Review of information reported using higher-tier methods and complex models in the LULUCF sector

Sandro Federici

-

Giacomo Grassi

Joint Research Centre of the European Commission,
Ispra (Italy)

OUTLINE

1. Background and objectives

2. Introduction

- Tier 3, models, and relationship between them

3. General aspects on the use and review of tier 3 methods in LULUCF

- Tier 3 but not models, and models but not tier 3
- Why using a complex model in LULUCF?
- Implementation of complex LULUCF model

4. Specific review issues for complex LULUCF models

- Issues related to transparency, consistency, comparability, completeness and accuracy

5. Conclusion

6. Examples

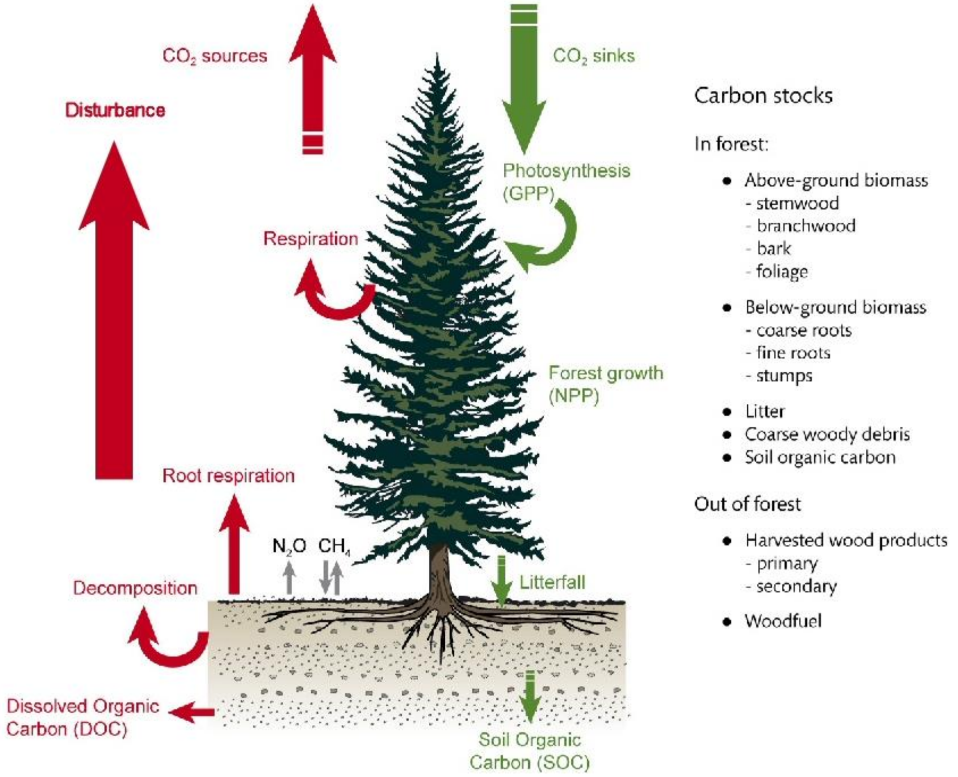
1.Background and objectives

Why LULUCF is complex?

Includes highly complex and dynamic ecological systems

... the influence of climate adds further variability

... the management of these system adds further complexity



This complexity translates in uncertainties of the estimates and of their attribution, e.g.:

- how to report complex and highly dynamic system (in space and time) with discrete categories and annual estimates?
- how to distinguish the human-induced effect from a highly variable background?
- how to separate the effect of current practices from the lagging effect of previous ones?

Background (cont)

Simple approaches for estimating GHG fluxes are often not fully satisfactory because they fail to capture this complexity and dynamic.

IPCC suggests to use the “managed land" concept as *proxy* for reporting anthropogenic net emissions, but recognizes the usefulness of higher tier methodologies to *refine* the estimation of anthropogenic emissions and removals.

For these reasons, in the last years more Annex I Parties are using complex tier 3 methods.

However, the complexity of tier 3 methods (especially models) represents a major challenge for the expert review teams (ERTs). The most challenging aspect is the apparent lack of transparency and the difficulty to assess the outputs.

Objectives of the course on review of complex models and higher-tier methods

Provide guidance on the general and specific aspects and procedures related to the review of emission/removal estimates performed using tier 3 methods, with the aim to improve the technical skills of experts participating in the review of GHG inventories and facilitate an harmonized application of criteria by the ERTs during the review of these estimates.

Objectives of this presentation

- Refresh general concepts
- Provide general guidance on issues to consider when reviewing complex models
- Stimulate the discussion among LRs and possibly collect additional examples

Background documents

- IPCC GPG for LULUCF and IPCC 2006Guidelines,
- Report of the IPCC expert meeting on the "Use of Models and Facility-Level Data in GHG Inventories", Sydney, Aug 2010
- Previous LRs meetings

1. Introduction

What is tier 3 ?

Models and inventory measurement systems tailored to address national circumstances, repeated over time, and driven by high-resolution activity data and disaggregated at sub-national to fine grid scales.

These higher order methods provide estimates of greater certainty than lower tiers and have a closer link between biomass and soil dynamics.

Such systems may be GIS-based combinations of age, class/production data systems with connections to soil modules, integrating several types of monitoring.

Pieces of land where a land-use change occurs can be tracked over time.

In most cases these systems have a climate dependency, and thus provide source estimates with interannual variability.

Models should undergo quality checks, audits, and validations, and a *thorough documentation* should be provided

What is a model ?

Modeling is a way to increase the power of data.

Although models are frequently used to assess complex systems or to generate data, models are means of data transformation and do not remove the need for the data to drive them

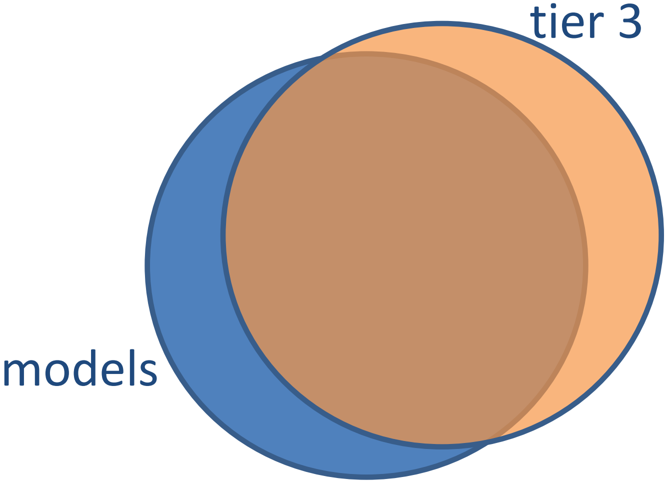
Every act of data interpretation has an underlying model. Even a simple calculation as Emissions = (activity data) × (emission factor)
is based on the assumption that units of activity individually or on the average, carry the same emissions burden. This assumption is the underlying model.

More complex models are used when this simple assumption is inadequate e.g., the sigmoid growth of a stand of trees means that one cannot simply multiply the removal rate by the stand area to get a removal from the atmosphere; the age of the stand also matters.

Relationship between models and tier 3

Not all tier 3 are models.
Not all models are tier 3.

E.g.:
- A stock-change approach may be tier-3 (if activity data is highly disaggregated e.g. a time-series of national forest inventories).
- A simple and coarse-resolution model is not tier 3



In the following slides, this presentation will shortly address :
- Tier 3, but not models
- Models but not tier 3

And then will analyse more extensively the issues to consider in the review of complex models

3. General aspects on the use & review of tier 3 in LULUCF

TIER 3, BUT NOT MODEL (short overview)

Tier 3 methods for estimating annual carbon stock changes may consist of a time-series of detailed datasets (with high resolution and disaggregation) provided by continuous or periodical stock inventory.

This usually apply only to forest land and often does not cover all carbon pools (although recent changes driven by the Kyoto Protocol requirements are transforming classical forest inventories toward carbon forest inventories).

When a time-series of inventory stock data is used the review consists in the review of the inventory design, i.e. the elements to be checked include the sampling design, the sample size, the plot size and shape and the inferences used for estimating the population attributes from those measured in the plots.

Moreover, the ERT should carefully check the consistency in definitions and methods applied between two consecutive inventories and, in case of identified inconsistencies, how these have been addressed by the Party.

<u>MODEL, BUT NOT TIER-3</u> (short overview)

If a Party uses coarse/simple data and regressions that deviate from the IPCC methods, strictly speaking cannot be considered neither tier 2 method (does not apply IPCC method) nor a tier 3 method (lack of proper disaggregation).

In this case the reviewer should check the robustness of the assumptions and inference applied to establish the accuracy of the estimates. Statistical tests (as r^2) should be provided for proving robustness of the inference while robust evidences/reasoning should be provided for the assumptions.

Example: linear regression between the aboveground biomass and another C pool in Forest land. This method assumes that increases/decreases in one pool correspond to increase/decrease in another pool. In practice, it assumes that older forests having more aboveground biomass contains also more C in other pools than younger forests. This is not universally true for all carbon pools and depends on forest typologies, morphology, geological substrate, climate, management practices.

If the model does not ensure accuracy of the estimates, it may be used in the accounting as “interim method” (if estimates are conservative), while the Party is developing a more accurate tier 3 method *or* dataset for applying an IPCC method.

COMPLEX MODELS

Why using a model in LULUCF?

- Higher accuracy, including: higher spatial and temporal resolution, better representation of complex dynamics (e.g. DOM and soil) and of climate/disturbances effects, etc.
- Completeness: improved coverage of land areas and/or C pools and GHG sources
- Cost-efficiency as compared to tier-2 (which may need extensive data collection)
- Improved uncertainty assessment
- Improved understanding of the processes
- Better assessment of the impacts of mitigation efforts/policy measures
- Improved time series consistency
- Can be used to estimate consistently past data and future projections

Models types and uses

In the LULUCF sector, models can be applied for two main issues:

1) For producing data, to be used as input data for estimating carbon stock changes and other emissions;

- By generating data, as for models used for producing activity data from remotely sensed images;
- By gap-filling in space and time discrete set of data, so reconstructing complete and consistent time-series and spatial datasets;

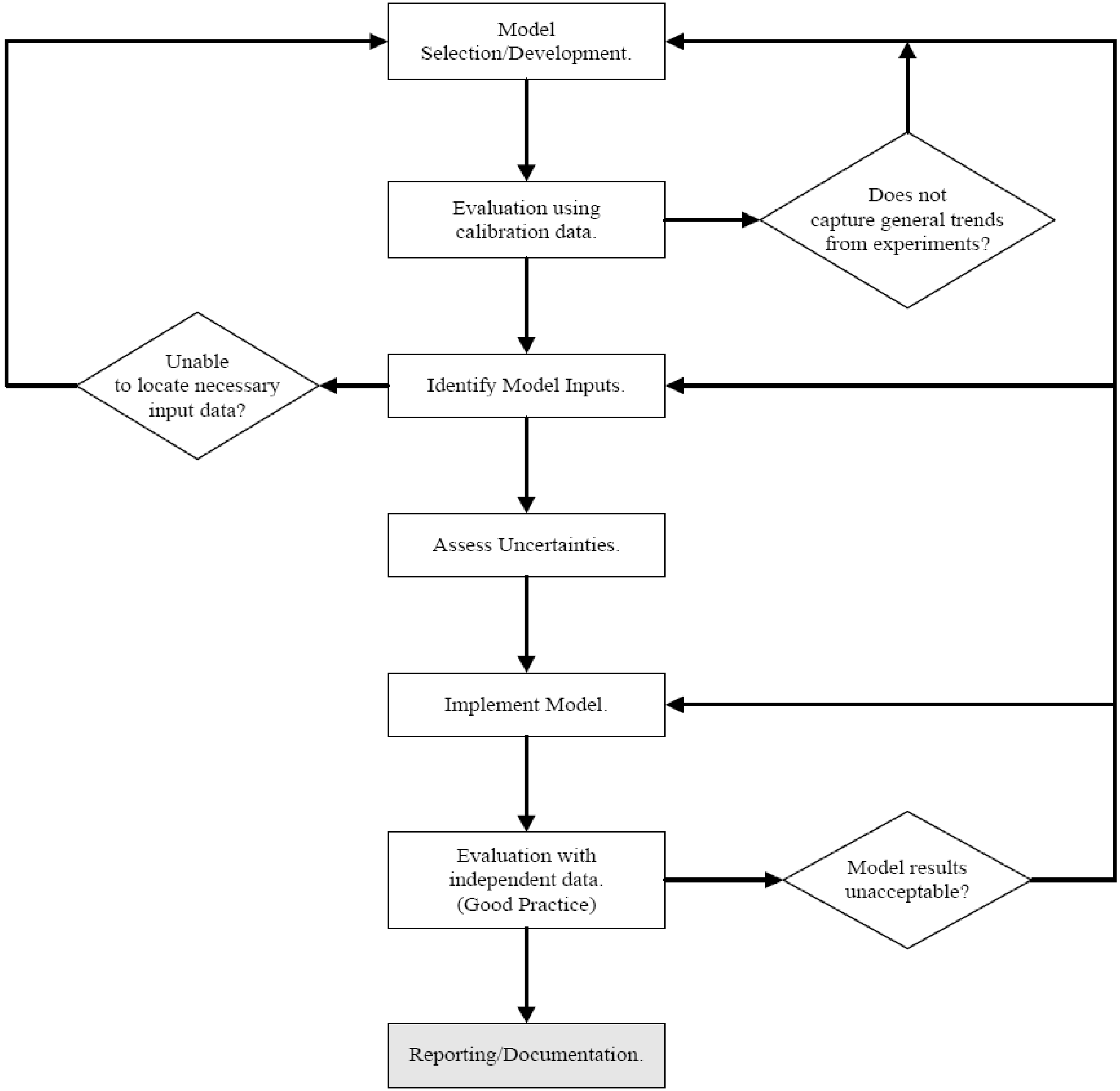
2) For estimating annual carbon stock changes, in each pool, and other emissions

- By statistical empirical method, an evolution of the IPCC gain and loss method, which are usually built on yield data and disturbances statistics
- By process-based method, which formalizes and simulates dynamic systems (i.e. systems that have a state, which is a description of the system at a given point in time, and processes that represent phenomena that occur in the system and cause the state to change over time)

Implementation of a complex LULUCF model

IPCC-GPG provides little general guidance on tier 3 models

IPCC 2006 GL in CH. 2.5.2 (AFOLU) describes 7 steps to implement a tier-3 model-based inventory:



Implementation of a complex LULUCF model (cont)

1. *Select/develop a model for calculating the stock changes and/or GHG emissions.*

   A model should be selected or developed to achieve higher accuracy than Tiers 1 or 2 approaches. It is good practice to consider the availability of input data (Step 3) and the resources needed to implement the model (Step 5).

2. *Evaluation with calibration data.*

   Model results are compared directly with measurements that were used for model calibration/ parameterization. Comparisons can be made using statistical tests and/or graphically, with the goal of demonstrating that the model effectively simulates measured trends for a variety of conditions in the category of interest.

   It is good practice to ensure that the model responds appropriately to variations in activity data and that the model is able to report results by relevant land-use category (or activity).

   Re-calibration of the model or modifications to the structure may be necessary if the model does not capture general trends or there are large systematic biases. Evaluation results are an important component of the reporting documentation, justifying the use of a particular model for quantifying GHG emissions.

Implementation of a complex LULUCF model (cont)

3. *Gather consistent timeseries of spatio-temporal data on activities and relevant environmental conditions that are needed as inputs to the model*

   These inputs may range from weather and soils data, forest types, natural disturbances or cropping management practices. It is good practice for the input data to be consistent with spatio-temporal scale of the model (i.e., algorithms).

   For example, if a model operates on a daily time step then the input data should provide information about daily variation in the environmental characteristic or activity data.

4. *Quantify uncertainties and perform sensitivity analysis*

   Uncertainties are due to imperfect knowledge about the activities or processes leading to GHG fluxes, and are typically manifested in the model structure and inputs.

   Then, uncertainty analyses are intended to provide a rigorous measure of the confidence attributed to a model estimate, based on uncertainties in the model structure and inputs, generating a measure of its variability.

   Sensitivity analysis, i.e. how the variability (uncertainty) in the output of a model can be apportioned, qualitatively or quantitatively, to different sources of variation in the input of the model.

Implementation of a complex LULUCF model (cont)

5. *Implement the model*

   Limitations in resources may constrain the complexity and range of spatial or temporal resolution that can be used in implementing the model at the national scale.

6. *Evaluation with independent data*

   While Step 2 involves testing model output with field data that were used as a basis for calibration (i.e., parameterization), the evaluation with independent data shall be done with a completely independent set of data from model calibration.

   Optimally, independent evaluation should be based on measurements from a monitoring network similar in principle to a series of sites that are used for a measurement-based inventory. However, the sampling does not need to be as dense because the network is not forming the basis for estimating carbon stock changes or other emissions, as in a purely measurement-based inventory, but is used to check model results.

   Problems may stem from one of three possibilities: errors in the implementation step, which typically arise from computer programming errors, poor input data, which are not representative of management activity or environmental conditions, or an inappropriate model (the model fails in formalizing real processes/dynamics).

7. *Reporting and Documentation*

4. Specific review issues for complex LULUCF models

Issues related to transparency

Among the 5 reporting principles (TACCC), transparency of the model output is of outmost importance. Without transparency, the other principles cannot be assessed. Indeed, where incomplete or unclear information is reported, it is not possible to assess whether a reported estimate:

- has no bias,
- is based on consistent data,
- completely covers the category to which refers,
- is consistent with IPCC (and therefore comparable with other Parties)

Consequently, the review will not ensure that the COP has adequate and reliable information and that the Party is assisted in improving the quality of its inventory.

In the context of reviewing a complex model, transparency essentially means providing all the information needed to understand and assess the model and its outputs. It is not just about the amount of information provided, i.e. thousands of pages (on model structure, development, calibration, etc....) may help to UNDERSTAND the model, but are not necessarily enough to ASSESS the outputs.

Here is reported a check list for documenting the Tier 3 Model-Based Inventory, which the ERT could use during the review:

1. Whether appropriate and sufficient information are reported on <u>Model Selection and Development</u>, and in particular on:
   a. Applicability of model and its adaptation to country's climate, vegetation and management conditions;
   b. Type of model, its conceptual approach (e.g. model represents statistical relationships or processes), and the mathematical formulation in general terms, (e.g. the model is process-based with a bottom-up approach to estimate emissions);
   c. Main processes and equations
   d. Key assumptions in model (e.g., first order approximation was assumed to represent soil organic matter decomposition for three kinetically-defined pools with a short, medium and long turnover time)
   e. Domain of application (e.g., all agricultural lands with arable crops grown on upland soils)

2. Whether appropriate and sufficient information are reported on Model Calibration and Evaluation, in particular on:
   a. Calibration of the model (i.e., parameterization);
   b. Model evaluation, as evaluation of calibrated model to measured emissions data or evaluation or other intermediate calculated variables of the model such as net primary production and respiration, litterfall, harvest transfers, or stock levels and transfers.

3. Whether appropriate and sufficient information are reported on Data Inputs, in particular on:
   a. Key inputs to the model. e.g., weather data were based on analysis of long-term precipitation and temperature data from the national weather service;
   b. Publications of the input data;
   c. Any key assumption that was necessary to use these data, such as representativeness of management data;
   d. Domain of the inventory application using the model given input data e.g.:
      - were different input data sets used in different parts of the domain, or
      - was the application of the model limited to specific parts of the country due to the domain of the input data

Check list for documenting the Tier 3 Model-Based Inventory (continue)

4. Whether appropriate and sufficient information are reported on Uncertainties and Sensitivity, in particular on:
   a. any sensitivity analysis conducted;
   b. derivation of uncertainties in the model inputs and model structure, as well as any other key uncertainties;

5. Whether appropriate and sufficient information are reported on Model Implementation, in particular on:
   a. computing framework including the hardware, databases and programs that were used to execute the inventory;
   b. description of key outputs variables from the model and any conversions or modifications made to derive the final emissions and removal estimates;
   c. summary of QA/QC procedures adopted to ensure the modelling systems performed appropriately, such as conservation of land area
   d. Optionally, examples of simple model calculations, such as emissions and removals by forest stands or landscapes in response to different forest management, natural disturbance, or mitigation scenarios.

Check list for documenting the Tier 3 Model-Based Inventory (continue)

6. Whether appropriate and sufficient information are reported on Evaluation of inventory results , in particular on:

   a. Evaluating inventory results which are determined by both the model and the input data, by comparison to:

      i. lower tier emission factors and/or expected ranges (out of range values may require further explanation),

      ii. to lower tier methods,

      iii. to independent measurements

   b. Evaluate the conservation of mass through the inventory analysis (carbon or nitrogen entering the system in combination with the existing storage, is accounted for through emissions and/or storage in the system). Note that:

      i) completeness shall be ensured, all inputs and outputs shall be accounted including those not related to GHG emissions (e.g., transfer between C pools)

      ii) storage in the system must be increasing or decreasing to balance the difference in inputs and outputs

Issues related to comparability

Tier 3 models do not follow, by definition, the IPCC method. Therefore, a check of comparability should not focus on how close it is the model to the IPCC method but on its capability to produce:

- consistent timeseries of estimates,
- a complete estimate of emissions and removals from the category,
- accurate estimates

If consistency, completeness and accuracy of the estimates is achieved then the comparability of method is ensured.

Another issue to be considered is the model capability to provide estimates according to IPCC/UNFCCC categories/activities and C pools.

While consistency with IPCC/UNFCCC categories/activities must always be ensured, some flexibility may be considered for C pools: if estimates do not follow IPCC/UNFCCC aggregation of C pools, then the ERT should check that the level of aggregation of reported estimates reflects the real and complete dynamics of carbon stocks. E.g. most models do not separate DOM and SOM.

A case of non-comparability among IPCC method and models - the SOM changes in mineral soils

- According to IPCC method, SOM changes, due to change in practices or land use, in mineral soils are estimated taking into consideration constant volume of soil (by default the upper 30 cm)

  This cause inaccuracy in SOM change estimates since

  - a partial accounting of increase of SOM due to accumulation of SOM in the upper layers, and

  - a partial accounting of decreases of SOM due to erosion (and leaching) of SOM in the upper layers;

- An accurate model for SOM should completely and accurately account for each loss and each gain so that estimates between the IPCC method and the model could hardly be comparable

Issues related to completeness and consistency

Completeness and consistency in land area should be assessed through:

- Checking the the "*land balance principle*", i.e. checking that area reported is constant along the whole time-series and matches the total area of the country;

- Additional area consistency checks, including:

a) a gross increases in the category "*land use X*" should correspond to the same gross increase in the area of subcategory of "*land converted to land use X*" (i.e. any land use change must transit to "land in conversion" category)

b) a gross increase in the subcategory "*land use X remaining land use X*" should correspond to the gross decrease in the sub category "*land converted to land use X*".

c) a gross decrease in the category "*land use X*" should correspond to the aggregate gross increase in the areas of subcategories of "*land use X converted to other land uses*".

More accurate checks should assess the consistency at the level of subcategories.

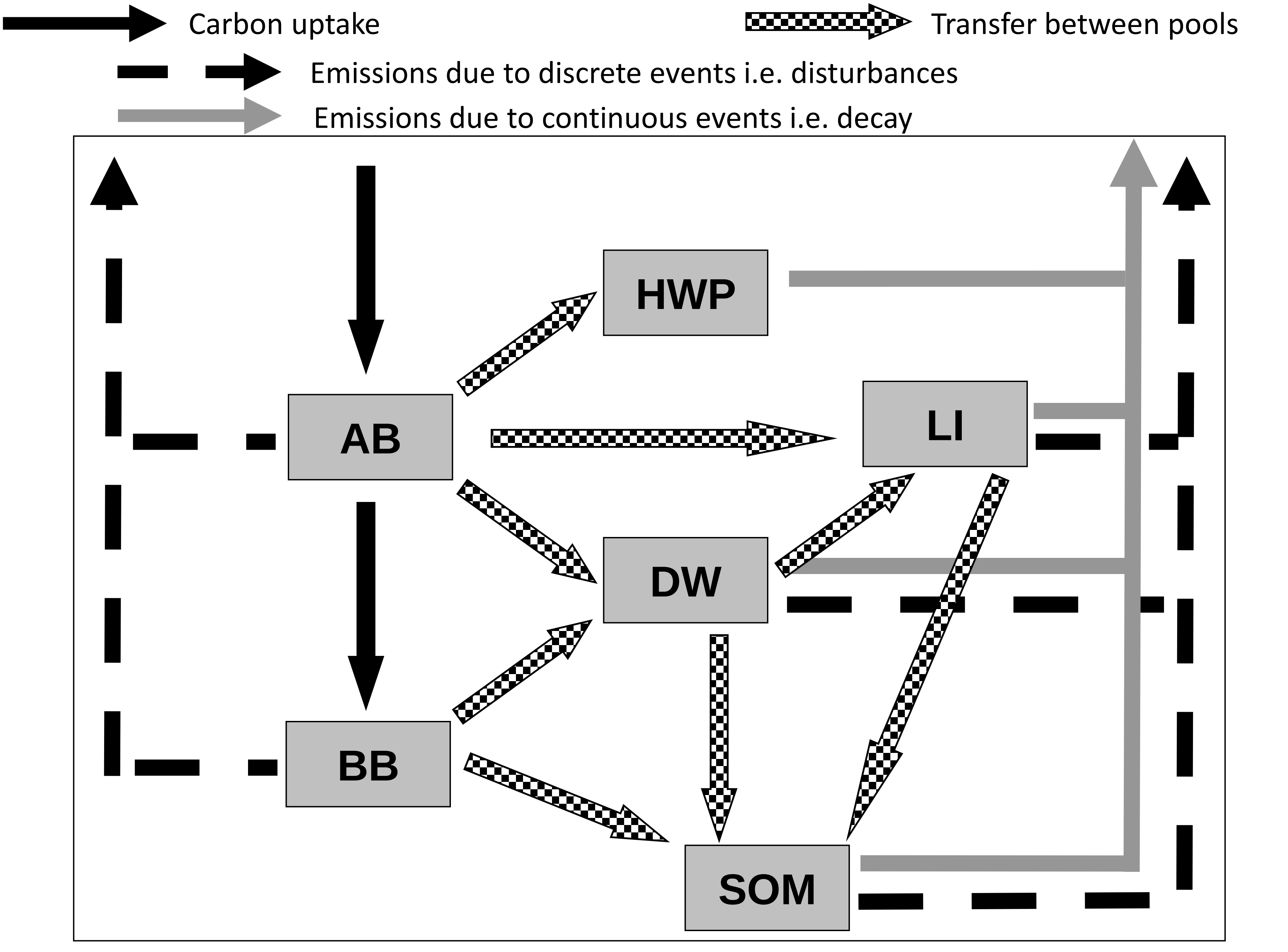
For managed lands of forest land, grassland and wetland, possible changes in unmanaged area should also be checked.

Issues related to completeness and consistency (cont)

Completeness and consistency in carbon stock change should be assessed through the "*mass balance principle*", i.e. checking with IPCC default method whether:

- total net change in stock in carbon pools corresponds to the total uptake of carbon from the atmosphere minus CO_2 emissions to the atmosphere and transfers to pools not included in the model e.g. solid waste disposal,
- for each biomass pool, net change in stock pairs uptake of carbon from the atmosphere minus CO_2 emissions to the atmosphere and transfers to other pools,
- for each pool not containing biomass, net change in stock pairs transfer of carbon to the pool minus CO_2 emissions to the atmosphere and transfers to other pools.

Moreover, mass balance should be checked along the time-series, i.e. at the onset of the new cycle carbon stocks levels correspond to those calculated at the end of the previous cycle



Carbon uptake
Transfer between pools
Emissions due to discrete events i.e. disturbances
Emissions due to continuous events i.e. decay
HWP
AB
LI
DW
BB
SOM

Issues related to consistency

The development complex models for the LULUCF sector is a process subject to frequent improvements, which may lead to recalculations. For example:

- The instruments used to collect activity data may change through time, and it is impossible to go back in time to apply the new instrument e.g. satellite imagery.
- Some data sources such as stock inventories may not be available annually because of resource constraints.
- Emissions and removals typically depend on past land use activity. Thus, data should cover a large historical period (20 years or more), and the quality of such data will often vary through time.
- Where changes are made to the data inputs or mathematical relationships in a model, the entire time series of estimates should be recalculated to maintain consistency.

The ERT should carefully check if the appropriate method to ensure consistency is applied, accordingly to IPCC (i.e. “overlap method” or variations of it, interpolation/extrapolation, etc.)

Issues related to accuracy

Checking the accuracy of a model means assessing its outputs. This is different from reviewing the full functioning and structure model itself (which is far more difficult and could potentially require much more time).

Indeed, a potential risk during the review is focusing only on model's functioning. Although this is very important to increase the confidence in the outputs, sometimes the (necessary) complexity of a model may preclude to fully assess the correctness of its functioning in one week of hard work.

For these reasons, the model's outputs should also be evaluated independently of the process which generated them.

What counts is the credibility of the results.

Issues related to accuracy (cont)

When assessing the accuracy of models' final outputs, the ERT may start with the "generalized approach to review tier 3 approaches", which means checking whether the outputs are "*what one would expect*", i.e.:

- Are within the lower Tier uncertainty range?
- Is the uncertainty decreased compared to a lower Tier estimate?

If not, can the deviation be explained?

Although this generalized approach is a valid *first step* for all sectors, in LULUCF some problem may arise due to:

- High uncertainties of lower tier estimates (tier 1 may not be a fully reliable source for comparison)
- For some C pool tier 1 assumes no change in C stock
- The outputs may not be directly compared to lower tier estimates

For these reasons, additional issues should be evaluated, including:

- verifying final model's output with real measurements
- comparing of intermediate models' outputs with lower tiers, with independent estimates and /or with field data;

Issues related uncertainties

A model may be thought of as a hypothesis regarding how the real-world system behaves. Thus, there are two key considerations in model uncertainty:

(1) has the correct, most relevant real-world system been identified, and have conceptualisations been constructed in a way that properly serve as the basis for model development?

(2) is the model has been developed as an accurate representation of the chosen system?

Uncertainties may be distinguished between 1 (conceptualisation uncertainty) and 2 (model uncertainty, i.e. lack of proper model development relative to the intended system).

The ERT should check that model's output uncertainties are identified and described accordingly to IPCC.

7. Conclusions

Among the reporting principles (TACCC), when reviewing a complex LULUCF models **transparency** is particularly important and challenging.

Transparency essentially means providing all the information needed to understand and assess the model and its outputs (in one week of hard work).

A clear and detailed description of models' structure and function is important and necessary to understand and build confidence on the model, but cannot substitute the assessment of the model's output (for their comparability, consistency, completeness and accuracy).

The ultimate scope of a review is to assess if models' output is credible (not to review the model itself).

The reviewer should thus starting assessing if the information provided allows for such assessment.

Conclusions (cont)

When reviewing a complex model, the following information should be considered:

(i) The basis and type of model;

(ii) The application and adaptation of the model;

(iii) The main equations and processes;

(iv) The key assumptions;

(v) The domain of the application;

(vi) How the model parameters were estimated;

(vii) A description of key inputs and outputs;

(viii) The details of calibration and model evaluation;

(ix) Uncertainty and sensitivity analysis;

(x) The quality assurance and quality control procedures adopted;

(xi) References to peer-reviewed literature;

Conclusions (cont)

When **assessing models' final outputs**, the "generalized approach to review tier 3 approaches" (compare with estimates from lower tier and check whether the estimates are "what one would expect") is a first useful and necessary step.

However, additional checks may be needed, including:

- verifying final model's output with measurements
- comparing of intermediate models' outputs with lower tiers, with independent estimates and /or with field data;
- land conservation check
- mass conservation check

some final thoughts on the relationship between reviewers and modellers...

Typically the reviewer starts saying that the model is a “black box”

... and the first reaction of modeller is to make the model more “transparent”

(reviewer)

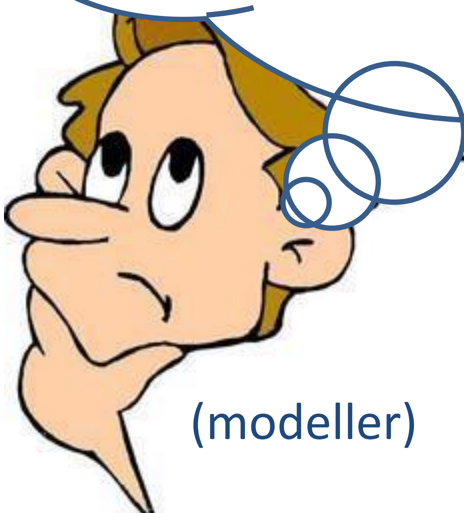
model

Then the reviewer asks to compare the model's results with results from a tier 2 approach

why hell they ask for such a comparison?

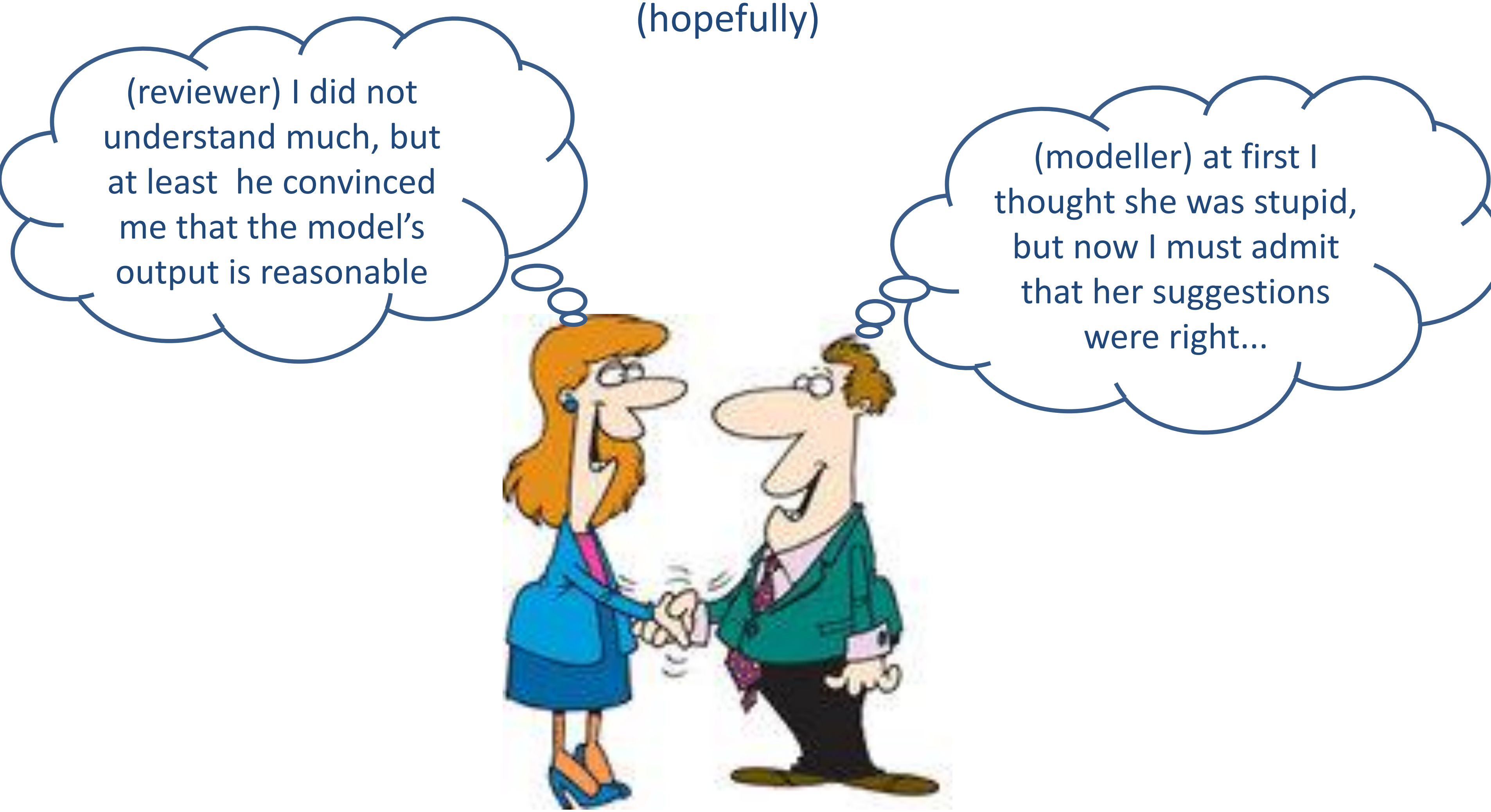
(tier 2)

(tier 3 model)



(modeller)

...and the last day of review...
(hopefully)



Thank you for the attention!

Issues for discussion

All important issues are covered?

Any specific example to be suggested? (best practices and /or problems / issues during review)

Other suggestions?