Appendix 3:

Foundations Working Group summaries

Cascade Climate

Table of Contents

Working Group overview

The *Foundations for Carbon Dioxide Removal Quantification in Enhanced Rock Weathering Deployments* document is the primary output of a multi-stakeholder community process undertaken from October 2023 through August 2024. The process spanned 14 technical Working Group discussions and 19 collaborative Working Group-industry targeted problem solving sessions. Iterative periods of drafting, reading, and constructive feedback were interlaced throughout the process.

The initial detailed assessment of the current state of science and practice of CDR quantification for ERW deployments was developed through facilitated discussions of **five technical Working Groups**. The focus areas for these Working Groups were designed to (as much as possible) comprehensively address the core components of CDR quantification, and target known challenges and open questions. Each Working Group met for three two-hour sessions from November 2023 through January 2024. Following each session, Cascade circulated a synthesis of the Working Group discussion to all process participants for feedback; this feedback was presented to Working Group members during the next session, and is shared in its original form below.

The five technical Working Groups were as follows:

- **Working Group 1**
    - **Scope:** Methods for measuring initial feedstock dissolution and alkalinity dynamics in the soil column, with a particular focus on sampling design and uncertainty characterization.
    - **Membership:** Thorben Amann, Antonio Carlos de Azevedo, Jelle Bilma, James Campbell, Rachael James, Mike Kelland, Frank McDermott, Isabel Montañez, Ian Power, Tom Reershemius, Radomir Schmidt, Jesper Suhrhoff, and Noah Planavsky (Working Group chair).
- **Working Group 2**
    - **Scope:** Details of the carbon cycle fluxes and biogeochemical processes in the soil column that need to be considered in a holistic accounting of netCDR, including interactions between ERW interventions and soil organic carbon cycling (as well as non-CO2 greenhouse gas fluxes).
    - **Membership:** Christiana Dietzen, Mathilde Hagens, Isabel Montañez, Noah Planavsky, Chris Reinhard, Rafael Santos, Radomir Schmidt, Eric Slessarev, Noah Sokol, Sara Vicca, Arthur Vienne, Bonnie Waring, and Kate Maher (Working Group chair).
- **Working Group 3**
    - **Scope:** Biogeochemical fluxes in downstream systems that should be considered as weathering products are transported from the application site

to the durable storage reservoir; constraining net changes to the carbon balance in the lower vadose zone and along groundwater flow paths, in surface water systems, and in the ocean as a result of an ERW deployment.
    - **Membership:** Yoshiki Kanzaki, Becca Neumann, Kirsty Harrington, Pete Raymond, Jim Saiers, Tao Wen, Chris Reinhard, and Shuang Zhang (Working Group chair).
- **Working Group 4**
    - **Scope:** Soil biogeochemistry models, with a particular emphasis on the current role of modeling in netCDR quantification and the infrastructure needed for evaluating, intercomparing, and validating models moving forward.
    - **Membership:** Salvatore Calabrese, Yoshiki Kanzaki, Maria Val Martin, Brian Rogers, Lyla Taylor, and Chris Reinhard (Working Group chair).
- **Working Group 5** (comprising two sub-groups)
    - First sub-group
        - **Scope:** Considerations for quantifying life cycle emissions as well as other considerations for assessing the human health and ecological impacts for ERW deployments.
        - **Membership:** Spyros Foteinis, Cara Maesano, Stephen McCord, Morimoto Shinichirou, and Yuan Yao.
    - Second sub-group
        - **Scope:** Human health and ecological impacts.
        - **Membership:** Maya Almaraz, Stephanie Grand, Fatima Haque, Anu Khan, and Charlotte Levy.

In constructing the technical Working Groups, all participants confirmed in writing that they do not own any equity stake in a for-profit company working in ERW; do not have any revenue-linked compensation agreement with any for-profit company working in ERW; do not work for a carbon removal buyer nor are involved in ERW purchasing decisions of any kind; and are not full-time employees of ERW suppliers, credit issuers, or MRV service providers.

Context and guidance for reading Working Group summaries and feedback requests

The following sections summarize the output of each successive cycle of Working Group meetings; each Working Group met three times between November 2023 and January 2024. The summaries reflect the evolving hypotheses, recommendations, and other ideas and considerations that emerged from Working Group discussions. This collection of output provided the core content for the development of the "Foundations" document. Note that certain aspects of the summaries in the sections below may represent moment-in-time, now-outdated perspectives that evolved or were adapted in the course of the writing and editing process, and should be considered as such. The "Foundations" document is the most current and comprehensive synthesis of inputs and perspectives gathered from the many participants in the process.

Cascade also solicited feedback from Market Focus Groups—including individuals working at project developers (sometimes referred to as “practitioners” or “suppliers”), buyers, and credit issuers—to tap into the wealth of applied scientific knowledge that resides among teams working on ERW in a commercial context. One goal in doing so was to help ensure that eventual "Foundations" recommendations incorporated an awareness of operational and financial considerations and constraints.

As a reminder, at the beginning of the community process in 2023, it was envisioned that the output of the process would be an industry-wide standard, with generalizable rules that would form the basis for ERW credit issuance. Over the course of the process, the Cascade team arrived at the conclusion that it is too early to articulate across-the-board rules that apply irrespective of deployment setting. Putting out a standard could create premature lock-in around an overly narrow or inflexible set of measurement approaches.

Instead, we believe that the optimal outcome is to help enable the most productive possible ‘pre-standardization’ phase. What is needed in the near term is more real-world deployments, across a variety of settings, that draw on the guidance and insights in the "Foundations" document while taking a “know your system and measure accordingly” approach. Critically, this must be combined with transparent reporting of site characterizations, quantification methods, and underlying data.

This shift resulted in the "Foundations" document taking the form of a technical reference for the field, rather than a prescriptive standard. More detail on this shift is described within the "Foundations" document and on the "Foundations" page of Cascade’s website. As such, language used regarding standardization in Working Group summaries may be out of date. Feedback has been left in the present tense to represent perspectives shared at the given moment in time.

In regards to how to review the below sections:

- Most of the below sections represent a syntheses of consensus Working Group opinions, structured by topic area discussed in the first, second, and third sessions of each of the five Working Groups.
- After each of the topic areas, we highlighted the feedback prompt for market participants. Feedback prompts fell under three categories:
    - **Blue** was where the Working Group had a **recommendation** on the sets of options that could go into the "Foundations" document. Feedback was framed to better understand whether the recommended options included approaches that were in line with current practices, and where any process participants might have concerns or reservations.
    - **Purple** suggested a **recommendation under development**, where the Working Group recognized a need for guidance in the eventual "Foundations" document, but hadn't yet settled on firm option sets. Feedback was framed to seek input from practitioners on their current approach to a given quantification component in order to help inform the technical Working Group's ability to come to a recommendation.
    - **Green** was where the Working Group had **ideas**, but no set opinion, on whether a given topic would/should go into the eventual "Foundations" document. Feedback was framed to seek input from the market ecosystem on whether scientific Working Groups should be prioritizing that topic in future sessions.

Working Group session 1: Summaries and initial requests for feedback

The first sessions for each of the five technical Working Groups were convened on November 6-8th, 2023.

Working Group 1 - Feedstock characterization / solid-phase measurements

Scope of this meeting

The focus of this Working Group will largely be to create rules, frameworks, and guardrails for how enhanced weathering suppliers conduct soil measurements in their deployments. We will cover how suppliers must address spatial and temporal heterogeneity, the depth at which measurements must be taken, the construction of baselines, soil carbon and cation fluxes that measurements address, common measurement pitfalls and how to avoid them, and uncertainty quantification. This meeting focused exclusively on solid-phase measurements, and we will turn to aqueous-phase and gas-flux measurements in session 2. We also spent the beginning of this meeting on how suppliers must characterize their feedstock.

1.1 Feedstock characterization

Status: **Purple / Recommendation under development**

What about the feedstock needs to be characterized?

- There was some consensus in the group that the mineralogy of the feedstock needs to be characterized, rather than just major oxides. Reasons discussed:
    - Major oxides alone don't capture whether they came from silicates vs. carbonates
    - Need to check for presence of asbestiform minerals
- It was also proposed that the BET surface area should be required to report
- Heavy metal concentrations must be reported as well.
- The group agreed that methods to predict mineralogy from major oxide concentrations do not work effectively and should not be used (e.g., CIPW Norm).
- Even in cases where quantification methods don't require feedstock mineralogy (e.g., aqueous measurements of alkalinity and DIC), this characterization should still be required as an important sanity check.

How frequently must feedstock piles be subsampled and characterized? How should uncertainty associated with feedstock heterogeneity be characterized?

- The working group agreed that many quarries have substantial heterogeneity in mineral composition (wt % amorphous phase, major oxide composition) over space and time
- Proposals for standards requirements:
  - Feedstock must be characterized once for every application event, unless feedstock heterogeneity is constrained and sufficiently low
  - Methods used to subsample a larger pile must be clearly outlined
  - For uncertainty characterization — need to build a representative sense of mineral heterogeneity within the feedstock pile through multiple samples

Feedback prompt: Are these reasonable requirements for the standard? For suppliers, how do these requirements compare to what you do today, both in terms of feedstock measurements and measurement frequency?

1.2 Soil column fluxes

Status: **Blue / Recommendation**

The group broadly aligned on the following list of fluxes within the soil column that quantification methods should attempt to constrain, emphasizing the importance of both cation and carbon fluxes (these will be addressed across Working Groups 1, 2, and 4 throughout the remaining meetings):

- Initial alkalinity release from feedstock dissolution
- Non-carbonic acid weathering
- Plant cation uptake
- Sorption time-lags / cation exchange
- Secondary mineral precipitation
- Degassing due to pH conditions
- Changes to soil organic matter / respiration

Procedural suggestion: When a supplier protocol proposes using a measurement approach (especially if it is new / unestablished), the protocol needs to clearly articulate which fluxes it claims to cover, and make geochemical arguments for how it does so. This way, knowledgeable auditors could at least be able to evaluate the geochemical claims being made.

- Comparisons drawn to the SOC space where private companies used their own inscrutable (largely model-driven) methods for quantification

Feedback prompt: Do you have any comments on this list of fluxes?

1.3 Analytical error in solid-phase measurements

Status: **Blue / Recommendation**

There was consensus that suppliers must report the analytical error of their measurements, and to do this robustly they must use a) geostandards, and b) duplicates. Analytical error must be propagated and included in the aggregate net CDR uncertainty characterization.

In order to establish that suppliers can pick up a resolvable signal from noise when comparing measurements through time to observe a dissolution rate, suppliers must transparently report the methods used to determine statistical significance.

Feedback prompt: Does this set of analytical error requirements seem reasonable? Anything you would add to ensure rigorous error quantification?

1.4 Spatial heterogeneity and sampling density

Status: **Purple / Recommendation under development**

Recapping relevant core principles:

- Soil systems are heterogeneous in space and time. Suppliers must characterize the baseline field heterogeneity prior to deployment, and sampling densities must take this heterogeneity into account.
- Suppliers must fully characterize the uncertainty of their net CDR quantification, including but not limited to propagated analytical error and uncertainty tied to feedstock and underlying soil variability.

Our goal is to figure out how the standard should operationalize these principles.

Meta-level point of consensus: At the current state of scientific understanding, suppliers should be required to quantify net CDR and associated uncertainty with respect to each field or deployment, rather than smearing out uncertainty over large aggregate regional scales.

What are the forms of baseline spatial heterogeneity that suppliers should be concerned with for solid-phase dissolution measurements?

- Soil cation concentrations, such as variability in baseline concentrations of immobile detrital elements if being used as tracers in a mass balance approach
- Other contextual variables that are heterogeneous, such that suppliers should have a handle on their variability in deployment fields:
  - Hydrological overview of the site, including soil texture to inform infiltration, and precipitation and evaporation rates
  - Cation exchange capacity, pH, and buffer pH

Recommendations for sampling density and design:

- It was suggested that sampling density should be driven by the in-field heterogeneity of variables being measured. For example, if measuring dissolution rates using mobile/immobile cation ratios, sampling density should be tied to baseline field variance in the immobile tracer concentration.
- Practically, suppliers should characterize field heterogeneity before deployments. Suppliers can use existing soil maps with sufficient spatial resolution (need more clarity on what sufficient spatial resolution entails), or take samples prior to deployment.
- Suppliers should be required to use a robust statistical framework to quantify uncertainty in field-scale estimates due to spatial variability, and propagate this error.

Feedback prompt: For suppliers: it would be helpful for the Working Group to get a sense of how you decide your sample density today, and what statistical frameworks you use to rigorously quantify uncertainty due to baseline heterogeneity. What should be required here?

1.5 Control Strips

Status: **Purple / Recommendation under development**

- It is important to have a representative control strip (unamended) for each deployment, and control strips must be representative of the field they're being compared to
- The group is still making a decision about:
    - Whether there should be a prescribed ratio of amended land to control land in the standard
    - What rules should be in place to ensure that control strips capture in-field heterogeneity

Feedback prompt: For suppliers—are you currently using control plots in your deployments? How should we ensure that control plots are representative of the rest of the field? Should they be required?

1.6 Solid-Phase Sampling Depth

Status: **Green / ideas**

- For solid-phase measurements of feedstock dissolution, required sampling depth should be tied to tillage depth
- Suppliers should be required to constrain cation fluxes throughout the entire soil column in addition to in the weathering zone (top 10-30cm)

  - Parking lot for future discussion—the role of 1m solid-phase soil cores. Even if operationally infeasible to do on every site, when might it make sense. See the discussion in WG4 as well!

Feedback prompt: For suppliers—how do you currently quantify fluxes in deeper soils below 30cm? Do you have thoughts on what the standard should require?

Working Group 2 - Grappling with pH, strong acids, and liming

Scope of the meeting

This was the first session digging into how cation, alkalinity, and dissolved inorganic carbon (DIC) dynamics within soil systems impact net CDR for enhanced rock weathering deployments. The session focused on considerations of pH in the soil profile, non-carbonic acid weathering, and accounting for the baseline of historical liming / agronomic pH control. The Working Group will continue with a discussion of secondary mineral precipitation and cation sorption in the next session, before tackling interactions with soil organic carbon cycling and the net greenhouse gas balance in session 3.

One of the emergent themes from this session's discussion was the importance of time and developing a common framework for dealing with processes that occur over different timescales. What constitutes rigorous accounting for many of these processes is dependent on the time horizon considered, and we need to be able to consistently account for temporal lags in the system. The working group will return to this next session with the conversation about cation sorption and exchange in soil systems.

2.1 Conceptual framing: conformance zones

Status: **Green / ideas**

The working group began with a discussion of the utility of considering two conformance zones: a "near-field" conformance zone that captures the flux of weathering products to groundwater, and a "far-field" conformance zone that captures the flux from groundwater to the ocean. Potential advantages of this framework include that:

- It separates the accounting: constrain the flux from the soil profile, then consider how that flux may lead to carbon loss or gain downstream. The two zones could have very different uncertainty profiles and time horizons.
- It enables potential separation of carbon from alkalinity
- It is similar to how nuclear waste regulations are developed.

The Working Group was broadly aligned with formalizing this separation, but emphasized the importance of being able to track and pass fluxes from one zone to another, particularly when considering the full system boundary for counterfactual calculations.

Feedback prompt: Any comments on the framing of two 'conformance zones' for net CDR accounting? For suppliers: Do you have a similar type of segmentation in your MRV approach?

2.2 Accounting for the influence of soil pH

Status: **Purple / Recommendation under development**

The carbonic acid weathering of silicate and carbonate minerals will consume protons and generate alkalinity. The extent to which that alkalinity is charge balanced by DIC is pH and pCO_2 dependent. This sets up a question of system boundaries / scope. Should suppliers only consider carbon that is captured as DIC by charge balancing alkalinity released from feedstock dissolution *within the soil profile* in the net CDR calculation, or can net CDR from downstream or 'far field' processes also be included?

It was proposed that the standard should utilize the charge balance definition of alkalinity, considering how an ERW deployment yields a change in the flux of excess cations (+ corresponding anion balance) from the soil profile, and the efficiency with which that charge balance equivalent is transported through the system. The Working Group largely aligned with this proposal, and on the question of what should be *measured* or accounted for within the soil profile, reached consensus that cation accounting alone is not sufficient because considering the cation flux alone does not account for potential time lags in net CDR.

Emerging recommendations from the working group include:

- Suppliers should track both the total alkalinity and DIC flux from the soil profile; quantification in this near-field zone is required at a minimum. The Working Group will return to the details of how to acceptably track those fluxes through measurements and models in consultation with Working Group 1.
- For a supplier to claim additional CDR in the 'far-field' zone downstream, a sufficiently rigorous catchment model of the site-specific counterfactual and ERW deployment scenario is required. The Working Group will return to requirements for this catchment-scaling modeling in consultation with Working Groups 3 and 4.

Feedback prompt: Do you have any feedback on treating soil pH impact on net CDR through this lens?

2.3 Explicitly accounting for non-carbonic acid weathering

Status: **Purple / Recommendation under development**

The core question of non-carbonic acid weathering is whether the impact of proton buffering by the weathering reaction can be considered in the quantification of net CDR. The Working Group considered two endmember options:

- Option A: Conservatively, strong acid weathering should be subtracted from net CDR, as released cations are not being charged balanced by bicarbonate, so no CO_2 removal has yet occurred.
- Option B: Even if released cations are initially charge balanced by e.g., nitrate or sulfate, this weathering has taken up protons that may have otherwise reacted with bicarbonate and caused CO_2 degassing. This has now been avoided, and thus in this view, CDR should still be credited.

The working group's emerging recommendation:

- The default path is for suppliers to conservatively remove strong acid weathering from net CDR quantification. This could be done by measuring nitrate and sulfate concentrations in the pore water at the base of the weathering zone, or by making a conservative estimate of strong acid production through nitrification based on fertilizer application rates (assuming 100% of ammonium converted). The fraction of feedstock dissolution driven by strong acid weathering is subtracted from net CDR.
- However, if suppliers are able to convincingly model a counterfactual scenario to quantify the impact of strong acid release in the absence of feedstock application, the additional carbon maintained by the system can be retained in the net CDR quantification. The uncertainty in the modeled counterfactual must also be explicitly quantified. It was also highlighted that the validity of this approach depends on how time and temporal lags are treated in the standard as a whole.
- The Working Group will return to the question of dealing with time lags in the next session, and will consider the question of what a 'convincingly modeled' counterfactual scenario of the full system (soil to long-term storage reservoir) would entail in consultation with Working Groups 3 and 4.

The Working Group also touched on the question of how to consider organic acids. The emerging consensus is that for the v1 standard, it is valid to assume that organic acids will be degraded following reaction with silicate minerals and generate an equivalent amount of DIC as carbonic acid weathering.

Feedback prompt: Do you have any feedback on these emerging recommendations for considering non-carbonic acid weathering?

2.4 Liming or agronomic pH control counterfactual

Status: **Purple / Recommendation under development**

For deployment sites where liming has previously been used to control soil pH, accounting for the impact of that historic lime use becomes an additionality question. Depending on the deployment site, liming can be an immediate source of CO_2 to the atmosphere (e.g., acidic soils), a transient sink of CO_2 but net neutral over longer timescales (e.g., in cases of significant carbonate precipitation in the soil profile or downstream), or represent net carbon removal over a timescale of centuries.

The Working Group has broad agreement that the standard will need to include requirements for accounting for the baseline of historic liming practices, but consensus was not reached on how to do so.

Two options were discussed for how a supplier should be required to account for the baseline liming practice, if the ERW deployment leads to a change in practice:

- **Option A**: For any change in liming practice, both the lifecycle emissions and net CDR associated with that baseline liming must be quantified. Only net CDR that is additional to this baseline may be credited.
  - Pros: If data is available, this is the most rigorous, and may allow for accounting of the temporal implications of substituting ERW with a silicate feedstock for liming.
  - Cons: Implications for early adopters and penalizing farmers who have been driving net CDR through their liming practices (analogs to cover cropping). Data on both lime application rates and counterfactual lime LCA details can be hard to access and in many cases must be approximated.
- **Option B**: The deployment must only demonstrate net additionality (net CDR from the ERW deployment > net CDR from the baseline liming practice), not explicitly subtract CDR driven by historic liming practices.

A subset of the Working Group members will take up the tradeoffs involved in these options, in collaboration with members of Working Group 5, before the next session.

The utility of requiring a control plot that maintains the historic agronomic pH control practice was also discussed; although multiple group members were in favor (at least for the early deployments covered by the v1 standard), full consensus was not reached.

Feedback prompt: Which of the two options for incorporating the counterfactual of historic liming practices into net CDR quantification would you advocate for and why? For suppliers, would you have the data required to calculate net CDR from representative baseline liming practices for your near-term deployments?

Working Group 3 - Catchment characterization and river evasion

Scope of the meeting

The alkalinity generated by enhanced rock weathering and associated CO_2 removed from the atmosphere as charge-balanced dissolved inorganic carbon (DIC) is not durably stored until it reaches a long-duration reservoir: in some cases a long residence time groundwater system, and in many cases the ocean. Multiple evasion pathways in surface-water systems can lead to CO_2 loss that should be accounted for in net CDR quantification. The goal of this working group session was to articulate the acceptable pathways for quantifying the magnitude of potential losses in open-water systems after the alkalinity/DIC flux exits the soil profile and enters the broader drainage catchment. Carbon dynamics in coastal and marine ecosystems will be considered in the next session. In Session 3 we'll return to considerations of catchment hydrology, including requirements of hydrologic characterization, and constraining loss pathways in the lower vadose zone and groundwater systems.

An important outcome of the initial Working Group 2 session was the need to articulate concrete guardrails for what constitutes a sufficiently rigorous catchment model to calculate counterfactual and ERW deployment scenarios for catchment-scale cation/proton balance arguments. This will be discussed by Working Group 3 in the next session.

3.1 Catchment characterization

Status: **Blue / Recommendation**

The working group started by discussing catchment characterization and how to determine whether a given deployment is draining into a sufficiently long residence time groundwater system to forgo consideration of carbon exchange in surface water systems.

For the purposes of the discussion, a mean residence time threshold of 1000 years to be considered 'sufficiently long' was suggested and generally accepted by the Working Group.

The group reached consensus that it is critical to demonstrate that fluids infiltrating through soils of a deployment predominantly drain into a flow path that actually has a long residence time if such an argument is going to be made. It is not sufficient to demonstrate that water from the deployment field is infiltrating the ground somewhere - many local groundwater systems have very short residence times. To the question of what would constitute sufficient evidence, the group highlighted that this is dependent on basin topography and geologic structure, and the best way to infer this in the absence of lots of data is groundwater flow models. There are national-scale models that have been used to estimate groundwater flow residence times for large watersheds, but local models would likely be needed to inform individual ERW deployments.

Feedback prompt: Is storage in long-residence time groundwater systems a consideration for any of your near-term deployments? If so, what data do you use to characterize drainage flow paths?

3.2 Fluxes in surface water systems

Status: **Green / ideas**

The working group then discussed and sought alignment on the key carbon fluxes to account for when considering potential losses from surface water systems. We started with a comprehensive overview of processes that the community would ideally be able to account for. That list is articulated here. We then homed in on the question of what processes are likely to be most important and the state of our understanding of these flux pathways + ability to parameterize them in quantitative models (next section).

The Working Group identified the following carbon system fluxes as potential components of downstream evasion quantification:

- Outgassing caused by DIC system equilibration
- Carbonate mineral precipitation and burial
- Changes to organic matter respiration and metabolic activity in the surface water system, including the influence of aquatic vegetation that directly takes up bicarbonate
- Authigenic clay formation / reverse weathering
- Bedrock interaction with solution
- Changes to nutrient fluxes and nitrogen cycling. Impactful from both the perspective of greenhouse gas fluxes and the charge balance implications of nitrate loss.

Feedback prompt: Are there any additional fluxes that you'd like to draw the Working Group's attention to?

3.3 Quantifying loss pathways in river systems

3.3.1 What fluxes do we currently need to account for?

Status: **Purple / Recommendation under development**

In discussing how/if each of the fluxes articulated above should be included in a process-based quantification of downstream evasion, the group was broadly aligned that the community needed to start with a sensitivity analysis to better determine what processes are likely to lead to large downstream losses. There was not consensus amongst the group that any processes have been demonstrated to have a high probability of driving large carbon losses. But the group highlighted that only outgassing due to DIC system equilibration and carbonate mineral precipitation have been considered in the literature in a rigorous way. Understanding how changes in DIC, alkalinity, and individual element fluxes associated with ERW deployments could modify many of these carbon system fluxes remains at the edge of science.

The Working Group's initial assessment of both the importance of and our ability to currently capture the potential loss pathways in quantitative models is as follows:

- Outgassing caused by DIC system equilibration
  - Requires river network data (surface area, discharge rate, flow paths), major ion chemistry, pH, pCO2, and water temperature
  - It is not clear that this would be a large net loss pathway under most circumstances. However, it is tractable to model in cases where regional-scale river network and stream chemistry data exist. The working group did not reach consensus that this could be ignored in process-based models and noted that it could be important regionally.
- Carbonate mineral precipitation and burial
  - Noted that it is the net loss due to carbonate precipitation that matters - considering carbonate precipitation without the potential for carbonate dissolution in sediments may provide an overly conservative estimate of loss
  - Relatively conservative constraints can be placed on this using process-based models
- Changes to organic matter respiration and metabolic activity in the surface water system, including the influence of aquatic vegetation that directly takes up bicarbonate.
  - Potentially quite important, but we don't yet have sufficient constraints - R&D stage.
- Authigenic clay formation / reverse weathering
  - Potentially important, but we don't yet have sufficient constraints - R&D stage.
- Changes to nutrient fluxes and nitrogen cycling.
  - Not further discussed by the working group, but likely out of scope for the v1 standard.

Feedback question: Do you have input on whether these (or other) carbon flux pathways have been demonstrated to be important avenues for carbon loss? Are there tractable approaches to quantify their impact?

3.3.2 Driving towards acceptable practices for articulation in the standard

Status: **purple / recommendation under development**

While the Working Group did not reach consensus on the minimum requirements the standard could articulate for accounting for downstream loss from river systems, the group was broadly aligned on the following counts:

- The potential for evasion in river systems cannot be ignored.
- Setting a single conservative loss estimate to be used across all deployments would entail making largely unconstrained assumptions.
- Hypothesis from the Working Group is that the standard could require the use of a process-based model that incorporates a designated suite of processes.
- There is potential to develop a set of empirical rules for estimating the loss magnitude in the absence of requisite river chemistry data, but more work and discussion is needed to reach consensus around satisfactory guardrails.

The Working Group also hypothesized that developing the models required to assess downstream evasion in riverine systems is likely outside the scope of individual suppliers considering individual deployments. These models and frameworks should be developed by the broader scientific community.

Feedback Prompt: It would be helpful for the Working Group to get a sense for how you are currently quantifying or planning to quantify the potential for evasion in riverine systems for your deployments. Are you currently using a spatially and temporally explicit process-based model? Assuming a set conservative 'haircut'? Or assuming that downstream losses before marine systems will be negligible?

Do you have comments for the Working Group on data availability for modeling riverine processes (hydrologic data, aqueous chemistry data, etc.) in your deployment regions of interest to inform the operational feasibility of requiring region-specific process-based models or empirical heuristics based on catchment characteristics?

Working Group 4 - Validating process-based reactive transport models

Scope of the meeting

The goal of this group is to lay the foundation for the creation of impartial systems for model validation. This first meeting was particularly focused on process-based reactive transport models (RTMs), but time will be spent in later meetings on black-box machine learning models and other approaches. This particular meeting focused mostly on a) modeling soil column cation transport and storage, as opposed to dissolution kinetics, and b) systems for validating models with test data, as opposed to e.g., model-model intercomparison as a form of validation. Future meetings will get further into the practical details of how a system for supplier model validation could work practically, and whose responsibility it is to collect test data.

4.1 Conceptual Framing

Status: **Green / ideas**

The working group built consensus around the following proposal: the representation of feedstock dissolution kinetics in a model could be **validated separately from** the representation of cation transport, storage, and exchange in the soil column.

- The standard will require suppliers to adequately constrain or conservatively estimate cation fluxes in lower soils.
- Thus, if suppliers have robust measurements to constrain dissolution rates, but no measurements for deeper soil cation dynamics taking place, sufficiently validated RTMs could potentially be used to quantify these deeper soil column fluxes. Dissolution rate measurements taken in the top 10-30 cm of the soil can be used as an empirical constraint for the model.

This working group meeting largely focused on how model representations of cation transport / storage in the soil column could be characterized.

Feedback prompt: Does this divide into model validation for dissolution kinetics vs validation for soil cation dynamics make sense? Any thoughts?

4.2 Required Processes in an RTM

Status: **Purple / Recommendation under development**

The working group had an emerging recommendation that sufficiently rigorous models should be able to represent:

- All three phases (solid, liquid, and gas)

- Particle size distribution tracking
- Strong acid dynamics (especially due to nitrogen fertilizer application) and pH effects
- Cation exchange and secondary site processes
- Secondary mineral precipitation—at least generic carbonate and clay phases
- Soil mixing (bioturbation and tillage)
- Fluid flow and infiltration
- Respiration and CO_2 production—at least first-order decomposition kinetics + diffusion + root respiration
- Plant cation uptake, and plant decomposition

The group also discussed the potential importance of being able to deal with:
- Organic acid dynamics
- Details of soil organic carbon cycling (including mineral-associated organic matter)

Feedback prompt: Does this seem like a reasonable list of required RTM processes? Any thoughts?

4.3 Data-model validation in the soil column

Status: **Green / ideas**

Emerging thinking about how will validation of models should work in practice:
- The standard will likely require a separate set of field data be used to test/validate RTMs than is used to spin-up and tune these models.
- Ideally, model validation will be performed by a non-financially conflicted third party, who can compare the model with a suite of test data.

The working group recognizes that unlike shallow soil dissolution measurements, deeper soil datasets like those described below will be very challenging for individual suppliers to regularly collect, especially beyond their designated "R&D fields". Thus, a community-wide effort is needed to build these datasets, both for model tuning and ultimately for validation.

In order to validate the performance of RTMs in their **representation of soil cation dynamics** (again, we will come back to dissolution kinetics modeling in a later session), the following types of data were brainstormed by the working group to be collected across sufficiently diverse and representative locations:
- Aqueous phase cation concentration at relevant depths (potentially at each soil horizon)
- Soil moisture or matric potential at multiple depths, to characterize water flow
- Measurements of pH, buffer pH, TA, DIC, and/or pCO2

- Cation exchange capacity and solid-phase exchangeable cation concentrations (at representative depths)
    - Q-XRD may be a compliment to solid-phase cation measurements
- Local constraints on precipitation and evapotranspiration
- Total inorganic carbon on a lower temporal frequency
- Root distribution

To assess model performance against this data, the group began to discuss what performance metrics could be applied. There were two broad categories of ideas:

- Statistical tests—e.g., models could be evaluated using a Bayesian approach with information criteria
- Key data benchmarks—models must reproduce the following soil measurements with enough precision across spatially and temporally heterogeneous conditions:
    - pH / buffer pH / base saturation
    - TA / DIC in lysimeters or (sufficiently well-constructed) mesocosms

Feedback prompt: Does this seem like a reasonable approach for validating model performance in deep soils against data? For suppliers, how have you thought about model validation?

4.4 Mesocosm Data

Status: **Purple / Recommendation under development**

The standard will also likely create rules around when, if ever, mesocosm data be used in model validation. The group's emerging hypothesis here:

- Mesocosms are frequently unrepresentative of in-field soil dynamics, especially if the soil column is not carefully preserved (affecting water-flow dynamics) or the mesocosm is too small (small pot, root-bound systems)
- Thus, potential guardrails on mesocosm data used to validate include sufficient size of the mesocosm and carefully intact preservation of the soil column structure

Feedback prompt: Any feedback on these initial thoughts about using mesocosm data for model validation?

Working Group 5 - ERW Life Cycle Assessments

Scope of the meeting

This life cycle assessment Working Group convened and discussed the topics covered below in Section 5.1-5.4. This Working Group will not meet again and will discuss asynchronously after market feedback is collected and synthesized.

5.1 System boundaries of ERW LCA

Status: **Blue / Recommendation**

The Working Group had a clear recommendation on what should be included within the system boundary for cradle-to-field LCA:

- All operations required for extracting the weathering material, including fuels, electricity, water, and other material and equipment inputs
- All operations required for processing of the weathering materials to desired particle size (crushing, grinding, milling), including fuels, electricity, water, and other material and equipment inputs
- All operations required for loading, transportation, storage of the weathering materials, including fuels, including fuels, electricity, water, and other material and equipment inputs
- All operations required for spreading the weathering materials, including fuels, electricity, water, and other material and equipment inputs

GHG fluxes from field to oceans are not covered in th Working Group 5, they are instead covered in Working Groups 2 and 3.

The Working Group also had a clear consensus to exclude emissions associated with monitoring and verification (traveling to site, sampling, analysis, model simulations) from the system boundary. ERW is still in very early phases of deployment where direct and sometimes redundant measurements are necessary to reduce uncertainty. Given such, inclusion of emissions associated with monitoring and verification could skew results and disincentivize robust MRV. Therefore, the recommendation is to exclude.

Feedback prompt: Are these LCA system boundaries in line with your current quantification approach? If not, what are your specific concerns or reservations?

5.2 Multifunctionality practices when weathering materials are waste, byproducts or secondary products

Status: **Blue / Recommendation**

The Working Group had a clear recommendation on an option set for multifunctionality best practices when the weathering material is a waste, byproduct or secondary product:

- It is okay to use a “cut off approach” approach if the weathering material is a waste or a natural byproduct of another operation with no price or economic value. In this case, the extraction of the weathering material is considered burden-free.
- It is better to use expand the system boundary or allocate burdens when the weathering material has a price or economic value:

    - If the “system expansion” approach is used, suppliers should transparently show how they define the expanded and substituted systems, what inputs and output flows are being accounted for, and underlying data and calculations
        - ***Special note: Section 5.4 discusses whether the mining, processing and spreading of agricultural lime can or cannot be chosen as the substituted system
    - If the “allocation” approach is used, suppliers should provide the inventory of input and output flows, and underlying data and calculations. It is also recommended that suppliers provide a sensitivity analysis of LCA results from different allocation approaches (by mass, by economic value, by energy input, etc.), and be conservative in their selection of allocation approaches.
- The Working Group has a strong recommendation for data and analysis transparency at the early stage of ERW deployment, so that meta-analysis can be performed to design future policies that incentivize the most climate and societally beneficial outcomes.

Feedback prompt: Are the above options for treating the multifunctionality problems of waste, byproducts and secondary products in line with your current quantification approach? If not, what are your specific concerns or reservations?

5.3 Economic co-benefits and/or leakage when weathering application results in yield and land use changes

Status: **Blue / Recommendation**

The Working Group thinks that it is premature to incorporate yield and land use changes in LCA at this stage. The synthesis of the discussions below:

- It is inherently difficult to attribute yield changes as a result of weathering application. Variance seen is significant, so is heterogeneity across soil types and climate regions.
- The consensus is to be very conservative with any yield increase claims. Sufficient proof would require multi-year statistically stable data, compared against a control or counterfactual that incorporates optimal present-day liming/fertilizing practices.
- The Working Group also shares that it is recommended for suppliers to take penalties if there is a reported yield decline or loss.

At the same time, the Working Group also strongly recommends to request reporting of year-over-year yield analyses from ERW deployments. Data collection starting now will make multi-year meta-analyses much easier down the road.

Feedback prompt: Do you track yield impacts as a result of your ERW deployment? Do you account for yield impacts as part of your LCA? Any comments on the Working Group recommendations?

5.4 LCA baseline when an ERW deployment replaces agricultural liming practices

Status: **Purple / Recommendation under development**

The hypothesis is that the standard will include some recommendations on how to conduct LCA baseline accounting when an ERW deployment replaces agriculture liming. Two different approaches are simultaneously discussed:

- One approach is system expansion + substitution - i.e., accounting for life cycle emissions of weathering materials plus the avoided life cycle emissions of agricultural lime.
    - Rationale here is to encourage more detailed meta-analyses to prove that weathering application is net additional to agricultural liming
- A more conservative approach is to not allow the subtraction of avoided cradle-to-field emissions of agricultural lime in v1 standard, when the scale of ERW deployment is small and limited.
    - The argument here is that the avoided emission accounting is problematic today because it's hard to verify the economy-wide impact of ag lime production decrease from a single deployment of weathering application.

Feedback prompt: Do you include avoided emissions of substituted-out agricultural lime in your LCA today? Why and why not?

Working Group session 2: Summaries and initial requests for feedback

The second sessions for each of the five technical Working Groups were convened on December 4-6th, 2023.

Working Group 1 - Aqueous measurements and sampling deep soils

Scope of the meeting

This session addressed 1) the creation of guardrails for taking and interpreting aqueous-phase measurements, and 2) the need for measuring cation fluxes at depths below the weathering zone (in both the solid and aqueous phase). The beginning of the meeting was spent discussing market stakeholders' feedback on the last session's material.

1.1 Market Feedback Theme #1: Feedstock Characterization Frequency

Status: **Purple / Recommendation under development**

In Session 1, the Working Group recommended that suppliers characterize their feedstock once per application event. Market participants stressed the need to refine the requirements for feedstock characterization frequency (as a function of time, volume, or other metric), and that once per application event is not feasible for suppliers.

Working Group discussion:

- The Working Group wants to work towards a tonnage-based requirement for frequency of feedstock characterization, suggesting a baseline requirement of feedstock characterization between [every 1000 and 5000 tons]. (These numbers are still under revision.)
- The working group also suggest that the frequency requirements be in part tied to the underlying heterogeneity of the feedstock composition and thus the error of supplier feedstock characterizations, such that demonstrating more vs less heterogeneity in feedstock properties like major oxide composition could allow for revisions to the above tonnage based requirement.

Feedback prompt: Does a feedstock characterization frequency requirement that includes a tonnage-based requirement along with a heterogeneity-based requirement make sense to you? For suppliers, how do you currently demonstrate the heterogeneity of your feedstock sources?

1.2 Market Feedback Theme #2: BET Surface Area vs PSD

Status: **Purple / Recommendation under development**

In Session 1, the working group included BET surface area characterization as a component of required feedstock characterization. Some market participants suggested that using BET as a feedstock characterization mechanism is operationally and commercially infeasible at scale, and that measuring Particle Size Distribution (PSD) should be sufficient.

Working Group discussion:

- The purpose of requiring BET data for quantification is mostly as a sanity check on empirical dissolution rates, as well as for parameterizing models.
- The relationship between PSD and SSA (specific surface area) is not well-constrained. PSD doesn't account for surface roughness or internal porosity, so if anything it will generally represent a conservative estimation.
- The working group was evenly split as to if PSD should be sufficient for reporting, or if SSA / BET data should also be required. Potentially, BET could be required at a lower frequency than PSD.

Feedback prompt: Feel free to surface any other thoughts on the need for BET characterizations if any, before an initial decision gets made on this.

1.3 Market Feedback Theme #3: Pre-Deployment Sampling

Status: **Purple / Recommendation under development**

Suppliers indicated that requiring high resolution sampling prior to deployment would be unmanageable and would lead to deployment delays because of planting/harvesting schedules being farmers' top priority.

While the working group acknowledged the existence of pre-existing agronomic measurement proxies (pH, soil type) and soil maps (e.g., SSURGO), they stressed that these will almost certainly not include immobile cation concentrations (e.g., titanium), and that rigorous compensatory claims require an initial sampling density high enough to capture the potentially large spatial variability.

- Most of the working group voted to require a round of pre-sampling when it represents the only way to understand baseline heterogeneity of trace elements, despite the operational challenges.

Feedback prompt: If suppliers feel strongly that operational constraints would make a round of pre-sampling infeasible, we would love to hear stronger arguments or alternative ideas as to how to ensure initial sampling density fully accounts for baseline heterogeneity of immobile trace elements in ways that don't require pre-deployment sampling. What other approaches could be sufficiently rigorous?

1.4 Aqueous Measurements

The following three sections relate to setting requirements and guidelines for how aqueous-phase measurements should be taken and interpreted, if suppliers choose to use them as part of their quantification approach.

1.4.1 Hydrologic Characterization and Spatial Heterogeneity

Status: **Purple / Recommendation under development**

For broader scale hydrologic characterizations, **suppliers must use water mass balance modeling** in a site-specific way. As important elements of calculating the water mass balance, the working group called out the following variables:

- Precipitation
- Irrigation
- Measurements needed to estimate evapotranspiration (wind speed, temperature, relative humidity, solar irradiance and sunlight hours, etc.)
- Soil infiltration capacity

1.4.2 Temporal Heterogeneity

Status: **Blue / Recommendation**

The working group agreed that suppliers must interpret aqueous phase measurements through time against a temporal baseline of counterfactual fluid and alkalinity flow through the soil profile.

The group largely agreed that the following two options, if performed well and with reasonable characterization of uncertainty, could be reasonable ways of doing this:

- **Option 1**: Take simultaneous aqueous-phase measurements on an unamended control plot that is sufficiently representative of the treatment soil, and use these measurements as the temporal counterfactual. Guardrails would be needed for the control plot surrounding size, representativeness, and being hydrologically isolated.
    - Other working group members noted that getting a control plot to work in practice has been quite challenging due to hydrological differences.
- **Option 2**: Estimate a temporal baseline using real-time precipitation and irrigation data (or alkalinity-discharge relationships if monitoring catchment drainage), without a control measurement.

The group also mentioned the potential of using radiogenic tracers as an approach to constructing time-varying baselines.

Surrounding uncertainty characterization:

- Suppliers must fully characterize the uncertainty of their aqueous measurements, including spatial variability, temporal variability and uncertainty associated with their dynamic baseline.
- This is a complex system with many parameters, so suppliers should think about using a Bayesian approach, bootstrapping, or resampling methods.

Feedback prompt: For suppliers that use aqueous measurements, how do you take spatial and temporal variability into account? How do you create a time-varying baseline to interpret measurements against? Do the requirements above seem reasonable?

1.5 Measurements of deep soils (deeper than 30 cm)

Status: **Green / ideas**

Feedback from the previous session was mixed on the topic of deep soil measurements:

- Some suppliers stated that deep measurements, especially meter-long soil cores, were a part of the measurements they perform at lower density.
- Other suppliers stated that sampling >40cm deep in the soil column lacks utility and was commercially nonviable. Some mentioned the operational constraint that many farms, especially on those outside of the U.S., do not till and the ground is too hard to get deep cores.

The group discussed two potential purposes for deep soil measurements:

- Actually picking up a resolvable signal of a change in the fluxes we care about (e.g., carbonate precipitation, sorption, etc) in soils below the weathering zone, deeper than 30 cm.
- Collecting empirical constraints to more accurately model deeper soils, but without expecting to measure the fluxes themselves at depth.

The Working Group agreed that operational challenges will in many cases make deep solid-phase or aqueous measurements challenging or impossible, such as hardpans. However, most group members were in favor of requiring at least some degree of deeper soil sampling, even if at a lower density, to close the mass balance and better trace the fate of cations after measuring a dissolution rate.

The group discussed the potential to create a flowchart to identify site-specific characteristics that might allow for elimination of particular depth measurement requirements on a site-specific basis. While this hasn't been fleshed out yet, it could include characteristics such as:

- Soils with a hardpan vs with deep profiles
- Water table depth
- Saturation states that suggest carbonate precipitation will be negligible

A smaller group will need to get together over the next month to dive deeper into how a standard should at a high-level prescribe deeper soil measurements.

Feedback prompt: 1) If you as a supplier routinely take deeper soil cores or deep aqueous measurements, what sort of analysis do you do and at what sample density? 2) If taking deeper measurements at all feels operationally infeasible, we'd love to hear more about why, and whether these constraints can be overcome.

Working Group 2 - Secondary mineral precipitation and dealing with time in net CDR quantification

Scope of the meeting

This was the second session digging into how cation, alkalinity, and DIC dynamics within soil systems impact net CDR for enhanced rock weathering deployments. After considering market practitioner feedback from the first session, the Working Group discussed acceptable avenues for constraining secondary mineral precipitation in the soil profile. For the second half of the session, the discussion turned to options for dealing with time in net CDR quantification. It should be emphasized that this was very much an initial discussion by the Working Group, and this question will be the subject of future Working Group <> market practitioner joint problem solving sessions and successive rounds of Working Group deliberations.

Session 1 feedback themes from market participants:
<u>Theme 1</u>: In the last session, two options were discussed for how a supplier should be required to account for baseline liming practices, if the ERW deployment leads to a change in practice:

- Option A: For any change in liming practice, both the lifecycle emissions and net CDR associated with that baseline liming must be quantified. Only net CDR that is additional to this baseline may be credited.
- Option B: The deployment must only demonstrate net additionality (net CDR from the ERW deployment > net CDR from the baseline liming practice), not explicitly subtract CDR driven by historic liming practices.

Market practitioners were evenly split on which of these two options was the best path forward, with many voicing that Option A is more technically rigorous, but would prove too constrictive at this point in time.

- This was not further discussed by the Working Group at this time, but instead will be considered in a small group convening / problem solving session alongside members of other Working Groups and the broader ERW community.

Theme 2: There was overall market practitioner agreement and support for considering separate near-field and far-field conformance zones.

Theme 3: Multiple respondents highlighted that *directly measuring* DIC at each deployment would prove too costly.

- This feedback was presented to both Working Groups 1 and 2, but primarily addressed during Working Group 1's discussion of aqueous-phase measurements. The importance of correcting DIC measurements for degassing was highlighted here.
- The Working Group discussed the tension between advocating for what's most scientifically robust vs. what's pragmatic. The group emphasized the utility of identifying areas and questions that need to be met by the academic community ("grand challenges or R&D priorities") versus what is needed now from suppliers on a deployment by deployment basis.

Theme 4: There was broad agreement amongst market practitioners with the suggested treatment of strong acid weathering (defaulting to subtracting a conservative estimate of strong acid weathering from net CDR). There was feedback to consider non-ammonia based fertilizers, the details of quantifying the charge balancing anions, and the importance of feedstock characterization (including minor carbonate phases).

- This was presented to the Working Group but not discussed further.

2.1 Accounting for secondary carbonate precipitation

Status: **Purple / Recommendation under development**

The Working Group discussed what would constitute an adequate set of measurements or heuristic rules to qualify the impact of carbonate precipitation on net CDR.

Four options were discussed in detail (A-D below), and there is a strong emerging consensus that a version of option D should be developed for the standard:

- **Option A:** Carbonate precipitation is already captured by the feedstock dissolution measurement.
    - If net CDR is quantified by measuring the time-integrated alkalinity flux at the base of the soil profile (e.g., monitoring of tile drainage,

    - deep lysimeters, or catchment drainage waters), carbonate precipitation in the soil profile is already captured.
    - If feedstock dissolution is quantified by solid or aqueous-phase measurements in shallow soils, carbonate precipitation within the measurement depth horizon is captured, but the potential for carbonate precipitation deeper in the soil column must be accounted for.
  - The Working Group was hesitant about this approach, both because the alkalinity flux would not account for the carbon stored as carbonates, and challenges with the measurement itself (including temporal dynamics and the influence of organic alkalinity).
- **Option B**: Carbonate precipitation throughout the soil profile is constrained by additional measurements of soil inorganic carbon content.
  - The group highlighted that directly measuring changes in soil inorganic carbon content could be very challenging given baseline spatial heterogeneity and the fact that even small changes to the soil inorganic carbon pool could be important to net CDR quantification. High spatial density sampling may be required.
- **Option C:** Carbonate precipitation throughout the soil profile over time is estimated with a reactive transport model, parameterized with site-specific soil physiochemical and hydrologic characteristics and climatic forcing.
  - The Working Group highlighted that relying on reactive transport modeling to constrain carbonate precipitation is also challenging, given the need to accurately capture soil water dynamics, the sensitivity to modeled pH, and the nuanced mechanisms of carbonate mineral formation not currently accounted for in most models. The group also highlighted that simple geochemical calculations can go a long way in many systems.
- **Option D**: A decision tree consisting of empirical heuristics / 'rules of thumb' is used to determine whether and how carbonate precipitation needs to be taken into account for a given deployment. These could include:
  - Rules based around soil pH. For example, if the average (post-application) soil pH remains below 6.5, the potential for carbonate precipitation can be assumed to be negligible. Above pH 8.0, carbonate precipitation must be directly measured. In between, some combination of direct measurements and geochemical modeling may be deployed.
    - In practice, this is complicated by spatial and temporal gradients in pH. In the very least, soil type may need to be an additional differentiator, and/or depth gradients in pH must be considered.
  - Rules based on the carbonate saturation state throughout the soil profile, or background soil conditions (e.g., does carbonate already form in the soil profile?).

    - Rules based on the feedstock used and deployment microclimate (precipitation vs. potential evapotranspiration, moisture regimes, etc.).

A subset of the Working Group, likely working in collaboration with market practitioners in a joint problem solving session, will work to develop the decision tree described in Option D in January.

Feedback prompt: Do you have any feedback on this proposed approach for determining how a supplier must quantify carbonate precipitation in the soil profile? Do you have any suggestions for the design of the 'guardrails' or branches in the decision tree used to determine whether and how carbonate precipitation needs to be taken into account?

2.2 Secondary Silicate Formation

Status: **Purple / Recommendation under development**

The Working Group discussed a similar suite of options for accounting for secondary silicate formation in the soil profile. This includes both authigenic clay formation and incongruent dissolution that leads to the formation of metastable Al/Fe (oxy)hydroxide phases, which can then become cation enriched.

While the Working Group did not reach consensus on the minimum requirements the standard could articulate for accounting for secondary silicate formation in the soil profile, the group was broadly aligned on the following counts:

- Direct measurements to constrain secondary silicate formation could be incredibly challenging, especially over short timescales.
- Some version of Option A articulated above could be viable, but ideally you would want to be sure you're capturing all of the processes that could be influenced by secondary silicate precipitation - including organic carbon dynamics. In addition, measurements in shallow soils or to 'agronomic depth' (10-30cm) would be insufficient, but as in Working Group 1, the group did not reach consensus on the depth to which suppliers should be required to regularly sample or monitor.
- Unlike carbonate precipitation, it is difficult to identify regimes where secondary silicate precipitation could be assumed to be fully negligible. Driving towards a version of Option D, where a combination of feedstock and soil characteristics could be mapped onto an estimate of the magnitude of CO_2 evasion driven by secondary silicate formation, could be a good priority target for future R&D. It was noted that iron and aluminum oxides can be tractably measured using sequential extraction techniques, and these measurements could be used to train models.
- As discussed by Working Group 4, there is potential to try to capture secondary silicate formation in reactive transport models, but more validation of these

post-dissolution cation dynamics is required before such models can be used for compensatory claims.

- One concrete proposal for potential inclusion in the v1 standard was requiring the development of a stoichiometric matrix (e.g., Morel and Hering Tableau method) to track the predicted acid-base reactions for a given feedstock application at a specific site. While potentially operationally difficult to complete, this would enable consistent, transparent accounting across projects.

Overall, the Working Group concluded that constraining the impact of secondary silicate formation should very much be a key R&D priority. The group will return to the question of consistent, practical paths forward for the v1 standard in subsequent discussions and problem solving sessions.

Feedback prompt: For suppliers: do you currently attempt to constrain secondary silicate formation in your deployment or research sites? Are there measurement techniques you would like to highlight for the Working Group? Do you have any specific feedback on the operational feasibility of developing a stoichiometric matrix for a specific feedstock application/deployment?

2.3 Frameworks for dealing with time in net CDR quantification

Status: **Green / ideas**

One of the emergent themes from Session 1 was the importance of time and developing a common framework for dealing with processes that occur over different timescales. What constitutes rigorous accounting for many of these processes is dependent on the time horizon considered, and we need to be able to consistently account for temporal lags in the system. In particular, following alkalinity release through feedstock dissolution, cation transport and (temporary) storage in the soil profile can lead to non-negligible lags between feedstock dissolution (potential CDR) and durably realized CDR.

Operationally, this leads to two critical and very much linked questions: When can net CDR be credited, and at that time of crediting, how are prior and future carbon gains/losses folded into the net CDR quantification?

As an initial foray into this topic, the Working Group discussed three representative frameworks for dealing with time in CDR quantification. Importantly, the question of how the v1 quantification standard should deal with time is not one that can be answered in a single 45 minute discussion session. We will return to this subject with representatives from across the ERW (and broader CDR) community in the coming weeks.

Option 1.0: ~Status quo. Ignore transient dynamics.

- Crediting occurs when the alkalinity is released through feedstock dissolution (potential for CDR). Prior and future permanent losses are folded into the CDR quantification at that time, with no temporal discounting framework - functionally assumes all upstream LCA emissions + downstream losses occur simultaneously with the alkalinity release.
    - Transient carbon losses (e.g., cation sorption in the soil profile) and potential downstream gains (e.g., additional CO_2 uptake due to carbonate system equilibration in a buffered, higher pH system downstream) are not considered.

Option 2.0: Temporal dynamics must be considered within a 'near-field' zone

- Detailed tracking of the temporal dynamics of net CDR is required until the weathering flux passes a spatial threshold (e.g., passes from the soil profile into groundwater; passes a soil depth outside of 'diffusive exchange' with the atmosphere; or passes a set depth in the soil - e.g., 1m).
    - The Working Group discussed potential options for this spatial boundary, but did not reach consensus. Attempting to define a depth at which the soil is no longer in diffusive exchange with the atmosphere is attractive, but likely operationally impractical - the diffusivity can change significantly through time. Using the groundwater depth could be operationally simpler, but would entail tracking fluxes to great depths in regions with low water tables. A hybrid approach, with an operationally defined spatial boundary for the 'near-field' zone that allows for flexibility based on the context of a specific site, is likely required.
- **Option 2.A.** Most conservative: wait until the carbon passes out of the near-field zone to credit.
    - Credit the carbonate alkalinity flux that crosses from the near-field to far-field conformance zone, and fold LCA emissions + downstream net predicted carbon loss at that time.
- **Option 2.B.** Credit the potential CDR associated with alkalinity release from the feedstock during a given reporting period, tempered by the *change* in soil stocks that represent a reduction in net CDR.
    - The *change* in carbon sink terms during the reporting period (e.g., change in base cations in the exchangeable fraction; inorganic carbon content; cations incorporated into secondary clays; plant uptake; potentially soil organic carbon pools) is calculated for the integrated soil profile in the near-field conformance zone. LCA emissions + downstream net predicted carbon loss are also incorporated with this crediting event.
    - Net CDR for a given reporting period = $CDR_{alkalinity\_released}$ + change in carbon sink terms - LCA emissions - net downstream carbon loss

        - At each reporting period, you'll have a new assessment of the amount of feedstock that's been dissolved and a new measurement/model of inorganic C content, cations in the exchangeable pool, etc. That tells you how those carbon pools have changed (how much alkalinity is tied up on sorption sites, how much net new carbonate is in the soil profile). That change could be due to alkalinity released from the feedstock during either this reporting period or from previous reporting periods. But the change gets assigned to CDR calculated for this reporting period, thereby accounting for the temporary sink (or release) of CO_2.
    - Where could this go wrong? Is there potential to need to 'clawback' credits here?
        - One example: Very little alkalinity is released from the feedstock during a reporting period, but cations released during prior reporting periods move into lower base saturation zones deeper in the soil profile and get tied up on exchange sites. Net CDR for that reporting period could be negative.
        - This could be addressed by 'holding back' unrealized potential CDR in the soil profile (e.g., integrated exchange sites in the soil column) until it has passed the near-field threshold.
- These options do not explicitly deal with the temporal dynamics of downstream processes. Instead, the largely conservative assumption is made that the net downstream loss occurs at the time of alkalinity release. But they *could*, in the form of a far-field crediting event for a reporting period. If models of the full catchment-scale proton balance are sufficiently rigorous to capture net additional CDR after the weathering flux has reached a surface water system, the supplier could credit *additional* net CDR at that time (when the CDR has actually occurred).

Option 3.0: Focus on the *time* it takes the weathering flux to transit the soil profile, and discount or credit accordingly.
- Using either a site-specific reactive transport model or a transfer function that translates measured soil properties into a 'transfer rate' or average residence time, constrain the lag time between feedstock dissolution and the released alkalinity exiting the near-field conformance zone.
- **Option 3A**: A CDR credit cannot be generated until the lag time has passed. However, the ERW credit could be coupled with a short-term warming credit or high quality but low permanence CDR credit to cover the CO_2 equivalent residing in the atmosphere for the duration of the lag time. Another version of this would be 'in-setting' the impact of avoided N_2O fluxes from a given deployment, if the net greenhouse gas balance can be sufficiently rigorously measured or modeled.

- **Option 3B**: Temporal discounting is used to decrease the value of the credit generated at the time of feedstock dissolution to account for the climate impact of CO_2 residing in the atmosphere for the duration of the lag period.

Overall thoughts from the Working Group:

- The Working Group sees promise in Option 2 (specifically Option 2B) as it would generate a consistent, transparent set of equations and account for temporal lags in the near-field zone. There was some hesitancy as to whether it is currently tractable to track all of these fluxes in practice, but agreement that this should be the long-term goal.
- A temporal discounting framework like that suggested in Option 3 could also be considered.
- Working Group members flagged the importance of considering how the framework developed for the v1 standard relates logically to current norms/practice in other CDR pathways that have temporal considerations (e.g., biomass pathways and ocean alkalinity enhancement).

Feedback prompt: Do you have any preliminary feedback on the operational feasibility of these different options for accounting for time in net CDR quantification?

Working Group 3 - Downstream evasion—catchments, rivers, ocean

Session 2: Accounting for ocean carbonate dynamics

Scope of the meeting

The alkalinity generated by enhanced rock weathering and the associated CO_2 removed from the atmosphere as charge-balanced dissolved inorganic carbon (DIC) is not durably stored until it reaches a long-duration reservoir: in some cases a long residence time groundwater system, and in many cases the ocean. The last Working Group session focused on quantifying potential carbon evasion pathways in surface water systems. In the first half of this session, the Working Group considered market practitioner feedback and discussed options for actionable guidance for the v1 standard. The Working Group then considered the quantification of CO_2 loss due to ocean carbonate dynamics.

Session 1 feedback themes from market participants:

Theme 1: Regarding storage in long-residence time groundwater systems. Most suppliers do not currently consider storage in long-term groundwater reservoirs; some are exploring. Market practitioners provided feedback that groundwater access may change in the next thousand years (droughts, new wells, etc), meaning that long residence time groundwater is still brought to the surface. In many places, field irrigation is also carried out with groundwater, meaning that the weathering products in the groundwater could be re-distributed over farmland multiple times.

- The Working Group agreed with the feedback provided.

Theme 2: In Session 1, the Working Group articulated a list of carbon system fluxes that should be accounted for in a v1 standard to determine potential losses in surface water systems. There was market consensus that the fluxes identified do need to be accounted for, but some market practitioners have significant reservations on how to do so in a commercially viable way. There are concerns about accounting for these fluxes in emerging markets lacking extensive river network + chemistry databases.

- The Working Group returned to the question of requirements for modeling potential carbon evasion in surface water systems, with a particular focus on data availability, in the first half of the session. See Section 3.1 below.

Theme 3: In the previous session, there was an emerging hypothesis from the Working Group that the standard could require a process-based model to quantify potential evasion from surface water systems. Feedback from suppliers reflected a roughly even split between those in favor of employing process-based models and those in favor of using a conservative haircut. Again, there were concerns about data availability to parameterize process-based models in countries that do not have extensive river network + aqueous chemistry databases built out.

- The group very much empathized with the concerns about data availability, and added that even in countries that we believe to be data rich, developing models that are sufficiently validated remains a formidable challenge. There is still a significant knowledge base limitation that the industry is working to overcome.

3.1 Quantifying potential loss pathways in surface water systems: Driving towards acceptable practices for articulation in the standard

Status: **Blue / Recommendation**

Working from this feedback, the Working Group returned to the question of actionable options for constraining potential riverine losses for the v1 standard. The group first reconsidered whether an appropriate conservative loss estimate could be developed, particularly for application in regions with limited data availability. Ultimately, the working group's discussion led to the synthesis of a "middle-ground" approach to the issue. This "middle-ground" attempts to walk the line between requiring all suppliers to use a process-based model and allowing for a conservative loss estimate that could be applied to any deployment. The Working Group particularly stressed that a catchment-specific approach must be taken, even when considering empirical loss estimates, as opposed to a single global conservative "haircut".

The two "middle-ground" options that the working group landed on are as follows:

Option A: The ERW community is given the task to develop a *global map* that visualizes and delineates the vulnerability of certain catchments to downstream losses. This vulnerability index could then be tied to different magnitudes of conservative loss estimates. This would ideally become a community data product available to all suppliers as a default conservative loss estimate for deployment in a given catchment.

Option B: Simplified models can be developed, extrapolating or drawing insights from catchment areas that are currently data rich, to estimate the magnitude of downstream evasion from a smaller subset of characteristics about a catchment. This would allow for the development of downstream loss estimates from a much wider range of 'data-poor' regions. Importantly, basic data from a deployment catchment would still be required (e.g., soil and water pH).

This led to a discussion of monitoring requirements and to what extent suppliers might be responsible for some degree of downstream monitoring of their deployments. There was emerging consensus that, in the very least, *some* data is required about a catchment. If a certain minimum viable dataset does not exist, the supplier may need to collect data on that catchment. With regards to monitoring after a deployment has taken place, the group was broadly aligned that it would be unrealistic to have a supplier do their own bespoke river network monitoring, but stressed that suppliers would in the very least need to be transparent with the details of how much, how often, and where they are deploying. There were discussions about the potential to use hydrologic calculations to define a 'zone of impact' that a supplier could be required to monitor (considering the volumes of water going from a small sub catchment into the main flow stream, and considering when the impact would be sufficiently diluted in a deployment and catchment-specific context), but this requires further discussion by the Working Group.

Feedback prompt: Given these two "middle-ground" options, which do you think would have more utility and/or traction? Do you have any feedback for the Working Group on the practical implementation of these options? For suppliers, are you currently doing any downstream characterization / monitoring of your deployments, and do you have any feedback on the operational feasibility of downstream monitoring in a spatially constrained 'zone of impact'?

3.2 Modeling the catchment-scale proton and alkalinity balance for quantifying potential net additional downstream CDR

Status: **Purple / Recommendation under development**

In Session 1, Working Group 2 highlighted a number of instances when less conservative assumptions could be made (e.g., that strong acid weathering in some instances could still be considered to drive net CDR) if a sufficiently rigorous catchment model of the site-specific counterfactual and ERW deployment scenario could be developed.

The following question was posed to the Working Group:
For the purposes of the v1 standard, what would constitute a sufficiently rigorous catchment-scale model to capture alkalinity transport + resulting DIC dynamics? Do sufficient validation frameworks and datasets exist to use catchment-scale process-based models to credit additional net CDR downstream? What guardrails would need to be put in place?

The working group briefly touched on this subject, but given time constraints decided to return to the discussion asynchronously, and potentially in a later session.

3.3 Quantifying losses in coastal & open ocean settings

The following three sections relate to creating requirements and guidelines for quantifying potential downstream CO_2 losses in coastal and open ocean settings.

3.3.1 Identifying priority carbon loss pathways in coastal and open ocean ecosystems

Status: **Purple / Recommendation under development**

In the previous session, the Working Group articulated a list of potential CO_2 loss pathways that should be considered in rivers / surface water systems for the v1 standard. The Working Group considered the same question for coastal / open ocean ecosystems in this session, focusing again on delineating what needs to be considered for the v1 standard vs. remains a R&D priority.

There was strong consensus amongst the Working Group that CO_2 evasion due to carbonic acid system equilibration should be considered in the v1 standard. There was consensus that the influence of carbonate mineral precipitation and burial is important to account for, but that there might be options to craft volume-dependent requirements, such that suppliers deploying low volumes of feedstock far away from the ocean need not explicitly account for carbonate burial in marine ecosystems. Changes to organic matter respiration/metabolic activity & authigenic clay formation were discussed, but the Working Group concluded that these are very much active research areas that fall outside the scope of the v1 standard.

Feedback prompt: Are there other carbon flux pathways that you would like to raise to the Working Group as important avenues for carbon loss in coastal and open ocean ecosystems? Are there tractable approaches to quantify their impact?

3.3.2 Homing in on requirements for models or empirical estimates for quantifying the magnitude of potential CO_2 evasion due to the processes identified

Status: **Blue / Recommendation**

The working group was asked to determine what requirements should be put in place for the use of models or empirical estimates to quantify the magnitude of potential CO_2 evasion due to the processes identified above.

Three options were discussed (A-C as follows). The Working Group reached a strong consensus that Option B should be presented in v1 of the standard as a baseline requirement, with Option A listed as something that suppliers should strive towards and/or use to demonstrate a 'lower than default' loss magnitude for a given deployment. The group agreed that using a single conservative loss estimate was not appropriate at this stage (Option C), and that more research is needed to determine whether results from relatively coarsely gridded Earth Systems Models (as used by Kanzaki et al. 2023) continue to hold with more finely resolved models.

> **Option A**: A region-specific coastal biogeochemistry model is run, using alkalinity and DIC outputs from the process-based model used to estimate riverine evasion. Uncertainty is quantified through model ensembles in which key parameter values are stochastically varied.
>
> **Option B**: A conservative assumption of evasion from carbonic acid system equilibration is derived by considering the thermodynamic storage efficiency as a worst-case scenario, assuming full equilibration at representative temperature, salinity, and current atmospheric pCO_2.
>
> **Option C**: A set conservative loss estimate (e.g., ~10%) is used in place of deployment-specific modeling.

Feedback prompt: Do you agree with using Option B as a baseline requirement in the v1 standard, with the option and recommendation to pursue Option A if possible? For suppliers, how are you currently accounting for losses in ocean ecosystems? Do you have the in-house capabilities to run coastal ocean biogeochemistry models or Earth System Models?

3.3.3: Potential netCDR losses due to ocean carbonate burial

Status: **Blue / Recommendation**

The working group was given 2 options to discuss regarding ocean carbonate burial over long timescales (100-1000yrs) and how it should be considered in the v1 standard. A consensus was reached that carbonate burial and carbonic acid system equilibration should be lumped together, and that a similar approach to Option B above should be followed. It is again noted that although Option B is not a perfect choice, that it is a strong step in the right direction for a v1 standard that will iterate in the future as the science and rigor bars develop.

Feedback prompt: Do you agree with the working group's consensus to do a similar "Option B" approach for accounting for ocean carbonate burial over 100-1000yr timescales? Do you have any other suggestions on how ocean carbonate burial should be accounted for in a v1 standard?

Working Group 4 - Uncertainty characterization and parameterizing a deployment-specific model

Session 1 feedback themes from market participants:

- Theme 1: Market/suppliers highlight the need for a model intercomparison project, and indicate that validation uncertainty should be fed into CDR uncertainty.
- Theme 2: Mesocosms are a critical tool for model development and parameterization, but all models should be validated using only field data.

4.1 What might the v1 standard say about RTMs?

Status: **Purple / Recommendation under development**

Some potential endpoints for the standard surrounding RTMs that the working group will continue to flesh out:

- The standard will definitely say that reactive transport models cannot be used end-to-end to quantify CDR without any measurements or empirical constraints, but operational challenges with measuring deep soils will require some use of models in conjunction with site-specific empirical constraints for cation sink quantification.
- Suppliers could be required to document their implementation of different processes in a standardized, legible format (e.g., used a "special grabby phase" vs more complex cation exchange implementation)
    - To discuss and start to flesh out in the next meeting.
- Suppliers could be required to report the parametric uncertainty of their models and the assumed prior distributions of their parameters.

    - The working group noted that there are too many parameters to report uncertainties, sensitivities, and prior distributions for all of them, and the group intends to work towards an initial list of most important driving parameters as a baseline.
- Suppliers could be required to show how their model performs against an empirical validation dataset that it hasn't seen before in the tuning or parameterization process, in a pipeline shared across the market.
    - The group is still developing details for how this could work in a v1 standard, versus over the coming year as the community has more time to develop structured, third-party model validation systems.
- Suppliers could be required to collect a variety of different types of pre-specified site-specific data as empirical constraints when parameterizing the model.
- Suppliers could be required to participate in shared model intercomparison efforts and in shared sensitivity analyses.

Feedback prompt: Do these types of requirements for model validation seem reasonable? What else could you imagine in terms of community goods and pipelines to understand RTM predictive skill? Do any of these feel unreasonable or unrealistic?

4.2 Parametric Uncertainty and Sensitivity Analyses

Status: **Purple / Recommendation under development**

The group started by articulating an incomplete, initial list of parameters that may be relevant for required parametric uncertainty analyses and sensitivity analyses. These included:
- Initial reactive surface area of the feedstock
- Kinetic coefficients for mineral dissolution
- Cation exchange capacity
- Soil porosity and soil Ksat (infiltration rate)
- Farmer practices - including cover crops (which might be affecting the nitrogen cycle), fertilizer usage, tillage
- Organic C decomposition rate (+ other biological processes)

However, the group also emphasized that depending on model implementation details and the set of processes included within the model, different parameters will be more vs less important.

For a v1 standard, the group intends to require suppliers to run and submit sensitivity analyses across a finalized set of parameters.

- The group agreed that this must be a **multivariate sensitivity analysis**, due to complex interactions between variables, and highlighted that inclusive stochastic analyses are a good way of doing this.

For the coming year or so, the group started to develop two ideas for community goods related to model parametric uncertainty:

- The group agreed that a shared platform for running the same automated sensitivity analyses for each model submitted would be highly beneficial. An example of this in the vegetation model community is the PEcAn Project.
- The group also wants to create a better matrix of more standard 1) input parameters, 2) output benchmarks, and 3) shared standard boundary condition runs for clear comparison of scenarios across models and how varying each standard input parameter affects performance against benchmarks. One member of the group made an analogy to the ILAMB project, which created a benchmarking system to score land models across a series of variables ("benchmarks") on their performance relative to empirical data.

Feedback prompt: If you use a model (especially to predict deeper soil cation dynamics), how do you estimate parametric uncertainty? What sorts of sensitivity analyses do you run?

4.3 Deployment-specific empirical constraints

Status: **Purple / Recommendation under development**

The standard intends to require that suppliers parameterize their models to be deployment-specific using empirical measurements and site-specific conditions. The group set out to answer the following three questions:

1. What variables are needed as input to parameterize a deployment-specific model?
2. For which variables are regional averages insufficient and field-scale measurements are required? Which variables can be aggregated or regional?
3. How can the uncertainty associated with using more aggregated data in a model be assessed?

The group recognized several forms of variability that might create error or uncertainty in the input values to the model, including spatial variability at a sub-field level, depth variability, and temporal variability (e.g., in water table height, meteorological variables like precipitation, and fertilizer application).

The group decided that:

- Meteorological variables (e.g., precipitation) may be regional averages, trading off spatial granularity for more temporal specificity

- Soil variables need to be measured (empirically constrained) at a field or even sub-field scale depending on the extent of heterogeneity. Important variables noted were:
    - Initial buffering capacity of the soil, including pH, buffer pH, CEC, and base saturation. The group noted that different lab extraction techniques used to obtain buffer pH and CEC can matter a great deal, so this information is likely needed as well.
    - Soil infiltration (including the precipitation vs infiltration relationship, and how much surface runoff is taking place)—with variables including soil type/texture,
    - Other variables mentioned: bulk density, crop type / crop rotation, respiration rates, water table depth

Over the coming meetings, the group will work to develop an understanding of which variables are truly dealbreakers that must be empirically constrained for each field, with rigorous rules for sensitivity analyses and understanding parametric uncertainty for variables that practically need to be gathered in more aggregated form.

Feedback prompt: For suppliers, what sort of in-field empirical constraints do you use to parameterize a deployment-specific model? Do the above level of empirical data requirements make sense to you?

Working Group 5 - Best Practices for Ecosystem Impact Assessment

* The sets of recommendations developed by this Working Group will not turn into quantification requirements in v1.0 standard, but rather be positioned as supplier best practices.

Scope of Meeting

Various efforts have built a strong foundation for ecosystem impact assessments for ERW, including but not limited to:

- Environmental and Social Safeguards requirements in Puro's ERW Methodology
- Health & Ecosystem Impact Assessment Rubric used in the diligence process leading to Frontier's first ERW offtake
- A forthcoming perspective paper by Carbon 180 on the environmental impacts of terrestrial enhanced weathering

The goal of this session is to distill a supplier-facing best practices on monitoring, mitigating and accelerating public learning of the environmental impacts of ERW.

5.1 Categories of ERW ecosystem impacts

The Working Group identified the following categories as ecosystem impacts from weathering applications. This list is intended to be all possible ecosystem impacts.

Terrestrial:

- Releases of toxic trace metals into the soil, groundwater, and uptake by crops
- Changes to soil porosity and physical properties
- Changes to soil chemistry, microbiome, macrofauna, and plants
- Formation of hexavalent chromium
- Agronomic benefits from mineral uptake by crops

Air and Aerosol:

- Human health impacts from particulate matter inhalation
- Soil chemistry, plant damage and ecosystem shifts from dust deposition
- Production of air pollution, NOx, and ozone
- Production of silicate dust materials in scenarios where mining and transport is increased by demand for ERW feedstock

Aquatic:

- Changes in freshwater alkalinity and pH
- Effects of fertilization, such as hypoxic “dead zones”
- Release of toxic trace metals in freshwater environments
- Shifts in marine chemistry
- Sedimentation in rivers
- Reduced nitrogen runoff from increased plant uptake and changing nutrient load in downstream water environments

Feedback prompt: Any new categories are we missing? Any existing categories that you’d suggest we edit?

5.2 Where monitoring and/or mitigation should be required as supplier activities

Of these impact categories, the Working Group almost unanimously agreed that the following categories could and should be actively monitored and/or mitigated by suppliers.

- Releases of toxic trace metals into the soil, groundwater, and uptake by crops
- Formation of hexavalent chromium
- Human health impacts from particulate matter inhalation
- Agronomic benefits from mineral uptake by crops

Beyond the above list, there are several more impact categories where the Working Group suggested that we should also begin to develop monitoring and/or mitigation requirements, all/some of which should be the responsibilities of suppliers.

- Changes to soil porosity and physical properties
    - The Working Group recommends monitoring soil hydraulic conductivity when high amounts of alkaline material is applied.
- Production of air pollution, NOx, and ozone
    - The Working Group recommends monitoring dust production, smog and NOx where increasing mining, transporting and spreading activities are occurring as a result of ERW demand.
- Changes in freshwater alkalinity and pH / Effects of fertilization, such as hypoxic “dead zones” / Release of toxic trace metals in freshwater environments
    - The Working Group recommends continuous, multiple-time-stamp measurements of heavy metal load in freshwater streams and after heavy storm events.
    - ... and considering regulatory enforcement of such monitoring through existing drinking water quality and fish / aquatic life quality standards.

Feedback prompt: *For suppliers*: Which of the above ecosystem impacts are you already monitoring for and/or mitigating? Can you describe on a high level how you are monitoring and/or mitigating these impacts? *For registries and buyers*: Which of the above ecosystem impacts do you require suppliers to monitor for and/or mitigate as part of protocol review or project due diligence process? Can you describe what mechanisms you use to ensure supplier compliance against your ecosystem safeguard requirements?

5.3 Where the ecosystem perturbations could cause the highest risks or considered as potential “showstoppers”

We asked the Working Group to highlight the ecosystem impact categories that, in their view, are potential “showstoppers” (i.e., where the ecosystem risks are considered to be of high enough magnitude that we should put thresholds on the volumes of weathering applications, the types of weathering materials used, the location where weathering applications should be permitted, etc). Here are the categories that the Working Group has highlighted:

- Releases of toxic trace metals into the soil, groundwater, and uptake by crops
    - Why a “showstopper”: Uptake of toxic metals in soils pose a risk to crop safety and human health. And the current state of knowledge is still relatively poor.
    - Recommendations for research priorities and path forward: a clear framework for mapping allowable feedstock materials (at allowable application rates) for different soil types.
- Changes to soil porosity and physical properties

- Why a “showstopper”:
  - Negative changes to soil porosity and physical properties could be a “showstopper” for farmer acceptance. Impacts on soil porosity are even less understood than impacts of metal release in soils.
  - For sophisticated crop production systems in the Global North where soil testing is routine, any major soil porosity changes could halt further weathering deployments. However, the same is not true for crop production systems where data and soil testing are less / not available, particularly in the Global South.
- Recommendations for research priorities and path forward:
  - Threshold limits on application rates, or on the grain size of the weathering materials applied.
  - ... and/or considering co-application of silicate materials with biochar or compost - to help alleviate soil texture changes and replenish organic matter.

- Changes to soil chemistry, microbiomes, macrofauna, and plants
  - Why a “showstopper”: Degradation of chemical fertility will be a major concern for farmers. Specifically, in areas where sodium is already high in local soils, ERW could pose a risk as sodium is the first cation exchanged.
  - The Working Group doesn’t yet have a recommendation for research priorities and path forward.
- Formation of hexavalent chromium
  - Why a “showstopper”: There are existing EPA regulations on maximum contaminant level on chromium in drinking water. The formation of hexavalent chromium through weathering application could trigger regulatory and social acceptance concerns.
  - It is generally agreed that hexavalent chromium should be monitored within soils and water by suppliers, and that their threshold limits be met.
- Effects of fertilization, such as hypoxic “dead zones” and release of toxic trace metals in freshwater environments are also highlighted as potential “showstoppers”.

Feedback prompt: What are on your mind as the highest-risk ecosystem impact categories where the community should be proactively developing monitoring and/or mitigation requirements?

Working Group session 3: Summaries and initial requests for feedback

The third sessions for each of the five technical Working Groups were convened on January 16th-22nd, 2024.

Working Group 1 - Sampling strategy, cation exchange characterization, and gas-phase measurements

Scope of this meeting:

This session addressed sampling design and the representativeness of control plots, characterizing cation sorption as a transient alkalinity sink, and gas-phase measurements.

Feedback from the previous session:
Theme 1 (feedstock characterization): The working group proposed a tonnage-based requirement (requiring a feedstock characterization event between every 1000-5000 tons) as well as a heterogeneity-based requirement. The market feedback indicated that many suppliers already require sampling at frequencies less than 5000 tonnes, so this is in line with current practices.

Theme 2 (pre-deployment sampling): The working group concluded in session 2 that a round of pre-deployment sampling must be required in the standard. Market feedback agreed with this conclusion and noted that there are currently no sufficient proxies and no sufficiently high-resolution soil maps for relevant soil parameters available to rely on to allow pre-deployment sampling to not be a requirement.

Theme 3 (aqueous measurements): In the discussion of aqueous measurements, working group members aligned that:

- Suppliers must use water mass balance modeling in a site-specific way
- Suppliers must either take simultaneous aqueous-phase measurements on an unamended control plot that is sufficiently representative of the treatment soil and use these measurements as the temporal counterfactual **OR** estimate a temporal baseline using real-time precipitation and irrigation data (or alkalinity-discharge relationships if monitoring catchment drainage), without a control measurement.

The market feedback reported that no supplier currently uses water mass balance models. There were also points raised about the feasibility of commercially scaling aqueous
measurements, but general market agreement that these requirements could be put in place in the v1 standard for any supplier who chooses to use aqueous measurements.

Theme 4 (deep sampling): In session 2, the working group agreed that operational challenges will in many cases make deep solid-phase or aqueous measurements challenging or impossible, such as hardpans. However, most group members were in favor of requiring at least some degree of deeper soil sampling, even if at a lower density, to close the mass balance and better trace the fate of cations after measuring a dissolution rate. The group discussed the potential to create a flowchart to identify site-specific characteristics that might allow for the elimination of particular depth measurement requirements on a site-specific basis. Market feedback on this point was generally positive, and suppliers agreed that the flowchart approach is a reasonable path forward for deciding whether a particular situation requires deep measurements and at what frequency they are required.

1.1 Representativeness of control plots

Status: **Blue / Recommendation**

The working group aligned in a previous meeting that suppliers must include control plots in their deployment fields as a sanity check and potentially a temporal baseline against which to compare outcomes in treatment fields. This session explored how we ensure that control plots must be representative of the variability of the rest of the field.

Some key parameters pointed out by the working group that matter to ensure control plots are representative include:

- Any parameters being measured as key proxies / parameters for CDR quantification
- Input additions (especially nitrogen fertilizer for nitrification)
- Water throughput in irrigated systems
- Field topography / slopes that might affect water flow

The group agreed that demonstrating that quantitative parameter distributions are sufficiently similar between control plots and deployment field areas requires pre-deployment sampling to estimate these distributions, and statistical tests must be done to ensure these distributions have similar means and variances. One working group member also noted that areas being considered in these tests should be of similar size (e.g., compare the control plot to subsets of the deployment field that match the control plot in area).

Feedback prompt: How do you currently ensure that control plots are sufficiently representative of the deployment field area? What parameters should be most heavily taken into account?

1.2 Big-Picture Sampling Design

Status: **Purple / Recommendation under development**

The working group broadly agreed that a core goal of any sampling design strategy should be to capture the in-field mean and variability of soil parameters being used to quantify net CDR, and that standard guidance on sampling methods should flow from this goal. The challenges of doing this well were emphasized by a working group member via this Ramsey et al 1995 paper, showing significant differences between different soil sampling designs (sampling bias was present).

The working group discussed a variety of sampling designs: gridded sampling, zig-zag patterns, simple random sampling within a deployment area, stratified random sampling, or sampling along transects (see study here). This book-length reference was provided for further examination of how to ensure each of these sampling designs is performed well. The v1 standard will not prescribe a particular sampling pattern and likely will not rule any of these out, but one of the Problem Solving Sessions may be used to refine requirements for particular patterns above, such as creating rules surrounding sample compositing.

Finally, the group discussed the suggestion of creating a subplot of the full field that is very heavily sampled, with then a lower density sampling regime in the rest of the field. The group didn't rule this out unequivocally, but pointed out ways that this system could be potentially gamed in the process of selecting the subplot area, and the need for strong statistical frameworks and norms to ensure the subplot is representative of the larger field. Problem solving groups will be tasked with articulating rules and guardrails for such statistical frameworks.

Feedback prompt: For suppliers, how do you currently design your sampling regime, and how do you ensure that this design fully captures in-field variability of key parameters being measured?

1.3 Characterizing cation sorption as a transient alkalinity sink

Status: **purple / recommendation under development**

Base cation sorption on exchange sites has important implications for the temporal dynamics of carbon dioxide removal from an ERW deployment - particularly in low base saturation soils. In order to determine when credits can be issued, the group agreed that suppliers need to be able to empirically constrain the proportion of base cations released from the alkaline feedstock that are held up on exchange sites at any given time. The working group was given two potential options for how to characterize the impact of cation sorption within the soil column:

- **Option A:** Aqueous cation concentration or alkalinity measurements are taken at the required depth, which will exclude cations on exchange sites. In this case, we don't need to care about the exact composition of the exchangeable fraction.
- **Option B:** Require soil cores taken down to a required depth, segregate the cores by depth, and measure cation-specific base saturation over time.

There was broad agreement that both options were workable, and it was argued that it could be left up to the supplier to use one or the other, based on their situation.

Feedback prompt: Do you see any issues in adopting an approach that requires constraining cation exchange lags but allows either approach, allowing suppliers to choose either option A or option B? Do you have any different suggestions on how to constrain sorption?

1.4 Gas phase measurements

Status: **blue / recommendation**

The working group broadly aligned on the consensus that the gas phase measurement signal-to-noise ratio is too weak to allow suppliers to rely solely on gas-phase measurements for net CDR quantification. Suppliers who want to constrain CDR through solely gas-phase measurements should be required to prove the robustness of their CDR quantification through comparison with additional solid or aqueous phase measurements.

Feedback prompt: Do you agree with this consensus? Do you have a differing viewpoint?

Working Group 2 - Alkalinity losses via plant uptake, changes to SOC dynamics, and accounting for other soil GHG emissions

Scope of the meeting

This meeting focused on the living side of soils and interactions between ERW and organic matter dynamics. The Working Group discussed requirements for constraining alkalinity losses through plant uptake, accounting for potential changes to SOC cycling as a result of ERW deployments, and how to consider potential changes in other soil GHG emissions such as CH_4 and N_2O.

Session 2 feedback themes from market participants:
Theme 1 (secondary carbonates): To account for secondary carbonate precipitation, Option D was strongly supported in market feedback. Option D proposes a "decision tree"

approach where empirical heuristics/rules-of-thumb are used to determine how secondary carbonate precipitation is accounted for.

- A small group will work to develop this decision tree in a joint working group <> market practitioner problem-solving session and bring a proposal to the Working Group for Session 4.

Theme 2 (secondary silicates): No supplier that replied to the feedback form currently monitors for secondary silicate formation in the soil profile. Market practitioners overwhelmingly agreed that this is an important area for further research, and there was some support to consider secondary silicate formation from a theoretical perspective - e.g., considering what components of the feedstock mineralogy are likely to dissolve incongruently and lead to secondary silicate precipitation.

- The question of whether geochemical calculations should be used to constrain the potential for secondary silicate formation, or if secondary silicate formation is largely left as an R&D priority for the v1 standard, will be addressed in a joint working group <> market practitioner problem-solving session. A final decision will be made during Session 4 of this Working Group (with opportunities for asynchronous feedback/discussion).

Theme 3 (time): Dealing with time in net CDR quantification. Option 3A makes sense to many but there was hesitation that it would be too complex of a crediting situation for buyers to trust the system *(Option 3A: credit can't be generated until lag time has passed, but credit could be coupled with a short-term warming credit or high quality/low permanence credit to cover the CO_2 remaining in the atmosphere for the total lag time)*. Option 2B is recognized as the most scientifically rigorous option, but many were hesitant about being able to sufficiently track the fluxes involved *(Option 2B: Credit the potential CDR associated with alkalinity release from the feedstock during a given reporting period, tempered by the change in soil stocks that represent a reduction in net CDR)*. Option 1 is recognized as the most feasible, but least accurate option. No consensus amongst feedback on which option is the best option to move forward with *(Option 1: Status quo; crediting occurs when the alkalinity is released through feedstock dissolution)*.

- The question of how the v1 standard should consider temporal dynamics (importantly within the broader context of how time is treated in other CDR pathways) will be revisited by a joint working group <> market practitioner problem-solving group.

2.1 Alkalinity losses via plant uptake

Status: **Purple / Recommendation under development**

The session started with the question: What needs to be measured or constrained to sufficiently capture alkalinity losses from plant uptake across a wide range of crop types? The group discussed how accounting for base cation uptake differs between perennial crops vs annual crops, the current lack of sufficient synthesized data on cation uptake rates for different crop types (including as a function of growth rate, base ion concentration, and nutrient ratios), and whether suppliers should be required to take direct measurements of cation concentrations in harvested biomass or new growth in the case of permanent crops.

The group was in agreement that plants can take up a significant amount of base cations, particularly for certain elements such as potassium. With this knowledge, the group decided that *alkalinity loss due to plant uptake cannot be ignored in the net CDR calculation*. The requirements for constraining plant uptake will depend on the measurement technique being used to quantify net CDR. If net CDR is constrained by the time-integrated alkalinity flux at the base of the soil profile (e.g., monitoring of tile drainage, deep lysimeters, or catchment drainage waters), then the alkalinity loss from plant uptake has already been accounted for. If solid phase measurements are being taken, then the supplier would be required to separately consider plant alkalinity losses as a part of their mass balance equation.

Posed with the question of whether look-up tables either already existed or could be compiled to provide a reasonable estimate of element-specific plant uptake for different crop types, the group agreed that there is not currently enough information available about how base cation uptake varies across crop types, or as a function of growth rate, growing conditions, soil chemistry, etc. Once more is known about what levers affect plant base cation uptake across different crop types, such look-up tables could be developed. However, to create this lookup table, much more research and empirical data are needed from field trials.

The group thus decided that alkalinity loss via plant uptake needs to be informed by direct measurements of the base cation content of biomass harvested from annual crops, or by direct measurements of new growth in perennial crops or forestry systems. For perennial crops or forests, new growth estimates would be determined by comparing the growth in treatment areas to growth in areas of control where no alkaline feedstock was applied to ultimately determine the net base cation loss resulting from new growth. In annual crops, the direct measurement of total base cation content in harvested biomass could be coupled with an estimate of the total biomass removed from the field. If a control plot is being used, then the requirement would be to subtract the difference in total base cations removed between the control and the deployment areas.

The group also briefly discussed the complications in total carbon accounting presented by the fact that while increased growth and standing biomass in perennial or forestry systems would lead to alkalinity losses, it would also lead to increased carbon stored as biomass. Although this increased C sequestration would not be credited in this v1 standard (discussed in more detail below), this scenario highlights the need to navigate between different crediting regimes (soil organic carbon, forest biomass) to avoid overcounting while providing a more holistic accounting of the overall carbon balance.

Feedback prompt: For suppliers: How do you currently estimate plant cation uptake for the crops that you are deploying on and are there any data sources that you rely on? Is there sufficient data on base cation uptake for the crops in your deployment area to justify not taking direct measurements? Do you have any additional feedback on quantifying plant uptake and its influence on net CDR?

2.2 Accounting for potential changes to SOC cycling

Status: **purple / recommendation under development**

The evidence base for how shifts in soil biogeochemistry and physical conditions due to enhanced rock weathering deployments can influence different aspects of soil organic carbon cycling (productivity, respiration, stabilization, or destabilization of existing organic carbon stocks) is just emerging. Observations thus far have demonstrated variable responses of soil organic carbon (SOC) to alkaline feedstock addition. Importantly, this discussion began from the default position that *increased* soil organic carbon will not be credited as additional carbon removal under the v1 ERW standard due to the different permanence profile and monitoring requirements for carbon stored as SOC.

The working group discussion revolved around two posited options for how to measure potential changes to soil organic carbon dynamics. Option A involved a "decision-tree" approach, in which a supplier could determine whether a substantial *destabilization* of, reduction in accumulation rate of, or decline in SOC stocks is likely to occur given the characteristics of the deployment. The second option, Option B, involved the required direct monitoring of potential organic carbon losses for all deployments.

The working group determined that there is not currently enough information regarding the underlying mechanistic drivers or interactions between SOC and ERW to confidently say that potential SOC destabilization is not a risk. The consensus opinion of the working group was that, at this stage of our understanding, some monitoring of soil organic carbon stocks is required in the near term. However, the group aligned that the magnitude of risk scales with the size of existing SOC stocks, which means that the standard could provide guidance against deploying commercially in areas that are considered “high risk” until further studies are completed (e.g., wetlands and peatlands).

The details of soil organic carbon monitoring requirements will be discussed in a working group <> market practitioner session and brought back to Working Group 2 for final decision-making in Session 4.

Feedback prompt:
How are you currently monitoring organic carbon dynamics in your field trials or mesocosm experiments? Do you have any comments or recommendations for the Working Group on monitoring deployment-scale changes to soil organic carbon stocks?

2.3 How to consider potential changes in other soil GHG emissions (CH_4, N_2O)

Status: **Blue / Recommendation**

As with changes to soil organic carbon, for the v1 standard, the default position is that decreased N_2O or CH_4 emissions would not be credited as additional carbon removal. The Working Group discussion thus focused on protecting against potential downside risks of non-negligible increases in other greenhouse gas emissions due to an ERW deployment.

The group discussed two endmember options:

- **Option A:** The v1 standard does not contain requirements for considering non-CO_2 greenhouse gas emissions.
    - Better understanding interactions between ERW interventions and soil N_2O emissions and CH_4 production in methanogenic systems is a key R&D priority but does not lead to additional requirements for this v1 quantification standard.
- **Option B**: Direct monitoring of soil GHG emissions is required for certain systems.

There was general agreement on the fact that suppliers should be incentivized to measure gas fluxes (for example, on research plots) and that we do not know enough to discount this as a nonissue. However, the Working Group reached a strong consensus in favor of Option A; when put to a vote, the vast majority of group members voted to not require measurement of other soil GHG emissions at this juncture.

Feedback prompt: Do you agree with the working group’s consensus that measuring changes in other soil GHGs should not be required in this v1 standard?

Working Group 3 - Characterizing deployment hydrology and looking forward

Scope of the meeting

The alkalinity generated by enhanced rock weathering and the associated CO_2 removed from the atmosphere as charge-balanced dissolved inorganic carbon (DIC) is not durably stored until it reaches a long-duration reservoir: in some cases a long residence time groundwater system, and in many cases the ocean. In this session, the 'downstream evasion' Working Group focused on requirements for characterizing the field and watershed-scale hydrology of individual deployments; how to consider the lower vadose zone and groundwater dynamics; returned to the question of monitoring requirements; and ended the session with ideation on articulating the requirements for models used to simulate downstream loss pathways.

Session 2 feedback themes from market participants:
Theme 1: In session 2, the Working Group identified two potential options for deriving context-specific conservative loss estimates for evasion from surface water systems. Option A involved developing a *global map* that visualizes and delineates the vulnerability of certain catchments to downstream losses, which would be created by the ERW community as a whole. This vulnerability index could then be tied to different magnitudes of conservative loss estimates. Option B involved the development of simplified models that would extrapolate from catchment areas that are currently data-rich, to estimate the magnitude of downstream evasion from a smaller subset of characteristics about a catchment. This would then allow for the development of downstream loss estimates from a much wider range of 'data-poor' regions. The market feedback was positive surrounding both options, with some calling for a combination of the two. No option was strongly preferred over the other in supplier responses.

- A small group will produce the work product that will be used to calculate a default context-specific conservative loss estimate for the v1 standard in a joint working group/market practitioner problem-solving session. This could include proposing an interim option and defining the study required to produce a more robust version for v2 of the quantification standard. The proposed solution will be returned to the Working Group for Session 4.

Theme 2: To quantify the magnitude of potential CO_2 evasion in the coastal ocean, the working group agreed in the last session to use *Option B as a requirement in the standard with Option A as a "reach"* or a "nice to have" that could potentially become a part of a future iteration of the standard. Option B involved a conservative assumption of evasion from carbonic acid system equilibration, which is derived by considering the thermodynamic storage efficiency as a worst-case scenario, assuming full equilibration at representative temperature, salinity, and current atmospheric pCO_2. Option A was a region-specific coastal biogeochemistry model that would be run, using alkalinity and DIC outputs from the process-based model used to estimate riverine evasion. The market feedback agreed with the working group's hypothesis regarding the two options.

Theme 3: In the last session, the working group agreed that carbonate burial and carbonic acid system equilibration should be lumped together when quantifying potential netCDR losses and that an "Option B" approach like the one in theme 2 should be followed. The market feedback agreed with the working group's hypothesis that a conservative assumption should be required in the standard, with a "reach" being a region-specific biogeochemistry model.

3.1 Deployment hydrology characterization: from field to groundwater

Status: **Purple / Recommendation under development**

The Working Group started with a discussion of the requirements for characterizing hydrology at both the field and watershed scale for any given deployment. The group began by discussing the question of what practitioners need to do to constrain the water balance of the deployment area and to understand/document fluid flow at these different scales. To constrain the water balance (irrigation + precipitation relative to evapotranspiration), suppliers would need to define the hydrologic boundaries of the watershed and monitor (or have data on) meteorological conditions of the deployment area and its surrounding environment.

In addition, the Working Group was aligned that it is important to understand groundwater flow and the distribution of groundwater residence times, particularly if suppliers are either trying to credit for long-duration storage in groundwater or if they are monitoring drainage waters (e.g., at a stream that the field or watershed is draining into). An understanding of the distribution of transit times between the point of application and the monitoring point is required to understand when the weathering flux is expected to arrive, and how to deconvolve the influence of the mixing of different groundwater sources or the mixing of waters with different transit times in the measured signal.

Feedback prompt: For suppliers: How do you currently consider water flow dynamics at deployment sites? Do you simply consider the overall water balance for net transport + modeling? Or are you needing to grapple with complicated shallow-surface flow paths that complicate both deep-soil aqueous phase measurements and monitoring of the drainage waters? How do you currently characterize the hydrology of a given deployment site?

3.2 Alkalinity loss along the flow path through the lower vadose zone and groundwater systems

Status: **purple / recommendation under development**

The Working Group then discussed how to think about potential alkalinity losses as weathering products travel through the lower vadose zone and along the groundwater flow path. The discussion started with the prompt that what we care about within the soil profile and in downstream surface water systems doesn't go away along the flow paths in the vadose zone and groundwater systems (specifically alkalinity sinks: sorption, secondary mineral precipitation, etc.). The net impact of those reactions thus matters for CDR quantification once the water mass returns to a surface water system.

The group discussed the potential utility of installing shallow groundwater wells at select sites for informing and validating models and debated whether this should be considered simply an R&D priority - appropriate for sentinel sites akin to critical zone observatories - or a requirement for commercial deployments. The group consensus was that this should *not* be a requirement for all deployments, but installing groundwater monitoring wells in a select subset of commercial deployments (e.g., where groundwater and/or catchment drainage monitoring is operationally feasible) could be highly beneficial from an R&D perspective. This could potentially be incentivized through additional funding specifically earmarked for this 'R&D grade MRV', or through research collaborations.

Feedback prompt: Do you have any feedback on characterizing water-rock interactions and potential alkalinity loss along lower vadose zone and groundwater flow paths? For suppliers, do you currently do any groundwater monitoring at your deployment sites? Do you have any feedback on the proposed model of providing separate R&D funding for groundwater monitoring systems to improve our understanding and modeling of deep critical zone processes that impact net CDR quantification?

3.3 Driving towards decisions on requirements for monitoring of downstream systems

Status: **Blue / Recommendation**

The Working Group also circled back to the question of whether suppliers should be required to do any direct monitoring of surface water systems in a specific 'zone of impact' downstream of their deployment area. The group debated the potential utility of this type of monitoring, identifying two key purposes: to simply demonstrate that there is a weathering signal appearing downstream, and for model validation. The group was largely aligned that at the scale of current deployments, most would not be able to detect a weathering signal against baseline variability, and thus downstream monitoring should not be required. It was discussed that there could be utility in doing pre- and post-deployment monitoring of any zero-order streams within a deployment site, but downstream monitoring should not be required as a general rule given current deployment volumes.

Feedback prompt: Do you have any final feedback on monitoring downstream surface water systems in the context of the v1 quantification standard?

3.4 Returning to the question of 'sufficiently rigorous' catchment-scale models and sharpening requirements for models used in downstream loss assessment

Status: **Purple / Recommendation under development**

Last session, the Working Group discussed options for arriving at a context-dependent empirical loss estimate for CO_2 evasion in riverine/surface water systems. In this session, the Working Group returned to the question of requirements for models used to characterize downstream evasion, if default context-dependent empirical loss estimates are not used. What does the standard need to constrain about the empirical or process-based models used to ensure rigorous quantification?

The group discussed the utility of having benchmarked models that have been developed and/or validated by the academic community (see the Working Group 4 discussion for more details about potential benchmarking frameworks). Until such community benchmarked models and a process for rigorously validating new models exists, the group discussed what types of guardrails should be in place for the v1 standard. The group consensus was that you 'can't just use any old process-based model', and that requiring peer review of models could be reasonable but not sufficient (the group is also cognizant of the timescales involved with the peer review process and limitations there). The group recommended that the v1 standard also include a suite of basic requirements, and that model use be critically assessed as part of the standard compliance / credit verification process. Potential requirements could include:

- Reporting: documentation of model function and uncertainty characterization
- Demonstration of ability to reconstruct background fluctuations (e.g., spatial and temporal fluctuations in carbonate system parameters)
- For process-based models, a minimum suite of processes that must be considered (and associated functional requirements - e.g., must track major ions and carbonate system parameters)

These requirements will be further discussed and refined in a joint working group <> market practitioner problem-solving session, and brought back to the working group for final decision making in Session 4.

Feedback prompt: Do you have any feedback or suggestions on requirements for models used in downstream evasion assessment? Is there anything you would specifically like to see in the reporting requirements?

Working Group 4 - Benchmarking and validation of enhanced weathering models

Scope of the meeting

The third session of this working group dove into the details of what an enhanced weathering model validation and benchmarking system could look like. The working group discussed how a model's imperfect predictive skill and structural uncertainty interfaces with typical methods of model uncertainty quantification. This session was underpinned by the definitions of model validation and uncertainty quantification as follows:

- Model validation - assessing a model's predictive skill through comparison with empirical field data. Goals include avoiding overfitting / over-tuning and determining over what parameter space the model is sufficiently validated.
- Uncertainty quantification - determining error bars for the particular model output variable that is being modeled. For a process-based model, this often takes the form of an aggregate sensitivity analysis, using Monte Carlo simulations over parameter distributions to derive an output.

Session 2 feedback themes from market participants:
Theme 1: The working group was asked in session 2 what the v1 standard might say about the use of RTMs in an MRV framework. The working group put forth a set of proposed potential requirements for the use of RTMs in the v1 standard such as: requiring a standardized format for supplier reporting of their RTM processes, not using RTMs as an end-to-end net CDR quantification tool without empirical constraints, shared sensitivity analyses, and assumed prior distributions of parameters, among other requirements. The market participants largely agreed with the proposed set of requirements that the working group produced.

Theme 2: In session 2, the working group articulated a list of parameters that could potentially be required in sensitivity analyses and used to calculate model parametric uncertainty. The market feedback participants held a general perspective (with a few exceptions) that the presented parameter list and set of parametric uncertainty expectations would be acceptable. No market feedback participants indicated how they estimated parametric uncertainty or if they ran sensitivity analyses on their models.

4.1 Model benchmarking and validation

Status: **purple / recommendation under development**

The working group discussed at length the potential to develop a benchmarking system for ERW models, based on comparing model output to high-resolution ERW datasets shared across the community. This benchmarking system might be inspired by existing soil RTM benchmarks (e.g., see here), but likely requires something new outside of this, given the need for new modeling capabilities to handle the enhanced weathering context (e.g., variable particle size tracking). The group proposed that different datasets should be collected in distinct regions to test predictive skill across heterogeneous soil conditions.

This benchmarking discussion also included some thinking about how we could draw analogies to CMIP—including the use of intercomparison to ensure outputs produced by different models lie in similar ranges, and hindcasting exercises, where a suite of climate models are each used to simulate the period from 1850 to the present day, and predictive skill is evaluated through comparison of model output to historical data.

Working group members also discussed how to determine whether a model has sufficient predictive skill over specific parameter ranges (operating under the assumption that we have perfect knowledge of input parameters). One point of view brought up in the discussion was that most deployments would initially be in working/managed lands that have generally similar parameter ranges, so baseline conditions may not affect model predictive skill very much. It was pointed out that the parameters that usually can throw off a model such as salinity, extreme temperatures, and extreme pressures, are not present in most managed lands scenarios that are being modeled. However, the agricultural soil systems being modeled are also definitely not in steady-state due to years of tampering from human activities, creating the potential for variation in predictive skill and meaning that the details of model spin-up likely need to be included in a benchmarking system.

Feedback prompt: What do you think the elements of a successful enhanced weathering model benchmarking system could look like? Would you be interested in participating in such an initiative or helping co-develop it?

4.2 Assessing model goodness-of-fit and avoiding over-tuning / overfitting

Status: **Purple / Recommendation under development**

The group first agreed that using a root mean squared value is usually what is reported for goodness-of-fit, but acknowledged a desire and willingness to improve its statistical practice. The group plans to investigate distance metrics available to compare field data to model data before reconvening in Session 4.

To prevent the over-tuning of RTMs and "distribution gaming", the group proposed a few ideas. First, using separate datasets for model tuning and model validation is critical. Second, having a clear articulation of which parameters are being tuned and which are initialized to a particular value needs to be shared with a verifier. Third, a benchmarking system could provide fixed parameter values, rather than letting modelers tune their own models, preventing any distribution gaming. Lastly, transparency over the process used to tune a model could allow an external verifier to determine the extent to which over-tuning has occurred.

For a verifier to evaluate the predictive skill of a supplier's model, the supplier should participate in the community benchmarking system described above, showing the fit between model outputs and community datasets, as well as a list of the parameters used and how each parameter was used in the model run (i.e., which for tuning, which for initializing, etc.).

Feedback prompt: What other statistical frameworks have you used to assess model goodness-of-fit? How would you recommend that a system check against overfitting? What are we missing here?