Dealing with uncertainty and sensitivity issues in process-based models of carbon and nitrogen cycles in northern forest ecosystems

G.R. Larocque[1], J.S. Bhatti[2], A.M. Gordon[3], N. Luckai[4], J. Liu[5], S. Liu[5], P. A. Arp[6], C.-F. Zhang[6], A. Komarov[7], P. Grabarnik[7], M. Wattenbach[8], C. Peng[9], Jianfeng Sun[9], T. White[10]

[1]Natural Resources Canada, Canadian Forest Service, Laurentian Forestry Centre, 1055 du P.E.P.S., Ste-Foy, Qc, Canada, G1V 4C7. Guy.Larocque@NRCan.gc.ca.
[2]Natural Resources Canada, Canadian Forest Service, Northern Forestry Center, 5320-122 Street, Edmonton, Alberta, T6H 3S5, Canada. jbhatti@nrcan-rncan.gc.ca
[3]Department of Environmental Biology, University of Guelph, Guelph, Ontario, N1G 2W1, Canada.
[4]Faculty of Forestry and the Forest Environment, Lakehead University, Thunder Bay, Ontario, P7B 5E1.
[5]Scientific Applications International Corporation (SAIC), U.S. Geological Survey Center for Earth Resources Observation and Science (USGS EROS), Sioux Falls, SD 57198, USA.
[6]Faculty of Forestry and the Forest Environment, University of New Brunswick, Fredericton, New Brunswick, E3B6C2.
[7]Institute of Problems in Soil Science, Institutskaya, 2, Pushchino, Moscow Region, 142290, Russia.
[8]Aberdeen University, Cruickshank Buidling, Aberdeen, Scotland, Ab24 3UU, United Kingdon.
[9]University of Quebec at Montreal, Case Post 8888, Succ. Centre-Ville, Montreal, Qc, H3C 3P8, Canada.
[10]Natural Resources Canada, 506 Burnside Road, Victoria, B.C., V8Z 1M5, Canada.

Abstract: Many process-based models on carbon (C) and nitrogen (N) cycles have been developed for northern forest ecosystems. These models are widely used to evaluate the long-term decisions in forest management dealing with effects like particulate pollution, productivity and climate change. Regarding climate change, one of the key questions that have sensitive political implications is whether northern forests will sequester atmospheric C or not. Whilst many process-based models have been tested for accuracy by evaluating or validating against observed data, few have dealt with the complexity of the incorporated procedures to estimate uncertainties associated with model predictions or the sensitivity of these predictions to input factors in a systematic, inter-model comparison fashion. In general, models differ in their underlying attempts to match natural complexities with assumed or imposed model structure and process formulations to estimate model parameters, to gather data and to address issues on scope, scale and natural variations. Uncertainties may originate from model structure, estimation of model parameters, data input, representation of natural variation and scaling exercises. Model structure relates to the mathematical representation of the processes modelled and the type of state variables that a model contains. The modelling of partitioning among above- and below-ground C and N pools and the interdependence among these pools remain a major source of uncertainty in model structure and error propagation. Most soil C models use at least three state variables to represent the different types of soil organic matter (SOM). This approach results in creating three artificial SOM pools, assuming that each one contains C compounds with same turnover rate. In reality, SOM consists of many different types of C compounds with widely different turnover rates. Uncertainty in data and parameter estimates are closely linked. Data uncertainties are associated with high variations in estimating forest biomass, productivity and soil organic matter and may be incomplete for model initialization, calibration, validation and sensitivity analysis of generalized predictor models. The scale at which a model is being used also affects the level of uncertainty, as the errors in the prediction of the C and N dynamics differ from the site to the landscape levels and across climatic regions. If the spatial or temporal scale of a model application is changed, additional uncertainty arises from neglecting natural variability in system variables in time and space. Uncertainty issues are also intimately related to model validation and sensitivity analysis. The estimation of uncertainties is needed to inform decision process, in order

to detect the possible corridor of development. Uncertainty in this context is an essential measure of quality for stakeholder and decision makers**.**

Keywords**:** northern forest ecosystems; models; structure process formulations, error propagation, uncertainties, sensitivity analyses

1. INTRODUCTION

Process-based models designed to simulate the dynamics of carbon (C) and nitrogen (N) cycles in northern forest ecosystems are increasingly being used in concert with other tools to predict the effects of environmental factors on forest productivity and forest-based C and N pools. Among the environmental factors, we include everything from intensive management practices to climate change, from local to global and from hours to centuries, respectively. Policy makers expect that well-calibrated process-based models will be at the center of rational and sustainable forest management policies and planning and prioritization of research efforts within the context of global change. In this context, it is important for policy makers to understand that any model is a simplified representation of reality. The predictive capacity of models is generally limited by an incomplete understanding of the complexity of the C and N cycles, and by imprecise and incomplete representations of processes and conditions needed to drive and calibrate the model calculations and estimates of natural variation. The last limitation is critical, as the majority of existing models are deterministic. There is, therefore, a need to examine (or re-examine) empirical robustness and uncertainty and sensitivity issues associated with model formulation, initialization, parameterization, calibration and error propagation from local to global scales.

2. SOURCES OF UNCERTAINTY

Three main sources of uncertainty have generally been recognized in models on C and N cycles in forest ecosystems, and in biological models in general [O'Neil and Rust, 1979; Medlyn et al., 2005]:

- Model structure, and lack of understanding of the biological processes;
- The plasticity that is associated with estimating model parameters, due to the general interdependence of model variables and parameters; related to this is the search to determine the least set of independent variables required to span the most important system states and responses from one extreme to another, e.g., from frozen to non frozen, dry to wet, hot to cold, calm to stormy, etc;
- The range of natural variations associated with each biological system under study.

Data uncertainties and incorrect assumptions about these data lead to additional errors. Uncertainties and errors also result from switching model scales, e.g., from daily to monthly to annual, and from stand- to catchment- to landscape-levels [Wu et al., 2005]. Data on C and N pools and flows in forest ecosystems, such as those reported and documented in Johnson and Lindberg [1991], are scarce and invaluable, but nevertheless incomplete in several ways: pool sizes and transfers among pools are generally for old growth conditions, and mortality assessments are inconsistent across the sites.

Process-based forest models tend to be complex, simulating many different processes and feedback mechanisms by integrating information on the underlying processes in trees, soil and atmosphere. Their strength is their complexity, which allows them, in principle, to reproduce the complex dynamics of forest ecosystems in detail. However, it also remains a weakness in that it makes their use and evaluation difficult. There is a need to quantify output uncertainty and identify key parameters and variables. The uncertainties are linked: uncertain parameters imply uncertain predictions and uncertainty about the real world implies uncertainty about model structure and parameterization. Because of these linkages, model parameterization, uncertainty analysis, sensitivity analysis, prediction, testing and comparison with other models need to be based on a consistent quantification of uncertainty.

2.1 Model structure and complexity

Many problems are generated by model structure alone. Two issues can be related to model structure: (1) mathematical representation of the processes and (2) description of state variables. For example, several types of models can be used to represent the effect of temperature variation on processes, including the Q_{10} model, the Arrhenius function or

other exponential relationships. The degree of uncertainty in the predictions of a model can increase significantly if the relationship representing the effect of temperature on processes is not based on accurate theoretical description [see Kätterer et al., 1998; Thornley and Cannell, 2001; Davidson and Janssens, 2006]. Most C and N models contain a relatively simple representation of the processes governing soil C and N dynamics, including simplistic parameterization of the partitioning of litter decomposition products between soil organic C and the atmosphere. For example, the description of the mineralization (chemical, physical, and biological turnover) of C and N in forest ecosystems generally addresses three major steps: (1) splitting of the soil organic matter into different fractions, which decompose at different rates, (2) evaluating the robustness of the mineralization coefficients of the adopted fractions, and (3) initializing the model in relation to the fractions [Wander, 2004].

Table 1 gives a cross-section of a number of recent models (or sub-components of models) used to determine litter decomposition rates. The entries in this Table illustrate how the complexity of the C and N modeling approach varies, even in describing a basic process such as forest litter decomposition: the number of C and N components by each model ranged from 5 to 10 (or 6 times the number of cohorts considered in SOMM, starting with 3 C and 3 N organic matter components associated with the decay of single-species leaves). The number of processes considered varied from 5 to 32 and the number of C and N parameters ranged from 7 to 54. The number of additional parameters used for describing the N mineralization process, once the organic matter decomposition process is defined, is particularly interesting: it ranged from 1 to 27.

Most soil C models use three state variables to represent different types of soil organic matter (SOM), the active, slow and passive pools. Even though it is assumed that each pool contains C compounds with about the same turnover rate, this approach remains nevertheless conceptual and may lead to uncertainty in the predictions [Davidson and Janssens, 2006]. Also, these conceptual pools do not directly correspond to measurable pools. In reality, SOM contains many types of complex compounds with very different turnover rates and amplitude of reaction to change in temperature [Davidson and Janssen, 2006]. There were many attempts to find relations between model structure and real world either by measuring different decomposition rates of various different soil fractions [Zimmerman et al., 2006] or by restructuring the models pools [e.g., Fang et al, 2005].

Process-based C and N models are generally referred to as being deterministic or stochastic. These models may be formulated for the steady state (for which inputs equal outputs), or the dynamic situation, where model outcomes depend on time, in relation to time-dependent variations of the model input, and in relation to state-dependent component responses. Models are either based on empirical or theoretical derivations, or a combination of both (semi-empirical considerations). Process-based modeling is cognizant of the importance of model structure: the number and type of model components are carefully chosen to mimic reality, and to help minimizing the introduction of modeling uncertainties.

Process-based forest models vary from simple to complex, involving various process and feedback mechanisms for the purpose of integrating ecosystem-based process information about the functioning of trees, of their soil, and of the atmosphere. Simple model formulations often suffer from being too simplistic, but can nevertheless be illustrative and educational in terms of ecosystem thinking. Simple model formulations are generally aimed at quickly estimating the order of magnitude of C and N quantities associated with particular ecosystem processes, such as C and N uptake and stand-internal C and N allocations. In doing this, there is special need to validate model outcomes, through identifying key ecosystem components, processes, variables, and parameters.

Complex models have the challenge of being more precise and/or accurate than simple models. This being so, data requirements for the initialization and calibration of complex models need to be tightly controlled, and need to stay within the range of current field experimentation and exploration. The degree of model complexity also needs to be controlled, because this degree affects the overall model transparency and communicability, as well as affordability and practicality. Also, making models more complex can increase their structural uncertainty simply by increasing the number of parameters that are uncertain.

Modellers must carefully consider the trade-off between the potential uncertainty that may result from adding additional variables and parameters and the gain in accuracy by doing so. It may be argued as well that existing models on the C cycle are still in their infancy. It is not evident that modellers involved in the development of process-based models have considered all the tools, including

mathematical development, systems analysis and programming, to deal with complexity.

2.2 Data and parameters

Many methodologies have been used to better quantify the uncertainty of model parameters. Traditionally, these methodologies include simple trial-and-error calibrations, fitting model calculations with known field data using linear or non-linear regression techniques and assigning pre-determined parameter values, generated empirically through various means, in the laboratory, the greenhouse, or the field. For example, Wang et al. [2001] used nonlinear inversion techniques to investigate the number of model parameters that can be resolved from measurements. Braswell et al. [2005] and Knorr and Kattge [2005] used a stochastic inversion technique to derive the probability density functions for the parameters of an ecosystem model from eddy covariance measurements of atmospheric C. Williams et al. [2005] used a time series analysis to reduce parameter uncertainty for the derivation of a simple C transformation model from repeated measurements of C pools and fluxes in a young ponderosa pine stand and Dufrêne et al [2005] used the Monte Carlo technique to estimate uncertainty in net ecosystem exchange by randomly varying key parameters following a normal distribution.

Erroneous parameter assignments can lead to gross over- or under-predictions of forest-based C and N pools. For example, Laiho and Prescott [2004] pointed out that Zimmerman et al. [1995], using an incorrect C/N ratio (of 30) for coarse woody debris in the CENTURY[1] model, greatly overestimated the capability of a forest system to retain N. Prescott et al. [2004] also suggested that models that do not parameterize litter chemistry in great detail may represent long-term rates of leaf litter decay better than those models which do.

The success or failure of a model depends to a large extent on determining whether or not expected model outputs depend on particular values used for model compartment initialization. Models that are structured to be conservative, by strictly following the rules of mass, energy and electrical charge conservation, and by describing transfer processes within the ecosystem by way of simple linear differential or difference equations, lead to an eventual steady-state solution within a constant input-output environment, regardless of the choice of initial conditions. The particular parameter values assigned to such models determine the rate at which the steady state is approached. One important way to test the proper functioning of model parameterization and initialization is to start the model calculations at steady state, and then impose a disturbance pulse, or a series of disturbance pulses (harvesting, fire events, spaced regularly or randomly). This is to see whether the ensuing model calculations will correspond with known system recovery responses, and whether these calculations will eventually return to the initial steady state. In this, the empirical process formulation is crucial, in that each calculation step must feasibly remain within the physically defined solution space. For example, in the hydro-thermal context of C and nutrient cycling, this means that special attention needs to be given to how variations of "independent" variables, such as soil organic matter, texture, coarse fragment content, phase change (water to ice), soil density and soil wettability, combine deterministically and stochastically to affect subsequent variations in heat and soil moisture flow and retention [Balland and Arp, 2005].

2.3 Data uncertainties and natural variation

Data uncertainties are linked to:

- The high spatial and temporal variations associated with forest soil organic matter and corresponding dynamics of above- and below-ground C and N pools. For example, Johnson et al. [2002] noted that soil C measurements from a controlled multi-site harvesting study were highly variable within sites following harvest, but that there was little lasting effect of this variability after 15-16 years. Soil C estimates of an individual pedon are often determined by the combination of measurements of concentration, bulk density, soil depth, and rock content [Homann et al., 1995]; errors in anyone of these can lead to errors in the overall estimate.
- How to determine parameters needed to define, e.g., forest and vegetation type, climate, soil, productivity, and allocation transfers, and to know, whether these parameters are truly time and/or state-independent. Calibration parameters are, as a rule, fixed within models. They are usually obtained from other models, derived from theoretical considerations or estimated from the product of combinatorial exercises.

[1] http://www.nrel.colostate.edu/projects/century/nrel.htm

- Data definitions and sampling procedures, especially those that are vague and open to interpretation. For example, Gijsman et al. [2002] discussed an existing metadata confusion about determining soil moisture retention in relation to soil bulk density.
- Inadequate sampling strategies, in the context of capturing existing micro- and macro-scale C and N pool variations within forest stands, and across the landscape, at different times of the year. On a regional scale, failure to account for the spatial variation across the landscape and the vertical variation with horizon depth due to microrelief, animal activity, windthrow, litter and coarse woody debris input, human activity and the effect of individual plants on soil microclimate and precipitation chemistry may lead to uncertainty.
- Knowing how errors propagate through the model calculations. For example, soil C and N estimates of individual pedons are generally determined by the combination of measurements of C and N concentration, soil bulk density, soil depth, and rock content [Homann et al., 1995]; errors in anyone of these add to the overall estimation uncertainties.

By definition, process-based models should be capable of reflecting the natural range of variation (NRV) that exists in ecosystems of interest. This is an important issue in forest management. In boreal forest ecosystems, quantifying NRV has become a practical goal because forest managers must provide evidence that justifies their proposed use of silviculture (e.g. harvesting, planting, tending) as a stand replacing agent. NRV has been defined by Landres et al. [1999] as "the ecological conditions, and the spatial and temporal variation in these conditions, that are relatively unaffected by people within a period of time and geographical area to an expressed goal". Assuming that reasonable boundaries of time period, geography and anthropogenic influence can be identified, the manager or scientist must then decide which metrics will be used to quantify NRV. Common metrics include mean, median, standard deviation, skewness, frequency, spatial arrangement and size and shape distributions [Landres et al., 1999]. The adoption of NRV as a guiding principal of forest resource management is well-suited to boreal systems because 1) large, stand destroying natural disturbances continue to dominate in much of the boreal and 2) such disturbances may be reasonably emulated by forest harvesting [Hauessler and Kneeshaw, 2003].

The boreal forest is a region where climate change is predicted to significantly affect the survival and growth of native species. Consequently, policies and social pressures (e.g. Kyoto Protocol, Certification) may intensify efforts to improve forest C sequestration by reducing "low-value" wood harvesting. However, high prices for crude oil and loss of traditional pulp and paper wood markets may do the opposite by identifying "low-value" forest biomass as a readily available and profitable energy source. Quantifying NRVs therefore becomes practical as companies and communities charged with forest management have the obligation to provide evidence to justify proposed choices and use of silviculture (e.g. harvesting, planting, tending) as stand-replacing procedures. However, including variables that account for NRV increases the number and costs of required model calibrations, even for simple C and N models. An operational NRV definition is also needed, for practical reasons [Ride, 2004].

Structurally, process-based models often include a choice for the user – "stochastic or mean values". Stochastic runs usually require an estimate of the variation in some aspect of the system of interest. For example, CENTURY has a series of parameters that describe the standard deviation and skewness values for monthly precipitation as main drivers of ecosystem process calculations. This allows the model to vary precipitation, but not air temperature. Another option in CENTURY allows one to write weather files that provide monthly values for temperature and precipitation. However, neither of these options allows for stochasticity in stand replacing events that subsequently affect drivers, such as moisture or temperature, and processes, such as decomposition or photosynthesis.

From a philosophical point of view, it makes sense to build NRV into model function. Boreal systems are highly stochastic, the evidence of which can be found in the high level of beta and gamma diversity often reported. From a logistic point of view, however, including variables that account for NRV increases the number of required calibration values and subsequently the cost of calibrating even a simple model. Data describing the natural range of variation is itself hard to come by. More importantly, an operational definition of NRV has not been widely adopted [Ride, 2004].

2.4 Scaling

Natural variation, like many other ecological concepts, is scale dependent. For example, at the landscape level, it may be possible (1) to estimate the range of stand compositions and ages and therefore structures, (2) to determine a reasonable range of climatic conditions (mainly minimum and maximum temperatures and precipitation) for timeframes as long as a few rotations (i.e. several hundred years) and (3) to identify the successional pathways that reflect the interaction of (1) and (2). This information could then be used to provide a framework of stand and weather descriptions within which functional characteristics, such as SOM turnover, growth and nutrient cycling, could be modeled. Assuming that we have reasonable mathematical descriptions of key biological, chemical and physical processes – such as photosynthesis and decomposition, weathering and complexation, soil moisture and compaction – we could then "nest" our models one inside of another. This approach implies that NRV in the pools and fluxes normally included in process-based models is externally driven (i.e. by weather or disturbance) rather than by internal dynamics.

One example of such a model dealing with NRV scaling issues is the General Ensemble Biogeochemical Modelling System (GEMS), which is used to upscale C and N dynamics from sites to large areas with uncertainty measures [Reiners et al., 2002; Liu et al., 2004a, 2004b; Tan et al., 2005; Liu et al., 2006]. GEMS consists of three major components: one or multiple encapsulated ecosystem biogeochemical models, an automated model parameterization system (AMPS), and an input/output processor (IOP). Plot-scales models such as CENTURY [Parton et al., 1987] and EDCM [Liu et al., 2003] can be encapsulated in GEMS. GEMS uses an ensemble stochastic modeling approach to incorporate the uncertainty and variation in the input databases. Input values for each model run were sampled from their corresponding NRV spaces usually described by their statistical information (e.g. moments, distribution, etc.). This ensemble approach enables GEMS to quantify the propagation and transformation of uncertainties and errors from inputs to outputs. The expectation and uncertainty of the model output are given as:

$$E[p(X_i)] = \frac{1}{W}\sum_{j=1}^{w} p(X_{ij}) \quad (1)$$

$$V[p(X_i)] = \frac{1}{W-1}\sum_{j=1}^{w} (p(X_{ij}) - E[p(X_i)])^2 \quad (2)$$

where W is the number of ensemble model runs, and X_{ij} is the vector of EDCM model input values for the jth simulation of the spatial stratum i in the study area, p is a model operator (e.g., CENTURY or EDCM), and E and V are the expectation and variance of model ensemble simulations for stratum i, respectively.

3. MODEL VALIDATION

Model validation remains a subject of debate and is often used interchangeably with verification [Rykiel, 1996]. Rykiel [1996] differentiated both terms by defining verification as the process aiming at demonstrating the consistency of the logical structure of a model and validation as the process that examines the degree to which a model is accurate relative to the goals desired with respect to its usefulness. Validation does not necessarily consist in demonstrating the logical consistency of causal relationships underlying a model [Oreskes et al., 1994]. Other authors have argued that validation can never be fully achieved as models, like scientific hypotheses, can only be falsified, not proved, and have promoted the more neutral term "evaluation" for the process of testing the accuracy of a model's predictions [Smith et al., 1997]. Although model validation can take many forms or include many steps [e.g., Rykiel, 1996], the method that is most commonly used consists in comparing predictions with statistically independent observations. Using both types of data, statistical tests can be performed or indices can be computed. Smith et al. [1997] and Von Gadow and Hui [1999] provide a summary of the indices most commonly used:

$$\text{Mean residual} = \left(\sum (y_i - \hat{y}_i)/n\right) \quad (3)$$

$$\text{Root mean square error} = \left(\sqrt{\sum (y_i - \hat{y}_i)^2 / n - 1 - p}\right) \quad (4)$$

$$\text{Model efficiency} = \left(\sum (y_i - \hat{y}_i)^2 / \sum (y_i - \bar{y}_i)^2\right) \quad (5)$$

$$\text{Variance ratio} = \sum (\hat{y}_i - \bar{\hat{y}}_i)^2 / \sum (y_i - \bar{y}_i)^2 \quad (6)$$

Several examples exist in the literature on the comparison of predictions with observations or field determinations [Smith et al., 1997; Morales et al., 2005], but mostly for traditional empirical growth models in forestry, as part of the procedures used to determine the annual allowable cut within specific forest management units [e.g., Canavan and Ramm, 2000; Smith-Mateja and Ramm, 2002; Lacerte et al.,

2004]. In contrast, reports on a systematic validation of C and N cycle models are rare [e.g., deVries et al., 1995; Smith et al., 1997] and needed. The validation of C and N cycle models based on the comparison of predictions and observations has been more problematic than the validation of traditional empirical growth models. This is principally because forest inventories conducted by government forest agencies or private industry have been conducted for many decades. Therefore, long-term growth and yield data are available for model validation. This is why model testing has been largely based on growth variables, such as annual volume increment [Medlyn et al., 2005]. Although volumetric data can be converted to biomass and C, direct measurements of C and N pools and flows in forest ecosystems have been collected mainly for research purposes and historical datasets are relatively rare. Therefore, it is often difficult to conduct a validation exercise of C and N models based on the comparison of predictions with statistically independent observations.

So, what options exist for the validation of forest-based C and N cycle models? The most logical avenue is the establishment and maintenance of long-term ecological research programs and site installations to generate the data needed for both model formulation and validation. However, these remain extremely costly and don't receive much political favour in this day and age. One alternative consists in using short-term physiological process measurements [e.g., Davi et al., 2005; Medlyn et al, 2005; Yuste et al., 2005], although care should be given to the long-term behavior of the models in predicting C stocks in vegetation and soils [e.g., Barswell et al., 2005]. Recent technological advances in micrometeorological and physiological instrumentation have been significant, such that it is now possible to collect and analyze hourly, daily, weekly or seasonal data under a variety of forest cover types, experimental scenarios and environmental conditions at relatively low cost. For example, Medlyn et al. [2005] validated a model of CO_2 exchange using eddy covariance data. Davi et al. [2005] also used data from eddy covariance measurements for the validation of their C and water model, and closely monitored branch and leaf photosynthesis, soil respiration, and sap flow measurement throughout the growing season for additional validation purposes. The age factor, the effect of which takes so long to study, can be integrated by using a chrono-sequence approach (using stands of different ages on similar sites as a surrogate for time), which deals with validating C and N models by comparing model output with C and N levels and processes in differently aged forest stands of the same general site conditions. There is also the need to develop new methodologies that are able to integrate the above approaches to allow for model validation at fine and coarse time resolution.

4. SENSITIVITY ANALYSIS

Sensitivity analysis consists in analyzing differences in model response to changes in parameter values. This exercise is relatively easy when the model contains few parameters, but can become cumbersome for complex process-based models. It is beyond the scope of this paper to review all the different methods that have been used, but one of the best examples of sensitivity analysis for process-based models may be found in Komarov et al. [2003], who carried out the sensitivity analyses for EFIMOD 2. These authors showed that the tree sub-model is highly sensitive to changes in the reallocation of the biomass increment and tree mortality functions while the soil sub-model is sensitive to the proportion and mineralization rate of stable humus in the mineral soil. The model is very sensitive to all N compartments, including the N required for tree growth, N withdrawal from senescent needles, soil N and N deposition from the atmosphere. For example, the prediction of stem biomass is sensitive to the N concentration in needles after abscission (Figure 1), reflecting the degree to which the plant (tree) controls growth by retention and internal N reallocation [Nambiar and Fife, 1991]. However, although uncertainty surrounds initial stand density (often unknown), modelled soil C and N and tree stem C (major source of carbon input to the soil sub-model) are not very sensitive to initial stand density (Figure 2).

5. CONCLUSIONS

Many approaches have been developed and used to calibrate and validate process-based models. Models of the C and N cycles are generally based on sound mathematical representations of the processes involved. However, as previously mentioned, the majority of these models are deterministic. As a consequence, they do not represent adequately the error that may arise from different sources of variation. This is important, as both the C and N cycles (and models thereof) contain many sources of variation. Much can be gained by improving the use of calibration and validation methodologies both for

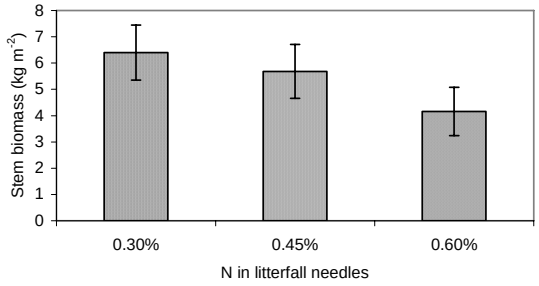

Figure 1. Sensitivity of simulated stem biomass to N content in needles after abcission.

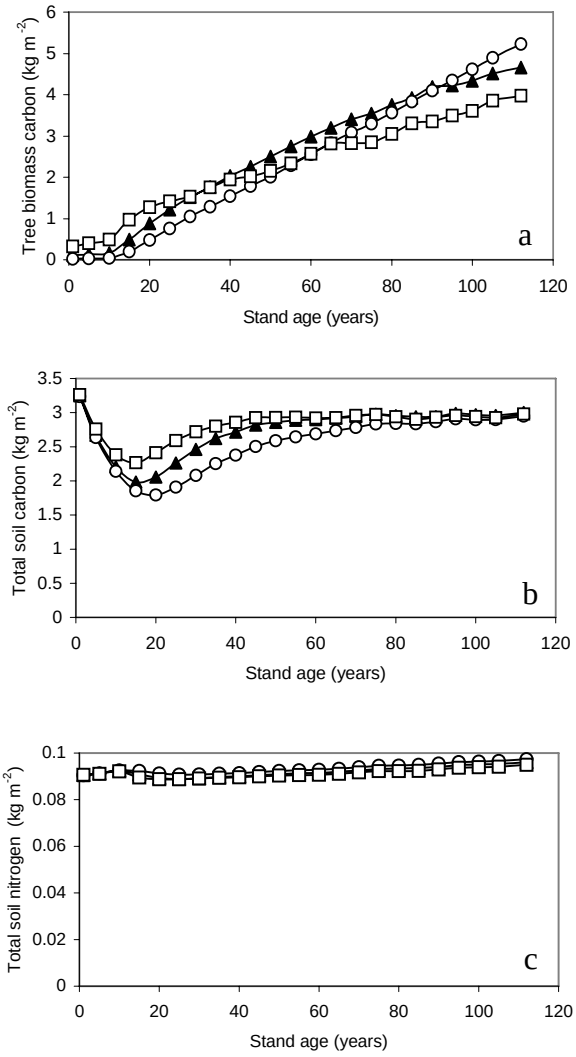

Figure 2. Sensitivity of simulated a) tree biomass carbon, b) total soil carbon and c) total soil nitrogen by EFIMOD 2 to initial stand density.

scientists involved in the modelling of these cycles and forest managers who utilize the results.

Upscaling C dynamics from sites to regions is complex and challenging. It requires the characterization of the heterogeneities of critical variables in space and time at the scales that are appropriate with the ecosystem models, and the incorporation of these heterogeneities into field measurements or ecosystem models to estimate the spatial and temporal change of C stocks and fluxes. The success of upscaling depends on a wide range of factors, including the robustness of the ecosystem models across the heterogeneities, necessary supporting spatial databases or relationships that defined the frequency and joint frequency distributions of critical variables, and the right techniques that incorporate these heterogeneities into upscaling processes. Natural and human disturbances of landscape processes (e.g., fires, diseases, droughts, and deforestation), climate change, as well as management practices will play an increasing role in defining the courses of carbon dynamics at local to global scales. Therefore, methods must be developed to characterize the change of these processes in time and space.

6. ACKNOWLEDGEMENTS

The authors wish to thank the organizing committee of IEMSs 2006 meeting for providing us with the opportunity to organize the workshop.

7. REFERENCES

Balland, V., and P.A. Arp, Modeling soil thermal conductivities over a wide range of conditions, *Journal of Environmental Engineering Science*, 4(6), 549-558, 2005.

Braswell, B.H., W.J. Sacks, E. Linder, and D.S. Schimel, Estimating diurnal to annual ecosystem parameters by synthesis of a carbon flux model with eddy covariance net ecosystem exchange observations. *Global Change Biology*, 11(2), 335–355, 2005.

Canavan, S.J., and C.W. Ramm, Accuracy and precision of 10-year predictions for Forest Vegetation Simulator – Lake States. *Northern Journal of Applied Forestry*, 17(2), 62-70, 2000.

Chertov, O.G., and A.S. Komarov, SOMM: A model of soil organic matter dynamics. *Ecological Modelling*, 94(2-3), 177-189, 1997.

Currie, W.S., and J.D. Aber, 1997, Modeling leaching as a decomposition process in humid Montane forests, *Ecology*, 78(6), 1844-1860, 1997

Davi, H., E. Dufrêne, A. Granier, V. LeDantec, C. Barbaroux, C. Francois, and N. Bréda, Modelling carbon and water cycles in a beech forest: Part II. Validation of the main processes from organ to stand scale, *Ecological Modelling*, 185(2-4), 387-405, 2005.

Davidson, E.A., and I.A. Janssens, Temperature sensitivity of soil carbon decomposition and feedbacks to climate change, *Nature*, 440, 165-173, 2006.

De Vries, W., M. Posch, T. Oja, H. van Oene, H. Kros, P. Warfvinge, and P.A. Arp, Modelling critical loads for the Solling spruce site, *Ecological Modelling*, 83(1-3), 283-293, 1995.

Dufrêne, E., H. Davi, C. François, G. Le Maire, V. Le Dantec, and A. Granier, Modelling carbon and water cycles in a beech forest: Part I: Model description and uncertainty analysis on modelled NEE, *Ecological Modelling*, 185(2-3), 407-436, 2005.

Fang, C., P. Smith, and J.U. Smith, A simple equation for simulating C decomposition in a multi-component pool of soil organic matter, *European Journal of Soil Science*, 56(6), 815-820, 2005.

Franko, U., B. Oelschlagel, and S. Schenk, Simulation of temperature-, water- and nitrogen dynamics using the model CANDY, *Ecological Modelling*, 81(1-3), 213-222, 1995.

Gijsman, A.J., S.S. Jagtap, and J.W. Jones, Wading through a swamp of complete confusion: how to choose a method for estimating soil water retention parameters for crop models, *European Journal of Agronomy*, 18(1-2), 77-105, 2002.

Haeussler, S., and D. Kneeshaw, Comparing forest management to natural processes. Pp. 307 – 368. *In* Burton, P.J., C. Messier, D.W. Smith and W.L. Adamowicz (eds.) Towards sustainable management of the boreal forest. NRC Research Press, Ottawa, ON. 1039 pp., 2003.

Homann, P.S., P. Sollins, H.N. Chappell, and A.G. Stangenberger, Soil organic carbon in a mountainous forested region: relation to site characteristics, *Soil Science Society of America Journal*, 59(5), 1468–1475, 1995.

Johnson, D.W., J.D. Knoepp, W.T. Swank, J. Shan, L.A. Morris, D.H. Van Lear, and P.R. Kapeluck, Effects of forest management on soil carbon: results of some long-term resampling studies, *Environmental Pollution*, 116(1), S201-S208, 2002.

Johnson, D. W., and S.E. Lindberg. (Eds.), Atmospheric Deposition and Forest Nutrient Cycling. A Synthesis of the Integrated Forest Study, *Ecological Studies*, 91. 707 pp, 1991.

Kätterer, T., M. Reichstein, O. Andrén, and A. Lomander, Temperature dependence of organic matter decomposition: a critical review using literature data analyzed with different models, *Biology and Fertility of Soils*, 27(3), 258-262, 1998

Knorr, W., and J. Kattge, Inversion of terrestrial ecosystem model parameter values against eddy covariance measurements by Monte Carlo sampling, *Global Change Biology*, 11(8), 1333-1351, 2005.

Komarov, A., O. Chertov, S. Zudin, M. Nadporozhskaya, A. Mikhailov, S. Bykhovets, E. Zudina, E. Zoubkova, EFIMOD 2 - A model of growth and cycling of elements in boreal forest ecosystems, *Ecological Modelling*, 170(2), 373-392, 2003.

Lacerte, V., G.R. Larocque, M. Woods, W.J. Parton, and M. Penner, Testing the Lake States variant of FVS (Forest Vegetation Simulator) for the main forest types of northern Ontario. *Forestry Chronicle*, 80(4), 495-506, 2004.

Laiho, R., and C.E. Prescott, Decay and nutrient dynamics of coarse woody debris in northern coniferous forests: a synthesis, *Canadian Journal of Forest Research*, 34(4), 763-777, 2004.

Landres, P.B., P. Morgan, and F.J. Swanson, Overview of the use of natural variability concepts in managing ecological systems, *Ecological Applications*, 9(4), 1179-1188, 1999.

Liu, J., S. Liu and T. Loveland, Temporal evolution of carbon budgets of the Appalachian forests in the U.S. from 1972 to 2000, *Forest Ecology and Management*, 222(1-3), 191–201, 2006.

Liu, S., N. Bliss, E. Sundquist, and T.G.Huntington, Modeling carbon dynamics in vegetation and soil under the impact of soil erosion and deposition, *Global Biogeochemical Cycles*, 17(2), 1074–1097, 2003.

Liu, S., M. Kaire, E. Wood, O. Dialio, and L. Tieszen, Impacts of Land Use and climate Change on Carbon Dynamics in South Central Senegal, *Journal of Arid Environments*, 59(3), 583-604, 2004a.

Liu, S., T.R. Loveland, and R.M. Kurtz, Contemporary carbon dynamics in terrestrial ecosystems in the Southeastern Plains of the United States, *Environmental Management*, 33(S1), S442–S456, 2004b.

Medlyn, B.E., A.P. Robinson, R. Clement, and R.E. McMurtrie, On the validation of models of forest CO_2 exchange using eddy covariance data: some perils and pitfalls, *Tree Physiology*, 25(7), 839-857, 2005.

Morales, P., M.T. Sykes, I.C. Prentice, P. Smith, B. Smith, H. Bugmann, B. Zierl, P. Friedlingstein, N. Viovy, S. Sabaté, A. Sánchez, E. Pla, C.A. Gracia, S. Sitch, A. Arneth, and J. Ogee, Comparing and evaluating process-based ecosystem model predictions of carbon and water fluxes in major European forest biomes, *Global Change Biology*, 11(12), 2211–2233, 2005.

O'Neill, R.V., and B. Rust, Aggregation error in ecological models. *Ecological Modelling*, 7(2), 91-105, 1979.

Oreskes, N., K. Shrader-Frechette, and K. Belitz, Verification, validation, and confirmation of numerical models in the earth sciences, *Science*, 263, 641-646, 1994.

Parton W.J., D.S. Schimel, C.V. Cole, and D.S. Ojima, Analysis of factors controlling soil organic matter levels in Great Plains grasslands, *Soil Science Society of America Journal*, 51(5), 1173-1179, 1987.

Prescott, C.E., L. Vesterdal, C.M. Preston, S.W. Simard, Influence of initial chemistry on decomposition of foliar litter in contrasting forest types in British Columbia, *Canadian Journal of Forest Research*, 34(8), 1714-1729, 2004.

Reiners, W.A., S. Liu, K.G. Gerow, M. Keller, and D.S. Schimel, Historical and future land use effects on N_2O and NO emissions using an ensemble modeling approach: Costa Rica's Caribbean Lowlands as an example. *Global Biogeochemical Cycles*, 16(10), 223-240, 10.1029/ 2001GB001437, 2002.

Ride, K.R., Estimates of ecoregional forest composition derived using modeled bounds of natural variation in northwestern Ontario. Ontatio Ministry of Natural Resources, Northwest Sci. and Info. Thunder Bay, ON. NWSI Tech. Rep. TR-136, 2004.

Rykiel, E.J. jr., Testing ecological models: the meaning of validation, *Ecological Modelling*, 90(3), 229-244, 1996.

Smith, P., J.U. Smith, D.S. Powlson, W.B. McGill, J.R.M. Arah, O.G. Chertov, K. Coleman, U. Franko, S. Frolking, D.S. Jenkinson, L.S. Jensen, R.H. Kelly, H. Klein-Gunnewiek, A.S. Komarov, C. Li, J.A.E. Molina, T. Mueller, W.J. Parton, J.H.M. Thornley, and A.P. Whitmore, A comparison of the performance of nine soil organic matter models using datasets from seven long-term experimental, *Geoderma*, 81(1-2), 153-225, 1997.

Smith-Mateja, E.E. and C.W. Ramm, Validation of the Forest Vegetation Simulator growth and mortality predictions on red pine in Michigan. In Crokkston, N.L., Havis, R.N. (comps.). Second Forest Vegetation Simulator Conference, February 12-14, 2002, Fort Collins, CO, pp. 38-44. USDA Forest Service, Rocky Mountain Research Station, Proceedings RMRS-P-25. 208p, 2002.

Tan, Z., S. Liu, C.A. Johnston, T.R. Loveland, L.L. Tieszen, J. Liu, and R. Kurtz, Soil organic carbon dynamics as related to land use history in the northwestern Great Plains, *Global Biogeochemical Cycles*, 19(3), GB3011, doi: 10.1029/2005GB002536, 2005.

Thornley, J.H.M., and G.R. Cannell, Soil carbon storage response to temperature: an hypothesis, *Annals of Botany*, 87(5), 591-598, 2001.

Van Gadow, K. and G. Hui, Modelling forest development. Kluwer Academic Publishers, Boston, 1999.

Wander, M., Soil organic matter fractions and their relevance to soil function. In Magdoff, F., Weil, R. (Eds.), *Advance in Agroecology*. CRC Press. pp. 67-102, 2004.

Wang, Y-P, R. Leuning, H.A. Cleugh, and P.A. Coppin, Parameter estimation in surface exchange models using nonlinear inversion: how many parameters can we estimate and which measurements are most useful, *Global Change Biology*, 7(5), 495–510, 2001.

Williams, M., P.A. Schwarz, B.E. Law, J. Irvine, and M.R. Kurpius, An improved analysis of forest carbon dynamics using data assimilation. *Global Change biology*, 11(1), 89-105, 2005.

Wu, J., B. Jones, H. Li, and O.L. Loucks, Scaling and uncertainty analysis in ecology. Methods and applications, Columbia University Press, New York, 2005.

Yuste, J.Ca., A. Janssens, and R. Ceulemans, Calibration and validation of an empirical approach to model soil CO_2 efflux in a deciduous forest, *Biochemistry*, 73(1), 209-230, 2005.

Zhang, C.F., F.-R. Meng, J.A. Trofymow, and P.A. Arp, 2006. Modelling mass and nirogen remaining in litterbags for Canadian forest and climate conditons, *Canadian Journal of soil Science* (Submitted), 2006.

Zimmerman, J.K., W.M. Pulliam, D.J. Lodge, C.E. Quinones-Orfila, N. Fletcher, S. Guzman-Grajales, J.A. Parrotta, C.E. Asbury, L.R.

Walker, and R.B. Waide, Nitrogen immobilization by decomposing woody debris and the recovery of tropical wet forest from hurricane damage. *Oikos*, 72(3), 314-322, 1995.

Zimmermann, M., J. Leifeld, M.W.L. Schmidt, P. Smith, and J. Fuhrer, Measured soil organic matter fractions can be related to pools in the RothC model. *European Journal of Soil Science* (in review), 2000

Table 1. Examples of models used for estimating rate of forest litter decomposition.

Model name	Reference	Predicted variables	Initialization variables	Predictor variables	Compartment number	Compartment type	Flows	Parameters	Comments
SOMM	Chertov and Komarov 1997	C and N remaining	Initial C, N, ash content	Annual, monthly or daily soil moisture and temperature estimates	3x C, 3x N; (x represents number of cohorts considered)	C & N litter, fermentation and humus cohorts (leaves, roots, coarse woody debris, etc)	7C, 7N	58 C, 3N	Parameters common across locations, initialized by species (cohort); C/N ratios prescribed per compartment
CENTURY	Parton et al. 1987	C and N remaining	Initial C, N, C/N ratios, lignin	Monthly precipitation and air temperature estimates	5 C, 5 N	Structural, metabolic, active, slow and passive C & N compartments	13 C, 13 N	20 C, 5 N	Parameters common across locations; intialized by species
CANDY	Franko et. al.1995	C and N remaining	Initial C, N	Monthly or daily soil moisture and temperature estimates	3 C, 3 N	Active, metabolic and stable C & N compartments	3 C, 3 N	5 C, 1 N, + 2 climate parameters (differs from original)	Parameters common across locations; decomposition not species specific
DOCMOD	Currie and Aber 1997	C and N remaining; dissolved organic C and N	Initial C, N, by compartment	Annual actual evapotranspiration estimates	5 C, 5 N	Lignin-cellulose, unprotected cellulose, extractives, microbial and humus C & N compartments	11 C, 10 N	17 C, 4 N	Parameters common across locations; C/N of humus prescribed
FLDM	Zhang et. al. 2005	Mass, C and N remaining	Initial mass, C, N; initial ash and acid and non-acid hydrolyzable fractions, or lignin fraction	January & July air temperatures and annual precipitation, by year; or monthly or daily soil moisture and temperature estimates	3 mass, 2 N	Fast C; slow and very slow C & N compartments	3 C, 2 N	11C, 1 N	Parameters common across locations; CIDET calibrated; C/N ratios process determined