

Aggregated monitoring of enhanced weathering on agricultural lands

Tim Jesper Suhrhoff^{1,2}, Anu Khan³, Shuang Zhang⁴, Beck Woollen^{1,3}, Tom Reershemius⁵, Mark A. Bradford^{1,6}, Alexander Polussa^{1,6}, Ella Milliken², Peter A. Raymond⁶, Christopher T. Reinhard⁷, Noah J. Planavsky^{2,1}

1: Yale Center for Natural Carbon Capture, New Haven, CT, USA

2: Department of Earth & Planetary Sciences, Yale University, New Haven, CT, USA

3: Carbon Removal Standards Initiative, Washington D.C., USA

4: Department of Oceanography, Texas A&M University, College Station, TX, USA

5: School of Natural and Environmental Sciences, Newcastle University, Newcastle-upon-Tyne, UK

6: The Forest School, Yale School of the Environment, Yale University, New Haven, CT, USA

7: Department of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA, USA

Abstract: Terrestrial enhanced weathering (EW) on agricultural land is a promising carbon dioxide removal (CDR) pathway with high potential to scale. Enhanced weathering also has the potential to provide significant agronomic co-benefits to farmers and producers. Today, most EW field trials are funded through the voluntary carbon market (VCM) with the purpose of generating carbon removal credits for corporate sustainability goals. As a result, monitoring, reporting, and verification (MRV) frameworks for EW are designed for attribution of tons of removal via weathering to individual fields. Here, we describe approaches for aggregation of weathering indicators across multiple fields using aqueous and solid phase measurements. First, we demonstrate that larger agricultural catchments are at least as suitable as smaller ones for detecting weathering signals in river chemistry, and in some cases may even offer advantages due to lower variability and background weathering fluxes. Second, we assess quantification uncertainty from in-field solid phase soil measurements at increasing scales and show that errors in CDR quantification can be reduced by aggregating signals over many fields. Critically, we also highlight that aggregation sets must be defined in advance and all plots included, as biased selection of fields can generate apparent removal signals out of statistical noise. Taken together, we find that aggregated monitoring of EW—quantifying CDR over multiple fields at once—can both improve existing MRV frameworks and support integration of EW practices with a broader array of government policies, unlocking funding and public support to achieve climate-relevant scale.

1 Introduction

Deep and immediate emissions reductions are needed to prevent the worst harms of climate change (UNEP 2024, IPCC 2018). In addition to mitigation, there is growing consensus that atmospheric carbon dioxide removal (CDR) will be necessary to stay within Paris Agreement temperature targets (Luderer *et al* 2018, Rogelj *et al* 2018, IPCC 2022). Multiple CDR approaches will be required to achieve the gigaton-scale drawdown proposed in net-zero and net-negative climate scenarios (IPCC 2022, Geden *et al* 2024, Lamb *et al* 2024), depending on local energy, land, infrastructure, and mineral resources.

One promising CDR approach is terrestrial enhanced weathering (EW) on agricultural land. Crushed cation-rich rocks applied to fields react with dissolved atmospheric carbon in water, forming aqueous bicarbonate ions. Carbon is durably stored for 1,000s of years (Renforth and Henderson 2017) in oceans as bicarbonate or in soils and sediments as solid carbonates. Global EW removal potential is estimated at 0.5–2 Gt CO₂ yr⁻¹ (Beerling *et al* 2020), or 64–217 Gt cumulatively by 2080 (Baek *et al* 2023), meeting up to ~20% of expected CDR needs. EW can also deliver agronomic benefits, including soil pH management and improved yields (Levy *et al* 2024), increasing the likelihood of adoption in alignment with climate goals.

Growing demand for CDR, especially from voluntary carbon market (VCM) buyers, has rapidly increased investment in EW. Over 25 companies now operate globally, with ~600,000 t CO₂ credits sold, though <2% are delivered (CDR.fyi 2025). This expansion raises critical questions about the ability to accurately estimate CDR via EW. Experience shows that VCM incentives can favor low-cost, inflated claims, leading to systematic quantification failures (Gill-Wiehl *et al* 2023, Badgley *et al* 2022, Sanders-DeMott *et al* 2025). Achieving large-scale, verifiable CDR via EW will require balancing accuracy and cost, using a mix of methods (e.g., solid, and liquid phase measurements, as well as models) with context-appropriate study design, across spatial and temporal scales (Clarkson *et al* 2024, Almaraz *et al* 2022). Here, we examine the benefits of monitoring EW across multiple fields or regions—through physical aggregation of weathering products in streams and statistical aggregation of field data—to reduce uncertainty in carbon quantification, and discuss cost and scalability implications from market and policy perspectives.

2 Methods

This study evaluates the potential of aggregated monitoring—quantifying carbon removal by jointly monitoring multiple field sites—of EW on croplands. We focus on silicate rock applications and consider both in-field and downstream signals of weathering products, including cations and carbon species. To illustrate the utility of aggregation, we analyze monitoring approaches in streams and soils using simple models that capture first-order system behavior. These models are not meant to yield precise forecasts but to demonstrate how aggregation can improve detectability and reduce uncertainty in EW monitoring.

Two distinct challenges frame this analysis: (i) reducing statistical noise from spatial heterogeneity, where aggregation across multiple fields or watersheds provides a tractable solution, and (ii) addressing system-level processes that affect permanence and transport of weathering products, which require moving beyond near-field soil measurements to downstream integration in rivers or groundwater. Recent work shows that solute export reflects not only soil-scale weathering but also subsurface redox structure, mineral buffering, and hydrological residence times (Shaughnessy and Brantley 2023, Shaughnessy *et al* 2023), underscoring the need to capture processes integrating across critical zone compartments. We attempt to address (i) through analyses and simulations of water and soil datasets, while (ii) is inherently only represented in the stream-water approach, which integrates across subsurface transport, buffering, and residence times.

2.1 Watershed analysis

Enhanced weathering can be quantified by tracking weathering products such as cations or alkalinity in the aqueous phase (Clarkson *et al* 2024, Almaraz *et al* 2022, Sutherland *et al* 2024). Monitoring bicarbonate alkalinity in effluent water captures CDR after losses to secondary phase formation and other soil or upstream processes, while cation fluxes can serve as alkalinity proxies (Bijma *et al* 2025). Detecting such riverine changes against baseline conditions is challenging, yet multiple studies document measurable shifts in river chemistry from agricultural liming (Hamilton *et al* 2007, Oh and Raymond 2006, Barnes and Raymond 2009, Duan *et al* 2025), demonstrating the feasibility of this approach.

Here, we assess the potential of downstream water monitoring to estimate CDR from large-scale silicate application. Using a continental US river database (USGS 2016), we establish baseline fluxes from stations with ≥ 10 years of data since 1990 and ≥ 10 measurements per year across all seasons ($n = 95$ for alkalinity, 81 for Ca, 80 for Mg; 120 stations total). After excluding catchments with $< 10\%$ agricultural land (cropland and pasture; USGS 2024), 51 sites remained (50 for alkalinity, 26 for Ca, 26 for Mg; Figure 1a). We then calculate the concentration increase in alkalinity, Ca, and Mg required for a detectable 2σ deviation from baseline, and translate the resulting fluxes into agricultural-area-normalized basalt application rates using average US basalt

composition (Lehnert *et al* 2000), watershed area (USGS 2016), and cropland or pasture extent (USGS 2024). For alkalinity, this is calculated based on charge balance from basalt base cation content. Some watersheds are nested; not all datapoints are independent.

Our goal is not to define exact detectable application rates but to evaluate how the utility of stream-based MRV scales with catchment size. The analysis assumes steady-state conditions—rapid dissolution, transport, and downstream detection of weathering products. While these assumptions underestimate required absolute application rates (Kanzaki *et al* 2025, Kirchner and Neal 2013, Godsey and Kirchner 2014, Calabrese *et al* 2022, Power *et al* 2025), they should not vary systematically with catchment size. Hence, the trends derived here remain informative for assessing the feasibility of river-based MRV as a function of scale (see Section 4.1).

2.2 Soil analysis

An alternative approach to quantify CDR via EW is using soil mass balance approaches (Kantola *et al* 2023, Reershemius and Suhrhoff 2023, Reershemius *et al* 2023, Clarkson *et al* 2024, Suhrhoff *et al* 2024, 2025), which assess differences in feedstock concentration before and after weathering by comparing mobile cations (e.g., Ca^{2+} , Mg^{2+}) to an immobile tracer (e.g., Ti). Combined with supporting measurements and assumptions (Campbell *et al* 2023, Reershemius *et al* 2023, Clarkson *et al* 2024, Suhrhoff *et al* 2024, 2025), the resulting dissolution fraction serves as a proxy for CDR. We demonstrate that aggregating results across multiple fields substantially improves the accuracy of soil-based quantification of rock powder dissolution, consistent with findings from soil organic carbon studies (Potash *et al* 2025, Bradford *et al* 2023). This analysis tests the ability of monitoring approaches to overcome high within- and among-field variability in cation concentrations.

We use elemental composition data from agricultural soils (Smith *et al* 2013) and basalt rock powder (Lehnert *et al* 2000), supplemented by new data on in-field compositional variance from five densely sampled sites (Suhrhoff et al 2025; see supplement S2). Using these datasets, we perform Monte Carlo simulations to estimate average error in detected dissolution fractions under varying conditions: rock powder applications of 50 and 100 t ha⁻¹, dissolution fractions of 0.25 and 0.5, and sampling frequencies from 1–20 samples ha⁻¹. For each simulation, we compare calculated dissolution fractions to known inputs to quantify (1) mean absolute error for single-field MRV and (2) accuracy after averaging across 10, 50, or 100 fields. Simulations were repeated 100 times to derive average errors and the frequency of >20% overestimation. The model assumes non-paired sampling (baseline and post-weathering samples at random locations) but includes a paired-sampling comparison. Detailed workflow and assumptions are provided in Supplement S2 (see Figure S8).

3 Results

Our analysis indicates that basalt addition is more detectable (i.e., requires lower application rates) in watersheds with larger total agricultural area (Figure 2). Required basalt application rate to cause a 2σ increase compared to baseline river concentrations forms a significant trend for alkalinity ($p < 0.001$) though at low R^2 (0.25). There is no significant trend with catchment size for Ca and Mg. Excluding loss and lag effects from slow weathering or solute retention, the average basalt dissolution and transport rates ($\text{t ha}^{-1} \text{ yr}^{-1}$) for watersheds with $>1 \text{ km}^2$ of agricultural area and $>20\%$ agricultural land cover (USGS 2024) are 0.63 ± 0.68 (1σ) for alkalinity ($n = 45$), 0.97 ± 0.87 (1σ) for Ca ($n = 23$), and 0.53 ± 0.31 (1σ) for Mg ($n = 23$).

Aggregating soil mass balance results across multiple fields substantially increases the robustness of CDR estimates from EW. Robustness is measured as (i) the average absolute error between simulated and calculated dissolution fractions (τ_j) and (ii) the frequency of overestimating dissolution by $>20\%$ (Figure 3). Both metrics improve markedly with aggregation: at a sampling density of 10 samples ha^{-1} (100 t ha^{-1} application, $\tau_j = 0.25$), average error declines from $>20\%$ for single fields to $<10\%$ (10 fields) and $\approx 5\%$ (100 fields). Over-crediting frequency collapses from $>20\%$ for individual fields to near zero when data from ≥ 50 fields are aggregated.

4 Discussion

4.1 Watershed

Our analysis indicates that within the existing set of USGS stream gage stations, detectable EW signals are impacted by catchment agricultural area (Figure 2) or catchment size (Figure S1) to some extent. For alkalinity, we find a statistically significant but relatively weak relationship between required basalt application rates and total agricultural area ($p < 0.001$, $R^2 = 0.25$; Figure 2), indicating that less basalt may be required in larger watersheds for viable stream-water MRV. No significant trends are observed for Ca or Mg, suggesting that signal detectability is shaped by a combination of factors.

One likely contributor to the alkalinity trend is that weathering rates (area-normalized alkalinity fluxes) tend to be lower in large agricultural watersheds ($p < 0.05$, $R^2 = 0.09$; Figure 4a), potentially reflecting lower runoff with increasing catchment area ($p < 0.001$, $R^2 = 0.35$; Figure 4b). Runoff is a well-established control on weathering rates (White and Blum 1995, Gaillardet *et al* 1999, Gislason *et al* 2009, Hartmann 2009, West 2012). By contrast, we find no significant relationship with erosion rates ($R^2 = 0.03$, $p = 0.19$; Figure 4c). Another factor that may help explain lower required basalt application in larger agricultural catchments could be that variability in stream chemistry appears to decline with catchment size, as indicated by decreasing relative standard deviation in annual baseline data ($p < 0.001$, $R^2 = 0.21$; Figure 4d).

Taken together, these results suggest that larger agricultural watersheds are at least as suitable—and often more advantageous—than smaller ones for aggregated monitoring of EW. When new monitoring stations are located directly within agricultural catchments, watersheds of varying sizes may be similarly suited for signal detection (see Figure S7 for stations with $>50\%$ agricultural land). Although this conclusion is limited to the specific set of USGS stations analyzed, it highlights that large catchments can be promising candidates for stream-based MRV. The added advantage is that the same infrastructure can monitor broader agricultural areas, lowering MRV costs per ton (see Section 4.4) and making large catchments particularly attractive for deployment.

We use this signal-to-noise analysis to explore how area-normalized basalt application rates relate to catchment size and to compare the relative utility of different watershed contexts. We do not aim to provide definitive estimates of the absolute application rates required for signal detection. The analysis assumes steady-state conditions: basalt dissolves rapidly, and dissolution products are transported to streams where they influence river chemistry. We did not model loss processes such as secondary phase formation, cation exchange, biomass uptake, or strong-acid weathering (Clarkson *et al* 2024), nor lag times between weathering and solute export due to interactions with exchangeable acidity (Kanzaki *et al* 2025). Consequently, the absolute rates reported here underestimate true requirements, but we argue the trends remain informative for assessing the utility of river-based MRV as a function of scale. The results presented here suggest it is going to

be as feasible to accurately detect the same area-normalized application rates in catchments with large areas (and proportions) of agricultural land.

This conclusion is based on the two assumptions that neither loss processes nor lag times co-vary with catchment size. Relating to the first, in reality, larger watersheds may experience some transport limitations where solute export is constrained by longer hydrological residence times or subsurface saturation, leading to higher retention of weathering products in soils, groundwater, or riparian zones and thus a dampened downstream signal. This possibility may be reflected in slightly higher baseline concentrations in some larger catchments ($p < 0.05$, $R^2 = 0.11$; Figure 4e), potentially suggesting slower solute turnover. Such scale-dependent effects are not captured by our steady-state framework. Consistent with the second assumption, data in Figure 4f show that watersheds with greater agricultural area do not have longer average distances to the nearest river, supporting the assumption that the primary transport-limiting step—cation movement through topsoils and from topsoils to rivers (Kanzaki et al 2025)—should not increase with catchment size.

The impact of these two assumptions—and the challenges they represent—can furthermore be mitigated through deployment choices. Experimental studies (generally <6 months) have shown that cation storage in secondary phases can alter effluent water composition (Renforth *et al* 2015, Pogge von Strandmann *et al* 2022, Iff *et al* 2024, te Pas *et al* 2025, Vienne *et al* 2025), while models indicate that interactions with soil exchangeable cation pools can extend lag times beyond 30 years in high-CEC regions such as the U.S. Corn Belt (Kanzaki *et al* 2025). Because this process is largely governed by soil cation exchange capacity (CEC) and base saturation, lag times are expected to be shorter in the tropical soils of the southeastern U.S. (Kanzaki *et al* 2025). Similarly, precipitation of secondary carbonates is promoted by high pH and carbonate content, more common in western soils (Smith *et al* 2013, Wiczorek 2019), suggesting lower losses in the east. Catchments with low CEC, high base saturation, high infiltration, and short residence times are therefore most suitable for minimizing lag and losses. In practice, watersheds favorable for EW from a geochemical perspective are also those best suited for detecting downstream weathering signals, emphasizing the potential to optimize monitoring efficacy through informed site selection.

In summary, we find no evidence that larger watersheds are less effective than smaller ones for stream-based MRV at equal application rates. In many cases, they appear more suitable, offering the additional benefit of broader spatial coverage and lower per-ton monitoring costs. Provided EW signals can be reliably detected in streams—as supported by previous studies (Hamilton *et al* 2007, Oh and Raymond 2006, Barnes and Raymond 2009, Duan *et al* 2025, Shao *et al* 2016, Taylor *et al* 2021)—this approach shows strong potential for large-scale monitoring of EW.

4.2 Soil

In statistics it is well known that individual-level “noise” can obscure intervention effects, whereas representative sampling across a population yields a robust average treatment effect (Holland 1986, Rothman *et al* 2008, Rubin 1974). We refer to this as “aggregation” and suggest it is a useful

tool for estimating treatment effects and other parameters relevant to EW in agricultural systems, such as pH or base saturation. This has been well demonstrated for soil organic carbon management (Potash *et al* 2025, Bradford *et al* 2023), where it has been postulated that detection may not be reliable at a field level, but aggregation may facilitate identifying population-level trends.

We demonstrate that the same principle applies to soil-based MRV for EW: aggregating data from multiple fields reduces both the average error in detected dissolution fractions and the frequency of overestimating weathering by >20% (Figure 3). Within the level of variability modeled here, high accuracy in CDR estimates is only achievable through multi-field aggregation. As demonstrated in other soil mass balance analyses, detectability improves over time as cumulative application and weathering progress (Suhrhoff *et al* 2024). Even at higher feedstock applications (Figure 5a–b) and dissolution fractions (Figure 5c–d), aggregation reduces both error and over-crediting frequency.

Our analysis focuses primarily on unpaired sampling, representing a conservative estimate of potential errors in heterogeneous soils (Rogers and Maher 2025). For comparison, we also evaluate paired sampling (supplement S2.3). When paired sampling is implemented reliably—via high-precision GPS or permanent markers—soil mass balance approaches can become feasible at the field scale (Figure S12; Suhrhoff *et al* 2025). Importantly, aggregation can achieve comparable accuracy at larger scales without depending on the success of paired sampling.

The statistical framework used here can inform sampling protocols given target uncertainty levels. If baseline data on soil heterogeneity are available, the model can estimate expected errors as a function of sampling frequency and field aggregation, similar to approaches proposed by Rogers and Maher (2025; preprint) but in our case agnostic to what is an “acceptable” level of uncertainty. Minimum sampling requirements should be defined before deployment, and uncertainties in CDR quantification propagated, for example via Monte Carlo simulations (Derry *et al* 2025). While we have only modeled “sampling frequency”, provided that field-scale heterogeneity is captured (i.e., sub sample radius > wavelength of in-field heterogeneity), the required sampling frequency can also be achieved by pooling sub-samples. In this context, a large body of literature exists on how to accurately sample heterogeneous media from the context of soil pollution remediation (i.e., incremental sampling methodologies; (Clausen *et al* 2013b, ITRC 2020, Clausen *et al* 2013c, 2013a, Hadley *et al* 2011, Hewitt *et al* 2007).

While our modeled results indicate that clear gains for MRV robustness may be achieved by aggregating over relatively small numbers of fields (e.g., 10), in practice, generating an accurate aggregate value for CDR presents several challenges. Given the unreliability of single-field estimates, simple area-weighted multiplication of τ_j by application rate, field size, and assumed carbon losses (incl. LCA emissions) can bias results toward the largest fields with the most applied material. A more rigorous approach is to define subsets of fields with similar characteristics (e.g., size, application amount, feedstock, soil type, pH, base saturation, and any parameter used for

MRV and its resolvability; Suhrhoff et al 2024) so that aggregating within these subsets yields meaningful signals. Defining such subsets requires large datasets that enable clustering and defensible grouping of comparable fields, rather than aggregating a random set of fields. Moreover, true multi-field aggregation should combine multiple same-sized independent fields rather than subdividing single fields into smaller sub-fields to benefit from the increase in sampling numbers that true multi-field aggregation entails. Furthermore, EW projects typically include control plots, introducing additional variability and requiring larger aggregation sets to maintain robustness (Bradford *et al* 2023). In practice, unless extraordinary effort is devoted to scouting fields of comparable starting conditions, the need for large sets of fields to define meaningful aggregation subsets may naturally align more closely with frameworks suited to monitor the impact of pay-for-practice policies than with today's VCM protocols.

4.3 Accuracy, Cost, and Scale

Deploying EW as a climate solution requires optimizing accuracy, cost, and scale. Accuracy ensures genuine climate impact, while cost reductions are essential to make projects affordable for voluntary carbon market buyers and, ultimately, feasible as public investments. Current credit prices from first-of-a-kind field trials remain high (~\$300–400 t⁻¹ CO₂; CDR.fyi 2025), and MRV is one of the largest cost drivers (Mercer *et al* 2024). Novel measurement approaches are therefore needed to maintain acceptable accuracy at lower cost over time.

Watershed monitoring is one promising option, providing direct, lower-bound measurements of removal rates over large areas and integrating across soil chemical processes that may otherwise lead to unobserved carbon losses. Although individual gage stations can be expensive (\$20–100k plus recurring lab costs; Harmel et al 2023), each can represent vast agricultural areas, yielding substantially lower per-area monitoring costs. Because station costs do not scale with catchment size, larger watersheds achieve lower per-ton MRV costs and improved signal-to-noise ratios (see Section 4.1). Moreover, gage stations serve multiple public functions—such as water-quality monitoring—allowing costs to be shared among stakeholders.

In-field soil sampling and dissolution estimation via soil mass balance are already common in commercial EW projects (Puro.Earth 2024, Sutherland *et al* 2024). Accuracy is a challenge for this method, given the high spatial variability of soil composition. Building off a similar analysis for soil organic carbon (Potash *et al* 2025, Bradford *et al* 2023), we demonstrate that averaging over 10-100 fields significantly reduces estimation error. In the context of carbon markets, this translates to reduced risk of over crediting, or generating carbon removal credits that do not reflect a real change in atmospheric CO₂ concentration. It is important to note that our results do not support trading off in-field sampling with multi-field averaging. Low in-field sampling densities (less than 2 samples/ha) consistently result in high error rates even when averaged over an increasing number of fields (Figure 3a). Furthermore, suppliers would need to ensure that fields are sufficiently similar to warrant grouping in aggregation sets and must not simply group any set of given fields such that the required high total numbers may be prohibitive for VCM suppliers.

An alternative scaling approach—intensive sampling on a small plot (e.g., 0.1 ha) and linear extrapolation to larger areas—is unlikely to yield accurate results because soil composition, hydrology, and management vary widely across fields, even within the same farm or watershed. Such heterogeneity affects both weathering rates and detectability of products. Evidence from soil-organic-carbon projects shows that extrapolation from one or two small plots per field produces unreliable estimates (Heikkinen *et al* 2013, Maillard *et al* 2017, Poeplau and Don 2015). In EW, this means that a few intensively monitored fields could overestimate removals if scaled to thousands of hectares. Robust estimates therefore require methods that capture landscape variability—either through direct aggregation of field measurements or integration with downstream monitoring.

Aggregation of commercial deployment data also requires clear statistical protocols. Fields must be assigned to aggregation sets prior to post-treatment data becoming available, and all plots must be included in the final analysis for credit delivery. Because soil heterogeneity and measurement error can yield apparent CDR even when true removal is zero, selectively excluding low or negative values would bias results upward (Figure 6). Importantly, quantifying CDR for individual fields rather than aggregation sets causes more frequent overestimation of CDR (Figure 3 and Figure 5), i.e. landing on the right-hand side of the distribution in Figure 6a. In such occurrences there will be no statistical indication that CDR is overestimated. However, while suppliers are not required to deduct CDR emissions from fields with no or negative CDR signals from project deliveries where signals are apparent, the ability to deliver on such individual field sites will invariably inflate credit deliveries. Hence, EW crediting must move beyond individual fields and assess sets of comparable deployments. Strict criteria should define when plots can be added or excluded (e.g., land-management changes), as in existing land-use crediting systems that exclude control plots from dynamic baselines (Shoch *et al* 2024). Such guidance will be essential as the industry scales.

Lastly, we note that integration of modeling into monitoring frameworks can further optimize cost efficiency, though at present models are not a substitute for empirical approaches, neither at watershed nor field scale (Zhang *et al* 2025, Kanzaki *et al* 2025). Modeling, when paired with distributed sensor networks and targeted sampling, can strengthen robustness without dominating budgets. For example, the New York City watershed program invests ~\$6.7 million annually in monitoring, with 10–15 % allocated to modeling tools that enable real-time forecasting (NASEM 2020). Similarly, aggregated EW monitoring that combines field sampling, watershed instrumentation, and calibrated models to constrain downstream losses could offer a credible, lower-cost pathway for large-scale MRV across heterogeneous landscapes.

4.4 Policy implications

Government policy can support accurate estimation of CDR via EW at the scales discussed here—catchments and hundreds of agricultural fields—and beyond, including eco-regions or jurisdictions encompassing thousands of fields. In the U.S., this analysis builds directly on publicly

funded, openly available data from the U.S. Geological Survey (USGS) stream-gage network, which supports economic and environmental decision-making for governments, communities, and commercial users. While the U.S. network is uniquely extensive, other countries maintain comparable public water-monitoring programs (Barker *et al* 2022).

Expanding water-quality monitoring networks serves two critical purposes for EW. Before large-scale deployment, they enable accurate assessment of background weathering rates, guiding optimal site selection and watershed pairing for counterfactual comparisons. After deployment, watershed-scale monitoring provides a direct measure of carbon removal and storage in the aqueous bicarbonate pool, allowing estimation of CDR and offering a conservative check on soil-based measurements.

When multiple projects introduce alkalinity into the same waterway, attribution cannot rely solely on deployed rock amounts or treated area, as deployment strategies affect both storage efficiency and potential outgassing. A more rigorous approach would standardize the use of publicly available reactive-transport models—expanded and cross-calibrated from frameworks such as SCEPTER (Kanzaki *et al* 2022, 2025, 2024), CrunchFlow (Steefel and Molins 2009), or PFLOTRAN (Mills *et al* 2007, Hammond *et al* 2007) amongst others (Taylor *et al* 2017)—to allocate watershed-scale CDR proportional to modeled realized fluxes. Where additional alkalinity inputs occur (e.g., wastewater treatment or other engineered CDR methods), attribution frameworks must adjust EW-derived fluxes accordingly to prevent overestimation. Watershed-level EW monitoring may ultimately require new governance mechanisms beyond the current VCM (Woollen and Planavsky 2024).

Policy can likewise advance soil-based MRV by enabling extensive sampling and aggregation of soil data. Large, publicly maintained datasets—such as national soil censuses—could establish baselines for site selection and monitoring (Smith *et al* 2013, USGS 2024, 2023). Systematic collection of soil pH and related parameters would substantially improve EW assessment accuracy. We stress, however, that large datasets are not in themselves a panacea because robust and accurate estimation at population scales requires representative sampling of individuals (Bradley *et al* 2021). Beyond public programs, substantial amounts of valuable soil data already exist within commercial laboratories—Waypoint (US) and the Tentamus Group (global) each analyze >1.5 million soil samples annually. Policy could unlock these data through incentives for sharing, following models from medicine, energy, and other public-private data partnerships (Susha *et al* 2023). In the U.S., agricultural extension officers could further support this effort by advising farmers on sampling timing and locations, improving regional coverage while minimizing redundancy. Together, such measures would enable robust, aggregated soil datasets, reduce uncertainty in CDR estimates, and create shared public goods benefiting both carbon markets and agricultural management.

Aggregated monitoring, in turn, enables policies that accelerate EW deployment across regions by supporting lower-cost, large-scale assessment of removal fluxes. These fluxes could contribute to

national greenhouse-gas inventories toward meeting Nationally Determined Contributions under the Paris Agreement. A relevant precedent is the widespread use of subsidies for agricultural liming (CRSI 2025). Similar mechanisms could extend to silicate or mixed-feedstock applications, structured as pay-for-practice (area-based payments for spreading material, independent of MRV) or pay-for-results (base payments plus performance incentives following verified removal via soil- or water-phase MRV).

Policy effectiveness will vary regionally and may not align with existing monitoring infrastructure. In the U.S., for instance, regions with the highest weathering potential (primarily the Southeast; Moosdorf et al 2011, Kanzaki et al 2025) do not coincide with areas of dense USGS baseline data (Figure 1a), indicating a need to expand the stream-gaging network. Moreover, use of alternative policy mechanisms (beyond carbon crediting) would benefit from consistent IPCC guidance on the accounting of EW practices in national inventories, a process that is currently underway (IPCC 2024).

In summary, there is a synergistic relationship between policies that support aggregated monitoring (e.g., expanded water-quality networks, public-private data sharing) and those that promote EW deployment. Optimizing across this opportunity space requires balancing CDR potential, measurement accuracy, and cost.

5 Conclusion

Enhanced weathering is a promising strategy for atmospheric CO₂ removal, but its scalability remains constrained by the accuracy, cost, and robustness of MRV. Our analysis shows that aggregated monitoring—via physical integration of weathering products in streams and rivers and statistical aggregation of field-level measurements—offers a credible, lower-cost pathway to MRV at scale. By leveraging existing infrastructure and established methods, aggregation helps overcome central barriers to EW deployment while aligning with approaches already used in soil carbon and forestry frameworks.

At the watershed level, we find no evidence that larger agricultural catchments have a lower utility of detecting EW signals in stream waters based on equal rock application, dissolution, and transport rates. In contrast, in some cases signal detection may be favorable in larger catchments due to lower variability in stream chemistry and lower background weathering fluxes. Because installation and operation costs are largely fixed, MRV costs per hectare and per ton decrease with catchment size, making watershed monitoring an attractive option for large-scale or jurisdictional deployment. Results from such monitoring can serve both as direct CDR measurements and as “top-down” validation for field-based estimates. At the field scale, soil sampling and mass-balance analysis remain standard MRV tools (Clarkson *et al* 2024) but are limited by spatial heterogeneity. Aggregation across multiple fields markedly improves accuracy and robustness, lowering both average error and the risk of over-crediting, particularly when paired with sufficient in-field sampling density.

Together, these findings indicate that aggregation is not merely a technical workaround but a foundational principle for robust EW monitoring. Whether through hydrological integration at the catchment scale or statistical integration across fields, aggregation lowers variance, mitigates systematic bias, and enables credible CDR estimates. Aggregation sets must be defined *a priori*, with all predesignated plots—controls and treatments—included in final analyses. Selectively excluding low or negative results risks inflating average removals and undermining integrity. Importantly, aggregation also shifts MRV cost structures: rather than scaling linearly with the number of participating fields, costs can be amortized across larger areas and multiple stakeholders, making EW more feasible for inclusion in both VCMs and national greenhouse gas inventories.

More work is needed to refine not only the economic and policy frameworks that could support aggregated monitoring, but also the statistical foundations underlying it. This includes developing guidance on how similar fields must be to form valid aggregation groups, defining sufficient field sets as a function of within- and among-field variance, and establishing stratification approaches that account for potential co-variance among parameters influencing field-level CDR. In parallel, quantitative comparisons of soil- and water-based MRV costs can help identify optimal deployment strategies across agronomic and hydrological contexts. Expanding public water and

soil monitoring infrastructure—together with protocols mandating inclusion of all fields within predefined aggregation sets and clear criteria for group formation—can enhance transparency and crediting rigor while generating wider societal benefits. Integration with existing agricultural support mechanisms, such as liming subsidies or soil census programs, offers a practical path for embedding EW monitoring within established governance structures.

In sum, while uncertainties remain about absolute CDR rates, our results demonstrate that aggregated monitoring provides a viable route to accurate, scalable, and cost-effective MRV for EW. By coupling hydrological integration in watersheds with statistical integration across landscapes, aggregation can anchor EW's credibility as a climate solution that benefits farmers, reduces costs, and facilitates policy adoption. As deployment scales up, aggregation may prove to be the key enabling principle that bridges the gap between scientific rigor, economic feasibility, and climate impact.

6 Acknowledgements

The authors would like to thank Prof. Dr. Susan L. Brantley, Dr. Abby Lunstrum, and Dr. Lucia D. Simonelli for helpful feedback and discussion of this manuscript.

7 Funding

TJS acknowledges funding by the Swiss National Science Foundation (Grant P500PN_210790). TJS, MAB, AP, and NP acknowledge support from the Yale Center for Natural Carbon Capture. AK acknowledges funding from Advocates for Climate Innovation. SZ acknowledges support from College of Arts & Sciences Environment and Sustainability Initiative (ESI) at Texas A&M University. NP and CR acknowledge funding from the United States Department of Agriculture (USDA) and the Grantham Foundation for the Environment.

8 Conflict of interest

PAR, CTR, and NJP are scientific advisors to CREW carbon, a company that generates alkalinity in wastewater treatment plants using lime addition but does not work on terrestrial EW.

NJP and CTR were co-founders of the EW supplier Lithos Carbon but have no financial ties to the company.

All other authors declare no conflicts of interest.

9 Data availability statement

All data used for modeling in this study are publicly available as of 22 September 2025 and/or contained within this manuscript. These datasets include: U.S. Geological Survey (USGS) stream gage and river chemistry records (USGS 2016), agricultural land cover data from the National Land Cover Database (NLCD) (USGS 2024), GEOROC database for calculation of US-average basalt composition (Lehnert *et al* 2000), Geochemical and mineralogical data for soils of the conterminous United States to constrain the composition of a large set of agricultural fields (Smith *et al* 2013), data on in-field heterogeneity for 5 field sites (Suhrhoff et al 2025; see supplement S2), constraints on soil CO₂ emissions and variance thereof (Kazula and Lauer 2023, Milliken *et al* 2025).

All analysis and simulation code used in this study is attached to this publication and will be made permanently available via Zenodo upon final publication.

10 Figures

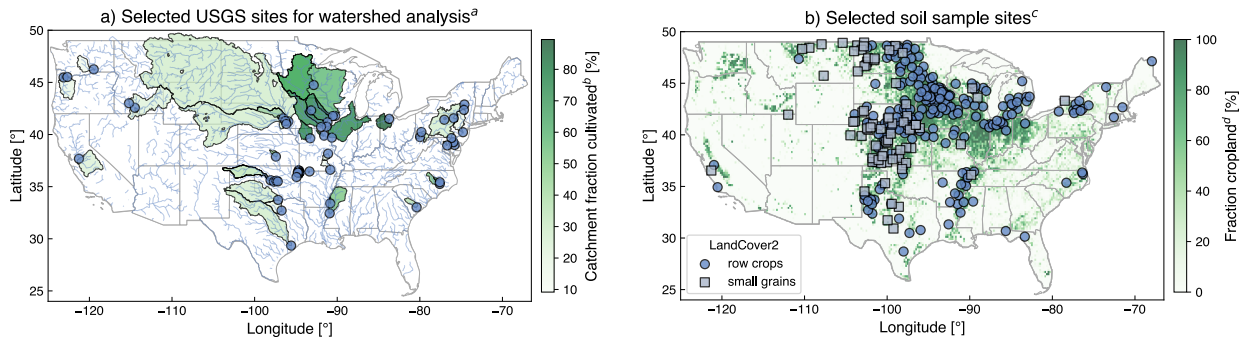


Figure 1: Map of the selected USGS sites for the watershed analysis (a; some watersheds smaller than marker size). Watersheds are color coded by the fraction of land that is classified as cultivated vegetation (Tuanmu and Jetz 2014). Panel (b) shows soil samples used to model aggregated monitoring of soil-based quantification approaches.

^a(USGS 2016) ^b(Tuanmu and Jetz 2014) ^c(Smith *et al* 2013) ^d(Potapov *et al* 2022)

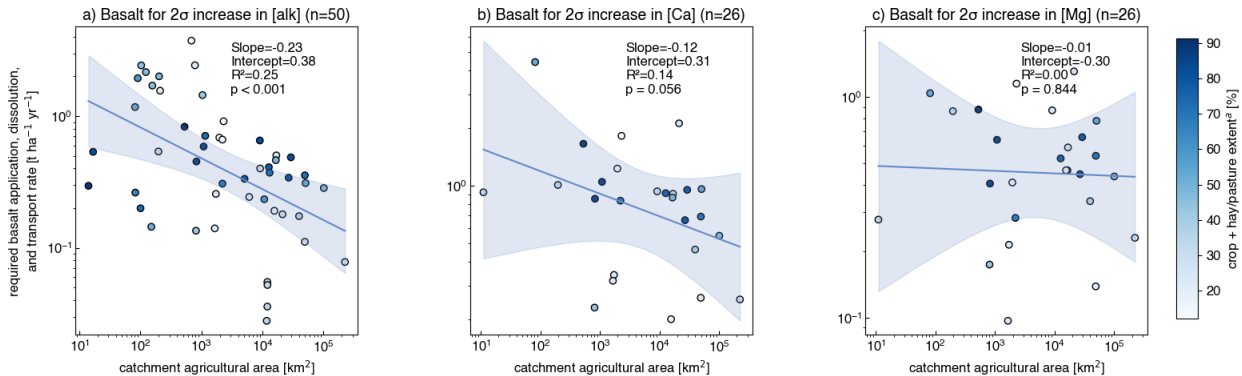


Figure 2: Required agricultural area-normalized basalt dissolution and transport rates to cause a 2σ increase in river alkalinity (a), Ca (b) and Mg (c) concentrations assuming steady state. A version of this Figure but normalized to total catchment area can be found in the supplement (Figure S1)

^a(USGS 2024)

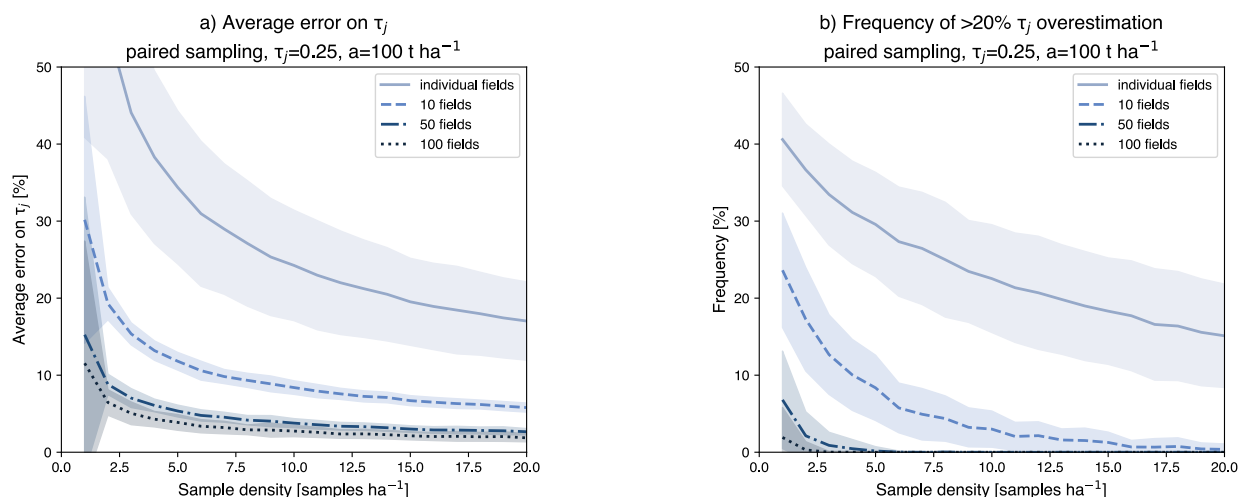


Figure 3: Panel a) shows the error on detected mass transfer coefficients τ_j one would get on average if one applied a soil mass balance approach to quantify rock powder dissolution once for individual as well as sets of fields. The frequency of overestimating τ_j by at least 20% as a function of sampling density and number of aggregated fields is shown in panel b). The plots for the remaining combinations of application amounts and dissolution fractions can be found in the supplement (Figures S3-7). All simulations assume unpaired sampling (see supplement S2.3 for paired sampling).

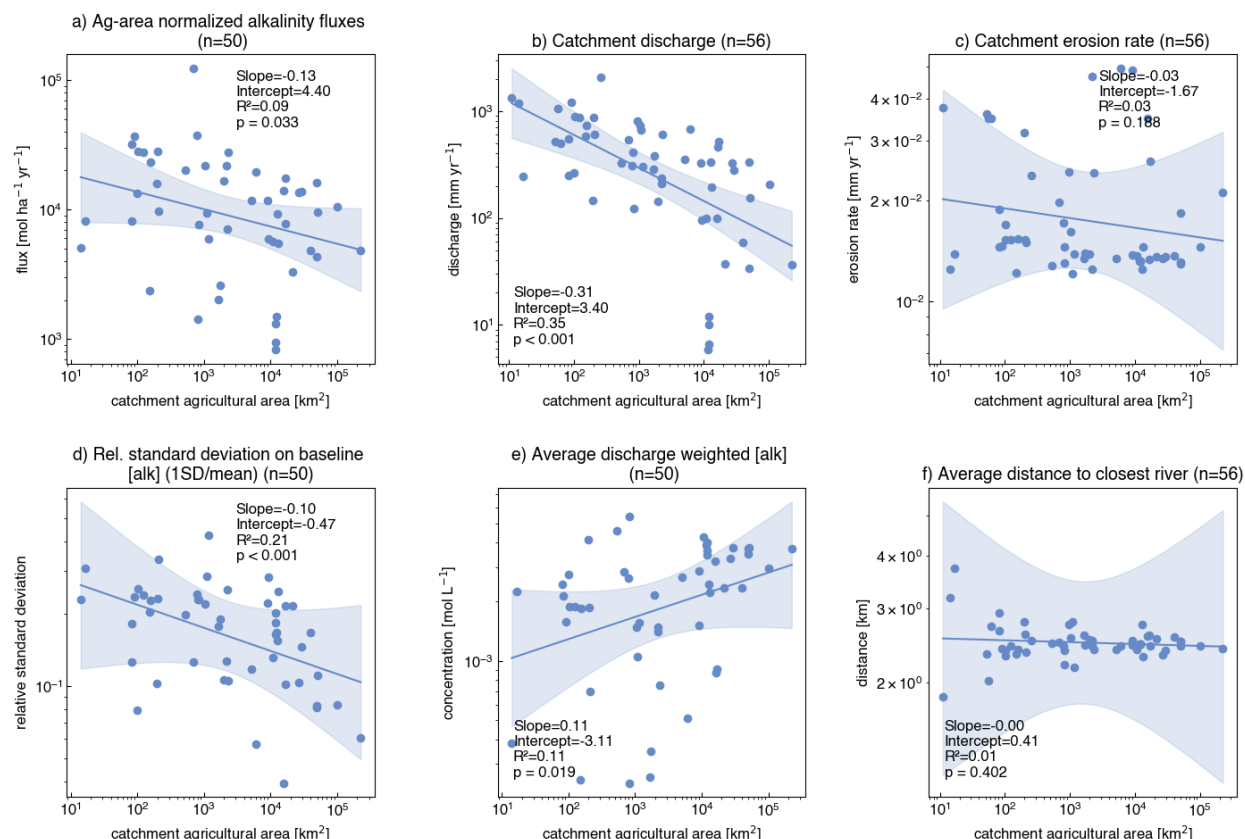


Figure 4: The top row shows agricultural area normalized alkalinity fluxes (a), average annual alkalinity concentration (b), and the variability of baseline alkalinity data (c) for the set of rivers that have sufficient alkalinity baseline data (n=89). These trends are generally similar for Ca and Mg (Figure S3). The bottom row shows area normalized runoff (calculated from discharge and catchment area) (d), average distance to the closest river (e) and erosion rates (f) for all stations that have sufficient baseline data for alkalinity, Ca, or Mg as well as a non-zero proportion of agricultural land (n=56). A version of this figure based on total catchment area, not catchment agricultural area, is included in the supplement (Figure S5).

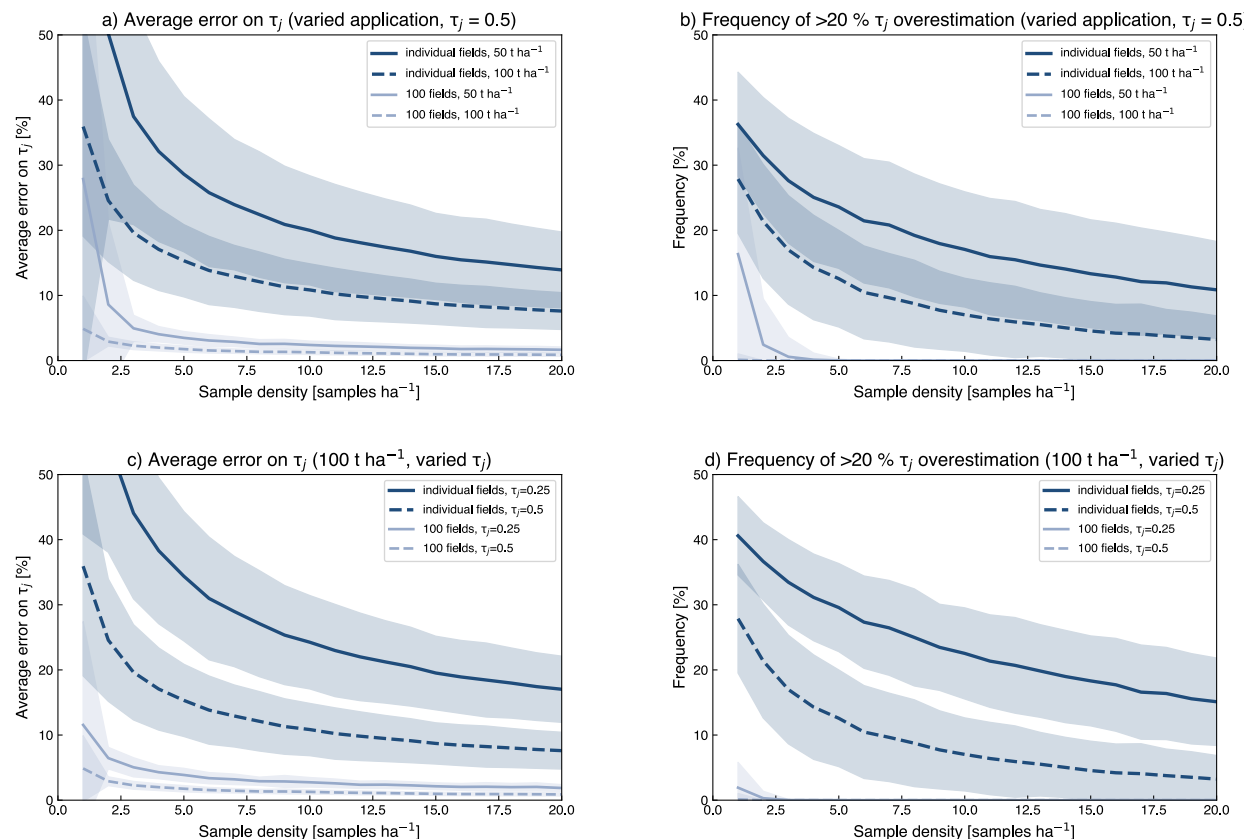


Figure 5: Average error on detected mass transfer coefficients (i.e., dissolution fractions; a and c) as well as the frequency of overestimating mass transfer coefficients by more than 20% (b and d) for constant τ_j but variable application amounts (a and b) as well as at constant application amount but variable τ_j (c and d). These simulations are based on un-paired sampling (see supplement S2.3 for paired sampling).

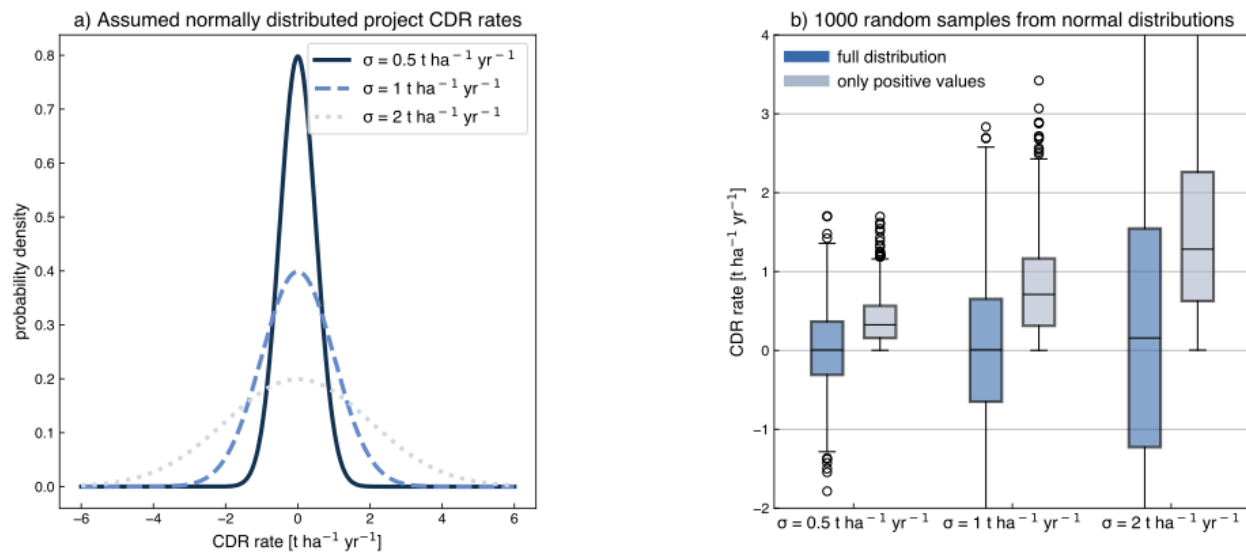


Figure 6: Panel a) shows hypothetical EW deployments where no CDR has occurred. Fields of different heterogeneities (or possibly temporal variability in case of water-based approaches) have varying spread around the mean when samples are used to constrain CDR. If only positive realizations of these random distributions are considered, CDR rates are generated from noise (b). This effect increases the more noise is in the system. Collectively, this demonstrates the necessity to include all fields included in a set when issuing credits as well as that field sets may not be defined based on apparent signal emergence (or exclusion from lack thereof). This effect is not only relevant for no-CDR cases but as long as “negative CDR”, i.e., CO₂ emissions, is within uncertainty of detected CDR rates.

11 References

- Almaraz M, Bingham N L, Holzer I O, Geoghegan E K, Goertzen H, Sohng J and Houlton B Z 2022 Methods for determining the CO₂ removal capacity of enhanced weathering in agronomic settings *Frontiers in Climate* **4**
- Badgley G, Freeman J, Hamman J J, Haya B, Trugman A T, Anderegg W R L and Cullenward D 2022 Systematic over-crediting in California’s forest carbon offsets program *Glob Chang Biol* **28** 1433–45
- Baek S H, Kanzaki Y, Lora J M, Planavsky N, Reinhard C T and Zhang S 2023 Impact of Climate on the Global Capacity for Enhanced Rock Weathering on Croplands Earth’s Future *Earths Future* **11** 1–12
- Barker L J, Fry M, Hannaford J, Nash G, Tanguy M and Swain O 2022 Dynamic High Resolution Hydrological Status Monitoring in Real-Time: The UK Water Resources Portal *Front Environ Sci* **10** 1–15
- Barnes R T and Raymond P A 2009 The contribution of agricultural and urban activities to inorganic carbon fluxes within temperate watersheds *Chem Geol* **266** 318–27 Online: <http://dx.doi.org/10.1016/j.chemgeo.2009.06.018>
- Beerling D J, Kantzas E P, Lomas M R, Taylor L L, Zhang S, Kanzaki Y, Eufrazio R M, Renforth P, Mecure J F, Pollitt H, Holden P B, Edwards N R, Koh L, Epihov D Z, Wolf A, Hansen J E, Banwart S A, Pidgeon N F, Reinhard C T, Planavsky N J and Val Martin M 2025 Transforming US agriculture for carbon removal with enhanced weathering *Nature* **638** Online: <http://dx.doi.org/10.1038/s41586-024-08429-2>
- Beerling D J, Kantzas E P, Lomas M R, Wade P, Eufrazio R M, Renforth P, Sarkar B, Andrews M G, James R H, Pearce C R, Mercure J, Pollitt H, Holden P B and Edwards N R 2020 Potential for large-scale CO₂ removal via enhanced rock weathering with croplands *Nature* **583** 242–8 Online: <http://dx.doi.org/10.1038/s41586-020-2448-9>
- Bijma J, Hagens M, Hammes J, Planavsky N, von Strandmann P A E P, Reershemius T, Reinhard C T, Renforth P, Suhrhoff T J, Vicca S, Vienne A and Wolf-Gladrow D A 2025 Reviews and syntheses: Carbon vs. cation based MRV of Enhanced Rock Weathering and the issue of soil organic carbon *Biogeosciences* 1–29 Online: <http://dx.doi.org/10.5194/egusphere-2025-2740>
- Bradford M A, Eash L, Polussa A, Jevon F V., Kuebbing S E, Hammac W A, Rosenzweig S and Oldfield E E 2023 Testing the feasibility of quantifying change in agricultural soil carbon stocks through empirical sampling *Geoderma* **440** 116719 Online: <https://doi.org/10.1016/j.geoderma.2023.116719>
- Bradley V C, Kuriwaki S, Isakov M, Sejdinovic D, Meng X L and Flaxman S 2021 Unrepresentative big surveys significantly overestimated US vaccine uptake *Nature* **600** 695–700
- Calabrese S, Wild B, Bertagni M B, Bourg I C, White C, Aburto F, Cipolla G, Noto L V. and Porporato A 2022 Nano- to Global-Scale Uncertainties in Terrestrial Enhanced Weathering *Environ Sci Technol* **56** 15261–72
- Campbell J, Bastianini L, Buckman J, Bullock L, Foteinis S, Furey V, Hamilton J, Harrington K, Hawrot O, Holdship P, Knapp W, Maesano C, Mayes W, Pogge von Strandmann P, Reershemius T, Rosair G, Sturgeon F, Turvey C, Wilson S and Renforth P 2023 *Measurements in Geochemical Carbon Dioxide Removal* (Heriot-Watt University)
- CDR.fyi 2025 *Keep Calm and Remove On - The CDR.fyi 2024 Year in Review*

- Clarkson M O, Larkin C S, Swoboda P, Reershemius T, Suhrhoff T J, Maesano C N and Campbell J S
2024 A Review of Measurement for Quantification of Carbon Dioxide Removal by Enhanced
Weathering in Soil *Frontiers in Climate* **6** 1–20
- Clausen J, Gerogian T and Bednar A 2013a *Incremental Sampling Methodology (ISM) for Metallic
Residues*
- Clausen J L, Georgian T and Bednar A 2013b *Cost and Performance Report of Incremental Sampling
Methodology for Soil Containing Metallic Residues Engineer Research and Development Center*
- Clausen J L, Georgnian T, Bednar A, Perron N, Bray A, Tuminello P, Gooch G, Mulherin N, Gelvin A,
Beede M, Saari S, Jones W and Tazik S 2013c *Demonstration of Incremental Sampling
Methodology for Soil Containing Metallic Residues*
- CRSI 2025 Policy database *Zenodo* Online: <http://www.reecp.org/policy-database>
- Derry L A, Chadwick O A and Porder S 2025 Estimation of Carbon Dioxide Removal via Enhanced
Weathering *Glob Chang Biol* **31** 1–3
- Duan S, Kaushal S S, Rosenfeldt E J, Murthy S and Fischel M H H 2025 Chemical weathering and land
use control river alkalization and dissolved inorganic carbon in the Potomac River, USA *Applied
Geochemistry* **191** 106530 Online: <https://doi.org/10.1016/j.apgeochem.2025.106530>
- Gaillardet J, Dupré B, Louvat P and Allègre C J 1999 Global silicate weathering and CO₂ consumption
rates deduced from the chemistry of large rivers *Chem Geol* **159** 3–30
- Geden O, Gidden M J, Lamb W F, Minx J C, Nemet G F and Smith S M 2024 The State of Carbon
Dioxide Removal
- Gill-Wiehl A, Kammen D and Haya B 2023 Pervasive over-crediting from cookstoves offset
methodologies *Res Sq* 1–23
- Gislason S R, Oelkers E H, Eiriksdottir E S, Kardjilov M I, Gisladottir G, Sigfusson B, Snorrason A,
Elefsen S, Hardardottir J, Torssander P and Oskarsson N 2009 Direct evidence of the feedback
between climate and weathering *Earth Planet Sci Lett* **277** 213–22 Online:
<http://dx.doi.org/10.1016/j.epsl.2008.10.018>
- Godsey S E and Kirchner J W 2014 Dynamic, discontinuous stream networks: Hydrologically driven
variations in active drainage density, flowing channels and stream order *Hydrol Process* **28** 5791–
803
- Hadley P W, Crapps E and Hewitt A D 2011 Time for a Change of Scene *Environ Forensics* **12** 312–8
- Hamilton S K, Kurzman A L, Arango C, Jin L and Robertson G P 2007 Evidence for carbon sequestration
by agricultural liming *Global Biogeochem Cycles* **21** 1–12
- Hammond G, Lichtner P and Lu C 2007 Subsurface multiphase flow and multicomponent reactive
transport modeling using high-performance computing *J Phys Conf Ser* **78**
- Harmel R D, Preisendanz H E, King K W, Busch D, Birgand F and Sahoo D 2023 A Review of Data
Quality and Cost Considerations for Water Quality Monitoring at the Field Scale and in Small
Watersheds *Water (Switzerland)* **15** 1–19
- Hartmann J 2009 Bicarbonate-fluxes and CO₂-consumption by chemical weathering on the Japanese
Archipelago — Application of a multi-lithological model framework *Chem Geol* **265** 237–71
Online: <http://dx.doi.org/10.1016/j.chemgeo.2009.03.024>
- Heikkinen J, Ketoja E, Nuutinen V and Regina K 2013 Declining trend of carbon in Finnish cropland
soils in 1974–2009 *Glob Chang Biol* **19** 1456–69

- Hewitt A D, Jenkins T F, Walsh M E, Walsh M R, Bigl S R and Ramsey C A 2007 *Protocols for Collection of Surface Soil Samples at Military Training and Testing Ranges for the Characterization of Energetic Munitions Constituents*
- Holland P W 1986 Statistics and causal inference *J Am Stat Assoc* **81** 945–60
- Iff N, Renforth P and Pogge von Strandmann P A E 2024 The dissolution of olivine added to soil at 32°C: the fate of weathering products and its implications for enhanced weathering at different temperatures *Frontiers in Climate* **6** 1–18
- IPCC 2022 *Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*
- IPCC 2018 Global warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, *Ipcc - Sr15* **2** 17–20
Online: <https://www.ipcc.ch/sr15/>
- IPCC 2024 *IPCC Expert Meeting on Carbon Dioxide Removal Technologies and Carbon Capture Utilization and Storage* ed M Enoki, T., Hayat (Report of the IPCC Expert Meeting, Pub. IGES, Japan) Online: https://www.ipcc-nggip.iges.or.jp/public/mtdocs/pdfiles/2407_Background_CDR_CCUS.pdf
- ITRC 2020 *Incremental Sampling Methodology (ISM) Update The Interstate Technology & Regulatory Council (ITRC)*
- Kantola I B, Blanc-Betes E, Master M D, Chang E, Marklein A, Moore C E, von Haden A, Bernacchi C J, Wolf A, Epihov D Z, Beerling D J and Delucia E H 2023 Improved net carbon budgets in the US Midwest through direct measured impacts of enhanced weathering *Glob Chang Biol* 1–17
- Kanzaki Y, Chiaravalloti I, Zhang S, Planavsky N J and Reinhard C T 2024 In silico calculation of soil pH by SCEPTER v1.0 *Geosci Model Dev* **17** 4515–32
- Kanzaki Y, Planavsky N, Zhang S, Jordan J, Suhrhoff T J and Christopher T 2025 Soil cation storage is a key control on the carbon removal dynamics of enhanced weathering *Environmental Research Letters* **20** 1–11
- Kanzaki Y, Zhang S, Planavsky N J and Reinhard C T 2022 Soil Cycles of Elements simulator for Predicting TERrestrial regulation of greenhouse gases: SCEPTER v0.9 *Geosci Model Dev* **15** 4959–90
- Kazula M J and Lauer J G 2023 Greenhouse gas emissions from soils in corn-based cropping systems *J Environ Qual* **52** 1080–91
- Kirchner J W and Neal C 2013 Universal fractal scaling in stream chemistry and its implications for solute transport and water quality trend detection. *Proc Natl Acad Sci U S A* **110** 12213–8 Online: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3725102&tool=pmcentrez&rendertype=abstract>
- Lamb W F, Gasser T, Roman-Cuesta R M, Grassi G, Gidden M J, Powis C M, Geden O, Nemet G, Pratama Y, Riahi K, Smith S M, Steinhauser J, Vaughan N E, Smith H B and Minx J C 2024 The carbon dioxide removal gap *Nat Clim Chang* **14** 644–51 Online: <http://dx.doi.org/10.1038/s41558-024-01984-6>
- Lehnert K, Su Y, Langmuir C H, Sarbas B and Nohl U 2000 A global geochemical database structure for rocks *Geochemistry, Geophysics, Geosystems* **1**

- Levy C R, Almaraz M, Beerling D J, Raymond P, Reinhard C T, Suhrhoff T J and Taylor L 2024
Enhanced Rock Weathering for Carbon Removal – Monitoring and Mitigating Potential
Environmental Impacts on Agricultural Land *Environ Sci Technol* **58** 17215–26
- Luderer G, Vrontisi Z, Bertram C, Edelenbosch O Y, Pietzcker R C, Rogelj J, De Boer H S, Drouet L,
Emmerling J, Fricko O, Fujimori S, Havlík P, Iyer G, Keramidas K, Kitous A, Pehl M, Krey V,
Riahi K, Saveyn B, Tavoni M, Van Vuuren D P and Kriegler E 2018 Residual fossil CO₂ emissions
in 1.5-2 °C pathways *Nat Clim Chang* **8** 626–33 Online: <http://dx.doi.org/10.1038/s41558-018-0198-6>
- Maillard É, McConkey B G and Angers D A 2017 Increased uncertainty in soil carbon stock
measurement with spatial scale and sampling profile depth in world grasslands: A systematic
analysis *Agric Ecosyst Environ* **236** 268–76 Online: <http://dx.doi.org/10.1016/j.agee.2016.11.024>
- Mercer L, Burke J and Rodway-dyer S 2024 *Towards improved cost estimates for monitoring , reporting
and verification of carbon dioxide removal* (London School of Economics and Political Science)
- Milliken E, Woodley A and Planavsky N J 2025 Direct Measurement of Carbon Dioxide Removal Due to
Enhanced Weathering *Environ Sci Technol Lett*
- Mills R T, Lu C, Lichtner P C and Hammond G E 2007 Simulating subsurface flow and transport on
ultrascale computers using PFLOTRAN *J Phys Conf Ser* **78**
- Moosdorf N, Hartmann J, Lauerwald R, Hagedorn B and Kempe S 2011 Atmospheric CO₂ consumption
by chemical weathering in North America *Geochim Cosmochim Acta* **75** 7829–54 Online:
<http://dx.doi.org/10.1016/j.gca.2011.10.007>
- NASEM 2020 *Review of the New York City Watershed Protection Program* (National Academies of
Sciences, Engineering, and Medicine;The National Academies Press, Washington, DC)
- Oh N H and Raymond P A 2006 Contribution of agricultural liming to riverine bicarbonate export and
CO₂ sequestration in the Ohio River basin *Global Biogeochem Cycles* **20** 1–17
- te Pas E E E M, Chang E, Marklein A R, Comans R N J and Hagens M 2025 Accounting for retarded
weathering products in comparing methods for quantifying carbon dioxide removal in a short-term
enhanced weathering study *Frontiers in Climate* **15** 4998 1–10
- Poeplau C and Don A 2015 Carbon sequestration in agricultural soils via cultivation of cover crops - A
meta-analysis *Agric Ecosyst Environ* **200** 33–41 Online:
<http://dx.doi.org/10.1016/j.agee.2014.10.024>
- Pogge von Strandmann P A E, Tooley C, Mulders J J P A and Renforth P 2022 The Dissolution of
Olivine Added to Soil at 4°C: Implications for Enhanced Weathering in Cold Regions *Frontiers in
Climate* **4** 1–11
- Potapov P, Turubanova S, Hansen M C, Tyukavina A, Zalles V, Khan A, Song X P, Pickens A, Shen Q
and Cortez J 2022 Global maps of cropland extent and change show accelerated cropland expansion
in the twenty-first century *Nat Food* **3** 19–28
- Potash E, Bradford M A, Oldfield E E and Guan K 2025 Measure-and-remeasure as an economically
feasible approach to crediting soil organic carbon at scale *Environmental Research Letters* **20**
- Power I M, Hatten V N J, Guo M, Rausis K and Klyn-hesselink H 2025 Are enhanced rock weathering
rates overestimated? A few geochemical and mineralogical pitfalls 1–9
- Puro.Earth 2024 Enhanced Rock Weathering Methodology for CO₂ Removal 1–68 Online:
[https://7518557.fs1.hubspotusercontent-na1.net/hubfs/7518557/Supplier Documents/ERW
methodology.pdf](https://7518557.fs1.hubspotusercontent-na1.net/hubfs/7518557/Supplier Documents/ERW methodology.pdf)

- Reershemius T, Kelland M E, Davis I R, D'Ascanio R, Kalderon-Asael B, Asael D, Suhrhoff T J, Epihov D E, Beerling D J, Reinhard C T and Planavsky N J 2023 Initial Validation of a Soil-Based Mass-Balance Approach for Empirical Monitoring of Enhanced Rock Weathering Rates *Environ Sci Technol* **57** 19497–19507 Online: <http://arxiv.org/abs/2302.05004>
- Reershemius T and Suhrhoff T J 2023 On error, uncertainty, and assumptions in calculating carbon dioxide removal rates by enhanced rock weathering in Kantola et al ., 2023 *Glob Chang Biol* 1–3
- Renforth P and Henderson G 2017 Assessing ocean alkalinity for carbon sequestration *Reviews of Geophysics* **55** 636–74
- Renforth P, Pogge von Strandmann P A E and Henderson G M 2015 The dissolution of olivine added to soil: Implications for enhanced weathering *Applied Geochemistry* **61** 109–18 Online: <http://dx.doi.org/10.1016/j.apgeochem.2015.05.016>
- Rogelj J, Popp A, Calvin K V., Luderer G, Emmerling J, Gernaat D, Fujimori S, Strefler J, Hasegawa T, Marangoni G, Krey V, Kriegler E, Riahi K, Van Vuuren D P, Doelman J, Drouet L, Edmonds J, Fricko O, Harmsen M, Havlik P, Humpenöder F, Stehfest E and Tavoni M 2018 Scenarios towards limiting global mean temperature increase below 1.5 °c *Nat Clim Chang* **8** 325–32 Online: <http://dx.doi.org/10.1038/s41558-018-0091-3>
- Rogers B and Maher K 2025 A Framework for Integrating Spatial Uncertainty into Critical Zone Models : Application to Enhanced Weathering *CDRXIV*
- Rothman K J, Greenland S and Lash T L 2008 *Modern epidemiology* vol 3 (Wolters Kluwer Health/Lippincott Williams & Wilkins Philadelphia)
- Rubin D B 1974 Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol* **66** 688
- Sanders-DeMott R, Hutrya L R, Hurteau M D, Keeton W S, Fallon K S, Anderegg W R L, Hollinger D Y, Kuebbing S E, Lucash M S, Ordway E M, Vargas R and Walker W S 2025 Ground-Truth: Can Forest Carbon Protocols Ensure High-Quality Credits? *Earths Future* **13**
- Shao S, Driscoll C T, Johnson C E, Fahey T J, Battles J J and Blum J D 2016 Long-term responses in soil solution and stream-water chemistry at Hubbard Brook after experimental addition of wollastonite *Environmental Chemistry* **13** 528–40
- Shaughnessy A R and Brantley S L 2023 How do silicate weathering rates in shales respond to climate and erosion? *Chem Geol* **629**
- Shaughnessy A R, Forgeng M J, Wen T, Gu X, Hemingway J D and Brantley S L 2023 Linking Stream Chemistry to Subsurface Redox Architecture *Water Resour Res* **59** 1–20
- Shoch D, Swails E, Sonne Hall E, Belair E, Rifkin B, Griscom B and Latta G 2024 VM0045 - Improved Forest Management Using Dynamic Matched Baselines From National Forest Inventories 1–70
- Smith D B, Cannon W F, Woodruff L G, Solano F, Kilburn J E and Fey D L 2013 Geochemical and mineralogical data for soils of the conterminous United States *U.S. Geological Survey Data Series* **801** 1–26
- Steefel C I and Molins S 2009 CrunchFlow *Software for modeling multicomponent reactive flow and transport. User's manual* 12–91
- Suhrhoff T J, Reershemius T, Jordan J, Li S, Zhang S, Kalderon-asael B, Ebert Y, Nyateka R, Reinhard C T, Planavsky N J, Haven N, Sciences P, Haven N, Carbon M, Sciences A and Kingdom U 2025 Updated framework and signal-to-noise analysis of soil mass balance approaches for quantifying enhanced weathering on managed lands *CDRXIV* 1–59

- Suhrhoff T J, Reershemius T, Wang J, Jordan J S, Reinhard C T and Planavsky N J 2024 A tool for assessing the sensitivity of soil-based approaches for quantifying enhanced weathering: a US case study *Frontiers in Climate* **6** 1–17
- Susha I, Rukanova B, Zuiderwijk A, Gil-Garcia J R and Gasco Hernandez M 2023 Achieving voluntary data sharing in cross sector partnerships: Three partnership models *Information and Organization* **33** 100448 Online: <https://doi.org/10.1016/j.infoandorg.2023.100448>
- Sutherland K, Holme E, Savage R, Gill S, Matlin-Wainer M, He J, Marsland E and Patel C 2024 Isometric Enhanced Weathering in Agriculture v1.0 Online: <https://registry.isometric.com/protocol/enhanced-weathering-agriculture>
- Taylor L L, Beerling D J, Quegan S and Banwart S A 2017 Simulating carbon capture by enhanced weathering with croplands: An overview of key processes highlighting areas of future model development *Biol Lett* **13**
- Taylor L L, Driscoll C T, Groffman P M, Rau G H, Blum J D and Beerling D J 2021 Increased carbon capture by a silicate-treated forested watershed affected by acid deposition *Biogeosciences* **18** 169–88
- Tuanmu M N and Jetz W 2014 A global 1-km consensus land-cover product for biodiversity and ecosystem modelling *Global Ecology and Biogeography* **23** 1031–45
- UNEP 2024 *Emissions Gap Report 2024: No more hot air ... please! With a massive gap between rhetoric and reality, countries draft new climate commitments* (United Nations Environment Program, Nairobi, Kenya)
- USDA 2024 Farms and Farmland 2022 *Census of Agriculture Highlights*
- USDA 2022 Farms and Land in Farms 2021 Summary *United States Department of Agriculture: National Agricultural Statistics Service* 1995–2004 Online: https://www.nass.usda.gov/Publications/Todays_Reports/reports/fnlo0222.pdf
- USGS 2023 CroplandCROS 2022 Online: <https://pdi.scinet.usda.gov/portal/apps/sites/#!/cropcros/pages/download-data>
- USGS 2024 *National Land Cover Database (NLCD) 2021 Land Cover (CONUS)* Online: <https://www.mrlc.gov/data>
- USGS 2016 National Water Information System data available on the World Wide Web (USGS Water Data for the Nation) Online: <http://waterdata.usgs.gov/nwis/>
- Vienne A, Frings P, Rijnders J, Suhrhoff T J, Reershemius T, Poetra R P, Hartmann J, Niron H, Estrada M P, Steinwider L, Boito L and Vicca S 2025 Weathering without inorganic CDR revealed through cation tracing *EGU sphere*
- West J A 2012 Thickness of the chemical weathering zone and implications for erosional and climatic drivers of weathering and for carbon-cycle feedbacks *Geology* **40** 811–4
- White A F and Blum A E 1995 Effects of climate on chemical weathering in watersheds *Geochim Cosmochim Acta* **59** 1729–47
- Wieczorek M E 2019 Area- and Depth- Weighted Average of Soil pH from STATSGO2 for the Conterminous United States and District of Columbia
- Woollen B J and Planavsky N J 2024 Common-Pool Resource Management Problems in Enhanced Weathering Deployments for Carbon Dioxide Removal *Workshop on the Ostrom Workshop – WOW7 Conference* Online: <https://hdl.handle.net/10535/10968>
- Zhang S, Reinhard C T, Liu S, Kanzaki Y and Planavsky N J 2025 A framework for modeling carbon loss from rivers following terrestrial enhanced weathering *Environmental Research Letters* **20**

770

771

Supplementary information to “Aggregated monitoring of enhanced weathering on agricultural lands”

Tim Jesper Suhrhoff^{1,2}, Anu Khan³, Shuang Zhang⁴, Beck Woollen^{1,3}, Tom Reershemius⁵, Mark A. Bradford^{1,6}, Alexander Polussa^{1,6}, Ella Milliken², Pete Raymond⁶, Christopher T. Reinhard⁷, Noah J. Planavsky^{2,1}

1: Yale Center for Natural Carbon Capture, New Haven, CT, USA

2: Department of Earth & Planetary Sciences, Yale University, New Haven, CT, USA

3: Carbon Removal Standards Initiative, Washington D.C., USA

4: Department of Oceanography, Texas A&M University, College Station, TX, USA

5: School of Natural and Environmental Sciences, Newcastle University, Newcastle-upon-Tyne, UK

6: The Forest School, Yale School of the Environment, Yale University, New Haven, CT, USA

7: Department of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA, USA

S1 Supplementary Figures

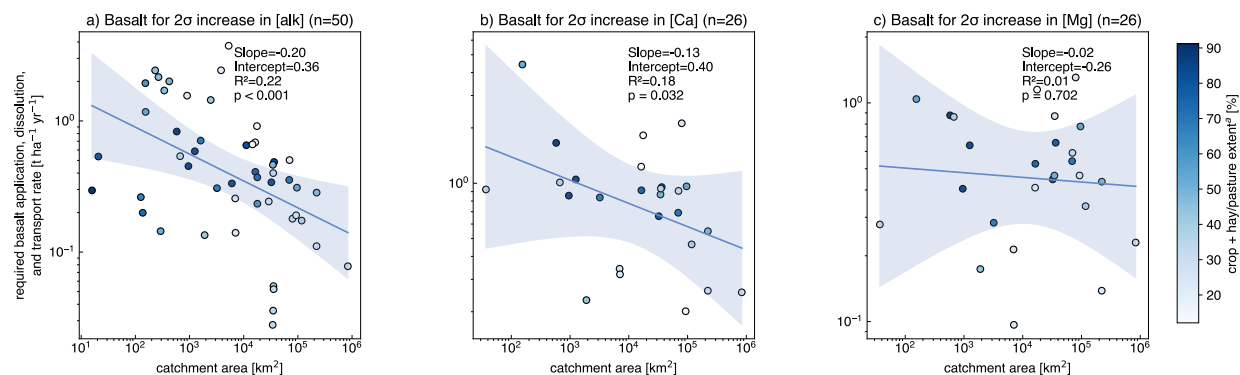


Figure S1: Required total catchment area normalized basalt dissolution and transport rates to cause a 2σ increase in river alkalinity (a), Ca (b) and Mg (c) concentrations assuming steady state.

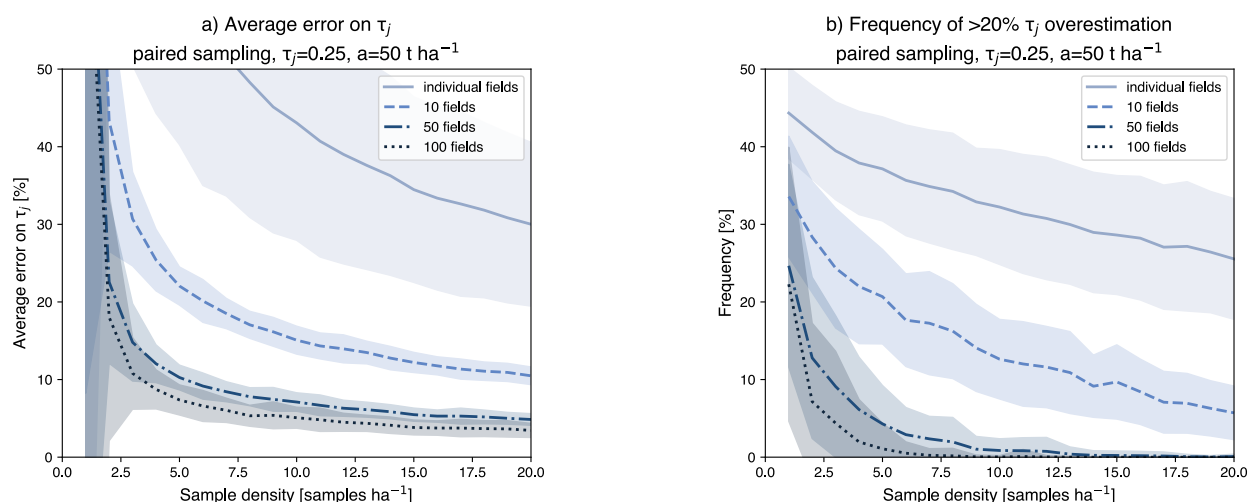


Figure S2: Equivalent to Figure 3 but for $\tau_j = 0.25$ and $a = 50$ t ha $^{-1}$. Panel a) shows the error on detected mass transfer coefficients τ_j one would get on average if one applied a soil mass balance approach to quantify rock powder dissolution once for individual as well as sets of fields. The frequency of overestimating τ_j by at least 20% as a function of sampling density and number of aggregated fields is shown in panel b). The plots for the remaining combinations of application amounts and dissolution fractions can be found in the supplement (Figures S3-7). All simulations assume unpaired sampling (see supplement S2.3 for paired sampling).

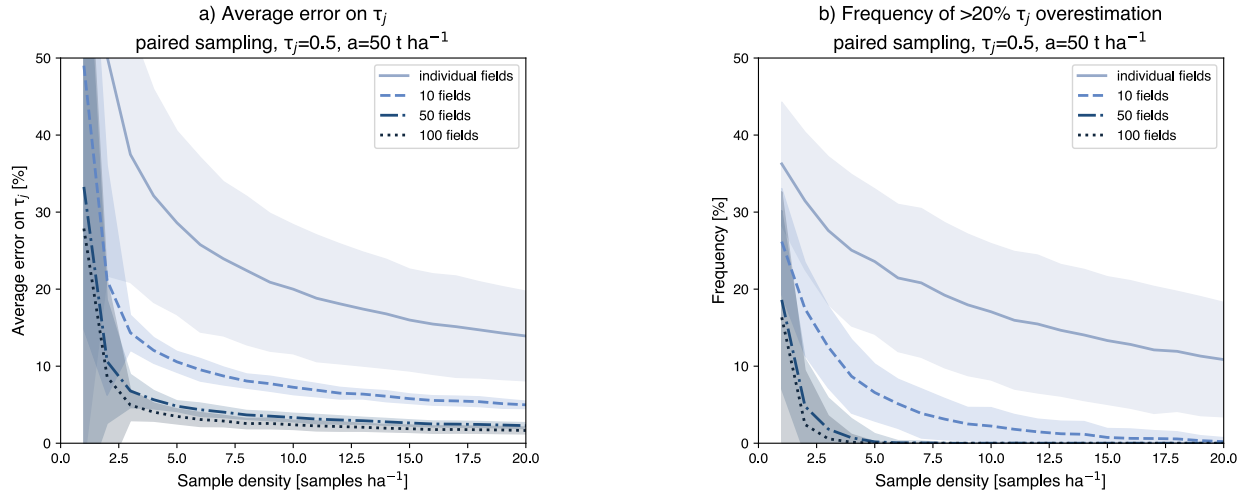


Figure S3: Equivalent to Figure 3 but for $\tau_j = 0.5$ and $a = 50 \text{ t ha}^{-1}$. Panel a) shows the error on detected mass transfer coefficients τ_j one would get on average if one applied a soil mass balance approach to quantify rock powder dissolution once for individual as well as sets of fields. The frequency of overestimating τ_j by at least 20% as a function of sampling density and number of aggregated fields is shown in panel b). The plots for the remaining combinations of application amounts and dissolution fractions can be found in the supplement (Figures S3-7). All simulations assume unpaired sampling (see supplement S2.3 for paired sampling).

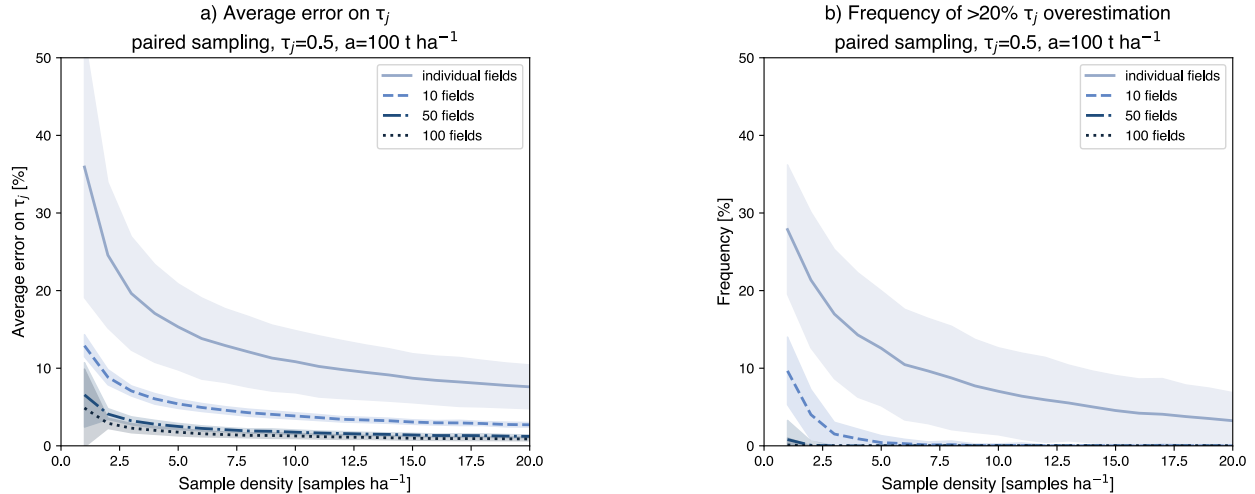
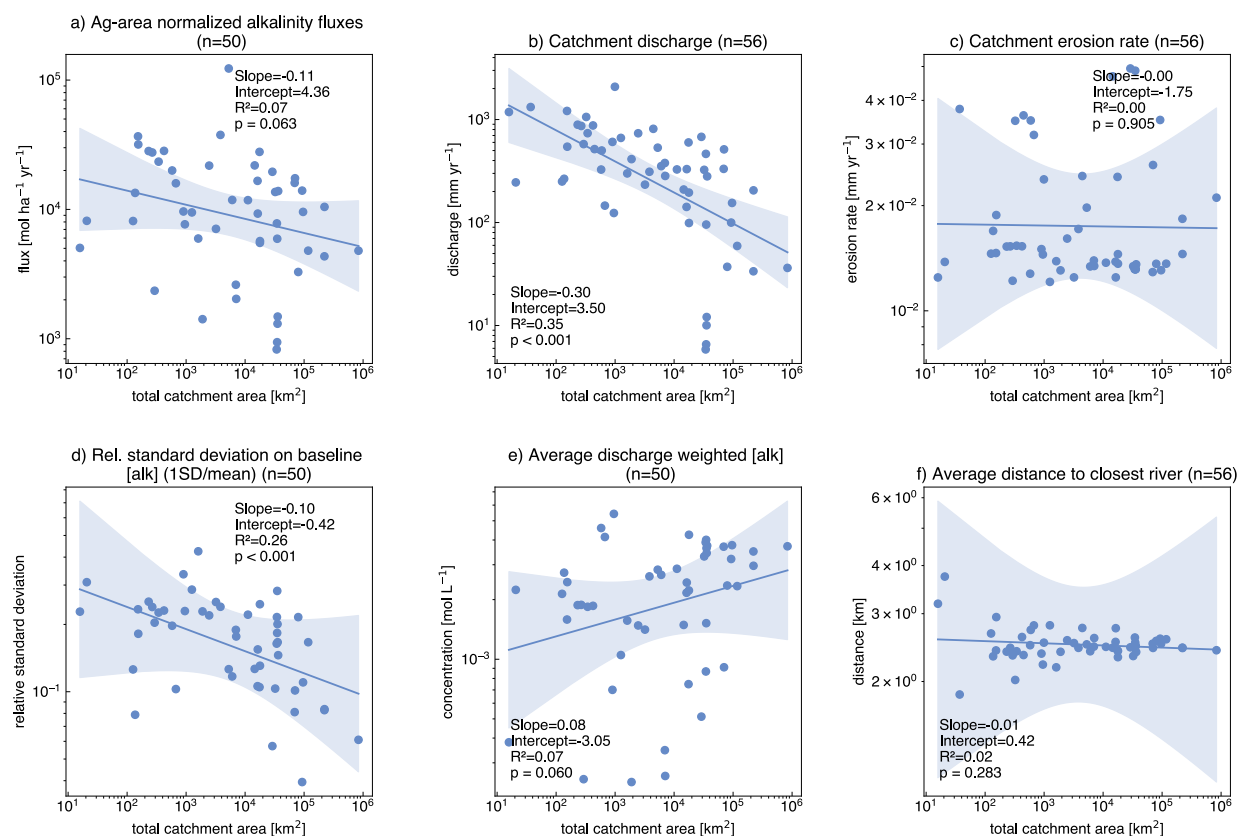


Figure S4: Equivalent to Figure 3 but for $\tau_j = 0.25$ and $a = 50 \text{ t ha}^{-1}$. Panel a) shows the error on detected mass transfer coefficients τ_j one would get on average if one applied a soil mass balance approach to quantify rock powder dissolution once for individual as well as sets of fields. The frequency of overestimating τ_j by at least 20% as a function of sampling density and number of aggregated fields is shown in panel b). The plots for the remaining combinations of application amounts and dissolution fractions can be found in the supplement (Figures S3-7). All simulations assume unpaired sampling (see supplement S2.3 for paired sampling).



819

820 Figure S5: The top row shows total catchment area normalized alkalinity fluxes (a), average annual
 821 alkalinity concentration (b), and the variability of baseline alkalinity data (c) for the set of rivers that have
 822 sufficient alkalinity baseline data (n=89). The bottom row shows area normalized runoff (d), average
 823 distance to the closest river (e) and erosion rates (f) for all stations that have sufficient baseline data for
 824 alkalinity, Ca, or Mg as well as a non-zero proportion of agricultural land (n=118). The equivalent
 825 information of panels a, d, and e for Ca and Mg can be found in Figure S6.

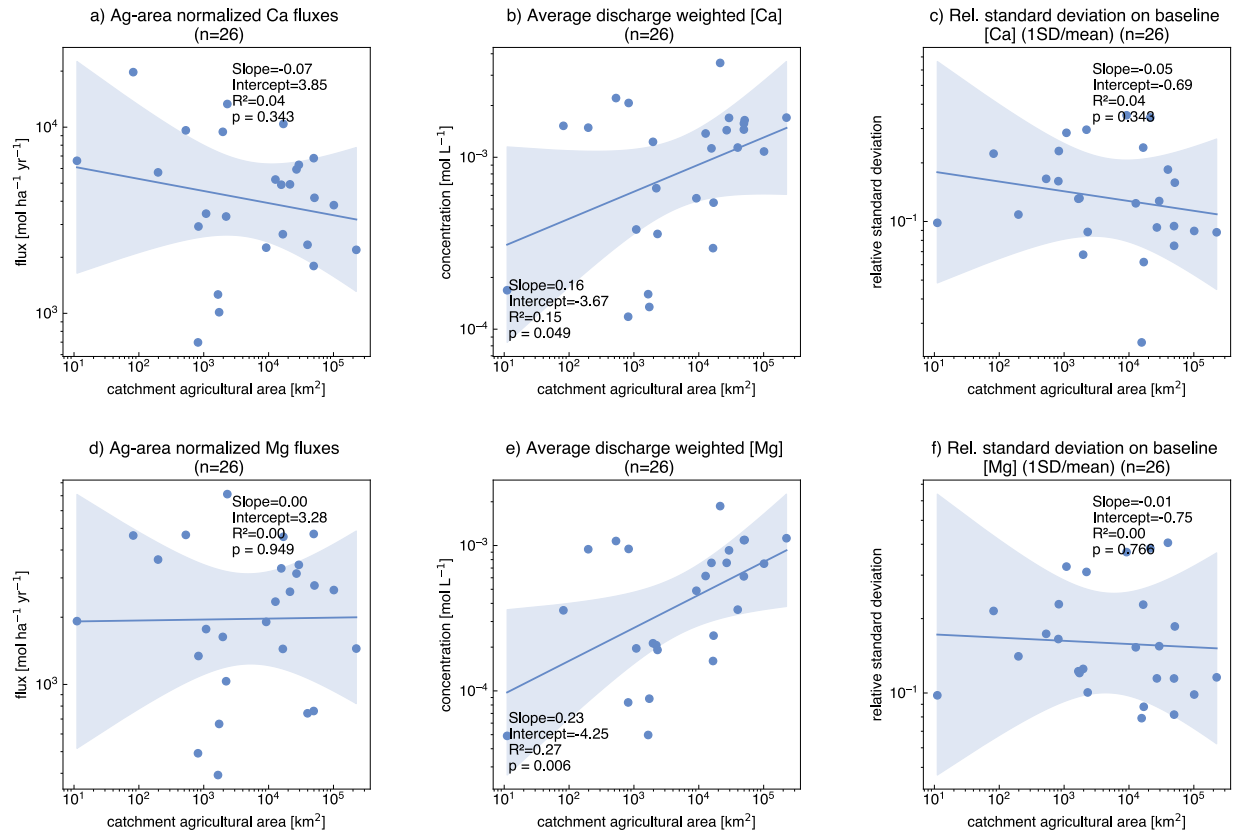


Figure S6: The top row shows agricultural area normalized Ca fluxes (a), average annual Ca concentration (b), and the variability of baseline Ca data (c). The bottom row shows the same type of data but for Mg (d-f).

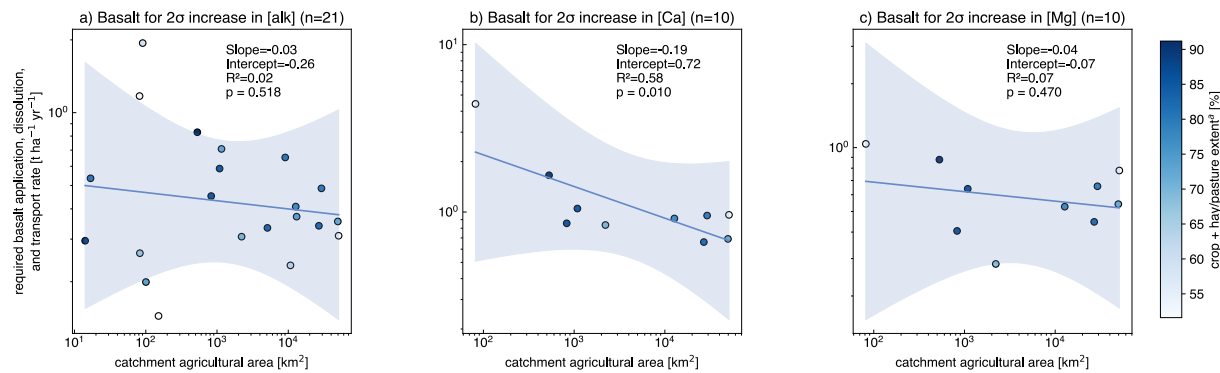


Figure S7: Required basalt application rates to cause a 2σ increase in river alkalinity (a), Ca (b) and Mg (c) concentrations in catchments where more than 50% of the area is classified as crop or hay/pasture land.

^a(USGS, 2024)

S2 Soil signal-to-noise Monte Carlo simulations

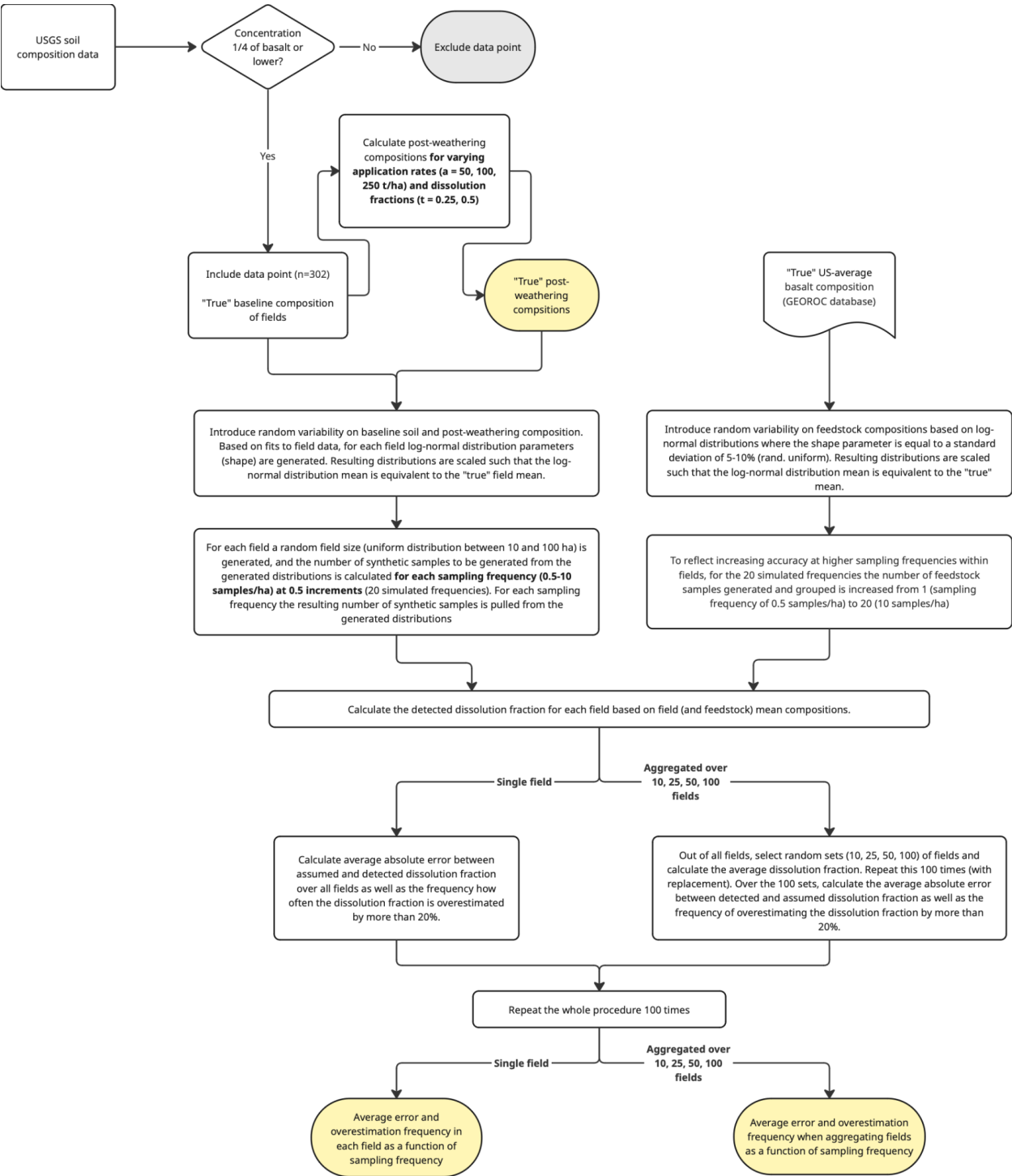
S2.1 Method details

To complement the simplified description of our soil analysis framework presented in the main text, we provide here a more detailed account of the data sources, assumptions, and modeling steps underlying the Monte Carlo simulations. This expanded description also outlines how baseline and post-weathering sample compositions are generated, how heterogeneity is parameterized, and how error metrics are derived for both single-field and aggregated-field applications.

The analysis uses multiple data sources to constrain the elemental composition of agricultural fields, rock powder, and expected in-field heterogeneity. To ground this analysis in realistic data of soil composition, we use US soil data classified as “Row Crops” and “Small Grains” (LandCover2) within the “Geochemical and mineralogical data for soils of the conterminous United States” database (Smith *et al* 2013). These samples are treated as the “true” baseline composition of fields, each datapoint in the database being used as one representative field composition. The composition for rock powder is based on the average composition of all samples contained in the GEOROC database that are classified as basalt and contained within the conterminous US (Lehnert *et al* 2000). Because the framework only works for feedstock-soil combinations whose composition is sufficiently different (Suhrhoff *et al* 2024), we only consider soil compositions where both base cation (here Ca^{2+} and Mg^{2+}) as well as immobile element (Ti) concentrations are at least 4 times lower than for basalt ($n=302$; see Figure 1b).

For each field, we calculate a “true” post-weathering soil-feedstock mix composition based on assumed application amounts ($a=50$ and 100 t ha^{-1}) and dissolution fractions ($\tau_i=0.25$ and 0.5) (Suhrhoff *et al* 2025). Note that application amount corresponds to the total cumulative amount deployed. Many EW studies apply $40 \text{ t ha}^{-1} \text{ yr}^{-1}$ such that even the highest rate modeled here may be realistic after multiple years of deployments (Beerling *et al* 2020, 2025). Furthermore, for each field a size between 10 and 100 ha is randomly generated (uniform distributions), which is a compromise between skewed US farm size distributions with most farm land being in farms larger than 2,000 ha but most farms being smaller than 72 ha (USDA 2022, 2024).

To constrain variance on field-level sample compositions resulting from spatial heterogeneity, we use a new dataset based on high-density spatial sampling (Suhrhoff *et al* 2025; cf. S2.2 below). This dataset of soil heterogeneity is based on new ICP-MS soil composition measurements (residual phase after exchangeable cations were leached with 1M ammonium acetate) from 5 field sites in the US with spatial sampling frequencies ranging from 0.6 – 19.8 samples ha^{-1} (7.1 – 39.6 pooled sub samples ha^{-1}). We fit log-normal distributions to field data (using the Python `scipy.stats` module), and use fitted shape parameters representing the standard deviations (σ) of the underlying normal distribution to model in-field variance—see section S2.2 for more detail on log-normal fits to field data. The shape parameters corresponding to field data are shown in Figure S9 and Figure S10, and uniform distributions between the range of observed shape parameters is used to generate synthetic σ values for Monte Carlo simulations of baseline soils where the resulting distributions are scaled such that the mean of the log normal distribution is equivalent to the field mean (see supplement S2.2). For feedstocks, heterogeneity is introduced by generating shape parameters of 5–10% (uniform distribution, i.e. σ values of 0.05 to 0.1 for generated log normal distributions).



876
877 Figure S8: Flow chart of the Monte Carlo simulation algorithm used for the main analysis. This approach
878 assumes non-paired sampling
879 We use these data to generate a simple statistical model based on nested Monte Carlo simulations (see
880 Figure S8) to assess expected errors on calculated dissolution fractions. The simulated in-field
881 heterogeneity is used to generate baseline and post-weathering soil sample compositions for a range of
882 sampling frequencies. To reflect increasing thoroughness of the sampling approach, as soil sampling

frequency increases from 1 to 20 samples ha^{-1} we also increase the number of total samples that the composition of the feedstock endmember is calculated from (from 1 to 20 samples). Next, the average composition of the realized samples is used to calculate the corresponding dissolution fraction. This is compared to the assumed, true value. This is done for each field 100 times, and we 1) calculate the average error over all fields for each sampling frequency as a realistic error estimate given US soil composition and realistic spatial heterogeneity, and 2) based on first averaging the calculated dissolution fraction based on realized sample composition over multiple fields (10, 25, and 50, done 100 fields based on random pulls from the 302 fields; 100 times) before computing the average error on the dissolution fraction (see Figure S8 for details). These reflect the error that one would get if applying such soil-based mass balance framework as the basis to quantifying CDR on average if applying it once either to an individual field or an aggregated set of fields at the same time. We use these error rates to assess the frequency with which rock powder dissolution is overestimated by 20% or more.

While the dissolution fraction is only a proxy for CDR, this approach can be translated to CDR as well using generated field sizes, application amounts, dissolution fractions, and assumed loss fractions, but we focus here on the dissolution fraction (or mass transfer coefficient) as it is the primary measured quantity. Note that by independently generating baseline and post-weathering samples, we model a non-paired sampling approach reflecting the worst case scenario where paired sampling is either not attempted or prevented due to bad GPS accuracy (typical GPS accuracy of 5 m). We also include additional analysis of paired sampling (cf. S2.3; Figure S11 and Figure S12).

S2.2 Implementation of soil heterogeneity in Monte Carlo simulations

We use soil composition data from five novel field sites sampled at high spatial densities to constrain in-field heterogeneity for the Monte Carlo signal-to-noise analysis. The data are normalized by the field mean concentration (Figure S9) before we fit log-normal distributions to make sure the population means are 1. The use of log-normal (rather than normal) distributions is intentional because samples generated from log-normal distributions always have positive values, preventing the occurrence of non-physical negative soil concentrations in the signal-to-noise analysis without having to filter some data. For normal distributions, this could be achieved by simply filtering out negative model occurrences, but this would change the mean of generated sample distributions and cause a systematic error in calculated dissolution fractions. In addition, using log-normal compared to normal fits also represents a conservative choice for the signal-to-noise analysis due to the generally higher variance, as well has overall better fits compared to normal distributions (R^2 better for 11 out of 20 elemental field distributions).

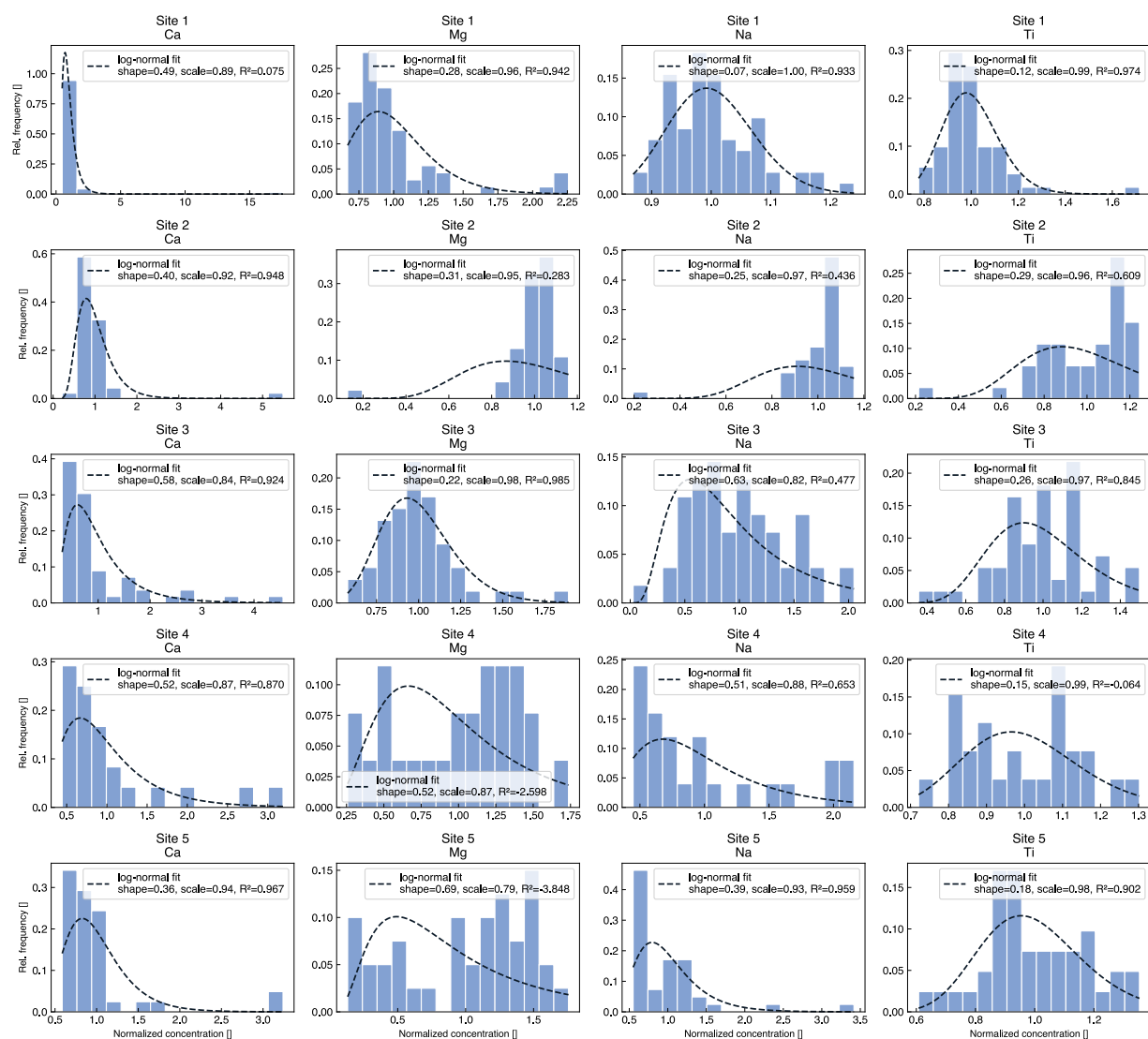


Figure S9: Distributions of baseline data for the 5 field sites (see Table S1 and Suhrhoff et al (2025) for more details) including log-normal fits to the data. The shape parameters, corresponding to the standard deviation of the normal distribution of the logarithm of the data, are plotted in Figure S10.

Table S1: Information on the field sites used to constrain spatial heterogeneity in the signal-to-noise analysis. The number of pooled cores corresponds to the number of-sub sample cores that were combined for each measured sample. Soil heterogeneity refers to the σ of log-normal fits to soil concentration distributions normalized to the field mean such that the resulting distribution has a mean of 1 (Figure S9). Site names are anonymized and location data are rounded to one decimal degree to protect farmer privacy.

Site name	Lat [°]	Lon [°]	size [ha]	# samples	# pooled cores	sample density [ha ⁻¹]	core density [ha ⁻¹]	soil heterogeneity (σ ; log-normal)			
								Ca	Mg	Na	Ti
								[]	[]	[]	[]
Site 1	45.3	-87.6	6.42	40	2	6.23	12.46	0.493	0.278	0.072	0.120
Site 2	42.3	-73.6	5.08	41	2	8.07	16.14	0.395	0.309	0.250	0.288
Site 3	31.3	-84.4	2.02	40	2	19.80	39.60	0.582	0.218	0.630	0.264
Site 4	35.8	-78.2	42.4 4	25	12	0.59	7.07	0.519	0.523	0.510	0.154
Site 5	35.8	-78.2	26.8 5	38	12	1.42	16.98	0.355	0.687	0.391	0.177

Generally, a random variable is log-normally distributed if:

$$X \sim \text{LogNormal}(\mu, \sigma) \quad 1$$

Which means that:

$$\ln(X) \sim N(\mu, \sigma^2) \quad 2$$

where μ is the mean, σ the standard deviation, and σ^2 the variance of the respective distributions, with log-normal distributions conventionally defined via the standard deviation of the underlying normal distribution. The expected value (mean) of a log-normal variable X can be calculated as:

$$E[X] = e^{\left(\mu + \frac{\sigma^2}{2}\right)} \quad 3$$

Hence, when using the parameters of log-normal fits to populations with a given mean (Figure S9) to generate synthetic data for the Monte Carlo simulations, if generating μ and σ independently, the mean of the resulting populations will not be the same as of the initial distribution (i.e., 1). Or said differently, if we want the mean of a synthetic distribution to be a specific value, μ and σ are not independent—only one can

be randomly generated. We implement this into the Monte Carlo simulation by randomly generating shape parameters (σ_{syn}) and then calculating μ_{syn} such that $E(X) = 1$:

$$E[X] = e^{\left(\mu_{syn} + \frac{\sigma_{syn}^2}{2}\right)} = 1 \quad 4$$

Now, taking the natural logarithm:

$$\ln\left(e^{\left(\mu_{syn} + \frac{\sigma_{syn}^2}{2}\right)}\right) = \ln(1) \Rightarrow \mu_{syn} + \frac{\sigma_{syn}^2}{2} = 0 \Rightarrow \mu_{syn} = -\frac{\sigma_{syn}^2}{2} \quad 5$$

The empirically constrained simulated μ_{syn} and σ_{syn} describe log-normal distributions with a mean of 1 and σ (shape) parameters constrained from field data (with a mean of 1), and are used to randomly generate sets of samples by multiplying these in-field variance factors with true “true” sample compositions.

Because the σ values from the fit to field data (Figure S9) are neither normally nor log-normally distributed (negative R^2 ; Figure S10), in the Monte Carlo simulations we generate synthetic σ_{syn} values by randomly pulling from uniform distributions set out by the minimum and maximum observed σ values observed in field data (for Ca, Mg, and Na the used values are 0.072402 and 0.687422, and for Ti 0.119775 and 0.288003).

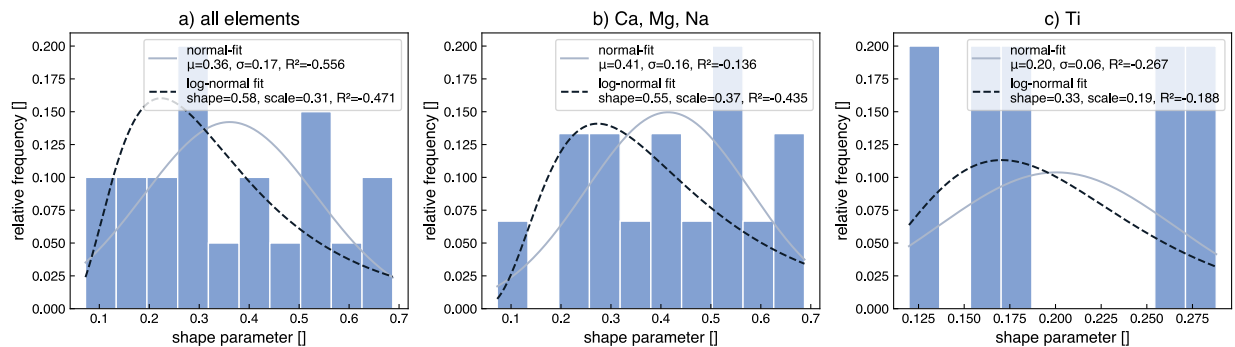


Figure S10: Histograms as well as normal and log-normal fits to the shape parameters from log-normal fits to soil data. The signal-to-noise analysis and related Monte Carlo simulations use uniform distribution set out by the minimum and maximum Ca, Mg, and Na shape values (b) as well as Ti shape values (c) due to low fit of both normal and log-normal distributions.

S2.3 Implementation of paired sampling in Monte Carlo simulations

In addition to the Monte Carlo signal-to-noise analyses that assume un-paired sampling (i.e., independently generate log-normal distributions for baseline and post-weathering soil compositions), we also simulate expected errors of the soil MRV approach when using paired samples.

Here, the work flow is adjusted for the generation of post-weathering samples (Fig S10, flowchart). As before, field data is used to generate synthetic log-normal distribution parameters for baseline samples (see methods and S2). However, for paired sampling, for each individual synthetic baseline sample the “true” post weathering composition based on the simulated feedstock application amount and dissolution fraction is calculated first. Next, we generate variance around “true” post-weathering compositions of baseline samples by generating a multiplier for each sample based on the generated log-normal shape parameter for this simulation, scaled such that the mean of the generated factors is 1 (eq. S51 (update from above)). Compared to the shape parameters used to simulate the baseline samples, the value of the shape parameter is reduced by 50% reflecting the efficacy of a paired sampling approach to reduce sampling variance. While arbitrary, this can be tested for any real deployment.

As expected, the paired sampling approach drastically reduces expected errors (Fig S11). While for individual fields, paired sampling is necessary to yield adequate errors, for aggregated monitoring lower errors are possible even when using a non-paired approach. Hence, the analysis suggests that paired sampling may not be necessary when using an aggregated monitoring approach.

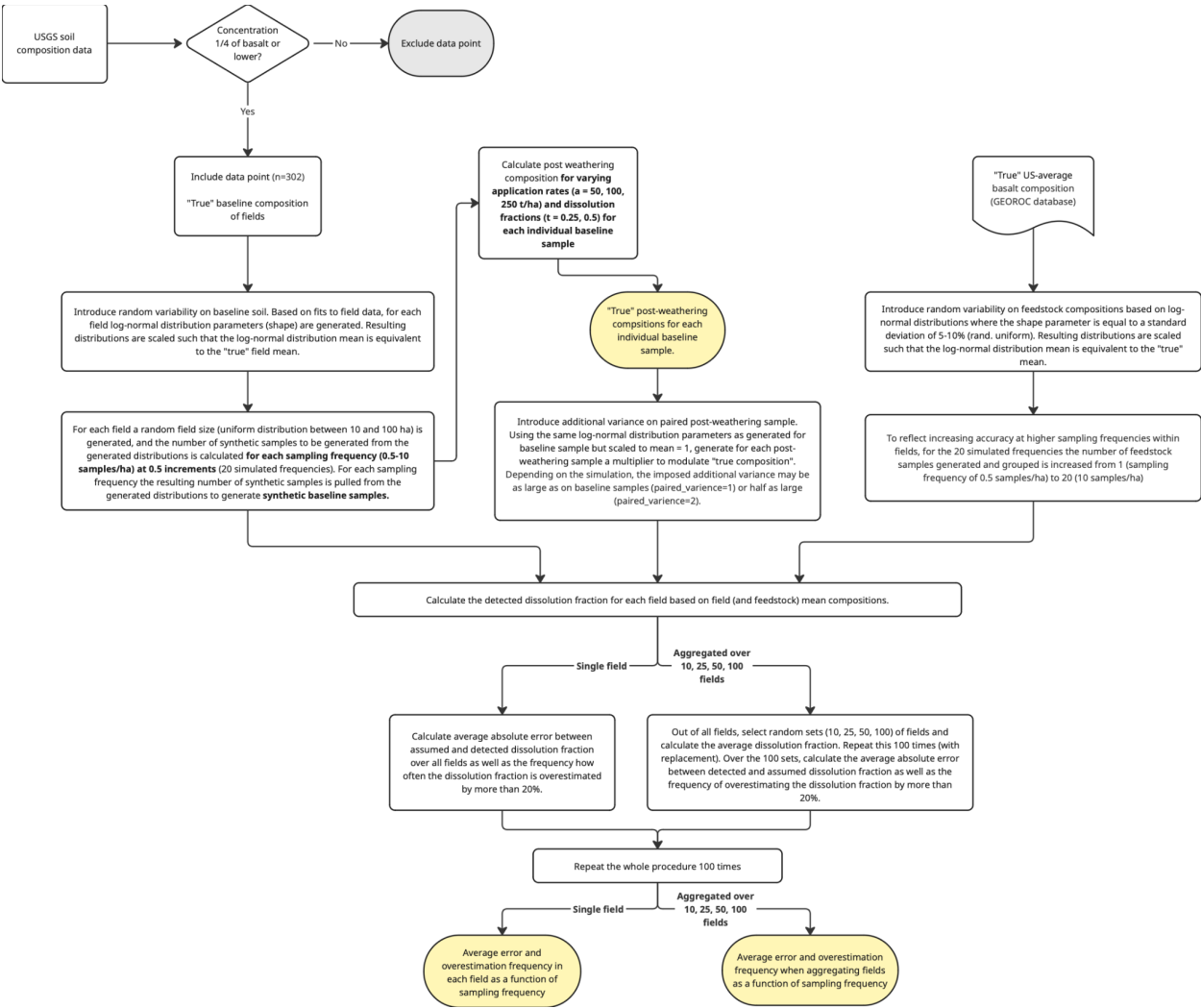
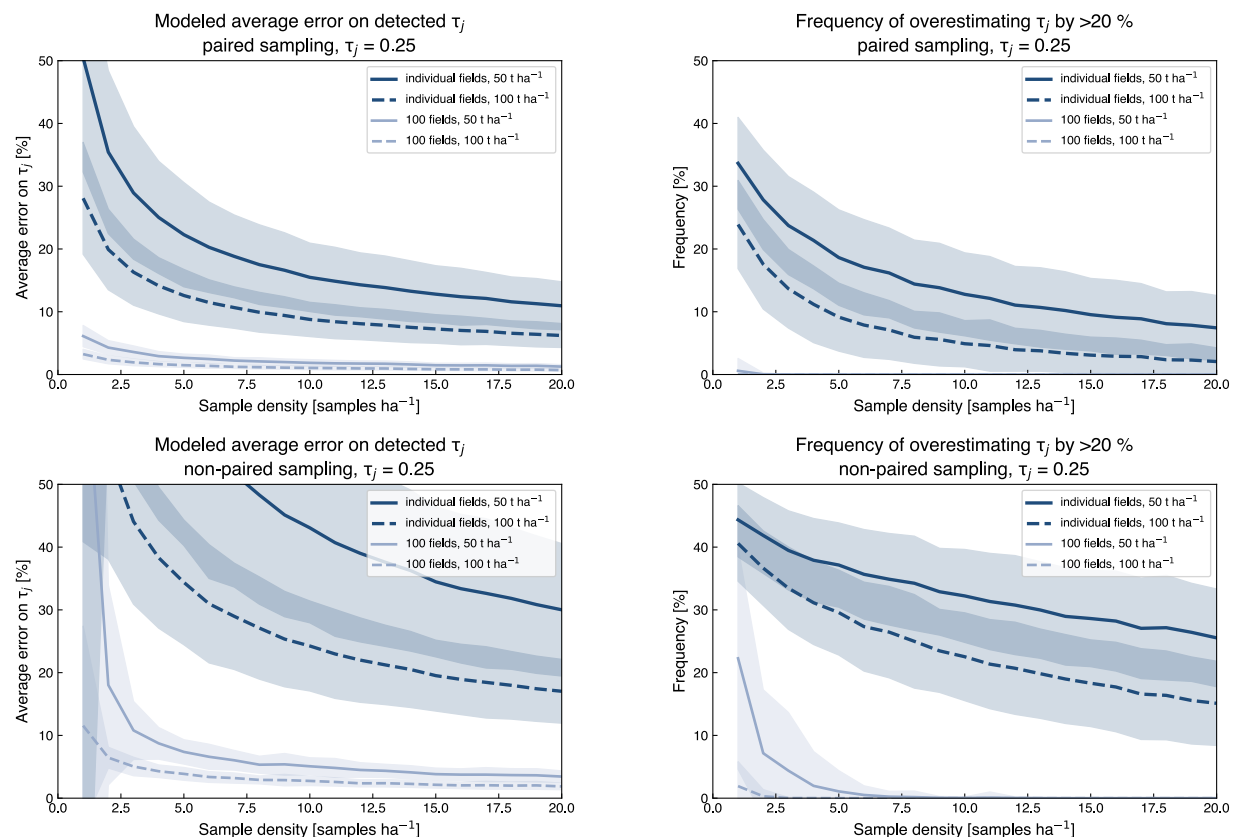


Figure S11: Flow chart for the Monte Carlo type signal-to-noise analysis using a paired sampling approach where post-weathering samples are taken at the same sites as baseline samples.



988

989 Figure S12: Comparison of selected signal-to-noise simulations between simulated to paired sampling (top
 990 row) and non-paired sampling (bottom row). Panels a and c show average error on detected mass transfer
 991 coefficients (i.e., dissolution fraction), b and d the frequency of overestimating mass transfer
 992 coefficients by more than 20% for constant τ_j but variable application amounts (a and b) as well as at constant application
 993 amount but variable τ_j (c and d).