Biogeosciences, 10, 1717–1736, 2013
www.biogeosciences.net/10/1717/2013/
doi:10.5194/bg-10-1717-2013

Biogeosciences Open Access

Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations

K. E. O. Todd-Brown[1], J. T. Randerson[1], W. M. Post[2], F. M. Hoffman[1,6], C. Tarnocai[4], E. A. G. Schuur[5], and S. D. Allison[1,3]

[1]Department of Earth System Science, University of California, Irvine, CA 92697, USA
[2]Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831–6335, USA
[3]Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697, USA
[4]Research Branch, Agriculture and Agri-Food Canada, Ottawa, Ontario, Canada K1A 0C6, Canada
[5]Department of Biology, University of Florida, Gainesville, FL 32611, USA
[6]Computer Science & Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831–6335, USA

Correspondence to: K. E. O. Todd-Brown (ktoddbro@uci.edu)

Received: 28 September 2012 – Published in Biogeosciences Discuss.: 18 October 2012
Revised: 7 February 2013 – Accepted: 13 February 2013 – Published: 13 March 2013

Abstract. Stocks of soil organic carbon represent a large component of the carbon cycle that may participate in climate change feedbacks, particularly on decadal and centennial timescales. For Earth system models (ESMs), the ability to accurately represent the global distribution of existing soil carbon stocks is a prerequisite for accurately predicting future carbon–climate feedbacks. We compared soil carbon simulations from 11 model centers to empirical data from the Harmonized World Soil Database (HWSD) and the Northern Circumpolar Soil Carbon Database (NCSCD). Model estimates of global soil carbon stocks ranged from 510 to 3040 Pg C, compared to an estimate of 1260 Pg C (with a 95 % confidence interval of 890–1660 Pg C) from the HWSD. Model simulations for the high northern latitudes fell between 60 and 820 Pg C, compared to 500 Pg C (with a 95 % confidence interval of 380–620 Pg C) for the NCSCD and 290 Pg C for the HWSD. Global soil carbon varied 5.9 fold across models in response to a 2.6-fold variation in global net primary productivity (NPP) and a 3.6-fold variation in global soil carbon turnover times. Model–data agreement was moderate at the biome level (R^2 values ranged from 0.38 to 0.97 with a mean of 0.75); however, the spatial distribution of soil carbon simulated by the ESMs at the 1° scale was not well correlated with the HWSD (Pearson correlation coefficients less than 0.4 and root mean square errors from 9.4 to 20.8 kg C m^{-2}). In northern latitudes where the two data sets overlapped, agreement between the HWSD and the NCSCD was poor (Pearson correlation coefficient 0.33), indicating uncertainty in empirical estimates of soil carbon. We found that a reduced complexity model dependent on NPP and soil temperature explained much of the 1° spatial variation in soil carbon within most ESMs (R^2 values between 0.62 and 0.93 for 9 of 11 model centers). However, the same reduced complexity model only explained 10 % of the spatial variation in HWSD soil carbon when driven by observations of NPP and temperature, implying that other drivers or processes may be more important in explaining observed soil carbon distributions. The reduced complexity model also showed that differences in simulated soil carbon across ESMs were driven by differences in simulated NPP and the parameterization of soil heterotrophic respiration (inter-model $R^2 = 0.93$), not by structural differences between the models. Overall, our results suggest that despite fair global-scale agreement with observational data and moderate agreement at the biome scale, most ESMs cannot reproduce grid-scale variation in soil carbon and may be missing key processes. Future work should focus on improving the simulation of driving variables for soil carbon stocks and modifying model structures to include additional processes.

1 Introduction

Soil organic carbon is the largest carbon pool in the terrestrial biosphere (Jobbagy and Jackson, 2000), and losses of soil carbon due to climate change could contribute to rising atmospheric CO_2 concentrations. Loss rates for soil carbon through heterotrophic respiration depend on temperature (Davidson and Janssens, 2006; Lloyd and Taylor, 1994), moisture (Orchard and Cook, 1983; Ryan and Law, 2005), and disturbance regimes such as land use change (Post and Kwon, 2000) and fire (Harden et al., 2000). The sensitivity of many of these drivers to climate change creates the potential for feedbacks that may accelerate or decelerate the buildup of greenhouse gases in the atmosphere (Young and Steffen, 2009).

Although field studies and atmospheric–ocean carbon measurements suggest that the terrestrial biosphere is currently a net sink for carbon dioxide (Houghton, 2007; Lund et al., 2009; Le Quéré et al., 2009), it is unclear if this sink will persist as climate changes. Projections from recent Earth system models (ESMs) suggest that the magnitude of the sink is likely to decline in response to climate change over the 21st century (Cramer et al., 2001; Friedlingstein et al., 2006; Koven et al., 2011). However, this response is highly uncertain (Friedlingstein et al., 2006) and depends, in part, on the strength of feedbacks from the nitrogen cycle (Thornton et al., 2009) and the impact of drought stress on net primary production (NPP), tree mortality, and fires (Goulden et al., 2011; Huntingford et al., 2008; Phillips et al., 2009).

In northern ecosystems, permafrost soils contain large stocks of carbon (Tarnocai et al., 2009) that are particularly vulnerable to loss with climate change (Schuur et al., 2008; Zimov et al., 2006), given the large temperature increases expected for the region (Giorgi, 2006). Models of permafrost soil carbon have only recently been integrated into ESMs (Koven et al., 2011) and further improvements in the representation of thermokarst dynamics, peat accumulation, and soil hydrology are needed to reduce uncertainties related to climate–carbon feedbacks in northern biomes.

Because future climate projections depend on the carbon cycle, ESMs must be capable of accurately representing the pools and fluxes of carbon in the biosphere, particularly in soils that store a large fraction of terrestrial organic carbon. However, there have been few quantitative assessments of ESM skill in predicting soil carbon stocks, contributing to uncertainty in model simulations. To help reduce this uncertainty, we analyzed simulated soil carbon from ESMs participating in the 5th Climate Model Intercomparison Project (CMIP5). If ESMs can accurately represent current soil carbon stocks, then we might have more confidence in their ability to predict future stocks under a changing climate (Luo et al., 2012).

Our analysis had three specific goals: (1) quantify the variation in ESM representation of soil carbon stocks, (2) understand the driving factors regulating soil carbon distribution in ESMs, and (3) compare the ESM soil carbon stocks to empirical data. We conducted these analyses at grid ($1^\circ \times 1^\circ$), biome, and global scales across models to assess spatial variability in the data and model simulations. We compared model outputs to the global Harmonized World Soil Database (FAO/IIASA/ISRIC/ISSCAS/JRC, 2012) and the Northern Circumpolar Soil Carbon Database (Tarnocai et al., 2009). We used an additional data set at high latitudes because these areas contain a large percentage of global soil carbon but are difficult to model and to measure empirically. We expected ESMs to represent high latitude soils poorly because many of the terrestrial decomposition models were developed for mineral soils, as opposed to the organic soils found in many high latitude ecosystems (Koven et al., 2011; Neff and Hooper, 2002; Ping et al., 2008). More generally, we expected that the global distribution of soil carbon in the ESMs would be primarily driven by NPP, soil temperature, and soil moisture. We also anticipated that ESMs with more soil carbon pools would be capable of representing a wider range of soil carbon dynamics, and thus would yield more accurate simulations when compared to observations.

2 Materials and methods

We examined soil carbon stocks in 16 ESMs (Tables 1 and S1 in Supplement) from the 5th Climate Model Intercomparison Project (CMIP5). The model simulations were compared with the Harmonized World Soil Database (HWSD) (FAO/IIASA/ISRIC/ISSCAS/JRC, 2012) and high latitude soil carbon stocks from the Northern Circumpolar Soil Carbon Database (NCSCD) (Tarnocai et al., 2009). We analyzed the underlying drivers of soil carbon variability with a set of reduced complexity models.

2.1 Earth system models

ESMs from CMIP5 use common simulation and output protocols, enabling direct comparisons between models. One of the goals of CMIP5 is to facilitate benchmarking of ESMs through the *historical* simulation protocol, which has a prescribed time series of atmospheric CO_2 mixing ratios and land use change (Taylor et al., 2011). ESMs were selected from the CMIP5 repository based on the availability of soil carbon and other key output variables for the *historical* simulation, as well as consultation with modeling centers.

The model structure for soil carbon across ESMs was relatively uniform (Table 1). The soil carbon sub-models in all ESMs represented decomposition as a first-order decay process involving 1–9 dead (soil or litter) carbon pools. The temperature sensitivity of decomposition in most ESMs was described by the Q_{10} or Arrhenius equations, which are functionally similar (Davidson and Janssens, 2006; Lloyd and Taylor, 1994). In the Q_{10} form of the temperature sensitivity function, the decomposition rate is modified as a function

Table 1. Summary of soil carbon models including Earth system model names, history of model development, number of litter and soil pools, temperature and moisture functions, and representation of nitrogen cycling.

Model name	Soil model history	Litter	Soil	Temperature	Moisture	Nitrogen
BCC-CSM1.1 Wu et al. (2013)	AVIM2; Huang et al. (2007), Ji et al. (2008) CEVSA; Cao and Woodward (1998) CENTURY; Parton et al. (1987, 1993)	2	6	Hill[a]	Hill	Yes
CanESM2 (CMIP5 output)	CTEM1; Arora and Boer (2005), Arora (2003), Arora et al. (2011)	1[b]	1[b]	Q_{10}^{c}	Hill	No
CCSM4 Gent et al. (2011)	CLM4; Oleson et al. (2008) CN; Thornton et al.(2007) Biome-BCG 4.1.2 Kimball et al. (1997), Thornton and Rosenbloom (2005), Thornton (1998), Thornton et al. (2002), Olson (1963), Veen and Paul (1981), Veen et al. (1984)	3	3	Arrhenius	Increasing	Yes
GFDL-ESM2G (CMIP5 output)	LM3 (LM3p7_cESM, M45) Shevliakova et al. (2009) ED; Moorcroft et al. (2001), Bolker et al. (1998) CENTURY; Parton et al. (1987)	–	2	Hill	Increasing	No
GISS-E2-H GISS-E2-R (N. Kiang, personal communication, 2012)	NCAR-CSM1.4; Doney et al. (2006) NASA-CASA; Potter et al. (1993), Randerson et al. (1997)	–	9	Increasing	Increasing	No
HadGEM2-ES HadGEM2-CC Jones et al. (2011)	Martin et al. (2011), Collins et al. (2011) TRIFFID; Cox (2001)	–	4	Q_{10}	Hill	No
INM-CM4 Volodin et al. (2010)	Volodin (2007) LSM; Bonan (1995, 1996), Bunnell et al. (1977)	–	1[b]	Q_{10}^{c}	Hill	No
IPSL-CM5A-LR IPSL-CM5B-LR (http://icmc.ipsl.fr, 2012)	ORCHIDEE (http://orchidee.ipsl.jussieu.fr/) STOMATE; Krinner et al. (2005) CENTURY; Parton et al. (1988)	3	4	Q_{10}	Increasing	No
MIROC-ESM MIROC-ESM-CHEM Watanabe et al. (2011)	SEIB-DGVM; Sato et al. (2007) Roth-C; Coleman and Jenkinson (1999) DEMETER-1; Foley (1995) CENTURY; Parton et al. (1987, 1992)	–	2	Arrhenius	Increasing	No
MPI-ESM-LR (CMIP5 output)	JSBACH; Raddatz et al. (2007) BETHY; Knorr (2000) CENTURY; Parton et al. (1993)	1	1	Q_{10}	Increasing	No
NorESM1-ME NorESM1-M Tjiputra et al. (2012)	CLM4; Oleson et al. (2008) CN; Thornton et al. (2007) Biome-BCG 4.1.2; Kimball et al. (1997), Thornton and Rosenbloom (2005), Thornton (1998), Thornton et al. (2002), Olson (1963), Veen and Paul (1981), Veen et al. (1984)	3	3	Arrhenius	Increasing	Yes

[a] We define a hill function as a function that increases to a maximum and then decreases. [b] Turnover parameterization dependent on biome or vegetation type. [c] Q_{10} value dependent on temperature.

of temperature (T) relative to a baseline (T_0), such that $f(T) = Q_{10}^{(T-T_0)/10}$. In this equation the Q_{10} value is often set to between 1.5 and 2.5 based on estimates inferred from ecosystem flux measurements (Mahecha et al., 2010; Raich and Schlesinger, 1992) or the annual cycle of atmospheric CO_2 (Kaminski et al., 2002; Randerson et al., 2002). In some of the models, the temperature sensitivity of decomposition follows neither a Q_{10} nor Arrhenius relationship. For example, in BCC-CSM1.1 and GFDL-ESM2G, decomposition rate increases up to some optimal temperature and then decreases (Ji et al., 2008; Parton et al., 1987; Shevliakova et al., 2009). For the GISS-E2 model, the soil respiration response to temperature is a linear fit to data from Del Grosso et al. (2005) up to 30 °C, with a plateau above 30 °C. In all of the models, decomposition either increases monotonically with increasing soil moisture or increases up to some optimum moisture level and then decreases. Three ESMs include nitrogen interactions with soil carbon: CCSM4, NorESM1, and BCC-CSM1.1.

We downloaded soil organic carbon, litter carbon, annual NPP, 2 m air temperature, soil temperature, and total soil water from the *historical* simulation, where available, for each ESM (*cSoil*, *cLitter*, *npp*, *tas*, *tsl*, and *mrso*, respectively, from the CMIP5 variable list). The monthly means for all variables from 1995–2005 were averaged for each grid cell to generate an overall mean for comparison to the HWSD and to use as drivers for our reduced complexity models (see below). We combined litter and soil carbon for our analysis and refer to the sum as soil carbon. Coarse woody debris (*cCwd* from the variable list) was not included in the sum since there is no respiration from this pool in the two models that report it (CCSM4 and NorESM1). Global turnover times for soil carbon were calculated by dividing total soil carbon by total terrestrial NPP from each ESM. INM-CM4 did not report NPP directly, so we derived NPP from gross primary production and autotrophic respiration (*gpp* and *ra* from the variable list). Soil temperatures were reported for each soil layer, but only the top 10 cm mean was used in this analysis. Land area was calculated from the grid area modified by the land cover for each model (*areacella* and *sftlf* from the variable list, respectively). All ensemble members were averaged for each model; however, not all variables were available for each ensemble member. For example, GISS-E2-R reported *cSoil* but not *tsl* for ensemble member r1i1p1 at the time of download.

We performed a hierarchical cluster analysis and found that ESMs from the same climate center generated very similar distributions of soil carbon for 1995–2005 (see Supplement Fig. S1). Clusters were constructed using complete linkage of the Euclidian distances between the global 1°-gridded soil carbon distributions for each model. Models from the same climate center always showed more than 90 % relative similarity and included the following pairs: GISS-E2 H and R, HadGEM2 ES and CC, IPSL-CM5 (LR) A and B, MIROC-ESM and MIROC-ESM-CHEM, and finally NorESM1 ME and M. As a result, these model pairs were averaged. We were left with 11 independent simulations, one representing each modeling center, that were evaluated using the approaches described below.

ESMs do not report the depth of carbon in the soil profile to CMIP5, making direct comparison with empirical estimates of soil carbon difficult. Although many soil models were originally constructed to represent C dynamics at an approximate depth range of 0 to 20 cm (e.g., Kelly et al., 1997), we assumed that all simulated soil carbon was contained within the top 1 m to simplify comparison with data sets. We recommend that future model intercomparison projects request soil carbon output from model simulations with specific depth ranges (for example, soil carbon above 1 m and below 1 m) to allow for more accurate and direct comparison to survey data.

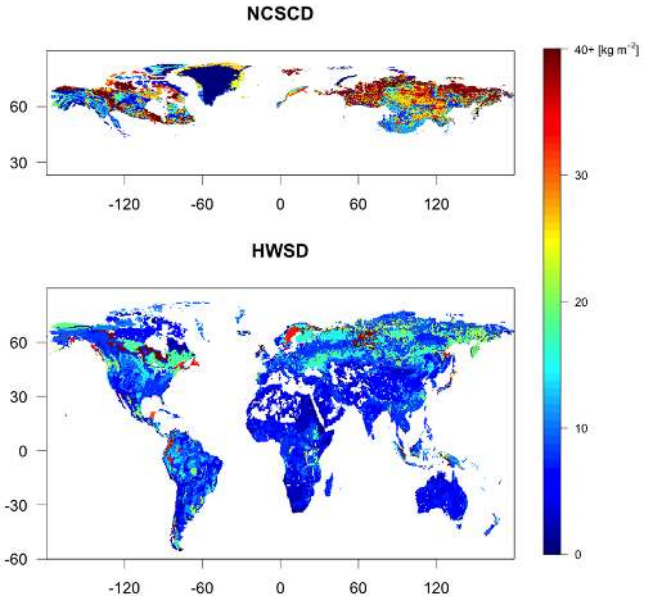

Fig. 1. Carbon density [kg m^{-2}] in the top 1 m of soil from the Northern Circumpolar Soil Carbon Database (NCSCD) (Tarnocai et al., 2009) and Harmonized World Soil Database (HWSD) (FAO/IIASA/ISRIC/ISSCAS/JRC, 2012).

2.2 Data sets

2.2.1 Soil carbon

The HWSD provided empirical estimates of global soil carbon stocks to validate ESM simulations. The HWSD is a product of the Food and Agriculture Organization of the United Nations and the Land Use Change and Agriculture Program of the International Institute for Applied Systems Analysis. The HWSD aggregates data from the European Soil Database (ESDB, 2004), the Soil Map of China (Shi et al., 2004), regional soil and terrain databases (Sombroek, 1984), and the FAO-UNESCO Soil Map of the World (FAO/UNESCO, 1981). Soil carbon stocks were calculated from bulk densities and organic carbon concentrations given in the HWSD for the top 1 m of soil at a $0.5° \times 0.5°$ resolution (Fig. 1). Bulk density estimates in the HWSD were derived from soil texture; however, this approach is not appropriate for high carbon soils (FAO/IIASA/ISRIC/ISSCAS/JRC, 2012; Saxton et al., 1986). Therefore, we replaced Histosol and Andisol bulk densities with values from the World Inventory of Soil Emission Potentials (Batjes, 1996).

Because high latitude soils contain a large fraction of global soil carbon, we also validated ESM simulations of soil carbon in high latitudes with the NCSCD, which is an independent survey of soil carbon (Tarnocai et al., 2009). The NCSCD covers 18.8×10^6 km^2, including areas with different geological histories and stages of soil development. We used the $1° \times 1°$ soil carbon data product for the first meter

of soil (Fig. 1). The spatial and soil data used to develop this database were collected during the last 60 yr and originated from a variety of sources.

Quantitative uncertainty analyses for the HWSD and NCSCD have not been performed and would be a challenge to construct because of the diverse data sources involved. However, some estimate of uncertainty was essential to enable quantitative comparisons with the CMIP5 models. To generate such a range for total soil carbon from both data sets, we constructed preliminary 95 % confidence intervals (CI_{95}) using a qualitative approach. These estimates must be interpreted with caution because they are not based on formal error propagation methods. Furthermore, these estimates only apply to database totals, and uncertainties for individual grid cells are likely to be larger.

For the HWSD, the major sources of error were related to analytical measurement of soil carbon, variation in carbon content within a soil type, and mapping of soil types. Analytical measurements of soil carbon concentrations are generally precise, but measurements of soil bulk density are more uncertain and may contribute to CI_{95} values that are ± 15 % of the mean carbon content for a given soil profile. Soil types in the HWSD are defined based on Food and Agriculture Organization soil taxonomic units that are assumed to experience similar histories of soil forming factors such as climate, vegetation, disturbance, topography, and parent material. Batjes (1997) reported quartiles of soil carbon content for 23 soil taxonomic units based on 18 to 1270 soil profiles per unit. These quartiles suggest that soil carbon content is approximately log-normally distributed, allowing for calculation of CI_{95} values for each soil unit following log transformation. When back transformed, CI_{95} ranged from 6 to 33 % below the median to 6 to 48 % above the median, with an average CI_{95} of 14 % below to 17 % above the median across all 23 units.

Another major source of HWSD uncertainty was related to the mapping of soil units and scaling of soil maps to 0.5°. Soil taxonomic units and associated carbon contents were spatially extrapolated using expert knowledge informed by topography, geology, and vegetation (usually based on aerial photography). Original soil maps were scaled up in the HWSD by classifying each 0.5° grid cell according to its dominant soil unit. We assumed that the uncertainty associated with mapping and scaling was similar in magnitude to measurement error and spatial variation, with a CI_{95} of approximately ± 15 % of the mean. To estimate an overall CI_{95} for the HWSD, we assumed that variation in soil carbon content within soil taxonomic units already included analytical error, and that median carbon content within a soil unit is extrapolated by multiplying by the area of the unit. Thus the CI_{95} values representing variation in soil carbon content and mapping uncertainty were summed to yield an overall CI_{95} of 29 % below the mean to 32 % above the mean, or a range of 890 to 1660 Pg C with a mean of 1260 Pg C. This estimate is broadly consistent with other empirical estimates of global soil carbon (Eswaran et al., 1993; Jobbagy and Jackson, 2000; Sombroek et al., 1993).

For the NCSCD, the uncertainties varied by geographic region. The North American portion of the data set was based on analysis of 1169 pedons producing a medium to high confidence rating (66–80 %). Thus we estimated the CI_{95} for the North American portion of the NCSCD to be 165 ± 17 Pg C, corresponding to ± 10 % of the mean. In Eurasia, soil carbon estimates were based on fewer pedons (591) plus 90 peat cores, producing a low to medium confidence rating (33–66 %). Therefore we estimated the CI_{95} for the Eurasian region to be 331 ± 99 Pg C, or ± 30 % of the mean. Carbon in Yedoma deposits and river deltas was estimated independently using surveyed depth information where available. This deeper soil carbon had the lowest confidence rating but contributed only ~ 1 % or 5 Pg of the database total; therefore we allowed for a CI_{95} of 5 ± 5 Pg C on this estimate. Together, these uncertainty estimates yielded an overall CI_{95} of roughly 500 ± 120 Pg C for the first meter of soil.

2.2.2 Net primary productivity and temperature

To assess model skill in simulating key driving variables that could affect soil carbon stocks, we compared ESM outputs to temperature data from the Climate Research Unit (CRU) and to NPP data from a literature synthesis and from the Moderate Resolution Imaging Spectrometer (MODIS). The CRU and MODIS data also were used in parameter estimation for the reduced complexity models (Eqs. 1–2) to explain the spatial variation in observed global soil carbon with observed temperature and NPP. We used a $0.5° \times 0.5°$ gridded air temperature data set from the CRU, specifically the 1995–2005 mean of the *tmp* variable from CRU_TS 3.10 (Jones and Harris, 2008). For NPP, we used the $0.008° \times 0.008°$ gridded MODIS product MOD17A3 from 2000–2011 (Zhao and Running, 2010). We also compared ESM-simulated NPP to Ito's (2011) value of 54 ± 11 Pg C yr^{-1} (mean $\pm$ standard deviation). This estimate was based on empirical models that used environmental parameters to extrapolate field measurements of NPP to the global scale. We considered ESM-simulated NPP values to be consistent with empirical data if they fell within 2 standard deviations of Ito's (2011) estimate.

2.2.3 Biome map

To evaluate ESM soil carbon across biomes, we aggregated HWSD estimates and model simulations of soil carbon within biomes. The biome map was based on the MODIS land cover product MCD12C1 (Friedl et al., 2010; NASA LP DAAC, 2008) (Fig. S2). We assigned one of 16 land cover types to each $1° \times 1°$ grid cell by taking the most common land cover from the original underlying $0.05° \times 0.05°$ data. Each $1° \times 1°$ grid cell was assigned to one of 9 biomes: tundra, boreal forest, tropical rainforest, temperate forest, desert and shrubland, grasslands and savannas, cropland and urban,

Table 2. Soil carbon totals across all grid cells in each ESM, grid cells present in the HWSD and ESM, and grid cells present in the NCSCD and ESM. Database totals include 95 % confidence intervals based on qualitative uncertainty analysis, shown in brackets. Values are rounded to the nearest 10 Pg C.

				Soil carbon [Pg C]		
Database or model name	Original grid size ° lat. × ° long.	Number of model versions	Ensembles per version	Global total	HWSD and ESM shared	NCSCD and ESM shared
HWSD	0.5 × 0.5	–	–	1260 [890, 1660]	1260	290
NCSCD	1 × 1	–	–	500 [380, 620]	480	480
CCSM4	0.94 × 1.25	1	6	510	510	60
NorESM1	1.89 × 2.50	2	3, 1	550	530	60
BCC-CSM1.1	2.81 × 2.81	1	3	1050	990	240
HadGEM2	1.25 × 1.88	2	1, 2	1120	1090	200
IPSL-CM5	1.89 × 3.75	2	5, 1	1310	1250	390
GFDL-ESM2G	2.01 × 2.50	1	1	1410	1360	690
CanESM2	2.79 × 2.81	1	5	1540	1460	370
INM-CM4	1.50 × 2.00	1	1	1680	1630	280
GISS-E2	2.00 × 2.50	2	15, 16	1900	1820	520
MIROC-ESM	2.79 × 2.81	2	3, 1	2570	2490	810
MPI-ESM-LR	1.86 × 1.88	1	3	3040	2930	340

snow and ice, or permanent wetland. Details for the biome construction can be found in Fig. S2 in the Supplement.

2.3 Regridding approach

All model outputs and data sets were regridded to 1° × 1° for biome- and grid-scale comparisons. Our regridding approach assumed conservation of mass and that a latitudinal degree is proportional to distance for close grid cells. Regridding the outputs to 1° × 1° downscaled the models while upscaling the data (Table 2).

2.4 Assessing driving variables for soil C stocks

We developed several reduced complexity models to evaluate the drivers of simulated soil carbon variability and facilitate comparisons between ESMs. These reduced complexity models consisted of a single pool of soil carbon at each grid cell driven by locally varying NPP, soil temperature, and soil moisture (Figs. S3, S4, S5 in Supplement) and globally uniform parameters including a decomposition rate constant, Q_{10} value, and moisture coefficient. By applying the same simple model, we were able to compare parameters across ESMs and assess which variables had the strongest control over soil carbon. Driving variables for the reduced complexity models were taken from ESM annual means of NPP, soil temperature (T, top 10 cm mean), and total soil water content (W) over the period 1995–2005.

Our reduced complexity models assumed that the soil carbon pool (C) in grid cell i was at steady state, such that NPP inputs equaled outputs from heterotrophic respiration (R):

$$0 = \frac{\mathrm{d}C_i}{\mathrm{d}t} = \mathrm{NPP}_i - R_i.$$

Carbon pools were not expected to be exactly at steady state for 1995–2005, and mean grid differences between NPP and R across the ESMs ranged from 0.01 to 0.12 kg m^{-2} yr^{-1}, or between 1 % and 20 % of the mean grid NPP for this period. Thus the ESMs were close to steady state, and we assumed steady state to simplify our analysis. For the simplest reduced complexity model, we assumed that soil heterotrophic respiration was directly proportional to the soil carbon pool with a spatially uniform decomposition rate constant k (Olson, 1963; Parton et al., 1987):

$$R_i = kC_i.$$

Combining the two above equations yielded the simplest reduced complexity model, Eq. (1), in which soil carbon was proportional to NPP and inversely proportional to a global decomposition rate (k):

$$C_i = \frac{\mathrm{NPP}_i}{k}. \quad (1)$$

We formulated a second reduced complexity model, Eq. 2, in which soil respiration in each grid cell also depended on soil temperature (T) according to a Q_{10} function with a baseline temperature of 15 °C (Lloyd and Taylor, 1994):

$$C_i = \frac{\mathrm{NPP}_i}{kQ_{10}^{(T_i-15)/10}}. \quad (2)$$

A third reduced complexity model, Eq. (3), included a moisture modifier that increased with total soil water content (W_i) normalized to maximal soil water content for each ESM (W_x) according to an exponential function, where b was a positive scaling exponent:

$$C_i = \frac{\mathrm{NPP}_i}{k Q_{10}^{(T_i-15)/10} \left(\frac{W_i}{W_x}\right)^b}. \quad (3)$$

The parameters k, Q_{10}, and b in each reduced complexity model were optimized using ESM soil carbon and driving variables across all grid cells. We used a constrained Broyden–Fletcher–Goldfarb–Shanno optimization algorithm (Byrd et al., 1995), a quasi-Newtonian method, as implemented in R 2.13.1 (R Development Core Team, 2012). This algorithm was selected for parameter fitting because of its robust convergence and short run time. We ran the optimization with the following constraints: $k \in \left(10^{-4}, 10\right)$, $Q_{10} \in (1, 4)$, and $b \in (0, 3)$. The initial parameter estimates were $k = 0.1$, $Q_{10} = 1$, and $b = 0$. We used the root mean square error (RMSE) as the measure function. The optimization was performed only on grid cells with non-zero soil carbon values.

We conducted an additional analysis to assess the causes of variation in simulated soil carbon across ESMs (Eqs. 4–7). For this analysis, we used a modified version of Eq. (2) to predict total global soil carbon (C) for each ESM:

$$C = \sum_i \frac{\mathrm{NPP}_i}{k Q_{10}^{(T_i-15)/10}}, \quad (4)$$

where the Q_{10} and k parameters were derived from fitting Eq. (2) to the spatial distribution of soil carbon (at 1° resolution) from each ESM as described above. Grid-scale NPP outputs (NPP_i) and soil temperatures (T_i) from each ESM were used as drivers in Eq. (4) to calculate soil carbon in each grid cell i. Soil carbon was then summed across all grid cells in each ESM to calculate the global soil carbon pool (C). Thus Eq. (4) represents the contribution of both model parameterization (Q_{10} and k) and soil carbon drivers (NPP_i and T_i) to the global soil carbon pool. To isolate the effect of ESM parameterization on C, we substituted multi-model mean values for NPP ($\overline{\mathrm{NPP}_i}$) and temperature ($\bar{T}_i$) into Eq. (4) for each grid cell i:

$$C = \sum_i \frac{\overline{\mathrm{NPP}_i}}{k Q_{10}^{(\bar{T}_i-15)/10}}. \quad (5)$$

To isolate the effect of ESM driving variables on C, we substituted multi-model mean values for Q_{10} ($\bar{Q}_{10}$) and k ($\bar{k}$) into Eq. (4):

$$C = \sum_i \frac{\mathrm{NPP}_i}{\bar{k} \bar{Q}_{10}^{(T_i-15)/10}}. \quad (6)$$

Finally, we substituted only the multi-model mean temperature into Eq. (4) to isolate the effect of NPP on inter-model variation in C:

$$C = \sum_i \frac{\mathrm{NPP}_i}{k Q_{10}^{(\bar{T}_i-15)/10}}. \quad (7)$$

Using regression analysis, we compared the predicted C from Eqs. (4)–(7) to the totals simulated by the ESMs. These regressions measure the contribution of parameterization (Eq. 5) versus driving variables (Eqs. 6 and 7) to variation in soil carbon totals across ESMs. We excluded GISS-E2 from this inter-model analysis because Eq. (2) could not be fit to this model, and therefore Q_{10} and k were not available.

2.5 Statistical analyses

ESM simulations were compared to data sets using Pearson correlation coefficients, RMSE, and Taylor scores in R 2.13.1 (R Development Core Team, 2012). The Taylor score (T_S) combines the Pearson correlation coefficient (r) and standard deviation (σ) of the model results (m) compared to the data (d):

$$T_S(d, m) = \frac{4[1 + r(d, m)]}{\left[\frac{\sigma(m)}{\sigma(d)} + \frac{\sigma(d)}{\sigma(m)}\right]^2 [1 + r_{\max}]},$$

where $r_{\max}$ is the maximum correlation attainable, assumed to be 1 in this case (Taylor, 2001). Biome-aggregated totals were compared to observations using linear regression.

3 Results

3.1 Global soil carbon stocks and turnover times

The mean (±SD) global soil carbon reported across all ESMs was 1520 ± 770 Pg, whereas the global soil carbon in the HWSD was 1260 Pg with a CI_{95} of 890 to 1660 Pg (Table 2, Fig. 2). CCSM4 reported the lowest total at 510 Pg C and MPI-ESM-LR the highest at 3040 Pg C. Examining only the area shared by all ESMs and the HWSD reduces the global carbon totals but does not substantially change the rank order of the models (Table 2). CCSM4 and NorESM1 underestimated global soil carbon stocks by about 50 %, whereas GISS-E2, MIROC-ESM, and MPI-ESM-LR overestimated global soil carbon stocks anywhere from 50 % to 140 %. The other models predicted global soil carbon totals that were within 35 % of the HWSD global mean and fell within its preliminary CI_{95}.

Soil carbon in the high northern latitudes (as defined by grid cells in the NCSCD) was underestimated by most ESMs relative to the NCSCD (Table 2, Fig. S6). In addition, the rank order of ESM soil carbon totals in this region differed from the rank order based on global totals. CCSM4 and NorESM1 simulated just over 10 % of the total soil carbon

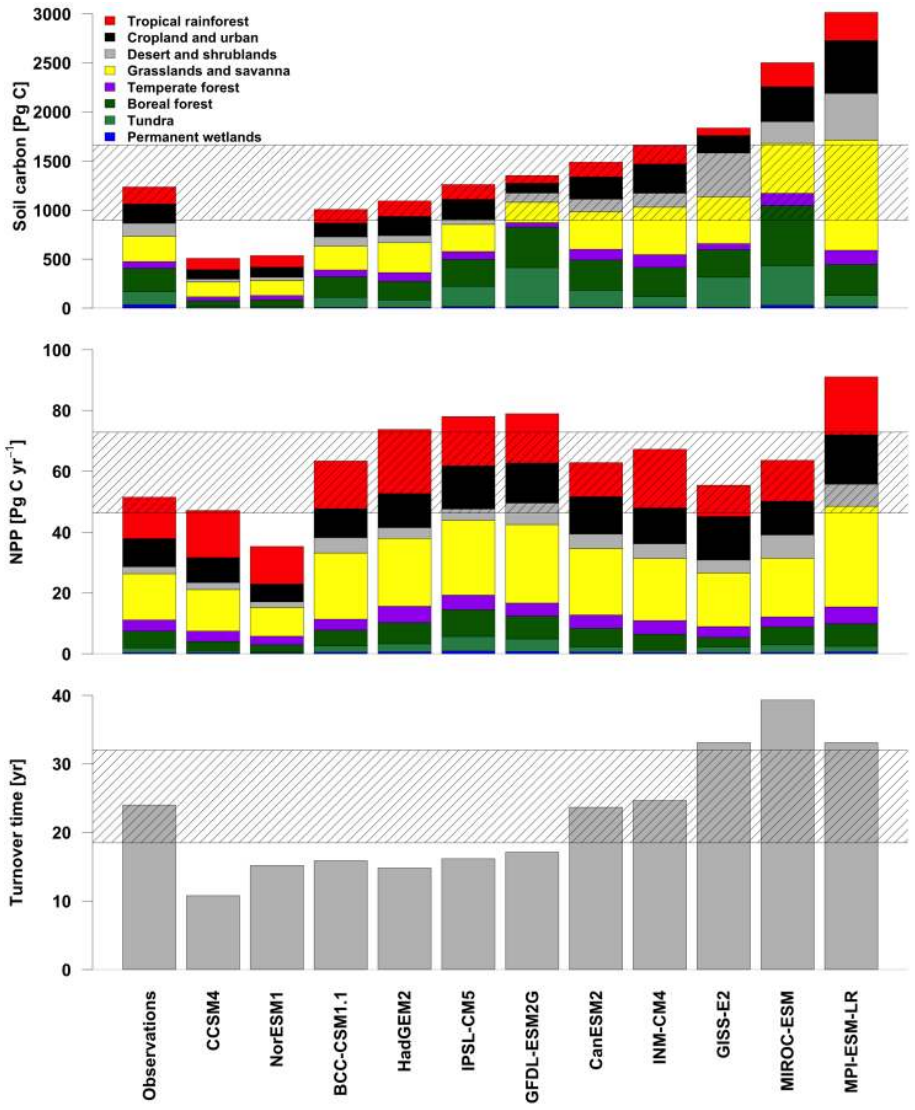

Fig. 2. Global soil carbon (top), net primary production (middle), and soil carbon turnover times (bottom) for observations and ESMs. Turnover times were calculated as HWSD carbon divided by MODIS NPP for the observations, and simulated global soil carbon divided by simulated global NPP for the ESMs. The gray hashed area on the top panel represents the 95 % confidence interval for global soil carbon in the HWSD based on a qualitative uncertainty analysis (see text). The hashed area on the middle panel represents ±2 standard deviations around the mean global NPP estimate from Ito (2011) based on empirical models. The hashed area on the bottom panel indicates the range of turnover times for global soil carbon found in the literature (Amundson, 2001; Raich and Schlesinger, 1992). For soil carbon and NPP, each global estimate is separated into individual biome components according to the legend shown in the top panel.

observed in the NCSCD. HadGEM2, BCC-CSM1.1, INM-CM4, MPI-ESM, and CanESM2 also simulated soil carbon totals below the preliminary CI_{95} for the NCSCD. In contrast, GFDL-ESM2G and MIROC-ESM overestimated high latitude soil carbon stocks by 45–60 %. Only IPSL-CM5 and GISS-E2 soil carbon fell within the CI_{95} for the NCSCD.

Variation in global soil carbon stocks simulated by ESMs could be driven by variation in modeled NPP, and we found that global terrestrial NPP varied by a factor of 2.6 across the models (Fig. 2). CCSM4, BCC-CSM1.1, CanESM2, INM-CM4, GISS-E2, and MIROC-ESM all predicted global NPP values within 2 standard deviations of the Ito (2011) estimate of 54 Pg C yr^{-1}, ranging from 46 to 73 Pg C yr^{-1}, whereas the remaining 5 models fell outside this range. NPP from MODIS was similar to Ito (2011) at 52 Pg C yr^{-1}. At high northern latitudes, NPP estimates from the ESMs were more variable (1.7 to 10.1 Pg C yr^{-1}), compared to a MODIS estimate of 4.7 Pg C yr^{-1} (Fig. S6 in Supplement).

Turnover times for global soil carbon from the ESMs varied by a factor of 3.6, between 10.8 and 39.3 yr, using global

Table 3. Goodness-of-fit measures by grid cell for each ESM including soil carbon versus the HWSD, soil carbon versus NCSCD, soil carbon versus HWSD without NCSCD grid cells (HWSD − NCSCD), NPP versus MODIS NPP, and land 2 m air surface temperature versus CRU 2 m air temperature. T_S = Taylor score; r = Pearson correlation coefficient; RMSE = root mean square error. RMSE has units of kg C m^{-2} for the soil carbon comparisons, kg C m^{-2} yr^{-1} for NPP, and °C for temperature.

	HWSD			NCSCD			HWSD − NCSCD			MODIS NPP			CRU temperature		
	T_S	r	RMSE	T_S	r	RMSE	T_S	r	RMSE	T_S	r	RMSE	T_S	r	RMSE
HWSD	NA	NA	NA	0.60	0.33	20.0	NA	NA	NA	NA	NA	NA	NA	NA	NA
CCSM4	0.21	0.18	11.3	0.05	−0.04	27.7	0.29	0.26	9.1	0.87	0.75	0.25	0.98	0.96	3.8
NorESM1	0.24	0.17	11.2	0.05	−0.07	27.7	0.33	0.25	9.1	0.81	0.71	0.28	0.97	0.96	4.3
BCC-CSM1.1	0.51	0.33	9.4	0.21	0.03	22.2	0.53	0.34	8.0	0.85	0.71	0.25	0.97	0.95	4.4
HadGEM2	0.56	0.31	9.7	0.18	−0.03	23.5	0.63	0.31	8.6	0.71	0.73	0.35	0.97	0.95	4.5
IPSL-CM5	0.68	0.38	9.7	0.33	0.19	18.1	0.67	0.39	8.1	0.77	0.64	0.34	0.97	0.95	4.7
GFDL-ESM2G	0.50	0.26	16.0	0.51	0.04	23.6	0.61	0.22	10.1	0.79	0.64	0.32	0.97	0.95	4.4
CanESM2	0.62	0.25	12.3	0.50	0.03	22.6	0.62	0.25	10.2	0.75	0.56	0.34	0.97	0.94	4.6
INM-CM4	0.65	0.29	11.3	0.34	−0.10	23.4	0.66	0.36	10.3	0.81	0.74	0.28	0.96	0.95	4.9
GISS-E2	0.34	0.06	20.7	0.46	−0.07	25.0	0.29	0.01	19.4	0.72	0.48	0.36	0.97	0.93	4.8
MIROC-ESM	0.48	0.39	19.8	0.52	0.06	27.7	0.54	0.37	14.3	0.77	0.55	0.30	0.95	0.93	5.1
MPI-ESM-LR	0.41	0.06	20.7	0.44	−0.03	23.2	0.34	0.10	22.0	0.70	0.67	0.39	0.97	0.94	4.5

stocks and NPP estimates from each model (Fig. 2). Using MODIS NPP, we calculated a turnover time of 24 yr for soil carbon in the HWSD. This estimate is consistent with the range of 18 to 32 yr reported in other studies (Amundson, 2001; Raich and Schlesinger, 1992). However, CanESM2 and INM-CM4 were the only two ESMs with turnover times that also fell within this range (Fig. 2). At high northern latitudes, 5 of the 11 models had turnover times that were considerably lower than the observations, whereas only 2 of the models had turnover times exceeding observational estimates (Fig. S6 see Supplement). Turnover times for high northern latitudes were 101.2 yr for the NCSCD and 60.8 yr for the HWSD.

3.2 Spatial distribution of soil carbon

The spatial distribution of soil carbon stocks varied widely among the ESMs (Fig. 3). CCSM4 and NorESM1 had the lowest overall soil carbon densities, but showed relatively high densities in northern South America, central Africa, eastern Asia, and eastern North America. HadGEM2, BCC-CSM1.1, and INM-CM4 showed a broader range of soil carbon densities with high densities in North America, western South America, central Africa, Southeast Asia, and north-central Eurasia. HadGEM2 also showed elevated soil carbon in southeastern South America. CanESM2 predicted high soil carbon in northeastern North America, northern Europe, northeastern Asia, central Africa, and eastern South America. GFDL-ESM2G and MIROC-ESM showed uniformly high carbon densities across all high northern latitudes and around the Tibetan Plateau. GISS-E2 predicted a region of high soil carbon across the northern latitudes of North America and Europe, as well as another area of high soil carbon from northeastern to southwestern Asia. MPI-ESM-LR showed an inverse pattern compared with the other ESMs; soil carbon peaked in the mid-latitudes across Asia, western North America, eastern Africa, southern South America, and southern coastal Australia.

There was generally poor agreement between the ESMs and the HWSD soil carbon distribution at the 1° scale (Table 3). Compared to the HWSD, ESMs had Pearson correlation coefficients between 0.06 and 0.39, RMSE between 9.4 and 20.7 kg C m^{-2}, and Taylor scores ranging from 0.21 to 0.68. Omitting the high latitude portion of the HWSD that overlapped with the NCSCD modestly improved these performance metrics for most but not all ESMs (Table 3). Model agreement with NCSCD soil carbon was poor with Pearson correlation coefficients between −0.10 and 0.19, RMSE between 18.0 and 27.7 kg C m^{-2}, and Taylor scores between 0.05 and 0.52. Agreement between the HWSD and NCSCD also was also low in the areas where the two data sets overlapped (Pearson correlation coefficient of 0.33, RMSE of 20.0 kg C m^{-2}, and Taylor score of 0.60), although better than the agreement between any individual ESM and the NCSCD.

ESM agreement with the HWSD generally improved at the biome level (Fig. 4). BCC-CSM1.1 and CanESM2 stood out as being highly correlated with the HWSD ($R^2 > 0.90$, $p < 0.01$), though CanESM2 overestimated soil carbon in boreal forests and grasslands and savanna. Biome simulations from HadGEM2, IPSL-CM5, INM-CM4, and MIROC-ESM also had relatively high levels of agreement with the HWSD ($0.90 > R^2 > 0.75$, $p < 0.01$), but most regression slopes and intercepts diverged from 1.0 and zero, respectively (Fig. 4). HadGEM2 overestimated soil carbon in grasslands and savanna. IPSL-CM5 generally overestimated tundra but underestimated desert and shrublands. INM-CM4 overestimated grasslands and savanna, boreal forests, croplands, and urban. MIROC-ESM overestimated all biomes except wetlands. Both CCSM4 and NorESM1 were moderately

Table 4. Coefficients of determination (R^2) and global-scale parameters ($1/k$ and Q_{10}) from reduced complexity models of ESM soil carbon distributions. Eq. (1): dependence on NPP; Eq. (2): dependence on NPP and soil temperature; Eq. (3): dependence on NPP, soil temperature, and soil moisture. Parameters are shown from Eq. (2) with $1/k$ analogous to turnover time. HWSD-MODIS-CRU represents reduced complexity models based on observational data. The reduced complexity model for CanESM2 was the only one improved by including soil moisture and had a turnover time ($1/k$) of 8.20 yr, Q_{10} of 1.48, and moisture exponent of 0.46 based on Eq. (3). All R^2 values were statistically significant ($R^2 > 0.05$, $p < 0.01$) unless otherwise indicated (NS).

	R^2			$C = \text{NPP}/(k\ Q_{10}^{(T-15)/10})$ Eq. (2)	
Model name	C= NPP/k Eq. (1)	C= NPP/k(T) Eq. (2)	C= NPP/k(T, W) Eq. (3)	$1/k$ (yr)	Q_{10}
HWSD-MODIS-CRU	NS	0.10	–	16.0	1.75
CCSM4	0.65	0.88	0.88	11.5	1.55
NorESM1	0.61	0.88	0.88	14.9	1.65
BCC-CSM1.1	NA	0.89	0.89	16.8	2.05
HadGEM2	0.27	0.79	0.85	13.5	1.52
IPSL-CM5	NS	0.93	0.93	13.2	1.61
GFDL-ESM2G	NS	0.85	0.89	10.9	2.61
CanESM2	NS	0.56	0.73	22.5	1.74
INM-CM4	NS	0.72	0.72	20.7	2.19
GISS-E2	NS	NS	NS	–	–
MIROC-ESM	NS	0.62	0.62	37.1	1.98
MPI-ESM-LR	NS	0.32	0.32	29.8	1.45

correlated with the HWSD ($0.70 > R^2 > 0.65$, $p < 0.01$), but consistently underestimated biome totals particularly in tundra, boreal forest, and desert and shrubland. Biome totals from MPI-ESM-LR were also moderately correlated with the HWSD ($R^2 = 0.62$, $p < 0.01$), but this model overestimated most biome totals, particularly grasslands and savanna. GFDL-ESM2G and GISS-E2 were weak to moderately correlated with the HWSD on the biome level ($R^2 = 0.38$ and $R^2 = 0.51$, respectively). GFDL-ESM2G overestimated biome totals from tundra and boreal forests while underestimating the other biomes. GISS-E2 overestimated biome totals in desert and shrublands, grasslands and savanna, tundra, and boreal forests while underestimating tropical rainforests.

3.3 Drivers of soil carbon distributions and global stocks

The spatial variability in 8 of the 11 ESMs was well explained by the reduced complexity model driven by NPP and soil temperature (Eq. 2) with R^2 values between 0.72 and 0.93 (Table 4). Consistent with our global-scale calculations (Fig. 2), the turnover times ($1/k$) for global soil carbon inferred from Eq. (2) (at a baseline temperature of 15 °C) varied from 11 to 37 yr, and Q_{10} values ranged from 1.5 to 2.6 (Table 4). The reduced complexity model for CanESM2 was the only one significantly improved by the addition of soil moisture (Eq. 3), with the R^2 value increasing from 0.56 to 0.73. Soil carbon outputs from GISS-E2 ($R^2 < 0.01$) and MPI-ESM-LR ($R^2 = 0.32$) were not well explained by any of the reduced complexity models.

Given the strong relationships between soil carbon, temperature, and NPP illustrated by our reduced complexity models, model skill in simulating driving variables could strongly influence simulated soil carbon stocks (Table 3). ESMs varied in their ability to capture the observed 1° spatial distribution of NPP (Pearson correlation coefficients from 0.48 to 0.75, biome regression R^2 values from 0.86 to 0.99; Table 3, Fig. S7 in the Supplement). In contrast, models performed better at simulating surface air temperature observations (correlations from 0.93 to 0.96, biome regression R^2 values from 0.93 to 0.97; Fig. S8 in the Supplement). Although air temperature is not directly comparable to soil temperature, particularly in areas with thick organic soils, the biome level correlation between soil and air temperature was high across all ESMs (R^2 values higher than 0.97; Fig. S9). INM-CM4, GISS-E2, BCC-CSM1.1, CCSM4, and NorESM1 all showed warmer soil temperatures compared to air temperatures in northern biomes (Fig. S9). In contrast to the strong relationships we found between soil carbon, NPP, and temperature in the ESMs, the 1° spatial distribution of soil carbon from the HWSD was not well explained by MODIS NPP and CRU surface air temperature data using the same reduced complexity model (R^2 value of 0.10 for Eq. (2); Table 4).

The reduced complexity models that explained within-model spatial variation of soil carbon also captured most of the variation in global soil carbon totals across ESMs

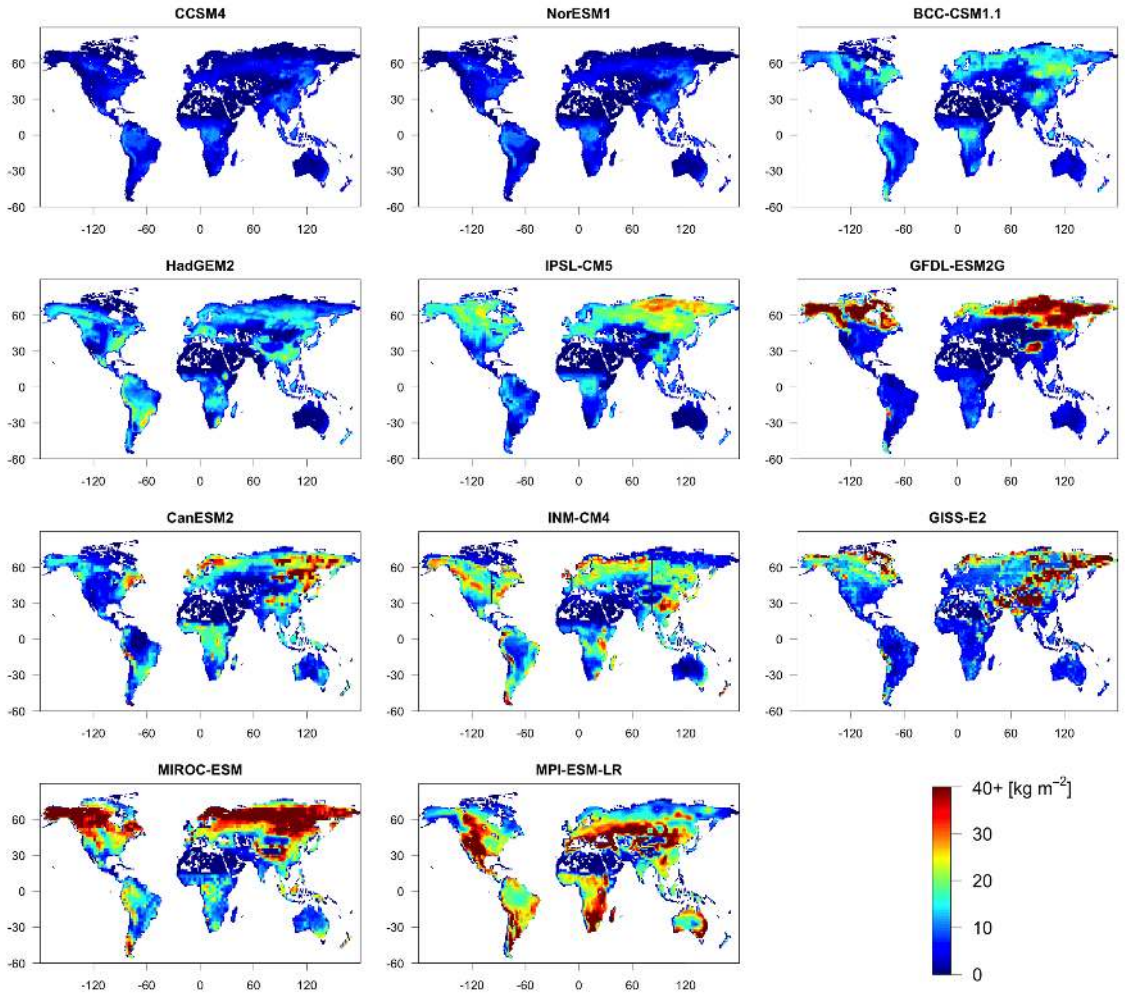

Fig. 3. Soil carbon densities [kg m^{-2}] from Earth system models. These soil carbon densities represent 1995–2005 means from the *historical* simulations of the Climate Model Intercomparison Project 5.

(Fig. 5). Reduced complexity models using ESM-specific values for k and Q_{10} and ESM-derived driving variables were able to explain 98 % the variation in total soil carbon across ESMs (Fig. 5a). Most of the variation in soil carbon across ESMs could be attributed to differences in parameterization as represented by Eq. (5) ($R^2 = 0.64$, Fig. 5b). There was no significant cross-ESM variation due to differences in driving variables alone as represented by Eq. (6) (Fig. 5c). However, driving variables must interact with parameters, since the variances explained by drivers alone (13 %, not significant) and parameters alone (64 %) did not sum to 98 % (Fig. 5a–c). NPP was likely the main driving variable in this interaction, since using the multi-model mean temperature at each grid cell (Eq. 7) still allowed us to explain 93 % of the variation in global soil carbon across ESMs (Fig. 5d).

4 Discussion

Accurate models of the soil carbon cycle are essential for predicting carbon–climate feedbacks in the future because soil carbon stocks are sensitive to climate change and large relative to the atmospheric CO_2 reservoir. As far as we know, our analysis is the first to benchmark soil carbon outputs from ESMs against empirical data at the global scale, and the first to explore the possible factors contributing to differences among models. We found that although some models simulated reasonable global soil carbon totals, fewer were able to match biome totals, and none were able to reproduce grid-scale distributions of soil carbon. There are a number of factors that may have contributed to the divergence between ESM simulations and observational data. These factors include (1) uncertainties in the data, (2) incorrect representation of environmental drivers in the models (e.g., NPP, temperature, and soil moisture), and (3) incorrect model structure

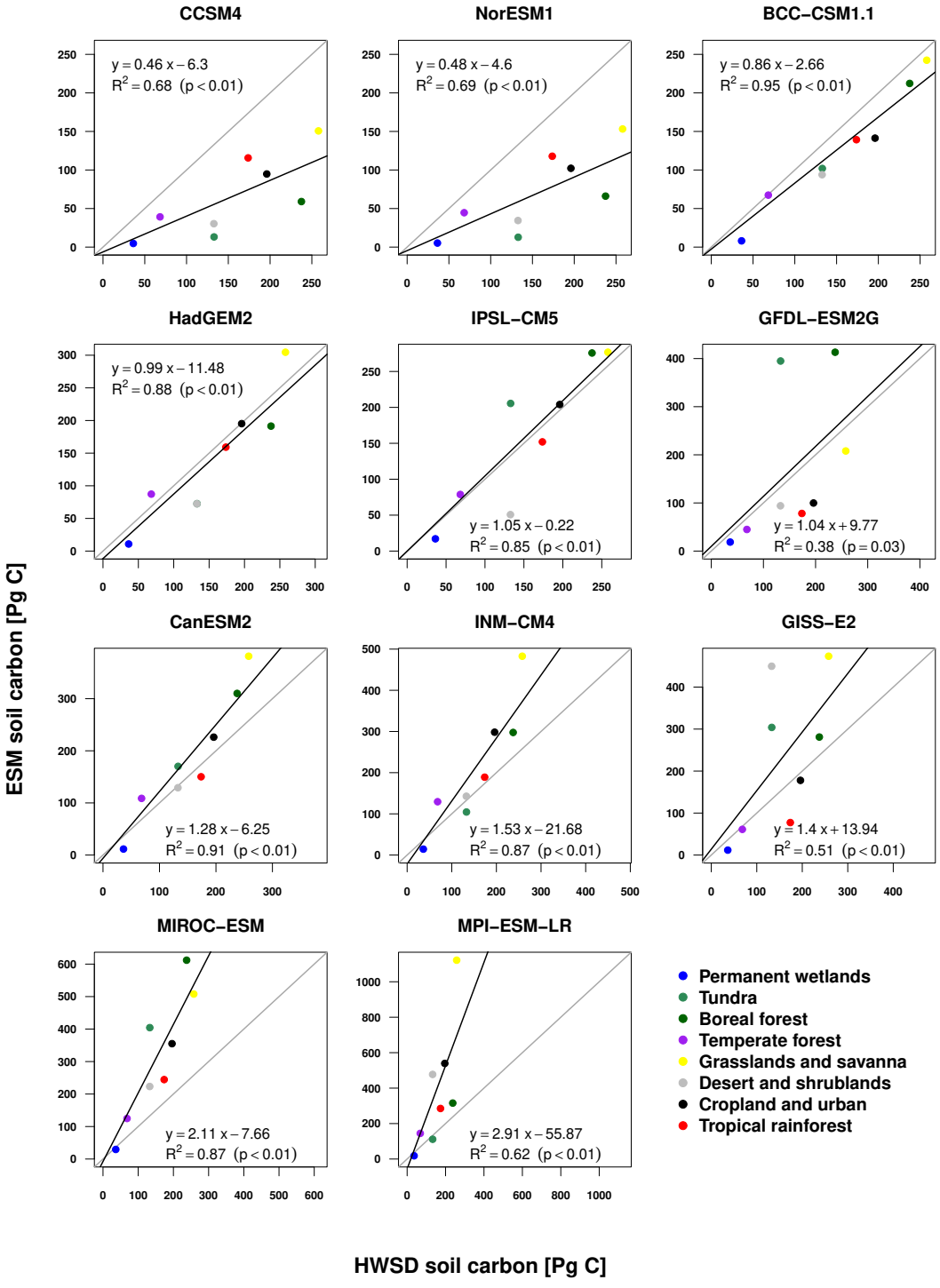

Fig. 4. Linear regression of ESM versus HWSD soil carbon totals [Pg C] for the 8 major biomes. The gray line indicates a 1 : 1 relationship and the black line is the linear regression.

or parameterization of the decomposition response to driving variables. Better performance at the global and biome scales may be due to aggregation of environmental variation that was not captured by the models at finer spatial scales. For instance, topographic controls on soil texture, moisture, and anoxia regulate soil carbon accumulation in peatlands and other organic soil types (e.g., Fan et al., 2008) but are poorly represented in many ESMs. Addressing these issues will be essential for increasing confidence in ESM simulations of terrestrial carbon in the future.

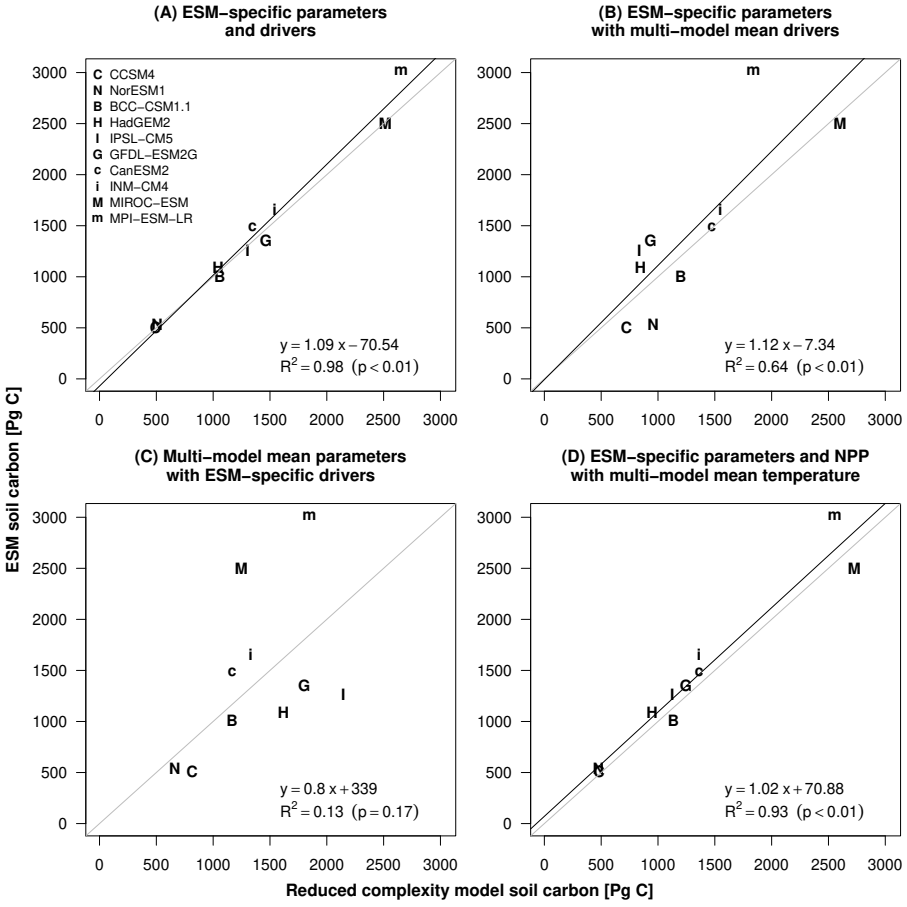

Fig. 5. Relationship between global soil carbon totals from ESMs and global soil carbon totals predicted by reduced complexity models (Eqs. 4–7). The reduced complexity model in **(A)** used ESM-specific parameters and drivers (Eq. 4); **(B)** used ESM-specific parameters with multi-model mean NPP and soil temperature (Eq. 5); **(C)** used multi-model mean parameters and ESM-specific NPP and soil temperature (Eq. 6); **(D)** used ESM specific parameters and NPP but multi-model mean soil temperature (Eq. 7). The gray line indicates a 1 : 1 fit and the black line is the linear regression.

4.1 Data uncertainties

Our ability to evaluate model performance relies on high quality empirical data with associated estimates of uncertainty. Whether model simulations diverge from the data is difficult to assess without a formal analysis of uncertainty in the data. Despite their comprehensiveness, the HWSD and NCSCD lack quantitative uncertainty estimates, thereby constraining our ability to use these data sets for benchmarking ESMs. Our preliminary analyses based on a qualitative assessment indicated that the uncertainty in empirical estimates of soil carbon stocks could exceed 770 Pg C at the global scale, an amount similar to the entire pool of atmospheric carbon.

At high northern latitudes, there was substantial disagreement between the two data sets. NCSCD estimates of CI_{95} were between 380 and 620 Pg C, whereas the corresponding HWSD estimate was only 290 Pg C. However, the HWSD did not include regional uncertainty information, meaning that the two estimates may agree once a formal uncertainty analysis has been performed. Such an analysis requires quantification of uncertainty in both measurement and scaling processes used to construct the spatial distribution of soil carbon. Uncertainty in the measurement of soil properties such as bulk density and carbon concentration must be integrated with errors involved in extrapolating data from individual soil profiles to the regional scale. Detailed analysis of the accuracy of soil maps will likely be essential for quantifying the uncertainty in this extrapolation process.

4.2 Driving variables

Accurate observational data are important but will not resolve the differences in simulated soil carbon that we observed among the ESMs. These differences must be due to differential model skill in simulating soil carbon drivers or in representing the response of soil carbon to drivers through model parameterization and structure. Our reduced

complexity models showed that differences in NPP contributed significantly to differences in soil carbon across ESMs (Fig. 5). NPP was also a significant driver of soil carbon within ESMs (Table 3), suggesting it should be a focal variable for improving soil carbon estimates. Given the large range of global NPP across ESMs (Fig. 2) and the low Taylor scores observed for some models in comparison to MODIS NPP (Table 3), it may be possible to improve soil carbon simulations by revising photosynthesis and autotrophic respiration algorithms in some of the ESMs so that NPP is more consistent with contemporary observations.

In contrast to NPP, differences in soil temperature did not contribute significantly to differences in soil carbon stocks across ESMs (Fig. 5). However, our reduced complexity models indicated that soil temperature was important for explaining soil carbon variation within models (Table 4). Therefore, simulation of soil temperature could also be a focal area for model improvement, especially since other studies suggest that soil temperature is not consistently well represented in ESMs. For example, the physical coupling between surface air temperature and soil temperature at high latitudes differs considerably across ESMs, which influences the spatial distribution of permafrost (Koven et al., 2012; Slater and Lawrence, 2013). Even if ESMs can simulate average air temperatures consistent with observations (Table 3, Fig. S8), fine-scale differences in the ability to represent soil temperatures (Fig. S9) and permafrost could have consequences for soil carbon distributions.

Soil moisture did not play an important role as a driving variable for soil carbon in our reduced complexity models, indicating that for most models this variable did not strongly control spatial patterns in soil carbon stocks (Table 4) or differences among models (Fig. 5). This result was unexpected because soil moisture affects decomposition rates in all ESMs (Table 1). Furthermore, other studies have shown that soil carbon stocks depend on the response of heterotrophic respiration to soil moisture in global models, although NPP and soil temperature were also important drivers of soil carbon (Falloon et al., 2011). It is possible that soil moisture influences soil carbon stocks in ESMs, but our reduced complexity model was unable to statistically distinguish the soil moisture effect from the NPP effect because these two drivers often covary.

Alternatively, the exponential form of the moisture function in our reduced complexity model might have been inappropriate if decomposition rates decline at high soil moisture. Based on empirical data, a substantial fraction of global soil carbon likely resides in areas where poor soil drainage impedes organic matter oxidation (Gorham, 1991). It is likely that the interaction of topographic controls and soil texture with soil moisture is not well represented in the current generation of ESMs. New approaches may be needed to determine which grid cells are poorly drained, and the rate at which organic soils form in these area (Ise et al., 2008). We also recommend that future CMIP archives require soil moisture information for different soil layers to facilitate benchmarking studies on the response of carbon to moisture in the soil profile.

Although our reduced complexity models indicated that soil carbon simulated by ESMs was driven primarily by NPP and temperature, this relationship was much weaker with observational data. According to the same reduced complexity model used for the ESMs, MODIS NPP and CRU surface air temperatures were only able to explain 10 % of the spatial variation in HWSD soil carbon stocks. Thus the variables that drive soil carbon stocks in ESMs may differ from those that determine observed carbon stocks. For example, soil temperatures in organic-rich soils may be weakly coupled to air temperatures (Koven et al., 2011; Slater and Lawrence, 2013), in contrast to the strong coupling predicted by most ESMs (Fig. S9). Alternatively, the mathematical form of our reduced complexity model may have been inappropriate for the observational data, even though it was appropriate for most ESM outputs. Regardless, our results imply that ESMs may need to incorporate a broader range of environmental drivers and processes to improve model–data agreement. Even if simulations of NPP and soil temperature can be improved in ESMs, these drivers have a limited ability to explain spatial patterns in global soil carbon with current model structures.

4.3 Parameterization and model structure

We found that parameterization was a major source of variation in ESM soil carbon simulations (Fig. 5). In some of the ESMs such as CCSM4, soil carbon turnover may be too fast, whereas in other models such as MIROC-ESM, turnover may be too slow (Fig. 2). To address these issues, ESMs could use terrestrial radiocarbon observations (both the total inventory and vertical distribution of ^{14}C in different biomes) to help constrain rates of soil carbon turnover (Torn et al., 1997; Trumbore, 2009). Another avenue of improvement for ESM parameterization could focus on processes that operate on fine spatial scales. Differences in soil texture and topography may lead to non-linear effects on soil carbon storage that are not well described by the average characteristics of a grid cell. For instance, relatively small-scale topographic variations are associated with peatland formation, and it is unclear how to scale these effects globally (Gorham, 1991; Koven et al., 2011). A multi-scale approach is required to determine which processes are important at the global scale and how to represent them.

Improving empirical data sets, model driving variables, and model parameterization could substantially increase model–data agreement for present-day soil carbon stocks. However, matching current soil survey data is a necessary but not sufficient condition for validating the accuracy of Earth system models. In order to have confidence in future simulations, the models must correctly represent the mechanisms and drivers of soil carbon change, such that they are right for

the right reasons. For example, models with incorrect mechanisms or drivers could be tuned to correctly simulate current soil carbon stocks, but might incorrectly simulate soil carbon stock changes in the future.

We initially hypothesized that models with more pools would have greater flexibility and capture more of the spatial variation in soil carbon. However, the structural features that we examined did not clearly relate to differences in ESM agreement with empirical data (Tables 1, 3). We saw no pattern in ESM-data agreement with respect to number of soil carbon pools, temperature and moisture sensitivity functions, or presence of a nitrogen component. Furthermore, our reduced complexity model (Eq. 3) explained most of the spatial variation within 9 of 11 models ($0.62 < R^2 < 0.93$). This result confirmed that, despite different simulated stocks of soil carbon, most of the models share a similar underlying structure. Such similarity means that the models likely make similar assumptions about the mechanisms regulating soil carbon cycling. If these underlying assumptions are incorrect or incomplete, the resulting errors will be present in all of the models.

CanESM2, MPI-ESM-LR, and GISS-E2 are three exceptions that were not well explained by our reduced complexity model (Eq. 2) driven by NPP and soil temperature (Table 4), and thus may be examples of models with structural differences. CanESM2 was the only model in which soil water content contributed to the explanatory power of the reduced complexity model (R^2 value improved from 0.57 to 0.74). This dependency on soil water content may be partly explained by the biome-specific turnover time in CanESM2. Since biomes are partially determined by precipitation and soil moisture status, biome-specific turnover times might have resulted in a tighter relationship between soil carbon and moisture in our reduced complexity model. Outputs from MPI-ESM-LR were only moderately explained by our reduced complexity models (R^2 value 0.32). We do not have a good explanation for the poor fit since there was no significant deviation in documented model structure, and driving variables were roughly in line with other model simulations. GISS-E2 outputs were not explained at all by the reduced complexity model (R^2 value less than 0.01). Unlike other models, GISS-E2 showed a unique disconnect between NPP and soil carbon which could be due to differences in the way plant biomass is allocated to litter in the model. However, we cannot offer a definitive explanation for the poor fit.

All ESMs may be missing key processes governing long-term carbon storage that affect model–data agreement. Decomposition models currently used in all ESMs are built on the assumption that carbon substrates have intrinsic chemical decomposition rates (Parton et al., 1993). However, there is an emerging consensus that key abiotic and biotic factors have a stronger governing role in decomposition than the carbon compounds themselves (Schmidt et al., 2011). These key governing components may include aggregate interactions (Six et al., 2000), microbial dynamics (Todd-Brown et al., 2012), cryoturbation (Koven et al., 2011), syngenetic soil formation (Fan et al., 2008; Shur et al., 2004), extracellular enzyme dynamics (German et al., 2011), and rare substrate formation (Allison, 2006). Representing these processes in the structure of soil carbon models remains a major challenge. However, smaller-scale decomposition models have begun to explore several of these mechanisms (Manzoni and Porporato, 2009).

Recent advances in the theory of microbial decomposition could provide a foundation for major changes in the structure of soil carbon models used in ESMs. Schimel and Weintraub (2003) proposed a model in which decomposition was mediated by soil enzymes and microbial biomass. Later models expanded this framework to include microbial functional groups that preferentially decompose specific substrate types (Moorhead and Sinsabaugh, 2006). In contrast to current substrate pool models used in ESMs, biomass-mediated decomposition models would likely include non-linear processes such as Monod uptake or Michaelis–Menten enzyme kinetics. These non-linear effects could produce very different behaviors at daily, annual, and centennial timescales. Compared to substrate pool models, models driven by microbial biomass predict smaller losses of soil carbon under warming due to declines in microbial growth efficiency with higher temperature (Allison et al., 2010).

5 Conclusions

Overall, we found that some ESMs simulated soil carbon stocks consistent with empirical estimates at the global and biome scales. However, all of the models had difficulty representing soil carbon at the 1° scale. Despite similar overall structures, the models do not agree well among themselves or with empirical data on the global distribution of soil carbon. Differences across ESMs are primarily due to differences in the representation of NPP and the parameterization of soil decomposition sub-models, not due to differences in model structure. However, all model structures may have serious shortcomings since NPP and temperature strongly influenced soil carbon stocks in ESMs but not in observational data. Fully reconciling this disagreement will require a range of approaches, including better prediction of soil carbon drivers, more accurate model parameterization, and more comprehensive representation of critical biological and geochemical mechanisms in soil carbon sub-models. There is also a need for better quantification of the uncertainty in empirical estimates of soil carbon stocks that are used to benchmark ESMs. If this uncertainty is too high for rigorous model comparison, additional measurements of soil carbon stocks may be required in some regions of the world. Addressing these issues will improve our ability to predict the response of the carbon cycle to climate change and inform policymakers about the potential impacts of carbon emissions.

Supplementary material related to this article is available online at: http://www.biogeosciences.net/10/1717/2013/bg-10-1717-2013-supplement.pdf.

Acknowledgements. We thank Shishi Lui and Yaxing Wei for assistance with the HWSD, as well as Yufang Jin for assistance with the NCSCD data set. This research was funded by grants from the NSF Advancing Theory in Biology program, the Decadal and Regional Climate Prediction using Earth System Models (EaSM; AGU-1048890) program, and the Office of Science (BER), US Department of Energy.

We acknowledge the World Climate Research Programme's Working Group on Coupled Modeling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table S1) for producing and making available their model output. For CMIP, the US Department of Energy's Program for Climate Model Diagnosis and Intercomparison provided coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

The authors have no potential conflicts of interest to declare.

Edited by: U. Seibt

References

Allison, S. D.: Brown ground: a soil carbon analogue for the green world hypothesis?, The American Naturalist, 167, 619–627, doi:10.1086/503443, 2006.

Allison, S. D., Wallenstein, M. D., and Bradford, M. A.: Soil-carbon response to warming dependent on microbial physiology, Nat. Geosci., 3, 336–340, doi:10.1038/ngeo846, 2010.

Amundson, R.: The carbon budget in soils, Ann. Rev. Earth Planet. Sci., 29, 535–562, doi:10.1146/annurev.earth.29.1.535, 2001.

Arora, V. K.: Simulating energy and carbon fluxes over winter wheat using coupled land surface and terrestrial ecosystem models, Agr. Forest Meteorol., 118, 21–47, doi:10.1016/S0168-1923(03)00073-X, 2003.

Arora, V. K. and Boer, G. J.: A parameterization of leaf phenology for the terrestrial ecosystem component of climate models, Glob. Change Biol., 11, 39–59, doi:10.1111/j.1365-2486.2004.00890.x, 2005.

Arora, V. K., Scinocca, J. F., Boer, G. J., Christian, J. R., Denman, K. L., Flato, G. M., Kharin, V. V, Lee, W. G., and Merryfield, W. J.: Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases, Geophys. Res. Lett., 38, L05805, doi:10.1029/2010GL046270, 2011.

Batjes, N. H.: Total carbon and nitrogen in the soils of the world, Eur. J. Soil Sci., 47, 151–163, doi:10.1111/j.1365-2389.1996.tb01386.x, 1996.

Batjes, N. H.: A world dataset of derived soil properties by FAO-UNESCO soil unit for global modelling, Soil Use Manage., 13, 9–16, doi:10.1111/j.1475-2743.1997.tb00550.x, 1997.

Bolker, B. M., Pacala, S. W., and Parton Jr., W. J.: Linear analysis of soil decomposition: Insights from the Century model, Ecol. Appl., 8, 425–439, 1998.

Bonan, G. B.: Land-atmosphere CO_2 exchange simulated by a land surface process model coupled to an atmospheric general circulation model, J. Geophys. Res., 100, 2817–2831, doi:10.1029/94JD02961, 1995.

Bonan, G. B.: A land surface model (LSM version 1.0) for ecological, hydrological and atmospheric studies: technical discription and user's guide, NCAR technical note NCAR/TN-417 + STR, available at: http://nldr.library.ucar.edu/collections/technotes/asset-000-000-000-229.pdf (last accessed 1 January 2013), 1996.

Bunnell, F. L., Tait, D. E. N., Flanagan, P. W., and Van Clever, K.: Microbial respiration and substrate weight loss – I, Soil Biol. Biochem., 9, 33–40, doi:10.1016/0038-0717(77)90058-X, 1977.

Byrd, R. H., Lu, P., Nocedal, J., and Zhu, C.: A limited memory algorithm for bound constrained optimization, SIAM J. Sci. Comput., 16, 1190–1208, doi:10.1137/0916069, 1995.

Cao, M. and Woodward, F. I.: Net primary and ecosystem production and carbon stocks of terrestrial ecosystems and their responses to climate change, Glob. Change Biol., 4, 185–198, doi:10.1046/j.1365-2486.1998.00125.x, 1998.

Coleman, K. and Jenkinson, D. S.: ROTHC-26.3, A model for the turnover of carbon in soils, Model description and windows user guide, available at: http://www.rothamsted.ac.uk/ssgs/RothC/mod26_3_win.pdf (last accessed 1 January 2013), 1999.

Collins, W. J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Halloran, P., Hinton, T., Hughes, J., Jones, C. D., Joshi, M., Liddicoat, S., Martin, G., O'Connor, F., Rae, J., Senior, C., Sitch, S., Totterdell, I., Wiltshire, A., and Woodward, S.: Development and evaluation of an Earth-System model – HadGEM2, Geoscientific Model Development, 4, 1051–1075, doi:10.5194/gmd-4-1051-2011, 2011.

Cox, P. M.: Description of the "TRIFFID" dynamic global vegetation model, Hadley Centre technical note 24, 2001.

Cramer, W., Bondeau, A., Woodward, I. F., Prentice, I. C., Betts, R. A., Brovkin, V., Cox, P. M., Fisher, V., Foley, J. A., Friend, A. D., Kucharik, C., Lomas, M. R., Ramankutty, N., Sitch, S., Smith, B., White, A., and Young-Molling, C.: Global response of terrestrial ecosystem structure and function to CO_2 and climate change: results from six dynamic global vegetation models, Glob. Change Biol., 7, 357–373, doi:10.1046/j.1365-2486.2001.00383.x, 2001.

Davidson, E. A. and Janssens, I. A.: Temperature sensitivity of soil carbon decomposition and feedbacks to climate change., Nature, 440, 165–173, doi:10.1038/nature04514, 2006.

Del Grosso, S. J., Parton, W. J., Mosier, A. R., Holland, E. A., Pendall, E., Schimel, D. S., and Ojima, D. S.: Modeling soil CO_2 emissions from ecosystems, Biogeochemistry, 73, 71–91, doi:10.1007/s10533-004-0898-z, 2005.

Doney, S. C., Lindsay, K., Fung, I., and John, J.: Natural variability in a stable, 1000-yr global coupled climate–carbon cycle simulation, J. Clim., 19, 3033–3054, doi:10.1175/JCLI3783.1, 2006.

ESDB: European Soil Database (vs 2.0), available at: http://eusoils.jrc.ec.europa.eu/ESDB_Archive/ESDBv2/index.htm (last accessed 1 January 2013), 2004.

Eswaran, H., Van Den Berg, E., and Reich, P.: Organic carbon in soils of the world, Soil Sci. Soc. Am. J., 57, 192–194, doi:10.2136/sssaj1993.03615995005700010034x, 1993.

Falloon, P., Jones, C. D., Ades, M., and Paul, K.: Direct soil moisture controls of future global soil carbon changes: An important source of uncertainty, Glob. Biogeochem. Cy., 25, GB3010, doi:10.1029/2010GB003938, 2011.

Fan, Z., Neff, J. C., Harden, J. W., and Wickland, K. P.: Boreal soil carbon dynamics under a changing climate: A model inversion approach, J. Geophys. Res., 113, G04016, doi:10.1029/2008JG000723, 2008.

FAO/IIASA/ISRIC/ISSCAS/JRC: Harmonized World Soil Database (version 1.10), FAO, Rome, Italy and IIASA, Laxenburg, Austria, 2012.

FAO/UNESCO: The FAO-UNESCO Soil Map of the World, UNESCO, Pairs, Legend and 9 volumes, available at: http://www.fao.org/geonetwork/srv/en/metadata.show?id=14116 (last accessed 1 January 2013), 1981.

Foley, J. A.: An equilibrium model of the terrestrial carbon budget, Tellus B, 47, 310–319, doi:10.1034/j.1600-0889.47.issue3.3.x, 1995.

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., and Huang, X.: MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets, Remote Sens. Environ., 114, 168–182, doi:10.1016/j.rse.2009.08.016, 2010.

Friedlingstein, P., Cox, P., Betts, R., Bopp, L., von Bloh, W., Brovkin, V., Cadule, P., Doney, S., Eby, M., Fung, I., Bala, G., John, J., Jones, C., Joos, F., Kato, T., Kawamiya, M., Knorr, W., Lindsay, K., Matthews, H D., Raddatz, T., Rayner, P., Reick, C., Roeckner, E., Schnitzler, K.-G., Schnur, R., Strassmann, K., Weaver, A. J., Yoshikawa, C., and Zeng, N.: Climate–carbon cycle feedback analysis: Results from the C4MIP model Intercomparison, J. Clim., 19, 3337–3353, doi:10.1175/JCLI3800.1, 2006.

Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., Lawrence, D. M., Neale, R. B., Rasch, P. J., Vertenstein, M., Worley, P. H., Yang, Z.-L., and Zhang, M.: The Community Climate System Model version 4, J. Clim., 24, 4973–4991, doi:10.1175/2011JCLI4083.1, 2011.

German, D. P., Chacon, S. S., and Allison, S. D.: Substrate concentration and enzyme allocation can affect rates of microbial decomposition, Ecology, 92, 1471–1480, doi:10.1890/10-2028.1, 2011.

Giorgi, F.: Climate change hot-spots, Geophys. Res. Lett., 33, L08707, doi:10.1029/2006GL025734, 2006.

Gorham, E.: Northern peatlands: role in the carbon cycle and probable responses to climatic warming, Ecol. Appl., 1, 182–195, doi:10.2307/1941811, 1991.

Goulden, M. L., Mcmillan, A. M. S., Winston, G. C., Rocha, A. V., Manies, K. L., Harden, J. W., and Bond-Lamberty, B. P.: Patterns of NPP, GPP, respiration, and NEP during boreal forest succession, Glob. Change Biol., 17, 855–871, doi:10.1111/j.1365-2486.2010.02274.x, 2011.

Harden, J. W., Trumbore, S. E., Stocks, B. J., Hirsch, A., Gower, S. T., O'neill, K. P., and Kasischke, E. S.: The role of fire in the boreal carbon budget, Glob. Change Biol., 6, 174–184, doi:10.1046/j.1365-2486.2000.06019.x, 2000.

Houghton, R. A.: Balancing the global carbon budget, Ann. Rev. Earth Planet. Sci., 35, 313–347, doi:10.1146/annurev.earth.35.031306.140057, 2007.

Huang, M., Ji, J., Li, K., Liu, Y., Yang, F., and Tao, B.: The ecosystem carbon accumulation after conversion of grasslands to pine plantations in subtropical red soil of South China, Tellus B, 59, 439–448, doi:10.1111/j.1600-0889.2007.00280.x, 2007.

Huntingford, C., Fisher, R. A., Mercado, L., Booth, B. B. B., Sitch, S., Harris, P. P., Cox, P. M., Jones, C. D., Betts, R. A., Malhi, Y., Harris, G. R., Collins, M., and Moorcroft, P.: Towards quantifying uncertainty in predictions of Amazon "dieback", Philosophical transactions of the Royal Society of London. Series B, Biol. Sci., 363, 1857–1864, doi:10.1098/rstb.2007.0028, 2008.

Ise, T., Dunn, A. L., Wofsy, S. C., and Moorcroft, P. R.: High sensitivity of peat decomposition to climate change through water-table feedback, Nature Geosci., 1, 763–766, doi:10.1038/ngeo331, 2008.

Ito, A.: A historical meta-analysis of global terrestrial net primary productivity: are estimates converging?, Glob. Change Biol., 17, 3161–3175, doi:10.1111/j.1365-2486.2011.02450.x, 2011.

Ji, J., Huang, M., and Li, K.: Prediction of carbon exchanges between China terrestrial ecosystem and atmosphere in 21st century, Sci. China Ser. D, 51, 885–898, doi:10.1007/s11430-008-0039-y, 2008.

Jobbagy, E. G. and Jackson, R. B.: The vertical distribution of soil organic carbon and its relation to climate and vegetation, Ecol. Appl., 10, 423–436, 2000.

Jones, P. and Harris, I.: CRU Time Series (TS) high resolution gridded datasets, NCAS British Atmospheric Data Centre, available at: http://badc.nerc.ac.uk/view/badc.nerc.ac.uk__ATOM__dataent_1256223773328276 (last accessed 1 January 2013), 2008.

Jones, C. D., Hughes, J. K., Bellouin, N., Hardiman, S. C., Jones, G. S., Knight, J., Liddicoat, S., O'Connor, F. M., Andres, R. J., Bell, C., Boo, K.-O., Bozzo, A., Butchart, N., Cadule, P., Corbin, K. D., Doutriaux-Boucher, M., Friedlingstein, P., Gornall, J., Gray, L., Halloran, P. R., Hurtt, G., Ingram, W. J., Lamarque, J.-F., Law, R. M., Meinshausen, M., Osprey, S., Palin, E. J., Parsons Chini, L., Raddatz, T., Sanderson, M. G., Sellar, A. A., Schurer, A., Valdes, P., Wood, N., Woodward, S., Yoshioka, M., and Zerroukat, M.: The HadGEM2-ES implementation of CMIP5 centennial simulations, Geosci. Model Dev., 4, 543–570, doi:10.5194/gmd-4-543-2011, 2011.

Kaminski, T., Knorr, W., Rayner, P., and Heimman, M.: Assimilating atmospheric data into a terrestrial biosphere model: A case study of the seasonal cycle, Glob. Biogeochem. Cy., 16, 1066, doi:10.1029/2001GB001463, 2002.

Kelly, R. H., Parton, W. J., Crocker, G. J., Graced, P. R., Klír, J., Körschens, M., Poulton, P. R., and Richter, D. D.: Simulating trends in soil organic carbon in long-term experiments using the Century model, Geoderma, 81, 75–90, doi:10.1016/S0016-7061(97)00082-7, 1997.

Kimball, J. S., Thornton, P. E., White, M. A., and Running, S. W.: Simulating forest productivity and surface-atmosphere carbon exchange in the BOREAS study region, Tree Physiology, 17, 589–599, doi:10.1093/treephys/17.8-9.589, 1997.

Knorr, W.: Annual and interannual CO_2 exchanges of the terrestrial biosphere: process-based simulations and uncertainties, Glob. Ecol. Biogeogr., 9, 225–252, doi:10.1046/j.1365-2699.2000.00159.x, 2000.

Koven, C. D., Ringeval, B., Friedlingstein, P., Ciais, P., Cadule, P., Khvorostyanov, D., Krinner, G., and Tarnocai, C.: Permafrost carbon-climate feedbacks accelerate global warming., Proc. Natl. Ac. Sci. USA, 108, 14769–14774, doi:10.1073/pnas.1103910108, 2011.

Koven, C. D., Riley, W. J., and Stern, A.: Analysis of permafrost thermal dynamics and response to climate change in the CMIP5 Earth System Models, J. Clim., in press, doi:10.1175/JCLI-D-12-00228.1, 2012.

Krinner, G., Viovy, N., De Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, Glob. Biogeochem. Cy., 19, GB1015, doi:10.1029/2003GB002199, 2005.

Le Quéré, C.,Raupach, M. R., Canadell, J. G., Marland, G., Bopp, L., Ciais, P., Conway, T. J., Doney, S. C., Feely, R. A., Foster, P., Friedlingstein, P., Gurney, K., Houghton, R. A., House, J. I., Huntingford, C., Levy, P. E., Lomas, M. R., Majkut, J., Metzl, N., Ometto, J. P., Peters, G. P., Prentice, I. C., Randerson, J. T., Running, S. W., Sarmiento, J. L., Schuster U., Sitch, S., Takahashi, T., Viovy, N., van der Werf, G. R., and Woodward, F. I.: Trends in the sources and sinks of carbon dioxide, Nat. Geosci., 2, 831–836, doi:10.1038/ngeo689, 2009.

Lloyd, J. and Taylor, J. A.: On the temperature dependence of soil respiration, Funct. Ecol., 8, 315–323, 1994.

Lund, M., Lafleur, P. M., Roulet, N. T., Lindroth, A., Christensen, T. R., Aurela, M., Chojnicki, B. H., Flanagan, L. B., Humphreys, E. R., Laurila, T., Oechel, W. C., Olejnik, J., Rinne, J., Schubert, P., and Nilsson, M. B.: Variability in exchange of CO_2 across 12 northern peatland and tundra sites, Glob. Change Biol., 16, 2436–2448, doi:10.1111/j.1365-2486.2009.02104.x, 2009.

Luo, Y. Q., Randerson, J. T., Abramowitz, G., Bacour, C., Blyth, E., Carvalhais, N., Ciais, P., Dalmonech, D., Fisher, J. B., Fisher, R., Friedlingstein, P., Hibbard, K., Hoffman, F., Huntzinger, D., Jones, C. D., Koven, C., Lawrence, D., Li, D. J., Mahecha, M., Niu, S. L., Norby, R., Piao, S. L., Qi, X., Peylin, P., Prentice, I. C., Riley, W., Reichstein, M., Schwalm, C., Wang, Y. P., Xia, J. Y., Zaehle, S., and Zhou, X. H.: A framework for benchmarking land models, Biogeosciences, 9, 3857–3874, doi:10.5194/bg-9-3857-2012, 2012.

Martin, G. M., Bellouin, N., Collins, W. J., Culverwell, I. D., Halloran, P. R., Hardiman, S. C., Hinton, T. J., Jones, C. D., McDonald, R. E., McLaren, A. J., O'Connor, F. M., Roberts, M. J., Rodriguez, J. M., Woodward, S., Best, M. J., Brooks, M. E., Brown, A. R., Butchart, N., Dearden, C., Derbyshire, S. H., Dharssi, I., Doutriaux-Boucher, M., Edwards, J. M., Falloon, P. D., Gedney, N., Gray, L. J., Hewitt, H. T., Hobson, M., Huddleston, M. R., Hughes, J., Ineson, S., Ingram, W. J., James, P. M., Johns, T. C., Johnson, C. E., Jones, A., Jones, C. P., Joshi, M. M., Keen, A. B., Liddicoat, S., Lock, A. P., Maidens, A. V., Manners, J. C., Milton, S. F., Rae, J. G. L., Ridley, J. K., Sellar, A., Senior, C. A., Totterdell, I. J.,Verhoef, A., Vidale, P. L., and Wiltshire, A.: The HadGEM2 family of Met Office Unified Model climate configurations, Geosci. Model Develop., 4, 723–757, doi:10.5194/gmd-4-723-2011, 2011.

Mahecha, M. D., Reichstein, M., Carvalhais, N., Lasslop, G., Lange, H., Seneviratne, S. I., Vargas, R., Ammann, C., Arain, M. A., Cescatti, A., Janssens, I. A., Migliavacca, M., Montagnani, L., and Richardson, A. D.: Global convergence in the temperature sensitivity of respiration at ecosystem level, Science, 329, 838–840, doi:10.1126/science.1189587, 2010.

Manzoni, S. and Porporato, A.: Soil carbon and nitrogen mineralization: Theory and models across scales, Soil Biol. Biochem., 41, 1355–1379, doi:10.1016/j.soilbio.2009.02.031, 2009.

Moorcroft, P. R., Hurtt, G. C., and Pacala, S. W.: A method for scaling vegetation dynamics: The Ecosystem Demography model (ED), Ecol. Monogr., 71, 557–586, 2001.

Moorhead, D. L. and Sinsabaugh, R. L.: A theoretical model of litter decay and microbial interaction, Ecol. Monogr., 76, 151–174, 2006.

NASA Land Processes Distributed Active Archive Center (LP DAAC): Land Cover Type Yearly L3 Global 0.05Deg CMG (MCD12C1), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, available at: https://lpdaac.usgs.gov/products/modis_products_table/land_cover/yearly_l3_global_0_05deg_cmg/mcd12c1, 2008.

Neff, J. C. and Hooper, D. U.: Vegetation and climate controls on potential CO_2, DOC and DON production in northern latitude soils, Glob. Change Biol., 8, 872–884, doi:10.1046/j.1365-2486.2002.00517.x, 2002.

Oleson, K. W., Niu, G.-Y., Yang, Z.-L., Lawrence, D. M., Thornton, P. E., Lawrence, P. J., Stöckli, R., Dickinson, R. E., Bonan, G. B., Levis, S., Dai, A., and Qian, T.: Improvements to the Community Land Model and their impact on the hydrological cycle, J. Geophys. Res., 113, G01021, doi:10.1029/2007JG000563, 2008.

Olson, J. S.: Energy storage and the balance of producers and decomposers in ecological systems, Ecology, 44, 322–331, 1963.

Orchard, V. A. and Cook, F. J.: Relationship between soil respiration and soil moisture, Soil Biol. Biochem., 15, 447–453, doi:10.1016/0038-0717(83)90010-X, 1983.

Parton, W. J., Schimel, D. S., Cole, C. V., and Ojima, D. S.: Analysis of factors controlling soil organic matter levels in great plains grasslands, Soil Sci. Soc. Am. J., 51, 1173–1179, doi:10.2136/sssaj1987.03615995005100050015x, 1987.

Parton, W. J., Stewart, J. W. B., and Cole, C. V.: Dynamics of C, N, P and S in grassland soils: a model, Biogeochemistry, 5, 109–131, doi:10.1007/BF02180320, 1988.

Parton, W. J., Running, S. W., and Walker, B.: A toy terrestrial carbon flow model, University Corp. for Atmospheric Research, Modeling the Earth System, SEE N94–30616 08–45, 303–314, available at: http://hdl.handle.net/2060/19940026125 (last accessed 1 January 2013), 1992.

Parton, W. J., Scurlock, J. M. O., Ojima, D. S., Gilmanov, T. G., Scholes, R. J., Schimel, D. S., Kirchner, T., Menaut, J.-C., Seastedt, T., Moya, E. G., Kamnalrut, A., and Kinyamario, J. I.: Observations and modeling of biomass and soil organic matter dynamics for the grassland biome worldwide, Glob. Biogeochem. Cy., 7, 785–809, doi:10.1029/93GB02042, 1993.

Phillips, O. L., Aragão, L. E. O. C., Lewis, S. L., Fisher, J. B., Lloyd, J., López-González, G., Malhi, Y., Monteagudo, A., Peacock, J., Quesada, C. A., van der Heijden, G., Almeida, S., Amaral, I., Arroyo, L., Aymard, G., Baker, T. R., Bánki, O., Blanc, L., Bonal, D., Brando, P., Chave, J., de Oliveira, A. C. A., Cardozo, N. D., Czimczik, C. I., Feldpausch, T. R., Freitas, M. A., Gloor, E., Higuchi, N., Jiménez, E., Lloyd, G., Meir, P., Mendoza, C., Morel, A., Neill, D. A., Nepstad, D., Patiño, S., Peñuela, M. C., Prieto, A., Ramírez, F., Schwarz, M., Silva, J., Silveira, M., Thomas, A. S., Steege, H., Stropp, J., Vásquez, R., Zelazowski, P., Dávila, E. A., Andelman, S., Andrade, A., Chao, K.-J., Erwin, T., Di Fiore, A., Honorio C., E., Keeling, H., Killeen, T. J., Laurance, W. F., Cruz, A. P., Pitman, N. C. A., Vargas, P. N., Ramírez-Angulo, H., Rudas, A., Salamão, R., Silva, N., Terborgh, J., and Torres-Lezama, A.: Drought sen-

sitivity of the Amazon rainforest., Science, 323, 1344–1347, doi:10.1126/science.1164033, 2009.

Ping, C.-L., Michaelson, G. J., Jorgenson, M. T., Kimble, J. M., Epstein, H., Romanovsky, V. E., and Walker, D. A.: High stocks of soil organic carbon in the North American Arctic region, Nat. Geosci., 1, 615–619, doi:10.1038/ngeo284, 2008.

Post, W. M. and Kwon, K. C.: Soil carbon sequestration and land-use change: processes and potential, Glob. Change Biol., 6, 317–327, doi:10.1046/j.1365-2486.2000.00308.x, 2000.

Potter, C. S., Randerson, J. T., Field, C. B., Matson, P. A., Vitousek, P. M., Mooney, H. A., and Klooster, S. A.: Terrestrial ecosystem production: A process model based on global satellite and surface data, Glob. Biogeochem. Cy., 7, 811–841, doi:10.1029/93GB02725, 1993.

R Development Core Team: R: A Language and Environment for Statistical Computing, Vienna, Austria, 2012.

Raddatz, T. J., Reick, C. H., Knorr, W., Kattge, J., Roeckner, E., Schnur, R., Schnitzler, K.-G., Wetzel, P., and Jungclaus, J.: Will the tropical land biosphere dominate the climate–carbon cycle feedback during the twenty-first century?, Clim. Dynam., 29, 565–574, doi:10.1007/s00382-007-0247-8, 2007.

Raich, J. W. and Schlesinger, W. H.: The global carbon dioxide flux in soil respiration and its relationship to vegetation and climate, Tellus B, 44, 81–99, doi:10.1034/j.1600-0889.1992.t01-1-00001.x, 1992.

Randerson, J. T., Thompson, M. V, Conway, T. J., Fung, I. Y., and Field, C. B.: The contribution of terrestrial sources and sinks to trends in the seasonal cycle of atmospheric carbon dioxide, Glob. Biogeochem. Cy., 11, 535–560, doi:10.1029/97GB02268, 1997.

Randerson, J. T., Still, C. J., Ballé, J. J., Fung, I. Y., Doney, S. C., Tans, P. P., Conway, T. J., White, J. W. C., Vaughn, B., Suits, N., and Denning, A. S.: Carbon isotope discrimination of arctic and boreal biomes inferred from remote atmospheric measurements and a biosphere-atmosphere model, Glob. Biogeochem. Cy., 16, 1028, doi:10.1029/2001GB001435, 2002.

Ryan, M. G. and Law, B. E.: Interpreting, measuring, and modeling soil respiration, Biogeochemistry, 73, 3–27, doi:10.1007/s10533-004-5167-7, 2005.

Sato, H., Itoh, A., and Kohyama, T.: SEIB–DGVM: A new dynamic global vegetation model using a spatially explicit individual-based approach, Ecol. Model., 200, 279–307, doi:10.1016/j.ecolmodel.2006.09.006, 2007.

Saxton, K. E., Rawls, W. J., Romberger, J. S., and Papendick, R. I.: Estimating generalized soil-water characteristics from Texture, Soil Sci. Soc. Am. J., 50, 1031–1036, doi:10.2136/sssaj1986.03615995005000040039x, 1986.

Schimel, J. and Weintraub, M.: The implications of exoenzyme activity on microbial carbon and nitrogen limitation in soil: a theoretical model, Soil Biol. Biochem., 35, 549–563, doi:10.1016/S0038-0717(03)00015-4, 2003.

Schmidt, M. W. I., Torn, M. S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I. A., Kleber, M., Kögel-Knabner, I., Lehmann, J., Manning, D. A. C., Nannipieri, P., Rasse, D. P., Weiner, S., and Trumbore, S. E.: Persistence of soil organic matter as an ecosystem property, Nature, 478, 49–56, doi:10.1038/nature10386, 2011.

Schuur, E. A. G., Bockheim, J., Canadell, J. G., Euskirchen, E., Field, C. B., Goryachkin, S. V., Hagemann, S., Kuhry, P., Lafleur, P. M., Lee, H., Mazhitova, G., Nelson, F. E., Rinke, A., Romanovsky, V. E., Shiklomanov, N., Tarnocai, C., Venevsky, S., Vogel, J. G., and Zimov, S. A.: Vulnerability of permafrost carbon to climate change: Implications for the global carbon cycle, Bioscience, 58, 701–714, doi:10.1641/B580807, 2008.

Shevliakova, E., Pacala, S. W., Malyshev, S., Hurtt, G. C., Milly, P. C. D., Caspersen, J. P., Sentman, L. T., Fisk, J. P., Wirth, C., and Crevoisier, C.: Carbon cycling under 300 years of land use change: Importance of the secondary vegetation sink, Glob. Biogeochem. Cy., 23, GB2022, doi:10.1029/2007GB003176, 2009.

Shi, X. Z., Yu, D. S., Warner, E. D., Pan, X. Z., Petersen, G. W., Gong, Z. G., and Weindorf, D. C.: Soil database of 1:1 000 000 digital soil survey and reference system of the chinese genetic soil classification system, Soil Survey Horizons, 45, 129–136, 2004.

Shur, Y., French, H. M., Bray, M. T., and Anderson, D. A.: Syngenetic permafrost growth: cryostratigraphic observations from the CRREL tunnel near Fairbanks, Alaska, Permafrost Periglac., 15, 339–347, doi:10.1002/ppp.486, 2004.

Six, J., Paustian, K., Elliott, E. T., and Combrink, C.: Soil structure and organic matter, Soil Sci. Soc. Am. J., 64, 681–689, doi:10.2136/sssaj2000.642681x, 2000.

Slater, A. G. and Lawrence, D. M.: Diagnosing present and future permafrost from climate models, J. Clim., in press, 2013.

Sombroek, W. G.: Towards a global soil resources inventory at scale 1:1 million, in Discussion Paper, ISRIC, Wageningen, The Netherlands., 1984.

Sombroek, W. G., Nachtergaele, F. O., and Hebel, A.: Amounts, dynamics and sequestering of carbon in tropical and subtropical soils, Ambio, 22, 417–426, 1993.

Tarnocai, C., Canadell, J. G., Schuur, E. A. G., Kuhry, P., Mazhitova, G., and Zimov, S.: Soil organic carbon pools in the northern circumpolar permafrost region, Glob. Biogeochem. Cy., 23, GB2023, doi:10.1029/2008GB003327, 2009.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res. Atmos., 106, 7183–7192, doi:10.1029/2000JD900719, 2001.

Taylor, K. E., Stouffer, R. J., and Meeh, G. A.: A summary of the CMIP5 experiment design, available at: http://cmip-pcmdi.llnl.gov/cmip5/docs/Taylor_CMIP5_design.pdf, 2011.

Thornton, P. E.: Regional ecosystem simulation: combining surface- and satellite-based observations to study linkages between terrestrial energy and mass budgets, University of Montana, Missoula., 1998.

Thornton, P. E. and Rosenbloom, N. A.: Ecosystem model spin-up: Estimating steady state conditions in a coupled terrestrial carbon and nitrogen cycle model, Ecol. Model., 189, 25–48, doi:10.1016/j.ecolmodel.2005.04.008, 2005.

Thornton, P. E., Law, B. E., Gholz, H. L., Clark, K. L., Falge, E., Ellsworth, D. S., Goldstein, A. H., Monson, R. K., Hollinger, D., Falk, M., Chen, J., and Sparks, J. P.: Modeling and measuring the effects of disturbance history and climate on carbon and water budgets in evergreen needleleaf forests, Agr. Forest Meteorol., 113, 185–222, doi:10.1016/S0168-1923(02)00108-9, 2002.

Thornton, P. E., Lamarque, J.-F., Rosenbloom, N. A., and Mahowald, N. M.: Influence of carbon-nitrogen cycle coupling on land model response to CO_2 fertilization and climate variablity, Glob. Biogeochem. Cy., 21, GB4018, doi:10.1029/2006GB002868, 2007.

Thornton, P. E., Doney, S. C., Lindsay, K., Moore, J. K., Mahowald, N., Randerson, J. T., Fung, I., Lamarque, J.-F., Feddema, J. J., and Lee, Y.-H.: Carbon-nitrogen interactions regulate climate-carbon cycle feedbacks: results from an atmosphere-ocean general circulation model, Biogeosciences, 6, 2099–2120, doi:10.5194/bg-6-2099-2009, 2009.

Tjiputra, J. F., Roelandt, C., Bentsen, M., Lawrence, D. M., Lorentzen, T., Schwinger, J., Seland, Ø., and Heinze, C.: Evaluation of the carbon cycle components in the Norwegian Earth System Model (NorESM), Geosci. Model Dev. Discuss., 5, 3035–3087, doi:10.5194/gmdd-5-3035-2012, 2012.

Todd-Brown, K. E. O., Hopkins, F. M., Kivlin, S. N., Talbot, J. M., and Allison, S. D.: A framework for representing microbial decomposition in coupled climate models, Biogeochemistry, 109, 19–33, doi:10.1007/s10533-011-9635-6, 2012.

Torn, M. S., Trumbore, S. E., Chadwick, O. A., Vitousek, P. M., and Hendricks, D. M.: Mineral control of soil organic carbon storage and turnover, Nature, 389, 170–173, 1997.

Trumbore, S.: Radiocarbon and soil carbon dynamics, Ann. Rev. Earth Planet. Sci., 37, 47–66, doi:10.1146/annurev.earth.36.031207.124300, 2009.

Veen, J. A. Van and Paul, E. A.: Orgainc carbon dynamics in grassland soils. 1. Background information and computer simulation, Can. J. Soil Sci., 61, 185–201, doi:10.4141/cjss81-024, 1981.

Veen, J. A., Ladd, J. N., and Frissel, M. J.: Modelling C and N turnover through the microbial biomass in soil, Plant Soil, 76, 257–274, doi:10.1007/BF02205585, 1984.

Volodin, E. M.: Atmosphere-ocean general circulation model with the carbon cycle, Izvestiya, Atmos. Oc. Phys., 43, 266–280, doi:10.1134/S0001433807030024, 2007.

Volodin, E. M., Dianskii, N. A., and Gusev, A. V.: Simulating present-day climate with the INMCM4.0 coupled model of the atmospheric and oceanic general circulations, Izvestiya, Atmos. Oc. Phys., 46, 414–431, doi:10.1134/S000143381004002X, 2010.

Watanabe, S., Hajima, T., Sudo, K., Nagashima, T., Takemura, T., Okajima, H., Nozawa, T., Kawase, H., Abe, M., Yokohata, T., Ise, T., Sato, H., Kato, E., Takata, K., Emori, S., and Kawamiya, M.: MIROC-ESM 2010: model description and basic results of CMIP5-20c3m experiments, Geosci. Model Dev., 4, 845–872, doi:10.5194/gmd-4-845-2011, 2011.

Wu, T. , Li, W., Ji, J., Xin, X., Li, L., Wang, Z., Zhang, Y., Li, J., Zhang, F., Wei, M., Shi, X., Wu, F., Zhang, L., Chu, M., Jie, W., Liu, Y., Wang, F., Liu, X., Li, Q., Dong, M., Liu, Q., and Zhang, J.: Global carbon budgets simulated by the Beijing Climate Center Climate System Model for the last century, J. Clim., submitted, 2013.

Young, O. R. and Steffen, W.: The Earth System: Sustaining planetary life-support systems, in Principles of ecosystem stewardship, pp. 295–315., 2009.

Zhao, M. and Running, S. W.: Drought-induced reduction in global terrestrial net primary production from 2000 through 2009, Science, 329, 940–943, doi:10.1126/science.1192666, 2010.

Zimov, S. A., Davydov, S. P., Zimova, G. M., Davydova, A. I., Schuur, E. A. G., Dutta, K., and Chapin, F. S.: Permafrost carbon: Stock and decomposability of a globally significant carbon pool, Geophys. Res. Lett., 33, L20502, doi:10.1029/2006GL027484, 2006.