Towards Bayesian uncertainty quantification for forestry models used in the U.K. GHG Inventory for LULUCF

Marcel van Oijen & Amanda Thomson
Centre for Ecology and Hydrology (CEH-Edinburgh), Bush Estate, Penicuik EH26 0QB, U.K., mvano@ceh.ac.uk

ABSTRACT

The GHG Inventory for the U.K. currently uses a simple carbon-flow model, CFLOW, to calculate the emissions and removals associated with forest planting since 1920. Here we aim to determine whether a more complex process-based model, BASFOR, could be used instead of CFLOW. The use of a more complex approach allows accounting for spatial heterogeneity in soils and weather, but places extra demands on uncertainty quantification. We show how Bayesian methods can be used to address this problem.

INTRODUCTION

Quantifying a GHG Inventory is a problem of incomplete information. No amount of data collection will provide us with a full inventory, so additional calculations and assumptions are required. In the case of LULUCF in the U.K., process-based models are used to quantify net CO_2 emissions associated with afforestation, reforestation and deforestation, based on forestry data and soil type information. The model currently used for forests planted after 1920 is CFLOW. This is a simple compartmental flow model for the carbon cycle which uses measured wood productivity as input and calculates the flows of carbon to tree parts and soil, with different turnover rates for the various compartments. We are investigating the scope for replacing CFLOW with a more complex process-based model, BASFOR, that can better take into account the spatial distribution of climate and soil properties across the U.K. However, the use of the models is hampered by incomplete knowledge of input variables as well as model parameters. This causes uncertainty in the model outputs which needs to be quantified and reported in the Inventory. A highly effective means of quantifying uncertainty in inputs, parameters and outputs of process-based models is Bayesian Calibration (BC; Van Oijen et. al. 2005). The key strength of the method is that it not only propagates uncertainty in inputs and parameters to model outputs, but also uses data on output variables to reduce the uncertainty in inputs and parameters. Here we shall demonstrate the application of BC to BASFOR, and show predictions of carbon sequestration including their uncertainty.

METHODS

Model BASFOR

The BASic FORest simulator, BASFOR, is a process-based forest model that simulates carbon and nitrogen cycling in trees, soil organic matter and litter (Van Oijen *et al*., 2005). It simulates the response of trees and soil to radiation, temperature, precipitation, humidity, wind speed, atmospheric CO_2 and N-deposition, and thinning regime. The model has 11 state variables, representing carbon and nitrogen pools in trees and soil, and 32 parameters controlling the rate of physiological processes and morphological characteristics. The model is deterministic and is solved by Euler integration with a time step of one day.

Weather data

Weather data were taken from the UKCIP climate scenarios (Hulme & Jenkins, 1998). For future weather, only the "Medium-high" scenario was used. The data are given for a regular spatial grid of 655 cells of 20 by 20 km each. Current spatial gradients for temperature and precipitation are dominated by latitudinal and longitudinal effects, respectively. Future warming is expected to show a decreasing pattern from the South-East to the North-West.

Atmospheric CO_2

Atmospheric CO_2 concentration has increased from 300 ppm in 1920 to current levels of around 380 ppm, with an average for the period 1920-2000 of 325 ppm. For the average CO_2 level in the period 2000-2080, the Bern model (Joos et al., 1996) predicts a value of 480 ppm.

N-deposition

Early 20th century levels of N-deposition were low across Europe (< 3 kg N ha^{-1} yr^{-1}) (Galloway, 1985). Data and calculations by the Co-operative Programme for Monitoring and Evaluation of the Long-Range Transmission of Air Pollutants in Europe (EMEP) show increasing N-deposition values during most of the 20th century with maxima reached around 1990. The 1999 Gothenburg Protocol to Abate Acidification, Eutrophication and Ground-level Ozone sets emission ceilings for 2010 for NOx, ammonia and other pollutants. Hence we assumed continued reductions of N-deposition until the year 2010, with deposition remaining constant thereafter. These temporal patterns were spatially disaggregated using the 2004 UK deposition map (R.I. Smith, pers. comm.).

Soils

Data on soil nitrogen, carbon and plant available water content were taken from the global soils database produced by the Data and Information Services of the International Geosphere-Biosphere Programme (IGBP-DIS, Global Soil Data Task 2000).

Tree data from sites Dodd Wood and Rheola

Forest Research U.K. provided data on tree growth and soil characteristics from two Sitka spruce stands, for use in model calibration (R. Matthews & P. Taylor, pers. comm.). The sites were Dodd Wood (54.64 °N, 3.17 °W, alt. 381 m., indurated brown earth sandy soil) and Rheola (51.74 °N, 3.68 °W, alt. 220 m., brown earth soil). Trees were planted in 1927 and 1935, respectively, and management followed a 5-year thinning cycle on both sites.

Bayesian calibration and uncertainty quantification

The parameters of the BASFOR model were quantified by means of Bayesian calibration, using the Forest Research data for Dodd Wood and Rheola. The procedure began with quantifying the uncertainty about the parameter values in the form of a prior probability distribution, based on literature data on conifer growth. The Forest Research data on model output variables were used to update the parameter distribution by application of Bayes' Theorem [$p(\theta|D) \propto p(D|\theta)\ p(\theta)$, where θ is the parameter vector and D is the data]. This yielded a posterior, calibrated probability distribution for the parameters. The predictive uncertainty of the model was then quantified by running the model with different parameter settings, sampled from the posterior distribution (n=5), using Markov Chain Monte Carlo (MCMC) simulation (Van Oijen et al. (2005)). One limitation of the present study was that only the uncertainty in model parameters was quantified. Uncertainty in model drivers (climate, soils) was not quantified, nor was the uncertainty relating to the structure of the BASFOR model itself assessed.

RESULTS

Bayesian calibration and uncertainty quantification

Table 1 lists the major parameters of BASFOR, with their prior uncertainty before application of data from UK forests, and their posterior uncertainty after Bayesian Calibration. For most parameters, prior uncertainty was large, i.e. lower and upper limits were far apart. Figure 1 (black dotted lines) shows for four model output variables (tree and soil carbon, height and total produced wood volume) how the prior parameter uncertainty effected uncertainty in model outputs at the Dodd Wood site. For example, the uncertainty interval (2 standard deviations wide) for tree carbon at the end of the eighty-year rotation ranged from below 40 to above 80 ton carbon ha^{-1}. Table 1 and Figure 1 also show to what extent uncertainties were reduced by the Bayesian calibration using the data from the Dodd Wood and Rheola sites, described above. The marginal posterior probability distributions were much narrower than the prior distributions, as can be seen from the small coefficients of variation. The data from Dodd Wood were not equally informative for all parameters, with CVs for three parameters – initial leaf and stem carbon content and N/C ratio of wood – exceeding 20%. However, Figure 1, red unbroken lines, shows that overall parameter uncertainty had been reduced enough to significantly reduce output uncertainty for the four selected variables.

C-sequestration 1920-2000

The calibrated model was applied to calculate UK-wide C-sequestration between 1920 and 2000 for a standardized conifer rotation with a 5-yearly thinning interval (Figure 2). C-sequestration was defined as the average annual total accumulation of carbon in soil, standing biomass and wood removed at thinnings. Calculated sequestration rates were highest in the South-West of the country, which combines moderately high temperature and precipitation. The far North is identified by the model as an area of net C-source rather than a sink (Figure 2). The spatial pattern of C-sequestration was not closely related to the spatial distribution of atmospheric N-deposition and soil nitrogen. The propagation of parameter uncertainty to uncertainty about C-sequestration rates was calculated by taking five samples from the posterior parameter probability distribution (Table 1) and calculating the standard deviation for the five different results. Figure 3 shows the resulting map of sequestration uncertainty. The spatial pattern of sequestration uncertainty differs strongly from that of sequestration itself (Figure 2), indicating that the coefficient of variation varies between different growing conditions.

C-sequestration 2000-2080

The same calculations of C-sequestration were repeated for the environmental conditions expected for the period 2000-2080. Figure 4 shows the spatial distribution of expected changes in sequestration, relative to 1920-2000. The changes are not closely related to the magnitude of expected changes in temperature, as the spatial patterns differ. However, some degree of warming is expected across the whole country, causing C-sequestration to change mainly in the higher, colder regions of Wales, North-England and Scotland.

Analysis in terms of environmental change factors: climate, CO_2, N-deposition

The preceding UK-wide assessments of the effects of environmental change on expected C-sequestration rates in conifer forests did not separate out the effects of the different environmental factors subject to change. For the purpose of such analysis, we ran additional simulations for the Dodd Wood site with a range of temperatures, atmospheric CO_2 concentrations and N-deposition rates, in a full-factorial set-up. Average temperature was varied from 6.8 to 9.9 °C (which amounts to expanding the UKCIP-estimates for the site for 1920-2000 and 2000-2080 with one degree on either side of the range), atmospheric CO_2 was varied from 320 to 480 ppm (corresponding to changes estimated by the Bern model using the IS92a emissions scenario for 1920-2000 and 2000-2080), and N-deposition was varied from 0 to double the 1920-2000 average value of 8.0 kg N ha^{-1} y^{-1}. Table 2 summarizes the results of application of the model for these environmental conditions. The first data column of the table lists the average values of yield class and annual C-sequestration rate across the considered set of environmental conditions, with standard deviations indicating the uncertainty arising from both the variation in environmental conditions as well as the parametric uncertainty determined before. The final three data columns of Table 2 give the average effect on yield class and sequestration of changes in temperature, CO_2 and N-deposition, with uncertainties. On the examined site, Dodd Wood, changes in each of the three environmental factors has an effect on the output variables, but with the strongest effect (relative to its expected degree of change) for CO_2. The analysis further suggests that C-sequestration rates are likely to increase to similar extent in soils and in tree biomass.

DISCUSSION AND CONCLUSIONS

This study has tried out methods that may be used to improve the construction of the UK GHG inventory. The process-based forest model BASFOR was parameterised efficiently using Bayesian calibration, allowing for uncertainty quantification when using the model to calculate UK-wide conifer forest C-sequestration and yield class. However, the procedure likely suffered from low quality of some data, in particular those on soils.

Uncertainties

Throughout our study we found relatively little sensitivity of UK forest C-sequestration rates and yield class to soil nitrogen content and atmospheric N-deposition, as opposed to the calculated sensitivities to changes in temperature and atmospheric CO_2 concentration. This finding may be an artefact from the use of the IGBP-DIS dataset with its possibly overestimated values of nitrogen contents of UK soils, leading to apparent nitrogen saturation (Van Oijen & Jandl, 2004).

The impacts of changes in environmental factors

The use of a process-based model for calculating C-sequestration, rather than the semi-empirical model CFLOW currently used in the U.K. GHG Inventory, allowed us to analyse the contributions of changes in temperature, CO_2 and N-deposition to changes in sequestration. However, this analysis should be seen as a proof of concept for the methodology rather than as a high-probability identification of a key environmental variable – given the likely poor quality of the soils data and because the factor analysis should first be repeated for the whole of the UK. The spatial pattern of uncertainties, both expressed in absolute terms and as coefficients of variation showed distinct spatial trends across the country, so not only the calculation of main effects, but also uncertainty quantification needs to be calculated country-wide.

The use of process-based models

The presence of nonlinear individual and interactive effects limits the usefulness of response factors as calculated in Table 2. For example, the yield class temperature response factor of 0.18 ± 0.05 (m^3 ha^{-1} y^{-1}) (°C$)^{-1}$ does not necessarily apply outside the Dodd Wood area. This has implications for the way in which we can use results from the process-based modelling to derive modifiers for the yield class values that are used as input for the carbon inventory calculations using CFLOW. The yield class modifiers likely need to be complex multivariate functions of the set of different environmental factors. However, we can calculate such functions if we redo the current factor analysis at a UK-wide scale and with improved input information. This needs to be accompanied by quantification of the uncertainties from incomplete knowledge of parameters, environmental drivers and model structure.

ACKNOWLEDGMENTTS

We thank Robert Matthews and Paul Taylor (Forest Research, Alice Holt) for providing information about the Sitka spruce stands at Dodd Wood and Rheola.

REFERENCES

Galloway, J.N. (1985). The deposition of sulphur and nitrogen from the remote atmosphere. In: Biogeochemical Cycling of Sulphur and Nitrogen in the Remote Atmosphere (edited by J.N. Galloway, R.J. Charlson, M.O. Andrew and H. Rodhe). NATO ASI Series D. Reidel, Dordrecht. Pp. 143-175.

Global Soil Data Task (2000). Global Gridded Surfaces of Selected Soil Characteristics (IGBP-DIS). International Geosphere-Biosphere Programme – Data and Information Services. [http://www.daac.ornl.gov]

Hulme, M. & Jenkins, G.J. (1998). Climate Change Scenarios for the United Kingdom: Scientific Report. UKCIP Technical report Report No. 1. Climatic Research Unit, Norwich, 80 pp. [http://www.cru.uea.ac.uk/link/ukcip/ukcip_report.html].

Joos, F., Bruno, M., Fink, R., Siegenthaler, U., Stocker, T.F., Le Quere, C. and Sarmiento, J.L. 1996. An efficient and accurate representation of complex oceanic and biospheric models of anthropogenic carbon uptake. Tellus 48B: 397–416.

Van Oijen, M. & Jandl, R. (2004). Nitrogen fluxes in two Norway spruce stands in Austria: An analysis by means of process-based modelling. Austrian Journal of Forest Science 121: 167-172.

Van Oijen, M., Rougier, J. & Smith, R. (2005). On Bayesian calibration and evaluation of process-based forest models: bridging the gap between models and data. Tree Physiology 25: 915-927.

Table 1. Prior and posterior probability distributions for parameters of BASFOR. The prior is beta-distributed between specified lower and upper limits. The posterior, derived using data from Dodd Wood and Rheola, is not analytical and is characterized here by the mean values of the marginal parameter probability distribution and the coefficients of variation (CV = standard deviation / mean) (posterior correlation matrix not shown).

Parameter vector			Prior probability distribution		Posterior probability distribution	
Symbol	Unit	Meaning	Lower limit	Upper limit	Mean	CV
$C_{B,0}$	(kg m^{-2})	Initial value branch C	0.00005	0.005	0.0010	0.18
$C_{L,0}$	(kg m^{-2})	Initial value leaf C	0.0001	0.01	0.0015	0.38
$C_{R,0}$	(kg m^{-2})	Initial value root C	0.0001	0.01	0.0017	0.16
$C_{S,0}$	(kg m^{-2})	Initial value stem C	0.00005	0.005	0.00090	0.34
B	(-)	CO_2-response factor	0.4	0.6	0.52	0.06
$CO_{2,0}$	(ppm)	CO_2-response base level	320	380	362	0.02
f_B	(-)	Allocation to branches	0.25	0.30	0.29	0.02
$f_{L,max}$	(-)	Maximum allocation to leaves	0.27	0.37	0.29	0.03
f_S	(-)	Allocation to stem	0.25	0.3	0.28	0.01
Γ	(-)	Respiration fraction	0.4	0.6	0.48	0.06
k_{CA}	(m^2)	Crown area allometric normalisation constant	5	15	11	0.12
$k_{CA,exp}$	(-)	Crown area allometric exponent	0.3	0.45	0.36	0.07
k_h	(m)	Tree height allometric normalisation constant	4	12	7.5	0.07
$k_{h,exp}$	(-)	Tree height allometric exponent	0.2	0.3	0.26	0.04
LAI_{max}	(m^2 m^{-2} mm^{-1})	Maximum LAI	4	10	6.3	0.06
LUE_0	(kg MJ^{-1})	Light-Use Efficiency	0.001	0.003	0.0014	0.10
$NC_{L, max}$	(kg kg^{-1})	Maximum C/N ratio leaves	0.02	0.05	0.028	0.12
$NC_{R,con}$	(kg kg^{-1})	C/N ratio roots	0.02	0.04	0.023	0.06
$NC_{W,con}$	(kg kg^{-1})	C/N ratio woody parts	0.0005	0.002	0.00080	0.23
SLA	(m^2 kg^{-1})	Specific Leaf Area	5	40	6.0	0.05
T_{opt}	(◦C)	Temperature optimum	12	28	19	0.12
$TC_{L,max}$	(d)	Maximum survival time coefficient leaves	365	1460	1048	0.09
δ	(kg C m^{-3})	Wood density	150	250	182	0.04

Table 2. Simulated change in average yield class and annual C-sequestration at the Dodd Wood site due to changes in temperature, CO2 and N-deposition. The standard deviations are due to uncertainty in parameterisation and to variation in interacting environmental factors, but not including soil characteristics.

		Impact of environmental change		
Ecosystem variable	Dodd Wood value	Effect of temperature (per °C)	Effect of [CO_2] (per 100 ppm)	Effect of N-deposition (per 10 kg N ha^{-1} y^{-1})
Yield class (m^3 ha^{-1} y^{-1})	7.91 ± 1.11	0.18 ± 0.05	1.32 ± 0.38	0.74 ± 0.26
C-sequestration (t C ha^{-1} y^{-1})	3.99 ± 0.64	0.10 ± 0.03	0.76 ± 0.21	0.41 ± 0.14
C-sequestration, soil (t C ha^{-1} y^{-1})	1.58 ± 0.31	0.05 ± 0.01	0.36 ± 0.10	0.18 ± 0.07
C-sequestration, trees and products (t C ha^{-1} y^{-1})	2.41 ± 0.34	0.05 ± 0.02	0.40 ± 0.12	0.23 ± 0.07