git.sr.ht/~calgacus/iddin-annakam`, 2026a. All model code, parsers, and per-result writeups referenced in this draft; archival DOI to be minted at submission.

Conor Reid. Identifiable or not? A Bayesian reanalysis of the Uluburun tin cargo, 2026b. Part I of this series. Working paper, `https://calgacus.srht.site/iddin-annakam/`.

Alan Williams, Daniel Berger, et al. From Land's End to the Levant: did Britain's tin sources transform the Bronze Age in Europe and the Mediterranean? *Antiquity*, 99(405):708–726, 2025. doi: 10.15184/aqy.2025.41.