ISOMETRIC

Model and Measurement Based Quantification of Enhanced Weathering 1.0 Module

Public Consultation Summary

Published on 9th October, 2026

Context

Isometric held a public consultation on its Model and Measurement Based Quantification of Enhanced Weathering Module 1.0 Module to receive stakeholder input on this Module.

The public consultation was announced on the 8th of June, 2026. The period of consultation lasted 30 days, with the final day as the 8th of July, 2026.

After the initial public consultation, the feedback received was considered for incorporation into the Module. All stakeholders have received responses to the submitted feedback.

This document summarizes the feedback received during the public consultation and the revisions included as a result of the comments. Content in italics and brackets are excerpts from the public consultation version of the protocol to give the reader necessary context behind the comment.

We thank all participants for their time.

Summary of feedback received

Model and Measurement Based Quantification of Enhanced Weathering Module 1.0

Theme	Section	Comment	Resolution
Model framework philosophy: catching wrong answers; mix-and-match risk	1.0 Introduction	We strongly support the framing that this Module supplements rather than supplants measurement: the requirement that direct field measurements constrain and continuously validate any model used for quantification is the right posture for the current maturity of the field. We also appreciate the candor of stating that the bar is set deliberately high and that it may be difficult for any existing model to meet it today, which sets an honest expectation for developers and the field as a whole.	No change. Thank you for the positive feedback on these sections, which are helpful for knowing what to retain.
	1.0 Introduction	*[The application of models as a quantification tool for EW is nascent, and technical understanding of model performance and data requirements continues to evolve.]* We agree, and think acknowledging this reality is necessary. At a high level, we understand this module as trying to set a bar where models can produce CO2e_export,RP estimates (and associated uncertainty) of similar quality to the measurement-based approach. That bar, first and foremost, requires that models can convert input data into reliable CDR estimates. Critically, it doesn't say much about the models themselves — if you get reliable answers, omitting key processes and nonlinearities isn't a	No change. Thank you for the context, your framing and understanding of the Module's intent is aligned with our intention and we appreciate your feedback through this lens.

		dealbreaker. The key problem the module therefore needs to address is not how to get the right answer for the right reasons, but how to tell when an answer is right versus wrong. This is quite different from the more fundamental challenge of developing models that can help advance our scientific understanding of how, where, and when EW works. Our feedback focuses on whether the module is equipped to catch wrong, model-based answers. We don't focus on how the models represent physical reality unless shortcomings can contribute to getting the wrong answer without knowing it. Still, this criteria is hard to judge since the evidence base around key modeling decisions is currently thin. Our feedback focuses on just a few high-level points where we think wrong answers could be especially likely to slip through the safeguards required by the module.	
	3.0 Quantification of CO2_stored	*[Alternatively, separate models or model components may be used to predict the individual terms in Equation 2, which are then combined to calculate CO2eexport,RPCO2e{export, RP}CO2eexport,RP.* *If separate models are used to quantify individual terms, each model must individually meet the relevant validation requirements described for the models used during quantification under this Module.* *Project Proponents may also elect to individually model some terms and empirically quantify others through direct measurement, following the requirement for that term outlined in the relevant section of the Enhanced Weathering in Agriculture Protocol.]* Treating model components as separate, and allowing practitioners to mix and match data and model approaches, risks two bad outcomes. First, it means your answer is more likely to be built on the wrong reasons, since the individual terms don't have to be internally consistent if the models are built narrowly enough to make the comparison impossible. This raises the risk that you get the wrong answer. Second, it opens the door for practitioners to mix and match model and data approaches to get the best CDR	Changed. Thank you for the feedback. We've amended this section to remove the optionality for mixing and matching model and measurement approaches for quantifying each term in Eq. 2, to mitigate against the concern for cherry-picking. However we retained the option to quantify plant uptake loss using measurements and subtracting it from the modeled output as that would be a conservative approach. In addition, we've removed the option to use independent models to separately calculate each term in Equation 2. We instead require that a single model is used to quantify CO_2 export to maintain a level of internal consistency, although we note that individual processes may still be represented by distinct components within a single coupled model.

		outcome, or at least puts substantial pressure on Isometric and verifiers to identify and prevent such cherry picking.	
Framing in Protocol	1.0 Introduction	We thank Isometric for the opportunity to review and provide feedback through the public consultation process. We appreciate the thoughtfulness this protocol reflects in considering a path for model inclusion in quantification. At the protocol level, we recommend the Module state, and be designed toward, a near-term hierarchy: measurement for primary quantification, data-driven models as the more viable model class where a model is used to supplement empirical measurements (and with strengthened transparency), and process-based models confined to measurement support and time-accounting in the short- to medium-term. We believe that process-based models are much further off from defensible use in crediting than data-driven model or measurement-only approaches, due to issues with model identifiability and structural uncertainty.	No change. Thank you for the suggestion. We do not think it is necessary to explicitly state a hierarchy in the Module or Protocol. The purpose of including alternative quantification options is to allow other options, which are viable as long as the requirements for a particular type of model are met.
Independent expert panel: composition and when required	1.0 Introduction	*[Where proprietary models are used, this Module requires additional independent review by an Expert Panel of qualified third parties.]* This requires some clarification on what is considered proprietary -- for example, what if you have added functionality to an open-source model? Would this require review by an expert panel?	Changed. Thank you for the comments. We have revised Section 10.3 to address these points together. First, the trigger is now defined by the level of data sharing, see Section 10.3.1 for more details. Second, we have defined a structural revision (which triggers re-review at later verifications) as a change to model structure. For data-driven models, this includes the architecture or the type/composition of training data; for process-based models, this includes the represented processes, rate laws, or databases. Third, on composition: we have clarified that the Independent Expert Reviewer must have expertise in both the modeling approach and ERW field science, and multiple experts may serve this role if it cannot be filled by an individual. We also recognize the concern that a small panel cannot exhaustively interrogate every model, however, public release does not guarantee that scrutiny either. We therefore see the Expert Panel as a pragmatic
	10.3.2 Independent Expert Panel Review	Recommend defining clear, specific criteria for when Expert Panel review is required, suggest it apply to all proprietary models and proprietary calibrations of open-source models.	
	10.3.2 Independent Expert Panel Review	A clearer definition of what constitutes a 'structural revision' that would trigger re-review by the Expert Panel would be helpful.	
	10.3.2 Independent Expert	*[Panel composition must include at least one expert in the modeling approach used (process-based, data-driven, or hybrid) and at least one expert in ERW field science.]*	

	Panel Review	Can this expert be the same person?	compromise that secures dedicated, targeted review within what is feasible, rather than as a necessarily weaker form of scrutiny than open release, which in practice often delivers only the potential for review rather than the fact of it.
	10.3.2 Independent Expert Panel Review	*[Isometric will select the members of the expert panel for each project requiring an Independent Expert Panel.]* Is there a set number of panel members? If so, what is it?	
	10.3.2 Independent Expert Panel Review	Two concerns. First, diligence capacity: it is unclear how a small panel can perform meaningful diligence on model codes, across potentially many independently developed models, that take years to understand fully and are early in their development cycles; a panel of one modeler plus one field scientist is thin for interrogating a full reactive-transport codebase, and this is not a knock on anyone's technical prowess but an acknowledgment that dragons hide in the details of each model. Second, confidential review does not deliver the independent replication that public release enables, so the proprietary pathway remains a weaker form of scrutiny than the open-source pathway the Module otherwise recommends.	
Public disclosure scope (training data, control datasets, IP)	1.0 Introduction	*[To build community trust in models and encourage knowledge sharing across the field, this Module requires public disclosure of the following, regardless of model type]* Sections below explicitly require sharing of training data, but it's not required here. Suggest editing for consistency. Additionally, while we agree with the need for transparency to build trust in model-based MRV, we also want to ensure that protections are built in for business-sensitive information.	Changed. Thank you for all the feedback around data sharing and transparency. We have revised the disclosure requirements to balance transparency against protection of supplier IP, reflecting feedback on both sides of this question. Monitoring data remains fully public. As under the EW 1.2 Protocol, the underlying treatment and control monitoring datasets, including sampling dates, depths, compositing metadata, and analytical methods, are publicly disclosed, not only derived validation statistics. This is unchanged. For model-related disclosure, we have introduced flexibility. In addition to the baseline that applies to all models, a plain-language model description, summary validation statistics, and the model outputs used for crediting, the Project Proponent must publicly disclose one of: (a) the model code, or (b) all auxiliary training, calibration, and
	10.2 Reporting	Suggest that the model update report required under Section 10.2 be included in the public-facing verification report.	
	10.3 Transparency and Public Disclosure	We strongly welcome the strong data transparency provisions in this Module: requiring public disclosure of a plain-language model description, all training/calibration/validation datasets, model inputs and outputs including the full Monte Carlo distribution, and summary validation statistics is a strong posture that is essential to building	

		field-wide confidence in model-based results.	input data. This allows a Project Proponent to demonstrate that the model meets the Module's requirements and was not tuned to produce a particular result, without mandating full public release of a dataset that others could use to reconstruct a proprietary model. Proprietary models remain subject to independent Expert Panel review. The model update report is now included in the public-facing verification report. We have also aligned the Introduction with the Transparency and Public Disclosure section so that the data sharing requirements are all stated consistently.
	10.3 Transparency and public disclosure	We have reservations about remaining secrecy surrounding the actual model codes, despite the strong provisions for data transparency found in this Protocol. In our view, publicly available models will need to be validated against publicly available datasets and demonstrate reasonable similarity in CDR predictions before they will be accepted by buyers, scientists, and regulatory bodies.	
	10.3 Transparency and public disclosure	If a model's scientific architecture cannot be sufficiently described in public, peer-reviewed literature such that an independent expert can assess whether the validation evidence is consistent with how the model works, it should not be used for commercial crediting; there should be a clear justification and an extremely high bar for departing from that norm. While the Expert Panel idea might indeed subject these models to more scrutiny than they might otherwise receive if lost in the piles of peer-reviewed literature, there is still a major public perception risk associated with using the closed-door Expert Panel that may make credits difficult to sell in the VCM. ERW is being considered within the EU CRCF process and is not yet in IPCC National GHG Inventory Methodologies; governments are unlikely to accept the political and legal liability of backstopping methodologies reliant on proprietary quantification approaches, so direct-measurement-based evidence will be critical to long-term support.	
	10.3.1 Public Disclosure Requirements	*[All training, calibration, and validation datasets, or, where data sharing is restricted by third-party agreements, documented procedures for accessing the data]* We suggest clarifying that public disclosure should include the underlying treatment and control monitoring datasets used to support model validation and control correction, not only derived validation statistics or processed model inputs. For ERW, model confidence depends strongly on whether untreated/control plots show comparable movement in	

		tracers, cations, pH, CEC/base saturation, or other soil measurements over the same time period. Public access to raw or minimally processed control and treatment data, with sampling dates, depths, compositing metadata, and analytical methods, would materially improve confidence that modeled removals are not being driven by background drift or sampling artifacts.	
	10.3.1 Public Disclosure Requirements	Noted an inconsistency between the introduction (which requires public disclosure of model training and calibration code for all models) and Section 10.3.1 (which omits it), and suggested adding it to 10.3.1.	
	10.3.1 Public Disclosure Requirements	Currently, the module requires that calibration, validation, and training data are shared publicly. While we understand that this requirement was written in the spirit of transparency, it is a major risk to supplier IP, as others could train their own models off of that data if it is completely open source. The goal of model transparency is to demonstrate that the model a) meets the requirements of the module, b) can successfully perform the task it was designed to perform, and c) was not manipulated to yield a particular result. None of those goals require full public disclosure of the calibration, validation, and training datasets. We propose the following middle ground to promote transparency while protecting supplier IP: 1. Publicly disclose the model structure, configuration, and calibration/validation process. 2. Provide Isometric, the VVB, and the external committee (if necessary) with the training code, architecture code, and the training set/validation set. Then Isometric and the verifier can replicate the training process easily without requiring the full training and validation dataset to be public. We view the role of the registry, the VVB, and external committee in this regard as helpful to strike the right balance between rigor, transparency, and business practicality.	
	10.0 Reporting	*[Reporting]* The layout of this section is great -- super helpful and clear!	No change. Thank you for the feedback, we have retained the layout of this section.

Timeline for incorporating the Module into the ERW Protocol	1.0 Introduction	*[Introduction]* Under what timeline will this modular be incorporated into the ERW protocol, as the ERW protocol v1.2 is not currently modular in a way that would allow the model and measurement module to slot in.	No change. Thank you for the question. Upon certification of this Module, the EW protocol 1.2 will undergo restructuring which will not materially affect the content and requirements for the Measurement-based approach, and the new Model and Measurements Module will be included as an alternative option. This restructuring will occur during one of Isometric's bi-annual Protocol update weeks, which takes place in the spring and fall of each year.
Applicability envelope: climate/weather extremes	2.0 Applicability	*[The extremes of required meteorological parameters (temperature and precipitation) experienced in the Reporting Period must be within the validation envelope.]* Does this mean P and T considered jointly, or separately? What time period are these extremes being calculated over? Weekly, monthly? Since drought is possibly cumulative, how will you define an extreme drought? (e.g. if in one year you have 0 mm of precip in one month and then five wet months, but in another year you have six straight months of 1 mm of precip in each month, does the second year fall within the validation envelope?) Lots to unpack here.	Changed. Thank you for the comments. We agree that assessing temperature and precipitation extremes in isolation is too coarse to bound applicability and have revised this section to add more clarity. Temperature and precipitation must be assessed jointly, so that a combination with no analogue in the training data is identified even when each variable is individually within its range. Regarding time period, the assessment must be performed at a temporal resolution fine enough to resolve the seasonal cycle and sustained sub-seasonal anomalies, with monthly resolution recommended, and temperature is represented by its mean and extremes at that resolution. On the definition of an extreme and the treatment of drought, precipitation is assessed on a cumulative basis over the Reporting Period and relevant sub-periods rather than as an instantaneous value, so that sustained, duration-dependent conditions such as drought are captured. We retained the 12-month requirement for the minimum validation window, as requiring multiple years before crediting to capture inter-annual variability such as ENSO is not practical. Inter-annual variability is handled through the validated envelope requirement. A Reporting Period whose meteorological conditions fall outside the validated envelope triggers the out-of-envelope provisions. Anomalous years, including drought and pluvial years, are therefore
	2.0 Applicability	The 12-month minimum validation window captures one seasonal cycle but does not guarantee coverage of inter-annual variability (ENSO swings, drought/pluvial years). The validated envelope is defined only on temperature and precipitation extremes. Weathering responds nonlinearly to seasonal water balance (not precipitation alone) and to event structure; this should be framed in terms of seasonal water balance to include ET. Two sites with identical precipitation extremes can differ substantially in effective infiltration, so the envelope as defined is likely too coarse to bound applicability.	

			caught and handled by this mechanism.
	2.0 Applicability	*[Models must be developed and validated using data representing the full range of operational conditions experienced in the project area during the Reporting Period.]* What happens if models are validated using a year's worth of data, representing conditions from all seasons of that year, and in the next RP some extreme weather event that is outside of the training conditions occurs (e.g. heat wave, rainfall event, drought, etc)? Can you clarify if there are additional requirements for events like this?	No change. Thank you for the question. This scenario is addressed by the existing provisions in Section 2.0. Where forcing conditions during a Reporting Period exceed the validated envelope (including a novel extreme event such as a heat wave, heavy rainfall, or drought occurring in a later Reporting Period), the Project Proponent must either (a) provide a documented, process-based justification, supported by a sensitivity analysis, that model performance is not materially affected, or (b) exclude the affected period from crediting, re-initialize the model using field measurements collected after the excluded period, and re-validate the model before crediting resumes. Because re-validation is required at every Reporting Period using the latest data, conditions that were initially outside the validated envelope are incorporated into it once observed, so the model progressively covers the range of conditions the project experiences.
	2.0 Applicability	We endorse the requirement that, where a period is excluded from crediting, the model be re-initialized from direct field measurements of the state variables that may have changed (remaining feedstock mass, particle size distribution, soil cation inventory, and base saturation) and then re-validated before crediting resumes. This is a thoughtful approach that closes what would otherwise be a serious gap in how a model restarts after an out-of-envelope period.	No change. Thank you for the positive feedback, which is helpful for knowing what to retain.
Ion exchange and alkalinity (charge-balance) representation	3.0 Quantification of CO2_stored	*[CO2eNetSorption,RP]* While higher CEC values can retard the transport of Ca^{2+} and Mg^{2+} released by enhanced weathering, the implications of cation exchange for alkalinity export require further consideration. Adsorption of Ca^{2+} and Mg^{2+} onto exchange sites is accompanied by the release of equivalent charges of exchangeable cations such as Na^{+} and K^{+}. Therefore, alkalinity export from the NFZ should not necessarily be evaluated solely from Ca	No change. Thank you for your comment. This statement is applicable to field measurements where HCO_3^- or alkalinity is not measured directly. In such cases, bicarbonate can be estimated from the charge imbalance between measured cations and anions. However, this approach is unnecessary for ERW reactive transport models, as these models explicitly enforce mass balance and electroneutrality, resulting in no actual charge imbalance in the

		and Mg fluxes. Instead, it may be estimated from the charge balance between total cations and a sum of anions in leaching soil solution, with the imbalance representing leached bicarbonate. Under this approach, the effect of ion-exchange-induced retardation is reduced because charge is conserved during cation exchange. Consequently, both this Module and the Enhanced Weathering in Agriculture Protocol should consider alkalinity accounting approaches that explicitly incorporate ion-exchange processes and charge balance.	simulated aqueous chemistry. Instead, bicarbonate is calculated directly as part of the model's aqueous speciation and is available as a predicted output.
	5.3.2 Processes for Quantifying CDR Terms	*[Models must track the source of protons. This includes protons released from the pre-existing exchangeable acidity pool on soil mineral and organic surfaces (e.g., exchangeable H+ and Al3+) during displacement by weathering-derived base cations.]* I can't understand how this is a "must" but a cation exchange model is a "should". What could a model look like that does this proton accounting but doesn't include an exchange model?	Changed. Thank you for the feedback. Our intent is that the model must explicitly track the sources of protons, regardless of whether ion exchange is included. If nitrification is simulated, the protons generated by this process should be excluded from the CDR calculation, as they do not contribute to carbon dioxide removal. Regardless of which processes are represented, the model itself must maintain both mass balance and charge balance. We have revised this section to clarify these requirements.
	5.3.2 Processes for Quantifying CDR Terms	*[The soil solid phase, i.e. clay minerals and organic matter, and Fe/Al oxide coatings, acts as a capacitive buffer for base cations and acidic cations, complicating the relationship between mineral dissolution and alkalinity export]* Ion exchange only occurs on (pseudo-)permanently charged surfaces (e.g., clays, carboxylic acid functional groups) whereas the other flavor of sorption ie., surface complexation occurs on variably charged amphoteric surfaces (e.g., Fe/Al oxides). This distinction should be made here.	Changed. Thank you for your comment. It is true that permanent structural charge is the classical source of electrostatic ion exchange, whereas variable (pH-dependent) charge on amphoteric oxide surfaces is the classical setting for surface complexation. However, we believe the statement is somewhat overstated. Ion exchange is not exclusive to permanently charged surfaces. For example, carboxylic acid functional groups on soil organic matter are variably charged (i.e., pH-dependent) rather than permanently charged, becoming deprotonated as pH increases. These carboxyl groups can participate in both ion exchange and surface complexation. Likewise, clay minerals are not restricted to ion exchange alone and may also participate in surface complexation through pH-dependent edge sites.

			We include ion exchange as a required process and surface complexation as a recommended process and have revised accordingly in the Module.
	5.3.2 Processes for Quantifying CDR Terms	*[The following additional terms are not required to be quantified but are highly recommended for completeness in implementing kinetic rate laws for mineral dissolution.]* Suggest adding a clarification here that these are recommended if sufficient data exists to constrain them. Otherwise you can get into equifinality problems.	Changed. We agree with the reviewer's comment. And it should be noted that BET surface area is not the same as reactive surface area (RSA). BET surface area can be measured but reactive surface area can only be calibrated. We revised this section accordingly as shown below.
	5.3.2 Processes for Quantifying CDR Terms	*[Where ion exchange is represented in the model, the following processes are required]* Depending on the soil organic matter content, a substantial proportion of the cation exchange capacity (CEC) may be derived from soil organic carbon. Because the charge characteristics of soil organic matter are variable and strongly dependent on pH, the associated ion-exchange capacity can change as soil chemical conditions evolve during enhanced weathering. Consequently, it is not clear whether fixed values of CEC and selectivity coefficients can be assumed a priori. Their applicability should first be evaluated in relation to soil organic carbon content and through sensitivity analyses to determine the extent to which model predictions depend on these parameters.	Changed. Thank you for the comment, we agree that CEC is not a constant and changes over time. For highly weathered soils (e.g., Oxisols/Ultisols in southeast Brazil), literature indicates CEC can increase roughly up to 70% within 12 months after applying feedstock (Anda et al., 2015, Gillman 1980). For less weathered soils, Boniao et al. (2002) indicates no significant changes in soil CEC due to basalt treatment in acid inceptisols within 9-month incubation. It is reasonable to treat CEC as a constant for sandy loam (typical soil in southeast Brazil farmland) at field rate, but justification should be provided. We recommend applying a pH-dependent correction, e.g., consider a modest CEC(pH) function rather than a hard constant. This would help to capture the main effect without full surface-complexation modelling. We have revised this section in the module accordingly. Anda, M., Shamshuddin, J., and Fauziah, C.I. (2015). Improving chemical properties of a highly weathered soil using finely ground basalt rocks. Catena, 124: 147–161. https://doi.org/10.1016/j.catena.2014.09.012 Boniao, R.D., Shamshuddin, J., Van Ranst, E., Zauyah, S., and Syed Omar, S.R. (2002). Changes in chemical

			properties and growth of corn in volcanic soils treated with peat, ground basalt pyroclastic, and calcium silicate. Communications in Soil Science and Plant Analysis, 33(7–8): 1219–1233. Referenced directly in Anda et al. (2015)'s reference list. Gillman, G.P. (1980). The effect of crushed basalt scoria on the cation exchange properties of a highly weathered soil. Soil Science Society of America Journal, 44(3): 465–468. https://doi.org/10.2136/sssaj1980.03615995004400030005x
	5.6 Summary of Required and Recommended Processes	*[*Ion exchange]* Why is this a recommended only process? In a scenario where crediting is solely based on process-based modeling, this would be the only way to quantify CO2e_NetSorption. If NFZ CDR is being modeled, this needs to be required.	Changed. Thank you for your comment, this has been fixed in the Module, we agree that sorption is one of the 1st principle processes. A research in 2024 Frontiers study (Hasemer et al., 2024) on measuring enhanced weathering found that soil retains most enhanced weathering products through cation exchange reactions, and CDR estimated by cations is often greater than CDR estimated by inorganic carbon. Surface complexation matters, but not in the primary CDR accounting pathway So we decided to set ion exchange as a required process but surface complexation as a recommended process. Reference provided here Hasemer, H., Borevitz, J., and Buss, W. (2024). Measuring enhanced weathering: inorganic carbon-based approaches may be required to complement cation-based approaches. Frontiers in Climate, 6:1352825. https://doi.org/10.3389/fclim.2024.1352825
Direct mineralogical measurement (XRPD) for application rate and dissolution	3.0 Quantification of CO2_stored	*[CO2eWeathering,RP]* Direct quantitative measurement of feedstock minerals in soils at the beginning and end of the Reporting Period, e.g., using quantitative X-ray powder diffraction (XRPD) method (Kurokawa et al. 2024, Soil Science Society of America Journal,	No change. We thank the reviewer for sharing this paper. We will continue to require measurement on an ICP, which is the only analytical technique with low enough detection limits to measure tracer concentrations for EW quantification. We will continue to monitor scientific literature on XRPD

		https://doi.org/10.1002/saj2.20757), would provide an independent evidence for validating CO2eWeathering,RP. This is important because base cations used in mass balance approaches may also come from external sources, including fertilizers, irrigation water, volcanic ash, and aeolian dust. Recent field evidence shows that quantitative XRPD can directly detect decreases in feedstock minerals (i.e., basaltic minerals such as plagioclase and pyroxene) in cropland soils after applying basalt powder (Kurokawa et al. 2026, Applied Geochemistry, https://doi.org/10.1016/j.apgeochem.2026.106869). Therefore, the Module should recognize direct mineralogical measurement as a useful complementary approach for validating feedstock dissolution.	and similar mineralogical techniques to potentially add them to the list of approved approaches in future versions of this protocol.
	4.1.1 Monitoring Area	*[Project Proponents must determine the alkalinity application rate over the treatment area from both operational logs and direct measurement, in accordance with Section 10.4.4 of the Enhanced Weathering in Agriculture Protocol.]* According to the Protocol, several methods are recommended for determining the application rate. Direct measurement of feedstock minerals in soils, e.g., using quantitative X-ray powder diffraction (XRPD) method (Kurokawa et al. 2024, Soil Science Society of America Journal, https://doi.org/10.1002/saj2.20757), would provide an independent assessment of feedstock-derived mineral abundance and spatial distribution in soils after application and ploughing. Recent field investigation showed that feedstock mineral content determined by XRPD in the plots applied with 10 wt.% basalt powder after application and ploughing were consistent with the values expected from mixing 10 wt.% basalt powder with 90 wt.% control soil (Kurokawa et al. 2026, Applied Geochemistry, https://doi.org/10.1016/j.apgeochem.2026.106869). Therefore, the Module should recognize quantitative XRPD as a useful complementary approach for independently assessing	

		feedstock-derived mineral abundance and spatial uniformity.	
	5.3.2 Processes for Quantifying CDR Terms	*[Models must be initialized with site-specific baseline soil mineralogical conditions because baseline soil mineralogy controls the initial porewater chemistry and the baseline saturation state of the system, particularly the SI of calcite, dolomite, and secondary silicates]* Quantitative mineralogical information of both feedstocks (i.e., rocks and minerals) and baseline soils is an important factor controlling the efficiency of enhanced rock weathering (ERW). However, accurate quantitative mineralogical analysis of soils is not always achieved by standard laboratory X-ray diffraction analysis. Recent investigation into the reliability of X-ray powder diffraction (XRPD) showed that state-of-the-art full pattern fitting method can accurately quantify minerals in soils when appropriate reference library for the fitting analysis and appropriate analytical procedures are used (Kurokawa et al. 2024, Soil Science Society of America Journal, https://doi.org/10.1002/saj2.20757). Therefore, the Module should clarify that quantitative mineralogical analysis used in the Module should follow published and validated methods.	
	5.6 Summary of Required and Recommended Processes	*[Summary of Required and Recommended Processes]* The current module does not explicitly require verification that applied Ca- and Mg-bearing minerals are dissolving in situ. This is a key limitation because common validation metrics, such as water chemistry and alkalinity export, are indirect and may also reflect background soil processes or agricultural inputs. We recommend including mineralogical observations (e.g., quantitative XRPD) in Section 4.6. XRPD provides direct evidence of feedstock mineral dissolution and can strengthen attribution of carbon removal to enhanced weathering. Recent field studies have demonstrated the ability of XRPD to quantify depletion of basalt-derived minerals in agricultural	

		soils (Kurokawa et al., 2024; Kurokawa et al., 2026). At high application rates, XRPD should be encouraged to support quantification of mineral weathering and CDR. At lower application rates, it should at minimum verify that dissolution of applied minerals is occurring.	
NFZ definition, depth, and depth-resolved sampling	3.0 Quantification of CO2_stored	*[is defined as the depth of deepest soil sampling or the depth of porewater sampling.]* Why this doesn't align with ERW v1.2, which recommends 20cm unless shallower sampling can be justified.	Changed. We have clarified that the depth of deepest soil or porewater sampling is recommended to be 20 cm unless otherwise justified, as stated in the Enhanced Weathering Protocol. We will not require Project Proponents to partition soil samples to multiple depths, but we acknowledge that this practice is useful to represent vertical gradients, as a commenter notes.
	4.1 In-Field Monitoring Approach	*[Samples must be taken to the depth of the NFZ as defined in the model.]* As to this sentence, the Module does not appear to provide clear guidance on how the NFZ depth should be determined. I encourage Isometric to provide clearer guidance on the definition of the NFZ and on the minimum vertical sampling strategy. In addition, the process-based model requirements specify that models should represent vertical gradients within the NFZ, which suggests that depth-resolved characterization may be necessary. However, the monitoring requirements do not clarify whether soil samples should be collected by depth intervals. From a geochemical perspective, the dominant processes are likely to differ between the rock-amended mixing layer and the underlying soil layer. In the amended layer, mineral dissolution kinetics are expected to dominate, whereas below the mixing layer, equilibrium reactions and solute transport become increasingly important. This suggests that, at a minimum, separate sampling of the amended layer and the underlying soil should be considered.	
Monitoring requirements	4 Monitoring	While we strongly support the continued requirement for soil measurements, there are a few aspects of the monitoring requirements as written that limit flexibility for project developers.	Changed. Thank you for this feedback. We have updated the module to add flexibility in the deployment area and validation

	Require ments	Firstly, the monitoring requirements implicitly mandate use of a 3-plot model, as described in Section 9.1. The 3-plot model is not easy to implement everywhere and may not actually be the best way to collect requisite data in many cases. Requiring it limits our ability to tune our project design to our deployment context. Adding more flexibility in project design is also important for cost, as it enables us to focus our resources more efficiently. We are asking you to include language that is agnostic of project design as long as the requirements for control monitoring are met, allowing for more flexibility in how field resources can be allocated. Secondly, we disagree with the mandated minimum sampling density in the deployment plots. While we recognize that high-density measurements are, in many cases, required to capture variance in key soil parameters, this variability is highly site-specific, and the stated minimum (1 sample/10 ha) is unlikely to be appropriate everywhere. If a site requires fewer samples to characterize the variance, this is an unnecessary cost that the project developer must bear; conversely, the required minimum may give the false impression that this density is enough to guarantee acceptance, which will not be the case in an area where more samples are necessary. We recommend that, rather than having a set minimum, the deployment site sampling density is determined based on underlying soil variance, which is how the sampling density is determined in the main protocol and for the monitoring site.	check to allow for more flexibility in project design. We maintain an overall minimum sampling density of 1/10 ha in the deployment area, however, we have added more guidance on the breakdown between deployment area sampling for the validation check and for agronomic monitoring for matching criteria. This includes flexibility for the project proponent to aggregate deployment area samples in a small area to perform the validation check, so long as spatial coverage is maintained for agronomic variable monitoring. With that, we have clarified that Project Proponents must have statistical power in the samples used in the deployment area validation check, but not in the agronomic monitoring samples.
Validation targets: direct measurables vs CDR; is porewater required?	4 Monitori ng Require ments	The required analytes are solid-phase only; porewater is "strongly recommended" but not required, We struggle to understand how a process-based model with detailed process representation involving solid, aqueous and gas phase, could be successfully parameterized and validated with once-a-reporting-period measurements of solid-phase samples alone. Robustly parameterizing such a model is likely to require vastly more and	Changed. This concern has been addressed by other changes in the Module and the requirement that process-based models are required to represent hydrological state and the water budget, so the aqueous phase is modelled regardless of whether porewater is sampled. It is important to distinguish porewater sampling from porewater modelling. A process-based model represents the aqueous phase in all cases; porewater sampling only

		different kinds of data than a project developer budget can support.	determines whether there is a direct measurement available to validate the model's porewater predictions. Where porewater is measured (strongly recommended), the model must reproduce the measured porewater chemistry. Where it is not, the model is not required to reproduce porewater chemistry, and validation instead rests on the required solid-phase measurements, the immobile-tracer and mobile-cation mass balance from which CDR is derived, together with the cross-variable consistency requirement (Section 5.4). We have clarified the affected requirements so that validation targets are conditional on what the model predicts and what is measured, rather than implying porewater is a mandatory measurement.
	4.3.1 Monitoring Area Measurement Requirements	*[Porewater measurements are not required but are strongly recommended as input parameters for process based models and training data for data driven models to improve quantification accuracy and validate the soil phase method.]* Other sections would suggest they are required because it is a requirement for the model to reproduce porewater chemistry (Sections 4.3, 4.6, 7.0). Please clarify.	Changed. Thank you for the comment. We have provided further clarification on this. The model is required to reproduce porewater chemistry measurements if such measurements are taken. We have revised this so it is now consistent throughout the Module
	5.3 Process Representation in Models	*[The model must be capable of predicting the time-integrated evolution of porewater geochemistry and alkalinity export from the soil column as a function of site-specific inputs,]* If porewater measurements are not required, how can this be validated?	Changed. Thank you for your comment, we have revised the Module accordingly. The module now follows the idea that the model must be validated against the observables it actually relies on and that are actually measured, not against a fixed list of variables regardless of model type. Therefore, validation targets are conditional on what is measured: a model must reproduce porewater chemistry only when porewater is used as a model input/target or is measured, otherwise representing this process is recommended.
	8.0 Model Validation	*[porewater chemistry,]*	No change. The commenter is correct that porewater measurements are not required, but if measured, porewater

	Requirements	But porewater measurements are not required?	chemistry may be used to demonstrate resolution requirements are met.
	8.0 Model Validation Requirements	Asked whether all models must reproduce porewater chemistry measurements or only where such measurements are part of the project's MRV, and requested clearer requirements.	Changed. Thank you for the comment and we have provided further clarification on this that the model is required to reproduce porewater chemistry measurement, we only require this if such measurements are taken. We have revised this so it should be consistent throughout the Module.
	8.0 Model Validation Requirements	*[Project Proponents must demonstrate that the calibrated or trained model reproduces measured porewater chemistry, soil geochemistry, and other relevant monitoring observations from the project area, on a held-out subset of the monitoring data, to within the validation performance bounds.]* Not all of these are required, especially if a data-driven model aims to just predict CDR directly. Allow data-driven approaches to "reproduce CDR and other target variables as applicable"	Changed. We thank the reviewers for these comments on appropriate validation metrics. These sentences have been updated to require that validation metrics and reproducibility must be evidenced for CO2e Export, RP, or other suitable parameters identified in consultation with Isometric. This addresses differences in process based and data driven models, while allowing Project Proponents to suggest an alternative term where it is more appropriate for the type of model being used.
	8.3 Validation Metrics and Evaluation of Model Performance	*[Project Proponents must report validation metrics on the CDR estimate and any required calibration target variables computed using the held-out validation data.]* A critical point here is that models are validated against CDR estimates which are not measured directly. We have seen that statistical methods (e.g., outlier removal, averaging, and bootstrapping) can significantly alter the CDR estimate derived from directly measured parameters. Therefore, models should predict directly measurable metrics (e.g., alkalinity export in porewater, cation and tracer contents in soil, or similar), and not the CDR estimate.	
Matching and monitoring-area representativeness (AoA, similarity index)	4.1 In-Field Monitoring Approach	*[Deployment areas should be within a 50 km radius of the monitoring areas to which they are matched to ensure consistency in climate conditions.]* The distance constraint seems unnecessary here if the conditions are appropriately represented. Suggest that	Changed. We have removed this requirement because we agree it is redundant. Distance was used as a proxy for climate similarity, but consistency is already demonstrated directly through the required matching on temperature and precipitation, together with the applicability requirement that the temperature and precipitation

		this should instead be based on the matching AoA test.	experienced during each Reporting Period fall within the validated envelope.
	4.1 In-Field Monitoring Approach	*[Project Proponents hosting deployment areas outside of this radius from the matched monitoring area must demonstrate consistency in climatic conditions over their project area. In this case, Project Proponents must identify the climate variables identified for similarity matching and perform statistical similarity measures.]* How are project proponents expected to demonstrate consistency in climatic conditions? Which data are required? These should be universal factors that can be applied to each project location. Also, please specify which statistical similarity measures need to be performed.	
	4.2 Monitoring Area Representativeness	*[The following parameters must be matched between monitoring plots and deployment areas, or demonstrated that they don't need to be matched because they are: (1) explicitly simulated by a process-based model, (2) explicitly represented as a model input, or (3) homogeneous across the entire project region.]* Regarding criteria (1) in the quoted text — just because a model represents a process doesn't mean you can apply it to "unmatched" data (potentially outside of the validation envelope). For example, if a process-based model is calibrated to a monitoring area's base saturation, adjusting for the deployment base saturation is nontrivial. It's an emergent property, so you can't just plug in a new value without risk of violating mass and charge balance. You would likely need to create a brand new spinup run which, itself, would require new model calibration. Process representation should not be conflated with model applicability.	Changed. Thank you for the comment, we agree that process representation and model applicability are distinct, and have modified this section to eliminate criteria 1. We note that base saturation is not treated as transferable on the basis of being modeled, but is assessed between monitoring and deployment areas through the matching framework.
	4.2 Monitoring Area Representativeness	*[Matching Requirements]* Under this module, some models are likely to misrepresent the systems they're simulating (e.g., see S4.3 process inclusion criteria). But even if the model misrepresents the system, rigorous matching requirements could build	Changed. Thank you for the feedback. We agree that because this Module credits by extrapolating a model from the monitoring area to deployment areas, the rigor of the matching sites carries more weight than in the EW 1.2

		confidence that the model still reaches the right answer by confirming it's been applied only to conditions resembling those where it was validated. As written, we see two potential issues. (Outlined in the replies due to character limit). First, the only parameters that must match are the feedstock (type and grain size), temperature and precipitation (mean annual is allowed), cropping system, and farm management practices. These variables matter but the literature shows that they are very poor predictors of CDR outcomes on their own. As such, they don't go far in building confidence that the validation data captures deployment conditions. Additional similarity measures are left to the discretion of the project proponents and certifying bodies. Specifically, other matching parameters are caveated with "should" be included, rather than "must" (e.g. the list of pre-application soil properties considered). This puts pressure on Isometric and the verifier to hold the line in the face of unconvincing or narrowly-matched similarity data. We think a stronger approach would be to specify the set of continuous parameters that "must" be considered, and define a regular schedule for revalidating the match. The second problem is that all matching variables can be considered independently, despite the fact that many are tightly coupled. This means, for example, that a deployment's precipitation, temperature, and pH values could all "match" with the monitoring plot on their own, but the combination of values in the deployment, and how they covary, could have no analog in the validation data. The recommended DI approach of Meyer and Pebesma (2021), to our understanding, doesn't fix the problem. It ignores the covariances so a novel combination of variables could appear "similar" to the validation data if the individual values are within range. Meanwhile, covariance structures seem to be considered elsewhere in the module (particularly the validation process, S7.1, point 3). We think that consideration should be extended to the matching requirements.	Protocol's measurement only approach, so we have strengthened the matching requirements as suggested. The core set of continuous soil parameters is now required rather than recommended, so pH, cation exchange capacity, soil texture, base saturation must all be included in the similarity assessment. We have also added a requirement for the dissimilarity index to be based on a covariance-aware distance metric.

		With these shortcomings, it would not be surprising for deployment areas to pass the “match” test without reasonably conforming to the monitoring plot, raising the risk that models are applied to areas outside of their validation envelope. As discussed in S4.3, this is especially concerning in cases where the model is known to not explicitly simulate key processes that should be expected to change outcomes and are not “matched” directly.	
	4.2 Monitoring Area Representativeness	*[Project Proponents must document and provide justification for:* *The exact set of variables used to define the vector and how they are scaled* *The distance metric selected to compute similarity* *Any inclusion or exclusion thresholds]* There are many degrees of freedom here. It may be better to set some requirements to ensure all projects adhere to the same quality standards and principles.	
	4.2 Monitoring Area Representativeness	*[Feedstock grain class size.]* What about feedstock roughness factor and specific surface area	No change. Thank you for the comment. Feedstock roughness and specific surface area are already constrained by the required matching on feedstock type, composition, and Wentworth grain size class. Where a model uses them, they should be treated as model inputs and captured through feedstock characterization rather than the site similarity vector, so we have not added them as standalone matching parameters.
	4.2 Monitoring Area Representativeness	*[Categorical parameters to match:]* A model can be trained across multiple feedstocks, climatic regimes, and crop types. Suggest that this could be enabled by adding a requirement to use a similarity index where the model is trained across several contexts, and ensure representativeness of monitoring sites across deployment scenarios.	Changed. Thanks for the feedback. The similarity framework does support the scenario where a model is trained across multiple feedstocks, climatic regimes, crop types, etc. We have added the following language to make this more explicit: "Each deployment area must be matched to the monitoring area(s) whose data are used to train/validate the model, using the similarity framework outlined in Section 4.2. E.g. if a separate model is developed for each monitoring area, then the deployment area must be matched to at least one monitoring area; but if a

			single model is developed using many monitoring areas, then the deployment area must be matched against the pooled set of monitoring areas used to train the model."
	4.2 Monitoring Area Representativeness	*[Note that there may be additional matching requirements for data-driven models compared to process-based models–this is because process-based models can explicitly simulate processes such as crop growth and particle size evolution, while data-driven models require closer parameter alignment to ensure the trained models are applicable.]* Please clarify what the additional requirements are.	Changed. We have clarified that for data-driven models, any additional variable used as a model input must be included in the matching between monitoring and deployment areas, unless the variable is homogeneous across the project region. This is because a data-driven model interpolates within its training distribution and cannot extrapolate mechanistically.
	4.2 Monitoring Area Representativeness	*[Farm management practices. Tillage regime, irrigation practice, and fertilizer application should be consistent between monitoring and deployment areas, as these affect soil moisture, pH, and mineral–soil contact.]* These are qualitative features, how do you ensure / match them?	Changed. Thanks for the feedback. This list is meant to be recommended and not required, so we have updated the language accordingly, and clarified that they can be considered qualitatively (e.g. through a description and justification of their consistency between the monitoring and deployment areas).
Post-application (T0) baseline sampling	4.1 In-Field Monitoring Approach	Suggested addition: "Where soil solid-phase measurements, immobile tracers, or mobile-cation inventories are used for model initialization, validation, or CDR quantification, Project Proponents should collect an immediate post-application T0 sampling round after feedstock spreading and before substantial weathering is possible using the same sampling depth, compositing procedure, and analytical methods as subsequent monitoring rounds. If T0 sampling is not performed, the PDD should justify how feedstock carryover, placement heterogeneity, and extraction/digestion artifacts are separated from weathering during the Reporting Period."	No change. We appreciate this reviewer's feedback stressing the importance of correctly recovering the application rate in model based quantification. We agree with this sentiment but ultimately maintain that suppliers may select a method for the application rate check, which may not necessitate a sampling round immediately post-application, though we stress the importance of having enough statistical power.
	4.1 In-Field Monitoring Approach	*[A sampling round must be collected at baseline, prior to rock spreading. Additional sampling rounds must be conducted at the end of the Reporting Period.]* We suggest adding an explicit	

		requirement or strong recommendation for an immediate post-application T0 sampling round after spreading, before meaningful field weathering is possible, at least in the monitoring area and any deployment-area subset used for soil-phase validation. Including a T0 sample would i) provide in-field validation for the calculation of feedstock-associated base cation addition using the immobile tracer signal; ii) help distinguish weathering signal from analytical extraction/digestion artifacts; iii) improve estimates of uncertainty in soil cation changes.	
	4.3.1 Monitoring Area Measurement Requirements	*[Required at the beginning and end of the Reporting Period]* For the immobile-tracer/mobile-cation measurements used for weathering determinations, "beginning of the Reporting Period" should be clarified. If this means only pre-application baseline, it may be insufficient for tracer/cation mass-balance approaches because the model state after spreading is not directly observed. We recommend that the cadence for this row distinguish among: pre-application baseline, immediate post-application T0, and end-of-Reporting-Period measurements. This is particularly important when the same measurements are used to infer feedstock addition, local application rate, remaining feedstock mass, or cation loss.	
Treatment plot use across application rates/areas	4.1.1 Monitoring Area	*[Project Proponents may subdivide the treatment plot of a monitoring area to test multiple application rates. In this scenario, Project Proponents must clearly demarcate boundaries between different application rates, collect a number of samples that demonstrates adequate statistical power to quantify uncertainty, and perform the minimum monitoring requirements listed in Table 2 in each application rate area.]* The purpose of this isn't clear from the text -- is this using the same treatment plot for multiple project areas with different application rates, a test for variable rate applicators, using a higher rate to extrapolate where there isn't a	Changed. We thank the reviewer for this feedback. This paragraph has been deleted in the interest of concision.

		resolvable signal, or any/all of the above? There would be different requirements for each of these scenarios, especially if you're extrapolating a higher rate for an unresolvable signal.	
Monitoring sampling design: compositing and prescribed statistical tests	4.1.1 Monitoring Area	*[statistical tests]* The type of statistical test and significance level should be prescribed by the protocol.	No change. We refer the reader to the end of this paragraph, which directs readers to the discussion of rarefaction analysis and power analysis in the EW Protocol.
	4.3 Minimum Measurement Requirements	*[To account for in-field heterogeneity, soil samples should consist of multiple soil cores or subsamples. It is strongly recommended that soil samples be composed of 10-20 composited soil cores or subsamples randomly or arbitrarily distributed about a sample coordinate.]* Taking multiple soil cores is good scientific practice. Compositing should be a requirement and not a recommendation.	No change. We agree with this comment and will maintain that soil samples should consist of composited soil cores, but will not prescribe a number of cores to be composited.
Deployment area validation: coverage, randomization, failure handling	4.1.2 Deployment Area	*[Once the matching criteria check is passed, Project Proponents will not be required to analyze deployment area samples for geochemical parameters in subsequent Reporting Periods.]* Does this statement mean that Project Proponents can stop measuring deployment geochemical conditions after achieving a single reporting period match? If that's correct, then this decision raises the risk that deployment areas depart from their matched monitoring areas over time, which could cause modeled CDR outcomes to diverge from reality undetected. It would also create a strong incentive for matching which we think should be met by strict matching requirements. As it stands, the matching requirements appear pretty flexible (we elaborate more in S9.2 "Matching Requirements"). For example, since no specific geochemical parameters are required for matching and they are treated independently, it seems plausible that a Project Proponent could select the parameter(s) that yield the best match, even if other data show significant dissimilarities.	Changed. Thank you for the feedback. We agree that there is a risk that deployment areas depart from their matched monitoring areas over time, particularly at this early stage in EW where we have few examples of multi-year projects. We have strengthened this to require the Deployment Area Validation check to be conducted at every reporting period, and will revisit the frequency of this check in a future module update once we have more field data demonstrating how well matching holds over time. Since this check is now required every RP, there is no longer a need for triggering re-validation as a result of changes in farming management practices. However, reporting any changes in practices is still required under the EW Protocol. In addition, we have updated the check for the Deployment Area Validation to require demonstrating that the CDR distributions from the model and measurements are not statistically significantly different, to be consistent

	4.1.2 Deployment Area	*[Once the matching criteria check is passed, Project Proponents will not be required to analyze deployment area samples for geochemical parameters in subsequent Reporting Periods.]* In our opinion, there should be validation tests on field samples at least 2-3 times during the project lifetime, in randomly selected parts of the deployment area, covering a certain percentage of the full project (e.g., 10%). It should also be ensured that there are no changes in farming practices in parts of the deployment area that are not affecting the monitoring area. This should be confirmed by the farmer in written form. If the developer fails to get this confirmation from the farmer, the coverage of the deployment area should be increased (e.g., 30%) and time intervals for validation sampling should be increased.	with the representativeness check under EW v1.2.
	8.5 Deployment Area Validation	Requested clarification on what changes to deployment conditions would trigger another Deployment Area Validation.	
	4.1.2 Deployment Area	*[Project Proponents may choose the location within the deployment area, the total spatial coverage, and the sampling density.]* In previous statements, Isometric chooses the deployment area where validation will be performed. Suggest that this should be randomized and that the requirement is checked for consistency with the rest of the module.	No change. Thank you for the comment. We have kept this section as is to allow for flexibility in project and monitoring design. We also recognize that mandating randomized validation locations may be operationally infeasible for some projects. Note that any project using this Module is held to the same data reporting requirements in Section 11 of the EW 1.2 Protocol, which includes pre-registration of Removal Areas, arbitrary exclusion of fields, and triggers for data audits.
	8 Model Validation	The overall validation architecture is coherent and strong. We support the complementary checks against overfitting: cross-validation across spatial units, re-validation at each Reporting Period against data that did not exist during model development, deployment-area validation against a broader set of conditions, and validation of individual modeled terms in the net CDR equation.	No change. Thank you for the feedback.

	8.5 Deployment Area Validation	*[derived from concentrations of immobile tracers and mobile cations collected from the deployment area and must not have been used in model training.]* Should this be expanded to include total cation accounting, not just tracer-based methods?	Changed. Total Cation Accounting is acceptable, and we have amended the language to be more inclusive of that. This now reads "These measurements must be derived from concentrations of mobile cations and immobile tracers (if using)..."
	8.5 Deployment Area Validation	Asked whether failing the Deployment Area Validation results in no credits being issued for that reporting period.	Changed. Credits may not be issued if a Deployment Area Validation is failed. We have stated this more explicitly now to make it clear stage-gate for credit issuance.
	8.5 Deployment Area Validation	Raised a concern that there could be pressure to permit model-based crediting even when the validation check fails.	
	8.5 Deployment Area Validation	Suggested that for the first Reporting Period, projects use the higher-density measurement-based v1.2 approach alongside the model until the Deployment Area Validation is demonstrated.	No change. Thank you for the feedback, we agree that is a reasonable thing a Project Proponent may choose to do to mitigate their own risk, however that choice is left to the Project Proponent and will not be required in the Module.
	8.5 Deployment Area Validation	A pass threshold of deployment RMSE < 2× cross-validation RMSE is very capacious; a model performing at twice its monitoring-area error can still pass. The tolerance should be justified against the materiality threshold rather than set at an arbitrary multiple.	No change. Thank you for the feedback. The deployment RMSE check is a secondary backstop that is in addition to the main criteria that the modeled and measured distributions of weathered fraction of CDR density are not statistically significantly different. We have not tied the tolerance to the materiality thresholds in the Standard or GHG module, since those bound the aggregate credited quantity in validation and verification, whereas this is a per-point predictive error metric, so the two are not directly comparable.
	8.5 Deployment Area Validation	*[Model predictions must be compared against independent field measurements collected from a subset of the deployment area, as defined in Section 9.1.2. These measurements must be derived from concentrations of immobile tracers and mobile cations collected from the deployment area and must not have been used in model training.]*	No change. Thank you for the comment. Full coverage of the Deployment Area is already required for agronomic and site-characterization measurements to capture baseline heterogeneity. These full coverage measurements are also required at every Reporting Period as well. What is not required throughout

		This means that there are some project areas that are never sampled or tested. However, we think that full coverage [PO3.1]of the project area at least at baseline and reporting period 1 is essential. Baseline heterogeneity may exist due to farm history or soil type variation that is not previously known to the farmer and project developer. There is a significant risk of false crediting if large parts of the project area are never sampled. We agree that in later time points, sampling subsets of the project area selected by Isometric may be sufficient. If this re-validation fails, either the model needs to be re-calibrated, re-trained or further developed (and potential over-crediting needs to be corrected), or more data needs to be collected to test whether local artifacts were responsible for the validation failure.	the entire deployment area are the intensive geochemical measurements of CDR itself, as requiring that across the entire Deployment Area would be the same as the measurement-only approach of the EW 1.2 Protocol. Instead the validated model is meant to supplement a reduced set of measurements. On re-validation failure, the Module does require the model to be re-calibrated, re-trained, or further developed, consistent with the commenter's suggestion.
Sampling density basis (statistical power vs fixed minimum)	4.1.2 Deployment Area	*[The number of independent samples collected at each sampling event from the deployment area must hold adequate statistical power for quantifying uncertainty.]* It seems contradictory to base sampling density for the deployment on statistical power for the first reporting period and then set a minimum density for subsequent reporting periods; recommend just moving to statistical power.	Changed. Thank you for this feedback. We have updated the module to add flexibility in the deployment area and validation check to allow for more flexibility in project design. We maintain an overall minimum sampling density of 1/10 ha in the deployment area, however, we have added more guidance on the breakdown between deployment area sampling for the validation check and for agronomic monitoring for matching criteria. This includes flexibility for the project proponent to aggregate deployment area samples in a small area to perform the validation check, so long as spatial coverage is maintained for agronomic variable monitoring. With that, we have clarified that Project Proponents must have statistical power in the samples used in the deployment area validation check, but not in the agronomic monitoring samples.
	4.1.2 Deployment Area	*[Soil samples must be collected from the deployment area at a minimum sampling density of 1 sample per 10 hectares]* Why is there a sampling minimum for the deployment only? It would make more sense to make sampling density based on statistical power for the full project.	
Non-carbonic acid weathering (exchangeable Al, sulfur, strong acids)	4.3.1 Monitoring Area Measurement Requirements	*[Total sulfur content]* If this is not a measurement of reduced sulfur only, how does this inform non-carbonic acid weathering?	No change. We recognize that measurement of reduced sulfur is an intensive analysis to perform at scale. Instead, we take the conservative assumption that all sulfur in the system is from the oxidation of reduced sulfur.

	5.3.2 Processes for Quantifying CDR Terms	*[The Module requires that the fraction of weathering attributable to strong acids (nitric, sulfuric) be deducted from the gross CDR.]* This phrasing is a bit confusing -- in practice, the model would output the impact on DIC or alk export (or surface CO2 flux) relative to the counterfactual, rather a specific deduction that reflects 'fraction of weathering attributable to strong acids'.	Changed. Thank you for this feedback, we have updated the module text accordingly.
	5.3.2 Processes for Quantifying CDR Terms	*[Where non-carbonic acid weathering is represented in the model, the following processes are required:]* The current Module does not appear to adequately consider the role of exchangeable Al in acidic soils on Non-carbonic acid weathering. In soils with pH values of approximately 4 to 5, exchange acidity is often dominated by Al3+, and Ca^{2+} and Mg^{2+} released during enhanced weathering may displace exchangeable Al3+ from the soil exchange complex. The subsequent hydrolysis of dissolved Al releases H^{+}, which can contribute to non-carbonic acid weathering and reduce the net alkalinity benefit associated with enhanced weathering. Because this mechanism may be particularly important in acidic, highly weathered soils, explicit guidance on the treatment of exchangeable Al and Al hydrolysis should be included in both this Module and the Enhanced Weathering in Agriculture Protocol.	No change. Thank you for your comment. We agree with this comment and want to clarify that the consideration of this pool of acidity is already included in the Module and can be found in Section 5.3.2.
	5.6 Summary of Required and Recommended Processes	*[Recommended]* I encourage Isometric to improve the consistency between the monitoring requirements and the process-based model requirements regarding non-carbonic proton sources. The monitoring requirements specify that Total S should be measured as a required parameter, indicating that sulfur-related processes are considered relevant to enhanced weathering. However, the corresponding process-based model only recommends, rather than requires, representation of non-carbonic proton sources. This appears inconsistent, as measurements intended to characterize	Changed. Thank you for the feedback. We have edited this section to clarify that projects must account for non-carbonic acid weathering in the process based model.

		important weathering processes should also be reflected in the process-based model used for carbon removal quantification. Moreover, although total S is required to be monitored, total N, which is also a proton source in the process of nitrification, is not required in Monitoring. Given that N flow is drastic in agricultural management, I suggest the total N also must be required in Monitoring protocols.	
Parity between process-based and data-driven models	5 Process-Based Models	We have identified meaningful inconsistencies in requirements between process-based and data-driven models. The requirements are overall more stringent for process-based models and we are concerned that this may result in disincentivizing their use. The underlying theme of the disparity in requirements is that process-based models must consider the subsurface in a way that is not required for data-driven models. For example, per Section 4.3, process-based models must be capable of predicting the time-integrated evolution of porewater geochemistry, and to our read of the module this is not required for data-driven models. This is a heavy burden – across geographies, suppliers and researchers alike have identified that there are discrepancies between measured alkalinity export from soil and porewater. This observation contributed to the removal of the requirement for secondary validation measurements from the Enhanced Weathering in Agriculture Protocol v1.2, and the module is thus out of sync with the requirements of the protocol. Another example of process-based models being required to consider the subsurface where data-driven models are not is in the requirement to reproduce subsurface hydrology (per Section 4.4). Process-based models should not be held to a higher standard than data-driven models, and we ask that these requirements are removed to maintain parity between all model types.	Changed. Thank you for the comment. The requirements for process-based models have been updated to align more closely with the measurement requirements in the EW 1.2 Protocol, while retaining appropriate parity with data-driven models.
RTM benchmarking	5 Process-	Support for benchmarking and intercomparison of major ERW reactive transport models on common datasets to	Changed. We agree that benchmarking and model intercomparisons are crucial for

and intercomparison	Based Models	build confidence in model-based CDR predictions.	assessing structural uncertainty. We have added to the Model Validation section that external benchmarking is highly recommended and should be prioritized as a test, along with a suggested approach. However, we are cautious about making it a requirement at this stage, for two related reasons. First, the value it would add is uncertain at this stage. There are many control variables that are opaque to project developers but are effectively controlled for by developing a model within a confined geographic and climatic area. The Module now recommends demonstrating model performance against an externally held benchmark dataset (e.g. Beerling et al., 2024). But we are cautious that this dataset is most applicable to enhanced weathering in the US Corn Belt, and that at this early stage an external benchmark will likely bring less value than regionally specific datasets or a project proponent's own ERW site data. Second, ERW modelling is still a relatively new field, and publicly available datasets remain very limited for such tests. All that said, we accept the advice and have added benchtesting as a recommendation, with rules for how the test should be conducted.
	9.1 Sources of Uncertainty for Process-Based Components	Structural model uncertainty is well-appreciated as an equal or larger source of uncertainty than parameter uncertainty (Berkowitz et al. 2016; Skierszkan et al. 2021; Haase et al. 2013; Carrera and Neuman 1986). As written, the Module operationalizes structural uncertainty as within-code configuration perturbation (swapping thermodynamic database, rate law, secondary-mineral treatment) inside a single model, which conflates structural with parametric uncertainty and does not capture between-code structural spread. In our view the only way to address this problem is through multi-model ensemble approaches (sensu Steefel et al. 2015; see also Lu et al. 2015 for the Hanford uranium case), yet the Module requires structurally distinct ensembles for data-driven components (§8.2) while listing multi-model comparison as merely one option for process-based components. Until multiple models can be robustly benchmarked on very high quality solid- and aqueous-phase datasets from a variety of environmental contexts, there is no way to assess the degree of structural uncertainty, resulting in overstated certainty in CDR predictions. We recommend that a benchmarking prerequisite gate the process-based pathway.	
		We recognize that ERW currently exists in a data-poor state, and that creating a formal, externally held benchmark dataset is not currently possible. However, we feel strongly that this should be a prioritized area of development and that the module should reference its necessity, even if it cannot currently be required. In the interim, we suggest that all project proponents should be required to furnish their own external benchmark dataset (separate from the validation dataset) and must use ERW sites that are sufficiently different from their specific validation set (e.g., using public ERW studies like Beerling and Kantola) to show a minimum level of performance on the ERW task. This is to prevent overfitting to a specific developer’s own validation set	

		(consider the benchmark dataset as the floor of performance, and the validation set as what will determine crediting performance) and encourage robust performance. Because the goal of benchmarking is to ensure that the model has appropriately learned ERW processes, the defined minimum performance should be less strict than performance on validation (e.g. mean bias being within 2 times the pooled measurement uncertainty of the benchmark observations).The benchmark dataset would be provided to the registry and VVB and would be cited in the PDD.	
Data-driven models: framing, architectures, and interpretability	5.1 Introduction	*[Process-based models can be viewed as a "glass-box" modeling framework where the underlying mechanisms and relationships are transparent and interpretable]* This comes across as biased language, implicitly labeling data-driven models as 'black box' frameworks. Data-driven models can also be interpretable. Recommend removing this phrase.	Changed. Thank you for the comment. We agree, and have made these changes.
	6.4 Model Development Requirements	*[Project Proponents are recommended to train and report results for at least three model architectures drawn from distinct model families]* Suggest rewording to 'multiple model architectures' rather than setting a specific recommended number	
	6.0 Data-Driven Models	*[Data-Driven Models]* We would recommend to encourage data-driven approaches to use interpretability techniques e.g., sensitivity analysis with known scientific processes as done in our recent ICLR publication, Sparse Auto Encoders for neural networks as used for weather prediction, SHAP. We suggest adding in some text to this effect. Given data-driven approaches can differ, it is hard to specify the exact type, but we include few examples above that should cover most model architectures.	Changed. Thank you for your comment. Based on this, we have added a recommendation to use interpretability techniques.
Scope/definition of process-based	5.1 Introduction	*[By explicitly defining the governing partial differential equations for fluid flow and chemical reactions, these models provide a mechanism for scaling CDR estimates across varied climatic*	Changed. Thank you for catching this inconsistency, we have reworded the introduction to the process-based

models (PDE vs ODE/box)		*and agronomic domains.]* There seem to be conflicting statements here with respect to having a process-based model with explicit spatial discretization vs embracing all kinds of process-based models. If you are looking for models governed by PDEs, then that specifically applies only to the former or i.e. RTMs. Other mechanistic models (e.g., box models, coupled ecohydrological-geochemical models like SMEW) are based on a system of ordinary differential equations because there is only time variation for such models that lump the entire length scale into a single reservoir	model section to resolve this. The description of spatial discretization and governing partial differential equations is now tied specifically to spatially resolved models such as RTMs. We have also clarified that the introduction is a general overview of the class of process-based models, and that eligibility under this Module is determined by the requirements set out in the Module below.
Spatial heterogeneity within the NFZ (rootzone, erosion/deposition)	5.2 Model Set-up Requirements	*[The NFZ is characterized by strong geochemical gradients in moisture, pH, pCO2, cation concentration, and alkalinity, which drive the weathering reaction and loss processes. Models representing this zone must resolve these gradients at sufficient spatial and temporal resolution to capture their non-linear feedback.]* Module should account for spatial heterogeneity of enhanced rock weathering (ERW) within the NFZ (near-field zone), including difference between rootzone and outside the rootzone (non-rootzone). Recent field evidence using quantitative X-ray powder diffraction (XRPD) shows that the decrease in feedstock minerals is substantially greater in rootzone soils than in non-rootzone soils. Specifically, the decrease in basaltic minerals such as plagioclase and pyroxene was about 8.5 times greater in the rootzone than in the non-rootzone (Kurokawa et al. 2026, Applied Geochemistry, https://doi.org/10.1016/j.apgeochem.2026.106869). This indicates that plant roots induce localized condition that greatly enhance mineral weathering and that spatial heterogeneity of mineral weathering should be considered when estimating feedstock dissolution and CO2 removal. Therefore, the Module should recognize root-induced heterogeneity within NFZ as an important factor controlling the estimation of CO2 removal.	Changed. We agree that root-induced spatial heterogeneity within the NFZ is a real and mechanistically well-supported phenomenon: the rhizosphere concentrates the drivers of weathering the Module already identifies: elevated pCO_2 from root respiration, localized pH depression, and moisture cycling and can therefore enhance mineral dissolution relative to bulk soil. We have revised the Model Boundary requirements (Section 5.2) to explicitly recognize lateral rhizosphere-vs-bulk-soil heterogeneity alongside the vertical gradients already named. However, we will not mandate explicit geometric resolution of the rhizosphere as a distinct model domain, as we consider the current field evidence base insufficient to require this of all projects, and because for CDR accounting the quantity that must be unbiased is the NFZ-integrated flux rather than its within-zone spatial distribution. We will keep this under review as the evidence base develops.

	6.2 Model Scope Requirements	*[For data-driven models, the model boundary defines what input features the model is trained on, what outputs it predicts, and the range of conditions over which predictions are considered valid (the Area of Applicability]* The current treatment of spatial variability appears to focus primarily on uncertainty in soil properties while giving insufficient attention to systematic landscape-scale processes such as erosion and deposition. In large-scale enhanced weathering projects, redistribution of crushed-rock-applied surface soil by water erosion can substantially alter the amount and spatial distribution of reactive material within the NFZ. Consequently, weathering rates may be systematically underestimated in erosive environments and overestimated in depositional environments if these processes are not accounted for. This issue is particularly important for data-driven models. Therefore, the Protocol should require consideration of topography, erosion risk, sediment redistribution, and related hydrological processes when developing and applying data-driven models. Failure to account for these factors may limit the transferability and robustness of model predictions across heterogeneous landscapes	Changed. We agree that landscape-scale erosion and deposition represent a distinct source of systematic bias, redistribution of applied feedstock changes the mass of reactive material in the NFZ, which is not the same as the soil-property variability the representativeness framework primarily addresses, and which the Area of Applicability framework does not detect. We have added a recommendation (Section 4.2, Climatic and Hydrological Representativeness) that Project Proponents should assess erosion and deposition risk across monitoring and deployment areas using topographic and hydrological information, demonstrate geomorphic representativeness, and address any residual feedstock redistribution in the CDR estimate. We have made this recommendation model-agnostic rather than specific to data-driven models. We also note that the immobile-tracer mass-balance approach required under the EW Protocol provides partial correction for physical soil mass change; the new requirement addresses the residual case in which the applied feedstock is preferentially redistributed.
Model boundary requirements: clarity and structure	5.2 Model Set-up Requirements	*[grid convergence test]* Please provide more clarity on how this should be done and any metrics need to be quantified or abided by	Changed. Thank you for the comment, we have added specific suggestions on how to operationalize the requirement.
	5.2 Model Set-up Requirements	*[To be more specific, the NFZ is the zone over which the gross CDR signal (CO2eWeathering,RPCO2e{Weathering, RP}CO2eWeathering,RP) and the primary soil-column loss terms (CO2eBiomassLoss,RPCO2e{BiomassLoss, RP}CO2eBiomassLoss,RP, CO2eNetNewCarbonate,RPCO2e{NetNewCarbonate, RP}CO2eNetNewCarbonate,RP, CO2eNetNewSilicate,RPCO2e{NetNewSilicate, RP}CO2eNetNewSilicate,RP, CO2eNetSorption,RPCO2e{NetSorption, RP}CO2eNetSorption,RP, CO2eNonCarbonicNeut,RPCO2e{NonCarbonicNeut,*	Changed. Thank you for your comment, we have accepted this suggestion.

		RP}CO2eNonCarbonicNeut,RP) are determined, consistent with Equation 2.] This sentence feels a bit out of place here, suggest removing.	
	6.2 Model Scope Requirements	*[Boundary condition]* This section may need to be split - we became somewhat confused internally between model boundaries and "boundary conditions" (e.g. fixed value, fixed flux (Neumann), etc.). More concretely, this seems like a very difficult requirement to check.	Changed. Thank you for pointing this out, we agree the terminology was confusing. We have renamed "model boundary" to "model scope" throughout to more clearly distinguish between the spatial and temporal domains of the model/input features, etc., and boundary conditions. We have also clarified what constitutes a structural revision that constitutes a restart of the training and validation procedure.
Required vs recommended processes: criteria and 4.6 table clarity	5.3 Process Representation in Models	*[A process is recommended ("should") in a process-based model when:]* Does only one of the following three bullet points need to be true for the process to be recommended rather than required, or do all three need to be true?	Changed. Thank you for the comment, the bullet points were intended to show the rationale behind our thinking around categorizing the processes as required vs recommended. They are not criteria for the categorization to be true. We have streamlined this section to only list the criteria for when a process is deemed required, otherwise it is recommended.
	5.3 Process Representation in Models	*[The process is material to CDR for some project configurations but not others, such that universal representation would impose disproportionate burden.]* Does this mean that projects for which a recommended process is material will need to include that process in their process-based model? How will it be decided whether that recommended process is material for a given project?	No change. Thank you for your comment, if a process can be material for a project, recommended means we highly recommend that the process be included in the model. The idea behind the required vs recommended processes are that we set a bar for certain processes that we think are important to be included. As noted in the "required" process rationale, if sufficient data exist to constrain and parameterize the process at The Project scale, the Project should include the process in the model to better constrain the output
	5.3 Process Representation in Models	*[These terms may instead be accounted for through direct measurements from the Monitoring Area, in accordance with the relevant section of the Enhanced Weathering in Agriculture Protocol.]* Recommend that the language is tighter: "if these terms are not modeled, they	Changed. Thank you for the comment, we adjusted the language according to your suggestion.

		must be accounted for through direct measurements…" otherwise the language is too vague.	
	5.6 Summary of Required and Recommended Processes	We support the decision to structure process representation as a tiered set of required and recommended processes rather than mandating a single maximal model, as this appropriately acknowledges that no general geochemical code represents every relevant process at once. It's hard to envision a successful model that doesn't incorporate many of the processes recommended, but we think it remains an open question of which processes are absolutely necessary to represent in a model in order to get robust results.	Changed. Thank you for your comment. We agree it remains an open question of which processes are strictly necessary to represent at this stage, and have retained the overall framing of a tiered set of required vs recommended processes. We have refined which processes are classified as required vs recommended. The required set is now grounded in whether a process exerts a first-order control on EW efficiency and can be represented accurately with data sufficient to constrain it; the rationale for each classification is set out in Section 5.6 and Appendix B.
	5.6 Summary of Required and Recommended Processes	*[or]* The 'or' implies that recommended terms are not required to be measured. This table should indicate which terms need to be measured if they are not modeled (Non-Carbonic Acid Weathering, Net sorption, and Plant Cation Uptake)	Changed. Thank you for your comment, we have updated the table and the section. The table is meant to describe a model generating CO2e_export term, not a weathering-only model; the model spans all the processes generating the gross weathering signal. The "Required/Recommended" label refers to whether a process must be represented within the model, which is distinct from whether a term is required in the net CDR equation (governed by Equation 2 and the EW Protocol). For the terms the reviewer identifies, the table already reflects this: Non-Carbonic Acid Weathering and Ion Exchange (the primary Net Sorption pathway) are required to be modeled, and Plant Base Cation Uptake is marked as required to be measured if not modeled, so these loss terms cannot be omitted and must be quantified somehow. We have revised the caption wording to remove the earlier implication that recommended terms need not be accounted for, resolving the ambiguity between what the model must represent and what the net CDR accounting must include.
	5.6 Summary of Required and Recommended Processes	*[Summary of Required and Recommended Processes]* Is this section meant to represent the nuance between a model that represents only weathering vs. a model that represents net CDR? It would be useful to clarify this, because as is this table is a bit confusing with what is required vs. recommended.	
	7.2 Hybrid Model Development	*[Project Proponents must provide justification that this does not introduce data leakage or lead to overfitting]*	No change. Thank you for the comment. Data leakage as a general risk is addressed for all models in Section 8.1

	Require ments	Why is this only a requirement for hybrid models?	(Cross-Validation), which requires that data splits be determined from baseline measurements only and that cross-validation account for spatial and temporal autocorrelation to avoid information leakage between training and validation sets. The provision in the hybrid section addresses a distinct, hybrid-specific case: leakage between components when a single observational dataset is reused to both calibrate the process-based component and train the data-driven component. This cross-component failure mode does not arise for standalone process-based or data-driven models which have a single fitting step, so the requirement is scoped to hybrid models.
Risk of omitting first-order/uncer tain processes	5.3 Process Represe ntation in Models	*[A process is required ("must") in a process-based model when:* *The process is well understood and can be represented accurately in current process-based modelling frameworks.* *Sufficient data exist to constrain and parameterize the process at The Project scale.* *A process is recommended ("should") in a process-based model when:* *The process is known to influence EW efficiency, but current process representations in models are uncertain, simplified, or under active scientific development.* *The data needed to constrain and parameterize the process are not yet routinely available, or are only available for a subset of project conditions.* *The process is material to CDR for some project configurations but not others, such that universal representation would impose disproportionate burden.]* This section allows models to omit uncertain processes precisely because they are uncertain. Many of the "uncertain" processes that may be omitted are also known to be first-order processes in EW with decades of research in the critical zone modeling community (including surface area evolution, secondary silicate and amorphous phases, non-carbonic acid weathering, and ion exchange). That first-order processes may be omitted makes it more likely that wrong	Changed. Thank you for the feedback. We agree that the required set should be defined by whether a process exerts first-order control on the modeled result, and we have revised Section 5 with this in mind. We want to be explicit about how the Module draws the line, because two considerations are in tension and we have tried to balance them. On one hand, we understand the argument that poor constraint on a process and high parameter uncertainty is not necessarily on its own a sufficient reason to exclude it from quantification. A process that materially affects the result cannot be ignored simply because it is hard to model. On the other hand, mandating explicit representation of every first-order process regardless of how well it can be constrained does not necessarily produce a more accurate result: forcing a poorly-constrained process into the model can introduce equifinality and false precision, and can degrade rather than improve the CDR estimate. Our resolution is to classify processes by their first-order influence, but to make the classification conditional on demonstrated materiality rather than fixed. A process may sit in the recommended set where it is not

		answers are certified for two reasons. First, it could contribute to an artificially narrow uncertainty envelope that doesn't capture the true answer in the deployment area. Section 8 propagates uncertainty by varying retained model choices across Monte Carlo iterations. An omitted process contributes no spread to the resulting uncertainty envelope. The only provision for handling this is a clause instructing proponents to "assess this impact on the final CDR estimate and take appropriate measures to discount it from the total Credits" (S8.4). This should be clarified and strengthened, because assessing those impacts without dedicated data likely requires modeling the process (and its uncertainty) anyway. Second, omitting first-order processes pulls models away from reality, making it more likely that when they are calibrated to a given answer, they reach it for the wrong reasons. Getting the right answer for the wrong reasons is not inherently a problem given the apparent goals of the module, but it increases the risk of a wrong answer once the model is moved off the monitoring data and applied to the deployment area. The module states this risk directly. S8.2 acknowledges that held-out validation "can confirm a model is wrong but cannot rule out misspecification when validation and training data share the same blind spots" (S8.2). The module recommends (though does not require) an ensemble of structurally distinct architectures as the remedy for data-driven models (S5.4). We think the shared-blind-spot risk is serious enough to both require architectural diversity for data-driven models, and to require inclusion of processes based on materiality rather than tractability. As written, the ultimate backstop for mitigating both of these risks is the matching framework presented in S9.2, which aims to guarantee that the deployment conditions are within the validation envelope. As we discuss in our comments to that section, we think the matching process does not meet the high bar required to justify omitting key model processes.	universally first-order across configurations, but the required global sensitivity analysis (Section 5.5) serves as the backstop: where a Project's own sensitivity analysis shows that a recommended or omitted process materially affects the modeled CDR, on the order of a few percent of the result, across both near-term and long-term horizons, that process must then be represented in the model or quantified through direct measurement for that Project. Materiality is therefore established through sensitivity analysis rather than assumed a priori, and "recommended" does not mean a process may be disregarded where it is shown to matter. This approach keeps the Module operationally feasible in the near term while closing the risk that a material process is omitted, and it allows the recommended set to also signal priority areas for further research as the field's data and methods mature. On our present reading, the first-order processes are already in the required set; the principal case under active consideration for reclassification is soil pCO_2, which we address separately.

	5.3 Process Representation in Models	*[A process is recommended ("should") in a process-based model when:* *The process is known to influence EW efficiency, but current process representations in models are uncertain, simplified, or under active scientific development.* *The data needed to constrain and parameterize the process are not yet routinely available, or are only available for a subset of project conditions.* *The process is material to CDR for some project configurations but not others, such that universal representation would impose disproportionate burden.]* Isometric acknowledges in other occurrences in this module that models may not yet be ready to meet the high standards required in this protocol. Nevertheless, processes that are difficult to model or measure are simply excluded from the requirements. This does not stimulate development and efforts to better represent these processes in models but simply sets the bar lower for developers to neglect factors shown to influence CDR rates.	Changed. Thank you for the feedback. We have changed our reasoning for required vs recommended processes, and have made adjustments to this section.
	5.6 Summary of Required and Recommended Processes	*[Table 1]* One major concern is that if there are feedback loops between processes, accurately modeling the system is difficult when some of those processes are omitted. Meanwhile, if important processes are excluded, will tuning and calibration simply alter the modeled processes to compensate for the unmodeled ones?	Changed. Thank you for the comment, and we have made changes to the Module to mitigate both risks through a combination of requirements rather than a single provision. Process classification is grounded in first-order materiality, not tractability (rationale documented in [Appendix B / Summary of Required and Recommended Processes]), so the processes driving the principal feedbacks, dissolution kinetics, secondary carbonate precipitation, ion exchange, and CO_2 supply are required; and where the global sensitivity analysis (Section 5.5) shows a recommended or omitted process is material, it must be represented or measured, so omission is permitted only where it is shown not to matter. On calibration compensating for unmodeled processes: calibration must be joint against all measured components and reproduce them with a single consistent parameter set (cross-variable consistency), multi-parameter optimization is required precisely to expose

			compensating errors, calibrated degrees of freedom are minimized via the parameter-to-observation ratio, and, most importantly, models are validated against independent held-out data and directly measured deployment-area CDR, which a model that has calibrated its way to a fit through compensating errors will fail. We acknowledge these safeguards do not fully eliminate structural uncertainty, which remains a field-wide limitation to be addressed through multi-model comparison as the field matures.
Dispersivity: scale dependence and justification	5.3.1 Processes for Defining Boundary Conditions	*[longitudinal dispersivity]* Both small-scale laboratory experiments and field-monitoring data collected at spatial scales much larger than the NFZ may provide biased estimates of dispersivity. Because dispersivity is strongly scale-dependent, values derived from laboratory experiments often underestimate dispersivity at the NFZ scale, whereas values derived from large-scale field observations may overrepresent heterogeneity that is not relevant to the NFZ. Consequently, both approaches have the potential to yield dispersivity estimates that are inappropriate for the spatial scale of interest.	Changed. Thank you for the comment. Because this parameter is scale-dependent, we recommend that the Project Proponent obtain laboratory or field measurements at a scale close to that of the model (e.g., a 2 m vertical soil profile), as this represents the most rigorous approach. However, this should not be considered a strict requirement. Given that the modelling scale for a typical farmland soil profile is approximately 2 m—a scale that is well represented in the literature (e.g., Vanderborght & Vereecken, 2007)—using values reported in the literature is also a reasonable and scientifically defensible option. We have revised the Module accordingly, as shown below. Vanderborght, J., and Vereecken, H. (2007). One-Dimensional Modeling of Transport in Soils with Depth-Dependent Dispersion, Sorption and Decay. Vadose Zone Journal, 6(1): 140–148. https://doi.org/10.2136/vzj2006.0103
	5.3.1 Processes for Defining Boundary Conditions	*[Dispersivity values must be justified using either site-specific tracer test data or peer-reviewed scale-dependent compilations appropriate to the modelled depth interval.]* Dispersivity exhibits strong scale dependence, and similar soil textures do not necessarily result in similar dispersivity values. Therefore, dispersivity should preferably be determined from	Changed. We agree that this is the most rigorous approach. However, we do not consider it to be a strict requirement for modelling a typical ~2 m vertical farmland soil profile (e.g., sandy loam). First, for relatively homogeneous agricultural soils such as sandy loam, the use of literature values is generally well supported. Second, a modelling scale of approximately 2 m is well

		laboratory or field experiments conducted at a spatial scale comparable to the depth of the NFZ.	represented in the literature (e.g., Vanderborght & Vereecken, 2007), making scale-appropriate literature values readily available. In addition, model sensitivity often has a greater influence on simulation results than the exact choice of dispersivity within a reasonable range. Therefore, the use of literature values remains a reasonable and scientifically defensible option. We have revised this section accordingly.
Numerical scheme and spatial discretization	5.3.1 Processes for Defining Boundary Conditions	*[Numerical stability constraints for the advection-dispersion solver must be satisfied at the chosen Δz and Δt, and the resulting stability constraints must be reported.]* Please clearly define what you are looking for (i.e., Grid Peclet number and Courant-Friedrichs-Lew condition?) and what thresholds should be met; Also, "chosen Δz" seems to imply we are only talking about 1-dimensional models	Changed. Thank you for the comment. We would like to clarify that the protocol focuses on 1D simulation, and the parameters mentioned throughout the document are for 1D model. This is based on the assumption that lateral flow in NFZ soil profile simulation can generally be ignored. In the revised version, we include constraints for numerical stability and accuracy. Please refer to the Module for revision.
	5.3.1 Processes for Defining Boundary Conditions	*[Numerical stability constraints]* Since numerical stability can vary substantially depending on the numerical solving method, it is necessary to clearly specify whether the model employs an explicit or implicit temporal scheme, along with the corresponding numerical stability constraints.	Changed. Thank you for your comment, we have revised the section to include both numerical stability and numerical accuracy in the revised document.
	5.4 Calibration Requirements for Process-Based Models	*[Water flux accounting]* The Module identifies lateral flow and hydrological connectivity as sources of uncertainty in water flow. However, the types of flow processes that need to be considered depend on the underlying modeling framework (or spatial scheme). Therefore, before discussing specific flow pathways, the report of the model should clearly describe whether the model is one-dimensional or multi-dimensional, and whether water fluxes are dynamically simulated or represented by assumed constant fluxes. The rationale for adopting these assumptions should also be provided. Without this information, it is difficult to evaluate the relevance and completeness	Changed. Thank you for your comment. The parameters discussed throughout the document are intended for a one-dimensional (1D) model. In the revised version, we explicitly clarify that the protocol is developed for a 1D model and that lateral subsurface flow is not considered within its scope. We made revisions to the end of Section 5.1 introduction paragraph and the water flux accounting section

		of the uncertainty assessment regarding water flow processes.	
Data provenance and justification for parameter/dataset selection	5.3.1 Processes for Defining Boundary Conditions	*[Soil pCO2: Process-based models should account for CO2 entering the soil system from multiple sources: dissolved in rainwater, dissolved in irrigation water, and produced by root respiration or microbial respiration of organic matter within the soil profile.]* This is an overall comment, when there is a requirement for specific data to be used it would be helpful to clarify what data sources are acceptable.	Changed. Thank you for the comments. We have added a new requirement, "Data provenance and justification" to the Reporting requirements for Project Validation. This applies to all data and parameter values that are not from project-specific measurements (including from established databases, peer-reviewed literature, remote-sensing products, etc.), and explicitly requires documenting the source, an assessment of how representative it is of the project's conditions, and any assumptions involved in applying it beyond the domain in which it was measured. Where the source conditions differ materially from the project's conditions, the difference must be addressed through the sensitivity analysis and uncertainty propagation or by adopting a conservative value.
	5.3.2 Processes for Quantifying CDR Terms	*[Where experiments are used, the experimental design must characterize pH dependence across the relevant field pH range, reactive surface area and its evolution during dissolution, and surface passivation effects, and the experimental Protocol, data, QAQC and derived parameters must be documented in the PDD]* It would be useful to describe similar requirements for peer-reviewed values regarding what the studies they are derived from must have considered.	
	6.1 Introduction	*[Data-driven models are trained on observational datasets to learn relationships between observed inputs and outputs]* The current guidance does not appear to require sufficient justification for the selection of datasets and parameter values. Without a clear description of why specific data sources were adopted and how their representativeness was evaluated, there is a substantial risk of inappropriate spatial and temporal extrapolation. This is particularly important for parameters known to exhibit strong variability across scales, climates, soil types, and management systems. The Protocol should therefore require explicit reporting of data provenance, selection criteria, representativeness, and any assumptions underlying extrapolation beyond the original measurement domain.	
	5.3.3 Measurement	*[For each proxy-to-state-variable mapping, justify the appropriateness of the conversion for the specific soil*	No change.

	Proxies and Model Variables	*conditions of the project area, with reference to peer-reviewed literature, site-specific characterisation data, or paired direct measurements of the state variable in question (e.g., porewater extraction or in-situ pH measurement at a representative subset of monitoring locations).]* It would be useful to give an example of what an appropriate justification would look like here (beyond what needs to be referenced). Maybe a good point of comparison would be a justification that would not be considered sufficient.	Thank you for the comment. The requirement to justify the appropriateness of each proxy-to-state-variable mapping for the project's conditions remains in Section 5.3.3. We agree a worked example of a sufficient versus insufficient justification would be helpful and may provide illustrative examples separately through supporting guidance rather than in the Module text.
	5.3.3 Measurement Proxies and Model Variables	*[Several routinely-measured soil parameters used as model inputs are method-dependent proxies for the underlying state variables that govern weathering and ion exchange in the model.]* This is an important topic and we appreciate seeing an explicit treatment of it in the protocol.	No change. Thank you for the positive feedback, which are helpful for knowing what to retain.
Surface area evolution representation and calibration	5.3.2 Processes for Quantifying CDR Terms	*[Surface area evolution:]* Surface area evolution is an important factor for weathering rates in every project setting and should be required.	No change. Thank you for your comment. We have given this careful consideration and decided to keep the reactive surface area evolution term as recommended. Dissolution rate scales linearly with reactive surface area (rate = A·k·f(pH,...)), and as grains dissolve the effective surface area changes over time, so a static surface area will progressively over- or under-predict long-term dissolution flux. Representing this evolution is therefore advisable for simulations extending beyond a single season, though it is less critical for short-term screening estimates (Heřmanská et al. (2022); James et al. (2024)). Typically as in RTMs (e.g., CrunchTope, PFLOTRAN), surface area evolution for a given mineral is given by a power law that scales the initial surface area by the change in porosity or mineral volume fraction. Because empirical constraints on this evolution are sparse, keeping this as recommended preserves flexibility in adjusting the initial surface area that influences reaction rates via the TST rate law and

			can be reasonably constrained by using a measured BET surface area as the initial value.
	5.4 Calibration Requirements for Process-Based Models	*[Reactive surface area evolution terms.]* This is ambiguous. Please provide more details as to which parameters you have in mind that can be calibrated. Typically as in RTMs (e.g., CrunchTope, PFLOTRAN), surface area evolution for a given mineral is predicated by a power law based on the prescribed initial surface area scaled by the ratio of porosity or volume fraction change. Provided empirical constraints are sparse, is this more of a reference to flexibility in adjusting the initial surface area that influences reaction rates via the TST rate law?	Changed. Thank you for your comment, after further evaluation, we agree that reactive surface area does not need to be calibrated because reactive surface area can be constrained by BET surface area measurement and we think this falls within the definition of "directly constrained by laboratory experiments or field scale measurements." So this term is removed from the list.
Soil pCO_2 representation	5.4 Calibration Requirements for Process-Based Models	*[Soil pCO2 variation with temperature, moisture content and crop type]* With regard to soil pCO_2 estimation, approaches based on meteorologically derived net primary productivity (NPP), such as those proposed by Seino and Uchijima (1985), should also be considered as an alternative approaches based on agricultural production data, such as those used by Beerling et al. NPP estimates derived from crop yields or agricultural production are strongly influenced by management practices, including fertilization, irrigation, cultivar selection, and other technological inputs. Consequently, they may reflect anthropogenic factors rather than the environmental controls that determine soil CO_2 production. For this reason, agricultural production-based NPP estimates may introduce systematic bias into pCO_2 calculations and may not be appropriate for generalized modeling applications.	No change. Thank you for your comment and recommendation. The Module will not prescribe a specific method for estimating soil pCO_2 or soil respiration, and we have not added a method-level prescription, as this sits within an already-recommended component (dynamic soil pCO_2) and appropriate methods will vary by project and data availability.
	5.6 Summary of Required and Recommended Processes	*[Recommended]* I suggest that the representation of soil pCO_2 should be classified as a required, rather than recommended, component of process-based models. Soil pCO_2 is a fundamental control on carbonate chemistry, and therefore directly influences estimated carbon removal through enhanced weathering.	No change. Thank you for the feedback. We agree that soil pCO_2 is an important control on carbonate chemistry and weathering rates, and that treating it as a fixed constant risks systematic bias. We note that pCO_2 already enters the required components of a process-based model through its control on porewater pH in the mineral

		Unlike atmospheric CO_2, soil pCO_2 can vary substantially due to seasonal changes, root respiration, and microbial activity (see Zhu et al., 2025, https://doi.org/10.1002/saj2.70136). Consequently, assuming a constant or simplified pCO_2 may introduce significant uncertainty into model predictions. Recent experimental studies have also demonstrated that variations in pCO_2 can significantly affect weathering rates and alkalinity production under enhanced weathering conditions (see Amann et al., 2019, https://doi.org/10.3389/fclim.2022.929268). Given its strong mechanistic influence on the processes, omitting soil pCO_2 risks systematic discrepancies between modeled and actual carbon removal.	dissolution rate law. What the Module currently classifies as recommended is the dynamic mechanistic simulation of pCO_2 as a state variable driven by root and microbial respiration and gas diffusion. We will retain that as recommended, because mandating it requires parameterizing soil respiration, a data-intensive and weakly-constrained sub-model whose forced inclusion risks introducing false precision. However, in response to this comment we have clarified that soil pCO_2 must be represented and justified rather than assumed constant: where pCO_2 is not dynamically simulated, it must be carried with a defensible uncertainty range in the Monte Carlo propagation rather than fixed at a single value, and where the required global sensitivity analysis shows the CDR estimate is materially sensitive to pCO_2, dynamic representation or direct constraint becomes required for that project. This addresses the risk the commenter identifies, unquantified error from a simplified pCO_2 assumption, without mandating a mechanistic respiration model in all cases.
Flow and hydrology representation	5.4 Calibration Requirements for Process-Based Models	*[Effective hydrological connectivity]* What does this term mean? Is this the unsaturated hydraulic conductivity curve i.e. K(Psi)? I do not think this term is widespread in soil physics.	Changed. Thank you for the comment. This has been clarified and the terminology updated to hydrological conductivity
	5.4 Calibration Requirements for Process-Based Models	*[Examples of terms that are uncertain and can be calibrated include but are not limited to: the partitioning of infiltrated water between matrix flow, preferential flow, and lateral subsurface flow within the NFZ]* Resolving these different compartments of flow (e.g., matrix vs immobile water; vertical drainage vs lateral drainage) is arguably still a frontier in hydrology. Moreover, representing these processes in a model require a 2-dimensional domain at minimum	Changed. Thank you for your comment. The parameters discussed throughout the document are intended for a one-dimensional (1D) model. In the revised version, we explicitly clarify that the protocol is developed for a 1D model and that lateral subsurface flow is not considered within its scope.

Thermodynamic/ kinetic database recommendations	5.4 Calibration Requirements for Process-Based Models	*[Thermodynamic equilibrium constants]* It's worth mentioning Kinec_v3 database here (thermodynamics from CarbFix and kinetics from the Hermanska compilaton) https://doi.org/10.1016/j.chemgeo.2025.123041 as these databases seem tailored for low-temperature Earth surface conditions	Changed. Thank you for the recommendation, we have included it.
Crediting on noisy or non-significant measured signal	5.4 Calibration Requirements for Process-Based Models	*[add requirement for a statistically significant weathering signal in the monitoring plot measurements]* Suggest removing the second half of the phrase. We agree that calibrating against primary measurements is fine, but saying that it is 'preferable where the CDR signal is noisy or not yet statistically significant' may raise some flags that you didn't intend.	Changed. Thank you for the feedback. We have clarified that crediting is conditional on a statistically significant weathering signal in the treatment plot. We have added the following to the monitoring section "For a treatment plot to be eligible for matching and use in Credit issuance, the Project proponent must demonstrate a statistically significant weathering signal in accordance with Section 10.1.3 of the Enhanced Weathering in Agriculture Protocol." We have retained the option to calibrate a model against primary geochemical observables such as porewater cations, pH, and alkalinity, which helps when the derived CDR signal is noisy, but we have removed the phrase describing this as preferable where the signal is "noisy or not yet statistically significant," since significance is now a precondition for crediting.
	5.4 Calibration Requirements for Process-Based Models	We like the idea of permitting calibration against primary geochemical observables while treating the CDR estimate as an emergent validation metric, esp. when the CDR signal is noisy.	
	5.4 Calibration Requirements for Process-Based Models	*[Calibrating against primary measurements is acceptable and preferable where the CDR signal is noisy or not yet statistically significant. In this case, the CDR estimate should be treated as an emergent validation metric that is tested at the validation stage against independent data, rather than used as a calibration target.]* This means that model-based credits can be issued even if measured signals are non-significant. However, statistical significance should be a requirement for crediting, or at least the model should be eligible for issuance when non-significant calibration data are supported by independent validation and conservative uncertainty treatment to avoid false issuances. If the data do not show significant weathering and cation or alkalinity export, no credits should be issued. Non-significant differences should also not be used as an emergent validation metric, because these differences may not reflect a be true	

		signal and might just be caused by noise. In our opinion, models may be used to extrapolate to deployment areas or to extrapolate in time after initial calibration on a data set covering the full area and at least 2-3 years (see our comments later), but we do not support credit issuances based on models where the data do not support the same claim.	
Calibration: multicomponent approach and parameter selection	5.4 Calibration Requirements for Process-Based Models	*[For example, a model calibrated against porewater Ca2+ should be able to reproduce porewater Mg2+, pH, and alkalinity within acceptable performance bounds using the same parameter set (i.e., multi-parameter optimization).]* The problem at hand (i.e., weathering of feedstock and propagation of weathering signal across soil profile) is inherently a multicomponent one, so no model should be independently calibrated against a single geochemical component in the first place	Changed. Thank you for the comment. We agree that weathering and the propagation of the weathering signal through the soil profile are inherently multicomponent, and a model should not be calibrated against a single geochemical component. Our example unintentionally implied single-component calibration; this was not the intent, and it was inconsistent with the multi-parameter optimization recommended elsewhere in Section 5.4. We have revised the text accordingly.
	5.4 Calibration Requirements for Process-Based Models	*[Project Proponents must use published literature and laboratory values or project-specific experiments determined values for the relevant mineral phases.]* Where project-specific experiments are used, how will they be evaluated, particularly if the data is unpublished?	No change. Thank you for your comment, we consider project-specific lab experimental data as an acceptable data source, they will be reviewed the same as published literature values.
Calibration log verifiability	5.4 Calibration Requirements for Process-Based Models	*[The complete calibration record, including the full parameter trajectory. For automated methods, this includes the parameter estimation log. For manual methods this may include a log of each adjustment step with the motivating observation is acceptable.]* This will be very hard to verify. During model development, developers will test different parameters for calibration and revise the model iteratively. A full calibration log will in many cases not be available and if a calibration log is submitted, there is no way to test whether this is correct and complete.	Changed. Thank you for the feedback. It is a good point that a complete calibration trajectory is frequently not preserved during iterative model development, and that a submitted log cannot be independently confirmed to be correct and complete. We have revised the requirement to focus on what is verifiable. For automated calibration methods, Project Proponents must now provide the calibration code, objective function, parameter bounds, and solver settings such that Isometric and the VVB can re-run the calibration and reproduce the reported parameter values, replacing reliance on an attested log with direct reproducibility. For manual calibration, which cannot be independently reproduced, we have removed the requirement for a

			step-by-step adjustment log and instead rely on the narrative documentation of adjustments (retained), the cross-variable consistency requirement, and the held-out validation checks in Section 8, which together are the operative safeguard against overfitting regardless of calibration transparency.
Identifiability and equifinality of process-based models	5.4 Calibration Requirements for Process-Based Models	The level of process richness permitted dictates a long list of parameters, many of which remain unconstrained by empirical data (to highlight some intrinsically unconstrainable ones: field roughness, particle size evolution, surface passivation). This leaves a large number of unconstrained degrees of freedom, and the challenges with non-uniqueness / identifiability / equifinality will be immense: there will likely be many combinations of parameters that fit the calibration and validation data well but drive starkly different CDR predictions on deployment plots. The Hanford uranium case studies funded by DOE/NRC are informative, where no amount of data acquisition could produce truly identifiable models (Davis et al. 2004). With on the order of thirty observations in a monitoring plot, a process-based model may be strictly underdetermined, with an effectively infinite number of parameter sets providing an exact fit to the data. The measurement cap in §8.3 bounds the consequence of this but does not resolve the identifiability problem itself.	No change. Thank you for these comments, which raises a genuine and, at the project level, largely unresolvable challenge. We agree that an underdetermined process-based model can have multiple parameter sets that fit the calibration and validation data while diverging on deployment plots. The Module does not attempt to produce a uniquely identifiable model; its goal is to bound the consequences of non-uniqueness and reject models whose non-uniqueness materially affects the credited estimate. Several requirements act on this together. A global sensitivity analysis must identify influential parameters, including those that cannot be independently constrained, and these must be carried as sampled distributions in the Monte Carlo propagation. This way, equifinal solutions that diverge on deployment plots widen the credited uncertainty. Furthermore, model eligibility is set by out-of-sample performance, not calibration fit: models must pass spatial cross-validation and a deployment-area validation against measurements. Crediting is then taken at the conservative 16th percentile and further capped based on directly measured CDR in the matched monitoring area, so the model cannot exceed what measurement supports. As you note, these measures bound the consequence of non-identifiability rather than resolving it, which is its intended role. Structural uncertainty across defensible formulations is acknowledged as a field-wide limitation to be addressed through multi-model comparison as the literature matures.
	5.4 Calibration Requirements for Process-Based Models	The calibrated-parameter-to-observation ratio is reportable but unbounded; consider an explicit ceiling.	

			On an explicit ceiling for the parameter-to-observation ratio: the ratio must be reported, but we have chosen not to set a ceiling. A defensible ratio depends on model structure and on how independent the observations are (autocorrelated time series provide fewer effective constraints than their count suggests), so a fixed threshold would be arbitrary across model types. Over-parameterization is instead penalized through out-of-sample validation and uncertainty propagation, with the reported ratio available to reviewers as a screening signal.
Global sensitivity analysis	5.5 Sensitivity Analysis	We support the requirement for a global sensitivity analysis as much more robust approach than OAT. For the feature-rich, nonlinear models contemplated here, this is the technically correct standard, and it is one that is uncommon in existing crediting methodologies. As a further step, the parameters identified as most influential might be required to receive either direct measurement constraint or conservative substitution, not merely reporting.	No change. Thank you for the supportive comment, and the global sensitivity analysis is retained in the Module. The suggested further step is already in place in the Module: parameters identified as both influential and underdetermined must receive ensemble treatment rather than a single point value, while an influential but well-constrained parameter must be supported by a documented literature bound.
Soil organic matter representation	6.1 Introduction	*[agronomic soil measurements]* Soil organic matter is a fundamental soil property that influences multiple processes directly relevant to enhanced weathering. It affects soil pCO_2 production through microbial respiration and organic matter decomposition, contributes substantially to cation exchange capacity (CEC), and influences hydrological properties such as water retention, hydraulic conductivity, aggregation, and preferential flow. Despite its importance, soil organic matter does not appear to be explicitly considered in the current Module. This omission may limit the ability of the framework to accurately represent key controls on weathering reactions, solute transport, and alkalinity export. Therefore, soil organic matter should be included as a key parameter in the Module, or the rationale for its exclusion should be clearly justified.	No change. We thank the reviewer for this comment. While we appreciate the secondary impacts soil organic matter has on the weathering process, it is not a required parameter because it is not used as a direct input to any modeled terms.

Training data: representativeness and acceptable sources	6.3 Training Data Requirements	*[Hydrology]* can you be more specific on what hydrological parameters should be considered	Changed. Thank you for the comment. We have clarified that 'hydrology' in this context refers at minimum to slope and drainage class, consistent with Section 4.2.3 (Climatic and Hydrological Representativeness), and have added a cross-reference.
	6.3 Training Data Requirements	*[The dataset must include direct measurements from the project area of parameters relevant to enhanced weathering and CO2 removal calculation, at minimum spanning the following required measurements (see Section 9 for more details):]* For both approaches, where there are available remote sensing sources or literature values that are established to be well representative of that variable in that location e.g., SoilGrids for Soil Texture, those data should be allowed to be used.	No change. Thank you for the feedback. We will not allow literature-derived values at this time, but we will keep this under review.
	6.3 Training Data Requirements	*[Data-driven models must be trained on high-quality, representative data]* It would be really useful to have a functional definition of both high-quality and representative. This could even be at the protocol or standard level.	No change. Thank you for the comment. This is already covered by the numbered list that follows: point 2 addresses data quality and point 3 addresses representativeness.
	6.3 Training Data Requirements	*[Concentrations of immobile tracers and mobile cations in the soil for calculating CO2 removal.]* The quantity you are looking for is the cation loss, not the underlying concentrations. Recommend allowing for practitioners to train models on weathering constraints rather than raw concentration values.	Changed. Thank you for the comment. We have revised this to specify that the required training target is the weathering or cation-loss signal, which may be derived either from concentrations of mobile cations and, optionally, immobile tracers (the mass-balance approach) in the soil, or from porewater-based measurements (e.g., DIC, alkalinity), consistent with the measurement types recognized in Section 4.
	6.3 Training Data Requirements	*[The dataset must be representative of the full project area and range of conditions over which CDR is being quantified.]* This is very vague. Please specify how representativeness needs to be tested.	Changed. Thank you for the comment. This has been modified to refer back to Section 4.2 (Monitoring Area Representativeness), which details how the Area of Applicability framework is used to determine representativeness.
Model performance	6.4 Model	*[Project Proponents are therefore recommended to use mean squared*	Changed.

metrics and goodness-of-fit thresholds	Development Requirements	*error (MSE), mean absolute error (MAE), or mean average percent error (MAPE) in their regression algorithm.]* This is non-specific. There should be rules for which goodness-of-fit (GOF) criteria are required. It might even make sense to set lower limits on certain performance indicators. If this is not the case, there is no quality control on model performance and there is too much leverage for Isometric to decide which models are accepted and which ones are not, leaving a door open for potential bias towards certain project developers.	Thanks for this comment. The training objective function (MSE/MAE/MAPE) is left to the Project Proponent because the right choice depends on that project's own data characteristics (e.g. outlier sensitivity, presence of near-zero targets), and we don't think a single mandated objective function would suit all EW projects. We have provided more guidance around the choice of objective function in the updated text. Model acceptance is instead governed by fixed, non-discretionary thresholds in Section 8 (bias relative to measurement uncertainty, 90% CI coverage of held-out data, and deployment RMSE relative to cross-validation RMSE), which apply regardless of which objective function was used in training. These are computed from the submitted data itself, not selected by Isometric per project.
	8.3 Validation Metrics and Evaluation of Model Performance	*[The choice of metrics must be justified and must include, at the minimum, the following:* *Root mean squared error (RMSE)* *Mean bias (mean signed error across all held-out predictions)* *Mean absolute percentage error (MAPE), with documentation of how near-zero observations are handled* *A 1:1 scatterplot of model predictions against observations, with the line of perfect agreement shown]* This contradicts 5.4, where the choice is free between MSE, MAE, or MAPE. We think that lower thresholds for model performance might make sense.	Changed. Thank you for this comment. The validation metrics that are required to be reported on do overlap but are distinct from the objective function used to train data-driven models. We have updated the text to make this clearer.
Cross-validation method: standardization and definitions	8.1 Cross-Validation	*[The monitoring area is partitioned into spatial units, such that there are at least three units. Where the monitoring area comprises multiple monitoring sites, each site is a unit. Where the monitoring area comprises one or two monitoring sites, or where sites are spatially extensive, the site(s) must be partitioned into spatial blocks. Block size must be at least equal to the spatial autocorrelation range determined from the empirical variogram of the modelling residuals.]* Block and unit are used interchangeably.	Changed. Thank you for the feedback. This change has been made.

		Please make them consistent and define them at the start	
	8.1 Cross-Validation	*[Cross-validation must be performed, and the following method is recommended:]* The exact method for cross-validation should be a requirement standardized for all projects and models. Since not all projects are designed in the same way, there can be options, for example: if only one T/C unit is monitored, use method A; if the number of monitored T/C units is at least 3, use method B. However, for every project design the method of cross-validation should be pre-defined, this would ensure consistency across projects and that validation is selected before data analysis begins.	No change. Thank you for the feedback. While we understand the point around consistency, we intend to keep the exact cross-validation method as recommended for now. We may decide to upgrade the exact methods to be required once we have more real data and models from projects using this Module.
Crediting threshold: p50 cap vs v1.2 p40	9.3 Monte Carlo Uncertainty Propagation	*[50th]* Under ERW v1.2, even in the best case (i.e., with successful liquid validation), you can only credit at 40th percentile. Why here you can use 50th percentile?	Changed. Thank you for the comments. We've added language to make the crediting logic and the role of the cap explicit. To clarify: crediting under this Module occurs at the 16th percentile of the modeled CDR distribution over the deployment area. The 50th-percentile-of-the-treatment-plot value is a separate upper cap layered on top of that. By construction, it can only reduce the credited value below the model's p16, never raise it. The cap is a different construct from the p30/p40 crediting percentile in EW v1.2, because this Module and the EW v1.2 Protocol quantify the deployment area CDR differently. Under EW v1.2, geochemical measurements are collected across both the treatment and deployment areas and used to quantify weathering directly wherever it is measured. EW v1.2 never infers deployment weathering from the treatment plot. This Module instead uses models to extrapolate the weathering signal from small, densely-measured monitoring sites to larger deployment areas that are not required to be geochemically measured throughout. The p50 cap has no counterpart in v1.2; it's an additional constraint that ties the modeled deployment estimate back to what was directly measured in the matched
	9.3 Monte Carlo Uncertainty Propagation	This approach may permit a higher crediting upper bound than the measurement-based v1.2 approach (which credits up to the 40th percentile on a passed check) and sought justification.	
	9.3 Monte Carlo Uncertainty Propagation	Crediting at p16 (≈ −1σ) is reasonable for parameter and input uncertainty but is contingent on the uncertainty budget being complete. Given the structural uncertainty the framework does not fully capture (§8.1), p16 is not nearly as conservative as the language suggests. Nonetheless, we regard the crediting rule as an elegant and effective safeguard. Taking the credited value at the 16th percentile of the Monte Carlo distribution, subtracting any positive mean bias identified in cross-validation, and capping the result at the 50th percentile of the matched treatment-plot measurement anchors model output to direct measurement and meaningfully bounds the upside available to an overtuned model. We could see this resulting in widespread de facto crediting at the mean of measurements,	

		though, which could result in critiques of insufficient conservatism.	treatment plot. We'd also note that matched treatment and deployment areas are similar but not identical. Weathering rates can differ for physically real reasons (e.g., slight differences in pH), and extrapolating across such differences is exactly what the model is for. The cap therefore uses the treatment plot as a conservative proxy while acknowledging the deployment rate may legitimately differ, so it functions as an added safeguard against over-crediting rather than as a higher crediting percentile.
	9.3 Monte Carlo Uncertainty Propagation	*[the credited CDR value must not exceed the 50th percentile calculated from measurements in the matched Treatment plots, following the quantification approach outlined in the EW 1.2 Protocol.]* Allowing for crediting at p50 of the measured treatment provides an avenue for less conservative crediting than standard measure <> re-measure. Is this intended? We recommend reconciling this with the primary ERW module, lest this become a crediting loop-hole. We also note that the uncertainty discount should be tied to model accuracy as well as model precision.	
Transparent documentation of deviations and justifications	General	Any deviations from standard requirements, and their justifications (by the Project Proponent, Isometric, or the VVB), should be transparently documented.	No change. Thank you for the comment, we agree on the importance of transparently documenting any deviations. Our standard practice is that any deviations from Protocols and the justifications for them are documented in the publicly-available verification reports, and we are working on making these explanatory notes even clearer moving forward.
Reporting period definition and minimum duration	Reporting Period Definition	Is there a minimum Reporting Period duration under the protocol?	No change. This Module falls under the Enhanced Weathering in Agriculture Protocol, which does not have a minimum duration for a Reporting Period provided that all the verification requirements are met at each Reporting Period.
	Reporting Period Definition	Suggest specifying that the first-Reporting-Period Deployment Area Validation span at least a year to capture seasonal variation.	No change. Thank you for the comment. The Reporting Period is defined in the EW 1.2 Protocol and the Isometric Standard, and this Module inherits that definition rather than setting a separate minimum duration. The seasonal-variation concern is addressed at the model-validation stage. A model must be initially validated on at least 12 consecutive months of data capturing the full annual cycle before it can be used for crediting. Because that seasonal

			robustness is a property of the validated model, it holds for any Reporting Period period length the model is applied to (provided the temporal translation criteria are met). For that reason, we have not added a minimum duration to the Deployment Area Validation. The Deployment Area Validation is a supplementary check on the matching and extrapolation, and it spans the Reporting Period to which it applies and a separate one-year minimum on this check is not necessary.
Minimum data set size for modeling		As written, the module does not specify a minimum data set size either for training or validation. While benchmarking of individual process components provides an additional guardrail for process-based models, we think it is prudent to specify a minimum data set size, especially for data-driven models, to ensure they are not overfit to a small dataset. Specifying this minimum may be out of scope for this consultation process, and we would encourage Isometric to do this in collaboration with the broader ERW community, drawing inspiration/inference from other fields. An example approach we see that could address this is as follows: Training Dataset Minimum 10–20 paired observations per estimated model parameter based on a common regression rule of thumb (Harrell, F. E. (2015). Regression Modeling Strategies (2nd ed.). Springer.) Note this isn't saying that linear regression must be used, but ultimately this is a regression problem, so regardless of model architecture, this is a helpful heuristic. The relevant count is estimated parameters, not raw input features. A model with ~10 input features (soil pH, texture, CEC, base saturation, application rate, temp, precipitation, feedstock mineralogy, PSD, time since application) will typically expand well beyond 10 parameters once non-linear terms (splines or polynomials, which each add degrees of freedom per continuous predictor) and a few interactions are included, giving on the order of 30–50	Changed. Thank you for sharing, we agree that specifying a minimum data set size is out of scope for this consultation process and would be better suited as a best practice guidance developed with the broader EW community. We expect that models trained with insufficient data will not generalize well enough to pass the required validation checks in the Module, but this is something we can revisit in a future Module update. We have added language clarifying that no minimum dataset size is required, but this may be revisited in a future update as the research evolves in this area.

		effective parameters. At 10 observations per parameter, that implies roughly 300–500 paired observations as a practical minimum. In general, we recommend 500 to 1000 paired observations, distributed across X independent spatial blocks to prevent spatial autocorrelation, and recommend datasets that have paired observations stretching over multiple years (≥2 years). Drawing inspiration from SOC validated models in the CAR protocol, calibrated/validated models on the larger end had about ~1000 observations (DayCent-CR v1.1.0 used ~1,197 SOC pairs (plus ~1,381 N_2O pairs) and DNDC v11.0 used ~914 SOC + ~1,271 N_2O + 76 CH_4 pairs), while other used less. That being said, Foundation Models should require fewer observations than the above recommendation. There are examples of FMs being able to generalize with even just a few examples, so rather than prescribe a minimum here, we would again recommend the 300-500 paired observations minimum (note this is a recommendation rather than a requirement, unlike the Data Driven from scratch, as we could see in the future one-to-few shot learning from better general models). We would like to encourage the use of training data beyond just ERW that can help with weathering + CDR prediction (e.g. natural weathering, mesocosm experiments, other environmental/subsurface tasks). As we have learned from the scale up and training of LLMs, training on as diverse and vast data as possible helps with performance over time, and given the current limits of available ERW data using as much data as available should be encouraged by learning from similar processes. Paired observations can be derived from these sources. Validation Dataset A minimum of 100 paired observations, or at least 10% of the Training Dataset Size.	
Language around sorption and strong acids		As written, the module does not capture all of the complexity around sorption processes. The module does not distinguish between ion exchange and surface complexation, and these processes represent distinct ion pools	Changed. Thank you for the feedback. We have made the recommended language change throughout the module where it makes sense, and have included

		that respond to changes in soil conditions differently. Cation exchange is an electrostatic process in which cations are held to balance the pH-independent, fixed negatively charged on clay minerals like smectite and illite, with retention governed by charge and concentration. On the other hand, surface complexation is a chemical process by which ions bond to the variably charged reactive surface sites of metal oxides or organics, which can gain or shed protons as pH changes. Cation exchange is expected to dominate in younger, clay-rich soils, whereas surface complexation is anticipated to dominate in highly weathered, oxide-rich soils. Soils in most systems sit on a continuum where both processes operate and thus must both be represented. This mechanistic distinction matters for CDR accounting given that acidity in soil can be opened up with an increase in pH via these two processes. As more commonly recognized, the displacement of 1 mole of aluminum from the exchangeable phase can yield 3 moles of protons in solution through hydrolysis. In a different manner for surface complexation, ≥ 1 mole(s) of proton(s) can be released from reactive surface sites depending on its polyproticity. Ascribing empirical constraints exclusively to a single mechanism in a model without considering the balance between different types of reactive surfaces can result in erroneously estimating the impact of sorption and thus soil acidity on alkalinity generated by weathering. We are actively working on developing a means to appropriately parameterize these mechanisms in a reactive transport framework based on observations common to ERW deployments and would appreciate for this module to be revised to a state that appropriately reflects both processes. We suggest that the pragmatic near-term solution is to broaden the language used. Wherever this module (and ideally in other modules and the overarching ERW protocol) specifies “ion exchange”, it should read as “sorption”, defined to cover both cation exchange and surface complexation. This update preserves flexibility to represent surface complexation associated with variably	surface complexation on variable-charge surfaces as a recommended process to represent under the Sorption section.

		charged surfaces where it would matter without mandating full parameterization where it would not. Moreover, such a change would readily align the process with language this module and the ERW protocol already uses in "CO2eNetSorption".	
General / uncategorized	5 Process-Based Models	Suggest requiring, rather than recommending, that models be recalibrated every reporting period or at a regular interval.	No change. Thank you for the comment. We don't think a blanket requirement to re-calibrate the model at a set interval is necessary, since it may be the case that a model is still performing well and would not benefit from re-calibration. Instead we require re-validation at every Reporting Period, and the purpose of that is to assess whether the model still performs adequately or needs to be updated. Where a model fails re-validation, that is when re-calibration comes in so that under-performing models are caught and corrected, without mandating re-calibration when it isn't warranted.
	5.7 Reporting Requirements	*[Model updates: It is recommended to update and recalibrate the model with each additional Reporting Period. If the model's predictive performance improves, the uncertainty discount can be reduced for future credit issuances. Project Proponents must provide a model update plan.]* We think that it should be required to update the model every reporting period. Weathering and cation leaching are non-linear reactions and calibration & validation on early data may not be sufficient to represent long-term system behaviour. The model should be repeatedly validated throughout the project lifetime. In the case that validation for a later timestep is unsuccessful, the model should be re-calibrated on all data. If estimates are lower for the revised parameter set, there should be a mechanism to account for over-crediting in previous reporting periods.	Changed. Re-validation is already required for every RP, with re-calibration if a model fails. We have updated the wording to make this more clear: the model must be re-validated against newly collected data at every Reporting Period, and re-calibration is required where a model fails re-validation. Further proactive model improvements beyond those requirements to improve performance are recommended.
	8.4 Re-Validation and Model Updates	*[The model must be initially validated on at least 12 consecutive months of data to capture the complete annual cycle of seasonal variation. Models must then be re-validated at each Reporting Period,*	

		using updated data collected since the previous Reporting Period.] Re-validation is optional in 4.7. Please clarify if this is a recommendation or requirement.	
	8.4 Re-Validation and Model Updates	*[The model must be initially validated on at least 12 consecutive months of data to capture the complete annual cycle of seasonal variation.]* What does 12 months of consecutive data mean? ERW measurements by definiton are sparse/apply over a period of time. Can this be changed to a period spanning 12 months (e.g. you have a measurement at month = 0 and month = 12)?	Changed. We have added the following clarification to this requirement where it first appears in the Applicability section: "The requirement for 12 consecutive months applies to the continuous forcing data, such as temperature and precipitation, which must cover the full annual cycle without gaps. Soil geochemical measurements are sparse by nature and are not required to be continuous."
		Similarly, while we are very supportive of requiring validation against both dedicated monitoring sites and the full deployment area, we would push back on the requirement that the model be re-validated every reporting period in the case that models are calibrated/validated on sufficiently long time horizon datasets for a given deployment domain. Specifically: As written, the module requires 12 months of continuous data collection for model development to ensure that all extremes of the deployment conditions are captured. The model is then re-validated annually to ensure continued applicability. However, this does not address the situation in which a longer period of data collection was used for the initial model development. In this case, we propose that the re-validation period occurs at the median calibration/validation dataset length (for example, 3 years if the model was trained on 3 years worth of data), and suppliers be allowed to credit at a conservative floor for interim reporting periods. This could be enabled with a forward-looking true-up mechanism at the validation reporting period. This incentivizes long-duration, robust dataset generation in targeted deployment domains while enabling project financeability in the near-term. We would propose that this longer re-validation period apply to both monitoring site and deployment	Changed. Thank you for the comment. We recognize that re-validation every Reporting Period may be more frequent than necessary for models developed on longer, multi-year datasets. However we prefer to keep the existing requirements of re-validation every RP because this serves as a check on model performance using truly held-out data that was not present during the model development stage. We have however added language acknowledging that the re-validation cadence may be relaxed in a future Module update as the evidence base matures, which could be for instance demonstrating over multiple consecutive Reporting Periods that the model continues to reproduce independently measured CDR without material drift over time.

		validation. We note that in this scenario, if an interim RP/credit issuance happens before a revalidation is required on the monitoring site, it is worth thinking about how model bias might have to be adjusted to not have it overcompensate.	
Editorial: labels, links, references, phrasing	1.0 Introduction	*[Appendix X]* Missing appendix number	Changed. Thank you for flagging these, they have all been fixed.
	1.0 Introduction	*[spatially limited measurements]* This is a nitpick, but measurements aren't just limited spatially. Suggest rephrasing to just 'limited measurements'.	
	2.0 Applicability	*[River and Ocean Losses Module (link)]* No link	
	4.1 In-Field Monitoring Approach	*[Tables X and Y]* Labels needed	