A Framework for Integrating Spatial Uncertainty into Critical Zone Models: Application to Enhanced Weathering

Brian Rogers* and Kate Maher

Department of Earth System Science, Stanford University, Stanford, CA

E-mail: rogersdb@stanford.edu

Abstract

Spatial heterogeneity introduces uncertainty when characterizing the Critical Zone, especially when sampling is sparse or requires repeated measurements at the same locations. Here, we layout a probabilistic sequential framework to systematically account for spatial uncertainty when measuring Critical Zone transformations. First, we use measurement variance propagation and distance-based sensitivity analysis to determine measurement variance criteria for meeting overall uncertainty requirements. We then stochastically simulate spatial fields and composite sampling to infer a minimally sufficient sampling plan that meets these criteria. Throughout the study, we apply this framework to solid-phase measurement of enhanced weathering, an open-system carbon dioxide removal strategy. Results indicate that field-scale variance in baseline soil concentrations must be accurately estimated before designing a sampling plan and, even then, such variance is likely too high for element-element mixing models to be effective near-term constraints on enhanced weathering. We conclude with opportunities to extend this framework to other solid-phase mixing and stock models, multi-phase measurement models, and transient Critical Zone processes.

Keywords: sensitivity analysis; Bayesian; carbon dioxide removal; CDR; soil-based; monitoring, reporting, and verification; MRV; solid-phase

Synopsis: This study addresses the need for standardized uncertainty analysis and reporting in Critical Zone calculations, especially for open-system carbon dioxide removal applications.

Introduction

Earth's Critical Zone (CZ) extends from the heights of vegetative canopies to the depths of weathering bedrock, encompassing a layer of regolith that interacts with the atmosphere and supports terrestrial and aquatic life [1–4]. The CZ thus embodies a complex system of physical and biogeochemical states, transformations, and fluxes that are subject to dynamic atmospheric and anthropogenic forcings, resulting in spatial and temporal heterogeneity that introduces significant uncertainty into empirical characterization [5–8]. Notwithstanding such complexity, a multitude of CZ descriptors need to be quantified for a broad range of applications, including contaminant monitoring and remediation, agronomic operations, ecosystem preservation, geotechnical engineering, and, increasingly, atmospheric greenhouse gas removal.

Spatial uncertainty in certain physical and chemical measurements, such as hydraulic conductivity [9, 10] and soil compositions [11, 12], often exceeds their temporal variability. As such, spatial heterogeneity can be difficult to constrain through point measurements, which are usually limited by logistical considerations combined with lack of predefined metrics to guide appropriate sampling [13]. Furthermore, application of existing geostatistical frameworks is not always straightforward for complex, high-dimensional CZ models. For example, calculation of both chemical depletion fractions and mixing between different lithologies, using concentrations or ratios, requires error propagation for multiple variables through linear and non-linear governing equations. Analytical propagation methods involve potentially un-

realizable assumptions of normality and linear and independent error [14]. High-dimensional, non-additive calculations thus warrant some application of non-parametric, bootstrapping, stochastic, hierarchical, and/or Bayesian methods.

A current CZ application that relies on characterizing different soil compositions is solid-phase measurement of enhanced weathering (EW) [15, 16]. EW involves amending soils, usually agricultural, with a reactive "feedstock", such as basalt or Mg-silicate, to shift the alkalinity of the system and effectively dissolve additional CO_2 [17, 18]. Given the low initial enrichment of feedstock mass relative to native soil, detecting feedstock weathering beyond the "noise" of the baseline becomes a challenging problem [16], and a common current approach relies on the ratio of base cations to immobile elements to constrain this weathering. The resulting depletion and mixing equations (Eq. 1-2; Supp. 1) require analysis of multiple soil samples in sequence, from soil (baseline) to the initial mixture (soil + feedstock) to weathered compositions over multiple time points [19, 20]. These measurements are used to estimate the true fraction of feedstock dissolved (f_d), a potential proxy for CO_2 removal, calculated as

$$f_d = 1 - \frac{[M]_{mix} - [M]_{bsln}}{[M]^0_{mix} - [M]_{bsln}} \tag{1}$$

where $[M]$ is base cation concentration of the baseline ($_{bsln}$), initial mixture ($^0_{mix}$), and weathered mixture ($_{mix}$). A multiplier is used to compute gross carbon dioxide removal (CDR) from f_d, hence an accurate estimate ($\hat{f}_d$) is the focus here. A common approach for calculating $[M]^0_{mix}$ is to also measure an immobile tracer (T) and use the following element-element mixing equation with baseline and feedstock ($_{fs}$) endmembers,

$$[M]^0_{mix} = [M]_{bsln} + \frac{([T]_{mix} - [T]_{bsln})([M]_{fs} - [M]_{blsn})}{[T]_{fs} - [T]_{bsln}} \tag{2}$$

Here, heterogeneity can interrupt basic assumptions of the mixing model if samples are not representative of the same "system" (Fig. 1). Uncertainty in these assumptions was

initially considered in the context of analytical variance [20], and it has recently been demonstrated that such calculations are highly sensitive to spatially heterogeneous soil compositions [21, 22]. It is challenging, however, for standards development organizations (SDOs) to provide specific guidance to project developers (PDs) on constraining such uncertainty while also remaining logistically feasible and cost-effective, especially considering the broad experimentation with measurement approaches [15, 16].

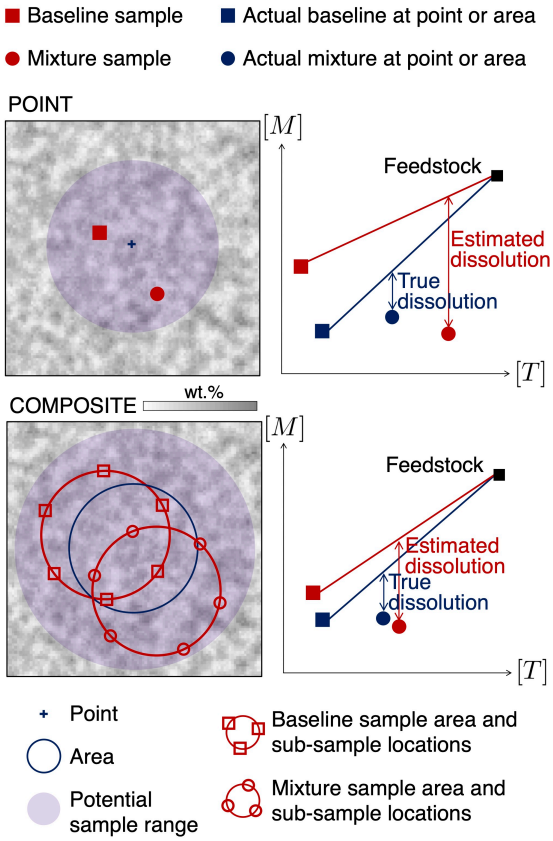


Figure 1: Illustration of how spatial uncertainty may introduce error into a feedstock dissolution calculation. In this example, the underlying mixing model relies on the assumption that baseline and mixture samples are representative of the same system. Considering spatial heterogeneity of soil compositions, this assumption may be interrupted by positioning error, tillage, and erosion, as well as sampling and preparation techniques and analytical precision. Composite sampling is commonly used to reduce the impact of spatial heterogeneity by reducing sample variance.

Here, our goal is to provide a prescriptive and robust approach to account for spatial uncertainty in high-dimensional measurement models of CZ transformations. We describe this

approach in the context of solid-phase measurement of EW, though it can be extended to other parameters that need to be precisely defined, for open-system CDR or other environmental considerations, such as organic carbon stocks [23, 24] or isotope signatures [25, 26]. The framework starts with definition of overall uncertainty requirements (*e.g.*, a maximum error and minimum confidence in $\hat{f}_d$), followed by variance propagation and sensitivity analysis (SA) to help define corresponding measurement variance requirements (*e.g.*, a maximum measurement variance and minimum confidence). We then use stochastic sampling simulations to infer a sampling approach that minimally meets these requirements and conclude with estimate and uncertainty reporting. This is detailed through the following steps:

1. **Define the uncertainty requirements and measurement model**, the latter including explicit relationships between input and output uncertainty using hierarchical Bayesian principles.

2. **Determine the maximum measurement variances** that fulfill the overall uncertainty requirements using variance propagation and SA.

3. **Define the measurement variance requirements and sampling model**, the latter involving stochastic simulation of spatial fields and composite sampling plans.

4. **Design a sampling plan** that minimally meets the measurement variance requirements using the sampling model and SA; if infeasible, reconsider the overall uncertainty requirements or measurement model.

   (Data collection)

5. **Report the final estimate and overall uncertainty**, with traceable and reproducible uncertainty quantification.

Integrated Methods and Results

For enhanced weathering, agricultural fields are typically chosen based on accessibility rather than a detailed understanding of soil properties and heterogeneity. The goal for the PD is to perform minimal sampling while still accurately calculating the amount of dissolved feedstock and attendant CDR.

In this example, we illustrate how early characterization can be integral to EW site selection and monitoring design to increase the likelihood of precisely quantifying CDR. We assume deployment of a basaltic feedstock, though the approach is generalizable to any amendment. For solid-phase verification of EW, the *measurement model* consists of equations (1) and (2), which are solved for f_d based on the measured baseline and, either the initial soil-feedstock mixture to determine application rates, or the mixture after some weathering has occurred [20]. Field trials [27–37] report feedstock application rates ranging from 5 to 100 tons per hectare (ha), resulting in relatively low mass enrichment of 0.1–3% after mixing within the upper 20 cm of soil (Supp. 2). Another important consideration is the chemical differentiation between the feedstock and the baseline, which we analyze using the feedstock-baseline ratio of mean cation concentration ($\mu_M^{fs:bsln}$) and mean tracer concentration ($\mu_T^{fs:bsln}$).

The *sampling model* outlines the planned configuration for sampling—whether through discrete point samples or carefully homogenized composite samples—which we stochastically analyze to infer the measurement variance associated with different sampling strategies. Because the measurement model depends on the baseline, the sampling plan is typically fixed after the feedstock is applied, underscoring the need for a robust baseline sampling strategy. The uncertainty requirements are defined by operational constraints (such as the need to present a compensatory claim) and reflect the probability that the resulting estimate will fall within a specified range of the true value.

1. Define the uncertainty requirements and measurement model

The goal of this first step is to define the problem mathematically to allow for rigorous variance propagation and SA. In EW, the baseline variance in soil elemental abundances is typically unknown prior to site selection, and recent work [21] suggests that variability in an immobile tracer element tends to exceed that of base cations, thereby dominating the total uncertainty. Consequently, the site-specific variance in these elements determines whether a given measurement approach is likely to fulfill the uncertainty requirements.

Uncertainty requirements

The uncertainty requirements are often defined by an SDO and, here, encapsulate:

- ϵ_{max}, the maximum relative error in $\hat{f}_d$ and thus CDR.
- p_{min}, the minimum probability that the relative error in $\hat{f}_d$ is less than ϵ_{max}.

For instance, an SDO specification might require 90% confidence ($p_{min} = 0.9$) that the reported CDR is within 10% of the true value ($\epsilon_{max} = 0.1$) [*e.g.*, 38]. There is no *a priori* guarantee, however, that any particular field deployment can meet these requirements for a specific site, due to the inherent variability in measurement conditions and system parameters.

We use ϕ to represent the outcome where the uncertainty requirements are fulfilled. Formally:

$$p(\epsilon \leq \epsilon_{max}) \geq p_{min} \implies \phi, \quad p(\epsilon \leq \epsilon_{max}) < p_{min} \implies \overline{\phi}, \tag{3}$$

where $\overline{\phi}$ indicates the requirements are not fulfilled.

Measurement model and parameter set

To calculate $p(\epsilon \leq \epsilon_{max})$ for different measurement approaches, we define a measurement model parameterized by θ. This model includes:

- a **measurement function** (*e.g.*, feedstock dissolution calculation, Eq. 1-2)
- **input parameters** (*e.g.*, M and T concentrations in the baseline, feedstock, and mixture)
- the **function response** (*e.g.*, fraction of feedstock dissolved, f_d),
- **measurement variance** of each input parameter (*e.g.*, spread of possible M and T measurement values given the point or composite sampling scheme), and
- **operational parameters** (e.g., feedstock-baseline differentiation, application rate, true fraction dissolved).

For this measurement function (Eq. 1-2), the input parameters are $[M]_{bsln}$, $[M]_{fs}$, $[M]_{mix}$, $[T]_{bsln}$, $[T]_{fs}$, and $[T]_{mix}$. Due to spatial heterogeneity, baseline and mixture measurements may be highly variable, interrupting their assumed comparability (Fig. 2). To account for the impact of spatial heterogeneity, we consider each input parameter to have a distribution of possible measurement values—characterized by a mean (μ) and variance (υ)— which we will propagate through the measurement function in step 2. While measurement variance represents aggregate spatial and analytical uncertainty, we only consider spatial uncertainty in this study, as analytical uncertainty can be made negligible if necessary [20]. Accordingly, we define μ and υ for the measurement distributions as follows:

Measurement means: We set μ_M^{bsln} and μ_T^{bsln} (baseline means) as constants at the simulation scale. We then specify feedstock-to-baseline ratios ($\mu_M^{fs:bsln}$, $\mu_T^{fs:bsln}$) to obtain mean feedstock concentrations. A uniform application rate (r_{app}) and uniform fraction dissolved (f_d) together determine the mean mixture concentrations after amendment and weathering. This f_d also serves as the "true" fraction against which estimation errors are calculated.

Measurement variances: We specify υ_M and υ_T, the measurement variances for baseline M and T. Due to feedstock mass enrichments of $< 3\%$, we assume the measurement variances

of the soil-feedstock mixture are equal to the baseline υ_M and υ_T. Also, since CZ parameters tend to span orders of magnitude, υ_M and υ_T are relative quantities (akin to coefficients of variation, CV). The feedstock itself is assumed homogeneous (negligible variance).

Collectively, these parameters form the set θ. Since we want to test the impact of different means, measurement variances, and operational parameters on fulfilling the uncertainty requirements, we initially consider a wide range of possible values for each parameter. These ranges are used as bounds for uniform cumulative density functions (CDFs), denoted $\hat{F}$, and $\hat{F}(\theta)$ is the multivariate CDF describing the entire parameter space (Table 1). The next step involves random sampling of this parameter space to rigorously evaluate the individual and joint impacts of each parameter on ϕ.

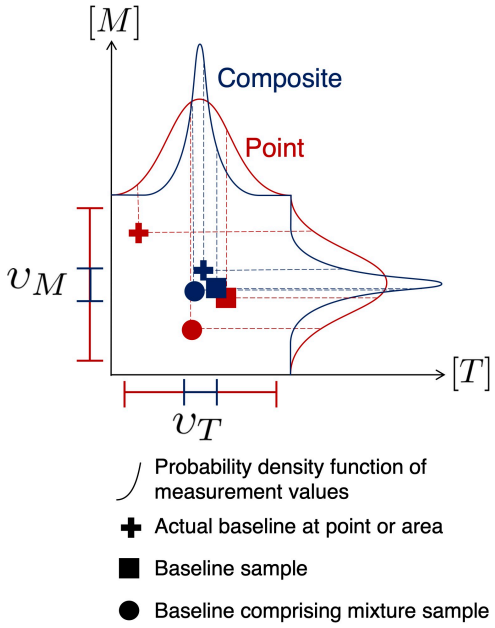


Figure 2: Illustration of how measurement variance (υ) in cation (M) or tracer (T) concentrations may interrupt the assumed comparability of baseline and mixture samples to differing degrees depending on sampling approach. The goal of step 2 is to calculate how small υ_M and υ_T must be, or how narrow each probability density function of measurement values must be.

Table 1: Exploratory $\hat{F}(\theta)$, uniform distributions used for variance propagation and SA to determine measurement variance requirements.

Parameter	**Symbol**	**Units**	**Range** (Transformation)	**Source**
Mean baseline concentrations	μ_M^{bsln}	mg/kg	$[7, 11]$ $(\ln[1100, 60000])$	Lower and upper bound are lowest p_{10} and highest p_{90} of, for M, Ca, Mg, Na, and K and, for T, Ti, Cr, and Ni concentrations in upper 5 cm of CONUS soils [39].
	μ_T^{bsln}	mg/kg	$[1, 9]$ $(\ln[2.72, 8100])$	
Measurement variances	υ_M	-	$[-9, -3]$ $(\ln[0.00012, 0.050])$	Ranges chosen such that resulting variance propagation includes significant amounts of realizations that do and do not fulfill the uncertainty requirements.
	υ_T	-	$[-9, -3]$ $(\ln[0.00012, 0.050])$	
Feedstock-baseline mean concentration ratios	$\mu_M^{fs:bsln}$	-	$[1, 75]$	Ranges computed by dividing concentrations of selected elements in six basalt compositions reported by Lewis et al. [40] by corresponding p_{10} and p_{90} CONUS soil concentrations [39].
	$\mu_T^{fs:bsln}$	-	$[1, 49]$	
Application rate	r_{app}	tons/ha	$[2, 100]$	Annual rate of 5-100 t/ha across eleven EW field trials [27–37]; lower bound extended to 2 t/ha based on conversations with EW PDs.
True fraction of feedstock dissolved	f_d	-	$[0.1, 0.3]$	Range chosen based on values reported by Beerling et al. [28], where $\hat{f}_d = 0.12$ using Mg, 0.32 using Ca, four years after initial amendment.

2. Determine the maximum measurement variances

With the measurement model and parameter ranges established, the next step is to analyze the sensitivity of the model and determine how small the measurement variances in M and T must be to meet the overall uncertainty requirements. This process involves propagating measurement variances through the measurement model (step 2.1), quantifying the influence of measurement variance on the accuracy of $\hat{f}_d$ (step 2.2), and constraining operational parameters to set maximum measurement variances (υ_{max}) for a specific deployment (step 2.3).

2.1 Propagate measurement variances through the measurement model

Here, we use nested Monte Carlo simulations to jointly vary the input means, measurement variances, and operational parameters encompassed by θ and, for each variation, compute the resulting $p(\epsilon \leq \epsilon_{max})$. This process begins with generating 10^4 *parameter realizations*, or samples of $\hat{F}(\theta)$. $\hat{F}(\theta)$ is a uniform multivariate distribution, meaning each parameter range in Table 1 is sampled from uniformly, and each realization represents a possible combination of baseline means (μ_M^{bsln}, μ_T^{bsln}), measurement variances (υ_M, υ_T), and operational parameters ($\mu_M^{fs:bsln}$, $\mu_T^{fs:bsln}$, r_{app}, f_d). $\hat{F}(\theta)$. For a given parameter realization, we use these values to construct Gaussian measurement distributions for $[M]_{bsln}$, $[M]_{mix}$, $[T]_{bsln}$, and $[T]_{mix}$, and we sample from these distributions to generate 10^4 *measurement realizations*. For each measurement realization, we compute $\hat{f}_d$ and its relative error (ϵ), such that $p(\epsilon \leq \epsilon_{max})$ for each parameter realization is the fraction of its measurement realizations where $\epsilon \leq \epsilon_{max}$.

Results of these simulations indicate that ϕ is highly dependent on keeping the measurement variances below critical thresholds (Fig. 3), while other parameters, such as μ_M^{bsln} and μ_T^{bsln}, have minimal impact. Specifically, the distribution of $p(\epsilon \leq \epsilon_{max})$ shows a clear divide (Fig. 3A), indicating that while many realizations achieve ϕ, a significant number fail. The p_{10}-p_{90} grey-shaded regions in Fig. 3B illustrate the spread of simulation outcomes across each parameter range—shaded regions that extend above p_{min} (red-dashed line) indicate pa-

rameter values for the realizations that achieved ϕ in Fig. 3A. Conversely, unshaded regions above p_{min}, most notable for high υ_M and υ_T and low r_{app}, correspond to parameter values that will not result in ϕ, effectively mapping the forbidden ranges for each parameter.

The correlations among the expected $p(\epsilon \leq \epsilon_{max})$—black p_{50} lines in Fig. 3B—and the individual parameters provide additional insight into sensitivities. For example, the conditional distributions of $p(\epsilon \leq \epsilon_{max})$ show a strong negative correlation with υ_M and υ_T, meaning high measurement variances make it unlikely to achieve $\leq 10\%$ error in $\hat{f}_d$ (Fig. 3B). Intuitively, $\mu_M^{fs:bsln}$, $\mu_T^{fs:bsln}$, r_{app}, and true f_d show moderate positive correlations with $p(\epsilon \leq \epsilon_{max})$, indicating that greater values tend to increase the expected accuracy in $\hat{f}_d$. Overall, the wide range of outcomes here emphasizes the importance of considering all possible outcomes early in site selection and monitoring design.

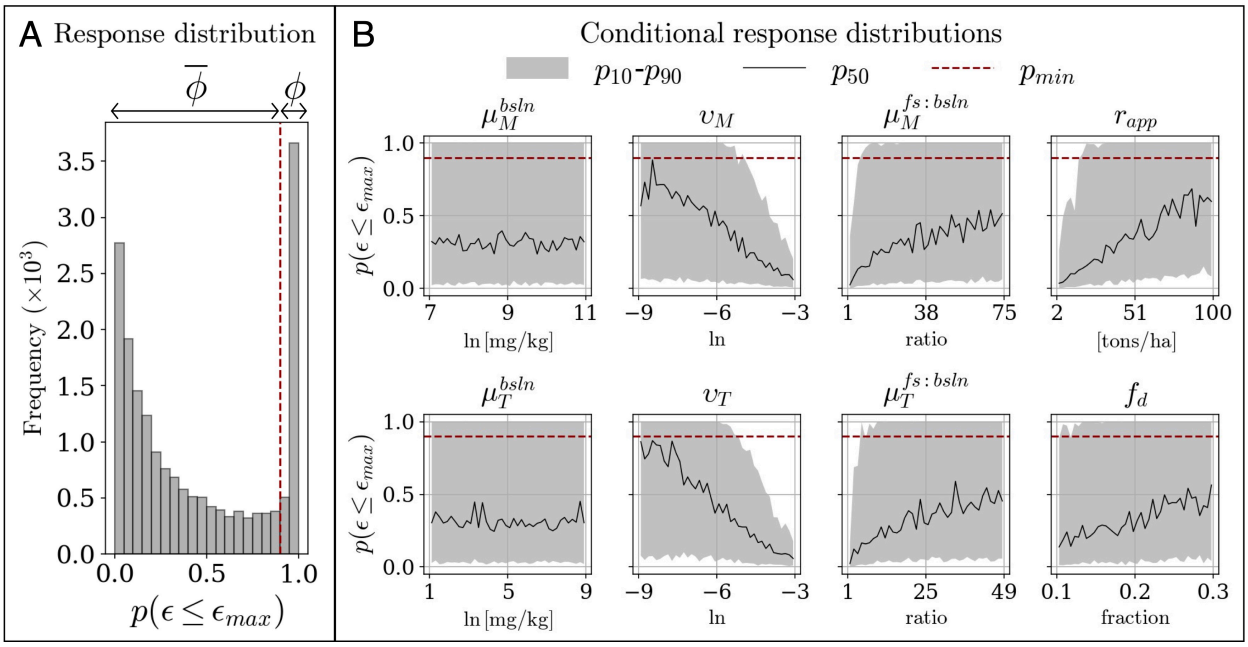


Figure 3: Exploratory SA of the influence of means (μ), measurement variances (υ), and operational parameters on $p(\epsilon \leq 0.1)$, the probability that the relative error in $\hat{f}_d$ is no greater than 10%, using the parameter ranges in Table 1. (A) shows the response across all 10^4 realizations, with a red-dashed line separating the realizations that do (ϕ) and do not ($\overline{\phi}$) fulfill uncertainty requirements of $p(\epsilon \leq 0.1) \geq 90\%$ (p_{min}). (B) provides the conditional response distribution for each parameter.

2.2. Quantify parameter influence on fulfilling the uncertainty requirements

To rigorously compare the sensitivity of ϕ to different parameters, we separate the 10^4 realizations of θ into one group that does fulfill the uncertainty requirements (ϕ) and one group that does not ($\overline{\phi}$). This can be represented by partitioning $\hat{F}(\theta)$ into two conditional distributions, $\hat{F}(\theta|\phi)$ and $\hat{F}(\theta|\overline{\phi})$, and we can analyze the differences between these distributions to determine which parameters most significantly influence the outcome. A common way to quantify such sensitivities [41–44] is to compute the "distance" between $\hat{F}(\theta|\phi)$ and $\hat{F}(\theta|\overline{\phi})$ for each parameter (Fig. 4A), normalizing the parameter ranges to $[-1, 1]$ so they do not influence comparison of the distances. The resulting sensitivity rankings (Fig. 4B) highlight that ϕ is most influenced by υ_M and υ_T, and less so by r_{app}, $\mu_T^{fs:bsln}$, $\mu_M^{fs:bsln}$, and true f_d. Collectively, this emphasizes the dominant role of measurement variances in determining success.

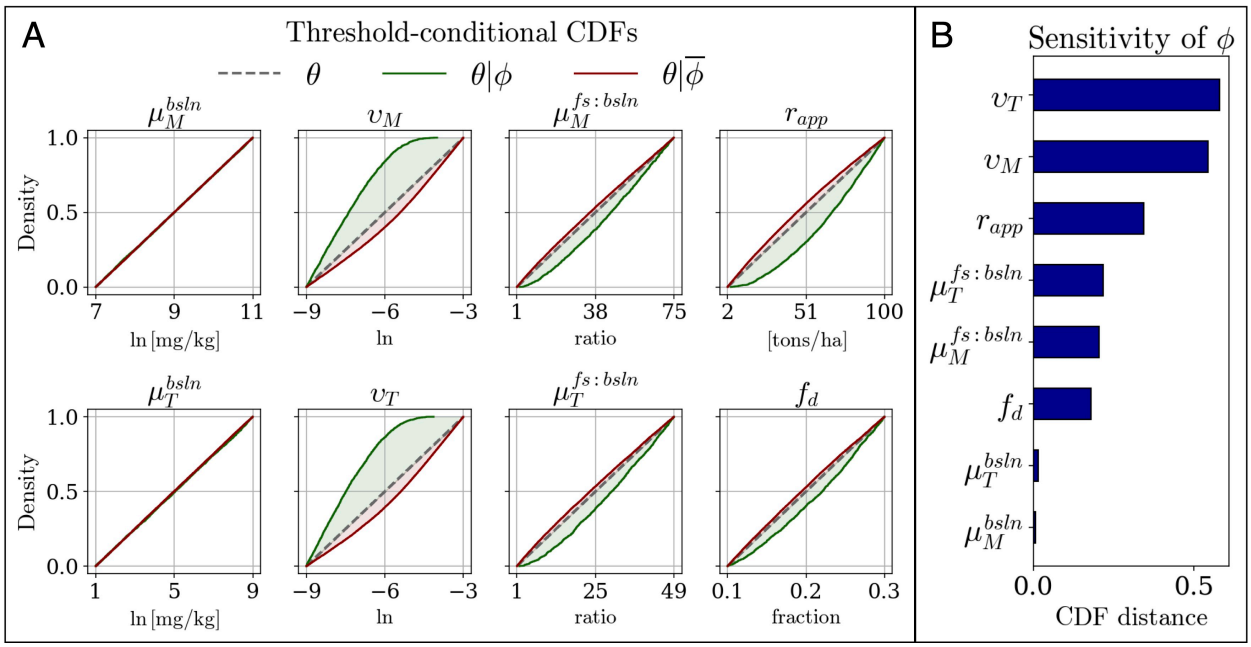


Figure 4: Distance-based sensitivity calculations for the exploratory SA in Fig. 3. (A) shows threshold-conditional CDFs (*e.g.*, partitions of the entire set of realizations (grey-dashed line) into realizations that did (green line) and did not (red line) fulfill uncertainty requirements, with shaded areas to visualize distances between CDFs. (B) provides a ranking of the parameters according to their influence on ϕ using this distance-based sensitivity metric.

2.3. Apply deployment-specific constraints and identify measurement variance limits

In practice, PDs can constrain certain parameters in θ, such as feedstock-baseline differentiation and application rate. For our theoretical deployment, we constrain $\mu_M^{fs:bsln}$ to 38 and $\mu_T^{fs:bsln}$ to 25 (midpoints from Table 1) and r_{app} to 40 tons/ha (median from Table 1). Performing the SA with these constraints (Fig. 5) reveals that the expected, or median, $p(\epsilon \leq \epsilon_{max})$ exceeds p_{min} for $\ln(\upsilon_M)$ and $\ln(\upsilon_T)$ less than approximately -7 (Fig. 5B). In contrast, the conditional response distributions for other parameters do not show an expected $p(\epsilon \leq \epsilon_{max})$ greater than p_{min}, as each distribution assumes values for all other parameters are randomly chosen from their respective ranges, thus incorporating effects from the entire ranges of υ_M and υ_T. While measurement variances are the primary control here, the true f_d will likely become significant after constraining υ_M and υ_T (Fig. 5C). This suggests that delaying intensive sampling, though also delaying return on investment to the PD, could be a key feature of profitable operations.

To determine specific measurement variance limits, we need to account for potential interactions between υ_M and υ_T by analyzing their joint conditional distribution (Fig. 6). It is also important to consider the entire $\hat{F}(\theta)$ when determining such limits. For example, using a wide, exploratory $\hat{F}(\theta)$ results in almost no combinations of υ_M and υ_T that achieve ϕ (Fig. 6A). Using the constrained $\hat{F}(\theta)$, however, indicates the expected outcome is ϕ when both $\ln(\upsilon_M)$ and $\ln(\upsilon_T)$ are greater than approximately -6.5 (Fig. 6B). Since, for this example, υ_M and υ_T exert similar influences on ϕ, we select a single υ_{max} of $e^{-6.5}$. The remaining analysis provides information on combinations of inherent site characteristics and sampling designs that could likely adhere to this maximum using stochastic simulations of spatial variability and composite sampling.

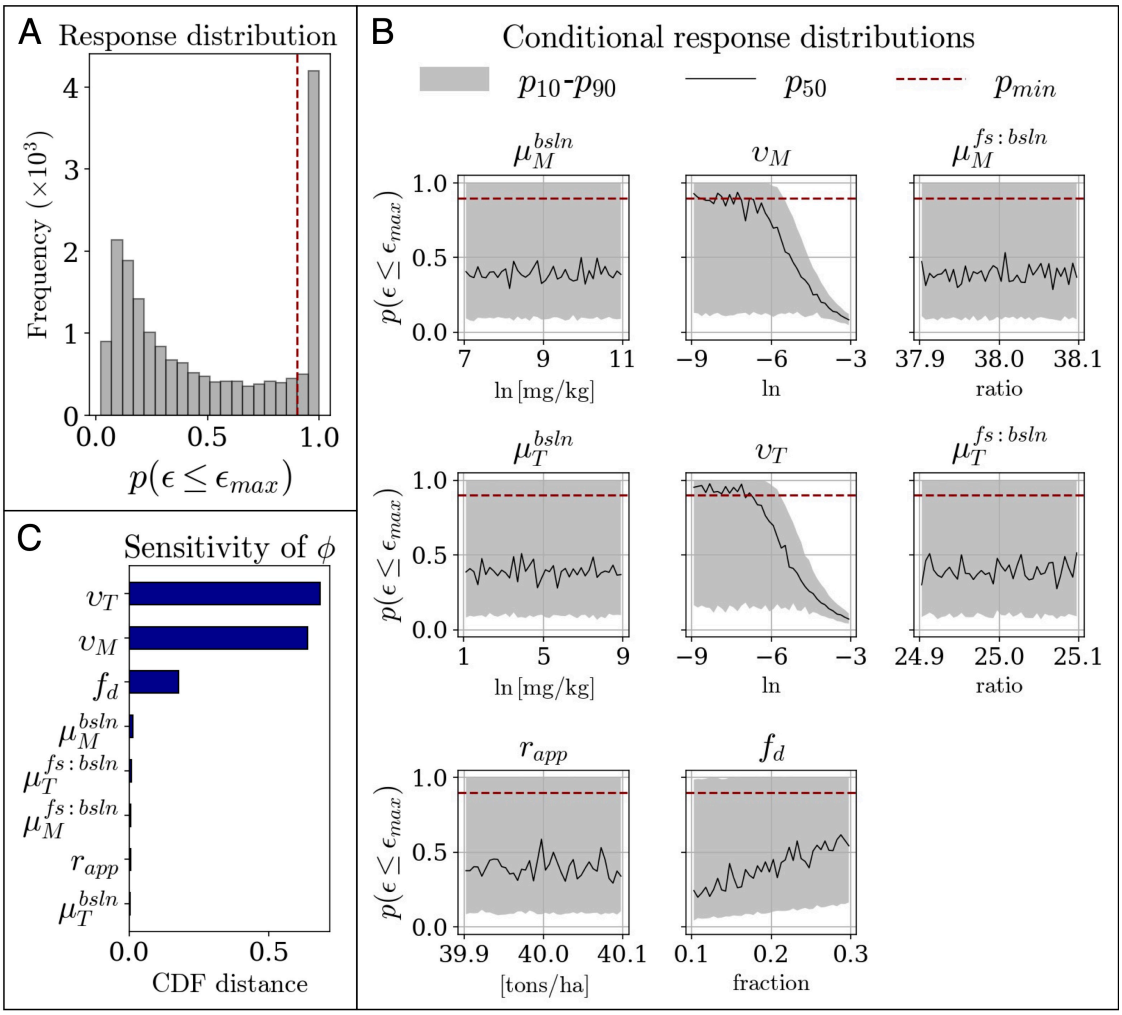


Figure 5: Deployment-specific SA where, relative to the exploratory SA in Fig. 3 and Fig. 4 and parameter ranges in Table 1, we apply constraints to soil-feedstock differentiation ($\mu_M^{fs:bsln} = 38$, $\mu_T^{fs:bsln} = 25$) and application rate ($r_{app} = 40$ tons/ha). (A) shows the updated response across all 10^4 realizations, (B) the updated conditional response distributions, and (C) the updated ranking of the parameters according to their influence on ϕ.

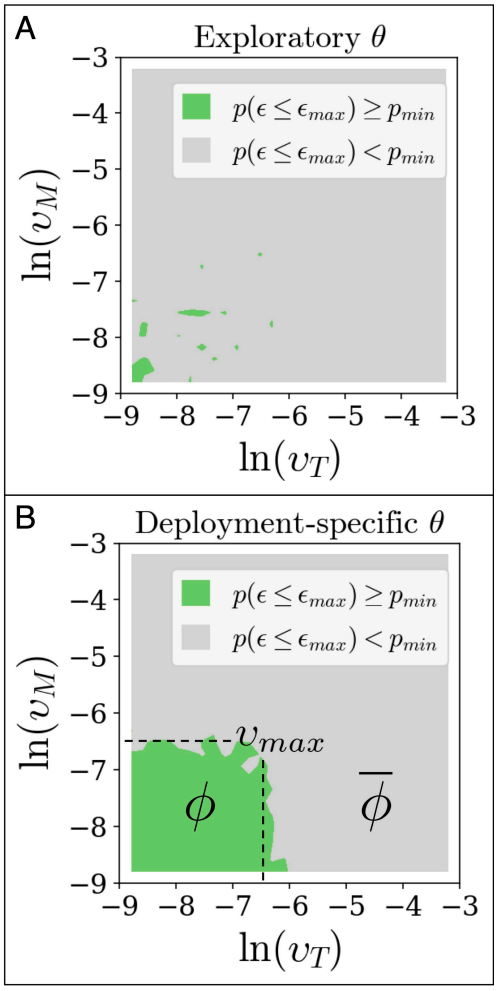


Figure 6: Combinations of base cation measurement variance (υ_M) and immobile tracer measurement variance (υ_T) that result in fulfilling uncertainty requirements (ϕ) of at least 90% likelihood of $\leq 10\%$ error in $\hat{f}_d$ for (A) loosely constrained, exploratory parameter ranges and (B) constrained parameter ranges for a theoretical deployment where $\mu_M^{fs:bsln} = 38$, $\mu_T^{fs:bsln} = 25$, and $r_{app} = 40$ tons/ha.

3. Define the measurement variance requirements and sampling model

With υ_{max} calculated, and before data collection, we need to design a suitable and efficient sampling plan. To determine sufficient sampling plans, one would in theory need to know the concentrations everywhere across the field site at high spatial resolution. Presumably, this would reveal lenses and patches of similar material, as opposed to a completely random distribution. Alternatively, we can create synthetic deployment fields based on models of spatial variability, an approach similar to that used in hydrogeology [45, 46], and sample them to develop measurement schemes that are robust across different types of spatial variability. In steps 3 and 4, the objective is to simulate spatial fields and composite sampling plans to determine approaches for achieving a measurement variance lower than υ_{max}, and then refine these approaches to roughly minimize the number of samples.

Given that high-density sampling over large deployment areas is not feasible, we assume identification of a representative 1-ha plot for high-density sampling with low-density sampling still performed across the remainder of the area, similar to plot designs recommended by SDOs [*e.g.*, 47]. Specifically, we are simulating 1 ha (10,000 m^2) at 0.1-m resolution, thus using a 1,000 by 1,000 structured grid, which could analogously be described as 100 ha at 1-m resolution or 10,000 ha at 10-m. Ideally the resolution or "support size" mimics physical sample collection, *e.g.*, individual core area when simulating at the sub-sampling scale, or compositing area if each sample is representative of a grid cell.

Measurement variance requirements

Since we have chosen the same υ_{max} of $e^{-6.5}$ for υ_M and υ_T, we can generally denote both $[M]$ and $[T]$ as an arbitrary spatial variable Z. Here, the measurement variance requirements for Z are defined by:

- υ_{max}, the maximum allowable measurement variance in Z,

- p_{min}, the minimum probability that the measurement variance in Z is below υ_{max}.

For a given spatial field and sampling plan, $p(\upsilon \leq \upsilon_{max})$ is the likelihood that the resulting measurement variance (υ) will be less than the maximum measurement variance (υ_{max}). The measurement variance requirements are fulfilled when $p(\upsilon \leq \upsilon_{max})$ exceeds the probability threshold p_{min}, and the corresponding outcome is denoted ϕ_Z. Formally:

$$p(\upsilon \leq \upsilon_{max}) \geq p_{min} \implies \phi_Z, \quad p(\upsilon \leq \upsilon_{max}) < p_{min} \implies \overline{\phi}_Z. \tag{4}$$

Sampling model and parameter set

To compute $p(\upsilon \leq \upsilon_{max})$ for different combinations of spatial field and sampling plan, we first define a sampling model with parameter set θ_Z that encompasses stochastic simulation of heterogeneous spatial fields and composite sampling plans.

A spatial field's heterogeneity can be characterized by its spatial covariance, or strength of correlation between values at different locations depending on the physical distances separating them, often analytically represented by a covariance or semivariogram function [48]. These functions involve distribution parameters, here μ and CV expressed as natural logarithms, and a correlation length, λ, which describes how distant two locations can be and still have correlated values, or the "size" of the heterogeneities (Fig. 7A). Different analytical forms (*e.g.*, exponential, circular, Gaussian) are distinguished by the "smoothness" of the heterogeneities (Fig. 7B).

The parameter set θ_Z encompasses these spatial field parameters, as well as parametrization of a composite sampling plan, including the number of composite samples (n) and sub-samples (n_{sub}), radius of each composite sample (r_{app}), and margin of error intrinsic to the positioning device (e_{pos}) (Table 2).

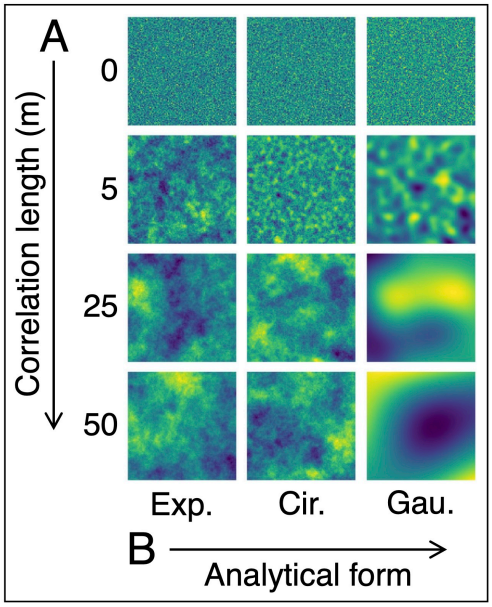


Figure 7: Examples of simulated 1-ha spatial fields with (A) increasing correlation lengths and (B) different analytical covariance models.

Table 2: Exploratory $\hat{F}(\theta_Z)$, uniform distributions used in stochastic spatial sampling simulations and SA to determine combinations of spatial field and sampling plan that result in measurement variances less than a target maximum.

	Parameter	**Symbol**	**Units**	**Transform**	**Range or Set**
Spatial field	True mean baseline concentration	μ_Z	mg/kg	ln	$\{1, 2, ..., 11\}$
	True baseline and mixture coefficient of variation	CV_Z	-	ln	$\{-8, -4, ..., -1\}$
	Distribution type	Dist. type	-	-	{normal, lognormal}
	Covariance model	Cov. model	-	-	{exponential, circular}
	Correlation length	λ_Z	m	-	{0, 10, 25, 50, 100}
Sampling plan	Radius of each composite sample	r_c	m	-	$[0.5, 5]$
	Number of randomly located composite samples	n	-	-	$\{1, 2, ..., 30\}$
	Number of sub-samples, collected at equal intervals along circumference of composite area	n_{sub}	-	-	$\{3, 4, ..., 15\}$
	Margin of error intrinsic to positioning device	e_{pos}	m	-	$[1, 10]$

4. Design the sampling plan

With the measurement variance requirements and sampling model established, the next step is to stochastically analyze the model to determine a minimally sufficient sampling plan. This involves quantifying the relative influence of spatial heterogeneity and sampling parameters on fulfilling the variance requirements (step 4.1) then applying deployment-specific constraints to identify sufficient sampling plans and refining them to a specific plan (step 4.2).

4.1. Quantify influence of spatial heterogeneity and sampling parameters on fulfilling measurement variance requirements

To partition the parameter space $\hat{F}(\theta_Z)$ into $\hat{F}(\theta_Z|\phi_Z)$ and $\hat{F}(\theta_Z|\overline{\phi}_Z)$ for sensitivity analysis, we use nested Monte Carlo simulations to compute $p(\upsilon \leq \upsilon_{max})$ for 10^4 realizations of θ_Z. After first generating the spatial field, we choose random locations for the n composite samples. For a single configuration of random locations, we simulate 100 rounds of composite sampling, computing the mean Z each time, and υ as the relative variance of the 100 means. Considering a PD would only sample a handful of times throughout the course of a deployment, these 100 rounds represent the theoretical variability introduced by random positioning error and inconsistent orientation of sub-samples over a heterogeneous field. In the context of solid-phase EW verification, this formulation assumes the sampling plan is fixed with baseline sampling, and the f_d calculation uses the mean of all n samples rather than each sample individually. Altogether, for a single parameter realization of θ_Z, we simulate 100 different configurations of random locations, and $p(\upsilon \leq \upsilon_{max})$ is the portion of configurations where the inferred measurement variance is less than υ_{max}.

Results of these nested simulations indicate that ϕ_Z is determined by the relative field-scale variance of Z, or coefficient of variation (CV_Z) (Fig. 8). Most realizations show an extremely low or extremely high likelihood of achieving a sufficiently small υ (Fig. 8A), and there is a clear $\ln(\mathrm{CV}_Z)$ threshold between -4 and -5 that dictates this behavior (Fig. 8B).

The exact $\ln(\mathrm{CV}_Z)$ threshold is dependent on the maximum number of samples considered in the SA, though additional results show that increasing the maximum n from 30 to 100 still results in a threshold below -4. Overall, this highlights that spatial heterogeneity not only needs to be accurately constrained before designing a sampling plan, but may also determine whether any monitoring strategy can succeed.

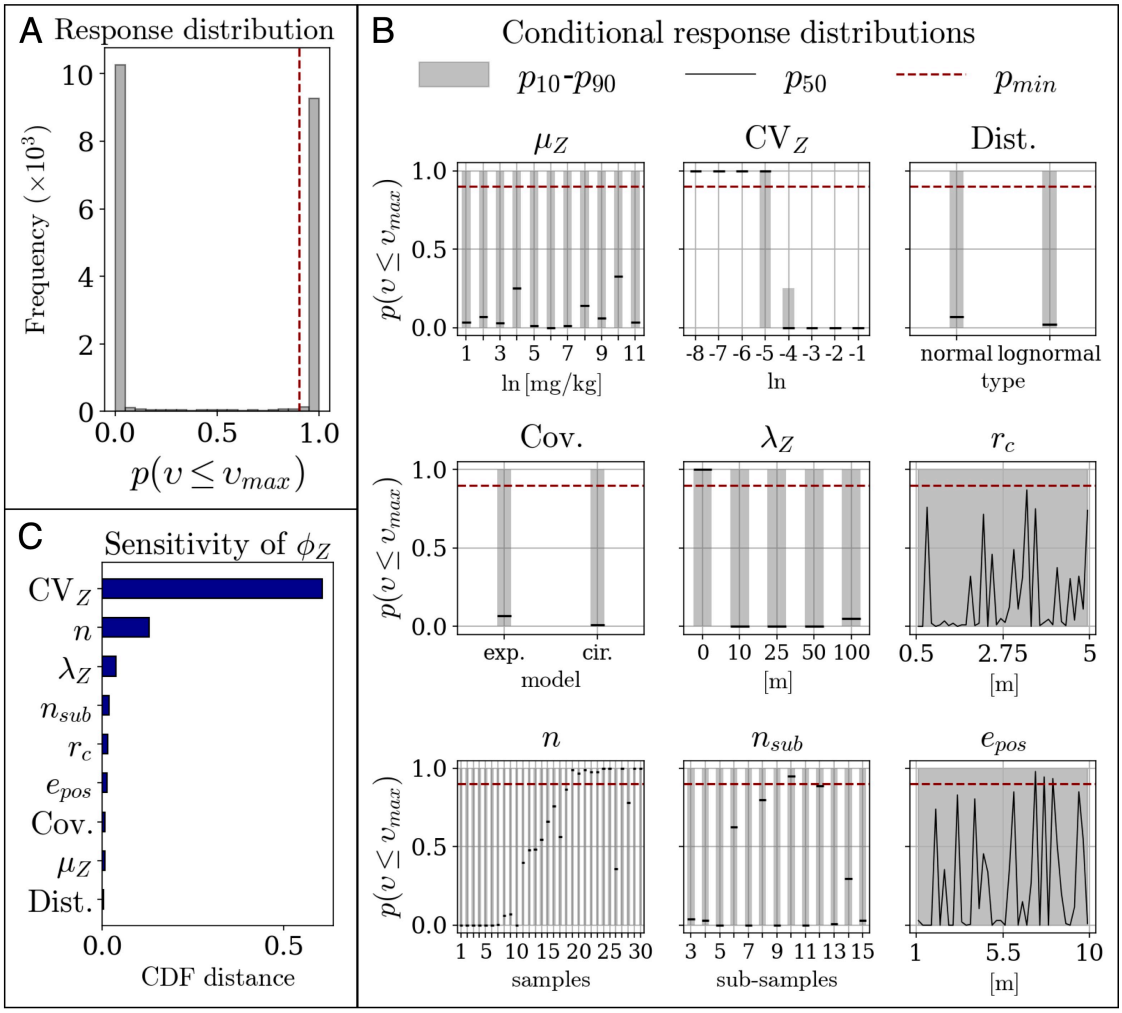


Figure 8: Exploratory SA of the influence of spatial heterogeneity and sampling plan on $p(\upsilon \leq e^{-6.5})$, the probability ($p$) that the inferred measurement variance (υ) is less than a maximum measurement variance (υ_{max}) of $e^{-6.5}$, using the parameter ranges in Table 2. (A) shows the response across all 10^4 realizations, with a red-dashed line separating the realizations that do and do not fulfill requirements of at least 90% (p_{min}) likelihood of measurement variance less than $e^{-6.5}$. (B) provides the conditional response distribution for each parameter. (C) provides the ranking of the parameters according to their influence on ϕ_Z, using the distance-based metric illustrated in Fig. 4.

4.2 Apply deployment-specific constraints and refine to a specific sampling plan

To narrow down to a specific sampling plan, we first need to constrain relative field-scale variance, the major control on ϕ_Z. Though point sampling is typically necessary to capture the true CV, such data are sparse for soil elemental composition at ha-scales. A few field studies [49–52] that did involve high-density point sampling of soil elemental concentrations (10^2-10^4 samples/ha) reported ha-scale variances of $-3 \leq \ln(\text{CV}) \leq -1$ for base cations (Ca, Mg, Na, K) and select trace elements (Ti, Ni, Al), which would not adhere to the threshold of -4 or -5 suggested by the analysis here (Fig. 9). For a PD interested in constraining site-specific variance, further stochastic point-sampling simulations indicate that, given observed ranges [49–52], only up to about 20 point samples are needed to estimate CV to the nearest ln with 90% confidence (Fig. 10). In theory, these suggested sample sizes are directly applicable to larger scales, assuming correlation length is scaled with grid resolution, and the random spatial fields tested here encompass patterns observed at larger scales. Overall, this suggests it would be feasible to collect the preliminary measurements needed to infer operational scalability for a robust array of potential empirical constraints.

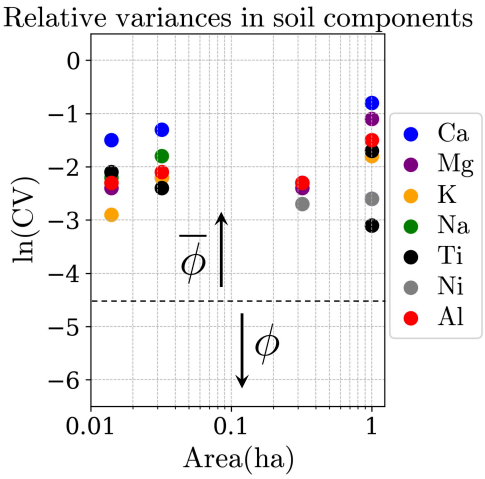


Figure 9: Relative variances of select soil elemental abundances reported across four studies [49–52] involving high-density sampling (10^2-10^4 samples/ha) of agricultural, grassland, and scrubland soils (upper 10 to 20 cm), here shown relative to the approximate field-scale variance threshold partitioning fields between those that likely can (ϕ) and cannot ($\overline{\phi}$) fulfill the uncertainty requirements considered in this study.

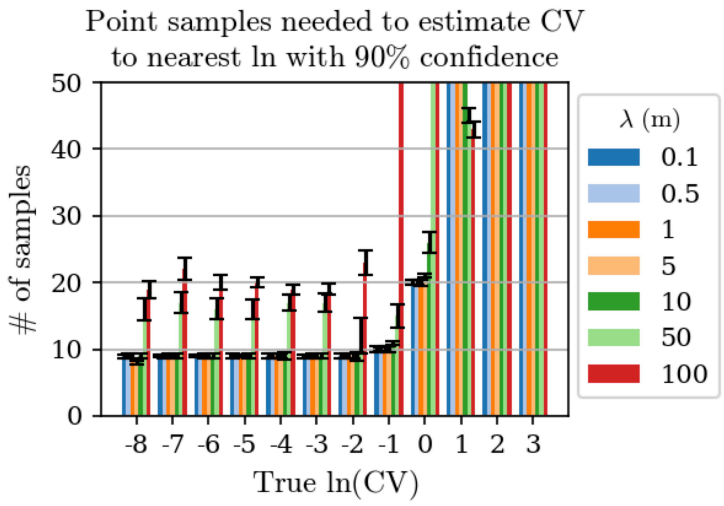


Figure 10: Simulation-based estimates of sample sizes needed to capture the coefficient of variation (CV) of a 1-ha (100 m x 100 m, 0.1-m resolution) lognormal spatial field to the nearest integer natural log (ln) with 90% confidence, considering different scales of spatial correlation (λ); errors bars represent the standard error across 10 spatial fields with different means.

While lower than reported for soils to-date [49–52], we constrain $\ln(\mathrm{CV}_Z)$ to -5 for our theoretical deployment and redo the SA (Fig. 11) to demonstrate next steps in monitoring design. Given this constraint, ϕ_Z becomes most sensitive to n, indicating at least 11 composite samples will result in $> 50\%$ likelihood, 18 samples $> 90\%$ likelihood, that υ will be sufficiently small (Fig. 11B). It may be important to further constrain λ (Fig. 11C, Fig. 10), and alternative technologies, such as remote sensing [53], may be necessary to control costs.

Altogether, the analysis here indicates that spatial heterogeneity in soil concentrations should be the foremost consideration when designing sampling plans for solid-phase verification of EW. Determining sufficient sampling plans requires preliminary constraints on relevant field-scale variances, and even minimally sufficient plans may be operationally infeasible, pointing toward reconsideration of the overall uncertainty requirements or measurement model.

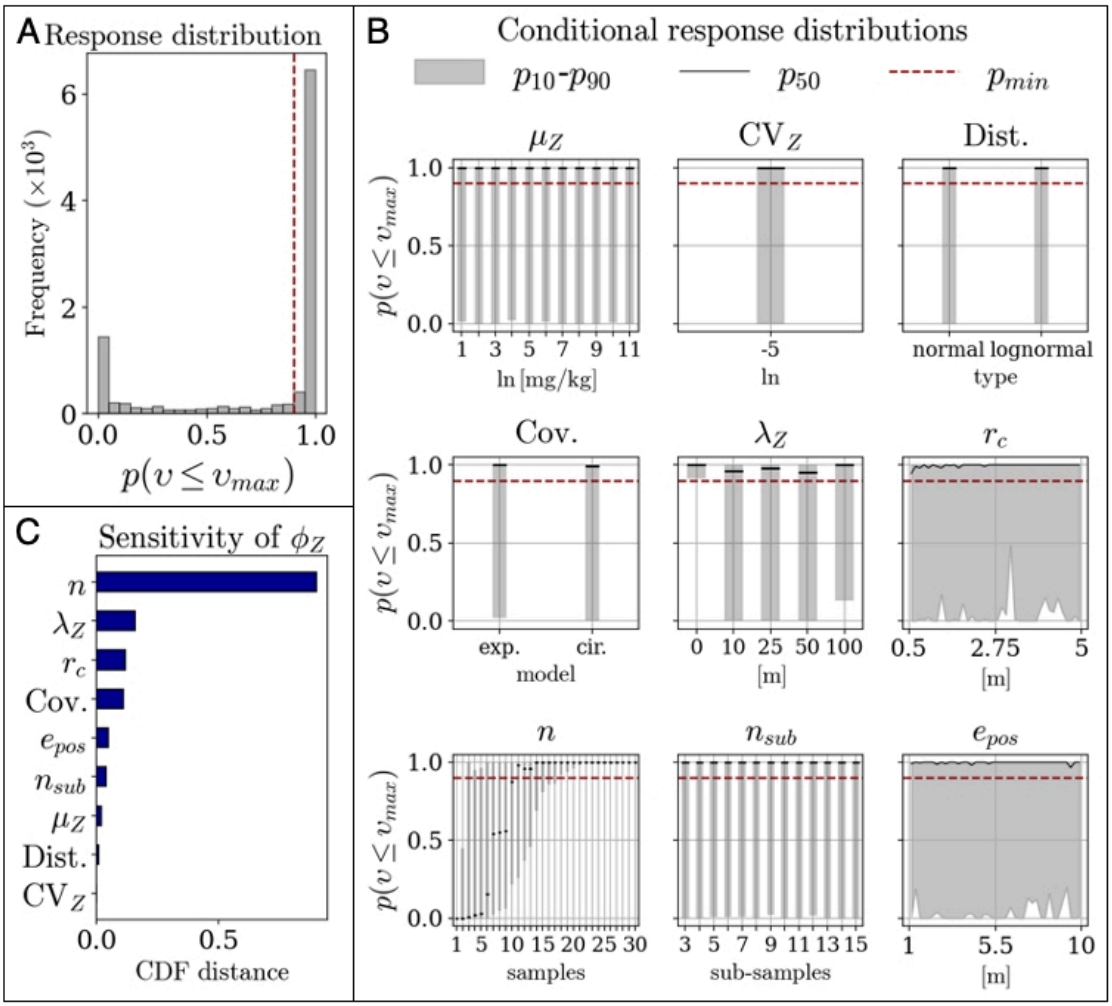


Figure 11: Deployment-specific SA where, relative to the exploratory SA in Fig. 8 and parameter ranges in Table 2, we constrain relative field-scale variance (CV_Z) to e^{-5}. (A) shows the updated response across all 10^4 parameter realizations, (B) the updated conditional response distributions, and (C) the updated ranking of parameters according to their influence on ϕ_Z.

5. Report the final estimate and uncertainty

In principle, reporting $\hat{f}_d$ and its uncertainty is relatively straightforward following sample collection and analysis, as we have predetermined the margin of error and confidence in the calculation. However, transparency, reproducibility, and traceability are critical for both scientific and compensatory applications. In particular, procedural compliance should be separated from the inherent scientific and environmental uncertainty in evaluating whether a project achieved a result within quantifiable confidence bounds. Thus, in addition to providing the underlying measurement data and appropriate metadata, ideally including an ISGN framework [54], the reporting framework should systematically capture key uncertainty targets and measurement distributions as constrained with each sampling event. Supp. 3 provides an example reporting format, where all sources of uncertainty are documented, including deviations from initial estimated field variance, and the final estimate is presented with clearly defined error and confidence.

Discussion

The methodology we develop here emphasizes the importance of constructing a measurement model in the context of uncertainty requirements, which contrasts with the way field studies are commonly designed, where the statistics are largely handled *ex post*. The probabilistic framework presented here is sequential in nature and starts with assessment of how input variability propagates through a measurement model, allowing us to then identify the critical parameters influencing the overall uncertainty (steps 1-2). For the case of EW, the measurement variances of base cation and tracer concentrations are critical determinants of the uncertainty in $\hat{f}_d$, indicating variance thresholds above which the uncertainty requirements will not be met. Stochastic spatial simulations (steps 3-4) are then used to identify composite sampling plans that adhere to these measurement variance thresholds, and iterative sensitivity analysis is used to refine these plans to a minimally sufficient plan.

Importance of spatial heterogeneity

This analysis further highlights that without the underlying assessment of spatial variance, even well-composited samples could lead to underestimation of the uncertainty, or worst-case, fail to achieve the minimum accuracy and confidence. Indeed, relative variances of soil elemental abundances from four studies [49–52] involving high-density sampling (10^2-10^4 samples/ha) of agricultural and grassland soils all exceed the maximum variance threshold identified in our analysis, suggesting that soil heterogeneity is too large for this mixing-model approach to produce reliable dissolution estimates (Fig. 9). While other tracers (*e.g.*, isotope ratios) or bulk cation stocks may exhibit less heterogeneity, the analysis here poses an important consideration for EW, namely the extent to which soil property distributions, and our ability to capture them with measurements, fall within the necessary ranges identified through this framework.

Here, we assume that spatial variance is the main contributor to measurement variance, though analytical uncertainty would be an additional factor for low-abundance chemical tracers. This could be directly incorporated into our framework by subtracting analytical variance from the maximum measurement variances identified in step 2, resulting in a lower target measurement variance in the sampling simulations. Another important gap to address is identification of intensively measured plots—simulated here at 1 ha, but can be larger—that are representative of up to tens-of-thousands of hectares. Collectively, simulating spatial variability and realistic sampling strategies can not only reduce logistical inefficiencies for EW, but also minimize the risk of unmet uncertainty requirements for CDR quantification.

Implications

For open-system CDR verification, SDOs face the challenge of balancing prescriptive requirements with flexibility. The approach presented here allows an SDO to define clear uncertainty requirements and, given preliminary estimates of spatial heterogeneity from the PD, recommend a deployment-specific sampling plan likely to fulfill the requirements. Be-

cause empirical measurement is typically a significant portion of open-system CDR costs, such a framework can reduce risk for PDs—if they provide reasonably accurate estimates of heterogeneity and adhere to the recommended plan, the SDO can confirm whether the requirements are fulfilled. The SDO would not necessarily need to be prescriptive with respect to the measurement types themselves, but rather, as this framework describes, how to mathematically relate the measurements and stochastically analyze their combined uncertainty. Transparent multi-stage reporting enables differentiation of procedural accountability and compliance from scientific and environmental uncertainties that affect removal performance. Through modular computational implementations, this framework could encompass a wide range of CZ processes and measurements while being tailored to soil systems and projects.

Acknowledgement

The authors gratefully acknowledge support from the Stanford Sustainability Accelerator (GHG-0012). BR acknowledges support from the Department of Energy Computational Science Graduate Fellowship (DE-SC0021110) and Stanford Data Science.

Supporting Information Available

- "Supp1_Derivations.pdf"
- "Supp2_Calculations.xlsx"
- "Supp3_ExampleReporting.xlsx"
- Simulation Data and Tools: https://github.com/rogersdb/cz-spatial-uq

References

[1] Sue Brantley et al. “Frontiers in Exploration if the Critical Zone: Report of a workshop sponsored by the National Science Foundation (NSF)”. In: (2006), p. 30.

[2] S. P. Anderson, F. Von Blanckenburg, and A. F. White. “Physical and Chemical Controls on the Critical Zone”. In: *Elements* 3.5 (Oct. 1, 2007), pp. 315–319. ISSN: 1811-5209, 1811-5217. DOI: 10.2113/gselements.3.5.315. URL: https://pubs.geoscienceworld.org/elements/article/3/5/315-319/137743 (visited on 11/04/2024).

[3] Suzanne Prestrud Anderson, Robert S. Anderson, and Gregory E. Tucker. “Landscape scale linkages in critical zone evolution”. In: *Comptes Rendus. Géoscience* 344.11 (Nov. 1, 2012), pp. 586–596. ISSN: 1778-7025. DOI: 10.1016/j.crte.2012.10.008. URL: https://comptes-rendus.academie-sciences.fr/geoscience/articles/10.1016/j.crte.2012.10.008/ (visited on 11/04/2024).

[4] Li Li et al. “Expanding the role of reactive transport models in critical zone processes”. In: *Earth-Science Reviews* 165 (Feb. 2017), pp. 280–301. ISSN: 00128252. DOI: 10.1016/j.earscirev.2016.09.001. URL: https://linkinghub.elsevier.com/retrieve/pii/S001282521630280X (visited on 11/04/2024).

[5] S. L. Brantley, M. B. Goldhaber, and K. V. Ragnarsdottir. “Crossing Disciplines and Scales to Understand the Critical Zone”. In: *Elements* 3.5 (Oct. 1, 2007), pp. 307–314. ISSN: 1811-5209, 1811-5217. DOI: 10.2113/gselements.3.5.307. URL: https://pubs.geoscienceworld.org/elements/article/3/5/307-314/137740 (visited on 11/04/2024).

[6] Craig Rasmussen et al. “An open system framework for integrating critical zone structure and function”. In: *Biogeochemistry* 102.1 (Jan. 2011), pp. 15–29. ISSN: 0168-2563, 1573-515X. DOI: 10.1007/s10533-010-9476-8. URL: http://link.springer.com/10.1007/s10533-010-9476-8 (visited on 11/04/2024).

[7] Steven A. Banwart et al. "Soil Functions: Connecting Earth's Critical Zone". In: *Annual Review of Earth and Planetary Sciences* 47.1 (May 30, 2019), pp. 333–359. ISSN: 0084-6597, 1545-4495. DOI: 10.1146/annurev-earth-063016-020544. URL: https://www.annualreviews.org/doi/10.1146/annurev-earth-063016-020544 (visited on 11/04/2024).

[8] Matthias Sprenger et al. "The Demographics of Water: A Review of Water Ages in the Critical Zone". In: *Reviews of Geophysics* 57.3 (Sept. 2019), pp. 800–834. ISSN: 8755-1209, 1944-9208. DOI: 10.1029/2018RG000633. URL: https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2018RG000633 (visited on 11/04/2024).

[9] Peter K. Kitanidis. "Parameter Uncertainty in Estimation of Spatial Functions: Bayesian Analysis". In: *Water Resources Research* 22.4 (Apr. 1986), pp. 499–507. ISSN: 0043-1397, 1944-7973. DOI: 10.1029/WR022i004p00499. URL: https://agupubs.onlinelibrary.wiley.com/doi/10.1029/WR022i004p00499 (visited on 06/04/2024).

[10] Dennis McLaughlin and Lloyd R. Townley. "A Reassessment of the Groundwater Inverse Problem". In: *Water Resources Research* 32.5 (May 1996), pp. 1131–1161. ISSN: 0043-1397, 1944-7973. DOI: 10.1029/96WR00160. URL: https://agupubs.onlinelibrary.wiley.com/doi/10.1029/96WR00160 (visited on 11/13/2024).

[11] Alecos Demetriades et al. "Continental, regional and local scale geochemical mapping". In: *Journal of Geochemical Exploration* 154 (July 2015), pp. 1–5. ISSN: 03756742. DOI: 10.1016/j.gexplo.2015.02.011. URL: https://linkinghub.elsevier.com/retrieve/pii/S0375674215000412 (visited on 05/17/2024).

[12] Clemens Reimann and Robert G. Garrett. "Geochemical background—concept and reality". In: *Science of The Total Environment* 350.1 (Nov. 2005), pp. 12–27. ISSN: 00489697. DOI: 10.1016/j.scitotenv.2005.01.047. URL: https://linkinghub.elsevier.com/retrieve/pii/S0048969705001889 (visited on 05/04/2024).

[13] Patrick G. Lawrence et al. “Guiding soil sampling strategies using classical and spatial statistics: A review”. In: *Agronomy Journal* 112.1 (Jan. 2020), pp. 493–510. ISSN: 0002-1962, 1435-0645. DOI: 10.1002/agj2.20048. URL: https://acsess.onlinelibrary.wiley.com/doi/10.1002/agj2.20048 (visited on 08/04/2024).

[14] John A. Rice. “Mathematical statistics and data analysis”. In: *Mathematical statistics and data analysis*. 3rd ed. Duxbury advanced series. Belmont, CA: Thomson/Brooks/Cole, 2007, pp. 181–191. ISBN: 978-0-534-39942-9.

[15] Maya Almaraz et al. “Methods for determining the CO2 removal capacity of enhanced weathering in agronomic settings”. In: *Frontiers in Climate* 4 (Dec. 16, 2022), p. 970429. ISSN: 2624-9553. DOI: 10.3389/fclim.2022.970429. URL: https://www.frontiersin.org/articles/10.3389/fclim.2022.970429/full (visited on 11/22/2023).

[16] Matthew O. Clarkson et al. “A review of measurement for quantification of carbon dioxide removal by enhanced weathering in soil”. In: *Frontiers in Climate* 6 (June 19, 2024), p. 1345224. ISSN: 2624-9553. DOI: 10.3389/fclim.2024.1345224. URL: https://www.frontiersin.org/articles/10.3389/fclim.2024.1345224/full (visited on 09/22/2024).

[17] Jens Hartmann et al. “Enhanced chemical weathering as a geoengineering strategy to reduce atmospheric carbon dioxide, supply nutrients, and mitigate ocean acidification: ENHANCED WEATHERING”. In: *Reviews of Geophysics* 51.2 (Apr. 2013), pp. 113–149. ISSN: 87551209. DOI: 10.1002/rog.20004. URL: http://doi.wiley.com/10.1002/rog.20004 (visited on 02/11/2023).

[18] Johannes Lehmann and Angela Possinger. “Atmospheric CO2 removed by rock weathering”. In: *Nature* 583 (July 9, 2020), pp. 204–205.

[19] George H Brimhall and William E Dietrich. "Constitutive mass balance relations between chemical composition, volume, density, porosity, and strain in metasomatic hydrochemkai systems: Results on weathering and pedogenesis". In: (1986).

[20] Tom Reershemius et al. "Initial Validation of a Soil-Based Mass-Balance Approach for Empirical Monitoring of Enhanced Rock Weathering Rates". In: *Environmental Science & Technology* (Nov. 14, 2023), acs.est.3c03609. ISSN: 0013-936X, 1520-5851. DOI: 10.1021/acs.est.3c03609. URL: https://pubs.acs.org/doi/10.1021/acs.est.3c03609 (visited on 11/30/2023).

[21] Tim Jesper Suhrhoff et al. "A tool for assessing the sensitivity of soil-based approaches for quantifying enhanced weathering: a US case study". In: *Frontiers in Climate* 6 (Apr. 12, 2024), p. 1346117. ISSN: 2624-9553. DOI: 10.3389/fclim.2024.1346117. URL: https://www.frontiersin.org/articles/10.3389/fclim.2024.1346117/full (visited on 04/16/2024).

[22] Mark Baum et al. *Mass-Balance MRV for Carbon Dioxide Removal by Enhanced Rock Weathering: Methods, Simulation, and Inference*. July 2, 2024. arXiv: 2407.01949[stat]. URL: http://arxiv.org/abs/2407.01949 (visited on 09/22/2024).

[23] Margaret S. Torn et al. "Mineral control of soil organic carbon storage and turnover". In: *Nature* 389.6647 (Sept. 1997), pp. 170–173. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/38260. URL: https://www.nature.com/articles/38260 (visited on 11/13/2024).

[24] Susan Trumbore. "Age of Soil Organic Matter and Soil Respiration: Radiocarbon Constraints on Belowground C Dynamics". In: ().

[25] Rosemary C. Capo, Brian W. Stewart, and Oliver A. Chadwick. "Strontium isotopes as tracers of ecosystem processes: theory and methods". In: *Geoderma* 82.1 (Feb. 1998), pp. 197–225. ISSN: 00167061. DOI: 10.1016/S0016-7061(97)00102-X. URL: https://linkinghub.elsevier.com/retrieve/pii/S001670619700102X (visited on 11/15/2024).

[26] Jennifer L. Druhan et al. "Timing the Onset of Sulfate Reduction over Multiple Subsurface Acetate Amendments by Measurement and Modeling of Sulfur Isotope Fractionation". In: *Environmental Science & Technology* 46.16 (Aug. 21, 2012), pp. 8895–8902. ISSN: 0013-936X, 1520-5851. DOI: 10.1021/es302016p. URL: https://pubs.acs.org/doi/10.1021/es302016p (visited on 11/18/2024).

[27] Ilsa B. Kantola et al. "Improved net carbon budgets in the US Midwest through direct measured impacts of enhanced weathering". In: *Global Change Biology* 29.24 (Dec. 2023), pp. 7012–7028. ISSN: 1354-1013, 1365-2486. DOI: 10.1111/gcb.16903. URL: https://onlinelibrary.wiley.com/doi/10.1111/gcb.16903 (visited on 05/04/2024).

[28] David J. Beerling et al. "Enhanced weathering in the US Corn Belt delivers carbon removal with agronomic benefits". In: *Proceedings of the National Academy of Sciences* 121.9 (Feb. 27, 2024), e2319436121. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.2319436121. URL: https://pnas.org/doi/10.1073/pnas.2319436121 (visited on 03/25/2024).

[29] Fuxing Guo et al. "Crop productivity and soil inorganic carbon change mediated by enhanced rock weathering in farmland: A comparative field analysis of multi-agroclimatic regions in central China". In: *Agricultural Systems* 210 (Aug. 2023), p. 103691. ISSN: 0308521X. DOI: 10.1016/j.agsy.2023.103691. URL: https://linkinghub.elsevier.com/retrieve/pii/S0308521X23000963 (visited on 11/22/2023).

[30] Christiana Dietzen and Minik T. Rosing. "Quantification of CO2 uptake by enhanced weathering of silicate minerals applied to acidic soils". In: *International Journal of Greenhouse Gas Control* 125 (May 2023), p. 103872. ISSN: 17505836. DOI: 10.1016/j.ijggc.2023.103872. URL: https://linkinghub.elsevier.com/retrieve/pii/S1750583623000427 (visited on 10/20/2023).

[31] Christina S. Larkin et al. “Quantification of CO2 removal in a large-scale enhanced weathering field trial on an oil palm plantation in Sabah, Malaysia”. In: *Frontiers in Climate* 4 (Aug. 30, 2022), p. 959229. ISSN: 2624-9553. DOI: 10.3389/fclim.2022.959229. URL: https://www.frontiersin.org/articles/10.3389/fclim.2022.959229/full (visited on 05/11/2023).

[32] Fangna Wang et al. “Wollastonite powder application increases rice yield and CO2 sequestration in a paddy field in Northeast China”. In: *Plant and Soil* (Feb. 28, 2024). ISSN: 0032-079X, 1573-5036. DOI: 10.1007/s11104-024-06570-5. URL: https://link.springer.com/10.1007/s11104-024-06570-5 (visited on 07/06/2024).

[33] Kirstine Skov et al. “Initial agronomic benefits of enhanced weathering using basalt: A study of spring oat in a temperate climate”. In: *PLOS ONE* 19.3 (Mar. 27, 2024). Ed. by Timothy Omara, e0295031. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0295031. URL: https://dx.plos.org/10.1371/journal.pone.0295031 (visited on 07/06/2024).

[34] Iris O Holzer, Mallika A Nocco, and Benjamin Z Houlton. “Direct evidence for atmospheric carbon dioxide removal via enhanced weathering in cropland soil”. In: *Environmental Research Communications* 5.10 (Oct. 1, 2023), p. 101004. ISSN: 2515-7620. DOI: 10.1088/2515-7620/acfd89. URL: https://iopscience.iop.org/article/10.1088/2515-7620/acfd89 (visited on 10/19/2023).

[35] Frank McDermott et al. “Enhanced weathering for CO2 removal using carbonate-rich crushed returned concrete; a pilot study from SE Ireland”. In: *Applied Geochemistry* 169 (Aug. 2024), p. 106056. ISSN: 08832927. DOI: 10.1016/j.apgeochem.2024.106056. URL: https://linkinghub.elsevier.com/retrieve/pii/S0883292724001616 (visited on 07/06/2024).

[36] Xavier Dupla et al. “Let the dust settle: Impact of enhanced rock weathering on soil biological, physical, and geochemical fertility”. In: *Science of The Total Environment*

954 (Dec. 2024), p. 176297. ISSN: 00489697. DOI: 10.1016/j.scitotenv.2024.176297. URL: https://linkinghub.elsevier.com/retrieve/pii/S0048969724064532 (visited on 12/18/2024).

[37] Noah W. Sokol et al. "Reduced accrual of mineral-associated organic matter after two years of enhanced rock weathering in cropland soils, though no net losses of soil organic carbon". In: *Biogeochemistry* 167.8 (July 23, 2024), pp. 989–1005. ISSN: 1573-515X. DOI: 10.1007/s10533-024-01160-0. URL: https://link.springer.com/10.1007/s10533-024-01160-0 (visited on 12/18/2024).

[38] The Earth Partners LLC. *VCS Module VMD0021: Estimation of Stocks in the Soil Carbon Pool*. Version 1.0. Nov. 16, 2012. URL: https://verra.org/methodologies/vm0021-soil-carbon-quantification-methodology-v1-0/.

[39] David Smith et al. *Geochemical and Mineralogical Maps, with Interpretation, for Soils of the Conterminous United States*. 2019. DOI: https://doi.org/10.3133/sir20175118.

[40] Amy L. Lewis et al. "Effects of mineralogy, chemistry and physical properties of basalts on carbon capture potential and plant-nutrient element release via enhanced weathering". In: *Applied Geochemistry* 132 (Sept. 2021), p. 105023. ISSN: 08832927. DOI: 10.1016/j.apgeochem.2021.105023. URL: https://linkinghub.elsevier.com/retrieve/pii/S0883292721001554 (visited on 02/06/2023).

[41] R Spear. "Eutrophication in peel inlet—II. Identification of critical uncertainties via generalized sensitivity analysis". In: *Water Research* 14.1 (1980), pp. 43–49. ISSN: 00431354. DOI: 10.1016/0043-1354(80)90040-8. URL: https://linkinghub.elsevier.com/retrieve/pii/0043135480900408 (visited on 11/27/2024).

[42] A. Saltelli, ed. *Sensitivity analysis in practice: a guide to assessing scientific models*. Hoboken, NJ: Wiley, 2004. 219 pp. ISBN: 978-0-470-87093-8.

[43] Céline Scheidt and Jef Caers. "Representing Spatial Uncertainty Using Distances and Kernels". In: *Mathematical Geosciences* 41.4 (May 2009), pp. 397–419. ISSN: 1874-8961,

1874-8953. DOI: 10.1007/s11004-008-9186-0. URL: http://link.springer.com/10.1007/s11004-008-9186-0 (visited on 11/17/2024).

[44] Darryl Fenwick, Céline Scheidt, and Jef Caers. "Quantifying Asymmetric Parameter Interactions in Sensitivity Analysis: Application to Reservoir Modeling". In: *Mathematical Geosciences* 46.4 (May 2014), pp. 493–511. ISSN: 1874-8961, 1874-8953. DOI: 10.1007/s11004-014-9530-5. URL: http://link.springer.com/10.1007/s11004-014-9530-5 (visited on 11/24/2024).

[45] George Christakos and Ricardo A. Olea. "Sampling design for spatially distributed hydrogeologic and environmental processes". In: *Advances in Water Resources* 15.4 (Jan. 1992), pp. 219–237. ISSN: 03091708. DOI: 10.1016/0309-1708(92)90008-P. URL: https://linkinghub.elsevier.com/retrieve/pii/030917089290008P (visited on 02/04/2025).

[46] Hwong-wen Ma and Chih-Chen Chang. "Assessment of the value of reducing uncertainty by sampling in a groundwater remediation system". In: *Science of The Total Environment* 402.1 (Aug. 25, 2008), pp. 9–17. ISSN: 00489697. DOI: 10.1016/j.scitotenv.2008.04.034. URL: https://linkinghub.elsevier.com/retrieve/pii/S004896970800452X (visited on 02/04/2025).

[47] Kevin Sutherland et al. *Enhanced Weathering in Agriculture*. Version 1.1.0. Jan. 22, 2025. URL: https://registry.isometric.com/protocol/enhanced-weathering-agriculture#determination-2.

[48] Richard Webster and Margaret A. Oliver. *Geostatistics for environmental scientists*. 2nd ed. Statistics in practice. Chichester: John Wiley & sons, 2007. ISBN: 978-0-470-02858-2.

[49] Akissa Bahri, Ronny Berndtsson, and Kenji Jinno. "Spatial Dependence of Geochemical Elements in a Semiarid Agricultural Field: I. Scale Properties". In: *Soil Science Society of America Journal* 57.5 (Sept. 1993), pp. 1316–1322. ISSN: 0361-5995, 1435-

0661. DOI: 10.2136/sssaj1993.03615995005700050026x. URL: https://acsess.onlinelibrary.wiley.com/doi/10.2136/sssaj1993.03615995005700050026x (visited on 05/03/2024).

[50] A. Gallardo. "Spatial Variability of Soil Properties in a Floodplain Forest in Northwest Spain". In: *Ecosystems* 6.6 (2003), pp. 564–576. URL: http://www.jstor.org/stable/3658998.

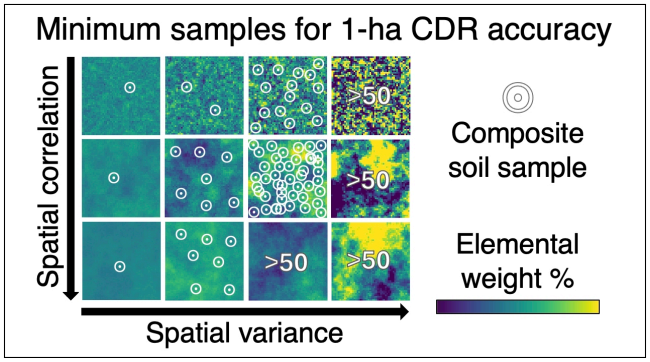
[51] Antonio Gallardo and Rocio Paramá. "Spatial variability of soil elements in two plant communities of NW Spain". In: *Geoderma* 139.1 (Apr. 2007), pp. 199–208. ISSN: 00167061. DOI: 10.1016/j.geoderma.2007.01.022. URL: https://linkinghub.elsevier.com/retrieve/pii/S0016706107000365 (visited on 11/30/2024).

[52] J. Spijker, S.P. Vriend, and P.F.M. Van Gaans. "Natural and anthropogenic patterns of covariance and spatial variability of minor and trace elements in agricultural topsoil". In: *Geoderma* 127.1 (July 2005), pp. 24–35. ISSN: 00167061. DOI: 10.1016/j.geoderma.2004.11.002. URL: https://linkinghub.elsevier.com/retrieve/pii/S0016706104002836 (visited on 11/30/2024).

[53] K. Dana Chadwick and Gregory P. Asner. "Landscape evolution and nutrient rejuvenation reflected in Amazon forest canopy chemistry". In: *Ecology Letters* 21.7 (July 2018). Ed. by Richard Bardgett, pp. 978–988. ISSN: 1461-023X, 1461-0248. DOI: 10.1111/ele.12963. URL: https://onlinelibrary.wiley.com/doi/10.1111/ele.12963 (visited on 02/20/2025).

[54] Joan E. Damerow et al. "Sample Identifiers and Metadata to Support Data Management and Reuse in Multidisciplinary Ecosystem Sciences". In: *Data Science Journal* 20.1 (Mar. 18, 2021), p. 11. ISSN: 1683-1470. DOI: 10.5334/dsj-2021-011. URL: https://datascience.codata.org/article/10.5334/dsj-2021-011/ (visited on 02/19/2025).

TOC Graphic


Minimum samples for 1-ha CDR accuracy
Spatial correlation
Spatial variance
>50
>50
>50
>50
Composite soil sample
Elemental weight %

Supporting Information 1:
Derivation of measurement model (Eq. 1-2)

1 Notation

- $[M]$ is solid-phase base cation elemental abundance
- $[T]$ is solid-phase immobile tracer elemental abundance
- $_{fs}$ denotes feedstock endmember
- $_{bsln}$ denotes baseline soil endmember
- $^{0}_{mix}$ denotes initial mixture between feedstock and baseline endmembers, *i.e.*, soil following feedstock amendment and tillage
- $_{mix}$ denotes weathered mixture, *i.e.*, following some feedstock dissolution

2 Defining f_d

The fraction of feedstock dissolved (f_d) is defined here as the complement of the fraction of feedstock remaining (f_r):

$$f_d = 1 - f_r$$

where f_r is the portion of feedstock cations still remaining in the solid-phase:

$$f_r = \frac{[M]_{mix} - [M]_{bsln}}{[M]^0_{mix} - [M]_{bsln}}$$

yielding Eq. 1 in the manuscript:

$$f_d = 1 - \frac{[M]_{mix} - [M]_{bsln}}{[M]^0_{mix} - [M]_{bsln}}$$

This formulation assumes negligible mass loss with feedstock dissolution, justified by relatively low feedstock mass fractions of 0.1-3% relative to the baseline endmember in the initial and weathered mixtures.

3 Calculating $[M]^0_{mix}$

Using an element-element mixing model with immobile tracer T, the increase in cation concentration can be calculated [1]:

$$[M]^0_{mix} - [M]_{bsln} = \frac{[M]_{fs} - [M]_{bsln}}{[T]_{fs} - [T]_{bsln}}([T]^0_{mix} - [T]_{bsln})$$

such that assuming $[T]^0_{mix} = [T]_{mix}$ and solving for $[M]^0_{mix}$ yields Eq. 2 in the manuscript:

$$[M]^0_{mix} = [M]_{bsln} + \frac{([M]_{fs} - [M]_{bsln})([T]_{mix} - [T]_{bsln})}{[T]_{fs} - [T]_{bsln}}$$

Reference: [1] Faure, G., & Mensing, T. M. (2005). Isotopes: Principles and Applications (3rd ed.). Chapter 16. Hoboken, NJ: John Wiley & Sons.

Example reporting format for measurement and verification of enhanced weathering
Method: Solid-phase mass balance with element-element mixing model

Values in this example are based on a theoretical deployment;
links to relevant data could go here.

Sampling parameters:

Plot area	1	ha
Compositing radius	5	m
Sub-samples per sample	5	cores
Positioning error margin	10	m

Sub-samples taken along circumference at approximately equal intervals.

Target uncertainty:

Maximum relative error	ϵ_{max}	10%
Minimum confidence	p_{min}	90%

Overall uncertainty in CDR calculation.

Target input variances:

		Measurement variance (ln(v))		
		Overall	*Analytical*	*Sampling*
Baseline cation conc.	$[M]_{bsln}$	-6.5	n/a	-6.5
Baseline tracer conc.	$[T]_{bsln}$	-6.5	n/a	-6.5
Mixture cation conc.	$[M]_{mix}$	-6.5	n/a	-6.5
Mixture tracer conc.	$[T]_{mix}$	-6.5	n/a	-6.5
Feedstock cation conc.	$[M]_{fs}$	n/a	n/a	n/a
Feedstock tracer conc.	$[T]_{fs}$	n/a	n/a	n/a

See Steps 1-2 for determining baseline and mixture target variances, specifically Step 2.3 for calculating the values entered here.
Feedstock assumed perfectly homogeneous.
Analytical variance assumed negligible, would be subtracted from *Overall* to determine *Sampling* variances.

For spatially explicit input:

		Plot-scale spatial variance (ln(CV))		
		Preliminary	*Baseline*	*Mixture*
Baseline cation conc.	$[M]_{bsln}$	-2	-1.8	-1.8
Baseline tracer conc.	$[T]_{bsln}$	-3	-3.3	-3.3
Mixture cation conc.	$[M]_{mix}$	-2	-1.8	-1.9
Mixture tracer conc.	$[T]_{mix}$	-3	-3.3	-3.7
Sample size:		*30*	*11*	*11*

See Steps 3-4 for determining minimum sample size, specifically Step 4.2 for calculating the *Baseline* and *Mixture* sample sizes entered here, as well as suggested sample sizes for *Preliminary* point sampling.
Plot-scale here is 1 hectare.
Dashed outline indicates where spatial variances for mixtures are assumed equal to baseline.

		Expected value (mg/kg)		
		Preliminary	*Baseline*	*Mixture*
Baseline cation conc.	$[M]_{bsln}$	3000	2645	2645
Baseline tracer conc.	$[T]_{bsln}$	30	32.12	32.12
Mixture cation conc.	$[M]_{mix}$	3500	3142	3271
Mixture tracer conc.	$[T]_{mix}$	35	36.11	34.21

Dashed outline indicates where mixture values are calculated using an ideal mixing model.

For other input:

		Expected value (mg/kg)
feedstock cation conc.	$[M]_{fs}$	50,000
feedstock tracer conc.	$[T]_{fs}$	20,000

Likely add another section "For temporally explicit input" for parameters with temporal variance.

Output estimates:

		Expected value (fraction)	
		Predicted	*Measured*
fraction of feedstock dissolved after 5 years	$\widehat{f_d}$	0.3	0.27
	Maximum error:		10%
	Confidence:		90%

Error and confidence will not necessarily match target uncertainty if heterogeneity estimates do not match initial estimates.