Appendix G. Methods for Validation and Uncertainty Assessment

G.1. Validation for the National Assessment

Within the context of the national assessment methodology, validation is defined as a quantitative evaluation of the quality of the input and (or) output data products upon which the assessment will be based. The validation strategy is designed to achieve two principal objectives: to identify, quantify, and document sources of error that underlie the assessment results; and to guide efforts to increase accuracy through improvements in data collection, model design, sampling design, and other elements of the methodology.

In terms of conducting validation for the assessment, potential errors underlying the assessment results can be attributed either to the input data products that are independent of the models or to model performance. The known accuracy of independent input data products, such as the National Land-Cover Datasets (NLCD), will be documented by referencing published reports—no new efforts are planned for validation of these existing data products. Instead, the validation strategy is focused on new data products generated from assessment models.

G.1.1. General Approach

The methods employed in the national assessment involve numerous input and output variables, each of which represents a potential target for validation. The validation strategy described herein is premised on recognition that individual variables are not equally effective as validation targets, and each target must be selected with consideration to its relative importance for the assessment results and the availability and quality of reference data. These considerations led to the selection of 14 variables as both appropriate and feasible targets for validation. The selected target variables and their characteristics (measurement units, spatial and temporal attributes) are listed in table G1. The set consists predominately of end-point data products from the modeling of terrestrial and aquatic systems, but also includes key, intermediate data products (land-use and land-cover change, ecosystem disturbance by fire).

Opportunities for validation fundamentally are constrained by the availability of suitable, existing reference datasets, and resources to support new dataset development and implementation of validation tasks. The key factors that affect the suitability of a reference dataset are its inherent data quality and the correspondence with the spatial and temporal attributes of the target variable. As a general rule, the validation will draw upon the best available (most suitable) existing datasets produced independently of the national assessment activity itself and additional monitoring data as they become available. The reference datasets to be employed for initial validation are listed in table G1. The reference data sources are identified as existing or prospective. The prospective data sources address data gaps or deficiencies and have strong potential to be realized in a timeframe that is sufficiently short to be considered feasible for application in the national assessment.

G.1.2. Sampling Strategy for Validation

A sampling approach to validation will be constructed to create a practical alternative to the time consuming and expensive option of a full coverage national validation using all potentially available validation data. The sampling approach will focus on a much smaller total area within which validation data can be selected, evaluated for quality, and processed for analyses. The probability sampling design underlying the validation will allow for rigorous inference to validate the full national assessment.

The rationale of the sampling approach is to spatially constrain the collection and processing of validation data. The candidate validation data from all carbon pools will be collected and the spatial co-location of these data will allow for analysis of associations among pools as well as within pools. The collection of validation data will not be restricted to the sample locations. For example, extremely valuable but sparse datasets, such as those available from FLUXNET, will be used in their entirety. Data of known quality that are available across a broad spatial extent (for example, Forest Inventory and Analysis (FIA) and National Resources Inventory (NRI)) also may be used in their entirety for certain validation analyses. The sampling approach primarily is targeted for potential validation data that require thorough scrutiny to establish fitness for use.

The sampling design for validation will be stratified with each of the assessment units serving as a stratum. This will allow validation results to be reported by these assessment units (U.S. Environmental Protection Agency Level II ecoregions, modified from Omernik, 1987). The sampling unit within each ecoregion will be a county, and a sample of counties will be selected within an ecoregion. A further stratification within each region will be constructed using criteria defining a priority of interest. That is, counties exhibiting large model uncertainties, high quantities of land-cover change, or containing rare conditions (for example, estuaries, wetlands, or impoundments) may be sampled with greater probability to increase the sample of validation data appropriate to address key questions. Potential validation data from any carbon pool then would be collected within the selected sample counties. In some cases, the validation data will have originated from a probability sampling design (for example, FIA and NRI reference data), and the desired probability-sampling feature of these validation data will be maintained. In other cases, the validation data will not have a rigorous sampling basis (for example, existing light detection and ranging (LIDAR) coverage), and the representation of the sample will be limited to the area of existing coverage.
Table G1. Major deliverable data products that will be targeted for validation and corresponding reference data sources for each deliverable.

[Abbreviations used in this table can be found in “Abbreviations, Acronyms, and Chemical Symbols” in the front of this report]

<table>
<thead>
<tr>
<th>Deliverable product</th>
<th>Measurement units</th>
<th>Temporal validation targets</th>
<th>Reference data source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modeled land-cover and land-use change(^1)</td>
<td>Thematic</td>
<td>2001–2010 estimates</td>
<td>LANDFIRE VCT</td>
</tr>
<tr>
<td>Ecosystem disturbance (fire)(^1)</td>
<td>Thematic</td>
<td>2001–2010 estimates</td>
<td>LANDFIRE MTBS, Consume outputs, NOAA CarbonTracker</td>
</tr>
<tr>
<td>Delivery of water to coastal areas(^1)</td>
<td>cubic meters per month; cubic meters per year</td>
<td>2001–2010 estimates</td>
<td>USGS streamgage network</td>
</tr>
<tr>
<td>Net ecosystem productivity by pools, ecosystems</td>
<td>grams of carbon per square meter per year</td>
<td>2001–2010 estimates</td>
<td>AmeriFlux, NEON, NOAA CarbonTracker</td>
</tr>
<tr>
<td>Net biome productivity</td>
<td>grams of carbon per square meter per year</td>
<td>2001–2010 estimates</td>
<td>AmeriFlux, NEON, NOAA CarbonTracker</td>
</tr>
<tr>
<td>Grain yields</td>
<td>grams of carbon per square meter per year</td>
<td>2001–2010 estimates</td>
<td>USDA NASS, ARS</td>
</tr>
<tr>
<td>Carbon stock by pools and ecosystems</td>
<td>grams of carbon per square meter</td>
<td>2001–2010 estimates</td>
<td>FIA/RPA, LTER, NEON, ARS, GRACEnet</td>
</tr>
<tr>
<td>Carbon emission by wildland fire(^1)</td>
<td>grams of carbon per square meter per year</td>
<td>2001–2010 estimates</td>
<td>LANDFIRE MTBS with Consume</td>
</tr>
<tr>
<td>Carbon removal by forest harvesting</td>
<td>grams of carbon per square meter per year</td>
<td>2001–2010 estimates</td>
<td>Forest Service FIA</td>
</tr>
<tr>
<td>Carbon pool size in lake and reservoir sediments</td>
<td>grams of carbon per square meter</td>
<td>2010</td>
<td>RESIS–II, ad hoc reports</td>
</tr>
<tr>
<td>Carbon accumulation rate in lake and reservoir sediments</td>
<td>grams of carbon per square meter per year</td>
<td>2010</td>
<td>RESIS–II</td>
</tr>
<tr>
<td>Methane emission by ecosystems</td>
<td>tons of CO(_2) per square meter per year</td>
<td>2001–2010</td>
<td>AmeriFlux, NEON, NOAA CarbonTracker, ad hoc reports</td>
</tr>
<tr>
<td>Nitrous-oxide emissions by ecosystems</td>
<td>tons of CO(_2) per square meter per year</td>
<td>2001–2010</td>
<td>AmeriFlux, ad hoc reports</td>
</tr>
<tr>
<td>Carbon delivery by rivers to coastal areas</td>
<td>grams of carbon per month; grams of carbon per year</td>
<td>2010</td>
<td>NWIS, SPARROW, NEWS</td>
</tr>
</tbody>
</table>

\(^1\)Intermediate product.
The validation sampling design will be constructed to allow continuous augmentation of the sample to build the validation database as resources become available; however, the sampling design will ensure that defensible estimates can be obtained from the sample at any stage during the procurement of the validation database.

G.1.3. Quality Control and Quality Assurance

Validation is the cornerstone for quality control and quality assurance (QC/QA) in the national assessment and subsequent monitoring. The validation and modeling teams will coordinate a continual review of information from the validation activity to identify problems or deficiencies in model results. Inconsistencies in model results and reference data that are deemed significant or anomalous in space or time will be flagged for further investigation. The modeling teams will seek to identify and understand the factors that underlie such inconsistencies, define explicit strategies to reduce or resolve them, and whenever feasible, promptly implement those strategies. Thus, the QC/QA process is realized through a dynamic feedback loop in which the validation leads to improved understanding of the methodology performance, which in turn leads to improvements in the methodology design and implementation (data, models, sampling).

G.1.4. Relation Between Validation and Monitoring

The strategies for validation and monitoring in the national assessment methodology are closely coupled. The data requirements addressed by the monitoring strategy encompass those for validation. Thus, validation and QC/QA will be sustained in parallel with monitoring subsequent to the initial national assessment.

G.1.5. Adaptability of Validation Strategy

The validation strategy will be adaptable to changes in data availability and information requirements. Individual data products from the assessment (including intermediate ones) may be added or removed from the list of validation targets in response to changes in model performance or specific issues that may arise. Additional or improved datasets will be incorporated as they become available and when deemed effective in support of validation objectives. The potential data sources identified in table G1 are recognized to have particularly strong potential for improving the reliability of the assessment results. In particular, implementation of LIDAR-based techniques for estimating aboveground biomass can be readily achieved through coordination of the growing set of planned and potential LIDAR-related activities of the U.S. Geological Survey (USGS) and other governmental or private organizations. The benefits and opportunities for incorporating LIDAR-derived biomass data in the national assessment are addressed below.

G.1.6. Addressing Data Gaps and Deficiencies: The Case of LIDAR

LIDAR technology of various configurations has been well demonstrated in the literature, beginning as early as the 1980s, to be effective in quantifying forest and nonforest structure (Lefsky and others, 2002; Lim and others, 2003). Current (2010) systems collect extremely accurate three-dimensional information at the meter level from airborne systems and at the centimeter level from ground-based systems. The wealth of commercial and research sensors are providing three-dimensional data of vegetation structure at an unprecedented rate; however, there is not yet a capability that can provide the spatial and temporal coverage that space-based optical and radar systems offer. Whereas LIDAR has been used successfully on disparate projects across the country to quantify vegetation structure and biomass (Nelson and others, 2003), there is not a coordinated, concerted effort to collect systematic, standardized LIDAR-derived structural information for a national-scale biomass estimation, validation, and quantification of change.

G.2. Uncertainty Assessment

Although the validation process evaluates the quality of output products based on comparison with existing data, uncertainty assessment builds on this by estimating confidence bounds on estimates that cannot be validated; for example, projections into the future or estimates for which there are no existing validation data. Assessing complex socioenvironmental systems generally contains some uncertainty resulting from data gaps, modeling capabilities, interactions between ecological phenomena, and our scientific understanding of the mechanics of these complex systems. It is essential for users of the national assessment to be aware of the many uncertainties inherent in methods and assumptions used. It is useful to distinguish between quantifiable uncertainties where some form of statistical information is available and nonquantifiable uncertainties where such information is not available. Because the latter are more difficult to analyze explicitly, the basic strategy will be to treat these uncertainties separately in terms of a two-level approach, based roughly on the Intergovernmental Panel on Climate Change (IPCC) guidelines (Intergovernmental Panel on Climate Change, 2006).

G.2.1. Communication About Unquantifiable Uncertainties

The first-level uncertainty assessment is designed to incorporate unquantifiable uncertainties (designated as
“unpredictabilities” by the IPCC) in terms of a representative set of scenarios to be modeled. These scenarios will be constructed to capture the relevant variability range in those factors deemed most important for carbon sequestration (including climate changes and population growth). There is no attempt to assign probabilities to different scenarios or storylines; they simply serve as examples of potential future conditions that might reasonably be expected to occur under different sets of assumptions about future environments and behavior; however, by adopting these different sets of assumptions, it is possible to model the carbon and greenhouse-gas (GHG) outcomes with measurable levels of uncertainty. Thus, these scenarios represent a set of uncertainty bounds on our assumptions about future conditions.

For this assessment, three of the major storylines proposed by the IPCC (Nakicenovic and others, 2000) will be followed, within which alternative management and mitigation scenarios are proposed (section 3.2 of this report). Uncertainty related to scenarios is considered unpredictable; hence, a strategy for communicating such uncertainty is needed. The communication effort will focus on sources of uncertainties and their potential effect.

G.2.2. Uncertainty Sources of Reference Scenarios and Potential Reduction Measures

Scenarios are useful tools to provide a range of potential future alternatives. This assessment will develop national and regional reference scenarios (section 3.2 of this report) that are consistent with IPCC storylines. The Special Report on Emission Scenarios (SRES) identified six primary sources of uncertainty within the scenarios framework (Nakicenovic and others, 2000). The same sources of uncertainty also are contained in the use of reference as well as alternative management scenarios developed for the assessment. The six sources identified are listed below.

Choice of storylines.—This category describes the uncertainty associated with the characteristics of the storylines and mostly is related to the combination of quantitative assumptions, such as increased population growth and decreased economic growth, used for each storyline.

Authors’ interpretation of storylines.—Differences in the translation of qualitative storylines into quantitative drivers can introduce uncertainty into the storylines. Uncertainty may be reduced for harmonized drivers (population, gross domestic product) if parameters for drivers are chosen consistently with the storylines.

Translation of understanding of linkages between driving forces into quantitative inputs for scenario analysis.—Scientific understanding between the linkages of drivers and quantifiable input parameters for models is limited and often results in an inconsistent application across modeling efforts.

Methodological differences.—These uncertainties arise from the modeling structure as well as the underlying uncertainties between drivers and their resulting effects.

Different sources of data.—Source data, such as land-use histories and baseline conditions, often are inconsistent in their availability (both spatially and temporally).

Inherent uncertainties.—Events considered “rare” are not included in modeling efforts because of their inherent unpredictability. Nonetheless, rare events have the ability to affect future trajectories and produce considerably different outcomes.

G.2.3. Uncertainties of Alternative Mitigation Scenarios and Potential Reduction Measures

As noted above, uncertainties around the IPCC SRES storylines and interpretation of them also apply to the alternative management scenarios for the same storylines. Here, a key uncertainty involves the design and implementation of future policies. Policy will affect eligibility for incentives, and policy instruments will motivate change to various degrees. Because of these concerns, direct prediction of future potential policies is not considered for the methodology to avoid increasing uncertainty. Instead, alternative management scenarios are linked with interpretation of plausible land-management activities. Sources of uncertainty and potential reduction measures relevant to alternative scenario development are given below.

Estimates of rates, suitable lands, and timing.—Uncertainties in the spreadsheet estimates can be improved by broadly incorporating expert knowledge at the regional and subregional level. Improving the spatial footprint by incorporating region-specific expert knowledge in the scenarios should (at least in theory) help improve uncertainties. Additionally, the estimates also can be improved by increasing the thematic precision of the land-use, land-cover, and land-management information (for example, managed loblolly pine forest versus softwood forest in the southeast). Increased thematic precision has ramifications for data requirements and availability and increases the cost of the analysis.

The evaluation of management activities for the criteria of cost, ecosystem effects, energy usage, and technological progress.—The uncertainties of a management activity’s performance are constrained by the use of relative, rather than absolute, estimates. Again, extensive consultation at the regional and subregional level will help.

Uncertainties related to behavioral responses assumed in developing the scenarios.—Such uncertainties may be captured or reduced by comparing the results of a scenario development with the results of a biogeochemical simulation, and more effectively, by monitoring or repeating assessments that revisit behavioral responses.

G.2.4. Estimation for Quantifiable Uncertainties

The second type of uncertainty treatment involves the explicit modeling of potential carbon sequestration for different scenarios. From a spatial perspective, the key objective of these models will be to scale down the parameters of each
scenario (such as overall climatic conditions and population pressures) to grid cells that are small enough to allow explicit carbon-sequestration modeling. At this level of modeling, it is deemed that all uncertainties should be treated in a quantifiable way. It is important to distinguish between “value” uncertainty of input data and model parameters, uncertainty of model structure and mathematical processes, and uncertainty affected by other technical components of the methodology such as land-use and land-cover change and disturbance modeling. All value uncertainties will be treated as probability distributions that can serve as inputs to model simulations. Where statistical data are available, such distributions will be estimated by standard statistical procedures based on IPCC recommendations (Intergovernmental Panel on Climate Change, 2006). Otherwise, such distributions will be elicited through expert judgments (typically triangular distributions based on elicited value ranges and most likely values).

G.2.5. Uncertainty of Input Data and Model Parameters

Following the IPCC (2006) guidance, uncertainty analysis will focus on random errors associated with input data and model parameters. The following steps will be followed during the assessment to evaluate and report this type of uncertainty.

Input data used for modeling—Such as biophysical data (climate, soil), modeled data (wildland fire, land-use change), and expert-knowledge-related data (mitigation activities)—will be assigned an uncertainty range, either expressed as a probability distribution function (PDF) curve or a probability look-up table. Example approaches include the following.

Land-cover data.—Uncertainty in initial land-use and land-cover data may be expressed as a contingency table, which can be used to develop empirical distributions of possible land-cover types for individual pixels, based on misclassification rates (Prisley and Smith, 1987; Fang and others, 2006). These empirical distributions can be translated into initial carbon-density distributions (Quaife and others, 2008).

Forest age and biomass.—Parameters used for initializing the biogeochemical modeling are based on the FIA program. Plot-level data can be aggregated based on location to the level of the Joint Frequency Distribution (JFD) unit used in biogeochemical methods. Aggregation can provide distributions for parameters such as forest age, biomass, species groups, site quality, canopy density, and so on.

Soil parameters.—Using the tables associated with the Soil Survey Geographic (SSURGO) database, it is possible to obtain ranges and distributions for important soil parameters based on the present soil components and their relative frequency.

Using approaches such as those outlined above or an empirical distribution (probability look-up table), similar to the fractile distribution shown in figure G1, input data distributions can be fit to mathematically defined statistical distributions, such as those described by IPCC (2006).

If a model parameter has a PDF, it can be evaluated using error propagation. When a parameter PDF is not available, it is possible to derive one using data-assimilation techniques. Some parameters may be obtained from expert judgment. For example, PDF parameters for remote-sensing-based fire-severity modeling may be obtained from table D2 in appendix D.

As input data are processed by models, additional opportunities arise to evaluate uncertainty. For example, the “forecasting scenarios of land-cover change” model (FORE–SCE) (appendix B of this report) uses logistic regression to predict probabilities of individual types of land-cover transitions. The result is a suite of probability surfaces representing the most likely locations for different types of changes to occur. Land-cover changes are then allocated across a landscape. During this process, information on uncertainty is available from fit statistics for the regression models (Hosmer and Lemeshow, 2000), as well as from the probability surfaces (Dendoncker and others, 2008). For example, comparing the probabilities for different types of change at a given pixel can indicate how much more likely one type of change is than another. When probabilities for several types of change are similar, there is greater uncertainty and ambiguity as to the type of change that will occur.

It should be noted, however, that the focus will be on quantifying the variability of end results, which will be carbon storage and GHG fluxes. Uncertainties may arise during many parts of the modeling process that may have little effect on the final outcome. For example, the specific location of land-cover changes across a homogeneous landscape may be extremely uncertain, but also may make a minimal difference in overall long-term carbon sequestration at the reporting-unit level.

Figure G1. Typical probability distribution (density) function (PDF) curves. From Intergovernmental Panel on Climate Change (2006, p. 3.25), used with permission.
G.2.6. Uncertainty of Model Structure and Component Interactions

Structural and other conceptual uncertainties will be treated using model-run evaluation and expert judgment. One issue that the assessment should address is which biogeochemical models to use for which ecosystems, and which key carbon-, nutrient-, and water-cycle elements will be treated by the models. This will be determined based on a criteria evaluation process outlined in chapter 3 of this report. Although it is possible to use alternative modeling forms in principle, it is deemed most practical from an operational perspective to consistently use the biogeochemical (BGC) models most recommended by experts. In doing so, it is vital that all assumptions be made explicit to model users. In addition, model sensitivities to key assumptions will be evaluated by simulations, and results made available to users. Both of these analyses can be tracked using a spreadsheet, which will be adopted for the entire assessment to track results and enhance user transparency.

For other conceptual uncertainties—such as the effects of climate change on disturbances and land-use and land-cover changes, or the interactions between the carbon, nutrient, and water cycles—statistical techniques such as the use of the IPCC-recommended PDF and Monte Carlo resampling methods may be used to understand the size of their uncertainties and relations between different ecosystem processes.

Because the General Ensemble Modeling System (GEMS) can encapsulate multiple models, and parameterize and drive these models with the same data, it provides an ideal environment or platform to identify and address issues of uncertainty related to model structure and mathematical representations of biophysical processes. For this assessment, model comparisons will be used within the GEMS structure and with other modeling groups via a national workshop. Additionally, to reduce biases in modeling, the models will be calibrated with in situ data (for example, flux-tower data, FIA data).

G.2.7. Uncertainty Related to Specific Methods

The assessment is required by the Energy Independence and Security Act (EISA) (U.S. Congress, 2007) to consider ecosystem-controlling processes, such as wildland fire, land-use change, lateral transport, and agricultural practices. The use of explicitly mapped and modeled ecosystem-controlling processes on a national basis for carbon sequestration and GHG fluxes should improve upon uncertainties in assessment results (Running, 2008), but it also is possible that incorporation of such information can introduce new uncertainties into the methodology. Uncertainties related to the ecosystem-controlling processes will be quantified and reported. The basic approach for assessing such uncertainties is related to developing synthesis information and data products in support of formulating mitigation strategies. This approach is discussed in chapter 3 and appendix F of this report.

G.2.8. Increasing User Confidence by Delivering and Comparing Results

This two-level approach focuses primarily on model inputs and model construction, but from the user’s perspective, uncertainties generally are most easily communicated in terms

![Figure G2. Diagram illustrating the recommended process for combining uncertainty from various sources in a carbon-sequestration assessment. PDF, probability distribution function.](image-url)
of model outputs. The validation process described previously, when communicated together with uncertainty analysis, will help enhance user confidence about the input scenarios and data or model uncertainties. An important aspect of communicating the results and their uncertainties will be the ability to draw comparisons between the results of this assessment and other published projections of terrestrial carbon sequestration. For example, the validation plan includes a comparison of assessment results with the spatial and temporal distribution of terrestrial sources and sinks of atmospheric carbon estimated by the biosphere and fire modules of the National Oceanic and Atmospheric Administration’s (NOAA) CarbonTracker system (table G1; Peters and others, 2007; additional information at URL http://www.esrl.noaa.gov/gmd/ccgg/carbontracker/). Highlighting differences between approaches, assumptions, data sources, and modeling techniques used in this assessment and other published works will help place the assessment results in perspective.

An additional means for communicating results to users is the delivery of assessment products in digital map format. The distribution of the maps through an online user interface (described in appendix I of this report) will allow users to obtain frequency distributions of the deliverables and an opportunity to explore these uncertainties in more depth (albeit, at different scales). It also is possible to provide summary measures of uncertainty based on all scenarios, such as overall value ranges (Intergovernmental Panel on Climate Change, 2006; fig. G2).

G.3. References Cited

[Reports that are only available online may require a subscription for access.]


This page intentionally left blank.