CLIMATE
ACTION
RESERVE

Model Calibration, Validation, and Verification Guidance

For Soil Enrichment Projects

Draft for Public Comment

August 2020

Table of Contents

1 Introduction 1

2 Model Calibration 2

2.1 Guidance on Model Calibration using Frequentist Approaches 4

2.2 Guidance on Bayesian Methods for Calibration, Validation, and Error 4

3 Model Validation 6

3.1 Declare Practice Categories Requiring Evaluation 6

3.2 Define the Project Domain 7

3.2.1 Declare Project Crop Functional Groups 7

3.2.2 Declare Project Land Resource Regions 8

3.2.3 Declare Project Soils 8

3.3 Gather Validation Data 8

3.4 Assessment of Bias for Each Practice Category 10

3.5 Linking Validation Data to Model Prediction Error 11

3.6 Reporting on Model Validation 11

4 Substitution for Missing Crop Types 13

5 Verification of Model Usage 13

6 References 15

1 Introduction

To be validated, a model must show acceptable fit and a lack of bias or conservative bias when being used to estimate soil organic carbon (SOC) stock change and, if applicable to the project, flux change of N_2O and CH_4, to quantify soil enrichment projects. Model validation must be specific to the model being used in the project, as well as how the model is being used to make project quantifications (i.e., for what cropping system and biophysical conditions). Model validation must also make appropriate use of published experimental datasets to compare modeled predictions to real-world change. To facilitate this process, the steps laid out in this document 2must be followed to evaluate model bias and fit, and demonstrate meeting model validation requirements.

A discussion regarding verification requirements for the proper implementation of this guidance can be found in Section 5 3.6of this document, as well as Section 8.3.1 of the Soil Enrichment Protocol (SEP).

All stakeholders making use of this guidance should contact the Reserve to ensure they are using the most up to date version of this guidance. Project developers and verifiers must use the version of this document that is in place as of the first date of the relevant reporting period. In the cases where multiple reporting periods are being verified at once, project developers and verifiers should seek Reserve guidance and approval as to which version of this guidance should be applied to which reporting period.

Model validation must be documented in a validation report. Options and requirements for submitting validation reports are outlines in Section 3.6. Validation reports must show that all requirements for a specific project have been met, including proof that the same model version and parameter sets are used, and that all project domain and crop functional group/practices category combinations have met minimum requirements for model validation. Validation reports must either be independently assessed by an approved 3rd-party entity or accepted for publication in one of the peer-reviewed publications listed in Section 3.6. Model validation reports will be public documents.

For each subsequent monitoring report, as long as a project area remains constant, or is only expanded to include new fields that already fit within the validated project domain, the existing validation report can be used. If the project is expanded to new practice effects, new crop functional groups, or the model is changed (using new model calibrations, changed parameters/parameter sets, or a different model, following guidance described in Section 6.5 of the SEP protocol), the validation report needs to be revised, reviewed by an approved 3rd party entity, and re-submitted.

The overall flow of the requirements for the use of models is illustrated in Figure 1.1, below.

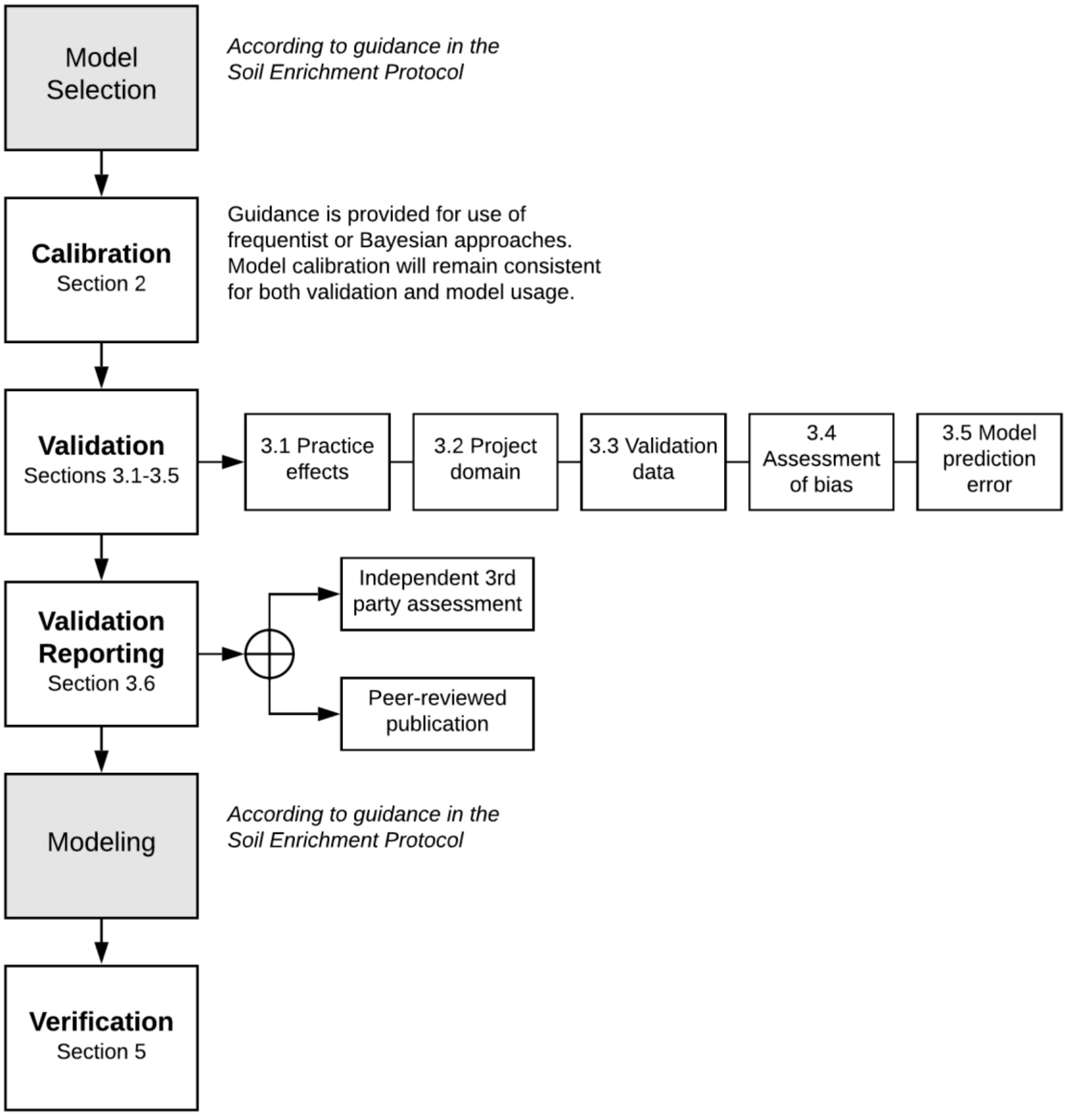

Figure 1.1 Steps related to the use of models for quantification in SEP projects

2 Model Calibration

Model calibration, parameterization and validation are often poorly defined concepts. For the purposes of this document, calibration is defined as any process involving the adjustment of parameters and constants within a model so that the model more accurately simulates measured values.

Model calibration is a variable and model-specific set of processes. Some examples include:

- statistical procedures to optimize rates of mass flow and the simulation of internal model pools (e.g. optimizing the allocation of daily net primary production to root growth to more accurately simulate observed root growth for a given crop);

- adjusting model parameters with directly measured values (e.g. setting the simulated fraction of plant residue left on the soil surface after a method of harvest using an average of observed values);
- 'tuning' a set of model parameters that may not be able to be measured directly using overall model performance and an understanding of model sensitivities (e.g. adjusting a constant downregulating the rate of soil biological processes under moisture-limited conditions using measures of soil respiration).

Deterministic models, where the same inputs always result in the same outputs, may have different calibration processes than stochastic models, which include random variability. Mechanistic models, which are based on mathematical representations of mechanisms within the modeled system, are more generalizable with fewer data than empirical models, which are based on statistical synthesis of observations and cannot be extended outside of where observations are available.

For any model used in a SEP project, model calibration must be a separate process and use separate datasets from model validation. Further, for either process the quality of measured datasets (i.e. rigor of the experimental design, accuracy of observations, applicability to the system that a model is being calibrated or validated to simulate) will determine the quality of the model, aka "garbage in, garbage out". Separation of datasets for model calibration and validation may be done manually, if it can be shown that both datasets are comparable in their applicability to the system that is being simulated. Separation may also be done statistically using a method like k-folding. For the purposes of this protocol, calibration and validation data should be demonstrably independent. This requirement can be met if datasets used for calibration and validation do not overlap in experimental research locations and are not taken from the same experimental study. If calibration and validation datasets do overlap in either experimental study or research location, independence between the datasets used for calibration and validation should be demonstrated at the crop functional group/practice category combination level (Section 3.2). For example, if root measurements in a tilled soybean/corn rotation experiment were used for model calibration, N_2O flux measurements from the same study could not be used as validation data for either the corn or the soy crop functional group and tillage practice effect combinations. However, if at the same research facility a demonstrably separate corn/soy rotation experiment occurred (separate in space or time, with separate experimental design or intention), those data would be permissible for inclusion in model validation.

The protocol does not prescribe a model calibration procedure. However, the calibration procedure must be reported to ensure model parameters and parameter sets were generated appropriately (see SEP protocol Section 6.5, item 3), as well as meet the requirements that

1. the parameter sets used when validating the model are the same used when the model is applied to simulate baselines and project practices, and
2. that the data used for model calibration and validation are separate.

It is encouraged for model parameters to be as generalizable as possible across the project domain, with minimal use of different parameter sets. However, it is acceptable for different parameter sets to be used as long as they are defined at scales no smaller than LRRs- i.e. the same parameter set is used for all simulations within a given LRR. Parameter sets must be declared for each project LRR, and should be used to simulate all crop functional groups and practice categories within that LRR. It is not acceptable to validate a model and then adjust model parameters when using the model to simulate project baselines and practices. All parameter sets must be validated following the guidance in section 3. If the minimums described in Section 3 do not result in all parameter sets being validated, and additional steps are not taken to validate all parameter sets, unvalidated parameter sets cannot be approved for use in the project.

Because biogeochemical models often contain a large number of parameters, different strategies can be employed to perform calibration. General guidance for frequentist and Bayesian approaches are provided in Sections 2.1 and 2.2.

Summary of model calibration requirements:

- The model calibration process and the datasets used for model calibration should be reported, including the experimental study and location where datasets were generated
- Calibration datasets should be demonstrably independent and separate from validation datasets.
- If any calibration and validation datasets overlap in experimental study or experimental location, their independence will need to be clearly demonstrated at the crop functional group/practice category combination level
- Once calibrated, the same parameter set can be used for multiple LRRs or for all LRRs in a project. Each LRR should have only 1 declared parameter set used for all model simulations within that LRR.
- To be approved for use in a project, all parameter sets must be validated for each crop functional group/practice category combination.
- Once calibrated and validated, model parameters may not be changed for with-project simulations.

2.1 Guidance on Model Calibration using Frequentist Approaches

Wallach et al. (2019) provide helpful guidance on common approaches to frequentist model calibration, including how to decide how many parameters to estimate, which parameters to estimate, whether to calibrate in stages, and how to avoid over-parameterization, i.e., where the model fits the data well but has poor predictive ability. Examples of model calibration are abundant in the peer-review literature and span a wide range of complexity and automation in their approaches (e.g. Bruun et al. 2003, Yeluripati et al. 2009, Liang et al. 2009).

2.2 Guidance on Bayesian Methods for Calibration, Validation, and Error

Model calibration can also be completed using Bayesian statistical methods, which apply a probabilistic approach to integrating existing knowledge and observed data (Wikle & Berliner, 2007). Bayesian statistical approaches are an emerging area of development in soil biogeochemical modeling. They typically require implementing Markov Chain Monte Carlo methods for sampling probability distributions. This can be computationally demanding with soil biogeochemical models, which can have dozens to hundreds or more parameters. Parameter values in these types of models can also be difficult to constrain, i.e., use data or existing knowledge to set limits on the range of values that a parameter may have, and define its probability distribution across that range. When there is little prior knowledge about a parameter value, ‘uninformative priors’ or ‘weakly informative priors’ are used to represent what is known or believed about the parameter. The resulting posterior distribution, or the distribution that represents the integration of prior knowledge and observed data, can be wide unless the observed data are strongly informative, i.e., have highly accurate and precise values. The following figure illustrates a strong prior belief (A) versus a weak prior belief (B).

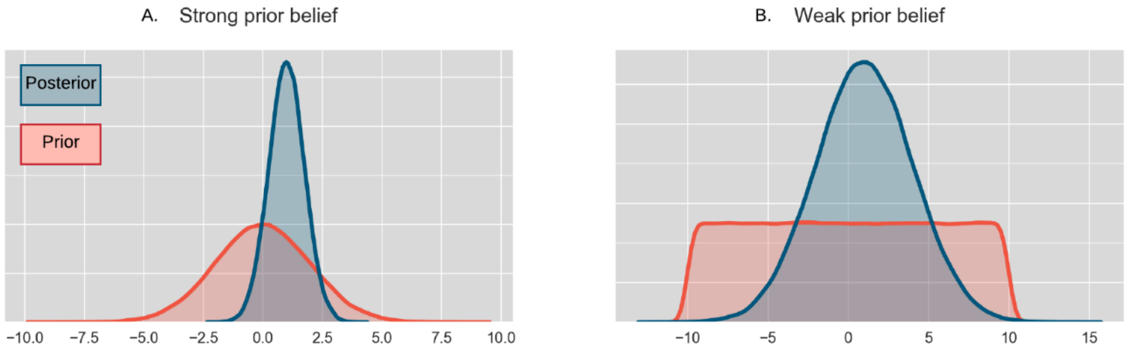

Figure 2.1 Comparison of prior and posterior distributions when there is strong prior belief (e.g. strong and consistent evidence and prior analyses, A), versus weak prior belief (e.g. weak or variable evidence or no prior analyses, B).

Across dozens or hundreds of parameters, Bayesian methods can be complex to implement and require large quantities of data. Despite these challenges, Bayesian methods provide a coherent mathematical framework to integrate diverse sources of information into model parameterization, as evidenced its central role in the developing field of Ecological Forecasting (Dietze, 2017), as well as in the Predictive Ecosystem Analyzer Project data-model integration system.[1] A Bayesian approach**Error! Reference source not found.** is encouraged for model validation and model prediction error, as the confidence intervals around model predictions will be directly based on the availability and variance of observed data. Figure 2.2 presents a conceptual workflow for a Bayesian approach to these analyses.

---

[1] Accessible at: pecanproject.org

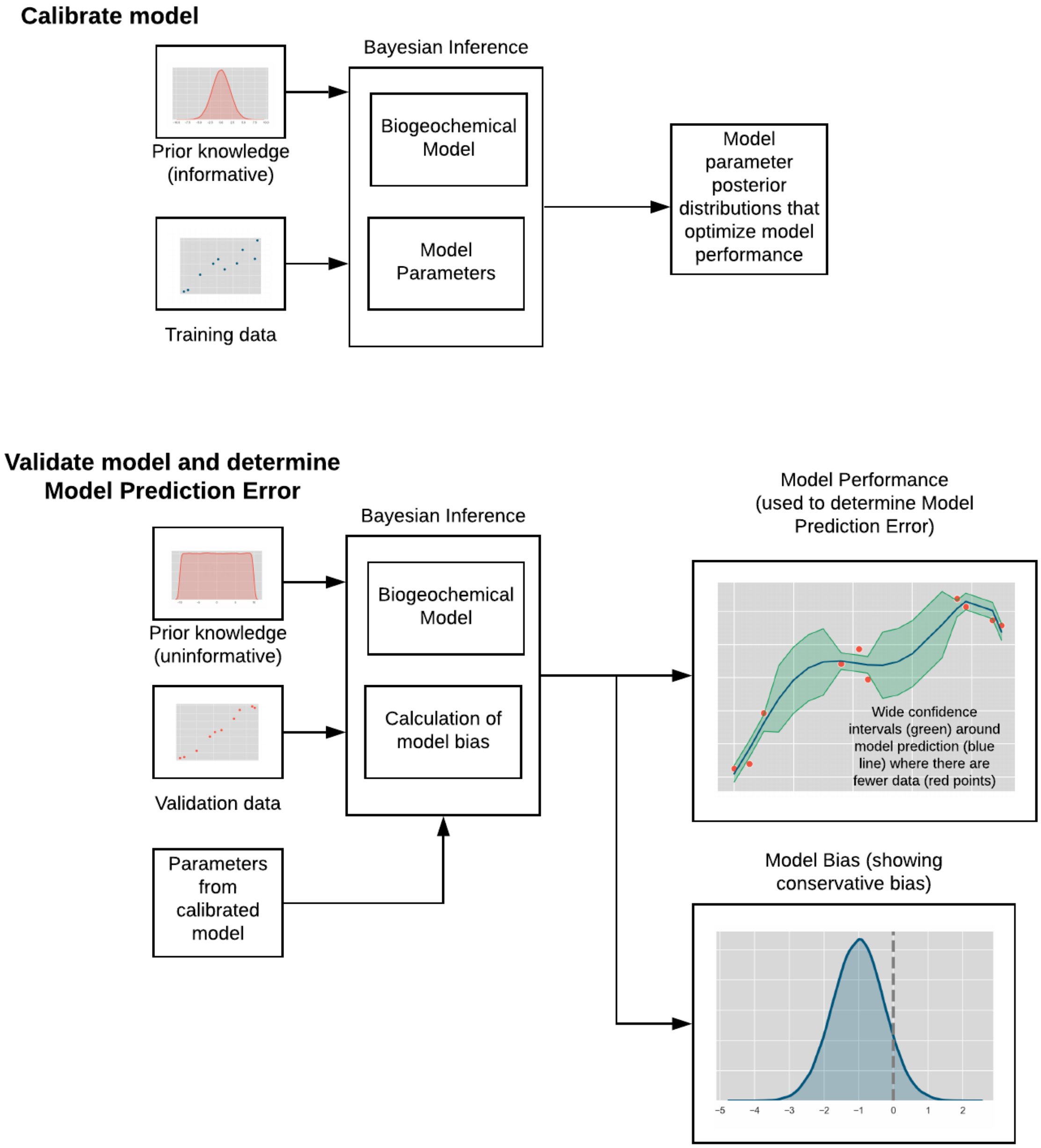

Figure 2.2 Conceptual framework for Bayesian approach to model calibration and validation.

3 Model Validation

3.1 Declare Practice Categories Requiring Evaluation

For every practice considered additional within the project, the model must be shown to have an acceptable fit and unbiased or conservatively biased representation of the underlying biogeochemical process governing the effect of that practice. To do so, each practice must be binned into the following categories to demonstrate the domain of practice effects and the categories requiring evaluation. Model validation of a practice category can use any practice effect in the category domain, evaluated using appropriate experimental data meeting requirements described below. Projects are encouraged to evaluate a range of practice effects in each practice category domain.

Table 3.1. Practice Categories and their Associated Practice Effects Requiring Biogeochemical Performance Evaluation

Practice Category Requiring Evaluation	Domain of Practice Effects
Inorganic nitrogen fertilizer application	Magnitude, form, timing, or method for nitrogen fertilizer applied, with form encompassing inorganic N fertilizers, and method encompassing surface, subsurface, or irrigation-based application
Organic amendments application	Magnitude, form, timing, method or variation in C:N ratio for organic amendments applied. Forms include and are not limited to biochar, mulch, compost, and manure, and methods encompass surface, subsurface, or irrigation-based application
Water management/irrigation	Magnitude, timing, source or method of irrigation water applied
Soil disturbance and/or residue management	Soil disturbance including tillage and compaction, , residue management encompassing soil exposure after harvest
Cropping practices, planting and harvesting (e.g., crop rotations, cover crops)	Variety of crops grown, increasing crop rooting depth, may include cover crops and soil preparations such as changing soil pH through liming
Grazing practices	Any of the following: presence/absence of grazing, stocking density, forage type or quality, species of grazers, mixed or single species herds, loading weight, grazing time, and rest/recovery periods

A project developer must declare all practice effects requiring evaluation for the project.

3.2 Define the Project Domain

For each practice category declared in the project description, the model must be evaluated in terms of its fit and bias in estimating emissions reductions. Evaluation of each category begins with defining the project domain in terms of its biophysical attributes. Specifically, the project developer must declare the unique crop functional groups, land resource regions, and soil attributes associated with each declared practice category.

3.2.1 **Declare Project Crop Functional Groups**

Crop functional groups for each practice category must be declared. Individual crop types can be grouped into functional groups across crops sharing unique combinations of the following attributes:

- N fixation (Y/N),
- annual/perennial (A/P) (defined in accordance with the NRCS Conservation Compliance categorization of crops[2]),
- photosynthetic pathway (C3/C4/CAM),

---

[2] Resource can be found here: https://www.nrcs.usda.gov/wps/portal/nrcs/detail/national/programs/farmbill/?cid=stelprdb1262733

- tree/shrub/herbaceous (trees and shrubs have woody plant growth, versus herbaceous species that do not grow woody plant material),
- flooded/not flooded

3.2.2 Declare Project Land Resource Regions

The full list of land resource regions (LRRs) associated with each practice category must be declared.[3] LRRs represent distinct combinations of climate, land resource use, and geographic features. For regions outside the US, IPCC climate zones must be declared for each practice category.

3.2.3 Declare Project Soils

Soils are to be declared for each practice category in terms of (1) soil textural class and (2) the associated clay content[4] of that class. NRCS soil texture classes include: sand, loamy sand, sandy loam, loam, silt loam, silt, sandy clay loam, clay loam, silty clay loam, sandy clay, silty clay, and clay.

3.3 Gather Validation Data

Requirement 1: Validation datasets for each declared practice category and crop functional group combination from Section 3.1 must include measurements for each modeled quantity, where the modeled quantity is the change in the flux of emissions to the atmosphere for SOC, N_2O, and/or CH_4 that results from the adoption of any practice associated with that effect. Datasets may include individual practice categories as well as combinations of practice categories (e.g. “stacked” practices), provided the practice category in question is experimentally varied and measured within the study. Some hypothetical examples of acceptable experimental treatments to evaluate practice categories are given in the following table:

Table 3.2. Examples of Acceptable Experimental Treatments to be Used in Evaluating Practice Categories

Practice Category	Experimental treatment
Inorganic nitrogen fertilizer application	Comparison of two different application rates of urea
Soil disturbance and/or residue management	Comparison of conventional tillage using moldboard plow to strip tillage.
Cropping practices, planting and harvesting (e.g., crop rotations, cover crops)	Comparison of single-crop rotation to double-crop rotation; comparison of no cover-crop to with cover crop.

Validation data must adhere to the following guidelines:

---

[3] Resource can be found here: https://www.nrcs.usda.gov/wps/portal/nrcs/detail/national/technical/nra/nri/?cid=nrcs143_013721

[4] See Table A-1 for clay contents of NRCS soil textural classes. https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/?cid=nrcs143_014055

- Measured datasets must be drawn from peer-reviewed and published experimental datasets with measurements of SOC stock change (and annualvmeasures of N_2O and CH_4 change if applicable) using control plots to test the practice category. All dataset sources must be reported. The same measurement dataset sources can be used for validating multiple practice categories, when appropriate. Datasets may be used from studies outside of the US. However, the associated IPCC climate zone where these datasets were collected should be shown to be relevant to LRR's in the project domain in terms of climate characteristics (mean annual precipitation, mean annual temperature).
- In the case of SOC stocks, repeat measurements of SOC stock change must be statistically robust to capture multi-year changes, as practice effects on SOC may combine short and long-term changes in soil biogeochemical processes. Statistically robust measurements from paired fields leveraging space-for-time analysis methods that approximate multi-year changes may also be used for SOC validation. Newer methods for SOC stock monitoring are becoming available that can observe changes with greater precision at shorter time intervals. These methods will be acceptable if there is peer-reviewed support or independent expert support approved by the Reserve for their use in SOC monitoring and demonstrate statistically robust evaluation of multi-year impacts on SOC stock changes. Measured datasets of SOC stock change may be made at any depth, but the model must also predict SOC stock change at the corresponding depth. Thus, a fully compiled validation dataset may contain different depths for SOC stock change measurements as long as the model is predicting SOC stock change at each corresponding depth.
- In the case of N_2O and CH_4 flux, any combination of measurements from chambers and/or eddy covariance flux towers are acceptable. Event-based or sub-annual measurements of N_2O and CH_4 must be aggregated to the annual timescale for use in validation using appropriate methods (e.g., Mishurov & Kiely, 2011; Turner et al., 2016).
- Datasets can be drawn from a benchmark database maintained by a third party, approved by the Global Soils Partnership (or comparable). The use of datasets from a benchmark database should include full citation of the database as well as a description of how datasets were extracted.
- Project developers are expected to use a process for selecting data for model validation that results in the assembly of validation datasets that are representative of the range of peer-reviewed observed results. Project developers must describe the methods, selection process, and data manipulations used to create the dataset applied in the model validation process. This includes describing search terms and databases used to identify available datasets, criteria used to select dataset sources, origin of extracted data (e.g. figures, tables, databases with DOI), original units of data and data uncertainty, and data manipulations used to convert original units into the units described above. The project developer should report the number of validation data measurements of each data type (SOC, N_2O and CH_4) for each project domain combination of practice category and crop functional group, and include a histogram showing the range of validation data values. In the case where validation data are unevenly distributed across the project domain, the method used to link validation data to model structural error (described in more detail in Section 3.5 below) should demonstrate that it addresses the discrepancy.

Requirement 2: Validating a practice category / crop functional group combination can only be completed at the scale of an individual sample/field if there are measurements of SOC stock and annual N_2O and CH_4 flux change (if applicable) meeting the above criteria that in total include the same soil textural class, or one that is within 30% of the same clay content, as well as the crop functional group and LRR relevant to that location.

Requirement 3: Validating a practice category / crop functional group combination for the entire Project Domain can only be completed if there are measurements of SOC stock and annual N_2O and CH_4 flux change (if applicable) that in total cover:

- At least three declared LRRs for projects within the US (or two IPCC climate zones per each required LRR for projects outside of the US)
- At least three declared soil textural classes
- A range in declared clay amount per unit of soil spanning at least 15 percentage points

It is in a project's interest to exceed these minimums and validate the model across more LRRs, soil texture classes, and clay contents, as model prediction error should use the same dataset as model validation, and will penalize the use of few data points (see Section 3.5). If the number of declared LRR's is less than 3, then all declared LRR's must have the above measurements to validate a practice category / crop functional group combination for the entire project domain.

Note that all model parameter sets must be validated for each crop functional group / practice category combination (see Section 2). If model parameter sets vary by LRR this may require additional measurement datasets beyond the minimum described above to ensure all parameter sets are validated.

3.4 Assessment of Bias for Each Practice Category

For each practice category declared in Section 3.1, the model must be shown to be unbiased or conservatively biased in estimating the change in SOC, N_2O, or CH_4 pools for the project domain defined in Sections 3.2, using measured data that meet the requirements of Section 3.3. This is done using the calculation of bias, a simplified version of average relative error (FAO, 2019), calculated between measured data and model predictions. The calculation of bias is defined as:

$$\boldsymbol{bias} = \left(\sum_{i=1}^{n} \boldsymbol{P_i} - \boldsymbol{O_i}\right)/\boldsymbol{n}$$

Where,

P_i = Predicted (modeled) value of change in SOC, N_2O, or CH_4 with the practice
O_i = Observed value of change in SOC, N_2O, or CH_4 with the practice
n = Number of values in study

Bias indicates the average tendency of the modeled estimates to be larger or smaller than their observed counterparts (Moriasi et al., 2007). An unbiased model will have *bias* = 0.0. Positive values indicate model overestimation bias, meaning that the model overestimates the practice effect and thus the credits earned. A negative value indicates model underestimation bias, or an underestimation of the credits earned.

Bias is evaluated in two ways. First, bias must be calculated for each individual experimental study since different studies may use different temporal units of aggregation, soil depths, or measurement techniques. Since observed values are measured with some error, it is sufficient for validation to show that model bias for a given study is less than the uncertainty of the observed value. Concretely, bias of an individual study must be shown to be ≤ pooled measurement uncertainty in all cases. Pooled measurement uncertainty is defined as the pooled standard deviation of all the measured values for a practice change:

$$\sigma_{meas} = \sqrt{\frac{\sum_{j=i}^{k} \sigma_j^2 (n_j - 1)}{\sum_{j=1}^{k} (n_j - 1)}}$$

Where

k = Number of studies examined

σ_j = Standard deviation of the observed change in SOC, N_2O, or CH_4 from practice in the *j*th study

n_j = Sample size of the *j*th study

Second, the model must be shown to be unbiased or conservatively biased *on average*, i.e., when considering all studies, the mean of the computed biases must be <= 0.

3.5 Linking Validation Data to Model Prediction Error

Validation data should be used to estimate the uncertainty of a model's predictions, i.e., the model prediction error (Section D.2) and evaluate model fit. The calculation of model uncertainty bounds associated with a particular prediction (i.e., the prediction interval) should account for where there are few validation data (e.g., by using a weakly informative prior if using a Bayesian framework, Fig 2.1.B) as well as account for data variability (i.e. with a wider posterior when data are more variable if using a Bayesian framework; see Equations D.4, D.5). These features enable the model to adequately estimate the confidence in its predictions, as described next.

In the model validation report, as a check that model uncertainty bounds have been appropriately set, measured versus modeled results should be compared for each crop functional group/practice category combination for changes in SOC, N_2O, and CH_4 (if relevant), and demonstrate a minimum confidence coverage of 90% for 90% prediction intervals (i.e., the 90% prediction intervals should contain the measured value for at least 90% of the validation data).

In the model validation report, the following should be also included for each crop functional group/practice category combination and for changes in SOC, N_2O, and CH_4:

- scatterplot of the model predictions versus measurements:
- histogram of residuals (the differences between predictions and measurements);
- mean squared error.

3.6 Reporting on Model Validation

A model validation report following the above guidance should be submitted to the Reserve either by the expert team member or a third-party expert employed by the project during a given reporting period. To be accepted by the Reserve, the report will need to have received review and approval by an independent expert entity as having followed guidance in this document. The independent expert entity will need to be approved by the Reserve, following an assessment of conflict of interest between the reviewing entity and the project developer. Model validation requirements must be satisfied and confirmed prior to the completion of verification activities. This validation report will need to be revised, re-reviewed, and re-submitted at any point changes are made to the model.

There are 3 options for model validation reports:

1. project-specific, that includes demonstration of model validation for a specific project's domain and combinations of crop functional groups/practices categories;

2. generalized to demonstrate overall performance of a given model, i.e. demonstrating where model performance is valid over a range of possible project domains and crop functional group/practices category combinations; or
3. project-specific and referencing an existing model validation report (type 1 or 2)

The above guidelines are written for type 1 model validation reports. A type 2 report would follow the same guidelines, but clearly identify all project domains and crop functional group/practice category combinations where a given model version and parameter set/s have been calibrated and validated. A type 3 report would follow the same guidelines, but clearly demonstrate that the referenced model validation report (type 1 or 2) meets model validation requirements for the specific project, including proof that the same model version and parameter sets are used, and that all project domain and crop functional group/practices category combinations have met minimum requirements for model validation. All types of model validation reports must either be independently assessed by an approved 3rd-party entity or accepted for publication in one of the peer-reviewed publications listed in Section 3.6.

For each subsequent monitoring report, as long as a project area remains constant, or is only expanded to include new fields that already fit within the validated project domain, the existing validation report can be used. If the project is expanded to new practice effects, new crop functional groups, or the model is changed (using new model calibrations, changed parameters/parameter sets, or a different model, following guidance described in Section 6.5 of the SEP protocol), the validation report needs to be revised, reviewed by an approved 3rd party entity, and re-submitted. All model validation reports will be public documents.

It is also acceptable that the validation report is submitted as a peer-reviewed journal article, provided that the journal is on the pre-approved list provided below. It is acceptable that the journal article has not yet been printed as long as it has passed peer review and has been accepted for publication with revisions that do not change any aspects of model validation following the guidance in this document. In this circumstance, the project should submit the peer reviewed publication and responses to all revisions that clearly demonstrate revisions do not impact model validation. Where the peer-reviewed publication option in pursued, it is additionally acceptable that model validation is completed using a different method than explicitly evaluating bias and goodness of fit as described above. The publication must demonstrate that separate datasets were used for model calibration and model validation (unless qualifying for a special exception; see Section 3.1). The model validation must demonstrate the model was found acceptable for use by the peer reviewers for a given biophysical domain and a given set of practices. Additionally, the biophysical domain and practices used in the publication must be shown to completely meet the same domain requirements laid out in Sections 3.2 and 3.3, as well as cover the practice categories and crop functional groups identified in Section 3.1. The same datasets used in the peer-reviewed model validation should be used to calculate model prediction error used in the project. The same model version and model parameter values/parameter set values must be used in the peer-reviewed publication as are used in the project. Lastly, as a means of enhancing transparency with the peer reviewers, the authors must clearly state the purpose of the paper as being to validate the model for use in generating verifiable carbon credits.

Pre-approved Journals:

- Agricultural and Forest Meteorology
- Agricultural Systems
- Agriculture, Ecosystems and Environment
- Agronomy Journal

- Atmospheric Environment
- Biogeochemistry
- Biogeosciences
- Ecological Applications
- Ecological Modeling
- Ecosystems
- Environmental Modelling and Software
- Environmental Pollution
- Field Crops Research
- Frontiers in Ecology and the Environment
- Geoderma
- Global Biogeochemical Cycles
- Global Change Biology
- Journal of Environmental Quality
- Journal of Geophysical Research - Biogeosciences
- Nutrient Cycling in Agroecosystems
- Plant & Soil
- PLoS ONE
- Science of the Total Environment
- Soil & Tillage Research
- Soil Biology & Biochemistry
- Soil Science Society of America Journal
- Soil Use & Management
- Vadose Zone Journal

4 Substitution for Missing Crop Types

If during the calibration and validation process no sufficient data are available for a declared crop type, a substitution may be made that entails specific replacements be made for the baseline and with-project simulations. This method depends on the availability of alternative crop types for a given practice effect that meet all of the above criteria; without any alternatives no substitution can be made.

- Baseline:
  - Replace the missing crop type with an unfertilized perennial grass
- Project:
  - Replace the missing crop type with an alternative crop type for which data are available that best matches the missing crop in terms of its attributes, i.e., N fixation, annual/perennial, photosynthetic pathway, plant form, and flooded/not flooded status. The acceptable alternative crop will have the most matching categories among all crop types available.
  - For multiple alternative crop types having the same number of matching attributes, the crop type that best accommodates the management practices of the missing crop should be selected.

5 Verification of Model Usage

Each verification team must include a person or persons who are expert in the particular biogeochemical model used to quantify emission reductions in that reporting period (if any). Guidance is provided in Section 2 for requirements that models must meet to be eligible. Verifiers will be required to confirm the requirements of Section 2 of this document are met.

Expert guidance is needed to ensure the given biogeochemical model is appropriately validated, parameterized, and calibrated for each reporting period. If the project employs the use of a third-party expert to undertake validation, parameterization, calibration, and/or running a biogeochemical model in a given reporting period, then there will be no need for the verification team to independently verify such activities have been done appropriately, provided the verification team: confirms that the use of such third-party has been approved by the Reserve, that the party in question has the requisite expertise, that all requisite steps as set out in Section 2 of this document have been followed, and provided the expert provides the verification team with a sensitivity analysis regarding the requisite data inputs for the given model.

In other words, the verifier is simply required to confirm approval from the Reserve, confirm the qualification of the third-party, and confirm the requisite validation steps have been followed, but the verifier does not independently need to run the model themselves to confirm results appear reasonable. The verification team will still be required to confirm the reasonableness of all data input into the given biogeochemical model, following the requirements for baseline modelling in Section 3.4.1.1 of the SEP, and following expert guidance on the sensitivity of the given model to the requisite data inputs.

6 References

Bruun, S., Christensen, B. T., Hansen, E. M., Magid, J., & Jensen, L. S. (2003). Calibration and validation of the soil organic matter dynamics of the Daisy model with data from the Askov long-term experiments. *Soil Biology and Biochemistry*, *35*(1), 67–76.

Climate Action Reserve. Expected adoption: June 10, 2020. Soil Enrichment Protocol. Available at http://www.climateactionreserve.org/how/protocols/soil-enrichment/.

Dietze, M. (2017). *Ecological Forecasting*. Princeton University Press

Food and Agriculture Organization of the United Nations FAO. (2019). Measuring and modelling soil carbon stocks and stock changes in livestock production systems. Available at http://www.fao.org/3/CA2934EN/ca2934en.pdf.

Intergovernmental Panel on Climate Change. (2006). 2006 IPCC Guidelines for National Greenhouse Gas Inventories. Available at https://www.ipcc-nggip.iges.or.jp/public/2006gl/.

Liang, Y., Gollany, H. T., Rickman, R. W., Albrecht, S. L., Follett, R. F., Wilhelm, W. W., … Douglas, C. L. (2009). Simulating soil organic matter with CQESTR (v. 2.0): Model description and validation against long-term experiments across North America. *Ecological Modelling*, *220*(4), 568–581

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *American Society of Agricultural and Biological Engineers*, *50* (3), 885−900.

Mishurov, M., & Kiely, G. (2011). Gap-filling techniques for the annual sums of nitrous oxide fluxes. *Agricultural and Forest Meteorology*, *151*(12), 1763–1767.

Natural Resources Conservation Service. 2014 Farm Bill - Conservation Compliance Crop List. United States Department of Agriculture. Available at https://www.nrcs.usda.gov/wps/portal/nrcs/detail/national/programs/farmbill/?cid=stelprdb1262733.

Turner, P. A., Baker, J. M., Griffis, T. J., & Venterea, R. T. (2016). Impact of Kura Clover Living Mulch on Nitrous Oxide Emissions in a Corn-Soybean System. *Journal of Environmental Quality*, *45*(5), 1782–1787.

Wallach, D., Makowski, D., Wigington, J., Brun, F. (2014). Working with Dynamic Crop Models: Methods, Tools and Examples for Agriculture and Environment, second ed. Elsevier Science, Oxford, UK.

Wikle, C. K., & Berliner, L. M. (2007). A Bayesian tutorial for data assimilation. *Physica D: Nonlinear Phenomena*, *230*(1–2), 1–16. https://doi.org/10.1016/j.physd.2006.09.017

Yang, J. M., Yang, J. Y., Liu, S., & Hoogenboom, G. (2014). An evaluation of the statistical methods for testing the performance of crop models with observed data. *Agricultural Systems*, *127*, 81‑89. Available at https://doi.org/10.1016/j.agsy.2014.01.008.

Yeluripati, J. B., van Oijen, M., Wattenbach, M., Neftel, A., Ammann, A., Parton, W. J., & Smith, P. (2009). Bayesian calibration as a tool for initialising the carbon pools of dynamic soil models. *Soil Biology and Biochemistry*, *41*(12), 2579–2583.