



An open global registry of primary-aluminium smelters and a derived gridded spatial prior for CF₄ (PFC-14) emissions

Zach Dissington¹

¹Independent researcher, ORCID: 0009-0007-0196-1371

Correspondence: Zach Dissington (zachdissington@gmail.com)

Abstract. CF₄ (PFC-14, tetrafluoromethane) is a long-lived greenhouse gas with a 100-year global warming potential of 7,380 and an atmospheric lifetime of 50,000 years (IPCC, 2021, AR6 WG1, Table 7.SM.7); its dominant anthropogenic source is the anode effect in primary-aluminium smelting. The smelters are a finite, public, mappable set of industrial facilities, yet the gridded inventories used as inversion priors distribute CF₄ spatially by population and built-up-area proxies. This paper describes an open registry of the world's primary-aluminium smelters operating in 2020 – 94 facilities across 36 countries (95 registry rows), with coordinates, nameplate capacity, operating status, and per-row source notes – and a set of gridded relative-weight fields derived from it: capacity-weighted, presence-only, and production-rescaled variants at 1° globally and 0.1° over Europe, formatted for use as the aluminium-sector component of a CF₄ prior in atmospheric inversions and inventory disaggregation. Registry capacity tracks USGS national production within roughly 25 % in every listed producing country except China, where 12 registered cluster-anchor plants stand for a fleet of roughly 120 and the production-rescaled variant rescales the national total (not the within-country placement). Evaluated against an observation-driven, flat-prior ICOS 2020 European CF₄ posterior ensemble, the smelter field correlates with the posterior better than the EDGAR v8.0 baseline in Iceland – 0.25 versus 0.00, significant under spatial tests in both (single-system, RHIME) ensemble members that cover Iceland – and directionally in five of six members over the pooled smelter countries, while France significantly favours the baseline in the same two members and Norway, Spain, and the UK are EDGAR-favoured; EDGAR assigns exactly zero CF₄ to 15 of the 22 European smelter cells in the analysis domain. The fields are relative weights for one sector, not emissions estimates, and Sect. 6 states the workflow and cautions required to use them. The dataset is archived at <https://doi.org/10.5281/zenodo.20673973> (Dissington, 2026a) under CC-BY-4.0.

1 Introduction

Fluorinated gases are the least spatially resolved corner of global emissions monitoring. The recent Climate TRACE methodology paper (Lancellotti et al., 2025) reports fluorinated gases at 2.4 % of the global total and places them under an “implicit estimation approach”, used “when there is insufficient information to reliably estimate facility-level emissions or to disaggregate coarser resolution emissions to individual assets” (their Supplementary Note 1.3); spatially uncertain emissions in that framework are distributed “using sector-specific proxy data (e.g., population, nightlight intensity, industrial activity)” (Supple-



mentary Note 1.2). The gridded inventories most widely used as inversion priors, EDGAR and GAINS, likewise allocate F-gas emissions with population and built-up proxies (Crippa et al., 2024).

For most F-gases this is a pragmatic compromise. For CF_4 it is wrong in a specific, correctable way. CF_4 is emitted almost exclusively during anode effects in primary-aluminium electrolysis, with a secondary contribution from semiconductor and electronics manufacturing; published attributions place the aluminium share at roughly 60–80 % of global emissions (Kim et al., 2021; Crippa et al., 2024), with the most recent global inversion at the top of that range (81 %; Püschel et al., 2025). The world’s primary-aluminium smelters number roughly 200 – the 94 registered here plus on the order of 110 further Chinese plants consolidated into registered clusters (Sect. 2.3) – and many sit far from population centres: beside hydropower in Iceland and Norway, on Siberian rivers, in western China’s coal and desert provinces. A population proxy therefore does not merely blur CF_4 ’s spatial distribution; it relocates it away from the true sources. In the evaluation below, the EDGAR field assigns exactly zero CF_4 to 15 of the 22 European smelter cells on the analysis grid, including all three Icelandic smelter cells (Sect. 4.3; the count is grid-dependent in one cell, Sect. 5.3).

Facility-located priors for these gases have precedent. Kim et al. (2021) hand-assembled the locations of 127 Chinese aluminium smelters (identified in 2011), together with semiconductor facilities, as the point-source member of a three-member prior ensemble for an East Asian inversion – describing it as potentially the most realistic spatial representation while cautioning that inaccurate prior locations can also harm an inversion. That registry was built for one study and is not available as a published dataset; recovering whatever subset of it remains current is part of the planned China completion (Sect. 2.3). What the community lacks is an open, documented, citable version of that input: a smelter registry with stated provenance and ready-gridded fields that any inversion group can evaluate, plug in as a sensitivity member, or cite as a location layer. The cost of the missing input is visible in current work: the referee discussion of the Püschel et al. (2025) global $\text{CF}_4/\text{C}_2\text{F}_6$ inversion flagged its fixed sectoral-attribution fractions as a possible bias source, and the authors’ own attribution checks rest on informal co-location with “known industry locations” – exactly the step a documented location layer makes citable.

Open facility data for the aluminium sector exists but does not fill this role. Climate TRACE publishes facility-level aluminium data (CC-BY) as part of its emissions inventory, but as modelled CO_2 -equivalent estimates per asset, without nameplate capacity, operating-status vintage, or per-row source documentation of the locations themselves; Global Energy Monitor’s heavy-industry trackers cover iron and steel, iron-ore mines, cement, and chemicals, with no aluminium tracker (its wiki’s “smelter” entries are the captive power stations co-located with some smelters, recorded as power assets); and the Oxford GeoAsset databases cover cement, steel, and petrochemicals, not aluminium. No existing open dataset provides documented smelter locations together with capacity and operating status – the registry described here does, while remaining a small, imperfect compilation whose construction, accuracy, and gaps Sects. 2 and 5 describe in detail.

This paper describes the registry (Sect. 2), the derivation of the gridded fields (Sect. 3), their evaluation against an observation-driven posterior (Sect. 4), the uncertainty in both registry and evaluation (Sect. 5), and the workflow and cautions for actually using the fields (Sect. 6).



2 The smelter registry

2.1 Sources and compilation

60 The registry (`smelters_global.csv`, 95 rows: 94 smelters operating in 2020, plus one European plant recorded as idle where compilation encountered it – the registry does not attempt a comprehensive idle-plant census) was compiled from public sources, with a deliberately simple hierarchy. The backbone is the public list of primary-aluminium smelters maintained on Wikipedia, cross-checked facility-by-facility against operator pages (Alcoa, Hydro, Rio Tinto, Trimet, Rusal, Chalco and others) for existence, ownership, and 2020 status. Coordinates were taken from Global Energy Monitor wiki pages where
 65 available (4 facilities – the captive-power-station entries co-sited with those smelters, Sect. 1), otherwise from public coordinate gazetteers or, for some European plants, the centroid of the host town; positional accuracy is discussed in Sect. 5.4. Nameplate capacities are as reported by the Wikipedia list and operator pages, of mixed vintage, and should be read as approximate (Sect. 5.4); national totals are anchored against the U.S. Geological Survey (USGS) Mineral Commodity Summaries production statistics (Sect. 2.3) rather than against the per-row capacities. Each row carries a free-text source note recording where its
 70 location and attributes came from. The European subset was additionally audited cell-by-cell against the evaluation domain of Sect. 4, because the 0.1° European fields demand the most of the coordinates.

This is a compilation of publicly documented facts about large, well-known industrial facilities, not a remote-sensing product: its value is consolidation and documentation, and its error modes are those of its sources – capacity vintage, status ambiguity in 2020 (Sect. 2.4), and coordinate granularity.

75 2.2 Schema

Each row carries: facility name, ISO-3 country code, latitude, longitude, nameplate capacity (kt yr^{-1}), 2020 operating status (operating/idle), and the source note. A technology-class field is present in the schema but unpopulated in this release: per-facility cell-technology assignments from a documented source were not obtainable openly at compilation time, and the column ships empty rather than filled from memory or inference. Consequences for use are discussed in Sects. 5.4 and 6.

80 2.3 Coverage and completeness

Figure 1 maps the registry. Coverage is documented per country against USGS national primary-aluminium production for 2020 (U.S. Geological Survey, 2021), the independent anchor (Table 1).

Registry capacity tracks national production within roughly $\pm 25\%$ everywhere except China. China is the disclosed structural gap: the 12 Chinese rows are named anchor plants for the country's major smelting clusters (Shandong/Weiqiao, Xinjiang, Inner Mongolia, Henan, Yunnan and others), standing for a fleet of roughly 120 plants; they carry 27.5 % of raw registry capacity against China's 56.7 % share of 2020 world production (the ratio column compares like with like; the capacity-vs-production mismatch is the point being measured). China's production share happens to coincide numerically with its estimated share of global CF_4 emissions ($\sim 56\%$, Sect. 4.4) – expected, since aluminium dominates the source, but the two figures are different



Table 1. Registry coverage per country against USGS 2020 estimated national primary-aluminium production (U.S. Geological Survey, 2021). Capacity exceeding production is expected, since nameplate capacity is not fully utilized.

Country (USGS-listed)	Rows	Registry capacity (kt yr ⁻¹)	USGS 2020e production (kt)	Ratio
China	12	11,765	37,000	0.32
India	8	4,463	3,600	1.24
Russia	10	4,064	3,600	1.13
Canada	9	3,127	3,100	1.01
UAE	2	2,540	2,600	0.98
Australia	4	1,698	1,600	1.06
Bahrain	1	1,540	1,500	1.03
Norway	7	1,418	1,400	1.01
Iceland	3	863	840	1.03
United States	4	847	1,000	0.85
All other countries (26)	34	10,497	9,000 (“Other”)	1.17
World	94	42,822	65,200 ^a	0.66

^a USGS reports the world total independently rounded; the listed entries sum to 65,240 kt.

quantities. The production-rescaled variant (Sect. 3) corrects this at the country level; within-China placement still rests on the 12 anchors, and Sect. 5.4 treats the consequences. Completing the per-plant China registry from Chinese-language sources is the highest-value extension of this dataset and is planned as a versioned update (Sect. 8).

2.4 Operating status in 2020

A facility is marked *operating* if public reporting indicates production during 2020; plants curtailed mid-2020 are retained as operating. The criterion is coarse – 2020 saw pandemic-related curtailments, and capacity utilization is not recorded – which is one reason the presence-only variant (Sect. 3) exists as a robustness alternative insensitive to capacity and utilization error.

3 From registry to gridded fields

The construction is deliberately simple and fully reproducible from the published code.

Weighting, three variants. (i) *Capacity*: each operating smelter weighted by nameplate capacity – the default where the registry is complete. (ii) *Presence-only*: each operating smelter weighted equally, robust to capacity error. (iii) *Production-rescaled*: country totals matched to USGS 2020 national production (USGS-listed countries directly; the 9,000 kt “Other” pool distributed across remaining registry countries by capacity), then distributed within each country by capacity. This is the recommended variant for global use because it corrects the China under-weight at the national level while preserving within-country placement (Fig. 2).



Gridding and scale convention. Smelter weights are rasterized onto the target grid: 0.1° for the European fields, and 1° globally, the latter matching the grid of the deposited emission fields of Püschel et al. (2025) so the two can be compared without regridding (their inversion itself runs on a variable-resolution internal grid; the 1° match is to the deposited product). Three additional files re-format the global variants with CF-1.8 metadata conventions. Every field is a relative weight: cell values are proportional to registered capacity (or rescaled production) and carry kt-equivalent magnitudes for interpretability, with `units = "1"` and an explicit rescale-before-use note in the file metadata. The absolute scale is arbitrary by construction – users renormalize to unit sum and multiply by their own sectoral total (Sect. 6) – and the fields must never be read as emission fluxes.

Validity window. The registry is a 2020 snapshot. The fields are most defensible for analyses within a few years of 2020; for years before ~ 2015 or after ~ 2022 the smelter fleet differs materially (Chinese build-out and relocation, European curtailments and closures), and a vintage-matched registry version – not this one – is the right input. The evaluation year (2020) coincides with the registry vintage, so the results of Sect. 4 are unaffected by this caveat.

Sectoral scope. The fields cover the aluminium sector alone, roughly 60–80 % of global CF_4 (Sect. 1). They are designed to be combined with an electronics/other component, or used alone as a documented aluminium-sector location layer for sectoral-attribution checks – for example, the registry records that South Korea and Taiwan host no primary smelters, so CF_4 inferred there cannot be aluminium; the point is not that practitioners do not know this, but that an attribution analysis can now cite a documented source for it rather than informal knowledge. The same smelters emit C_2F_6 from the same anode effects, so the registry doubles as the location layer for an aluminium-sector C_2F_6 prior; the gridded weights apply unchanged, with the caveats that C_2F_6 's aluminium share of global emissions is smaller than CF_4 's and its absolute fluxes are far smaller, so the European observational evaluation of Sect. 4 should not be assumed to transfer to C_2F_6 .

4 Evaluation against an observation-driven posterior

4.1 Design

The evaluation asks whether the smelter field predicts where an observation-driven inversion places CF_4 better than the population/built-up proxy does. The reference is the Integrated Carbon Observation System (ICOS) 2020 European CF_4 posterior from the PARIS project (Ganesan et al., 2026, ensemble methods described for HFCs by De Longueville et al., 2026): six members spanning three inversion systems (ELRIS, InTEM, RHIME), each run with two transport models (FLEXPART, NAME), all from a spatially uniform (“flat”) prior. The flat prior removes circularity with respect to both candidate fields – the posterior owes its spatial structure to neither – though not the influence of the observation network itself, whose footprint shapes where any inversion can place mass (Sect. 5.2).

Candidates are the smelter capacity field (designated in advance of the evaluation as the validation candidate; the production-rescaled variant is indistinguishable over Europe, where the rescaling barely acts) and EDGAR v8.0 CF_4 for 2020 (Crippa et al., 2023, 2024). Both are aggregated conservatively onto the posterior's native grid (0.234° latitude $\times$ 0.352° longitude). Regions are countries, defined per member as the cells whose country fraction is at least one half; the pooled domain is the



union of those cells over the smelter countries reported in Sect. 4.2; a “smelter cell” is a grid cell containing at least one registered operating smelter (the 27 operating European smelters, mapped with the 0.1° field in Fig. 3, occupy 22 cells of the analysis grid that fall inside the covered country domains). The exact European input is the cell-audited European subset
 140 `smelters_europe_killtest.csv` (Sect. 2.1), included in the archived MIT-licensed code (Sect. 9) so that this section is reproducible; its per-row capacities differ negligibly from the global registry’s European rows, and because the presence-only variant performs near-identically (Sect. 4.2), no result here depends on those differences. Comparison is by per-region Pearson correlation with the posterior, with three guards against spatial-autocorrelation artifacts: a fixed-tile spatial block bootstrap on the correlation difference (tile sides of 10 and 25 cells; reported only where the domain supplies enough tiles), a toroidal-shift
 145 permutation null for each candidate’s own correlation (a deliberately weak null – random translations of the candidate field – testing placement against chance rather than against alternative plausible source maps), and rank (Spearman, midrank-tied) correlation reported as a transparency check. One scoping note belongs in the design rather than the results: a point-source field is zero on almost all of the domain and carries no information about the spatial ordering of the low-emission background, so full-domain rank correlations favour any smooth field by construction; the skill claim of a point-source prior concerns the
 150 placement of flux mass, which the Pearson comparison and the enrichment metric (Sect. 4.3) measure.

Iceland merits a design-level remark. Before any posterior is examined, it is the region where the two candidate geographies diverge most: three smelters dominated by one large facility (Alcoa Fjarðaál) on a sparsely populated coast, so the smelter field and a population proxy disagree there almost completely. It is the natural discriminating test case for the mechanism and is treated as such – noting that only the two RHIME members cover Iceland (the ELRIS and InTEM domains do not include it at
 155 usable coverage), so the Icelandic result, however clean, is a single-region, single-system result (Sect. 5.1).

4.2 Correlation results

Table 2 summarizes the comparison at the posterior’s native grid.

The six members are three inversion systems $\times$ two transport models assimilating the same observation network – at most three quasi-independent results, not six (Sect. 5.1) – so the 5-of-6 directional consistency is corroboration, not replication
 160 (the dissenting member, one InTEM transport variant, sits at -0.012). Midrank Spearman correlations behave as the design paragraph predicts: over the pooled domain EDGAR’s smooth field tracks the background ordering (mean 0.35 versus 0.03 for the smelter field), and in Iceland both candidates are near zero (0.04 versus 0.01), confirming that the Icelandic Pearson result is carried by the placement of flux mass at the smelter cells rather than by domain-wide rank agreement – which is the point-source claim in correlation form.

165 **Transfer to the shipped 1° fields.** On the deposit’s 1° global grid the same pattern holds: Iceland 0.554 (smelter field) against -0.003 (EDGAR), with the shift-null passed ($p = 0.015$ – 0.027); pooled 0.195 (production variant 0.183) against 0.108, ahead in five of six members; Germany 0.504 against 0.189 (the German 1° domain supplies too few spatial tiles for a meaningful bootstrap, so this is reported descriptively); France 0.390 against 0.416, EDGAR ahead; Norway, Spain, and the UK remain EDGAR-favoured. The 27 European smelters occupy 25 distinct 1° cells (versus 22 in-domain cells on the finer analysis grid,
 170 where coverage masks exclude some), and all 25 are present in the shipped global field.



Table 2. Per-region spatial correlation with the flat-prior ICOS 2020 CF₄ posterior at its native grid ($0.234^\circ \times 0.352^\circ$), smelter capacity field versus EDGAR v8.0. Values are ensemble means unless noted.

Region	Smelter field	EDGAR	Reading
Iceland	0.25	0.00	Difference CI excludes zero at both tile sizes in both covering members; smelter field beats its shift-null ($p = 0.011$ – 0.016), EDGAR does not ($p = 0.09$ – 0.13). Two members only, both RHIME.
Pooled smelter countries	0.051	0.022	Direction favours the smelter field in 5 of 6 members (-0.012 to $+0.100$); no member individually significant; both correlations small in absolute terms, so the claim is the consistent direction of the difference, not the ratio. Presence-only variant 0.044 (near-identical): the result does not depend on capacity weighting.
Germany	0.06	0.015	Direction favours the smelter field in 5 of 6 members; never individually significant.
Norway, Spain, UK	~ 0	0.01–0.09	EDGAR ahead. Norway is a counterexample to any simple story (Sect. 5.1).
France	–	–	Only individually significant results favour EDGAR (both RHIME members with fully negative difference CIs, and one InTEM member at the larger tile size); the other InTEM member favours the smelter field strongly ($+0.37$) but not significantly. A system-dependent disagreement.

4.3 Enrichment at smelter cells

A complementary, distribution-free metric asks at what percentile of the pooled domain the smelter cells sit, for each field. All fields are compared on a common per-cell-mass basis on the posterior grid (densities multiplied by cell area; EDGAR aggregated conservatively, not point-sampled), with midrank tie handling: a cell tied with many others – for example one of
175 EDGAR’s many exact-zero cells – is ranked at the middle of its tie block rather than below it. Percentiles are averaged over the covering members.

Across the 22 in-domain smelter cells: the posterior places them at a mean 72.1st percentile; EDGAR places them at 59.3 and assigns exactly zero CF₄ to 15 of the 22, including all three Icelandic smelter cells (the zero cells sit at the midrank of EDGAR’s zero block, near the 41st percentile, which is what keeps EDGAR’s mean as high as 59). The inversion’s own flat
180 prior places the same cells at 76.8 – above the posterior – a number that mostly reflects field construction (land masking and cell area, since a uniform field has no source information) and shows that these percentile values partly measure construction rather than observation. With a 3×3 -neighbourhood allowance for the inversion smearing point mass into adjacent cells, the posterior figure rises to 81.9.

Both complications that follow are visible to anyone who runs the published code. The flat-prior comparison implies the
185 inversion’s absolute flux at most smelter cells goes down from the flat start: the flat prior’s overall level begins high and the observations cut the field almost everywhere, so the signed posterior-minus-prior increment is negative at most smelter



cells, and an increment-based reading would conclude the observations removed mass there. That reading conflates level with shape, and what a relative prior claims is shape. The prior-level-independent measure – the posterior/prior ratio, i.e. where the observations cut least and kept most – places the smelter cells at the 69.8th percentile of the domain, and the strongest single case, Alcoa Fjarðaál, at the 98.9th (99.4th in posterior mass). The defensible summary is deliberately modest: on a common basis the observation-driven posterior ranks smelter cells higher than the population proxy does (72.1 versus 59.3), the proxy zeroes most of them outright, and where the observations are strong the concentration is dramatic (Fjarðaál); because the percentile metrics partly reflect construction, the correlation analysis of Sect. 4.2, not this metric, carries the paper’s evaluation claims.

4.4 What the evaluation does and does not establish

The evaluation establishes that, over Europe in 2020, a smelter-location field is a better predictor of an observation-driven CF_4 posterior’s spatial structure than the population proxy in the region built to discriminate the two (Iceland, significantly, in the covering members) and directionally over the pooled smelter countries in five of six ensemble members, and that the population proxy assigns exactly zero CF_4 to most European smelter cells. Those are the established results. The boundary is equally clear: the smelter field does not win in weakly constrained or coastal-complex domains (Norway, Spain, UK, and France’s significant RHIME results), so the evaluation does not establish that this field should replace EDGAR as an operational European prior. And it cannot test the decisive domain: China holds roughly 56 % of global CF_4 emissions (2018–2023; Püschel et al., 2025) – a share whose rapid recent rise is attributed substantially to western-China aluminium (An et al., 2024) – and has no public observation-driven gridded posterior; the one public global deposit uses an EDGAR-based prior and cannot adjudicate between these candidates. The recommended use is as a documented sensitivity member alongside the standard prior (Sect. 6); the China test – a prior-sensitivity inversion or OSSE over East Asia – is the natural next experiment this dataset exists to make cheap, while the registry, the European evaluation, and the sectoral location layer stand on their own.

5 Uncertainty evaluation

5.1 Statistical uncertainty of the evaluation

The claim inventory, stated symmetrically. *Individually significant for the smelter field*: Iceland only – and Iceland is covered by only the two RHIME members, which share one inversion system, so the effective number of independent confirmations is close to one; the spatial tests (bootstrap, shift-null) quantify within-member spatial sampling uncertainty, not between-system structural uncertainty, which for Iceland is unsampled. *Individually significant against the smelter field*: France, in the same two RHIME members (and one InTEM member at the larger tile size). *Directionally favoured in five of six members, never individually significant*: the pooled domain and Germany (a different single member dissents in each). *EDGAR-favoured*: Norway, Spain, UK. Norway deserves emphasis as the counterexample to a clean mechanism story: Norwegian smelters are remote and EDGAR zeroes them, yet EDGAR wins the Norwegian correlation – the enrichment and correlation metrics disagree there,



plausibly because the coastal smelter cells border large weakly constrained interior and offshore areas, but no mechanism has been demonstrated and the loss stands as reported.

220 No multiplicity correction is applied, and the Icelandic shift-null p -values (0.011–0.016) would not survive, for example, a Bonferroni correction across the seven reported regions. The design defence is stated in Sect. 4.1 – Iceland is the ex-ante discriminating case for this mechanism, not a region discovered by scanning – and the per-region results are reported in full, wins and losses, so the reader can apply any correction they prefer. The bootstrap is a fixed-tile scheme whose tile counts are reported context: ~ 30 tiles (side 10) and 8 tiles (side 25) for Iceland – the side-25 interval is a sensitivity check on a small
 225 tile population, not an independent confirmation – and too few tiles at 1° in Germany for the procedure to be meaningful at all (dropped accordingly). Two further robustness items, both from a code audit of the published harness. First, the original year selection (`sel(time, method="nearest")`) silently returned the 2019 posterior field for the four ELRIS/InTEM members (their mid-year time stamps are equidistant from the start of 2020 and the tie resolves to the earlier index); all numbers in this paper use exact-2020 fields selected by slice with an assertion, and the Icelandic results are unaffected in any case
 230 because both covering members carry unambiguous 2020 stamps. Second, the native-grid correlations compare candidate per-cell masses against the posterior's flux density, whose cell area varies with latitude; recomputing with the posterior converted to per-cell mass changes no conclusion (pooled 0.045 versus the reported 0.051; Iceland 0.254, still significant at $p = 0.005$ –0.011 with the difference CIs excluding zero; all directional counts and losses identical – `audit_basis.py` in the code repository).

235 5.2 What the flat-prior reference can and cannot settle

The flat prior removes prior-induced circularity: the posterior's structure owes nothing to either candidate. It does not make the posterior truth. Posterior structure can only appear where the observing network has sensitivity; mass is smeared along footprint gradients; and weakly constrained cells retain near-prior values that dilute all correlations. Some of the smelter field's advantage could in principle reflect coincidental alignment with where the network can place mass rather than where sources
 240 are. The Fjarðaál case is the strongest counter-evidence – a 99th-percentile posterior concentration exactly at the one large facility on that coast, in both covering members – but a footprint-sensitivity diagnostic that would settle this quantitatively is not included in this version, and the Icelandic result should be read with that openly in mind.

5.3 Uncertainty of the enrichment metric

Reported in Sect. 4.3 alongside the result rather than separately: the metric's percentile values depend on basis (mass, with
 245 midrank ties, conservatively aggregated – the choices here), on tie handling (EDGAR's zero cells sit at the midrank of its zero mass), and on field-construction effects visible in the flat-prior column (76.8 for a field with no source information). The posterior/prior ratio at the 69.8th percentile is the construction-independent residue. EDGAR's zero count (15 of 22) is robust to these choices in substance though not in exact count: at EDGAR's native 0.1° the count is 15 with one Icelandic cell nonzero (the ISAL smelter borders the Reykjavík conurbation, where the population proxy happens to place mass), and on the analysis
 250 grid all three Icelandic smelter cells are zero.



5.4 Uncertainty of the registry and fields

Coordinates. Mixed granularity: facility-accurate where GEM or imagery-verifiable sources were available, town-centroid for some European rows. Adequate by construction for the 0.234° – 1° analysis grids and for the 0.1° European fields (the European subset was audited against the evaluation domain cell-by-cell); not adequate for sub- 0.1° use. *Capacities.* Nameplate values of mixed vintage from public lists and operator pages; per-row error of tens of percent is plausible. The country-level anchor (Table 1) bounds the aggregate error; the presence-only variant exists to bound sensitivity to capacity error entirely – over the European evaluation the capacity and presence variants perform near-identically, so no result in Sect. 4 depends on capacity accuracy. *Technology.* Unpopulated (Sect. 2.2). Capacity weighting therefore embeds an implicit uniform emission-per-tonne assumption across technologies, which is known to be false in magnitude terms – anode-effect emission factors differ strongly across cell technologies and operator practice (IAI anode-effect survey data) – and is tolerable here only because the fields are relative spatial weights whose magnitudes the user rescales; a technology-resolved variant is the planned use of the empty column once a documented per-facility technology source is secured. *China placement.* Within-China structure rests on 12 anchor plants for ~ 120 ; the production-rescaled variant fixes the national total but cannot add unregistered within-country structure. Users whose domain includes China should treat within-China placement as indicative only. *Vintage.* 2020 snapshot; see the validity window in Sect. 3.

6 Using the fields: workflow and cautions

The fields are one ingredient of a prior, not a prior. To run an inversion with them a user must supply: (1) an absolute sectoral total for the domain and year – e.g. national or global aluminium-sector CF_4 from a published attribution (Püschel et al., 2025; Kim et al., 2021) or from production statistics (U.S. Geological Survey, 2021) combined with IPCC Tier-1/2 anode-effect emission factors – by which the weight field, renormalized to unit sum, is multiplied; (2) the other sectors, principally electronics, as a separate component (this dataset deliberately does not invent one); (3) a prior uncertainty structure, which the weight field does not carry; and (4) a temporal profile if sub-annual structure matters – the fields are annual and time-invariant.

One caution is mandatory rather than advisory. The fields are exactly zero outside smelter cells. Used directly as the mean of a log-normal prior they are undefined at zero cells, and used with a Gaussian prior with small variance they forbid the inversion from placing mass anywhere the registry lacks a facility – including all electronics sources and the unregistered Chinese plants. Any total- CF_4 use must therefore blend the smelter component with a smooth background component (and/or apply a floor), with the blend weights set by the user’s sectoral split. Shipping the fields un-floored is deliberate – a floor encodes a sectoral assumption that belongs to the user – but the consequence is that the raw fields are an aluminium-sector shape, unusable alone as a total- CF_4 prior, and this is stated here once in bold terms so it cannot be missed: **do not use these fields as a standalone total- CF_4 prior.**

The lowest-risk uses, in increasing order of commitment: a citable location layer for sectoral-attribution and co-location checks; a sensitivity member run alongside the standard EDGAR-based prior to test how much posterior structure the prior choice drives; and the aluminium component of a composite prior in domains where the registry is strong (Table 1).



7 Conclusions

CF₄ is the F-gas for which facility-resolved spatial information is cheapest to obtain and most consequential to use: its sources are a few hundred large, public, documented industrial facilities, and the standard population proxy demonstrably misplaces it – zeroing most European smelter cells outright. This paper turns that observation into a reusable input: an open, provenance-documented global smelter registry and derived multi-variant gridded weight fields, evaluated against an observation-driven European posterior with wins and losses reported to the same standard. The evaluation supports the premise in the region built to discriminate the two candidate geographies (Iceland) and directionally across the pooled smelter countries, against it in France, and the recommended use is calibrated accordingly: a citable location layer and a prior-sensitivity member today, a component of operational priors where it earns that role. The decisive test is East Asia, and the data, code, and formats published here are designed to make that test a small effort for any group equipped to run it.

8 Data availability

The dataset is archived on Zenodo under CC-BY-4.0: version DOI <https://doi.org/10.5281/zenodo.20673973> (Dissington, 2026a); concept DOI <https://doi.org/10.5281/zenodo.20617485> resolves to the latest version. Registry coordinates are WGS84 (EPSG:4326) and all gridded fields are on regular latitude–longitude grids; the native weight fields carry the data variable `prior_weight`, and the CF-1.8 FLEXINVERT variants carry `emis_prior`. Twelve files: the registry (`smelters_global.csv`), provenance notes (`SOURCES.md`), three global 1° weight-field variants and two European 0.1° variants (NetCDF, relative weights, units “1”), three CF-1.8-formatted global variants on the grid of the Püschel et al. (2025) deposited fields, a per-file manifest (`README.txt`), and a one-page method note (PDF). Versions are released under the concept DOI; the registry is intended to be refreshed against each year’s USGS Mineral Commodity Summaries, with the per-plant China completion as the highest-priority planned update; corrections are accepted as issues or pull requests on the code repository. The complete provenance documentation and build code mean any researcher can reproduce, correct, or take over the dataset – for a single-maintainer dataset, that reproducibility is the substantive persistence guarantee.

9 Code availability

All construction and evaluation code is open (MIT) at <https://github.com/zachdissington/climate-solutions-research-kit> (`priors/pfc-aluminium-prior/src/registry-to-grid` builders for each variant and resolution, the EDGAR loader, the posterior fetch, the correlation and significance harness, and the CF re-formatting utility. The v1.2.0 tagged release is archived on Zenodo with its own DOI (Dissington, 2026b): <https://doi.org/10.5281/zenodo.20683507> (concept DOI; MIT). The European evaluation registry `smelters_europe_killtest.csv` (Sect. 4) is included in this archived code.

Author contributions. ZD compiled the registry, built and evaluated the gridded fields, and wrote the paper.



Competing interests. The author declares no competing interests.

Acknowledgements. This work uses the ICOS PARIS European inverse-modelling results (Ganesan et al., 2026) and the EDGAR Community
315 GHG Database (EDGAR F-GASES v8.0, European Commission JRC/IEA). The deposited global $\text{CF}_4/\text{C}_2\text{F}_6$ emission fields of Püschel et al.
(2025), which carry a CC-BY-NC licence, are used here only as a comparison reference in the 1° transfer evaluation (Sect. 4.2) and are
not redistributed; no value derived from them enters the deposited fields, which rest entirely on the open registry. The author thanks the
PARIS inversion teams for publishing flat-prior ensemble posteriors – the design choice that makes independent evaluations like this one
possible. During the preparation of this work the author used AI-based assistants (Anthropic’s Claude) to support manuscript drafting, code
320 development, and the validation analysis. The author reviewed and edited all content and takes full responsibility for the work.

Financial support. This research received no external funding.



References

- An, M., Prinn, R. G., Western, L. M., Yao, B., Zhao, X., Kim, J., Mühle, J., Chi, W., Harth, C. M., Hu, J., Ganesan, A. L., and Rigby, M.: Substantial increase in perfluorocarbons CF₄ (PFC-14) and C₂F₆ (PFC-116) emissions in China, *P. Natl. Acad. Sci. USA*, 121, e2400168 121, <https://doi.org/10.1073/pnas.2400168121>, 2024.
- Crippa, M., Guizzardi, D., Pagani, F., Banja, M., Muntean, M., Schaaf, E., Becker, W., Monforti-Ferrario, F., Grassi, G., Rossi, S., Brandao De Melo, J., Jacome Felix Oom, D., Branco, A., San-Miguel, J., and Vignati, E.: EDGAR v8.0 Greenhouse Gas Emissions, European Commission, Joint Research Centre (JRC) [data set], <https://doi.org/10.2905/b54d8149-2864-4fb9-96b9-5fd3a020c224>, 2023.
- Crippa, M., Guizzardi, D., Pagani, F., Schiavina, M., Melchiorri, M., Pisoni, E., Graziosi, F., Muntean, M., Maes, J., Dijkstra, L., Van Damme, M., Clarisse, L., and Coheur, P.: Insights into the spatial distribution of global, national, and subnational greenhouse gas emissions in the Emissions Database for Global Atmospheric Research (EDGAR v8.0), *Earth Syst. Sci. Data*, 16, 2811–2830, <https://doi.org/10.5194/essd-16-2811-2024>, 2024.
- De Longueville, H., Melo, D. B., Redington, A. L., Ramsden, A., Danjou, A., Andrews, P., Pitt, J., Murphy, B., Constantin, L., Stanley, K. M., O'Doherty, S., Wenger, A., Young, D., Engel, A., Schuck, T., Meixner, K., Wagenhäuser, T., Gad, F., Vollmer, M. K., Reimann, S., Maione, M., Arduini, J., Lunder, C., Schmidtbauer, N., Haszpra, L., Molnár, M., Frumau, A., Couret, C., Rigby, M., Henne, S., Manning, A., and Ganesan, A. L.: European HFC emissions evaluated with multiple atmospheric inverse models and UNFCCC national inventories, *Atmos. Chem. Phys.*, 26, 7647–7675, <https://doi.org/10.5194/acp-26-7647-2026>, 2026.
- Dissington, Z.: An open smelter-resolved spatial prior for CF₄ (PFC-14) emissions, Zenodo [data set], <https://doi.org/10.5281/zenodo.20673973>, 2026a.
- Dissington, Z.: Climate Solutions Research Kit: open spatial-prior method and smelter-resolved CF₄ (PFC-14) prior, Zenodo [software], <https://doi.org/10.5281/zenodo.20683507>, 2026b.
- Ganesan, A., Manning, A., Henne, S., De Longueville, H., Ramsden, A., Brito Melo, D., Danjou, A., Andrews, P., Murphy, B., Redington, A., and Pitt, J.: Inverse modelling results for European non-CO₂ greenhouse gas emissions, ICOS Carbon Portal [data set], <https://doi.org/10.18160/GR1Q-6SK4>, 2026.
- IPCC: The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity (Chapter 7; Forster, P. and others), Table 7.SM.7, in: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by Masson-Delmotte, V. et al., Cambridge University Press, <https://doi.org/10.1017/9781009157896.009>, 2021.
- Kim, J., Thompson, R., Park, H., Bogle, S., Mühle, J., Park, M.-K., Kim, Y., Harth, C. M., Salameh, P. K., Schmidt, R., Ottinger, D., Park, S., and Weiss, R. F.: Emissions of Tetrafluoromethane (CF₄) and Hexafluoroethane (C₂F₆) From East Asia: 2008 to 2019, *J. Geophys. Res.-Atmos.*, 126, e2021JD034 888, <https://doi.org/10.1029/2021JD034888>, 2021.
- Lancillotti, B. V., Malof, J. M., Davitt, A., McCormick, G., Anderson, S., Carbó-Mestre, P., Collins, G., Crane, V., Doctor, Z., Ebri, G., Foster, K., Gowdy, T. M., Guzzardi, M., Heal, J., Hunter, H., Kroodsma, D., Galib, K. M., Markakis, P. J., McDonald, G., Moore, D. P., Nguyen, E. D., Parvu, S., Pekala, M., Piatko, C. D., Piscopo, A., Powell, M., Raniga, K., Reilly, E. P., Robinette, M., Saraswat, I., Sicurello, P., Söldner-Rembold, I., Song, R., Underwood, C., and Bradbury, K.: Closing Gaps in Emissions Monitoring with Climate TRACE, *arXiv [preprint]*, <https://doi.org/10.48550/arXiv.2511.19277>, 2025.
- Püschel, B., Vojta, M., Kandler, L., Crotwell, M., Engel, A., Krummel, P. B., Lunder, C. R., Mühle, J., O'Doherty, S., Prinn, R. G., Stanley, K. M., Vimont, I., Vollmer, M. K., Wagenhäuser, T., Weiss, R. F., Young, D., and Stohl, A.: Global distribution of tetrafluoromethane (CF₄)



360 and hexafluoroethane (C₂F₆) emissions determined by inverse modeling, EGU sphere [preprint], <https://doi.org/10.5194/egusphere-2025-5656>, deposited emission fields: <https://doi.org/10.25365/phaidra.751>, 2025.
U.S. Geological Survey: Mineral Commodity Summaries 2021: Aluminum, U.S. Geological Survey, <https://doi.org/10.3133/mcs2021>, aluminum chapter; the volume DOI resolves to the full Mineral Commodity Summaries 2021, 2021.

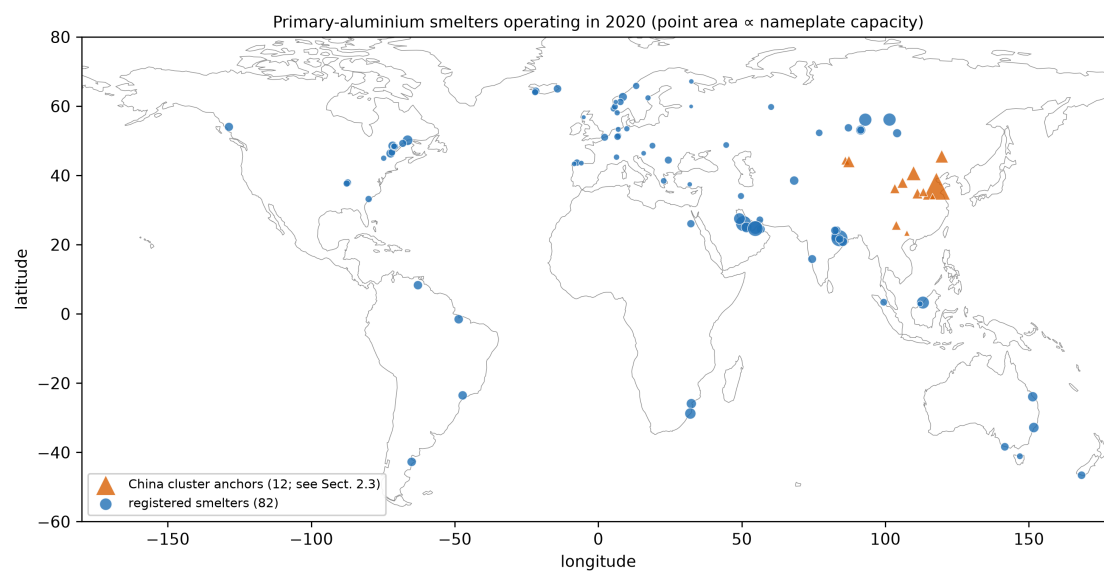


Figure 1. Primary-aluminium smelters operating in 2020 (point area proportional to nameplate capacity). Orange triangles: the 12 Chinese cluster-anchor plants (Sect. 2.3); blue circles: the 82 other registered smelters.

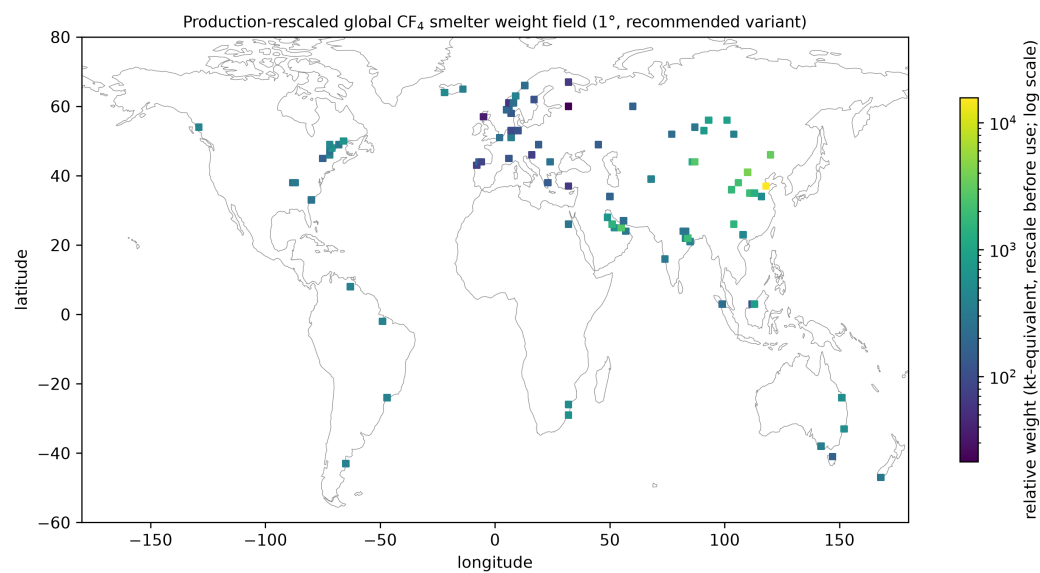


Figure 2. The production-rescaled global CF₄ smelter weight field (1°, recommended variant; log scale, relative weights to be rescaled before use).

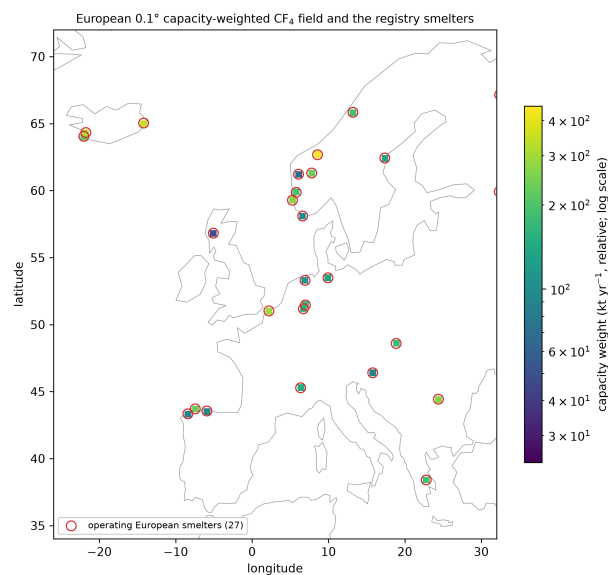


Figure 3. The European 0.1° capacity-weighted field (filled cells, log scale) and the 27 operating European registry smelters (red circles).