CORRECTIONS AND CLARIFICATIONS TO *VMD0053 VCS MODULE MODEL CALIBRATION, VALIDATION AND UNCERTAINTY GUIDANCE FOR THE METHODOLOGY FOR IMPROVED AGRICULTURAL LAND MANAGEMENT, V2.0*

Publication date: 22 January 2024

This document provides corrections and clarifications applicable to *VMD0053 Model Calibration, Validation and Uncertainty Guidance for the Methodology for Improved Agricultural Land Management, v2.0*. Such corrections and clarifications are effective on their issuance date. Project proponents and validation/verification bodies (VVBs) shall apply and interpret *VMD0053, v2.0* consistent with the clarifications set out in this document.

These updates will be incorporated into the next issued version of the methodology.

Correction/ Clarification	Description	Section Reference
Clarification 1	Clarified that VMD0053 must be applied when modeling baseline and project scenario	2 and 5
Clarification 2	Added reference to IPCC Climate Zones	3
Clarification 3	Clarified requirements for clay content range covered in the model validation dataset	5.2.3
Clarification 4	Clarified linkages to VM0042 for estimating model prediction error and added footnote clarifying characteristics of independent observations	5.2.5

Correction/ Clarification	Description	Section Reference
Correction 5	Corrections and clarifications to Figure 6	5.2.5., Figure 6
Clarification 6	Moved sentence from VM0042, Section 4	5.2.6
Clarification 7	Specification added that project proponent must provide all project documentation to independent modeling expert (IME)	Appendix 1

CLARIFICATION 1

Clarification:

2 SUMMARY DESCRIPTION OF THE MODULE

This module provides procedures for calibration, validation and verification of empirical or process-based models used to estimate stock change/emissions with the application of *VM0042 Methodology for Improved Agricultural Land Management*. It provides a standardized approach to test model performance as a component of greenhouse gas (GHG) quantification for agricultural land management (ALM) projects using VM0042. This module must be used for all GHGs and carbon pools for which models are employed following Quantification Approach 1 (Measure and Model) in VM0042, for modeling the baseline and the project scenario. Figure 1 shows the workflow for the use of models within the VCS project cycle.

[...]

5 PROCEDURES

The procedures for model calibration and validation described in this section must be adhered to when modeling the baseline and project scenarios following Quantification Approach 1 to determine SOC stock changes or GHG emissions in VM0042. As defined in Section 6 of VM0042, the historical look-back period for modeling the baseline must be at least three years and must include at least one complete crop rotation, where applicable.

Background:

To streamline assessment by independent modeling experts (IME), VMD0053 should clearly state that module guidance applies to baseline scenario modeling.

CLARIFICATION 2

Clarification:

3 DEFINITIONS

[...]

Climate zone
Geographic zone as defined in the *2019 Refinements to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories,* Volume 4, Chapter 3, Figure 3A.5.1[1] on page 3.47.

Background:

Clarify reference for IPCC climate zones.

[1] Consider the first Corrigenda to the *2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories* replacing the original climate zone map, available at: https://www.ipcc-nggip.iges.or.jp/public/2019rf/corrigenda1.html

CLARIFICATION 3

Clarification:

5.2.3 Gather Data to Validate Model Performance and Uncertainty

[...]

Requirement 2: Specific Dataset Requirements to Validate Model

The specific minimum requirements for validating model performance and uncertainty for a PC/CFG/ES combination are set based on the geographic extent of a project (i.e., the declared climate zones or nationally defined agricultural land regions), as well as the soil attributes encountered within the project (i.e., the declared soil textural classes and clay contents).

For all PC/CFG/ES combinations, each climate zone or nationally defined agricultural land region – depending on which is used – must be represented in the validation dataset. Additionally, the three soil textural classes expected to be most predominant in the project's geographic extent must be included in the dataset. The dataset must cover a range of soil clay content that spans 15 percentage points. Where clay content in the project area varies by less than 15 percentage points, the project proponent must demonstrate that the dataset covers a representative range of soil texture conditions likely to be encountered across the project site. Once validated, a PC/CFG/ES combination is approved for crediting within all declared climate zones/nationally defined agricultural land regions and for all declared soil textures.

[...]

Background:

Not all project areas will have soil clay content covering a range of 15 percentage points.

CLARIFICATION 4

Clarification:

5.2.5 Using Data to Evaluate Model Prediction Error

To evaluate the model for performance, the same datasets should be used to estimate the uncertainty of model predictions (i.e., the model prediction error) and evaluate model fit. The calculation of model uncertainty bounds associated with a particular prediction (i.e., the prediction interval) should account for cases where there are few validation data (e.g., by using a weakly informative prior when using a Bayesian framework, Figure 2B) as well as for data variability (i.e., with a wider posterior when data are more variable, where using a Bayesian framework). These features enable the model to adequately estimate the confidence in its predictions, as described next.

Two procedures for estimating model prediction error are outlined in VM0042:

- Analytical Calculation of Error: See procedures described in Section 8.6.1.1.1 in VM0042. Where sufficient data from experimental field trials with controls are available, model prediction error is equivalent to the variance of the errors comparing simulated and observed emissions reductions/removals.

- Monte Carlo Simulation: See procedures described in Section 8.6.1.2 in VM0042, particularly Equation 59 and procedures for the estimation of s^2_{model}. Models with parameter uncertainty are used to simulate observations in the validation dataset with posterior predictive distributions (PPDs). PPDs are then iteratively resampled to generate an estimate of model prediction error.

Users must select one approach or the other for each GHG being considered for validation. Where the Monte Carlo Simulation method is used, all model parameter distributions and other inputs (e.g., model hyperparameters) must be reported. Illustrative examples in this document are based on the Analytical Calculation of Error approach.

In the MVR, as a check that model uncertainty bounds have been appropriately set, measured versus modeled results should be compared for each PC/CFG/ES combination for changes in SOC, N_2O and CH_4 (where relevant). [...]

The prediction interval should be compared against independent observations[2] that were not used in calibrating model parameter distribution functions nor calculating the standard deviation of residuals. Leave-one-out or k-fold cross validation techniques are recommended to achieve this goal. Calculation of confidence coverage is then based on the total number of tests performed across all iterations.

It should be recognized that [...]

After demonstrating that the model uncertainty bounds have been appropriately set, the model prediction error is estimated using the entire dataset and reported in the MVR. The reported value should be used as the estimate of model prediction error in all uncertainty

[2] An example of observations that are not independent is when SOC is measured from 0 to 4 years and then again at 8 years. There is one observation from 0 to 4 years and another from 4 to 8 years, with the latter simulation also having to run through the 0–4 year period. The measurements over these two periods are not independent and must not be split between the validation and calibration subsets. Further, the period from 0–8 years is not eligible to be included, even if kept in the same subset, since it is just a combination of the two shorter periods so would distort the apparent model prediction error.

deduction calculations for the project per the procedures outlined in Section 8.6 of VM0042.

In the MVR, the following must also be included [...]

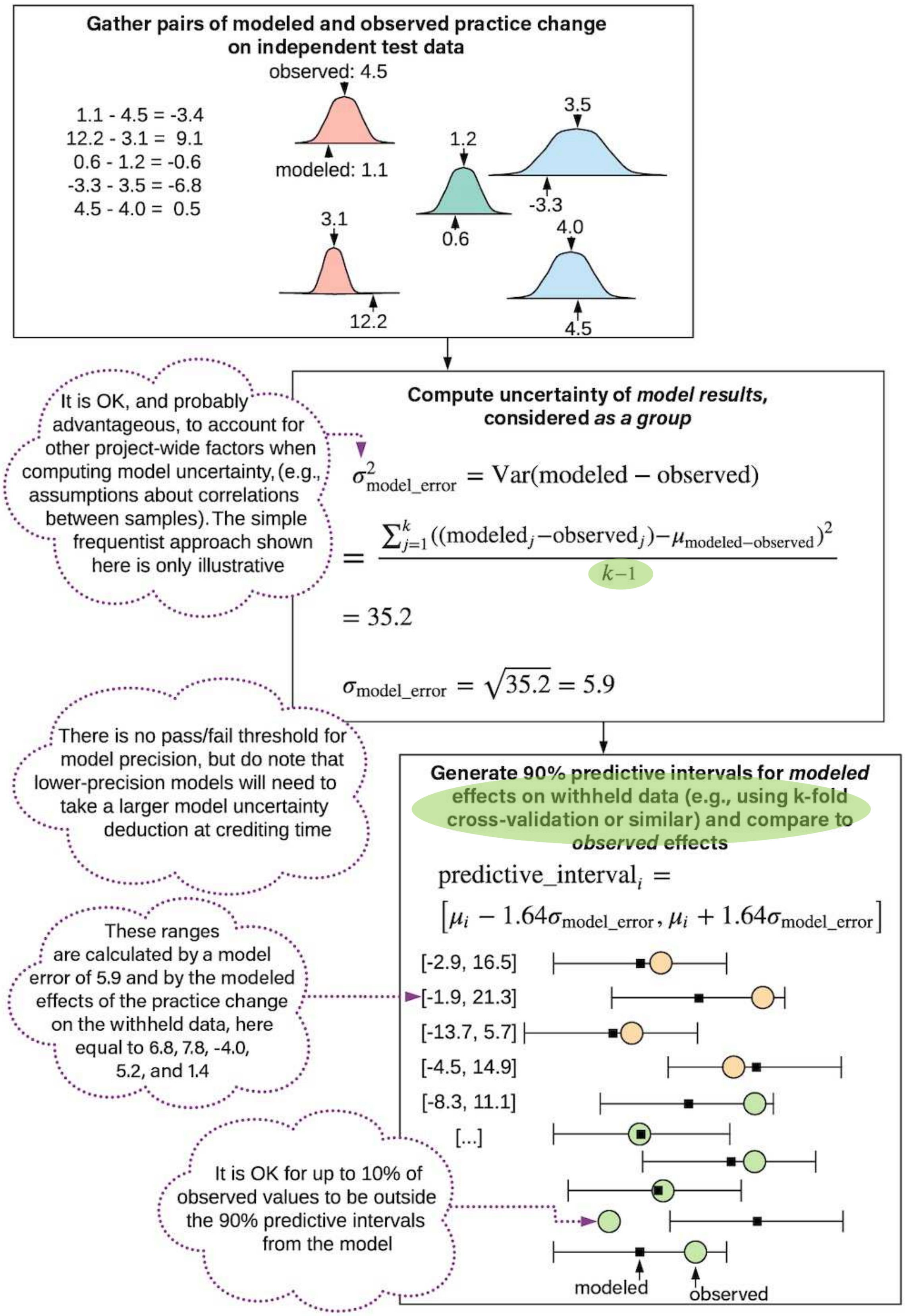
Background:

Improve explanations of linkages between VMD0053 and procedures in Section 8.6 of VM0042. Add clarifying example to provide further guidance on selecting experimental data for model calibration and validation. Note that footnote 2 in this document corresponds to footnote 7 in VMD0053.

CORRECTION 5

Correction:

Figure 6: Illustrative example of one possible approach to computing model prediction error and testing whether validation data are within predictive intervals. This figure represents one iteration in a k-fold or leave-one-out cross validation approach.


Gather pairs of modeled and observed practice change on independent test data
observed: 4.5
1.1 - 4.5 = -3.4
12.2 - 3.1 = 9.1
0.6 - 1.2 = -0.6
-3.3 - 3.5 = -6.8
4.5 - 4.0 = 0.5
modeled: 1.1
3.5
1.2
-3.3
3.1
0.6
4.0
12.2
4.5
It is OK, and probably advantageous, to account for other project-wide factors when computing model uncertainty, (e.g., assumptions about correlations between samples). The simple frequentist approach shown here is only illustrative
Compute uncertainty of *model results*, considered *as a group*
$\sigma^2_{\text{model\_error}} = \text{Var}(\text{modeled} - \text{observed})$
$= \frac{\sum_{j=1}^{k}((\text{modeled}_j - \text{observed}_j) - \mu_{\text{modeled-observed}})^2}{k-1}$
$= 35.2$
$\sigma_{\text{model\_error}} = \sqrt{35.2} = 5.9$
There is no pass/fail threshold for model precision, but do note that lower-precision models will need to take a larger model uncertainty deduction at crediting time
Generate 90% predictive intervals for *modeled* effects on withheld data (e.g., using k-fold cross-validation or similar) and compare to *observed* effects
$\text{predictive\_interval}_i =$
$[\mu_i - 1.64\sigma_{\text{model\_error}}, \mu_i + 1.64\sigma_{\text{model\_error}}]$
These ranges are calculated by a model error of 5.9 and by the modeled effects of the practice change on the withheld data, here equal to 6.8, 7.8, -4.0, 5.2, and 1.4
[-2.9, 16.5]
[-1.9, 21.3]
[-13.7, 5.7]
[-4.5, 14.9]
[-8.3, 11.1]
[...]
It is OK for up to 10% of observed values to be outside the 90% predictive intervals from the model
modeled
observed

Background:

Correct denominator for *k* to *k* − 1 to compute uncertainty of *model results*, considered as a *group*. Add clarification that generation of predictive intervals in this example applies when withholding data (e.g., using k-fold cross-validation or similar).

CLARIFICATION 6

Clarification:

5.2.6 Review and Approval of Model Validation Reports

[...]

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 MVR may be used. Models may be recalibrated or revised based on new data, or a new model may be applied, provided the same model version using the same parameters/parameter sets is applied in the baseline and project scenarios. Where the project is expanded to new PCs, CFGs or emissions sources, or the model is changed in a way that substantively affects model runs and the estimated ERRs, the MVR [...]

Background:

Appropriate position of information regarding updating models.

CLARIFICATION 7

Clarification:

APPENDIX 1: ASSESSMENT BY INDEPENDENT MODELING EXPERT (IME)

[...]

Where the peer-reviewed publication option is pursued instead of the MVR, the IME must assess the publication based on the requirements listed in Section 5.2.6(2).

Project proponents must provide all relevant project documentation to the IME. Project proponents must promptly respond to questions and findings from the IME and submit additional evidence and assist in arranging meetings with stakeholders, as requested. The burden of proof in the assessment process ultimately rests with the project proponent.

Background:

Streamline IME assessment by ensuring IMEs have all necessary information about a project.